PROCEEDINGS

Celebrating a decade of outstanding publications.

The Tenth International Conference on BioSignals, Images and Instrumentation

ICBSII 2024

March 20-22, 2024

DEPARTMENT OF BIOMEDICAL ENGINEERING

Sri Sivasubramaniya Nadar College of Engineering,
OMR, Kalavakkam, Chengalpattu, India

Proceedings of the 10th International Conference on Biosignals, Images and Instrumentation (ICBSII 2024)

Department of Biomedical Engineering

Sri Sivasubramaniya Nadar College of Engineering

March 20-22 , 2024

Editorial Board

Contents

Keynote Speakers

Research Papers

Track 1

Research Papers

Track 2

Research Papers

Track 3

Research Papers

Track 4

Research Papers

Track 5

Research Papers

Track 6

Research Papers

Track 7

Research Papers

Track 8

Research Papers

Track 9

Research Papers

Track 10

Research Papers

Track 11

Dedicated to

All the Staff and

Students

of the Department of Biomedical

Engineering

From the Chief Patron

Dr. Kala Vijayakumar

President, SSN Institutions

SSN Institutions (SSN) nurture the all-around development of the students, focusing not only on academic excellence but also on honing life skills such as leadership, discipline, team spirit, and time management. Students are encouraged to think critically and creatively. SSN prides itself on providing holistic education to its students.

Biomedical Engineering is a multi-disciplinary branch of study that brings together healthcare and technology, the front runners in the modern world. This dynamic field plays a crucial role in shaping the future of healthcare by driving innovation, improving medical technologies, and ultimately, enhancing the quality of life for individuals worldwide.

I congratulate the Department of Biomedical Engineering, SSN, for organizing the Tenth International Conference on Bio signals, Images, and Instrumentation (ICBSII 2024) in association with Drexel University, Philadelphia and Center for Healthcare Technologies, during 20-22, March 2024. This flagship conference of the Department of BME will provide the participants and the students with a unique opportunity to develop enriching perspectives by interacting with some of the renowned experts in these fields worldwide. These events facilitate collaboration, promote professional development, and provide a global perspective on biomedical engineering advancements. By staying updated on the latest research trends and connecting with peers from around the world, attendees can contribute to the advancement of the field and ultimately improve healthcare outcomes on a global scale. I appreciate the untiring, excellent teamwork carried out by the faculty of the Biomedical Engineering department towards organizing this conference.

I extend my felicitations to the BME department and wish the conference all success.

Dr. Kala Vijayakumar,
Chief Patron,
ICBSII – 2024

From the Patron

Dr. V. E. Annamalai

Principal, SSN College of Engineering

I am pleased that the Department of Biomedical Engineering, SSNCE in association with Drexel University, Philadelphia and Center for Healthcare Technologies is organising the Tenth International Conference on Bio signals, Images, and Instrumentation (ICBSII 2024) during 20-22, March 2024 in a manner befitting the biomedical engineer's community.

Biomedical engineering is a multidisciplinary field that encompasses the design, development, and implementation of innovative solutions ranging from medical devices and diagnostic tools to advanced therapies and healthcare systems. The department's involvement in a wide range of activities with students and faculty has strengthened it, allowing students to get a thorough full understanding of the industrial requirements possible.

This International Conference was conceived with the thought of bringing together scientists, engineers, and researchers from various domains all over the world. It has been a platform where some of the greatest minds of the country and abroad could interact, exchange ideas, and work together towards a common goal.

I congratulate the entire team of the Biomedical Engineering Department for structuring it to perfection and wish them all success.

Dr. V. E. Annamalai
Patron,
ICBSII – 2024

From the Conference Chair

Dr. A. Kavitha

Professor& Head,
Department of Biomedical Engineering,
SSN College of Engineering

Education serves as the cornerstone of progress, empowering individuals with the knowledge, skills, and perspectives needed to navigate an ever-evolving world. As conference chair, it is with great enthusiasm that I welcome you to this gathering of minds dedicated to advancing education. Our collective commitment to fostering learning, innovation, and collaboration is integral to shaping the future of education. Together, let us explore new ideas, share insights, and forge meaningful connections that will inspire transformative change in the realm of education and beyond. It gives me immense pleasure to present the Tenth International Conference on Bio-signals, Images, and Instrumentation (ICBSII 2024) in association with Drexel university.

Biomedical engineering stands at the intersection of healthcare, technology, and innovation, aiming to revolutionize the way we diagnose, treat, and understand medical conditions. With a blend of engineering principles and biological sciences, biomedical engineers design and develop cutting-edge medical devices, prosthetics, imaging systems, and therapies that enhance patient care and improve quality of life. As we delve into this dynamic field, we embark on a journey of discovery and transformation, driven by a shared commitment to harnessing science and technology for the betterment of humanity. The impact and significance of findings keeps growing progressively. Workshops, seminars, project exhibitions, and guest lectures on current trends relevant to the core and multidisciplinary disciplines in biomedical engineering are often offered by the department to help students obtain a thorough understanding of industry requirements.

The 10th International Conference on Biosignals, Images, and Instrumentation is being organised by the Centre for Healthcare Technologies, a multidisciplinary research initiative focusing on research through innovation in healthcare, in collaboration with Drexel University, with the goal of instilling research aptitude in students and providing a great platform for researchers to showcase their work in various domains of healthcare. It is rightly said, “A group becomes a team when each member is sure enough of himself and his contribution to praise the skills of others.” – Norman Shidle.

For the past Nine years, ICBSII has been successful in inculcating knowledge in all the participants. With confidence, we aim higher and higher, raising our bars toward the next success!!!

Dr. A. Kavitha
Conference Chair,
ICBSII – 2024.

From the Conference Co-Chair

Dr. Mark E Schafer
Fellow, AIUM, ASA, AIMBE, UIA
Past President, IEEE Ultrasonics, Ferroelectrics, and Frequency Control Society
Research Professor,
School of Biomedical Engineering, Science, and Health Systems
Drexel University

It is my great honor and pleasure to serve as the co-Chair of the Tenth International Conference on Biosignals, Images, and Instrumentation (ICBSII), a joint collaboration between The Department of Biomedical Engineering, SSNCE and Drexel University, Philadelphia. This conference shows the possibilities of worldwide cooperation in biomedical education, serving the goals of both institutions and countries.

As a student working on my PhD in Biomedical Engineering in 1984, I was often asked what it meant to have a "Biomed" degree, and why I was not in Electrical Engineering or some other more traditional field of study. Biomed was a new subject area, and I was fortunate to be at Drexel University, one of the first institutions to confer degrees in that field. Now so many years later, Biomedical Engineering is well established and one of the fastest growing areas of study and employment worldwide. The degree is recognized as combining fundamental understanding of biological systems with an engineering approach to real-world problem solving. Our collective goal is to provide our students with the education necessary to change lives through their development of new devices and treatments. It is one thing for the surgeon or physician to treat a patient; is it another thing to have developed tools and therapies that can treat thousands.

I join you all here in this wonderful setting to learn together from the assembled plenary and invited speakers, as well as your fellow student presenters. I hope that you will gain not only experience in presenting your work, and knowledge of other's efforts, but also contacts and connections with students, faculty, and researchers that will be a part of your educational and working environment for years to come. Turn contacts into colleagues and collaborators!

I want to thank the Department of Biomedical Engineering, SSNCE, the School of Biomedical Engineering, Science, and Health System, Drexel University, and the IEEE Ultrasonics, Ferroelectrics, and Frequency Control Society for their support. I wish you all a successful conference

Dr. Mark Schafer
Conference Co-Chair,
ICBSII – 2024

From the Organizing Chairs

Dr. S. Arun Karthick
Associate Professor,
Dept. of BME
SSNCE

Dr. K. Nirmala
Associate Professor,
Dept. of BME
SSNCE

Dr. N. Punitha
Assistant Professor,
Dept.of BME
SSNCE

Dr. Kambiz Pourrezaei
Professor,
School of BME
Drexel University

Biomedical Engineering is at the forefront of revolutionizing healthcare by integrating cutting-edge technology with medical science. This interdisciplinary field encompasses various specialities, including Biosignal processing, Medical Instrumentation, Biomechanics, Bioinformatics, Tissue Engineering, and Medical Imaging. Through collaborative efforts, Biomedical Engineers strive to develop innovative solutions to address pressing healthcare challenges, from personalized medicine and regenerative therapies to advanced diagnostic tools and wearable health monitoring devices. With the advancements in technology and the continuous evolution of the healthcare industry, establishing a conducive environment for the exchange of ideas and information is of paramount importance. So, we are extremely delighted to present our 10th International Conference on Biosignals, Images and Instrumentation (ICBSII-2024) which is to be held from 20th March to 22nd March at SSN College of Engineering, Department of Biomedical Engineering.

We have seen a tremendous response this time, with 194 papers received from the USA, the Middle East, Singapore, and Malaysia and speakers from different parts of the world will be sharing their expertise with us. Only 74 papers with good technical content, quality, and presentation style were accepted and will be forwarded to IEEE Digital Explore.

We extend our sincere gratitude to the management of SSN College of Engineering, our president Dr. Kala Vijayakumar SSN Institutions, and Dr. V. E. Annamalai sir - Principal, SSN College of Engineering for granting the department a delightful opportunity to organize this brilliant event that enables us to grow on a global level. We would like to thank our department Head and organizing chair Dr. A. Kavitha and Dr. Mark. E. Schafer (Drexel University), our colleagues involved in the ICBSII-2024 and our student volunteers for their support, constructive feedback, and meticulous efforts in conducting this event. Finally, we would also like to take this opportunity to thank all external reviewers and contributing authors for producing high-quality papers to be presented at the conference.

Conference Organizing Committee

Chief Patron

Dr. Kala Vijayakumar, President, SSN Institutions

Patron

Dr. V. E. Annamalai, Principal, SSN College of Engineering

Conference Chair

Dr. A. Kavitha, Professor and HOD/BME
Dr. Mark E Schafer (Drexel University)

Organizing Chair

Dr. S. Arun Karthick
Dr. N. Punitha
Dr. K. Nirmala
Dr. Kambiz Pourrezaei (Drexel University)

Conference Treasurer

Dr. S. Saranya
Dr. R. Subashini

Technical Program Chair

Dr. Kurtulus Izzetoglu (Drexel University)
Dr. N. Venkateswaran
Dr. S. Pravin Kumar
Dr. B. Geethanjali
Dr. M. Dhanalakshmi
Dr. J. Vijay

Publication Chair

Ms. B. Divya
Dr. R. Nithya
Dr. S.Allwyn

Information Contact

Dr. L. Suganthi

Registration Chair

Dr. R. Nithya
Dr. M. Dhanalakshmi

Supporting Chair

Ms. J. Suganthi
Mr. E. Sivananthan
Mr. A. Udayachandran
Mr. T. R. Sasidharan
Ms. Gayathri

Technical Advisory Committee

International Advisory Committee Members

- **Dr. Ivan corazza**, Bologna, Italy
- **Dr. Sriram Balasubramanian**, Drexel University, USA
- **Dr. J. Jesu Christopher**, ASTRAZENCA, Cambridge, UK
- **Dr.Yuvaraj Rajamanickam**, NTU, Singapore.
- **Dr. Justin Dauwels**, University of Delft.
- **Dr. Funminiyi Olajide**, Nottingham Trent University,UK.
- **Dr. V John Mathews**, School of Electrical Engineering & Computer Science, Oregon State University.
- **Dr. Jaume Anguera**, Universitat Ramon Llull University (URL),Barcelona, Spain.

National Advisory Committee Members

- **Dr. S. Ramakrishnan**, IIT Madras
- **Dr. Nilesh J Vasa**, IIT Madras
- **Dr. Renu John**, IIT Hyderabad
- **Dr.N. Sujatha**, IIT Madras
- **Dr. K. Mohanavelu**, DEBEL, DRDO, Bangalore
- **Dr. G. Sudhir**, Orthopaedic Spine Surgeon, SRMC, Chennai
- **Dr. M.Sasikala**, Anna University, Chennai
- **Dr. Venugopal**, APJ Abdul Kalam Technological University - Kerala
- **Mr. Balamurugan**, HCL
- **Dr. B. Banu Rekha**, PSG College of Technology, Coimbatore.
- **Dr. Dinesh Bhatia**, North-Eastern Hill University, Shillong
- **Dr. Hariharan**, NIT Uttarakhand
- **Dr Vaithiyanathan Dhandapani**, NIT Delhi
- **Dr. K K Deepak**, AIIMS Delhi
- **Dr. Gnanou Florence Sudha**, Puducherry Technological University
- **Dr. C. Arunachalaperumal**, Chairman, IEEE Photonics society
- **Dr. K. Porkumaran**, Chairman, IEEE Madras section, Principal, Sri Sairam Engineering College.
- **Dr. R. Hariprakash**, Secretary, IEEE Madras Section
- **Dr. S. Radha**, Treasurer, IEEE Madras Section
- **Dr. K. B. Jayanthi**, Chairman IEEE EMBS, K.S. Rangasamy College of Technology
- **Dr. R. Periyasamy**, Assistant Professor, NIT Tiruchirappalli

Review Committee

Technical Committee Members- IEEE Madras Section

- **Dr. O. Uma Maheswari,** Department of ECE, CEG Campus, Anna University, Chennai.
- **Dr. D. Vaidhyanathan ,** Department of ECE, NIT Delhi
- **Dr. S. Sasirekha,** SSN College of Engineering, Chennai
- **Dr. T. Sree Sharmila ,** Department of IT, Anna University, Chennai
- **Dr. J.K. Josephine Julina ,** SSN College of Engineering, Chennai
- **Dr. I. Joe Louis Paul ,** SSN College of Engineering, Chennai
- **Dr. E. Priya ,** Department of ECE, Sri Sairam Engineering College
- **Dr. A. K. Jayanthy,** Medical Engineering, Sri Ramachandra Faculty of Engineering and Technology, Porur, Chennai
- **Dr. K. Muthumeenakshi,**SSN College of engineering, Kalavakkam
- **Dr. K. Ravish,**Department of Medical Electronics, Dr. Ambedkar Institute of Technology, Bangalore
- **Dr.V.Vijayan,** Department of EIE St. Joseph's College of Engineering, Chennai
- **Dr. G. Kulanthaivel,** Department of ECE, NITTTR, Chennai
- **Dr. C. Vinoth Kumar,** SSN College of engineering, Kalavakkam
- **Dr. B. Banurekha Associate,** Department of BME, PSG College of Technology, Coimbatore
- **Dr. P. Rohini,** Department of ECE, IIITDM Kancheepuram
- **Dr. G. S. Uthayakumar,** Department of ECE, St.Joseph's Institute of Technology, Chennai
- **Dr. J. Samson Isaac,** Department of Biomedical Engineering Karunya Institute of Technology and Sciences, Coimbatore.
- **Dr.Kishore Balasubramanian,** Dr.Mahalingam College of Engineering and Technology
- **Dr. Gnanavel S,** Biomedical Engineering Department, SRMIST
- **Dr. M. Sridevi,** Department of Computer Science & Engineering, National Institute of Technology, Tiruchirappalli
- **Dr.R.Thilagavathy ,**Department of ECE, National Institute of Technology, Tiruchirappalli
- **Dr. Ashwin N ,**Biomedical Engineering Department, SRMIST
- **Dr. N.B. Muthu Selvan ,**SSN College of Engineering, Kalavakkam
- **Dr. V. Sherlin Solomi,** Hindustan Institute of Technology and Science
- **Dr. P. Palanisamy,** Department of ECE, NIT, Trichy
- **Dr. Varun P. Gopi ,**NIT, Trichy
- **Dr. Jino Hans W,** SSN College of Engineering
- **Dr. K.Usha,** SSN College of Engineering
- **Dr. Nithyavathy N,** Kongu Engineering College
- **Dr. Jawahar A,** SSN College of Engineering
- **Dr. Joseph Gladwin S,**BigCat Wireless, IITM Research Park, Chennai

International Review Committee Members

- **Dr.Vivek Padmanaabhan Indramohan,** Senior Lecturer Birmingham City University, UK
- **Dr. Kong Pui Wah,** Associate Professor Nanyang Technological University, Singapore
- **Dr. Teo EE Chon,** Professor, Nanyang Technological University, Singapore
- **Dr. R. Sivaramakrishnan,** Research Scientist, National Institutes of Health, USA
- **Dr. Rahuman Sheriff,** Project Leader Biomodels European Bioinformatics Institute, UK
- **Dr. R. Yuvaraj,** Research Scientist Nanyang Technological University, Singapore
- **Dr. Aleksandra Kawala- Janik,** Research scientist Opole University of Technology
- **Dr. Eko Supriyanto,** Professor Ijn-Utm Cardiovascular Engineering Centre, UTM
- **Dr. M. Murugappan,** Associate Professor Kuwait college of Science and Technology, Kuwait.
- **Dr. Md. Mamun Bin Ibne Reaz,** Professor Department of Electrical, Electronic and Systems Engineering, Universiti Kebangsaan Malaysia 43600 UKM, Bangi, Selangor Malaysia
- **Dr. Soon Xin Ng (Michael),** Associate Professor Electrical Engineering Department, University of Southampton.
- **Dr.Ali S Awad,** Associate Professor Faculty of Engineering and information Technology, Al-Azhar University-Gaz
- **Dr. Anita Singh,** Associate Professor Biomedical Engineering, Widener University, Chester, PA
- **Dr. Alejandro Riera,** Asia Commercial Manager Neuroelectrics Barcelona S.L. Barcelona Area, Spain.
- **Professor Dr Yazeed Al-Sbou,** Computer Engineering Department, Faculty of Engineering, Karak, Jordan
- **Dr. Deboleena Sadhukhan,** Postdoctoral researcher Institut LANGEVIN, ESPCI Paris - PSL University, France
- **Dr. Pauline John,** Research Scientist New York University, Abu Dhabi
- **Dr. S. Sathiyamoorthy,** Postdoctoral researcher Carleton University, Ottawa, Canada.
- **Dr. Srikrishnarka Pillalamarri,** Postdoctoral researcher Tampere University
- **Dr. Azhar Zam,** Associate Professor New York University Abu Dhabi
- **Dr. Sarathkumar Loganathan,** Post-Doctoral researcher Leeds University, UK.
- **Dr. Divagar Murugan,** Post-Doctoral researcher RWTH Aachen University, Germany
- **Dr. Ravi Dadsena,** Post-Doctoral researcher RWTH Aachen University Hospital
- **Dr. Yedukondala Rao,** Post-Doctoral researcher University of Connecticut, USA
- **Dr. Tamil Selvi,** Post-Doctoral researcher University of Texas, Austin
- **Dr. Kulandaisamy Arulsamy,** Post-Doctoral researcher Basic and Translational Research Division, Department of Cardiology, Boston Children's Hospital, Boston, USA.
- **Dr Jesin James,** Lecturer, Department of Electrical , Computer and software Engineering, University of Auckland, New Zealand

- **Dr.Clement Yuen,** Lecturer, Nanyang Technological University, Singapore
- **Dr.Kunio Kasugai,** Professor Aichi Medical University, Japan
- **Dyah Ekashanti Octorina Dewi,** Senior Lecturer Faculty of Biosciences and Medical Engineering. IJN –UTM Cardiovascular Engineering Center, Institute of Human Centered Engine

National Review Committee Members

- **Dr.A. Kavitha,** Ph.D. Professor & Head SSN College of Engineering, Chennai
- **Dr. N. Venkateshwaran,** Professor Department of BME, SSN College of Engineering, Chennai
- **Dr.Pravinkumar S,** Asso. Professor SSN College of Engineering, Chennai
- **Dr. L. Suganthi,** Asso. Professor SSN College of Engineering, Chennai
- **Dr. J. Vijay,** Asso. Professor SSN College of Engineering, Chennai
- **Dr. S. Arun Karthick,** Asso. Professor SSN College of Engineering, Chennai
- **Dr.Geethanjali,** B Asso. Professor SSN College of Engineering, Chennai
- **Dr.K.Nirmala,** Asso. Professor SSN College of Engineering, Chennai
- **Dr. R. Subashini,** Asso. Professor SSN College of Engineering, Chennai
- **Dr. Sachin G Sarate,** Asst. Professor SSN College of Engineering, Chennai
- **Dr. Sivaraman J,** Asso. Professor Department of Biotechnology & Medical Engineering, National Institute of Technology, Rourkela
- **Dr.C. M. Sujatha,** Asst.Prof Dept. of ECE, Anna University
- **Dr. A. Mythili,** Associate Professor Sensor and Biomedical Technology Division, School of Electronics Engineering (SENSE), VIT Vellore campus.
- **Dr. K. Kamalanand,** Asst. Prof MIT campus, Anna University
- **Dr. B. Minimol,** Ph.D. Associate Professor Department of Electronics & Biomedical Engineering, Model Engineering College, Kerala.
- **Dr. R. Tamilselvi,** Ph.D. Professor Department of ECE, Sethu Institute of Technology, Madurai.
- **Dr. A. K. Jayanthy,** Ph.D. Professor HOD, BME, SRMU, Chennai
- **Dr.S.Poonguzhali,** Ph.D. Associate Prof Department of ECE, CEG Campus, Anna University
- **Dr. S.Sasikala,** Ph.D. Professor Department of ECE, CEG Campus, Anna University
- **Dr. K. Adalarsasu,** Ph.D. Associate Prof School of EEE, Sastra University, Tanjavur
- **Dr. Ganesh Vaidhyanathan,** Ph.D. Principal, SVCE, Chennai.
- **Dr. R. Premkumar,** Professor Rajalakshmi Engineering College
- **Dr. T. R. Ganesh babu,** Ph.D. Professor Muthayammal Engineering College, Rasipuram
- **Dr. M. C. Jobin Christ,** Ph.D. Professor Department of Biomedical Engineering, Rajalakshmi Engineering College, Chennai
- **Dr. K. Vidhya,** Ph.D. Professor School of ECE, Saveetha University, Chennai.
- **Dr.B.Padmapriya,** Ph.D. Associate Professor Department of Biomedical Engineering, PSG College of Technology

- **Dr. Judith Justin,** Ph.D. Professor, Department of Biomedical Instrumentation Engineering, Avinashilingam University
- **Dr. M. Parisa,** Ph.D. Associate Professor Dept of ECE, Sethu Institute of Technology, Kariapatti, Virudhunagar.
- **Dr. P. Hosanna Princye,** Ph.D. Associate professor, Electronics and Communication Engineering department, SEA College of Engineering and Technology, KR Puram, Bangalore -560049
- **Dr. K. Bommanna Raja,** Ph.D. Professor, Electronics and Communication Engineering department, KPR Institute of Engineering and Technology, Coimbatore.
- **Dr. O. Saranya,** Ph.D. Asso. Professor Department of ECE, Government College of Technology, Coimbatore
- **Dr. T. Jayashree,** Ph.D. Asst. Professor Department of ECE, CEG Campus, Anna University
- **Dr.R.Amutha Professor,** Department of ECE, SSN College of Engineering, Chennai
- **Dr. P. Sakthivel,** Ph.D. Professor, Department of ECE, CEG Campus, Anna University
- **Dr. M. Dhanalakshmi,** Asso. Professor Department of BME, SSN College of Engineering, Chennai
- **Dr. R. Nithya,** Assi. Professor Department of BME, SSN College of Engineering, Chennai
- **Dr. S. Saranya,** Assi. Professor Department of BME, SSN College of Engineering, Chennai
- **Dr. N. Punitha,** Assi. Professor Department of BME, SSN College of Engineering, Chennai
- **Dr. S. Allwyn,** Assi. Professor Department of BME, SSN College of Engineering, Chennai
- **Dr. Satyavratan G,** Research Engineer vTitan Corporation Pvt. Ltd., Estancia IT Park New Woodlands Building, Vallancheri, Chengalpattu
- **Dr. Diptasree Maitra Ghosh,** Research Associate Centre for Programmable Photonic Integrated Circuits and Systems, IIT Madras
- **Dr. Lakshmi M Hari,** Assi. Professor Department of Electronics and Biomedical Engineering, Adi Shankara Institute of Engineering and Technology Kalady, Kerala
- **Dr. Logesh Kumar S,** Assistant Professor (Sl .G.) Department of Biomedical Engineering, KPR
- **Dr.P.Nirmala,** Assistant Professor School of Electronics Engineering,
- **Dr. V. Karthikeyan,** Professor Department of Electronics and Institute of Engineering and Technology, Coimbatore Vellore Institute of Technology Chennai Campus Communication Engineering Mepco Schlenk engineering college,Sivakasi
- **Dr. S.P.Angeline Kirubha,** Associate Professor Biomedical Engineering Department, School of Bioengineering, College of Engineering and Technology, SRM Institute of Science and Technology,
- **Dr.Balaji E,** Assistant Professor VIT University, Vellore.
- **Dr.R.Parameshwari,** Assistant Professor Dept of ECE, CEG, Anna University,
- **Dr.D.Santhi,** Associate Professor Mepco Schlenk Engineering College.
- **Dr. Kathirvelu D,** Assistant Professor Biomedical Engineering Department, SRMIST

- **Dr. Sudip Paul,** Associate Professor Department of Biomedical Engineering, NEHU, Shillong
- **Dr Bethanney Janney J,** Associate Professor Sathyabama Institute of Science and Technology
- **Dr. B. Devi,** Assistant Professor School of Bio and chemical Engineering Sathyabama Institute of Science and Technology
- **Dr.Jeeva J B,** Associate Professor Grade 1 School of Electronics Engineering Sensor and Biomedical Technology, VIT
- **Dr. Sasikumar K,** Assistant Professor (Senior) School of Electronics Engineering Sensor and Biomedical Technology, VIT Vellore.
- **Dr.Selvathi,** Senior Professor & Head, BioMedical Engineering Programme, Department of Electronics and Communication Engineering, Mepco Schlenk Engineering College
- **Dr S Esther Florence,** Associate Professor College of Engineering, Anna University, Guindy
- **Dr.P Prakash,** Professor Department of Electronics Engineering MIT Campus, Anna University
- **Dr. S. Ramprabhu,** Assistant Professor Department of Electronics Engineering MIT Campus, Anna University
- **Dr. M. Anbuselvi,** Associate Professor SSN College of Engineering
- **Dr.R.Periyasamy,** Assistant Professor Department of Instrumentation and Control, NIT Trichy
- **Dr R Hemalatha,** Associate Professor SSN College of Engineering
- **Dr.Uma Maheswari Sadasivam,** Associate Professor IT Dept, SSN College of Engineering
- **Dr.S.Saraswathi,** Associate Professor CSE, SSN College of Engineering Dr.D.Thenmozhi Associate Professor CSE, SSN College of Engineering
- **Dr. Suganthi,** Professor Department of Artificial Intelligence and Data Science, Sri Sai Ram Institute of Technology.
- **Dr.M.Kanthimathiu,** Professor Department of Electronics and Communication Engineering Sri Sai Ram Engineering College.
- **Dr.K.Sumathi,** Professor Department of Electronics and Communication Engineering Sri Sai Ram Engineering College.
- **Dr.Nelson.I** Associate Professor Department of Electronics and Communication Engineering, SSNCE
- **Dr. V Vaithianathan,** Associate Professor SSN College of Engineering
- **Dr. J. Sofia Bobby,** Associate Professor Department of Biomedical Engineering, Jerusalem College of Engineering
- **Dr.P.V.Pramila,** Associate Professor Saveetha School of Engineering
- **Dr. G. Kavitha,** Professor Department of Electronics, MIT Campus, Anna University, Chennai

Student Organizing Committee

Tracks:

- **Neelambari R A,** III Yr, UG
- **Gavishna R,** III Yr, UG
- **Kalpanadevi T,** III Yr, UG
- **Venkatesh M,** III Yr, UG
- **Farzana Fathima M K,** III Yr, UG
- **Ishwarya Ramachandran,** III Yr, UG
- **Hariharasudhan M,** III Yr, UG
- **Parvathi A,** III Yr, UG
- **Induja M C,** III Yr, UG
- **Krithiukha K B,** III Yr, UG
- **Sruthi Sankari S R,** III Yr, UG
- **Ranjana Sridhar,** III Yr, UG
- **Krishnaraj R,** III Yr, UG
- **Anchana Sri N,** II Yr, UG
- **Vijayalakshmi M,** II Yr, UG
- **Sujay,** II Yr, UG
- **Harishini H,** II Yr, UG
- **Priyadarshini Sivaraja,** II Yr, UG
- **Madhumitha B,** II Yr, UG
- **Aravinthan S,** III Yr, UG
- **Jeya Marshalin M,** III Yr, UG
- **Vaishali S,** III Yr, UG
- **Bruhati S,** III Yr, UG
- **Shakeel J,** II Yr, UG
- **Naveena N,** II Yr, UG
- **Augustine Wisely Bezalel,** II Yr, UG
- **Dineshkumar C,** II Yr, UG
- **Ezhil Meena A S,** II Yr, UG
- **Naveena N,** II Yr, UG
- **Sharadha G,** II Yr, UG
- **Rakshana,** II Yr, UG
- **Priyenga S,** II Yr, UG
- **Akshaya D,** II Yr, UG
- **Shrinidhi S K,** II Yr, UG

Hospitality :

- **Ishank Nandhan R,** III Yr, UG
- **Jothisa K,** III Yr, UG
- **Sibi Chakkaravarthi S,** III Yr, UG
- **Shivani Jasmin Kunjithapatham,** III Yr, UG
- **Dharini S,** III Yr, UG
- **Rishvega K R,** III Yr, UG

Accomodation :

- **Selvam P,** III Yr, UG
- **Venkatesh M,** III Yr, UG
- **Vaishali S,** III Yr, UG
- **Swaati S,** III Yr, UG

Operations :

- **Manimaran C,** III Yr, UG
- **Geetika H P,** III Yr, UG
- **Ilakkia Bharathi B,** III Yr, UG
- **Sharavan Kumar S,** II Yr, UG
- **Parvathi A,** III Yr, UG
- **Sruthi Sankari S R,** III Yr, UG
- **Aravinthan S,** III Yr, UG
- **Bruhati S,** III Yr, UG
- **Karishmaa Sathish Kumar,** III Yr, UG
- **Aishwarya S,** III Yr, UG
- **Subathira R,** PG I Yr
- **Pratheebha G,** PG I Yr
- **Krisha H,** PG I Yr

Media:

- **Jesse Samuel Stephen R,** III Yr, UG
- **Tarun Kumar S,** III Yr, UG
- **Hameem Ali,** III Yr, UG
- **Abdul Rahman,** III Yr, UG
- **Deepika K,** III Yr, UG
- **Jivethesh B,** III Yr, UG

Design & Editorial :

- **Pradiep Kumar G D**, IV Yr, UG
- **Renukka Shanmuga Sundari M,** III Yr, UG
- **Sarvajith P,** III Yr, UG
- **Nivedita Anugonda,** III Yr, UG
- **Jeevan J,** III Yr, UG
- **Krithiukha K B,** III Yr, UG
- **Hariharasudhan M,** III Yr, UG
- **Karthick Siva R,** III Yr, UG
- **Akshaya D,** II Yr, UG
- **Shrinithi S K,** II Yr, UG
- **Dinesh Kumar C,** II Yr, UG
- **Sharadha G,** II Yr, UG
- **Sujay,** II Yr, UG
- **Anagha Pradeep**, PG II Yr
- **Aishwarya M K,** PG I Yr

Registrations :

- **Sibiraj M,** III Yr, UG
- **Ishwarya Ramachandran,** III Yr, UG
- **Vayishnavee B,** III Yr, UG
- **Krithiukha K B,** III Yr, UG
- **Jeya Marshalin M,** III Yr, UG
- **Sreeniveatha P,** PG I Yr

Catering Co-ordination :

- **Suriyaprakash B,** III Yr, UG
- **Anjugam C,** III Yr, UG
- **Vasanthasurya E,** III Yr, UG
- **Bharathirpriya S,** III Yr, UG
- **Kishore S,** III Yr, UG
- **Nitheesh S,** II Yr, UG
- **Harish ,** II Yr, UG
- **Abinaya Priya L,** II Yr, UG
- **Lokiya V,** II Yr, UG
- **Mahalakshmi I,** II Yr, UG

Tenth International Conference on Biosignals, Images and Instrumentation (ICBSII 2024)

KEYNOTE SPEAKERS' PROFILE

Extended Reality in surgery: current experiences and future challenges

ABOUT THE SPEAKER:

Dr. Laura Cercenelli
Adjunct Professor
Dept. of Biomedical Engineering and Neuromotor Sciences,
University of Bologna, Italy

Dr. Laura Cercenelli obtained a degree in Biomedical Engineering in 2003 at Polytechnic University in Milano, Italy, and a Ph.D. in Electromagnetism and Bioengineering in 2006 at the Marche Polytechnic University, Italy.

She has long experience as a Post-doctoral Researcher in the Bioengineering field at the School of Medicine of the University of Bologna, and she is currently Adjunct Professor of Computer Science at the University of Bologna.

Her research interests include the field of sensors and devices for cardiac function monitoring, and more recently, the area of digital technologies applied to medicine. In particular, she has gained competencies and expertise in 3D modeling, 3D printing, Virtual and Augmented Reality applied to surgery for pre-operative planning, intraoperative guidance, and training. She has been working for 7 years at eDIMES Lab, an in-hospital laboratory of the University of Bologna focused on advanced digital technologies for supporting clinical research, practice, and medical education. She is the author/co-author of more than 80 scientific publications in peer-reviewed journals, and she is co-inventor of 5 US, 5 European, and 9 Italian patents.

ABSTRACT:

In the rapidly evolving landscape of healthcare, the integration of advanced image-based digital technologies is revolutionizing how we approach patient care. This talk delves deep into the potential of eXtended Reality (XR)—a spectrum that includes Virtual Reality (VR), Augmented Reality (AR), and Mixed Reality (MR)—as a key tool to enhance surgical precision and visualization in the field of computer-assisted interventions, as well as to offer a safe and exciting platform for surgical training.After an overview of the technological aspects that characterize different AR/MR vision systems, we will present the latest findings, insights, and experiences in using XR for pre-operative planning, intraoperative guidance, and surgical training, focusing on the potential and challenges still to be faced.

Designing ultrasound systems for brain research across different spatial scales

ABOUT THE SPEAKER:

Dr. Mark E. Schafer
PhD. Research Professor
School of Biomedical Engineering, Science and Health Systems,
Drexel University, Philadelphia, USA

Dr. Mark E. Schafer has had an extensive leadership career in the medical device industry, with hands-on development and program management experience from R&D through regulatory and commercialization. He is an internationally recognized expert in medical ultrasound, including design, development, intellectual property, regulatory, and applications. He has been a serial entrepreneur and is a named inventor on over 38 patents. He is a Fellow of the American Institute of Ultrasound in Medicine, the Acoustical Society of America, the American Institute for Biological and Medical Engineering, and the Ultrasonic Industry Association. After a thirty-five year career as an entrepreneur, consultant, and business leader, he was recently appointed Research Professor at Drexel University. He pioneered measurement devices, techniques, and standards ensuring the safety and efficacy of ultrasound imaging, therapeutic, and surgical systems. He played a crucial role in establishing domestic and international standards for ultrasound imaging systems. He is recognized as a leading authority on the acoustics of phacoemulsification and ultrasonic surgery, with extensive lecturing experience globally on ultrasonic cutting techniques. Recently, he assisted a startup company in developing a viable clinical research system for neuro-stimulation using ultrasound, which is now gaining traction as a drug-free alternative for disorders such as anxiety and coma. His current efforts include NIH-funded research into cardiac output measurement using piezopolymer ultrasonic catheters and the development of novel ultrasound sources for neuro-stimulation.

ABSTRACT:

Recently, there has been growing interest in using ultrasound for therapeutic purposes in neurological conditions such as anxiety, depression, and coma, as well as for facilitating drug delivery through the blood-brain barrier. However, there are questions about how ultrasound interacts with neural tissues. So, there is a need for the development of ultrasound sources tailored to specific experimental requirements across various scales, from fruit flies to rodents to humans.Each situation presents unique challenges in terms of source size, frequency, exposure volume, output control, dosimetry, and safety. For example, a sub-millimeter 100kHz exposure source was designed to stimulate fruit fly neurons, requiring a compact design to fit into a 300μm aperture. The final design drew from both surgical and industrial ultrasound domains, producing a uniform spherical sound field capable of inducing either excitatory or inhibitory effects in the fruit fly. Moving up in scale, a 2MHz transducer was developed to expose specific regions of a rodent brain while leaving nearby areas unexposed for control purposes. Simulations of the ultrasound beam were conducted and validated with hydrophone measurements. Experiments showed increased uptake of exosomes within the brain, without the need for microbubbles.

My Journey to System Safety

ABOUT THE SPEAKER:

Dr. Sandy Weininger
Senior Electrical/Biomedical Engineer, Center for Devices and Radiological Health, Food and Drug Administration, USA

Dr. Sandy Weininger is currently a senior engineer at the FDA's Center for Devices and Radiological Health (CDRH). He is keenly interested in the performance of sensors and actuators in assuring the safety of interoperable systems and electrical safety. He works with standards development organizations, including UL, AAMI, IEC and ISO to construct both horizontal and vertical safety standards. He is actively involved in developing and delivering courses on achieving safety in medical devices using systems engineering principles.

Dr. Weininger received his BSEE and MS/BME from Drexel University with a research thesis focusing on the properties of the electrode-tissue interface. He received his Ph.D. in Bioengineering from the University of Pennsylvania where he developed a system to evaluate the behavioral organization of the newborn nutritive sucking response, and specialized in signal processing and control systems.

He is a member of the AAMI Interoperability Working Group, AAMI-UL 2800 committee, chair of the ISO-IEC Pulse Oximeter Committee, and FDA's liaison to IEC TC 62 and SC 62A, committees responsible for safety of electromedical equipment. He has been an advisor to the MD PnP Program since its inception in 2004.

ABSTRACT:

System safety should be a primary objective of medical device product development. Students are not often exposed to the concepts of safety engineering until they enter the workforce. Medical device engineers – from early researchers through product development and deployment – should understand what it takes to build a safe product. This talk will focus on some personal history on how I became a safety engineer and what it means to me.without the use of microbubbles. Finally, a system for clinical (human) use had completely different design requirements that emphasized magnetic resonance imaging (MRI) and optical guidance, as well as safety and the capability to conduct double blind studies. flexibility This design used 650kHz transducers, as a compromise between focusing strength and attenuation loss through the skull. The design included internal MRI and external optical tracking markers along with a comprehensive approach to exposure control. Clinical data show positive results for several neurological conditions, such as anxiety, coma, and essential tremor, with no patient safety issues. In summary, ultrasound transducer technology has the flexibility to encompass a wide range of experimental requirements, supporting new areas of research and discovery in brain science.

Translating Assessment of Brain Function to Operational Environment via Functional Near Infrared Spectroscopy

ABOUT THE SPEAKER:

Dr. Kurtulus Izzetoglu
Associate Professor
School of Biomedical Engineering, Science and Health Systems,
Drexel University, Philadelphia, USA

Dr. Kurtulus Izzetoglu, PhD, Associate Professor in the School of Biomedical Engineering, Science & Health Systems, and affiliate faculty in the School of Education at Drexel University, Philadelphia, PA, USA. Dr. Izzetoglu's research and teaching interests are in neuroengineering, functional brain imaging, biomedical signal processing, medical device development, human performance, learning and training. He has a background in both electrical and biomedical engineering coupled with experience of developing and adopting highly portable optical brain imaging systems. His current research projects related to human performance assessment include integration of the fNIRS and other biosensors to enable personalized training in safety critical tasks; brain-in-the-loop studies to improve human autonomy teaming and training both in clinical (nurse, surgeon training) and in aviation (autonomous systems); effects of immersive VR manipulative interventions on learning. Dr. Izzetoglu has been the institutional principal investigator and the university lead in two U.S. Federal Aviation Administration (FAA) Centers of Excellence (COE) - COE for Technical Training and Human Performance and COE for Unmanned Aircraft Systems. He also serves on the boards and program committees, including International Conference on Augmented Cognition / International Conference on Human – Computer Interaction, International Conference for Research in Air Transportation (ICRAT) and International Symposium on Aviation Psychology.

ABSTRACT:

We have seen a marked increase in the availability of measurement techniques in assessing various aspects of human cognition that are prevalent in various operational environments. Wearable sensors and real-time analytic techniques have been widely used now to enable real-time monitoring of human operators' neural activity changes. Hence, this presentation will focus on know-how exchange and discussion on one of these emerging techniques, namely functional near infrared spectroscopy (fNIRS) for monitoring brain function. The speaker will share in-depth knowledge on pros and cons of fNIRS techniques and discuss signal processing and biomarker analysis methods required to provide end-users with an opportunity to apply such technologies in more contextually real and dynamic environments.

Multi-channel Non-stationary Biosignal Processing

ABOUT THE SPEAKER:

Dr. Ram Bilas Pachori
Professor, Department of Electrical Engineering, IIT Indore

Dr. Ram Bilas Pachori, Professor, Department of Electrical Engineering, IIT Indore earned his B.E. degree in Electronics and Communication Engineering from Rajiv Gandhi Technological University, Bhopal, India, in 2001, followed by M.Tech. and Ph.D. degrees in Electrical Engineering from IIT Kanpur, India, in 2003 and 2008, respectively. Before joining IIT Indore, he held positions as a Postdoctoral Fellow at the Charles Delaunay Institute, University of Technology of Troyes, France (2007-2008), and as an Assistant Professor at the International Institute of Information Technology, Hyderabad, India (2008-2009). At IIT Indore, he progressed from Assistant Professor (2009-2013) to Associate Professor (2013-2017) before becoming a Professor in 2017. His research interests include signal and image processing, biomedical signal processing, machine learning, and artificial intelligence in healthcare. He has affiliations with the Center for Advanced Electronics, IIT Indore, and has held visiting positions at institutions such as the University of Calabria, Italy, University of Malta, and Free University of Berlin, Germany. He is a Fellow of IETE, IEI, and IET, with 324 publications and eight patents. According to Google Scholar, his publications have an h-index of 68 and approximately 16,000 citations.

ABSTRACT:

In the last one or two decades, adaptive signal decomposition techniques have gained popularity for their broad applicability to almost all fields of science and technology. Empirical mode decomposition has been proposed to decompose the signal into amplitude-frequency modulated components (basis functions). Several methods have been proposed, followed by empirical mode decomposition for adaptive decomposition and to obtain improved signal representation. Empirical wavelet transform (EWT), Fourier-Bessel series expansion-based EWT (FBSE-EWT), iterative filtering, variational mode decomposition are a few popular techniques among adaptive decomposition techniques. Recent advancements in sensor technology make it easier to acquire signals from multiple sources simultaneously, which demands multivariate/multi-channel signal decomposition methods. The univariate iterative filtering has been extended for processing multichannel signals, which will be discussed in this talk. Also, applications of multivariate iterative filtering and machine learning in brain-computer interface and schizophrenia detection from multichannel electroencephalogram (EEG) signals will be presented. The obtained results show the effectiveness of the discussed multivariate/multi-channel adaptive signal decomposition techniques.

Ansys HFSS : Accelerating Engineering & Biomedical Technology Innovations

ABOUT THE SPEAKER:

Dr. Swetha Amit
Senior Application Engineer at Entuple Technologies Private Limited, Bangalore

Dr. Swetha Amit, a Senior Application Engineer at Entuple Technologies Private Limited, Bangalore is a motivated and passionate professional with over 15+ years of academic and Research on design and development of Antennas and RF components. She has worked on projects from Defense Innovation Organisation as part of Defense India Startup Challenge and won two challenges in design of Liquid Antennas and Conformal Antenna in 2021 and 2022. The projects were related to the Indian Navy and Indian Army with nodal office bases from Goa Shipyard Limited (GSL) and Bharat Electronics Limited (BEL) respectively. She was felicitated by Honourable Defense Minister Shri Rajnath Singh for these Defense projects. Her education includes Ph.D in Electronics Engineering from Jain University, Bangalore in 2018, M.Tech in Communication systems from R.V.College of Engineering, Bangalore securing Gold Medal in 2009 and B.E from AIT, Chikmagalur in 2005. She has worked as Assistant Professor in the Department of Electronics and Telecommunication Engineering, M S Ramaiah Institute of Technology, Bangalore from 2012 to 2021. She has developed a state of the art Anechoic Chamber working up to 40GHz with a self developed Fabrication unit for Antenna development and testing. Her research work is on Antenna Design, Wearable and Textile Antenna, SAR analysis and reduction in human body, Liquid antennas, Metamaterials with the extensive use of the Ansys HFSS tool. She has over 37 publications in journals, conferences, book chapters, in addition to funded projects from AICTE, VGST, consultancy from many Industries, guest lectures, Journal and Conference editor. She holds two Indian patents among which one patent is in collaboration with Semi Conductor Lab (SCL), Mohali.

ABSTRACT:

Engineering and biomedical technology innovations are being propelled forward with the aid of advanced simulation software like ANSYS HFSS. An electromagnetic field solver, enables engineers to design and optimize intricate biomedical devices such as implantable antennas for medical telemetry or wearable sensors for health monitoring. Its capabilities in simulating electromagnetic phenomena contribute to the development of innovative medical imaging systems, wireless communication devices for telemedicine, and electromagnetic therapy modalities. The technology emphasizes the exciting way of accelerating engineering and biomedical technology innovations, offering a captivating journey of discovery and impact through Ansys HFSS.

Unleashing the Potential of Deep Learning

ABOUT THE SPEAKER:

Dr. David Belo is a seasoned biomedical engineer specializing in the application of machine learning to biosignal analysis. With a Ph.D. in Biomedical Engineering from Faculdade de Ciências e Tecnologia – Universidade Nova de Lisboa, David's expertise encompasses deep learning applications in biosignal processing, particularly in electrocardiography (ECG) and electromyography (EMG). As the Founder and CEO of SAFE AI [4U], he leads the charge in developing responsible AI solutions for healthcare, emphasizing the ethical synthesis of data. David's diverse experience includes serving as an AI and biosignal expert at NASA's Frontier Development Lab (FDL) and managing projects such as AISym4Med under Horizon Europe. With a focus on promoting responsible AI practices and advancing data synthesis techniques, David is committed to driving innovation at the intersection of technology and healthcare.

Dr. David Belo
Founder and CEO
SAFE AI[4U], Portugal

ABSTRACT:

In an era marked by unprecedented advancements in technology, the convergence of deep learning and healthcare stands poised to revolutionize the landscape of medical diagnostics and treatment. "Unleashing the Potential of Deep Learning" embarks on a journey to explore the transformative power of artificial intelligence in the realm of healthcare. Nevertheless, even though enhanced diagnostic modalities and personalized treatment options hold immense potential, it is imperative to tread cautiously, ensuring that each stride forward is accompanied by a steadfast commitment to ethical and responsible AI deployment.

This presentation will navigate through the complexities of biosignal acquisition, delve into the theoretical constructs underpinning deep learning architectures, and confront the ethical considerations inherent in harnessing the full potential of AI in healthcare. The convergence of deep learning and healthcare has opened up a world of possibilities, enabling us to revolutionize medical science and transform the lives of patients. As we embark on this journey, let us remain steadfast in our commitment to responsible and ethical AI deployment, ensuring that patient welfare remains at the forefront of all our efforts. By leveraging cutting-edge technologies and embracing ethical imperatives, we can create a future where healthcare is not just about treating illnesses, but also about empowering patients to lead healthier, happier lives.

Rethinking Medical Image Compression in the Deep Learning Era

ABOUT THE SPEAKER:

Dr. Debdoot sheet
Assistant Professor, Electrical Engineering, Centre of Excellence, IIT Kharagpur

Dr.Debdoot Sheet is an Assistant Professor of Electrical Engineering and the Centre of Excellence in Artificial Intelligence at the Indian Institute of Technology Kharagpur. He received the BTech degree in electronics and communication engineering in 2008 from the West Bengal University of Technology, Kolkata; MS and PhD degrees from the Indian Institute of Technology Kharagpur in 2010 and 2014 respectively. His current research interests include multimedia compression, medical imaging, machine learning, image and multidimensional signal processing. He is a DAAD alumni when he was a visiting scholar at the Technical University of Munich during 2011-12. He is also recipient of the IEEE Computer Society Richard E. Merwin Student Scholarship in 2012, the Fraunhofer Applications Award at the Indo-German Grand Science Slam in 2012, the GE Edison Challenge 2013, IEM-UEM Distinguished Young Alumni Award 2016, MHRD Samadhan Award 2020, IEM Distinguished Alumni Award 2022. He is a Chartered Engineer, senior member of IEEE, member of MICCAI, life member of IUPRAI and BMESI. He has regularly offered the NPTEL MOOC on Deep Learning for Visual Computing since 2017. He was the organizing chair of the IEEE International Symposium on Biomedical Imaging (ISBI 2022), Kolkata, India, being held for the first time in India and second time in Asia since its inception in 2002.

ABSTRACT:

Limits of compression reached by traditional methods like JPEG or J2K are not sufficient to meet the real-world demand with growing visual data. This growth can be attributed to an increase in the sensor resolution, dynamic range of sensed pixels, frame rate of videos, etc. This has been impacting the ability to store some of the largest image data contributors,medical imaging industry, entertainment multimedia industry, and driving videos which get acquired in automobiles. The bottom line is to achieve the highest compression factor without loss of relevant information in the images. Compression of images from each industry has its own challenges. The prime challenge in medical imaging is that each modality is distinctively characterized by dynamic range and resolution of the signal and its spatial and statistical distribution. Thus, conventional natural image compression algorithms often fail to preserve clinically relevant details in medical images. We will explore how deep learning can be used to overcome these challenges, delving from theoretical foundations to their associated DevOps towards building softwarized Codecs.

Can Machine Learning Help in Assessing Upper-Limb Functioning Using Wearable Sensors?

ABOUT THE SPEAKER:

Dr. Sivakumar Balasubramanian
Professor and Head,
Department of BioEngineering,
Christian Medical College, Vellore

Dr. B. Sivakumar Balasubramanian Professor and Head, Department of Biomedical Engineering, Christian Medical College, Vellore received his Ph.D. in Bioengineering, from Arizona State University, USA. He was a postdoctoral research fellow in the Human Robotics Group, Imperial College, London, UK. Currently, he is the Professor and head of the department of Bioengineering at the Christian Medical College, Vellore. He is also affiliated with the School of Health and Rehabilitation at The University of Queensland, Brisbane, Australia, as an Honorary Associate Professor from Jan 2023. He is the head of the Biological Learning and Rehabilitation (BioRehab) group, which focuses on human motor learning and rehabilitation. His main research interests include development and validation of tools and training methods for delivery of neuro-rehabilitation, and quantitative human analysis.

ABSTRACT:

Assessing the extent of upper limb usage in daily life in patients with stroke is a crucial outcome measure of their rehabilitation process. While wearable sensors offer a means to obtain precise and objective data regarding upper limb usage in everyday activities, complexity arises due to the diverse movement signals captured by these sensors, the unpredictable nature of real-life environments, and the absence of contextual information. Traditional methods have not fared well in accurately measuring upper limb use, while machine learning algorithms have shown promise in solving this important problem. This talk will present some of our recent work in applying and understanding how machine learning algorithms for measuring upper limb use, and talk about the future directions of research to make quantitative, objective sensor-based upper limb functioning assessment a clinical reality.

Overview of the Steps for developing a Medical Device

ABOUT THE SPEAKER:

Dr. Kambiz Pourrezaei
Professor
School of Biomedical Engineering, Science and Health Systems,
Drexel University, Philadelphia, USA

Dr. Kambiz Pourrezaei, Professor, School of Biomedical Engineering, Science and Health Systems, Drexel has research interests in the areas of biomedical and pharmaceutical applications of nano- and micro-technology. He is actively researching near infrared (NIR) imaging of biological tissues for breast cancer and brain imaging. His research also involves the use of microtechnology for studying the attachment of proteins and cells to biomaterial surfaces.

Dr. Kambiz has been working in the area of bio-optics, bio-nanotechnology, and biomedical technology development for several decades and has made many innovative achievements. In addition, he has rich experiences in technology transfer and industrialization. In the past few years, Dr. Kambiz and his group have developed medical devices of which one particular device, Infrascan, is already FDA approved and commercialized. In addition to device development, he has the expertise in applying novel analysis methods to biological dataset for the purpose of increasing the prediction rate. Among several such successful instances, his team was the first group that applied machine learning methods for classification of painful and non-painful responses to noxious stimuli using the non-invasive and inexpensive measurement of hemodynamic response by near infrared spectroscopy (NIRS).

ABSTRACT:

Over the years the process of development of a medical device has reached a level of maturity and certain sequences of steps have been identified as the best practices. While these steps may vary from one country to another the underlying principles remain the same. In this talk I will focus on the process of " design control". I will also briefly discuss the " Bio Design Process" developed at Stanford University. Further, a quick introduction to prototyping, risk management
and standards will be provided.

Tenth International Conference on Biosignals, Images and Instrumentation, SSNCE, Chennai, March 20 – 22, 2024

TRACK 1

Comparison of DL model for retinal diseases classification using OCT images

Madhumithaa.S
Department of Biomedical Engineering
VelTech Rangarajan Dr Sagunthala R&D institute of science and technology
Chennai, India
vtd1152@veltech.edu.in

Dr Masoodhu Banu N.M
Department of Biomedical Engineering
VelTech Rangarajan Dr Sagunthala R&D institute of science and technology
Chennai, India
drmasoodhubanu@veltech.edu.in

Jehosheba Margaret M
Department of Biomedical Engineering
VelTech Rangarajan Dr Sagunthala R&D institute of science and technology
Chennai, India
jehosh17@gmail.com

A.Padmanabha Sarma
Department of Biomedical Engineering
VelTech Rangarajan Dr Sagunthala R&D institute of science and technology
Chennai, India
a.padmanabhasarmaa@veltech.edu.in

Abstract— Age-related macular degeneration (AMD) and Diabetic macular edema (DME) affect the macula of the retina for the person above 50 years. Accurate diagnosis of eye diseases in the initial stages prevents blindness. Manual detection needs the use of trained human experts. Hence, automated models are necessary. In this research paper, a novel approach for detecting retinal conditions is proposed. This involves the progress of three deep convolutional neural network (CNNs) models viz., AlexNet, ResNet50, and InceptionV3. These models are trained to classify Optical Coherence Tomography (OCT) retina images into three macular diseases viz., Choroidal neovascularization (CNV), DME, DRUSEN, and Normal. The publicly available Kermany dataset is apply to test the accuracy of each model. The outcomes show that the AlexNet model achieves an accuracy of 97.2% followed by the InceptionV3 model (95%) and the ResNet50 model (79.8%). This finding can help the Ophthalmologists for easy and fast diagnosis to ensure quality eye care for individuals who are at the risk of vision loss.

Keywords—CNV, DME, AlexNet, ResNet50, InceptionV3, Dataset.

Introduction

The macula, located in the retina and controls central vision, is affected by AMD. There are two types of AMD viz., dry or exudative (DRUSEN) and wet or non-exudative (CNV). For early AMD, drusen (polymorphous material and acellular deposition) is the key biomarker between Bruch's membrane and retinal pigment epithelium (RPE). Early AMD is asymptomatic, however it can progress to CNV, which is the final stage. This causes the irregular progress of blood vessels from the choroid due to drusen accumulation and abnormalities in RPE [1]

Another serious complication of diabetic retinopathy that affects the macula is DME causing damage to the retina. This results in an accumulation of fluid, either inside or outside the cells which in turn leads to a growth in the thickness of the retina [2]. The OCT images of Normal, CNV, DRUSEN and DME are shown in Fig.1

Three techniques are employed to diagnose AMD and DME as Fundus imaging, OCT imaging and Fluorescein angiography. By using a weak coherence interferometer, OCT provides detailed high-resolution transverse tomography of the retina and macula's internal microstructure [3]. Automatic diagnosis should be implemented due to doctor-patient imbalance.

Deep learning (DL) approaches can be employed to conduct automated analysis of DME and AMD. These techniques have been mainly employed for accurate classification and detection of objects in images. Deep learning architectures offer a significant advantage in early-stage disease detection. A well-designed DL algorithm will be useful to accurately detect even mild symptoms that may go unheeded by an ophthalmologist during a routine check-up. This enables the ophthalmologist to initiate the appropriate treatment before the disease progresses.

The core objective of this paper is to classify the three macular diseases and Normal through OCT images by DL networks viz., AlexNet, ResNet50 and InceptionV3Net. Alexnet is fine-tuned only by OCT images while the other two networks are used as pretrained networks by the ImageNet database. The paper is organised as follows. In this paper, existing research on disease classification using OCT images, the deep learning models and the database used to create and evaluate the model are discussed in detail.

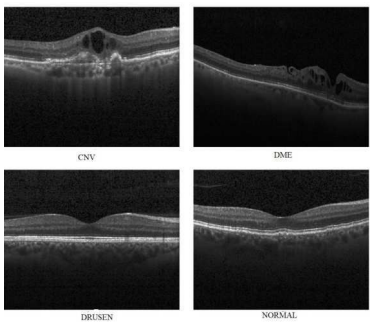


Fig. 1 OCT images of retinal diseases

I. LITERATURE SURVEY

Several studies have been carried out for the automated classification of retinal diseases using OCT images. Srinivasan *et al.,* [4] introduced an algorithm to identify three macular diseases and Normal through OCT

ISBN 979-8-3503-5094-4

imaging. The algorithm employed SVM with gradient descriptor for a fully automated detection process. Wang *et al*.,[16] introduced a CAD model to distinguish between diseases and healthy macula. Deng *et al.,* [5] developed a machine learning technique that could identify AMD and differentiate its various stages using choroidal images acquired OCT. This method extracts texture characteristics with the help of a Gabor filter bank and a non-linear energy transformation. These features were utilised to progress random forests, neural networks and SVM.

Kermany *et al*., implemented a preexisting CNN structure to categorise OCT retinal images to identify diseases through transfer learning [6]. Tayal *et al*., [7] developed a DL model that automatically identifies different types of eye abnormalities using OCT images. This model employed three powerful CNN models with various numbers of layers to accurately recognise different retinal layers, extract valuable information, identify new variations and predict diabetic DME, Drusen, CNV and healthy patient images. Slavesca *et al*.,8] proposed the binary classification and multiclass classification using AlexNet and compared the results with Kermany pre-trained model.

Choudhary *et al*., [9] developed a modified 19 layer deep CNN called VGG-19, coupled with transfer learning, to automatically distinguish three macular diseases. A data augmentation process was performed before training the model. Kayadibi *et al*.,[9] established a FD-CNN with dual preprocessing to diagnose three macular diseases and Normal from OCT images. This method employed a HSR filter to minimise the speckle noise from the OCT images. Once the speckle noise was minimised, the FD-CNN model was trained and the features were extracted. Finally, D-KNN and D-SVM were used to classify the parts.

Researchers led by Ibrahim [10] developed HyCAD-OCT to classify three macular diseases. In this method two preprocessing stages such as denoising B-scans with sparsity-based block matching and reducing speckle noise with BM3D filters were used. A Norm-VGG 16 neural network with a kernel regulariser was employed to segment regions of interest (ROI) and extract 512 features. Manual methods like DAISY and HOG segmented ROI and produced 8,500 features. Hy-CAD combined these features and entered into a three-level dense layer neural network to determine categorisation. Bhadra *et al*.,[11] used a classical process to segment the sphere of interest in OCT retinal images for improved performance. In this method, a binary transformation technique was applied to eliminate artefacts which converted all white pixels to dark ones. Then, a bounding box was placed around the dark pixels to segment the inner area of the OCT image. The result was resized and inputted into a deep multilayered CNN with multiple layers.

II. METHODOLOGY

The Block Diagram illustrates a DL model that employs a simplified approach for an automated early diagnosis of macula diseases in OCT retinal images, as explained in Fig.2

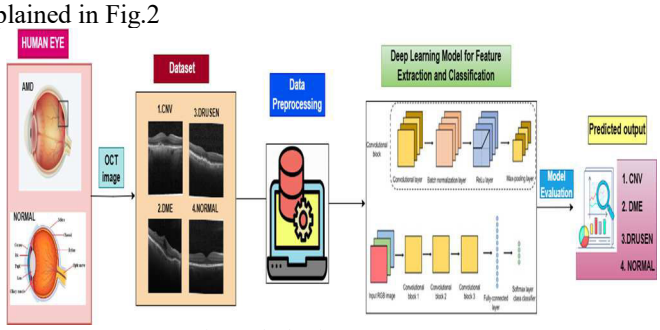


Fig. 2 Block Diagram

The architecture used in this study comprised four modules for diagnosing macular disease. The primary stage involved collecting the OCT image dataset which was then pre-processed in the second stage to enhance the images to high standard. The pre-processed data was then fed into the third stage, a DL model responsible for retrieving features and classifying the three macular diseases and Normal. In the final step, the model was evaluated and the output was predicted. This process can be beneficial for ophthalmologists for fast diagnosis. The detailed processes along with the dataset used and the DL architecture utilised are explained below.

A. Dataset

The labelled OCT dataset and Chest X-ray Images were openly available for research. The OCT images (31,789) were categorised into three folders (train, val, and test) and further split into CNV, DME, DRUSEN, and NORMAL directories. The images were labelled build upon the type of disease, randomised patient ID and image number. A three-tiered grading system was used for labelling the images involving undergraduate medical college students, four ophthalmologists, and two senior retinal specialists [6]. The OCT dataset and workflow of the experimental study were presented in Fig.3

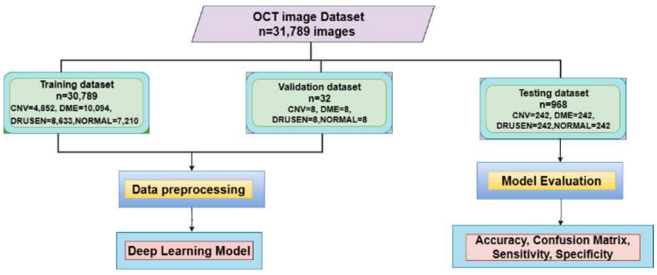


Fig. 3 Dataset and workflow

B. Data preprocessing

To enhance the model's efficiency, it was vital to preprocess the data before training it. Preprocessing was necessary to clean up the dataset to make it compact and get reliable results. The preprocessing steps of the present study was detailed below.

ISBN 979-8-3503-5094-4

- Input data synthetization: All classes of images in the dataset were analysed and found to be variable. Hence all the images were synthesised to a size of 256 x 256 to reduce the error.
- Rescaling: The dataset's image pixels were rescaled to fit in -1 to 1 value size.
- Data Balancing: The class weight argument in Keras assigns different weights to individual classes during training. This argument helped to balance the data of each class, which improved the model'efficiency for the minority class.
- Data Augmentation: Image augmentation was the process of applying different transformations on original images to create multiple copies of the same image. This technique helps to increase the size of the dataset and introduce variability, which enables the model to work on unseen data. Additionally, training on new and slightly altered images makes the model more robust and resilient.

Keras Image Data Generator was one of the easiest augmentation techniques for images with the following advantages.

- ✓ It provides standard augmentation techniques like horizontal flips, random shifts, random zoom, vertical flips and brightness.
- ✓ In the present work, classes = conditions, class_mode = "categorical", color_mode = "rgb", batch_size = 64 for training, 4 for validation, 22 for test dataset and target_size= (256, 256)) were used for augmentation.
- ✓ For each epoch, it provided real-time augmented images, which definitely prevented the model from overfitting.
- ✓ It is a memory saver.

C. *Model Architecture*

This section outlines the structure of the three DL models used in this study i.e AlexNet, ResNet50, and InceptionV3. The utilisation of various DL models is widespread and their application extends to various purposes within neural networks. These deep learning models provide precise and reliable results compared to traditional methods.

a) AlexNet

AlexNet is a CNN structure with five convolutional and three fully connected levels. The first level inputs a 227x227x3 colour image and convolves it using 96 kernels sized 11x11x3. The second to fifth levels convolve the output of the previous level with 256, 384, 384 and 256 kernels respectively. The fully connected (FC) layers consist of 4090 neurons each.[8]. The layers and their functions are described in Table I

TABLE I
LAYERS AND FUNCTION OF THE ALEXNET MODEL

LAYERS	FUNCTIONS
Convolution layer	Perform a convolution filtering operation to the input data.
Pooling layer	Reducing the size and magnitude of the input data
Dropout regularisation	During training, a certain proportion of neurons are randomly excluded.
FC level	A level of FC neurons
Softmax level	Perform a softmax function on the fully connected layer output.

b) ResNet50

Residual Networks, also known as ResNets, are a type of deep convolutional networks that can learn deeper architectures without encountering the vanishing gradient problem that makes it difficult to train very deep networks. Residual connections are implemented in ResNets, which include the result of a convolutional level to the input of the same level, effectively skipping over some of the intermediate layers. This approach allows ResNets to train deeper and more accurate models by preserving the gradient flow throughout the network. It comprises five stages, and each stage contains several residual blocks. A residual block comprises two or three convolutional levels, tailed by a batch normalisation level and an activation function ReLU. The last stage of the network consists of a global average pooling level and an FC level that produces 1024 outputs. These results then go through a softmax layer to determine the final classification.

c)InceptionV3

Google developed Inception V3 for image categorisation developed by Szegedy et al. in 2015[25]. It is unique for its use of factorised convolutions. In this technique, a high level of accuracy is maintained by minimising the quantity of parameters in the architecture. The model's accuracy is improved by factorised convolution which enables the network to learn both local and global features of an image. Inception V3 uses 1x1, 3x3, and 5x5 filters to extract image features while reducing dimensionality with 1x1 filters. Batch normalisation stabilises training and reduces internal covariate shift. [12]

The Inception-V3 neural network begins with 3 convolution level tailed by a max-pooling level, 2 convolution levels, and one more max pooling level. The next network stage employs inception convolution using different filter sizes for each convolution, combines their results and passes them through the network. The subsequent sections of the network are recurring several times, with some sections recurring up to 10 or 20 times towards the outcomes. To prevent overfitting, the architecture uses a stop-learning level that randomly drops weights). Finally, the last level is fully connected [12].

III. IMPLEMENTATION

The present work was implemented on an NVIDIA GPU computer. The computer could handle DL models like AlexNet, ResNet50 and InceptionV3. The programming was executed using Python 3.11 and TensorFlow in Jupyter Notebook. The model was stabilised in 100 epochs.

IV. RESULT AND DISCUSSION

The classification capability was evaluated for normal, CNV, DRUSEN and DME image categories. The DL model was trained using 30,789 OCT images. The images were in JPEG format. OpenCV was used to read them. These comprised 4,852 images with CNV, 10,094 images with DME, 8,633 images with DRUSEN and 7,210 normal images. The remaining 1000 images were split into three sets, 8 images for validation and 242 images for testing each category. Initially, the classification task was performed using AlexNet, and accuracy, sensitivity, specificity, and a confusion matrix were obtained. Then, the same task was done using ResNet50, followed by Inception V3, and the results were recorded.

ISBN 979-8-3503-5094-4

A. AlexNet

A typical AlexNet architecture had 3 fully connected levels, but in the proposed architecture, the fully connected levels were reduced to 2. This provided non-linearity to the model, reduced complexity, avoided fitting issues, prevented loss of spatial information and improved the learning of the feature patterns without affecting the quality of the result. Convolutional levels were used to secure sequential features from the input. Introducing dropout in the convolutional layers avoids overfitting and promotes the learning of diverse spatial hierarchies that make the model perform well for unseen data. This can also prevent overreliance on specific filters thereby resulting in robust representation of features in CNN layers. Using a dropout layer in the fully connected layer results in a 96% accuracy. The model had complied by the stochastic gradient algorithm optimiser with a learning rate (0.0027), categorical cross-entropy loss and accuracy metrics. The computational time of the model was approximately 127s per epoch or 265ms/step. The test accuracy of the AlexNet model was 97.2%, sensitivity 99.1%, specificity 97.4% and AUC score was 0.00041. The model results are shown in Table II, Fig. 4(a) and Fig.4(b).

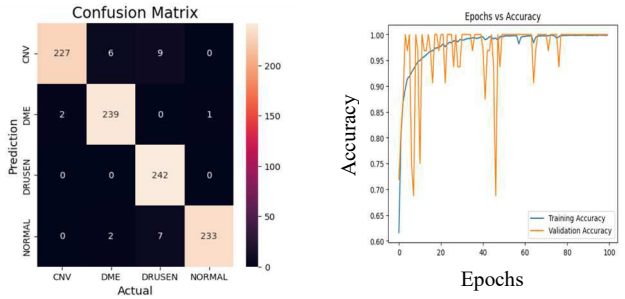


Fig. 4(a) Confusion matrix (b) Model accuracy graph

TABLE II
RESULT OF ALEXNET MODEL.

Class	Precision	Recall	F1 Score	Specificity	Accuracy (%)
CNV	0.99	0.94	0.96	0.99	93
DME	0.97	0.99	0.98	0.98	98
DRUSEN	0.94	1.00	0.97	0.97	100
NORMAL	1.00	0.96	0.98	0.99	96

The presented confusion matrix in Fig. 4(a) reveals that CNV had an accuracy rate of 93%, DME had 98%, DRUSEN had 100%, and normal had a 96% accuracy rate. These values suggest that the model can accurately categorise many images for each condition, with DRUSEN exhibiting the highest accuracy rate.

B. ResNet50

The model utilised transfer learning for image categorisation. The model was with pretrained weights obtained from ImageNet, which had a vast collection of labelled images. Global average pooling was used to lower the size of the feature map, making the network less susceptible to small spatial variations in the input data. The optimiser used to compile the model was Adam, while categorical cross-entropy loss and accuracy metrics were also utilised. The computing time of the model was approximately 148s per epoch or 306ms per step. The test accuracy of the ResNet50 model was 79.4%, sensitivity 79.4%, specificity 93.1% and AUC score was 0.00800. The model's results are shown in Table III, 5(a) and 5(b).

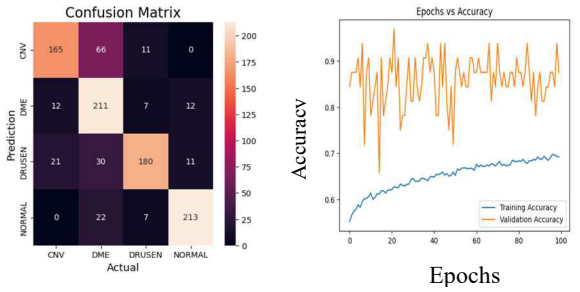


Fig. 5(a) Confusion Matrix (b) Accuracy graph for RESNET50 MODEL

TABLE III
RESULTS FOR THE RESNET50 MODEL

Class	Precision	Recall	F1 Score	Specificity	Accuracy (%)
CNV	0.83	0.68	0.75	0.95	68
DME	0.64	0.87	0.74	0.83	87
DRUSEN	0.88	0.74	0.81	0.96	74
NORMAL	0.90	0.88	0.89	0.96	88

The presented confusion matrix in Figure 5(a) shows that CNV had an accuracy rate of 68%, DME had an accuracy rate of 87%, DRUSEN had an accuracy rate of 74%, and normal had an accuracy rate of 88%.Based on the graph analysis, it can be observed that there was a notable discrepancy between the training and validation accuracy. The training accuracy seems to be high and improving, while the validation accuracy appears to plateau and eventually decrease. This suggests that the model could be too tailored to the training data and not performing well to new data. Further investigation and modifications to the model may be necessary to improve its performance.

C. Inception V3

The transfer learning model with pretrained weights from ImageNet was used. The optimiser used to compile the model was Adam, while categorical cross-entropy loss and accuracy metrics were utilized as well. The computational time of this model was approximately 127s per epoch or 257ms per step. The test accuracy of the Inception V3 model was 95%, sensitivity 99.1%, specificity 88.6% and AUC score was 0.00089. The model's results are presented in Table IV, Fig. 6 (a) and 6(b). The presented confusion matrix in Figure 6(a) shows that CNV had an accuracy rate of 84%, DME had an accuracy rate of 98%, DRUSEN had an accuracy rate of 100%, and normal had an accuracy rate of 99%. From the graph, the validation and training accuracy did not align as expected, indicating an issue with the fitting of the model. Further research may be necessary to identify any adjustments to improve the performance of this model.

TABLE IV
RESULTS FOR THE INCEPTIONV3 MODEL

Class	Precision	Recall	F1 Score	Specificity	Accuracy (%)
CNV	0.99	0.84	0.91	0.95	84
DME	0.90	0.98	0.94	0.83	98
DRUSEN	0.94	1.00	0.97	0.96	100
NORMAL	0.99	0.99	0.99	0.96	99

ISBN 979-8-3503-5094-4

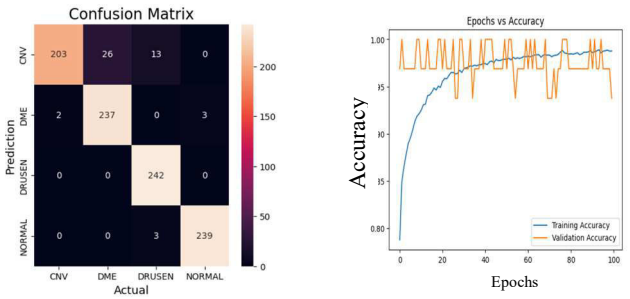


Fig. 6(a) Confusion Matrix (b) Accuracy graph for InceptionV3 model

The three model's performance was scrutinized and compared, taking into account various parameters such as computing timing, accuracy, sensitivity and specificity given in Table V. By carefully assessing each of these factors, this article was able to gain a comprehensive understanding of how each model performed and identify any strengths or weaknesses that may impact their overall effectiveness.

TABLE V
RESULT COMPARISON OF THREE MODELS

Models	Accuracy (%)	Sensitivity (%)	Specificity (%)	AUC score	Computational time (epochs/sec)
AlexNet	97.2	99.1	97.3	0.00041	127
ResNet50	79.4	79.4	93.6	0.00800	148
Inception V3	95	99.1	88.6	0.00089	127

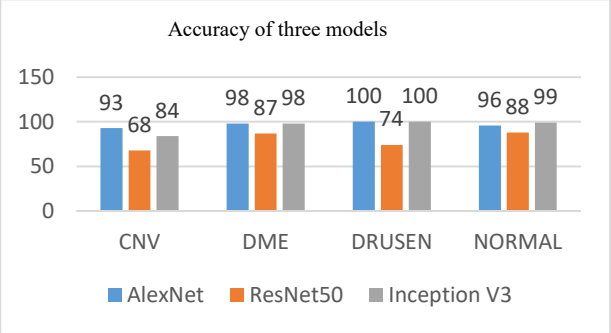


Fig. 7 Accuracy comparison of three models

After comparing the three models, AlexNet demonstrated superior accuracy in classifying CNV, DME, and Drusen (Fig. 7) compared to the other two networks. The size of the parameters for AlexNet was 41M, for ResNet50 25M, and for Inception V3 24M. The computational time for AlexNet, ResNet50 and Inception V3 was 127s, 148s, and 127s for each epoch respectively. Although AlexNet registered high parameters, its computational time was approximately the same as inception V3 and significantly lower than ResNet50. Inception V3 network has an auxiliary classifier, label smoothing and Factorized 7x7 convolutions. One of the benefits of these feature is that they help to proclaim label information to lower layer of the network, resulting in better performance than ResNet50.

V. CONCLUSION

This study compared the three DL models for classifying three macular diseases and Normal. Due to the smaller number of datasets, the study involved transfer learning. AlexNet performed significantly better than the other two transfer learning networks. It registered a superior accuracy of 97.2%, attributed to its advanced neural architecture and efficient learning algorithms, which enable it to make more precise predictions and minimise errors. This finding suggests that AlexNet may be a valuable tool for accurately classifying diseases, which could potentially assist in the creation of treatment plans and enhance patient outcomes. The limited availability of publicly accessible datasets has a negative impact on the progress of research.

References

[1] P. A. Keane, P. J. Patel, S. Liakopoulos, F. M. Heussen, S. R. Sadda, and A. Tufail, "Evaluation of Age-related Macular Degeneration With Optical Coherence Tomography," *Survey of Ophthalmology*, vol. 57, no. 5. pp. 389–414, Sep. 2012. doi: 10.1016/j.survophthal.2012.01.006.

[2] P. Vidal, J. de Moura, J. Novo, and M. Ortega, "Multivendor fully automatic uncertainty management approaches for the intuitive representation of DME fluid accumulations in OCT images," *Med Biol Eng Comput*, May 2023, doi: 10.1007/s11517-022-02765-z.

[3] T. He, Q. Zhou, and Y. Zou, "Automatic Detection of Age-Related Macular Degeneration Based on Deep Learning and Local Outlier Factor Algorithm," *Diagnostics*, vol. 12, no. 2, Feb. 2022, doi: 10.3390/diagnostics12020532.

[4] P. P. Srinivasan *et al.*, "Fully automated detection of diabetic macular edema and dry age-related macular degeneration from optical coherence tomography images," *Biomed Opt Express*, vol. 5, no. 10, p. 3568, Oct. 2014, doi: 10.1364/boe.5.003568.

[5] J. Deng *et al.*, "Age-Related Macular Degeneration Detection and Stage Classification using Choroidal OCT Images." [Online]. Available: http://csvision.swan.ac.uk

[6] D. S. Kermany *et al.*, "Identifying Medical Diagnoses and Treatable Diseases by Image-Based Deep Learning," *Cell*, vol. 172, no. 5, pp. 1122-1131.e9, Feb. 2018, doi: 10.1016/j.cell.2018.02.010.

[7] A. Tayal, J. Gupta, A. Solanki, K. Bisht, A. Nayyar, and M. Masud, "DL-CNN-based approach with image processing techniques for diagnosis of retinal diseases," in *Multimedia Systems*, Springer Science and Business Media Deutschland GmbH, Aug. 2022, pp. 1417–1438. doi: 10.1007/s00530-021-00769-7.

[8] R. R. Slavescu, Universitatea Tehnică din Cluj-Napoca. Computer Science Department, IEEE Romania Section, and Institute of Electrical and Electronics Engineers, *Proceedings, 2018 IEEE 14th International Conference on Intelligent Computer Communication and Processing (ICCP) : Cluj-Napoca, Romania, September 6-8, 2018*.

[9] İ. Kayadibi and G. E. Güraksın, "An Explainable Fully Dense Fusion Neural Network with Deep Support Vector Machine for Retinal Disease Determination," *International Journal of Computational Intelligence Systems*, vol. 16, no. 1, Dec. 2023, doi: 10.1007/s44196-023-00210-z.

[10] M. R. Ibrahim, K. M. Fathalla, and S. M. Youssef, "HyCAD-OCT: A hybrid computer-aided diagnosis of retinopathy by optical coherence tomography integrating machine learning and feature maps localization," *Applied Sciences (Switzerland)*, vol. 10, no. 14, Jul. 2020, doi: 10.3390/app10144716.

[11] R. Bhadra and S. Kar, "Retinal Disease Classification from Optical Coherence Tomographical Scans using Multilayered Convolution Neural Network," in *Proceedings of 2020 IEEE Applied Signal Processing Conference, ASPCON 2020*, Institute of Electrical and Electronics Engineers Inc., Oct. 2020, pp. 212–216. doi: 10.1109/ASPCON49795.2020.9276708.

[12] A. E. Minarno, L. Aripa, Y. Azhar, and Y. Munarko, "INTERNATIONAL JOURNAL ON INFORMATICS VISUALIZATION journal homepage : www.joiv.org/index.php/joiv INTERNATIONAL JOURNAL ON INFORMATICS VISUALIZATION Classification of Malaria Cell Image Using Inception-V3 Architecture." [Online]. Available: www.joiv.org/index.php/joiv

ISBN 979-8-3503-5094-4

A Computer Aided Detection System for Breast Cancer using Lightweight CNN Models for Smart Healthcare

Varun C, Ajay Kumar Reddy Poreddy, Thunakala Balakrishna, and Priyanka Kokil
Advanced Signal and Image Processing (ASIP) Lab,
Department of Electronics and Communication Engineering,
Indian Institute of Information Technology, Design and Manufacturing, Kancheepuram, Chennai-600 127.
Email: {ec20b1113, edm20d012, ec21d0004, priyanka}@iiitdm.ac.in.

***Abstract*—Breast cancer (BC) is a potentially life-threatening disease that occurs because of uncontrolled growth of abnormal corpuscles in the breast tissue. Pathologists analyze the tissue structures using histopathological whole slide images to identify cancerous anomalies. However, pathologists face severe challenges such as fatigue, subjectivity, and inter-observer variability in the early detection of BC. Understanding the intricacies of BC from molecular tissue structures is complex, and inexpertise leads to adverse outcomes. This paper proposes a computed aided detection (CAD) system that can assist histopathologists in the early detection of BC, potentially reducing the abnormalities and diagnostic time. Leveraging the power of convolutional neural networks (CNNs), a stacked ensemble-based model is developed to identify benign and malignant cancerous tissues using histopathological images. The ensemble models comprise three deep CNNs, namely MobileNetV2, ShuffleNet, and SqueezeNet, trained on the BreakHis dataset. Finally, individual CNNs predictions are fed to the average voting-based classifier to identify benign and malignant tissues. The stacked ensemble-based deep CNN model outperformed the individual CNN models in BC prediction, achieving superior accuracy and robustness.**

***Index Terms*—Convolutional neural networks, deep learning, hsitopathology, stacked ensemble.**

I. Introduction

Cancer is one of the most life-threatening diseases. Among many types of cancer, Breast Cancer (BC) has left an indelible mark by significantly impacting the lives of countless women worldwide. According to the World Health Organization (WHO) statistics, 3.19 million new cases of BC will be reported by the end of 2040, and 1.09 million deaths will occur because of BC every year [2]. Figure 1a demarcates the percentages of the global landscape of various cancer diseases in 2020. The figure shows that the number of women suffering from BC is very high compared to other cancers. Figure 1b shows the bar graph demarcating the mortality rates of various cancers in 2020. One may observe from the bar graph that the mortality rate of BC is very high, highlighting its paramount significance in the context of women's health.

BC occurs because of a wide range of changes in the DNA mutations and genetic alterations. These mutations significantly suppress the tumor-controlling genes, leading to the development of BC and other ovarian cancers. The major reasons and the risk factors that cause BC are dense breast tissue, hormonal imbalances, genetic predisposition, reproductive issues, etc. Early detection of BC is important for improving successful treatment and survival chances. The primary method for early detection of BC is to identify the symptoms associated with BC, such as the changes in breast size, shape, nipple alterations, color, texture, etc. The next approach for early detection of BC is to meet the professional health care provider (PHCP) every one to three years in their 20s and annually from the age of forty. The PHCP will conduct clinical breast exams, which help us to identify breast abnormalities.

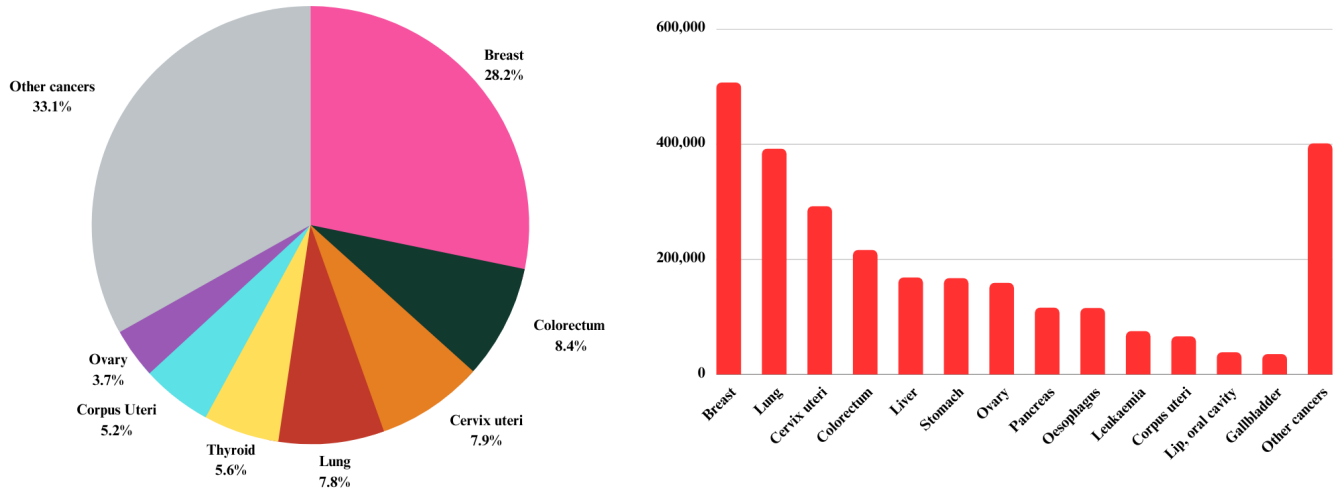
Histopathological images serve as the gold standard for the early identification and detection of BC [3], [4]. Unfortunately, early diagnosis using pathological images is hectic, time-consuming, and requires prior knowledge in detecting the BC. Elmore *et al.* [5] conducted a diagnostic concordance study using breast biopsies and found that the pathologists disagree with each other on an average rate of 24.7% in detecting the BC. This significant rate of misdiagnosis inspired us to develop a computer-aided BC technique to support pathologists in the early detection and diagnosis of BC.

The remaining sections are organized as follows: Section II demonstrates the existing works of BC detection. Section III give details about the proposed framework. Section IV illustrates the results of the paper and its description, and Section V concludes the paper.

II. Related works

Patel *et al.* [6] developed a deep CNN called 'GARL-Net' to classify BC images. The authors extracted the spatial attributes through graph CNN and optimized the features with the Laplacian basis function through an adaptive regularization mechanism. They achieved promising results by introducing the complement cross-entropy as a loss function, which helps to reduce the misclassifications. Wang *et al.* [7] leveraged a generative adversarial network (GAN) to generate images representing different tissue structures. Further, the generated images are transferred to the original dataset and classified using deep CNNs. Li *et al.* [8] developed a CNN architecture

ISBN 979-8-3503-5094-4



(a) (b)

Fig. 1: Visualization of various cancer diagnosis statitisics according to statistics of Cancer Today [1] in 2020: (a) Pie chart representing various cancer diagnosis. (b) Mortality rates caused due to various types of cancers

called 'IDSENet' to classify BC images. The authors integrated the channel attention module with DenseNet architecture to enhance the classification performance. Gupta *et al.* [9] computed the activation features from the final fully connected layer of ResNet50 architecture and fed to logistic regression to classify the BC images. Yamlome *et al.* [10] introduced a modified network based on the AlexNet architecture for image classification. They utilized the transfer learning technique by importing pre-trained layers from the AlexNet model and extended it with additional layers to tailor it to their specific task. The entire network was trained using high-resolution whole images, allowing it to learn and adjust its parameters for improved classification of the particular dataset. Saini *et al.* [11] utilized GAN-based architectures to solve the data imbalance of malignant class. Further, the balanced dataset and augmented images are given to the VGG16 network to classify the BC images. Liu *et al.* [12] developed a Siamese framework combined with an autoencoder to improve the breast cancer classification task. By integrating the Siamese framework with the proficient feature-learning capabilities of the autoencoder, they successfully reduced the inter-class variance between malignant and benign cases, ultimately enhancing the accuracy of breast cancer classification. Yari *et al.* [13] assessed the performance of ResNet50 and DenseNet121 in histological image classification. The author initialized the models with ImageNet weights and fine-tuned them using the BreakHis Dataset. The study involved detailed analysis by optimizing hyperparameters and rigorous evaluations in magnification-dependent and magnification-independent classification scenarios. Toğaçara *et al.* [14] proposed BreastNet, a CNN-based architecture designed explicitly for histopathological image classification. BreastNet incorporates attention mechanisms (CBAM), Residual Blocks, and the hypercolumn technique, which enhanced the model's ability to focus on critical areas, address gradient flow challenges, and capture features at various scales. Yang *et al.* [15] presents a novel approach called "DenseNet-AnoGAN," the system combines AnoGAN for identifying problematic regions and DenseNet for feature extraction to enhance the accuracy of breast cancer histopathological image classification. Using AnoGAN, the model identifies and screens out patches that might lead to misclassification, thus filtering out irrelevant features and improving the classification accuracy. Wang *et al.* [16] developed a feature selection method called dual-network orthogonal low-rank learning (DOLL) and utilized the ensemble technic for the classification of images. The authors extracted selective features with DOLL and trained E-SVM with fused features for classification. Kode *et al.* [17] evaluated three distinct feature extraction methods, which involved using a deep CNN, a knowledge-based approach, and the application of the transfer learning architecture VGG16. The authors utilized these features for classification with seven supervised classifiers for performance evaluation and provided a detailed analysis of the results. Amin *et al.* [18] designed the features agglomeration concept to optimize feature accumulation and utilizes a tree-like structure for efficient feature propagation across layers. The authors employed stain normalization techniques to address color variations in histopathology images, resulting in a comprehensive approach that enhances classification accuracy and network compactness. Amin *et al.* [19] proposed a modified architecture of VGGNet named "CBAM-VGGNet" designed explicitly for histopathology image classification. The authors effectively incorporated the global average pooling (GAP) layer and Convolutional block attention module (CBAM) into the VGGNet to enhance feature extraction capability.

To the best of our knowledge, none of the algorithms in the scientific literature on BC have provided a comprehensive overview of the CNN architectures. In this paper,

ISBN 979-8-3503-5094-4

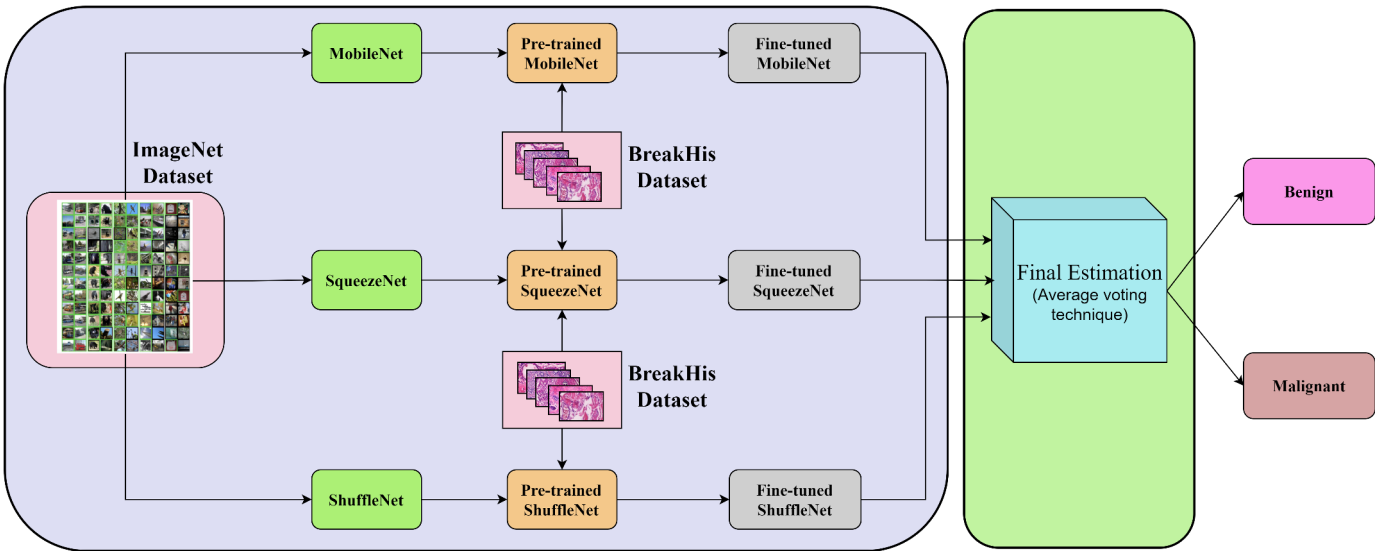


Fig. 2: Framework of the proposed stacked light weight ensemble model for BC detection.

we comprehensively trained all the available CNN models and developed a stacked ensemble model based on the top-performing architectures.

III. Proposed Methodology

The proposed model consists of two steps: In the first step, we train all the deep CNN models on the BreakHis dataset. In the next step, the predictions from the light weight CNN models are given to a average voting classifier module to classify benign and malignant cancer tissues. The framework of the proposed stacked ensemble model is depicted in Figure 2.

TABLE I: Classwise and Microscopic level BC image distributution of BreakHis Dataset [20].

Microscopic level	Benign	Malignant	Total
40 ×	625	1370	1995
100 ×	644	1437	2081
200 ×	623	1390	2013
400 ×	588	1232	1820
Total images	2480	5429	7909
# patients	24	58	82

TABLE II: List of hyperparameters used to fine tune the network.

Hyperparameter	Value
Optimizer	Adam
Batch size	16
# of epochs	50
Learning rate	1×10^{-4}
Regularization parameter	0.0004
Epsilon	0.9
Squared gradient decay factor	0.9

A. Stacked ensemble model

The stacked ensemble model developed for BC detection consists of three light weigh networks, i.e., MobileNet-V2, ShuffleNet, and SqueezeNet. The details of the aforementioned deep CNNs are discussed below.

1) **MobileNet-V2 [21]:** The architecture of MobileNetV2 consists of depthwise convolutional layers followed by 1×1 piecewise convolution layers. It contains inverted residuals along with shortcut and linear connections to improve the visual data flow through the network. These residuals help to capture the complex and hierarchical features of BC tissues and allow the network to avoid gradient overflow during training. The linear bottlenecks of Mobilenet-V2 helps to reduce the learning and trainable parameters, allowing the development of mobile-based applications.
2) **ShuffleNet [22]:** The architecture is well-designed to utilize less computational power without compromising accuracy. It consists of 50 layers in which it effectively leverages the two main operations, the pointwise group convolution and channel shuffling, that maintained the model accuracy with fewer learnable parameters, about 2.3 Million when compared to ResNet-50, which had 25.6 Million learnable parameters. This operation in design helps in parallel processing and cross-group information flow across the network. The ReLU activation function is employed to identify the subtle patterns differentiating benign and malignant tissues.
3) **SqueezeNet [23]:** SqueezeNet distinguishes itself as an exceptionally lightweight network compared to most networks. Using 1x1 convolutions in the network significantly reduces the learnable parameters to 1.2 Million while preserving the important features that make it a compact model. It contains "fire modules" which are a combination of 1x1 and 3x3 convolutions that boost the ability of the model to capture intricate details in tissues. The ReLU activation function helps to identify and differentiate the subtle tissue patterns. Its compact

ISBN 979-8-3503-5094-4

and minimal computational power makes it well-suited for deployment in environments where resources are limited or constrained.

B. Average voting classifier (AVC)

The predictions from the selected lightweight pre-trained models are given to AVC for final prediction. The mathematical description of AVC is as follows:

$$\mathrm{P} = \frac{1}{3}\left(P_1 \oplus P_2 \oplus P_3\right) \tag{1}$$

where P_1, P_2 and P_3 are the prediction scores obtained from MobileNetV2, ShuffleNet and SqueezeNet, and P is the final score obtained after AVC. Finally, the predicted scores are converted to categorical classes, i.e., Benign and Malignant, based on the thresholding mechanism.

IV. Results and Discussions

An automated deep-learning framework for classifying BC using histopathological images was developed. The proposed deep learning framework showed off-the-shelf results on the publicly available BreakHis dataset [20]. The section details the BreakHis dataset and a detailed analysis of the experimental results.

A. BreakHis dataset [20]

BreakHis dataset comprises of 7909 histopathological images collected from eighty-two patients from January 2014 to December 2014. These images are generated from breast biopsy slides containing images with four microscopic levels (40×, 100×, 200×, and 400×). Further, these images are labeled by pathologists and classified into two classes, namely benign and malignant. The details regarding the microscopic levels and class level details of the images are illustrated in Table I.

B. Experimental Details

The proposed model was implemented on MATLAB R2023a Software with system configurations: AMD EPYC 7452 32-core processor along with 128 GB RAM and Linux (64-bit) operating system. Table II provides the list of hyper-model parameters used to fine tune the proposed model and also the comparison models.

C. Performance metrics

Based on the evaluation metrics employed in the various medical image analysis tasks [24]–[26], the performance of the proposed framework for classifying BC using histopathological images is evaluated using the standard metrics: Accuracy, Recall, precision, F1-score, Matthews correlation coefficient (MCC), and Cohen-Kappa score (κ). Accuracy measures the proportion of correctly classified samples to the total number of instances. Precision measures the ratio of correctly classified positive predictions to the total number of positive predictions. Recall measures the model's ability to correctly identify the positive predictions relative to the overall number of positive predictions. F1-score is the harmonic mean of Prec and Rec. MCC is a statistical measure that measures the agreement between the predictions and the actual values based on TP, TN, FP, and FN. MCC value of '1' denotes a good rank of agreement, '0' represents predictions are random, and '-1' implies total disagreement between the prediction and actual values. Cohen-Kappa score (κ) is also a statistical measure that measures the predicted agreement and the actual labels while accounting for the possibility of the agreement occurring by chance.

D. Result analysis

In the preliminary breast cancer classification task analysis, the most widely employed deep learning (DL) models in medical image analysis are investigated. The DL models includes ResNet50, DenseNet, ResNet18, VGG19, ResNet101, VGG16, DarkNet53, InceptionNetV2, InceptionV3, DarkNet19, NasNet, GoogLeNet, AlexNet, MobileNet, SqueezeNet, and ShuffleNet. Among these deep CNN models, ResNet50 has achieved overall superior accuracy of 95.8%, recall of 97.8%, specificity of 94.9%, precision of 89.8%, F1 score of 93.6%, MCC of 90.7%, and Kappa-Cohen's score of 90.5%. specifically in the light-weight networks MobileNet achieved the best results of accuracy 95.2%, recall of 97.2%, specificity of 94.3%, precision of 88.6%, F1 score of 92.7%, MCC of 89.3%, and Kappa score of 89.1%. after MobileNet CNN model ShuffleNet and SqueezeNet given the best results in light-weight network category. Table III summarizes the comprehensive analysis of the breast cancer classification using various frequently utilized pre-trained deep CNN models in terms of widely used classification model metrics such as accuracy, recall, precision, specificity, F1 score, MCC, and kappa score. Table IV provides results of top performers in the lightweight networks

Furthermore, the best-performing light-weight individual CNN model results were combined to enhance the breast cancer classification results and the final results were estimated using the AVC. The proposed ensemble of the individual CNN model estimations produced superior classification performance compared to the individual CNN models. The proposed method achieved an accuracy, recall, specificity, precision, F1 score, MCC, and kappa score of 96.53%, 97.79%, 95.95%, 91.7%, 94.64%, 92.18%, and 92.08%, respectively. Moreover, the confusion matrices of the individual deep CNN model employed in the ensemble method and the proposed method are illustrated in Figure 3. The confusion matrix summarizes the class-wise classification performance of the CNN models, and the diagonal values in the confusion charts represent the accurate prediction by the classification models to the corresponding categories. From the confusion charts depicted in Figure 3, it is perceived that misinterpretations such as malignant as benign and vice versa are minimized, as well as overall classification accuracy has been improved. Therefore, it is noteworthy to state that integrating the individual CNN models can improve the overall classification performance.

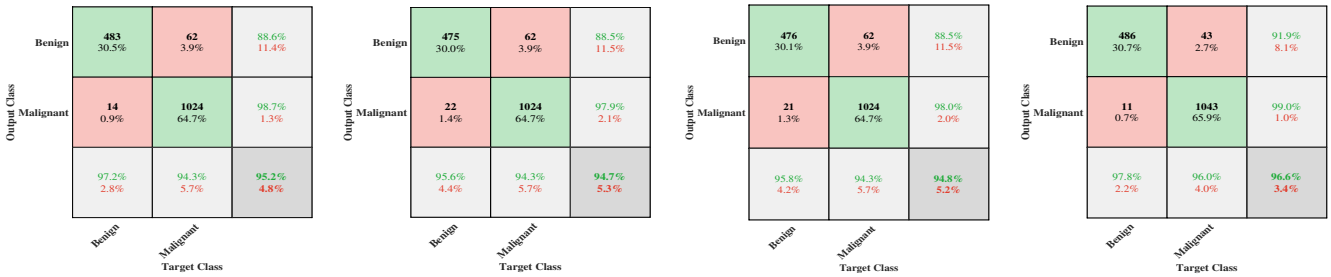
ISBN 979-8-3503-5094-4

TABLE III: Performance results of the deep CNN models in terms of various classification metrics.

Networks	Accuracy (%)	Recall (%)	Specificity (%)	Precision (%)	F1 Score (%)	MCC (%)	Kappa-Cohen's (%)
ResNet50	95.8	97.8	94.9	89.8	93.6	90.7	90.5
DenseNet	95.6	97.9	94.6	89.2	93.4	90.4	90.1
ResNet18	95.3	97.4	94.4	88.8	93.0	89.6	89.4
VGG19	95.3	97.6	94.3	88.7	92.9	89.7	89.4
ResNet101	95.1	97.8	93.9	88.0	92.7	89.3	89.0
VGG16	95.0	97.8	93.7	87.7	92.5	89.1	88.8
DarkNet53	94.9	99.2	93.0	86.6	92.5	89.2	88.7
InceptionNetV2	94.9	93.9	95.3	90.2	92.0	88.3	88.3
InceptionV3	94.7	96.6	93.8	87.8	91.9	88.2	88.0
DarkNet19	93.6	94.8	93.0	86.1	90.2	85.7	85.4
NasNet	94.4	93.8	94.8	89.1	91.4	87.3	87.3
GoogLeNet	93.7	94.6	93.4	86.7	90.5	86.0	85.8
AlexNet	93.4	95.9	92.3	85.0	90.2	85.6	85.3

TABLE IV: Performance results of light weight CNN modules along with stacked light ensemble model using different classification metrics.

Networks	Accuracy (%)	Recall (%)	Specificity (%)	Precision (%)	F1 Score (%)	MCC (%)	Kappa-Cohen's (%)
MobileNet	95.2	97.2	94.3	88.6	92.7	89.3	89.1
SqueezeNet	94.8	95.8	94.3	88.5	91.9	88.2	87.9
ShuffleNet	94.7	95.6	94.3	88.5	91.9	88.1	87.9
Average	**96.53**	**97.79**	**95.95**	**91.7**	**94.64**	**92.18**	**92.08**



(a) (b) (c) (d)

Fig. 3: Confusion charts of (a) MobileNetV2 (b) SqueezeNet (c) ShuffleNet (d) Proposed model.

V. Conclusion

This paper presents a CAD system that can identify BC images into two classes: benign and malignant. The CAD system consists of a stacked ensemble model comprised of three deep CNNs, MobileNetV2, ShuffleNet, and SqueezeNet, and a average voting technique for classifying BC tissues. The BreakHis dataset is used to evaluate the efficiency of the proposed ensemble model and the compared methods. The stacked ensemble model results surpassed the individual CNN models, highlighting its effectiveness in the early detection of BC. This integration of deep CNNs serves as a valuable tool for pathologists to provide enhanced diagnostic support in the early detection of BC.

In the future, we will develop federated learning (FL) based deep CNN models to improve BC detection by aggregating insights from decentralized BC datasets while ensuring the security and privacy concerns of the patients. Incorporation of quality assessment [27]–[29] and different preprocessing techniques that enhance BC's performance classification are possible future tasks to explore.

VI. Acknowledgement

This work was funded by the Department of Science and Technology (DST) under the Fund for Improvement of S&T Infrastructure (FIST), Govt. of India [Grant no. SR/FST/ET-I/2020/578], and Science and Engineering Research Board (SERB) [Grant no. EEQ/2021/000804.]

References

[1] "International Agency for research on Cancer," https://gco.iarc.fr/today/online-analysis-pie, [Online; accessed 2023-11-01].

[2] "Global breast cancer initiative implementation framework: Assessing, strengthening and scaling up of services for the early detection and management of breast cancer: executive summary — who.int," https://www.who.int/publications/i/item/9789240067134, [Accessed 12-10-2023].

[3] M. N. Gurcan, L. E. Boucheron, A. Can, A. Madabhushi, N. M. Rajpoot, and B. Yener, "Histopathological image analysis: A review," *IEEE Reviews in Biomedical Engineering*, vol. 2, pp. 147–171, 2009.

[4] V. Kumar, A. K. Abbas, N. Fausto, and J. C. Aster, *Robbins and Cotran pathologic basis of disease, professional edition e-book*. Elsevier health sciences, 2014.

ISBN 979-8-3503-5094-4

[5] J. G. Elmore, G. M. Longton, P. A. Carney, B. M. Geller, T. Onega, A. N. Tosteson, H. D. Nelson, M. S. Pepe, K. H. Allison, S. J. Schnitt, *et al.*, "Diagnostic concordance among pathologists interpreting breast biopsy specimens," *JAMA*, vol. 313, no. 11, pp. 1122–1132, 2015.

[6] V. Patel, V. Chaurasia, R. Mahadeva, and S. P. Patole, "Garl-net: Graph based adaptive regularized learning deep network for breast cancer classification," *IEEE Access*, vol. 11, pp. 9095–9112, 2023.

[7] D. Wang, Z. Chen, and H. Zhao, "Prototype transfer generative adversarial network for unsupervised breast cancer histology image classification," *Biomedical Signal Processing and Control*, vol. 68, p. 102713, 2021.

[8] X. Li, X. Shen, Y. Zhou, X. Wang, and T.-Q. Li, "Classification of breast cancer histopathological images using interleaved densenet with senet (idsnet)," *PloS one*, vol. 15, no. 5, p. e0232127, 2020.

[9] K. Gupta and N. Chawla, "Analysis of histopathological images for prediction of breast cancer using traditional classifiers with pre-trained cnn," *Procedia Computer Science*, vol. 167, pp. 878–889, 2020.

[10] P. Yamlome, A. D. Akwaboah, A. Marz, and M. Deo, "Convolutional neural network based breast cancer histopathology image classification," in *2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*. IEEE, 2020, pp. 1144–1147.

[11] M. Saini and S. Susan, "Deep transfer with minority data augmentation for imbalanced breast cancer dataset," *Applied Soft Computing*, vol. 97, p. 106759, 2020.

[12] M. Liu, Y. He, M. Wu, and C. Zeng, "Breast histopathological image classification method based on autoencoder and siamese framework," *Information*, vol. 13, no. 3, p. 107, 2022.

[13] Y. Yari, T. V. Nguyen, and H. T. Nguyen, "Deep learning applied for histological diagnosis of breast cancer," *IEEE Access*, vol. 8, pp. 162 432–162 448, 2020.

[14] M. Toğaçar, K. B. Özkurt, B. Ergen, and Z. Cömert, "Breastnet: A novel convolutional neural network model through histopathological images for the diagnosis of breast cancer," *Physica A: Statistical Mechanics and its Applications*, vol. 545, p. 123592, 2020.

[15] R. Man, P. Yang, and B. Xu, "Classification of breast cancer histopathological images using discriminative patches screened by generative adversarial networks," *IEEE access*, vol. 8, pp. 155 362–155 377, 2020.

[16] Y. Wang, B. Lei, A. Elazab, E.-L. Tan, W. Wang, F. Huang, X. Gong, and T. Wang, "Breast cancer image classification via multi-network features and dual-network orthogonal low-rank learning," *IEEE Access*, vol. 8, pp. 27 779–27 792, 2020.

[17] H. Kode and B. D. Barkana, "Deep learning-and expert knowledge-based feature extraction and performance evaluation in breast histopathology images," *Cancers*, vol. 15, no. 12, p. 3075, 2023.

[18] M. S. Amin and H. Ahn, "FabNet: A features agglomeration based convolutional neural network for multiscale breast cancer histopathology images classification," *Cancers*, vol. 15, no. 4, p. 1013, 2023.

[19] A. Ijaz, B. Raza, I. Kiran, A. Waheed, A. Raza, H. Shah, and S. Aftan, "Modality specific CBAM-VGGNet model for the classification of breast histopathology images via transfer learning," *IEEE Access*, vol. 11, pp. 15 750–15 762, 2023.

[20] F. A. Spanhol, L. S. Oliveira, C. Petitjean, and L. Heutte, "A dataset for breast cancer histopathological image classification," *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 7, pp. 1455–1462, 2016.

[21] K. Dong, C. Zhou, Y. Ruan, and Y. Li, "Mobilenetv2 model for image classification," in *2020 2nd International Conference on Information Technology and Computer Application (ITCA)*, 2020, pp. 476–480.

[22] X. Zhang, X. Zhou, M. Lin, and J. Sun, "Shufflenet: An extremely efficient convolutional neural network for mobile devices," in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2018, pp. 6848–6856.

[23] A. Elhassouny and F. Smarandache, "Trends in deep convolutional neural networks architectures: a review," in *2019 International Conference of Computer Science and Renewable Energies (ICCSRE)*, 2019, pp. 1–8.

[24] T. B. Krishna and P. Kokil, "Automated classification of common maternal fetal ultrasound planes using multi-layer perceptron with deep feature integration," *Biomedical Signal Processing and Control*, vol. 86, p. 105283, 2023.

[25] P. Kokil and T. B. Krishna, "Standard fetal ultrasound plane classification based on stacked ensemble of deep learning models," *Expert Systems with Applications*, vol. 238, p. 122153, 2024.

[26] T. B. Krishna and P. Kokil, "Automated detection of common maternal fetal ultrasound planes using deep feature fusion," in *IEEE 19th India Council International Conference (INDICON)*, 2022, pp. 1–5.

[27] A. K. R. Poreddy, P. A. Kara, R. R. Tamboli, A. Simon, and B. Appina, "CoDIQE3D: A completely blind, no-reference stereoscopic image quality estimator using joint color and depth statistics," *The Visual Computer*, pp. 1–11, 2023.

[28] A. K. R. Poreddy, R. B. C. Ganeswaram, B. Appina, P. Kokil, and R. B. Pachori, "No-reference virtual reality image quality evaluator using global and local natural scene statistics," *IEEE Transactions on Instrumentation and Measurement*, vol. 72, pp. 1–16, 2023.

[29] A. K. R. Poreddy, B. V. Atmakuru, T. B. Krishna, P. Kokil, and B. Appina, "Enhancing laparoscopic video quality assessment: A model addressing sensor and channel distortions," *IEEE Sensors Letters*, vol. 8, no. 3, pp. 1–4, 2024.

ISBN 979-8-3503-5094-4

A Multi-input Deep Neural Network Framework for Non-invasive Detection of Anemia using Finger Nail Images

Krithika S*, Ajay Kumar Reddy Poreddy†, Thunakala Balakrishna†, and Priyanka Kokil†
* Department of Computer Science and Engineering,
Amrita Vishwa Vidyapeetham, Chennai-601103, India
†Advanced Signal and Image Processing (ASIP) Lab,
Department of Electronics and Communication Engineering,
Indian Institute of Information Technology, Design and Manufacturing, Kancheepuram, Chennai-600 127.
Email: {krithikas0103@gmail.com}, {edm20d012, ec21d0004, priyanka}@iiitdm.ac.in.

***Abstract*—Anemia, characterized by a deficiency in red blood corpuscles or hemoglobin, poses a significant global health challenge, particularly affecting vulnerable populations. Traditional diagnostic methods often involve invasive procedures, posing challenges in resource-limited settings. This study aims to explore non-invasive anemia detection using fingernail images and convolutional neural networks (CNNs) as a promising alternative to conventional diagnostic approaches. The study utilizes a dataset of fingernail images collected from hospitals in Ghana, comprising both anemic and non-anemic cases. The dataset undergoes preprocessing, including selective enhancement of red components, conversion to the CIELAB color space, and feature extraction. A multi-input Deep Neural Network (DNN) framework employing pre-trained CNNs is proposed for anemia classification. The pre-trained CNN architectures include EfficientNet B1, EfficientNet B4, and MobileNet V3. The framework's performance was assessed using two methodologies: The first involved random shuffling of the dataset, followed by division into training, testing, and validation sets, with evaluation metrics including Accuracy, Precision, F1 scores, and a Confusion Matrix. The second employed five-fold cross-validation, measured using accuracy. The evaluation of the proposed DNN framework using both of the methodologies indicates that EfficientNet B4 achieved the highest testing accuracy (97.87%), precision (97.88%), recall (97.87%), and F1 score (97.88%) and a cross-validation accuracy of 97.37% for the first and second methodologies respectively making it best fit for the proposed DNN framework. The findings demonstrate that the proposed framework yields promising results, especially under the second approach, and opens avenues for further exploration in transfer learning, fine-tuning of deep neural networks for multi-input feature integration, and cross-validation.**

***Index Terms*—Anemia, CNN, EfficientNet, MobileNet, Cross Validation.**

I. Introduction

Anemia is a condition characterized by low red blood cell count or insufficient hemoglobin. Common causes include iron, vitamin deficiencies, chronic diseases, and genetic disorders like sickle cell anemia. Symptoms include paleskin, shortness of breath, fatigue and irregular heartbeat. Treatment varies based on the underlying cause and may involve dietary changes, supplements, or blood transfusions. Anemia can result from chronic infections, bone marrow disorders, and gastrointestinal bleeding. Timely diagnosis through blood tests is crucial for effective management. Addressing the root cause is essential to prevent complications. Pernicious anemia is linked to vitamin B12 deficiency. Inherited conditions like thalassemia and autoimmune disorders can also contribute to anemia. Early intervention and appropriate medical guidance are vital in managing and treating anemia effectively. This condition is especially common in young children in their early years and women undergoing menstruation, pregnancy, or postpartum recovery, accentuating its status as a severe public health concern.

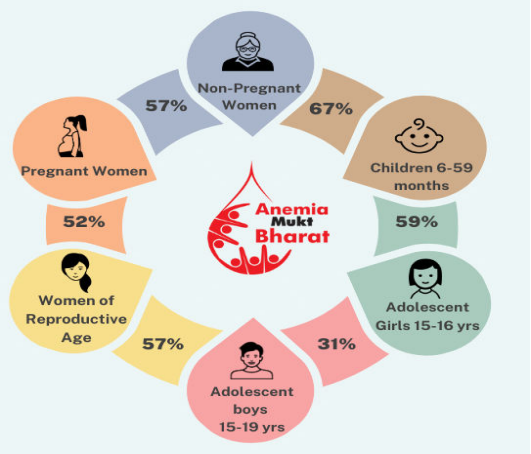


Fig. 1: Statistics of Anemia prevalence in different age groups of the people according to the statistics of NFHS conducted in 2019 [1].

Figure 1 demarcates the anemia prevalence in different age groups of the people in India conducted by the National Family Health Survey (NFHS) in 2019. The statistics clearly show that anemia is prevaling in different age groups of the people starting from children to aged persons. To curb this prevailance, Governement of India initiated a program called '*Anemia mukt Bharath*' in 2019. This program developed an $6 \times 6 \times 6$ strategy involving six groups of people by implementing six

ISBN 979-8-3503-5094-4

interventions in six institutions [1]. Traditionally, diagnosing anemia often involves invasive blood tests, posing risks such as infections and requiring substantial time, labor, and resources. To mitigate these issues, recent trends in medical research have turned towards non-invasive techniques, including machine learning algorithms and deep learning approaches [2]–[4], which hold promise for refining anemia diagnosis. These innovative approaches leverage the analysis of conjunctival pallor and other physical indicators that can be observed rapidly without blood draws. The advent of mobile technology and its application in health surveillance presents new frontiers in anemia screening. With their sophisticated imaging capabilities, smartphones represent a significant breakthrough in public health strategies by providing a convenient and easily accessible way to track possible anemia symptoms. A surge of research has been inspired by this shift toward technology-based solutions, with the goal of developing reliable, non-invasive techniques, utilizing computational analysis to identify and forecast anemia, thereby influencing public health interventions in the future.

In the current study, we propose a multi-input deep neural network (DNN) framework using deep learning that detects anemia from fingernail images. The top-performing model obtained an accuracy of 97% , proving to be a simple, non-invasive, cost-efficient approach for timely detection of anemia.

The rest of the paper is structured as follows: Section II provides the details of the existing literature on the classification task of anemia. The proposed model along with dataset details is discussed in Section III. Section IV discusses the results and discussions of the proposed model.

II. Related Works

Several machine learning algorithms have been introduced for identifying anemia through their cost-effectiveness, user-friendly operation, and non-intrusive procedures. These methods rely on the visual assessment of paleness in tissues that become apparent when hemoglobin levels are low and red blood cell count is reduced. The tissues under observation for these signs are the palpebral conjunctiva, the bed of the fingernails, the palm, and the tongue.

In [5], the authors present a non-invasive method for estimating hemoglobin levels. This method analyzes color and texture from photographs of the patient's inner lower eyelid using a neural network that considers eight distinct features to make its assessment. The technique achieved a sensitivity and specificity of of 71.42% and 89.47% respectively, indicating its effectiveness in identifying anemia. A smartphone application called HemaApp, which noninvasively measures hemoglobin levels in blood using the phone's camera and various light sources, is presented in [6], which achieved high sensitivity (85.7%) and precision (76.5%), comparable to an FDA-approved non-invasive hemoglobin measurement device.

In [7], the utilization of facial images was explored for anemia detection, showcasing a novel approach in clinical diagnostics. The study leveraged deep learning techniques to analyze facial features, potentially offering a non-invasive method for anemia prediction in healthcare settings. A deep learning algorithm (DLA) to non-invasively monitor anemia using Electro-cardiograms is discussed in [8], which achieved an AUROC of 0.923 for 12-lead ECGs and 0.901 during internal and external validation. A comprehensive sensor-based system for monitoring the health of expecting mothers during pregnancy is proposed in [9] that detects non-invasive anemia, glucose rate, heart rate, fall detection, and body temperature monitoring, with Wi-Fi and GSM modules for telemonitoring. The model obtained the highest accuracy of 82.1% and AUROC of 0.824. Hybrid models, GA-SAE and GA-CNN, combined with genetic algorithms with Stacked Autoencoders (SAE) and Convolutional Neural Networks (CNN), are proposed in [10] to address nutritional anemia. The GA-CNN model excelled with an impressive accuracy of 98.50%. A mobile application for detecting anemia in children is presented by [11], where YOLOv5 is used. The system achieved a sensitivity and a specificity of 0.71 and 0.89 respectively. In [12], the authors have developed a device that employs a unique combination of LEDs and fiber optics in a multiwavelength spectrophotometry sensing platform with mechanical lever-operated sensor. The device, augmented with an algorithm, demonstrated an RMSE and a correlation of $1.47 \pm 0.042\, g\, dL^{-1}$ and 0.79 ± 0.03, respectively, between predicted and actual hemoglobin levels. In [13], the authors propose an affordable paper-based sensor coupled with a mobile application for on-the-spot identification of concentration of haemoglobin using a minimal blood sample. The device correlated highly with pathological gold standard results and excellent sensitivity and specificity for detecting healthy and anemic samples. In [14], an autonomous and cost-effective technique to monitor blood hemoglobin levels is proposed that quantifies hemoglobin levels based on color spectroscopy from the paleness of conjunctiva. Classification of anemia is done by comparing the estimated hemoglobin level with the threshold $<11.5\, g\, dL^{-1}$. The model achieved accuracy and sensitivity of $\pm 0.32\, g\, dL^{-1}$ and 89% compared to measured blood hemoglobin levels, respectively. Yet another non-invasive approach to anemia detection is presented in [15], which uses photographs of nails to detect leukonychia, an indicator of iron deficiency anemia, achieving an 89% accuracy in anemia detection with high sensitivity (96%). A similar approach to the non-invasive detection of anemia using AI techniques is discussed in [16], which achieved an accuracy of 97.8%. An affordable smartphone-based Hb analyzer (SHbA) is discussed in [17] that utilizes the phone's built-in ambient light detector to gauge the absorption of colorimetric Hb chemical tests within a compact setup, delivering results in under one minute demonstrating high sensitivity and specificity for anemia diagnosis. In [18] five deep neural network architectures UNet, UNet++, FCN, PSPNet, and LinkNet, are compared against their performance in segmentation of the regions of palpebral conjunctiva wherein LinkNet architecture outperformed its counterparts, achieving accuracy, intersection-over-union (IoU), and Dice score metrics of 94.17%, 90.14%,

ISBN 979-8-3503-5094-4

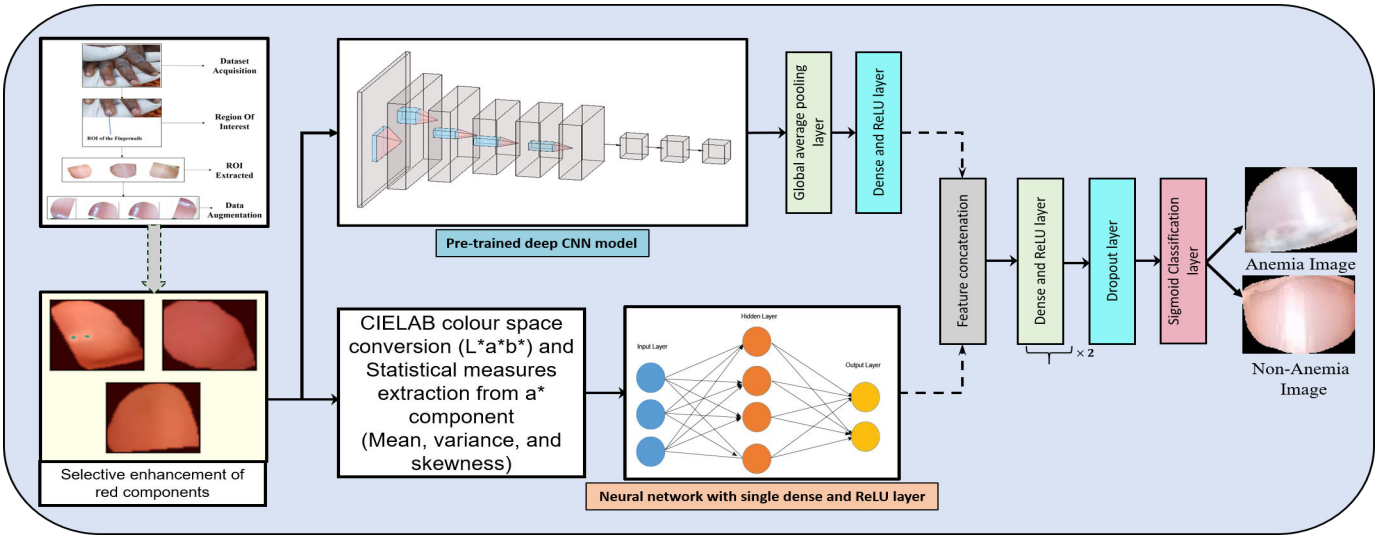


Fig. 2: Block diagram description of the proposed framework.

and 93.78%, respectively. In [19], the researchers propose a an artificial intelligence-enabled computational agent that can efficiently screen individuals for anemia, providing a cost-effective and accessible solution. The system employs segmentation and classification models, achieving an impressive Intersection Over Union (IOU) score of 0.922 for segmentation and high validation accuracy and recall for classification. In [20], a novel deep learning architecture was introduced that concurrently diagnoses anemia and determines hemoglobin concentration from images of the conjunctiva, offering a valuable resource for early anemia detection and treatment in resource-limited settings.

The current study proposes a multi-input deep neural network (DNN) framework that employs pre-trained convolutional neural network (CNN) models for detecting anemia from fingernail images. The framework was evaluated using two methodologies: (i) Random shuffling of the dataset, followed by division into training, testing, and validation sets, with evaluation metrics including Accuracy, Precision, F1 scores, and a Confusion Matrix. (ii) Five-fold cross-validation, measured using accuracy. The findings demonstrate that the proposed framework yields promising results, especially under the second approach, and opens avenues for further research in the current domain.

III. Proposed Methodology

This section includes a description of the dataset, preprocessing of the images acquired, model architecture selection, proposed training, and evaluation framework. Figure 2 describes the workflow of the proposed multi-input DNN framework using pre-trained CNN models for anemia classification.

A. *Dataset*

The dataset [21] used in this study comprises of fingernail images collected from several hospitals in Ghana and is publicly available under the Mendeley Dataset for detecting Anemia. The images were collected from young children aged five and below, and the region of interest was extracted using the threshold triangle approach. Initially, the dataset had 710 images, with 426 being anemic images and 284 non-anemic images. Later, the dataset was expanded using image data augmentation, which produced 4260 images. Of the total images produced, 2565 were anemic, and 1695 were non-anemic.

B. *Image Preprocessing*

The ROI of images was enhanced selectively using Gaussian blur to enhance the red component while blurring the green and blue components. The enhanced images undergo conversion to (CIE $L^*a^*b^*$) color space model. The features pertaining to a^* components: red components $a^*>0$ and green components $a^*<0$ are extracted using mean, standard deviation and skewness of the LAB images. These features along with enhanced images are given as input to the multi-input DNN framework for prediction. We are motivated by [22], [23] to utilize CIELAB color space in the proposed framework.

C. *Model Architecture Selection and DNN Framework*

The preprocessed images were trained using the multi-input DNN framework. The framework receives two feature variables as input - the preprocessed images, and a vector containing features such as mean, standard deviation and skewness pertaining to a* component extracted from the preprocessed images for enhanced feature selection and extraction. The images are fed into an ImageNet pretrained CNN model, followed by a global average pooling layer and a dense ReLu layer for fine-tuning while the numerical a* feature vector is passed through a dense layer with ReLu activation function. The features obtained from both the channels are concatenated and passed through a sequence of two dense layers with Relu activation followed by dropout layer with dropout rate of 0.5 to reduce overfitting. Sigmoid activation was applied to the final layer for classification into the two classes namely Anaemic and Non-anaemic. The framework was individually tested using three different pretrained models. The pretrained

ISBN 979-8-3503-5094-4

CNN architectures used for the proposed framework include ImageNet pre-trained models: EfficientNet B1, EfficientNet B4 from the family of efficientNet models and third - MobileNet V3 , computationally efficient version of MobileNet architecture , a light weight model.

1) ***MobileNet V3***

The MobileNet family of models are known for their lightweight wight architecture ,less complex model size and computationally efficiency. MobileNet V3 was considered the current study. The model employs depthwise separable convolutions like its predecessors for reduced computational load and incorporates squeeze-and-excitation blocks, a lightweight attention mechanism. Despite its compact nature, the model maintains a competitive performance highlighting the model's effectiveness in anemia detection, making it a practical and resource-efficient solution for healthcare applications.

2) ***EfficientNet B1***

Two of the seven architectures of Efficientnet family of models were considered for the current study. The EfficientNet B1 architecture, known for its efficiency and accuracy, comprises multiple convolutional layers with depth-wise separable convolutions and feature scaling using compound scaling techniques. Additionally, it also uses MBConv blocks (Mobile Inverted Bottleneck Convolution), similar to MobileNetV2 but with squeeze-and-excitation optimization thus, offering a good balance between accuracy and efficiency.

3) ***EfficientNet B4***

The EfficientNet B4 architecture is renowned for its depth and complexity, featuring a stack of convolutional layers with depth-wise separable convolutions, multiple expansion and squeeze-and-excitation blocks, and an increased number of parameters compared to its predecessors. The model demonstrates exceptional performance with remarkable efficiency and testing accuracy, precision, recall, and F1 score and cross validation accuracy in anemia classification, despite its enhanced model complexity. This makes the model best suited for high-performance servers or devices where computational resources are not a primary constraint, and maximizing accuracy is key.

D. Proposed training and Evaluation Framework

The multi-input DNN framework was trained using two set of methodologies: (i) In the first approach the dataset was randomly shuffled and was split into three sets - training, validation and test sets. Model training was carried out by using train set and was tested on validation set for improving the performance in each epoch. Reduce learning rate on plateau and early stopping were used to monitor the performance of the model on the validation set in order to prevent overfitting. The weights pertaining best performing model in validation set was saved and was used further on test set for further predictions validating the model's efficiency. The performance of each of the three models were compared using metrics of evaluation such as accuracy, binary cross-entropy loss,precision, recall, F1-score and confusion matrix. (ii) The second approach employs k-fold cross validation on the dataset for obtaining an unbiased estimate of the model's performance. Five fold cross validation was performed for the current study. The entire dataset was randomly divided into five equal-sized subsets. Each of the three models was trained on four subsets with remaining subset as validation set for evaluating the models. This procedure was carried out for five times, with every iteration using a distinct subset for validation. Thus results from each fold were averaged to produce a single estimation. This approach provides a more robust way to assess model performance compared to a single train-test split, especially with limited data, It also helps in reducing bias in the estimation of model performance. The evaluation metrics primarily used for this include Cross validation accuracy and Binary cross-entropy Loss.

IV. Experimental Setup and Results

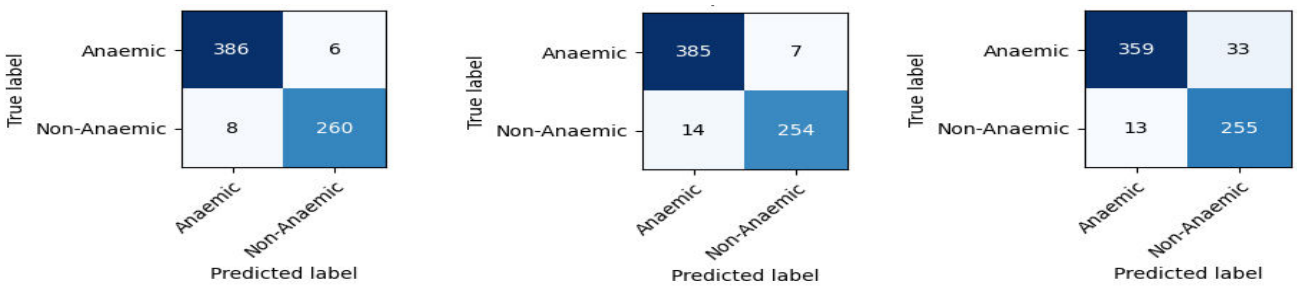
The proposed multi-input Deep Neural Network was trained in Kaggle environment with T4 GPU. The software requirements for implementing the multi-input Deep Neural Network framework utilizing the pretrained CNN architectures - EfficientNet B1, EfficientNet B4, and MobileNet are python 3.10, Keras and tensorflow for leveraging the pretrained CNN architectures and neural network fine-tuning, scikit learn library for leveraging different metrics of evaluation. The neural networks were trained for 30 epochs with a learning rate of 1e-4, batch size of 8. The multi-input Deep Neural Networks utilizing the pretrained CNN models - EfficientNet B4, EfficientNet B1 and MobileNet V3 would be referred to as DNN EffcientNet B4, DNN EfficientNet B1 and DNN MobileNet V3 for brevity. The results of the proposed framework using the two different methods of training and evaluation are as discussed below.

TABLE I: Results of the proposed DNN framework using first methodology against Evaluation Metrics

Model	Accuracy	Precision	Recall	F1-Score
MobileNet V3	93.03	93.27	93.03	93.06
EfficientNet B4	**97.88**	**97.88**	**97.87**	**97.88**
EfficientNet B1	96.82	96.83	96.82	96.81

Table I illustrates the outcomes of the proposed framework employing the first approach, benchmarked against four standard metrics: Accuracy, Recall, Precision, and F1-score. Meanwhile, Figure 3. depicts the corresponding confusion matrices obtained for the three models, providing a visual representation of their performance. Accuracy measures the proportion of correctly classified anemic and non-anemic cases among all instances. Precision indicates the proportion of correctly predicted anemic cases out of all instances predicted as anemic. Recall measures the proportion of actual anemic cases that were correctly identified by the model. The F1 score represents the harmonic mean of precision and recall, providing a balanced measure of model performance. The results clearly illustrate the impact of architectural variations on performance. The DNN EfficientNet B4, characterized by its model design and complexity, exhibited the highest testing accuracy of 97.87%, alongside precision, recall, and F1 score values of 97.88%, 97.87%, and 97.88%, respectively. The confusion matrix further validates this by providing a measure

ISBN 979-8-3503-5094-4



(a) (b) (c)

Fig. 3: Confusion charts of proposed DNN using (a) EfficientNet B4 (b) EfficientNet B1 (c) MobileNet V3.

of True positive, True Negative, False positive and False Negative. EfficientNet B4 3a shows highest True positive score and lowest true negative scores. The exceptional performance of EfficientNet B4 can be attributed to its extensive model capacity, which allows it to capture intricate patterns in the input data. On the other hand, The DNN EfficientNet B1, characterized by the architecture's balance between computational efficiency and model complexity achieved a testing accuracy of 96.82%, demonstrating competitive precision, recall, and F1 score values. The confusion matrix 3b gives more detailed insights that suggest the DNN performs well but with slightly more False Negative than DNN EfficientNet B4 making it bit more conservative in predicting anemia, preferring to minimize False Positive at the risk of increasing FN. In contrast,the deep neural network utilizing MobileNet V3, known for its lightweight design, achieved a testing accuracy of 93.03% with precision, recall, and F1 score values of 93.27%, 93.03%, and 93.06%, respectively. The confusion matrix 3c shows that the model despite achieving comparative True positive and true negative values, has the highest number of False postives indicating a tendency to over-diagnose anemia.

TABLE II: Results of the proposed DNN framework using second methodology against Evaluation Metrics

Model	Metrics	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Effective Score
EfficientNet B4	**Accuracy**	**97.88**	**98.01**	**96.36**	**96.59**	**98.01**	**97.37**
	Loss	**0.08**	**0.04**	**0.16**	**0.12**	**0.14**	**0.11**
EfficientNet B1	Accuracy	95.3	96.12	98.12	97.88	97.06	96.9
	Loss	0.23	0.17	0.04	0.06	0.11	0.12
MobileNet V3	Accuracy	96.47	96.94	97.41	95.18	95.3	96.26
	Loss	0.28	0.19	0.2	0.16	0.12	0.2

Table II, illustrates the results of the proposed framework using the second approach against the metrics: cross validation accuracy and loss. The proposed DNN EfficientNet B4, again surpasses the other models maintaining high consistency across all the folds with an effective cross validation accuracy of 97.37 and loss of 0.11. The DNN Efficient B1 and MobileNet V3 architecture maintain competitive performance with the former achieving effective cross validation accuracy of 96.90%with fold scores ranging between 95.30% and 98.12% and the latter with scores ranging from 95.18% to 97.41%, averaging at 96.26%. This performance, while strong, is slightly lower than that of EfficientNet B4. The effective loss score of DNN EfficientNet B1 indicates a slightly less optimal fit to the dataset due to slightly higher loss values compared to EfficientNet B4. However, the values are still relatively low, showing good model performance. The DNN MobileNet V3, despite having slightly higher loss values, still maintains good accuracy. However, higher loss values compared to both EfficientNet models, suggests that the model may not fit the data as closely as the EfficientNet models inspite of maintaining consistent accuracy.

TABLE III: Comparison of existing studies with the proposed framework

Methodology	**Results**
EfficientNet models [24]	Accuracy: 97.52%
Analysis of patient videos using Deep learning [25]	Sensitivity: 92.52%
Detection of Anaemia from palpebral conjuctiva using CNN [26]	Accuracy: 94%
Proposed DNN Framework	Accuracy: 97.87%

The findings clearly establish EfficientNet B4 as the most accurate model for the proposed DNN framework in the current study owing to its exceptional performance using both the approaches. The exceptional performance of family of EfficientNet models is also validated in [24]. Furthermore, consistent and an enhanced performance of the DNNs utilizing the pre-trained models: EfficientNet B1 and MobileNet V3 models using the second methodology indicate the efficiency of cross validation in assessing the performance of machine learning models paving way for further research and exploration in this field. Table III illustrates comparison of the results of the proposed DNN framework with the existing studies. The exceptional performance of the proposed multi-input DNN framework can be attributed to the transfer learning approach, coupled with enhanced feature extraction using multi-input channel and fine-tuning, thus emphasizing on the efficiency and reliability of the system for anemia detection and other health care applications.They also reiterate the effectiveness of machine learning approaches in non-invasive diagnosis.

V. Conclusion

This study explored the potential of machine learning and medical image analysis to address anemia detection, offering a non-invasive and cost-effective alternative to traditional

ISBN 979-8-3503-5094-4

blood-based diagnostic methods. A multi-input deep neural network framework empowered by pre-trained convolutional neural network models was proposed for classifying anemia based on fingernail images. The framework was trained and evaluated using two methodologies: (i) random shuffling and evaluation on test set using accuracy, precision, recall and confusion matrix (ii) five fold cross validation and evaluation using accuracy and loss. The proposed DNN framework utilizing EfficientNet B4 exhibited exceptional performance using both the methodologies, underscoring its suitability for high-precision applications, albeit with increased computational requirements. The DNN EfficientNet B1 provided a balanced trade-off between accuracy and efficiency, while MobileNet V3 proved adept at resource-constrained scenarios. These findings contribute to the ongoing efforts to combat anemia by harnessing the power of machine learning and medical imaging. Incorporation of quality assessment [27]–[29] and preprocessing modules along with deep CNNs are important future tasks to explore.

VI. Acknowledgement

This work was funded by the Department of Science and Technology (DST) under the Fund for Improvement of S&T Infrastructure (FIST), Govt. of India [Grant no. SR/FST/ET-I/2020/578], and Science and Engineering Research Board (SERB) [Grant no. EEQ/2021/000804.]

References

[1] "Anemia Mukt Bharath," https://anemiamuktbharat.info/, 2023, [Online; accessed 2023-11-21].

[2] T. B. Krishna and P. Kokil, "Automated classification of common maternal fetal ultrasound planes using multi-layer perceptron with deep feature integration," *Biomedical Signal Processing and Control*, vol. 86, p. 105283, 2023.

[3] P. Kokil and T. B. Krishna, "Standard fetal ultrasound plane classification based on stacked ensemble of deep learning models," *Expert Systems with Applications*, vol. 238, p. 122153, 2024.

[4] T. B. Krishna and P. Kokil, "Automated detection of common maternal fetal ultrasound planes using deep feature fusion," in *IEEE 19th India Council International Conference (INDICON)*, 2022, pp. 1–5.

[5] A. A. Khurshid and M. Das, "Image processing to quantitate hemoglobin level for diagnostic support," *Helix-The Scientific Explorer*, vol. 10, no. 01, pp. 228–234, 2020.

[6] E. J. Wang, W. Li, D. Hawkins, T. Gernsheimer, C. Norby-Slycord, and S. N. Patel, "HemaApp: Noninvasive blood screening of hemoglobin using smartphone cameras," in *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, 2016, pp. 593–604.

[7] A. Zhang, J. Lou, Z. Pan, J. Luo, X. Zhang, H. Zhang, J. Li, L. Wang, X. Cui, B. Ji, *et al.*, "Prediction of anemia using facial images and deep learning technology in the emergency department," *Frontiers in Public Health*, vol. 10, p. 3917, 2022.

[8] J.-m. Kwon, Y. Cho, K.-H. Jeon, S. Cho, K.-H. Kim, S. D. Baek, S. Jeung, J. Park, and B.-H. Oh, "A deep learning algorithm to detect anaemia with ECGs: A retrospective, multicentre study," *The Lancet Digital Health*, vol. 2, no. 7, pp. e358–e367, 2020.

[9] T. G. Troyee, M. K. Raihan, and M. S. Arefin, "Health monitoring of expecting mothers using multiple sensor approach:"preg care"," in *2020 2nd International Conference on Advanced Information and Communication Technology (ICAICT)*. IEEE, 2020, pp. 77–82.

[10] S. Kilicarslan, M. Celik, and Ş. Sahin, "Hybrid models based on genetic algorithm and deep learning algorithms for nutritional anemia disease classification," *Biomedical Signal Processing and Control*, vol. 63, p. 102231, 2021.

[11] M. Rivero-Palacio, W. Alfonso-Morales, and E. Caicedo-Bravo, "Anemia detection using a full embedded mobile application with yolo algorithm," in *IEEE Colombian Conference on Applications of Computational Intelligence*. Springer, 2021, pp. 3–17.

[12] R. D. Kumar, S. Guruprasad, K. Kansara, K. R. Rao, M. Mohan, M. R. Reddy, U. H. Prabhu, P. Prakash, S. Chakraborty, S. Das, *et al.*, "A novel noninvasive hemoglobin sensing device for anemia screening," *IEEE Sensors Journal*, vol. 21, no. 13, pp. 15 318–15 329, 2021.

[13] S. K. Biswas, S. Chatterjee, S. Bandyopadhyay, S. Kar, N. K. Som, S. Saha, and S. Chakraborty, "Smartphone-enabled paper-based hemoglobin sensor for extreme point-of-care diagnostics," *ACS sensors*, vol. 6, no. 3, pp. 1077–1085, 2021.

[14] S. Ghosal, D. Das, V. Udutalapally, A. K. Talukder, and S. Misra, "SHEMO: Smartphone spectroscopy for blood hemoglobin level monitoring in smart anemia-care," *IEEE Sensors Journal*, vol. 21, no. 6, pp. 8520–8529, 2020.

[15] S. Das and M. Bhattacharjee, "Image-based sensing of leukonychia for early diagnosis of anemia using a smartphone application," *IEEE Sensors Letters*, vol. 6, no. 11, pp. 1–4, 2022.

[16] S. Chakraborty, K. Kansara, R. Dinesh Kumar, D. Swaminathan, K. Aatre, and S. Acharya, "Non-invasive estimation of clinical severity of anemia using hierarchical ensemble classifiers," *Journal of Medical and Biological Engineering*, vol. 42, no. 6, pp. 828–838, 2022.

[17] Q. Fu, T. Qi, Z. Wu, Y. He, S. Guan, S. Luo, Q. Zhang, W. Luo, W. Xiao, B. Situ, *et al.*, "A portable smartphone-based hemoglobin point-of-care testing platform for accurate anemia diagnostics," *Biosensors and Bioelectronics*, vol. 217, p. 114711, 2022.

[18] S. Dhalla, J. Maqbool, T. S. Mann, A. Gupta, A. Mittal, P. Aggarwal, K. Saluja, M. Kumar, and S. S. Saini, "Semantic segmentation of palpebral conjunctiva using predefined deep neural architectures for anemia detection," *Procedia Computer Science*, vol. 218, pp. 328–337, 2023.

[19] Pallavi, B. Basumatary, R. Shukla, R. Kumar, B. Das, and A. K. Sahani, "A deep learning-based system for detecting anemia from eye conjunctiva images taken from a smartphone," *IETE Technical Review*, pp. 1–13, 2023.

[20] P. Appiahene, K. Chaturvedi, J. W. Asare, E. T. Donkoh, and M. Prasad, "Cp-anemic: a conjunctival pallor dataset and benchmark for anemia detection in children," *Medicine in Novel Technology and Devices*, p. 100244, 2023.

[21] J. W. Asare, P. Appiahene, and E. Donkoh, "Detection of anemia using colour of the fingernails image datasets from ghana," 2022.

[22] A. K. R. Poreddy, P. A. Kara, B. Appina, and A. Simon, "A no-reference 3d virtual reality image quality assessment algorithm based on saliency statistics," in *Optics and Photonics for Information Processing XV*, vol. 11841. SPIE, 2021, pp. 160–169.

[23] A. K. R. Poreddy and B. Appina, "BVRIQE: A completely blind no reference virtual reality image quality evaluator," in *2022 IEEE International Conference on Signal Processing and Communications (SPCOM)*. IEEE, 2022, pp. 1–5.

[24] A. A, A. Amrutesh, G. B. C G, A. R. K P, and G. S, "Efficientnet models for detection of anemia disorder using palm images," in *2023 7th International Conference on Trends in Electronics and Informatics (ICOEI)*, 2023, pp. 1437–1443.

[25] A. Zhang, J. Lou, Z. Pan, J. Luo, X. Zhang, H. Zhang, J. Li, L. Wang, X. Cui, B. Ji, and L. Chen, "Prediction of anemia using facial images and deep learning technology in the emergency department," *Frontiers in Public Health*, vol. 10, 2022. [Online]. Available: https://www.frontiersin.org/articles/10.3389/fpubh.2022.964385

[26] R. Magdalena, S. Saidah, I. D. S. Ubaidah, Y. N. Fuadah, N. Herman, and N. Ibrahim, "Convolutional neural network for anemia detection based on conjunctiva palpebral images," *Jurnal Teknik Informatika (Jutif)*, vol. 3, no. 2, pp. 349–354, Apr. 2022. [Online]. Available: https://jutif.if.unsoed.ac.id/index.php/jurnal/article/view/197

[27] A. K. R. Poreddy, P. A. Kara, R. R. Tamboli, A. Simon, and B. Appina, "CoDIQE3D: A completely blind, no-reference stereoscopic image quality estimator using joint color and depth statistics," *The Visual Computer*, pp. 1–11, 2023.

[28] A. K. R. Poreddy, R. B. C. Ganeswaram, B. Appina, P. Kokil, and R. B. Pachori, "No-reference virtual reality image quality evaluator using global and local natural scene statistics," *IEEE Transactions on Instrumentation and Measurement*, vol. 72, pp. 1–16, 2023.

[29] A. K. R. Poreddy, B. V. Atmakuru, T. B. Krishna, P. Kokil, and B. Appina, "Enhancing laparoscopic video quality assessment: A model addressing sensor and channel distortions," *IEEE Sensors Letters*, vol. 8, no. 3, pp. 1–4, 2024.

ISBN 979-8-3503-5094-4

Machine Learning-Based Early Seizure Detection: A Random Forest Classifier Approach for Pre-Ictal Stage Prediction in Epilepsy

M Anaz Mohammed Vishar
Division of Biomedical Engineering
Karunya Institute of Technology and Sciences,
Coimbatore, Tamil Nadu, India
manaz@karunya.edu.in

Hepsiba. D
Division of Biomedical Engineering
Karunya Institute of Technology and Sciences,
Coimbatore, Tamil Nadu, India
hepsiba@karunya.edu

R Prasanna
Division of Biomedical Engineering
Karunya Institute of Technology and Sciences,
Coimbatore, Tamil Nadu, India
prasannar@karunya.edu.in

L.D. Vijay Anand
Division of Robotics Engineering
Karunya Institute of Technology and Sciences,
Coimbatore, Tamil Nadu, India
vijayanand@karunya.edu

***Abstract*— Abrupt and irregular spikes in the brain's electrical activity are the cause of seizures. This could cause transient aberrant behaviors, sensations, or altered states of awareness or unconsciousness. Movements could include jerky, alternatively tensed arms and legs. One of the most prevalent neurological conditions worldwide, epilepsy affects almost 50 million individuals worldwide, according to a report. Over 25% of epileptic individuals experience uncontrollable seizures even with existing medication and surgical treatment options. Preventing Sudden Unexpected Death in Epilepsy (SUDEP) is aided by early seizure detection. A machine learning methodology for seizure detection was introduced by the proposed model.**

The recognized seizure is diagnosed using the Random Forest Classifier algorithm. Using datasets from the ECG (electrocardiogram) and EEG (electroencephalogram), a pre-trained model is created using supervised and machine learning techniques. Finding seizures in the pre-ictal period is the aim. Using the datasets that are available, the model is trained, demonstrating that there are several causes of epilepsy. The answer is the same despite these variations in the causes. The model predicts and alerts carers, nurses, and hospital administration to an upcoming seizure. For doctors, this early warning is very helpful in enabling timely treatment.

Keywords: Seizure, pre-ictal, epilepsy, SUDEP (Sudden Unexpected Death in Epilepsy), predict

I. INTRODUCTION

This Unintentional and abrupt electrical disruptions in the brain that might alter behaviour, movement, feeling, or awareness are known as seizures. It is brought on by aberrant brain electrical activity, which might have an impact on different bodily parts. The intensity of seizures can vary, ranging from slight confusion to complete convulsions. Numerous medical conditions, such as epilepsy, traumatic brain injuries, brain tumours, infections, fever, and genetic factors, can cause them. Depending on the underlying cause and severity of the seizures, management options for seizures may include medication, surgery, or other therapies. People with drug-resistant epilepsy (DRE) have disruptions in their everyday routines when they experience abrupt seizures. This subgroup, which accounts for 30% of all cases of epilepsy, is not helped by anti-epileptic drugs. This circumstance underscores the need for innovative approaches to control seizures [1-4].

The incorporation of seizure prediction models into a warning system could help reduce potential injury and anxiety caused by the unpredictability of epileptic seizures. Many research has presented seizure prediction systems that try to detect pre-ictal Electroencephalogram (EEG) patterns, which signify the shift from a normal brain state (inter-ictal) to a state of aberrant, hypersynchronous neuronal activity (ictal). The majority of these approaches rely on supervised learning techniques, which necessitate labelled data for each epoch (inter-ictal, pre-ictal, and ictal). As a result, precisely assessing these brain states is critical for the creation of successful seizure prediction models.

The global population currently stands at 783.66 billion individuals. Epilepsy stands out as one of the most widespread neurological disorders globally, impacting the lives of more than 50 million people. Epilepsy sufferers have an up to three time's greater chance of dying before their time than the general population. Based on research findings, an individual with untreated epilepsy may face the risk of fatality at a rate of one per 1,000 affected persons. When it comes to India, there are reportedly around 10 million PWE (Persons with Epilepsy). Additionally, the SMR was 4.25, the mortality rate was 5.73 per 100,000, and according to reports, incidence rates of SUDEP (Sudden Unexpected Death in Epilepsy) in children range from 0.36 to 0.43 per 1000 population per year.

Researchers investigated various machine learning algorithms on EEG recordings, which are routinely used to assess brain activity during seizures. Deep learning methods, including Recurrent Neural Networks (RNNs) and Convolutional Neural Networks (CNNs), have garnered considerable attention due to their ability to extract significant patterns from EEG data and precisely detect instances of seizures [8-11][24][25]. IoT (Internet of Things) device integration has also been critical in real-time seizure detection

ISBN 979-8-3503-5094-4

and monitoring. These devices, often equipped with sensors, capture and transmit data to centralized systems. This enables continuous monitoring of patients' health conditions and the activation of alerts during seizure episodes [12]. To enable remote monitoring and ensure seamless communication between devices, the usage of wireless sensor networks and IoT frameworks has been investigated. Despite advancements, challenges persist in this field. Ensuring the reliability and precision of seizure detection algorithms is of paramount importance, given that false positives and false negatives can have serious consequences for patients. Furthermore, data privacy and security issues associated to IoT devices and patient data must be carefully considered in order to protect sensitive medical information [13]. Overall, the literature review emphasizes the potential of machine learning and IoT technologies in seizure detection and monitoring applications, providing useful insights for researchers and practitioners trying to build effective and practical epilepsy treatment solutions. Continued research and technological improvements are projected to improve these treatments, ultimately benefiting those who have epilepsy [14].

An electrode-equipped functional headband for brain signal collection has been developed. The term "pre-ictal stage of seizure" refers to the ability to anticipate seizures before they occur by the processing of these signals using MATLAB/Python. The ECG and EEG datasets are combined to build a pre-trained multi-model. The normal cardiac variability rate and normal brain signals are used in conjunction with machine learning techniques to train the model. The model also includes the effects of seizures on brain signals and the variability of the heart rate.

After recording the heart variability rate and brain signal with an electrode, the physician measures the patient's brain signals in real time using ten to twenty electrodes known as scalp electrodes. Simultaneously, the ECG reading is taken during the procedure. Then, using IoT, the model evaluates brain and ECG signals to predict the pre-ictal period. If there is a possibility of a seizure, an emergency notice is sent to the medical staff and the carers right away. In this instance, the design is created with a wider scope in mind, not only for hospital patients but also for anyone who have been diagnosed with any kind of seizure.

In this instance, the family members and the ambulance will both receive the warning. ECG and EEG are used in conjunction with each other. An ECG looks at heart-rate signals to identify markers that may indicate an impending seizure. The added advantage of this combination is that in the event of one failing, the other takes over and produces the desired results. It is 100% definite that the individual in question is at danger of experiencing seizures if both of the symptoms point to a seizure-prone nature. The patient's likelihood of experiencing seizures will be shown if one signal is non-positive for seizures but the other suggests that there may be a probability Maintaining close observation during the process is crucial. The patient is deemed safe if none of the two indicators indicates a seizure risk, although monitoring is maintained during the procedure to make sure everything goes according to plan. Including seizure prediction models in a warning system could greatly help physicians manage epileptic patients. Predicting seizures beforehand enables medical professionals to take preventive action, which may save lives [15–19].

II. Related Work

This section displays research papers on electroencephalogram (EEG) data analysis and epilepsy. applying temporal and frequency domain signals, among other traditional and modern techniques [20], to look for epileptic activity in EEG recordings. The study, which employed a simple CNN framework, found that frequency domain signals performed better than time domain signals in both two- and three-class scenarios. Among the study's drawbacks are the huge continuous EEG recordings required for deep learning systems, as well as signal inconsistencies [21].

A novel cross-bispectrum characteristic-based approach for detecting epileptic seizure activity is put forth [22]. Multi-channel intracranial EEG recordings from 21 focal epileptic patients are used in the investigation. With a precise identification of 75 out of 78 seizures and an average detection latency of 9.42 seconds, this technique was proven to be effective. The wavelet transforms applied to individuals with absence seizures to analyse epileptiform discharges in EEG recordings [23]. The paper discusses the significance of frequency bands including delta, theta, alpha, beta, and gamma as well as the conventional visual interpretation of EEG recordings. According to the study, wavelet decomposition works well for precisely identifying and localising the transitory features of slow wave epileptic discharges and the 3 Hz spike in EEG in terms of both time and frequency domains.

The effects of epilepsy on people of all ages are covered, as well as the necessity of receiving the right medical care. The study emphasizes how common epilepsy is and how crucial it is to treat it in order to enhance patient outcomes. The paper suggests a cognitive architecture that gives epileptic patients sophisticated, real-time healthcare services using cloud computing and the Internet of Things. EEG data is transmitted and recorded using smart EEG sensors, and signals are classified as seizure or non-seizure using a cloud-based seizure detection system. By transforming IoT into a cognitive IoT powered by the brain, the suggested architecture hopes to help medical practitioners monitor and help patients. Deep learning-based EEG seizure detection has a sensitivity rate of 93.5 percent and an accuracy rate of 99.2%. Each research. Every study project has pros and cons of its own. To increase the precision and effectiveness of epilepsy diagnosis and treatment, further research is required

III. Proposed System

A. Methodology

The objective is to enhance representation through the analysis and segmentation of EEG and ECG signals. In addition, the goal is to find signals that are classified as normal or suggestive of a seizure by using a random forest classifier to extract features from the segmented data. The EEG and ECG signals will likely need to be preprocessed and filtered in order to remove noise and artefacts. This approach likely entails several phases. The signals are then segmented, or split into smaller, easier-to-manage parts from which properties like frequency, amplitude, and duration can be extracted. After the features have been retrieved, an EEG or ECG signal can be classified as normal or suggestive by training a random forest classifier with a collection of labelled data.

ISBN 979-8-3503-5094-4

It is imperative to guarantee that the data utilised for the classifier's training and testing is free from bias and plagiarism, as this could lead to inaccurate predictions and perhaps harmful medical decisions. Care should be taken in the selection and preparation of the data as well as in the building and validation of the classification model.

B. Dataset

As illustrated in Fig. 1, EEG and ECG signals are commonly utilised to study and diagnose neurological and cardiac issues, respectively. For this work, 1200 samples of EEG and ECG signals were taken from the NIH dataset and the data was divided into two folders for training and testing. The goal is to develop a machine learning model that uses EEG and ECG data to distinguish between people experiencing seizures and those who are not.

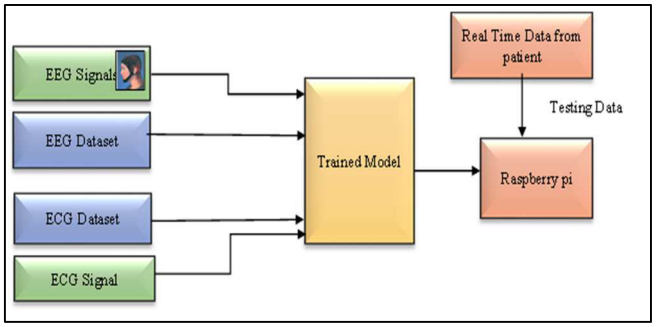


Fig .1. Process of seizure detection

- **Preprocessing:** In the preparation phase, the image will undergo scaling. The preprocessing stage is divided into segmentation preprocessing and classification preprocessing. Techniques applied like noise reduction and cropping to improve the image quality.
- **Segmentation:** Segmentation aims to simplify and modify the interpretation of an image, making it more coherent and easily understandable. The visual image is categorized into different sections that require analysis. Image segmentation results in either a series of segments covering the entire image or a set of contour lines derived from the image.
- **Feature Extraction:** Feature extraction involves identifying essential attributes or characteristics from data. This process, occurring during the data preparation stage, breaks down and condenses a large amount of raw data into smaller pieces, facilitating faster processing. As features, relevant structure in the image, such as lines, edges, or images, could be utilized.
- **Classification:** Before employing RFC for seizure detection, the ECG and EEG data must first be pre-processed utilising filtering, noise reduction, and feature extraction. It is possible to extract statistical parameters including mean, standard deviation, and skewness as well as time- and frequency-domain features. The preprocessed data is then split into testing and training sets. The testing set is used to evaluate the Random Forest Classification's (RFC) performance after it has been trained on the training set.

IV. Results and Discussion

Number Assessing the effectiveness of the proposed approach involves examining the discriminative capabilities of various input parameters through an in-sample classification. To achieve this, different preictal durations (5, 10, 20, 30, 45, and 60 minutes) and channels for each patient were systematically trained and tested using the algorithms. The optimal preictal time lengths for each patient were then determined based on the results obtained, enabling a comparison of the strategy's effectiveness across different input parameters. This process ensures the selection of the most efficient preictal duration and channels for the predictive algorithm.

TABLE I. SAMPLES OF EXAMINATION OF PATIENT FEATURES AND SEIZURES TYPES

Patient	Sex	Age	No. of scalp electrodes	No. of intracranial electrodes	No. of ECG electrodes	Seizures type	Seizures onset region	No. of analyzed seizures	No. of analyzed hours
Patient 1	F	21	27	0	2	CP	f	9	252
Patient 2	M	19	27	0	2	CP	t	6	173
Patient 3	F	22	27	0	4	SP	c	11	102
Patient 4	M	24	27	0	4	UC	f	9	174
Patient 5	M	32	27	0	2	SP	t	5	473

ISBN 979-8-3503-5094-4

These results highlight the peak performance of the prediction algorithm in the preictal phase, which is marked by a very high percentage of accurate classifications. There was no statistically significant variation in the categorization rates during preictal periods between the two recording methods, according to an investigation into whether the prediction algorithm's performance differed depending on the recording methods (scalp or combined intracranial/scalp EEGs). The analysis of how different EEG-ECG channels contribute to the classification process revealed that, for representative patients, the optimal classification rates are obtained when all channels are taken into account at once (Fig. 2). Additionally, a performance analysis of the predictive algorithm utilising only ECG channels was carried out.

After the predictive algorithm was developed, it showed that it could predict seizures using input from one or both datasets. The probability of a seizure occurring is shown by the algorithm through a percentage output. As an early warning system, this information is then sent as an alert to the appropriate medical specialists. By enabling prompt patient preparation and the use of preventive measures, this proactive alert system has the potential to save lives. Medical professionals can foresee seizures and take prompt action to properly manage them by utilising this prediction algorithm.

The effectiveness of these algorithms is assessed in the context of this study for the classification of EEG and ECG signals from epileptic patients. Preictal, interictal, ictal, and postictal phases as well as scalp and mixed intracranial/scalp EEG recording types were all included in the evaluation. Preictal times show notably higher categorization rates, especially when all possible routes are taken into account. When the model is evaluated on future seizures after being trained with preictal data, scalp EEG channels have been shown to have the best classification performances. In most situations, statistical significance was maintained even with reduced rates in out-of-sample classifications; in several cases, it even exceeded half of the values achieved for in-sample classification. There are no discernible relationships between patients' clinical features and good out-of-sample classifications.

The features of the patients' seizures are used to categorise them for the study, as shown in Table 1. There were three categories for seizure types: unclassified (UC), complicated partial (CP), and simple partial (SP). Seizures are classified as frontal (f), temporal (t), or central (c) depending on where they started. The seizure's lateralization was categorised as bilateral (b), right (r), or left (l).

The model's performance for prediction has been assessed using the Random Forest Classifier as a classifier. The outcomes demonstrated a high degree of accuracy, suggesting that the Random Forest Classifier was a useful tool for anticipating epileptic episodes. Table 2 shows the various classifier methods along with their accuracies.

TABLE II. DIFFERENT CLASSIFIER TECHNIQUES WITH ACCURACY

Model	Accuracy %
Logistic Regression	56.05
SVM	91
Random Forest	98.5
Decision Tree	89.76
K Neighbor	90.05
Gradient Boosting	82.7

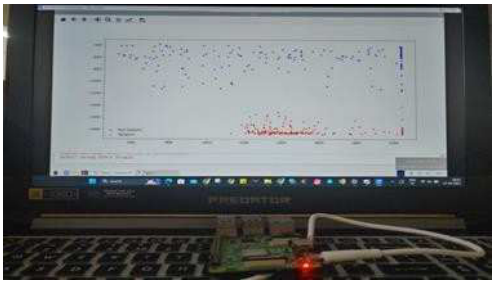


Fig. 2 .Raspberry pi output

Fig. 2 displays the outcomes of the machine learning model that was trained on a Raspberry Pi. By utilizing several EEG-ECG features, the model was able to predict epileptic seizures with a high degree of accuracy. When data from all available channels-scalp, intracranial, and electroencephalogram, or ECG—were taken into account simultaneously for classification, the highest classification rates were achieved. Electroencephalogram (ECG) channels by themselves performed well and, in certain cases, tended to provide the best categorization results. Despite being less than the in-sample results, the out-of-sample classification results nevertheless demonstrated good sensitivity in certain patients and significant rates in 75% of cases.

V. CONCLUSIONS

This paper presents an algorithm tailored to predict epileptic seizures in individual patients by leveraging machine learning and incorporating multiple EEG-ECG features. The algorithm demonstrated notable sensitivity and specificity, especially benefiting 25% of the subjects under study, making it a promising candidate for a seizure prediction alert system. By identifying physiological states linked to increased seizure risk and amalgamating various physiological traits, the algorithm holds potential for deployment as part of a multi-channel, multi-signal intelligent system for epilepsy monitoring in clinical settings. Future directions may include enhancing the algorithm's predictive accuracy and integrating it into wearable devices for real-time seizure prediction and management.

REFERENCES

[1] Sreekumar, Akshay, A. N. Sasidhar Reddy, D. Udaya Ravikanth, M. Chaitanya Chowdary, G. Nithin, and P. S. Sathidevi. "Epileptic Seizure Detection Using Machine Learning Techniques." In Advances in Communication and Computational Technology: Select Proceedings of ICACCT 2019, pp. 919-926. Springer Singapore, 2021.

[2] Siddiqui, Mohammad Khubeb, Ruben Morales-Menendez, Xiaodi Huang, and Nasir Hussain. "A review of epileptic seizure detection using machine learning classifiers." Brain informatics 7, no. 1 (2020): 1-18.

[3] Shoeibi, Afshin, Marjane Khodatars, Navid Ghassemi, Mahboobeh Jafari, Parisa Moridian, Roohallah Alizadehsani, Maryam Panahiazar et al. "Epileptic seizures detection using deep learning techniques: A review." International Journal of Environmental Research and Public Health 18, no. 11 (2021): 5780.

[4] Natu, Milind, Mrinal Bachute, Shilpa Gite, Ketan Kotecha, and Ankit Vidyarthi. "Review on epileptic seizure prediction: machine learning and deep learning approaches." Computational and Mathematical Methods in Medicine 2022 (2022).

[5] Lupión, Marcos, Juan F. Sanjuan, Javier Medina-Quero, and Pilar Martínez Ortigosa. "Epilepsy Seizure Detection Using Low-Cost IoT Devices and a Federated Machine Learning Algorithm." In International Symposium on Ambient Intelligence, pp. 229-238. Cham: Springer International Publishing, 2022.

ISBN 979-8-3503-5094-4

[6] Saminu, Sani, Guizhi Xu, Zhang Shuai, Isselmou Abd El Kader, Adamu Halilu Jabire, Yusuf Kola Ahmed, Ibrahim Abdullahi Karaye, and Isah Salim Ahmad. "A recent investigation on detection and classification of epileptic seizure techniques using EEG signal." Brain sciences 11, no. 5 (2021): 668.

[7] Siddiqui, Mohammad Khubeb, Ruben Morales-Menendez, Xiaodi Huang, and Nasir Hussain. "A review of epileptic seizure detection using machine learning classifiers." Brain informatics 7, no. 1 (2020): 1-18.

[8] Mohammadpoory, Zeynab, Mahda Nasrolahzadeh, and Javad Haddadnia. "Epileptic seizure detection in EEGs signals based on the weighted visibility graph entropy." Seizure 50 (2017): 202-208.

[9] Giannakakis, Giorgos, Vangelis Sakkalis, Matthew Pediaditis, and Manolis Tsiknakis. "Methods for seizure detection and prediction: an overview." Modern Electroencephalographic Assessment Techniques: Theory and Applications (2015): 131-157.

[10] Nicolaou, Nicoletta, and Julius Georgiou. "Detection of epileptic electroencephalogram based on permutation entropy and support vector machines." Expert Systems with Applications 39, no. 1 (2012): 202-209.

[11] Tang, Y., and Dominique M. Durand. "A tunable support vector machine assembly classifier for epileptic seizure detection." Expert systems with applications 39, no. 4 (2012): 3925-3938.

[12] Firpi, Hiram, Erik D. Goodman, and Javier Echauz. "Epileptic seizure detection using genetically programmed artificial features." IEEE Transactions on Biomedical Engineering 54, no. 2 (2007): 212-224.

[13] Mormann, Florian, Ralph G. Andrzejak, Christian E. Elger, and Klaus Lehnertz. "Seizure prediction: the long and winding road." Brain 130, no. 2 (2007): 314-333.

[14] Tzallas, Alexandros T., Markos G. Tsipouras, and Dimitrios I. Fotiadis. "Automatic seizure detection based on time-frequency analysis and artificial neural networks." Computational intelligence and neuroscience 2007 (2007).

[15] Subasi, Abdulhamit. "EEG signal classification using wavelet feature extraction and a mixture of expert model." Expert Systems with Applications 32, no. 4 (2007): 1084-1093.

[16] Mormann, Florian, Thomas Kreuz, Christoph Rieke, Ralph G. Andrzejak, Alexander Kraskov, Peter David, Christian E. Elger, and Klaus Lehnertz. "On the predictability of epileptic seizures." Clinical neurophysiology 116, no. 3 (2005): 569-587.

[17] Alonso-García, María, Rubén Fuente-Alonso, and Juan M. Corchado. "Machine Learning Based System for the Control and Evaluation of Programming Vulnerabilities." In International Symposium on Ambient Intelligence, pp. 179-186. Cham: Springer International Publishing, 2022.

[18] Natu, Milind, Mrinal Bachute, Shilpa Gite, Ketan Kotecha, and Ankit Vidyarthi. "Review on epileptic seizure prediction: machine learning and deep learning approaches." Computational and Mathematical Methods in Medicine 2022 (2022).

[19] Sreekumar, Akshay, A. N. Sasidhar Reddy, D. Udaya Ravikanth, M. Chaitanya Chowdary, G. Nithin, and P. S. Sathidevi. "Epileptic Seizure Detection Using Machine Learning Techniques." In Advances in Communication and Computational Technology: Select Proceedings of ICACCT 2019, pp. 919-926. Springer Singapore, 2021.

[20] Zhou, Mengni, Cheng Tian, Rui Cao, Bin Wang, Yan Niu, Ting Hu, Hao Guo, and Jie Xiang. "Epileptic seizure detection based on EEG signals and CNN." Frontiers in neuroinformatics 12 (2018): 95.

[21] Tzallas, Alexandros T., Markos G. Tsipouras, and Dimitrios I. Fotiadis. "The use of time-frequency distributions for epileptic seizure detection in EEG recordings." In 2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp. 3-6. IEEE, 2007.

[22] Lehnertz, Klaus, and Brian Litt. "The first international collaborative workshop on seizure prediction: summary and data description." Clinical neurophysiology 116, no. 3 (2005): 493-505.

[23] Khan, Y. U., and J. Gotman. "Wavelet based automatic seizure detection in intracerebral electroencephalogram." Clinical Neurophysiology 114, no. 5 (2003): 898-908.

[24] Ilias, Loukas, Dimitris Askounis, and John Psarras. "Multimodal detection of epilepsy with deep neural networks." Expert Systems with Applications 213 (2023): 119010.

[25] Pandya, Vandana, Urvashi P. Shukla, and Amit M. Joshi. "Novel Features Extraction From EEG Signals for Epilepsy Detection Using Machine Learning Model." IEEE Sensors Letters (2023).

ISBN 979-8-3503-5094-4

Integrating Custom GAN Segmentation with Advanced Deep Learning Classifiers for Enhanced Acute Lymphoblastic Detection

Naveen Prashanth G
Department of Electronics Engineering
Madras institute of Technology, Anna University Chennai, India
gnpaone@gmail.com

Lalith Kumar M
Department of Electronics Engineering
Madras institute of Technology, Anna University Chennai, India
kumarlalith0803@gmail.com

Abinaya S
Department of Electronics Engineering
Madras institute of Technology, Anna University Chennai, India
abinaya2srinivasan@gmail.com

Alagu S
Department of Electronics Engineering
Madras institute of Technology, Anna University Chennai, India
alagupugal@gmail.com

***Abstract*— Acute Lymphoblastic Leukemia (ALL) is a highly aggressive form of cancer that originates in the bone marrow and is characterized by uncontrolled growth of immature white blood cells. The proposed work utilizes the various public single-cell datasets and employs Generative Adversarial Networks to generate synthetic images (DCGAN) as well as GAN for semantic segmentation achieving an accuracy of 98.79% for healthy cell segmentation and 99.59% for blast cell segmentation. The work further explores the impact of segmentation on classification accuracy, integrating three classification models (RegNet-Y, ConvNeXt, Vision Transformer). Without segmentation, these models achieve accuracies of 67.16%, 82.19%, and 86.76%, respectively. Using the GAN segmentation outputs, notable improvements in accuracy are observed, with results reaching 72.63%, 93.79%, and 90.77%. This highlights the benefits of combining precise segmentation and robust classification models.**

Keywords—Generative Adversarial Networks (GAN), Deep Convolutional Generative Adversarial Networks (DCGAN), Semantic Segmentation, Transformer, Acute Lymphoblastic Leukemia (ALL).

I. INTRODUCTION

Leukemia is a type of blood cancer that affects the bone marrow and the blood-forming tissues of the body [1]. In this condition there is an uncontrolled proliferation of aberrant white blood cells, which disrupts the natural balance of blood cell production. In a healthy body, cells endure a tightly regulated process of growth, division, and death, ensuring a balance that maintains the normal functioning of tissues and organs. However, in leukemia, this balance is disturbed, leading to the excessive creation of immature white blood cells, known as blasts.

The uncontrolled multiplication of leukemia cells is often a outcome of genetic mutations that accumulate in the affected cells. These mutations can alter the normal governing mechanisms that regulate cell growth and division. In leukemia, the mutated cells often fail to respond to signals that would typically regulate their proliferation or programmed cell death (apoptosis).

Statistics shows that the 5-year survival rate of Acute lymphocytic leukemia (ALL) [2] is 71.3%. Timely detection of leukemia is vital for optimizing treatment outcomes. Recognizing the disease early allows for more effective and less aggressive treatments, increasing the chances of complete remission.

The content is structured as follows: Section 1 presents an overview of leukemia. Section 2 studies previous works conducted for the aim of this work. Section 3 contains the specifics of the suggested methodology. Section 4 describes the implementation and results acquired in this work. Section 5 contains the paper's conclusion and future scope, including the references covered in the last part.

II. RELATED WORKS

Saba Saleem et al. [3] encountered challenges in managing complex models (DarkNet-53 and ShuffleNet), affecting computational efficiency. The proposed methodology addresses this by simplifying the process through the utilization of GAN-generated images, with the potential to alleviate resource demands.

Li Ma et al. [4] utilized ResNet, which exhibits limitations in capturing long-range dependencies. The proposed method, however, opts for Transformer models, thereby enhancing the capability to discern intricate patterns within leukemia cell images. A modified loss function was also proposed to improve model discriminative power of learned features.

Chi-Hung Wei et al. [5] identified opportunities for improvement in classification accuracy. The proposed methodology strategically focuses on achieving high classification accuracy through addition of synthetic images generated by GAN, effectively addressing limitations related to data constraints.

Qile Fan et al. [6] proposed a novel deep learning model called QCResNet to analyze large dataset. Our proposed methodology adopts the Transformer model, surpassing QCResNet in its ability to discern complex structures within peripheral blood smear images and capturing of long-range dependencies and achieving state-of-the-art results.

Fatma M. Talaat and Samah A. Gamel [7] employed Optimized CNN (OCNN) with fuzzy optimization, introducing complexities and hyperparameter tuning dependencies. Conversely, the proposed methodology introduces Transformer models, offering a streamlined development process without extensive tuning, potentially leading to improved overall performance.

III. PROSPOSED METHODOLOGY

The proposed work introduces an innovative method for detecting ALL cells, utilizing a comprehensive dataset of microscopic images depicting both healthy and blast cells. The proposed work is illustrated in Fig. 1. The work involves

ISBN 979-8-3503-5094-4

acquiring and preprocessing images from three datasets, generating synthetic data using DC-GAN, applying a custom GAN for semantic segmentation, and employing top deep learning classifiers for image classification, then a meticulous performance analysis is conducted to compare the results with and without segmentation, enhancing our understanding of model effectiveness. The proposed method is implemented on a high-performance machine, utilizing a CPU: Intel Xeon Gold 5315Y and a GPU: Nvidia RTX A4000. This hardware configuration ensures efficient computation and accelerates the processing of large-scale microscopic images.

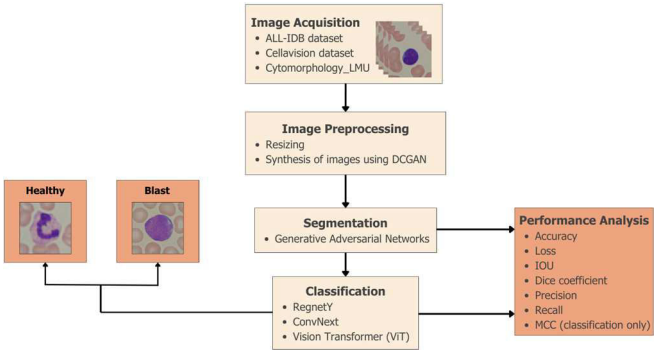


Fig. 1. Block diagram of the proposed method

A. Data acquisition

The ALL-IDB2 dataset [8], used in this work, comprises 130 images per class namely healthy and blast encoded in JPG format with a 24-bit color depth. These images, maintaining a uniform resolution of 257 x 257 pixels, ensure consistent and detailed visual information for comprehensive analysis.

The CellaVision dataset [9] features one hundred 300 x 300 pixels colour images collected from the CellaVision blog. Utilizing the DM96 system with an ocular lens of 22mm for capturing, these images capture distinct cell characteristics, presenting a unique perspective for analysis with a 24-bit color depth and consistent dimensions stored in BMP format.

The Cytomorphology LMU dataset [10], consisting of 18,365 expert-labeled single-cell images, was acquired using a M8 digital microscope/scanner. Images, stored in TIFF format, are 32-bit depth with a size of 400 x 400 pixels, allowing detailed classification of pathological and non-pathological leukocytes at 100-fold optical magnification.

B. Image pre-processing

In the preparation of images for the proposed work, a uniform resizing to 224 x 224 pixels from the original 257 x 257 pixels is done to align with the network requirements. To broaden the dataset, a Deep Convolutional Generative Adversarial Network (DC-GAN) [11] was employed. This process involves training a generator to produce realistic images from random noise or a seed, while a discriminator evaluates the authenticity of these images. The generator effectively transforms noise into data that closely resembles the original images, thereby enhancing the dataset. Through collaborative training, the generator is refined, resulting in the generation of more convincing and lifelike synthetic images. The utilization of DC-GAN proves to be valuable in augmenting the dataset with synthetic images that closely emulate the characteristics of the originals, thereby improving both diversity and data quality.

C. Segmentation

In the pursuit of precise leukemia cell segmentation, this research focuses on semantic segmentation to ensure accurate boundary delineation. This involves categorizing pixels to create distinct regions, which is crucial for defining cell boundaries. To address this, a novel approach is proposed—a custom Generative Adversarial Network (GAN)-based segmentation model. Leveraging GAN principles [12], this model generates synthetic images resembling actual segmented masks present in the ground-truth masks of the images in the dataset, thereby enhancing segmentation accuracy. Ground truth masks are generated by pixelwise color selection using HSV color space [13]. This facilitates the isolation of leukemia cell by creating a mask based on the purple color range within the HSV space. When overlaid on the original image, this mask provides a ground truth representation crucial for the next stages of the research.

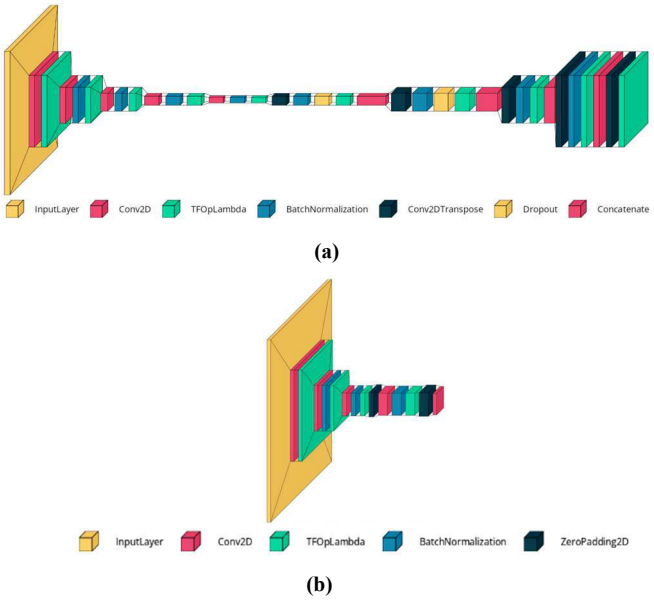


Fig. 2. Architecture of the proposed GAN segmentation model – (a) generator and (b) discriminator architectures

This tailored custom GAN model, illustrated in Fig. 2, is specialized for image segmentation tasks. Consisting of a generator and a discriminator, it employs downsample and upsample blocks within the generator to convert input images into segmented outputs. The downsample blocks capture hierarchical features, while the upsample blocks, with skip connections, restore spatial resolution for detailed segmentation. The generator learns from the input data distribution to create realistic segmented images. Simultaneously, the discriminator distinguishes between real and generated segmented images, enhancing the generator's performance through competitive learning. This GAN framework excels in producing high-quality segmented outputs, showcasing versatility for various segmentation tasks. The model summaries provide insights into the overall design and functionality of this segmentation-focused GAN.

In the process of generating segmented images from original counterparts using their respective custom GAN generated masks, this algorithm employs a series of image processing techniques. Initially, the GAN binary mask is considered and then the regions of interest are isolated through thresholding based on Otsu's method [14]. The formula for Otsu's threshold (t) can be expressed as follows:

ISBN 979-8-3503-5094-4

$$\sigma^2(t) = \omega_{bg}(t)\sigma_{bg}^2(t) + \omega_{fg}(t)\sigma_{fg}^2(t) \quad (1)$$

where $\omega_{bg}(t)$ and $\omega_{fg}(t)$ denotes the probability of number of pixels for each class at threshold t and σ^2 denotes the variance of color values. The weights are given by,

$$\omega_{bg}(t) \text{ or } \omega_{fg}(t) = \frac{P_{BG}(t) \text{ or } P_{FG}(t)}{P_{all}} \quad (2)$$

where P_{all} represents the total count of pixels in an image, $P_{BG}(t)$ is the count of foreground pixels at threshold t, $P_{FG}(t)$ is the count of foreground pixels at threshold t. The variance can be calculated using,

$$\sigma^2(t) = \frac{\sum(x_i - \bar{x})^2}{N-1} \quad (3)$$

The value of x_i represents the pixel at position i within the group (foreground and background). $\bar{x}$ denotes the mean of pixel values in the group (background or foreground), while N represents the total number of pixels.

The threshold obtained by maximizing $\sigma^2(t)$ effectively optimizes the distinction between the foreground and background classes, making it a robust method for automatic image thresholding.

D. Classification

In the comprehensive analysis of leukemia classification, three distinct models—RegNet-Y [15], ConvNeXt [16], and Vision Transformer (ViT) [17] employed, each offering unique insights into the intricate task of disease detection. RegNet-Y demonstrates efficiency and high performance through its architecture, characterized by a set of convolutional layers that gradually identify key information from input images. The inclusion of non-linear activation functions further refines its predictive capabilities. ConvNeXt, a convolutional neural network (CNN), harnesses the power of convolutional and pooling layers to grasp spatial hierarchies and global features, contributing to robust image classification. Vision Transformer (ViT) stands out with its innovative approach, relying on self-attention mechanisms to weigh different parts of input images differently. By dividing images into patches and employing self-attention, ViT excels in capturing long-range dependencies and spatial relationships, marking a paradigm shift in leukemia classification methodologies.

E. Performance analysis

The confusion matrix serves as a cornerstone in the evaluation of classification and segmentation models, delineating predictions into True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN) categories. Accuracy gauges overall correctness, while Cross-Entropy Loss ($H_p(q)$) measures the divergence between predicted and true labels. The Dice Coefficient and Intersection over Union (IoU) evaluate spatial overlap in segmentation. Precision and Recall assess the model's capability to properly recognize positive instances and capture all positives, respectively. F1 Score strikes a balance between precision and recall. Matthews Correlation Coefficient (MCC), considering all elements of the confusion matrix, proves valuable for imbalanced datasets. Together, these metrics provide a complete insight of the nuanced performance of classification and segmentation models.

$$Accuracy = \frac{(TP+TN)}{(TP+FP+TN+FN)} \quad (4)$$

$$H_p(q) = -\frac{1}{N}\sum_{i=1}^{N} y_i \cdot \log(p(y_i)) + (1-y_i) \cdot \log(1-p(y_i)) \quad (5)$$

$$Dice\ Coefficient = \frac{2|A \cap B|}{|A|+|B|} \quad (6)$$

$$IOU = \frac{|A \cap B|}{|A \cup B|} \quad (7)$$

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Positive} \quad (8)$$

$$Recall = \frac{True\ Positive}{True\ Positive + False\ Negative} \quad (9)$$

$$F1\ score = 2 \times \frac{(precision \times recall)}{(precision + recall)} \quad (10)$$

$$\text{MCC} = \frac{(TP \times TN) - (FP \times FN)}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}} \quad (11)$$

IV. Results and Discussion

The research initiative commenced by sourcing images from the datasets, capturing a diverse range of healthy and blast cells, as illustrated in Fig. 3. The subsequent focus shifted towards preparing these images for the network model by resizing them to adhere to specific criteria, refer Fig. 4. To advance the training accuracy, there was a deliberate expansion of the number of input images by generating synthetic counterparts.

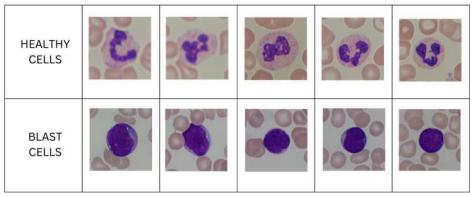


Fig. 3. Sample images from the dataset

Utilizing DCGAN, synthetic images are generated from noise over approximately 500 epochs for both healthy and blast cells, as illustrated in Fig. 5. The continuous refinement over the epochs allowed the generator to capture intricate patterns and details, resulting in realistic representations of both healthy and blast cell images. Fig. 5 visually showcases the outcome of this generative process, highlighting the evolving quality of the synthesized images as the training progressed.

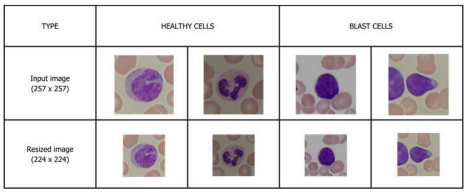


Fig. 4. Resizing of images

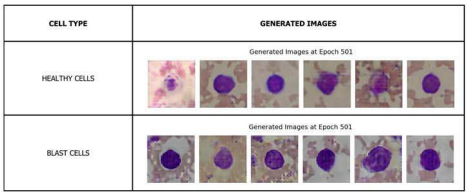


Fig. 5. Generation of images using DC-GAN

ISBN 979-8-3503-5094-4

Addressing challenges in generated image quality, a critical denoising phase explored various filters as in Fig. 6 [18], aiming to find the ideal balance between data integrity and effective smoothing. The Total Variation (TV) Chambolle filter proved to be the optimal choice. Its advantages lie in preserving data integrity while efficiently smoothing images. The filter's ability to strike this balance is evident through its high SSIM (Structural Similarity) score and achieving a superior ENL (Equivalent Number of Looks) score compared to other filters [19] as shown in Table 1, showcasing its proficiency in denoising and enhancing the overall quality of generated images.

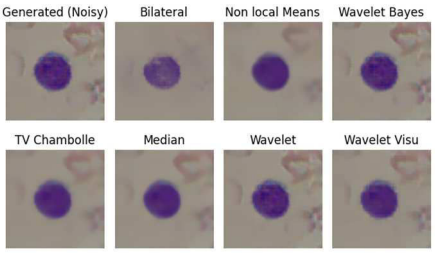


Fig. 6. Denoising of images

TABLE I. SIMILARITY SCORE OF THE FILTERED IMAGES

Filter	SSIM Score[a]	ENL Score
TV Chambolle	80.000%	13.41
Bilateral	40.000%	14.51
Median	60.000%	0.052
Non local means	50.000%	0.052
Wavelet	90.000%	0.051
Wavelet Bayes	90.000%	0.051
Wavelet Visu	40.000%	0.051

[a] SSIM score obtained when compared with DCGAN generated image.

The preprocessing resulted to a substantial increase in the quantity of images, which increased from 4460 to 8460. This increase was mostly caused by the addition of around 4000 photos using the DC-GAN procedure, as shown in Table 2. The subsequent segmentation process was built upon this enlarged dataset. Certain features within the photos were isolated and delineated using the custom GAN. The outcomes, represented in Fig. 7, underwent meticulous correlation analysis against ground truth annotations.

TABLE II. TOTAL NUMBER OF IMAGES AFTER PREPROCESSING

Cell Type	Sample Images	DCGAN Generated	Total
Healthy cells	2230	2000	4430
Blast cells	2230	2000	4430
Total cells	4460	4000	8460

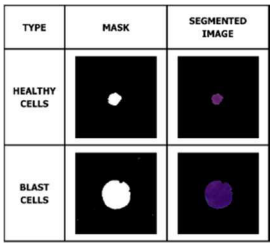
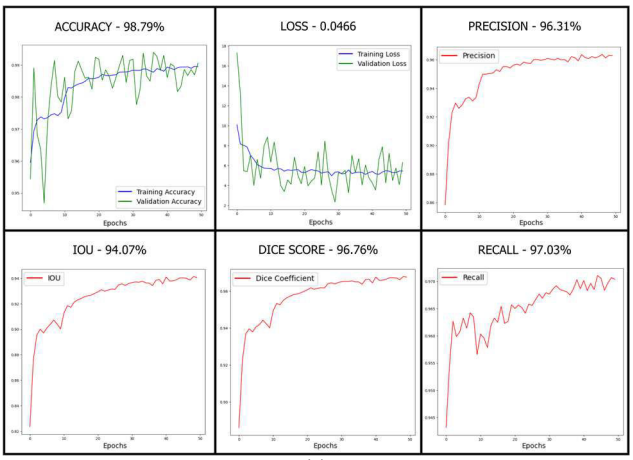


Fig. 7. Segmented image of the proposed custom GAN model generated mask

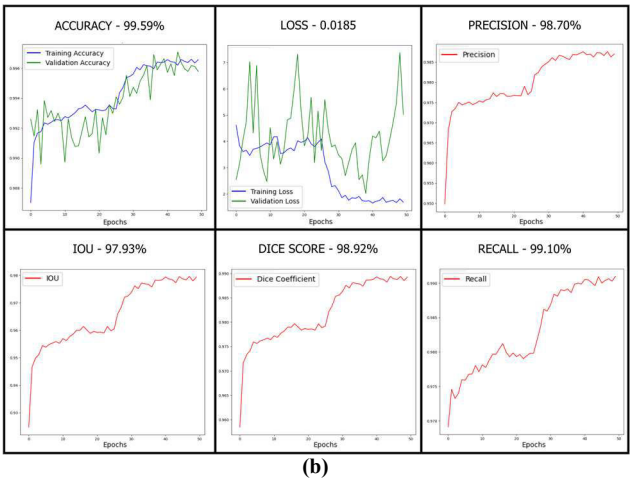
While comprehensively gauging the efficacy of the proposed model, an array of performance metrics was employed, including Accuracy, Loss, IoU, Dice coefficient, Precision, and Recall, each encapsulated within Equation 4,5,6,7,8,9. The segmentation models achieved the following accuracies: Custom GAN-based (Healthy 98.79%, Blast 99.59%), the metric values are tabulated in Table 3. The dynamic interplay of these metrics over epochs was visually presented through performance metric graphs as in Fig. 8, unveiling a pattern of diminishing loss and stabilizing metrics post a certain epoch count.

TABLE III. PERFORMANCE METRICS OF CUSTOM GAN

Metrics	Healthy Cells	Blast Cells
Accuracy	0.9879	0.9959
Loss	0.0466	0.0185
IoU	0.9676	0.9793
Dice	0.9407	0.9892
Precision	0.9631	0.9870
Recall	0.9703	0.9910



(a)



(b)

Fig. 8. Performance metrics graphs for (a) Healthy and (b) Blast cells

With segmented outputs, the focus shifts to classification using three classifiers—RegNet-Y, ConvNeXt, and Vision Transformer (ViT). Initially, classification without segmentation establishes a baseline, and then the same classifiers are tested with segmented images from the custom GAN.

ISBN 979-8-3503-5094-4

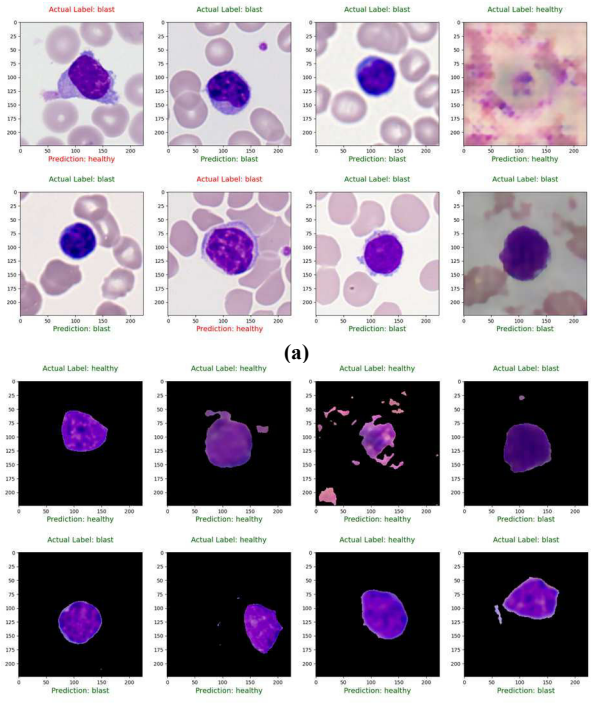


Fig. 9. Classification of healthy and blast cells (a) without segmentation and (b) with segmentation

TABLE IV. PERFORMANCE METRICS FOR CLASSIFICATION OF (A) WITHOUT AND (B) WITH SEGMENTATION CELLS

(A)

Mudel	Accuracy	Precision	Recall	F1 score	MCC
REGNET-Y	0.6716	0.7233	0.6716	0.6559	0.396
CONVNEXT	0.8219	0.8316	0.8219	0.8199	0.651
VIT	0.8676	0.8677	0.8676	0.8677	0.735

(B)

Model	Accuracy	Precision	Recall	F1 score	MCC
REGNET-Y	0.7263	0.7583	0.7263	0.7200	0.486
CONVNEXT	0.9379	0.9380	0.93799	0.9379	0.875
VIT	0.9077	0.9077	0.9077	0.9077	0.815

The classification predictions were depicted in Fig. 9, setting the stage for a meticulous evaluation of performance metrics, including accuracy score, precision score, recall score, F1 score, and Matthews correlation coefficient as in Equations 4,8,9,10,11. The culmination of this detailed analysis was encapsulated in Table 4.

The accuracy and loss trajectories of each model, both with and without segmentation, were visually represented in Fig. 10, Fig. 11, and Fig. 12. A comprehensive assessment was further facilitated through a radar chart as in Fig. 13 painting a vivid picture of the models' strengths and weaknesses. The absence of segmentation revealed accuracies of 67.16%, 82.19%, and 86.76% for the respective models. However, the introduction of the GAN-based segmentation outputs ushered in substantial accuracy boosts: a 5.47% increase for RegNet-Y (72.63% accuracy), an 11.60% increase for ConvNeXt (93.79% accuracy), and a 4.01% increase for Vision Transformer (ViT) (90.77% accuracy).

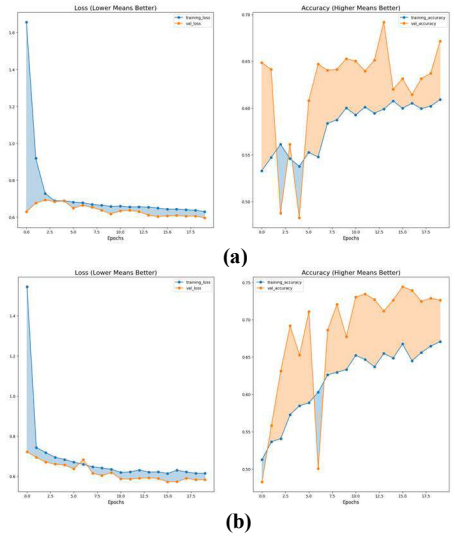


Fig. 10. Regnet-Y loss and accuracy plots for (a) without and (b) with segmentation

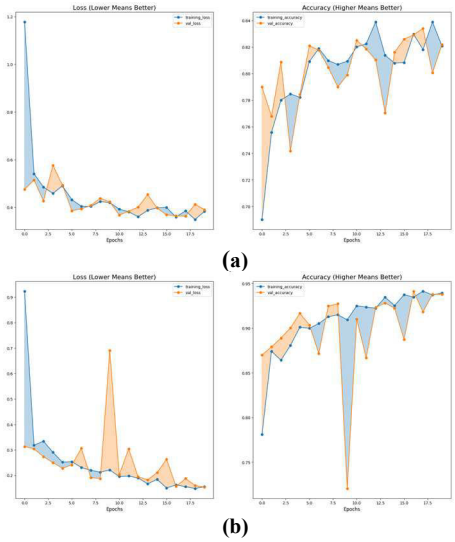


Fig. 11. ConvNeXt loss and accuracy plots for (a) without and (b) with segmentation

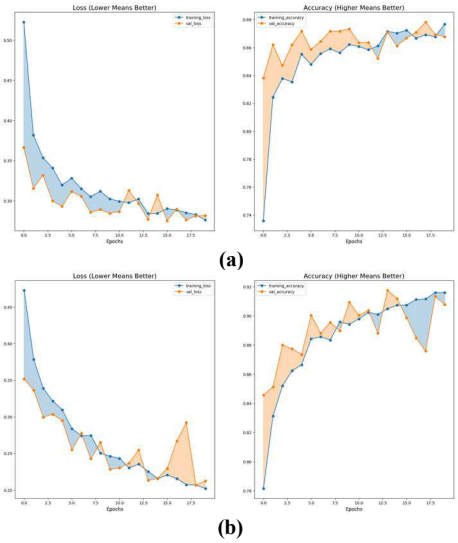


Fig. 12. ViT loss and accuracy plots for (a) without and (b) with segmentation

ISBN 979-8-3503-5094-4

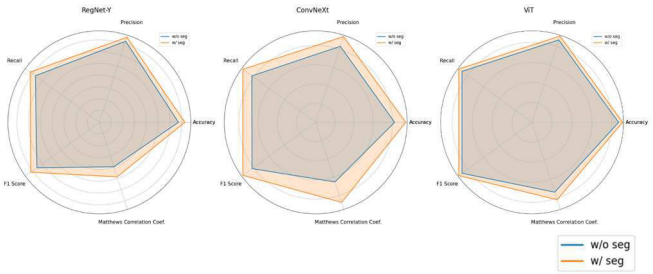


Fig. 13. Radar plots of the performance metrics for without and with segmentation

The diagnostic power of these models was further assessed through ROC curves as in Fig. 14 shedding light on the discriminatory capabilities, with RegNet exhibiting a relatively weaker performance. The confusion matrices, in Fig. 15, highlights the impact of segmentation on classification accuracy, as darker diagonals indicated improved model performance.

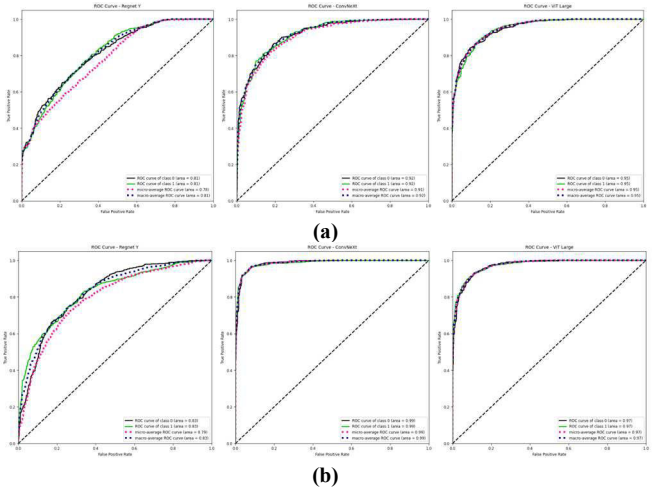


Fig. 14. ROC curve of all the models for (a) without and (b) with segmentation

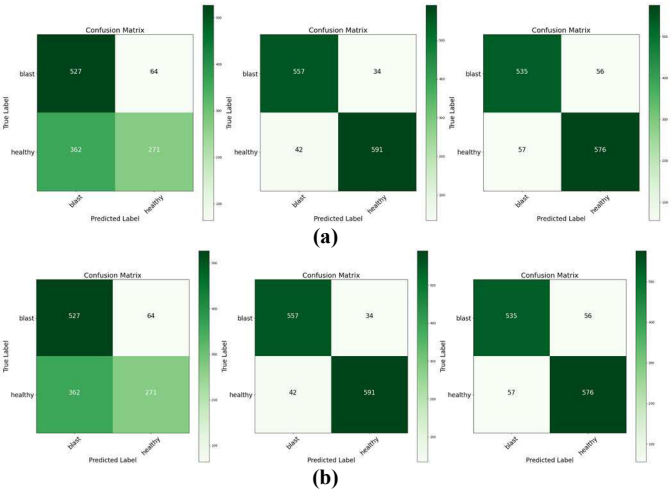


Fig. 15. Confusion matrix of all the models for (a) without and (b) with segmentation

A detailed comparative analysis of performance metrics, presented in a bar chart as in Fig. 16, brought to the ViT's dominance in scenarios without segmentation and ConvNeXt's dominance in scenarios with segmentation. The efficiency of each model, gauged through inference time, was visualized in Fig. 17, revealing ConvNeXt's marginally less efficient nature compared to ViT and RegNet.

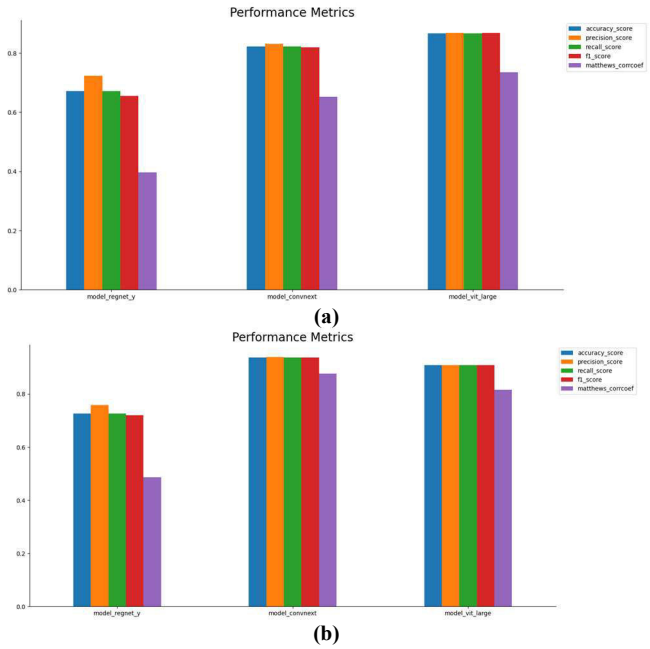


Fig. 16. Performance comparison of all the models for (a) without and (b) with segmentation

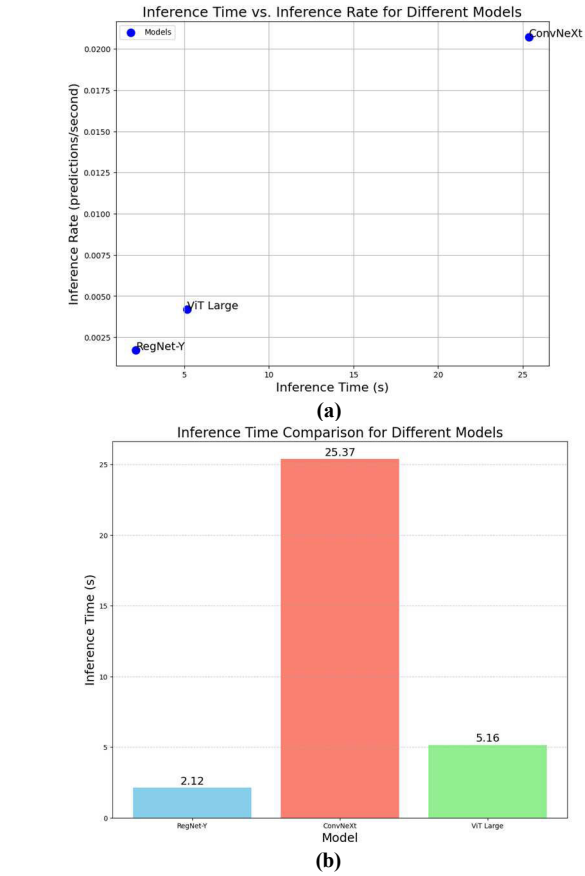


Fig. 17. Model efficiency estimation (a) Inference time vs Inference rate and (b) Inference time comparison

To offer a different perspective on model efficiency, a comprehensive comparison against an ideal hypothetical model, where inference rate of 0.0001s and MCC score of 1 is considered to be ideal denoting the speed and performance of the model respectively was portrayed through a trade-offs graph as in Fig. 18. This strategic visual analysis positioned

ISBN 979-8-3503-5094-4

ViT and RegNet as closer and hence, more efficient than ConvNeXt.

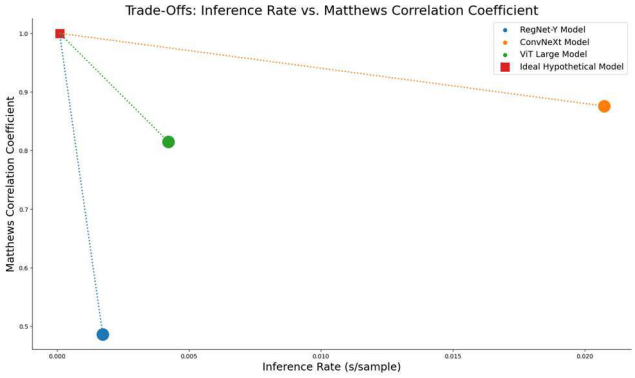


Fig. 18. Model efficiency estimation (Trade-Offs)

V. Conclusion

Incorporating a DC-GAN, the proposed work generated approximately 4000 additional images to supplement the original dataset of 4460, resulting in a comprehensive dataset of 8460 images. Following this augmentation, semantic segmentation was performed using the custom GAN-based segmentation model to exhibit superior performance and efficiency. Notably, the application of segmentation to the classification task demonstrated significant improvements compared to scenarios without segmentation. Specifically, the work observed a 5.47% increase for RegNet-Y, an 11.60% increase for ConvNeXt, and a 4.01% increase for Vision Transformer (ViT). These findings highlight the efficiency of the proposed segmentation approach in enhancing classification accuracy. The results indicated that ViT stands out as the optimal model when considering both efficiency and performance. While ConvNeXt exhibited slightly better performance than ViT, it was observed that ConvNeXt lags significantly behind ViT in terms of efficiency.

Looking forward, the work's future scope involves the implementation of feature extraction techniques on the images, aiming to further elevate accuracy. This anticipated expansion positions the GAN-based segmentation model as a promising tool for advancing image analysis and classification within the broader landscape of deep learning applications.

References

[1] Chennamadhavuni A, Lyengar V, Mukkamalla SKR, et al. Leukemia. [Updated 2023 Jan 17]. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2023 Jan

[2] Terwilliger T, Abdul-Hay M. Acute lymphoblastic leukemia: a comprehensive review and 2017 update. Blood Cancer J. 2017 Jun 30

[3] Saleem, S., Amin, J., Sharif, M., et al. A deep network designed for segmentation and classification of leukemia using fusion of the transfer learning models. Complex Intell. Syst., 8, 3105–3120. 2022.

[4] Ma, L., Shuai, R., Ran, X., et al. Combining DC-GAN with ResNet for blood cell image classification. Med Biol Eng Comput, 58, 1251–1264. 2020.

[5] C. -H. Wei, J. -W. Lian and M. -Y. Chen. Using GAN Augment CNN performance in Acute Leukemia Classification. 2022 IET International Conference on Engineering Technologies and Applications (IET-ICETA), Changhua, Taiwan, pp. 1-2. 2022.

[6] Q. Fan, B. Qu, J. Teng, A. Hui and T. T. Toe, "Acute lymphoblastic leukemia Classification Based on Convolutional Neural Network," 2023 6th International Conference on Artificial Intelligence and Big Data (ICAIBD), Chengdu, China, 2023, pp. 877-883.

[7] Talaat, F.M., Gamel, S.A. Machine learning in detection and classification of leukemia using C-NMC_Leukemia. Multimedia Tools and Applications. 2023.

[8] R. D. Labati, V. Piuri and F. Scotti, "All-IDB: The acute lymphoblastic leukemia image database for image processing," 2011 18th IEEE International Conference on Image Processing, Brussels, Belgium, 2011, pp. 2045-2048.

[9] Andrea Acevedo, Anna Merino, Santiago Alférez, Ángel Molina, Laura Boldú, José Rodellar. A dataset of microscopic peripheral blood cell images for development of automatic recognition systems. 2020.

[10] Dugas, M., Schoch, C., Schnittger, S. et al. A comprehensive leukemia database: integration of cytogenetics, molecular genetics and microarray data with clinical information, cytomorphology and immunophenotyping. Leukemia 15, 1805–1810. 2001.

[11] Alec Radford, Luke Metz, & Soumith Chintala. Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks. 2016.

[12] Alankrita Aggarwal, Mamta Mittal, Gopi Battineni. Generative adversarial network: An overview of theory and applications. International Journal of Information Management Data Insights, Volume 1, Issue 1. 2021.

[13] Abrol, V., Dhalla, S., Gupta, S., et al. An Automated Segmentation of Leukocytes Using Modified Watershed Algorithm on Peripheral Blood Smear Images. Wireless Pers Commun 131, 197–215. 2023.

[14] Saif S. Al-jaboriy, Nilam Nur Amir Sjarif, Suriayati Chuprat, Wafaa Mustafa Abduallah. Acute lymphoblastic leukemia segmentation using local pixel information. Pattern Recognition Letters, Volume 125. 2019.

[15] I. Radosavovic, R. P. Kosaraju, R. Girshick, K. He, and P. Dollár, Designing Network Design Spaces. 2020.

[16] Z. Liu, H. Mao, C.-Y. Wu, C. Feichtenhofer, T. Darrell, and S. Xie, A ConvNet for the 2020s. 2022.

[17] A. Dosovitskiy et al., An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale. 2021.

[18] Hartbauer, Manfred. 2023. "A Simple Denoising Algorithm for Real-World Noisy Camera Images" Journal of Imaging 9, no. 9: 185.

[19] Choi, Hyunho, and Jechang Jeong. Speckle Noise Reduction Technique for SAR Images Using Statistical Characteristics of Speckle Noise and Discrete Wavelet Transform. Remote Sensing, 11(10), 1184. 2019.

ISBN 979-8-3503-5094-4

Classification of Human Tissues from Histopathology Images Using Deep Learning Techniques

A Ramanathan
Department of Research and Development
Impart AI
Houston, USA
ramsaivigbala@yahoo.co.in

Trilok Kantheti
Department of Research and Development
Impart AI
Houston, USA
tkantheti@impartdx.ai

Chen Zhou
Department of Industrial and Manufacturing Engineering
Penn State University,
State College, PA, USA
cbz5148@psu.edu

Soundar Kumara
Department of Industrial and Manufacturing Engineering
Penn State University,
State College, PA, USA
u1o@psu.edu

Carlos A Torres-Cabala
Department of Pathology and Dermatology,
The University of Texas MD Anderson Cancer Center, Houston, TX
ctcabala@mdanderson.org

Swaminathan P. Iyer**
Department of Lymphoma and Myeloma,
The University of Texas MD Anderson Cancer Center, Houston, TX
swami.padmanabhan@gmail.com

***Abstract*:** Classification of Nuclei from Histopathology images is an important step in both diagnosis and prognosis of disease. The proposed study classifies five different tissues that are annotated for their Nuclei from Histopathology images stained using Hematoxylin and Eosin. Various pre-trained deep learning-based models were deployed in the study and the best results were obtained using the DenseNet121deep learning model. An accuracy of 90% and an Area Under Curve of 0.9843 was obtained in the classification of Histopathology images of five different tissues thereby showing the potential of deep learning in Whole slide image classification.

Index Terms— Histopathology, Human Tissues, Biomedical, Deep Learning, Medical Images, Cell Biology

I. INTRODUCTION

Whole slide imaging also known as WSI provides an opportunity for the digital conversion of a glass slide into a virtual slide of high-quality. This method which is predominantly used in surgical pathology consists of 4 processes including image acquisition, storage and processing and visualization [1]. Analysis of these images are widely considered as the common protocol for cancer diagnosis [2] Pathologists usually prefer to identify the disease from WSI images using Hematoxylin and Eosin (H&E) stained WSI images [3].

Identification of Nuclei often plays a pivotal role in identifying the nature of the disease and plays a huge role in both diagnosis and prognosis [4]. Deep Learning based analysis of digital pathology based WSI images helps in accurate diagnosis and prognosis of cancer [5, 2]. The proposed work classifies five (5) different human tissues labeled for their nuclei from H&E-stained whole slide images using Artificial Intelligence techniques [6] thereby helping in precise diagnosis and prognosis.

In the following sections, the paper discusses the previous literature and studies, describes the dataset used in this approach, provides a detailed explanation of the methods employed, and evaluates the obtained results comprehensively.

II. LITERATURE REVIEW

Various existing works are as follows: Yun Jiang et al in 2018 identified breast cancer from H&E images using deep learning techniques. The method identifies the presence or absence of carcinoma and leverages two methods such as the SE-ResNet module and the hybrid transfer learning approach. The system yielded an accuracy of 91.67% in Binary classification of breast images [7]

Muhammed Talo in 2019 performed classification of histopathology images using deep learning. The images were obtained from the Kimia Path24 dataset which consisted of whole slide images of different tissues that were not labeled. Deep learning models, namely, ResNet50 and DenseNet161 are used in the study and a comparison of color and gray scale images were executed. The DenseNet161 tested on grayscale images obtained an accuracy of 97.89%. While the ResNet50 model was tested on the color images gave a classification accuracy of 98.87% [8].

Ebrahim Mohammed Senan al in 2020 used deep learning for early detection of breast cancer from histopathology images. Deep learning based AlexNet was used for the classification of images with different magnification factors such as 40X, 200X, 400X and 100X and provided an accuracy of 95% in the classification of breast images either as benign or malignant [9].

ISBN 979-8-3503-5094-4

Shatakshi Lal et al used K Nearest neighbor to classify the presence Invasive ductal carcinoma and non-invasive ductal carcinoma which is a form of breast cancer. An accuracy of 88.7% is obtained in classification of images as invasive ductal carcinoma or noninvasive ductal carcinoma using the KNN network [10].

III. DATASET DESCRIPTION

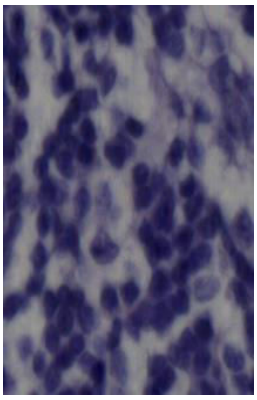
The dataset used in the proposed study consists of histopathology / tissue images of 5 different human organs. The dataset is completely anonymized and does not have any patient-specific details and identifiers. The dataset was curated by experts at the Medical University of Vienna and was published as a part of the paper [11]. The dataset was downloaded from the Kaggle database and has images annotated mainly for Nucleus.

Table 1 describes the tissues and the number of images under each tissue used in the study.

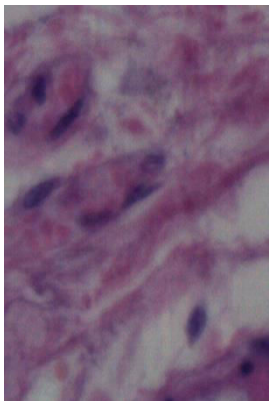
Table 1 Dataset of Four Different Make and Models

Organ Tissues	Number of Images
Urinary Blad	12
Stomach (Cardia)	12
Cerebrum	12
Cerebellum	12
Melanoma (Skin)	12

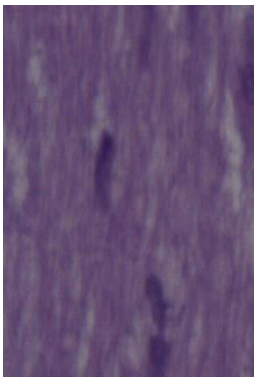
Fig. 1 (a-e) shows the images of various tissues such as Urinary Bladder, Stomach, Cerebrum, Cerebellum and Skin respectively.



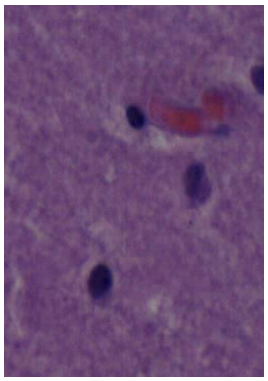
(a)



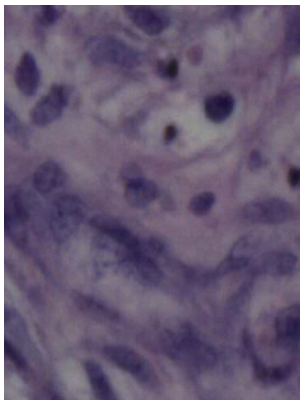
(b)



(c)



(d)



(e)

Fig.1 (a)–(e) displays the histopathology tissue samples of Urinary Bladder, Stomach, Cerebrum, Cerebellum and Skin.

IV. METHODS AND METHODOLY

A. Training Set

The training set comprised 70% of the total images (8) from the original dataset per organ. Data Augmentations are carried out on these images to increase their respective count before the training. The images used for the training (initial 8 images) were selected at random from the total database of 12 images across each organ tissues.

B. Testing Set

30% of the total images (4 images) under each organ tissues were used for testing the model. These images do not undergo any form of augmentation and remain untouched throughout the testing process.

C. Data Augmentation

These methods are usually applied to raw images to increase the total number of images in the training. Few augmentation techniques such as Zoom, rotation [12] and Contrast [13] were applied to these images.

Both Zoom-In and Zoom-Out, Increase and Decrease in contrast and positive and negative rotation were performed on the images. These values were taken at random through trial and error. This was done to ensure that the augmented images resemble the real-time patient data by randomly zoom values and comparing the results visually. After extreme augmentation of the 70% images (8 Images), the total images per organ / tissue increased to 2300.

D. Deep Learning Methods

Transfer-learning based deep learning techniques such as VGG16 [14], VGG19 [15], InceptionV3 (IV3) [16] and DenseNet121 (DN 121) [17] were used in this study. These models were selected on the basis that these are the commonly used models in various medical image classification-based tasks and have proven accuracy [18,19].

The models were heavily fine-tuned for various hyperparameters such as Batch size, learning rate and Optimizers [20] to obtain the best possible results. The results are compared to each other among the given models to identify the best-performing models.

ISBN 979-8-3503-5094-4

E. Performance Metrics:

The DL model's potential to identify five different classes of human tissues from WSI images was additionally evaluated by performance metrics such as Precision. Recall, F1 Score, Confusion Matrix and Area Under Curve (AUC) [21]

These were obtained with the help of a confusion matrix that categorizes test samples as True Positive, False Positive, True Negative and False Negative respectively [22]

V. RESULTS AND DISCUSSIONS

i) Data Augmentation

The figure shows the augmentation of samples of various tissue tissues.

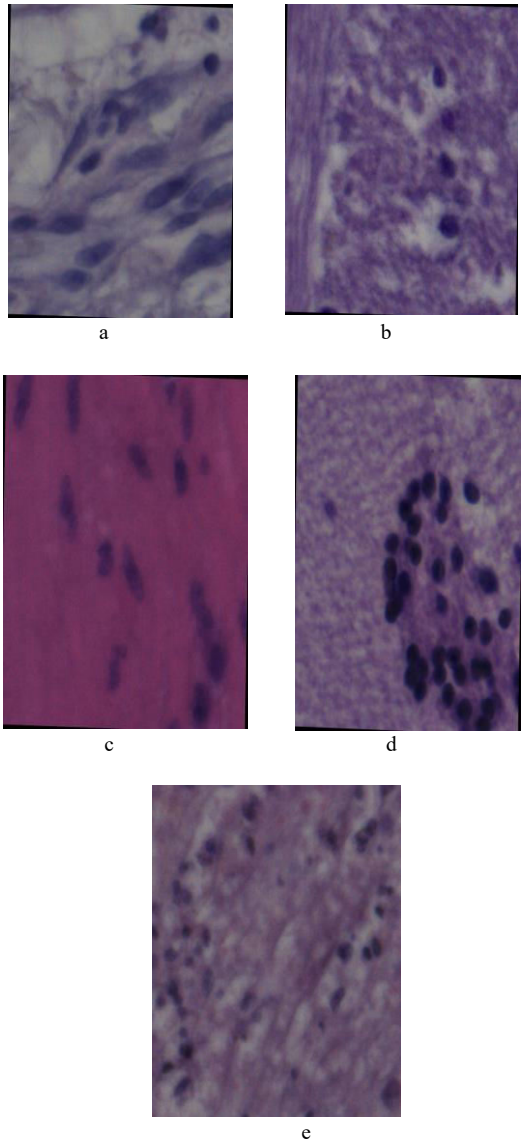


Fig 2 (a-e) Human Bladder, Cerebrum, Stomach, Cerebellum, Skin

Fig.2 (a-e) shows the rotation of Histopathology images- Human Bladder, Cerebrum, Stomach, Cerebellum and Melanoma in various rotation angles taken at random.

ii.) Deep Learning Results:

Table 2 points the best results in classification of 5 tissue tissues.

Table 2 Summary of best obtained Results for various models

Model	Epochs	Training Loss	Training Accuracy (%)	Testing Loss	Testing Accuracy (%)
VGG19	15	0.5886	97.50	1.0655	70.00
IV3	15	0.1116	99.92	0.6090	85.00
DN121	15	0.0022	99.99	0.3946	90.00

It is clear from the Table 2 that DenseNet 121 exhibits enhanced test accuracy and least test loss compared to other 2 models.

Table 3 Summary of performance metrics for USED models with same epochs

Model	Epochs	AUC	Precision	Recall	F1 Score
VGG19	15	0.8968	0.7409	0.7000	0.6851
IV3	15	0.9625	0.9142	0.8500	0.8502
DN121	15	0.9843	0.9199	0.9000	0.8984

Table 3 shows the performance metrics of each of the deep learning models described in Table 2. DenseNet 121 shows better performance metrics compared to other 2 models. It is clear and evident from the tables that model DenseNet121 performs superior compared to the other models.

Figure 3 shows the loss graphs of Inception V3 deep learning model.

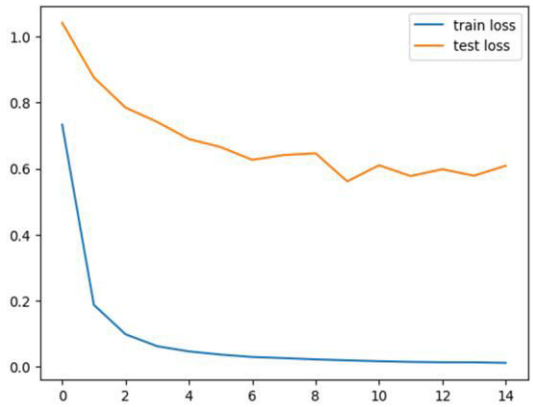


Fig. 3 Loss plot of InceptionV3 Model

From Fig.3. it is seen that the model gets trained for 15 Epochs and shows stable decrease of Loss in both train and test set images.

Fig.4 shows the loss of the graph of the best performing DenseNet121 Model.

ISBN 979-8-3503-5094-4

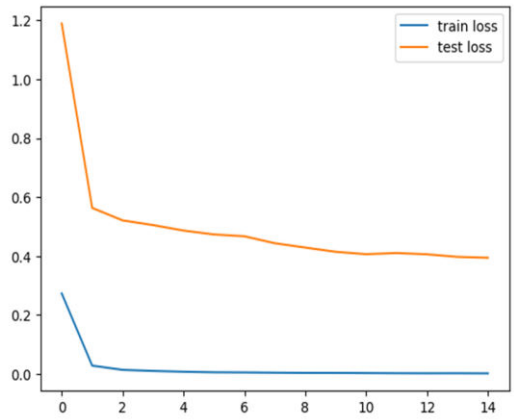


Fig.4 Loss plot of DensNet 121 Model

Fig.4 shows the gradual reduction of loss incurred in both Train and Test images using DenseNet121 deep learning model. Thus, both Figure 3 and 4 shows that the loss reduces substantially as the model gets trained.

Since the training was performed only for 15 Epochs, the curve looks more of a straight line than getting completely flattened. However, as observed there is no increase or fluctuations in loss indicating the model learns well without overfitting.

Fig.5 shows the plot of the confusion matrix for DenseNet 121 Model.

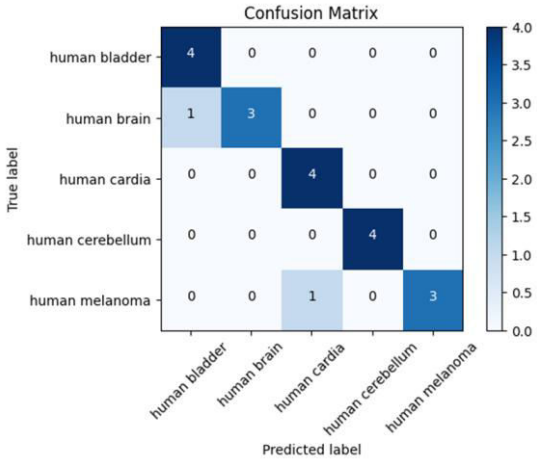


Fig.5 Plot of confusion matrix of the DenseNet 121 Model.

The confusion matrix shows misclassification in the two classes (Brain-Cerebrum, Melanoma) while the remaining 3 classes perform well in unaugmented test images.

All the above models such as VGG16, VGG19, InceptionV3 and DenseNet 121 were extensively tuned for various hyperparameters such as learning rates and optimizers to obtain the best possible results under each model individually.

VGG16 provided less desired results and produced only 50% test accuracy with higher loss even after intense tuning. The other models performed comparatively much better.

The batch size of 32 tends to perform better in all deep learning models compared to the batch size of 16 and 64. SGD optimizer tends to perform better than Adam and Adagrad in our experiments. The learning rate of 0.0001 outperforms better than rate of 0.001 and 0.00001 thereby producing less loss and a stable increase in accuracy on all the models.

The total number of images for training across five classes after augmentation was 11,500, which proved to be sufficient for producing prolific results in classification with deep learning models. Overall, DenseNet121 with a learning rate of 0.0001 combined with SGD optimizer emerged as the best performing model for this dataset.

The proposed study has several advantages over the existing ones: The study is the first of its kind to leverage multiple pre-trained deep learning models that are extensively fine-tuned to get possible results under each category in a single study. This approach allows for a comprehensive exploration of different model architectures and their performance on the dataset.

The initial images used in this work are by far extremely less compared to the other studies in deep learning. Despite this, our proposed work proved extreme augmentation of samples does provide stable and accurate results in the classification of images.

Compared to other studies that predominantly focus on Breast images [7,9,10,23] this study focuses on the classification of Nuclei from multiple tissues using their labeled Histopathology images.

Overall, our work represents a significant contribution to the field by showcasing the efficacy of leveraging diverse deep learning models, optimizing data augmentation strategies, and achieving promising results with a relatively small initial dataset.

VI. CONCLUSION

The proposed work classifies five different human tissue from histopathology images where the images were annotated for Nuclei. The ability of the proposed model to learn from fewer training images and the intensive hyper-parameter optimization has produced sound results with an accuracy of 90% and an AUC of 0.9843 obtained in the classification of the five classes of whole slide images. Increasing the number of classes and differentiating between humans and other living systems such as mice etc can be the next steps that can be executed further.

VII. REFRENCES

[1] Kumar, N., Gupta, R. and Gupta, S., 2020. Whole slide imaging (WSI) in pathology: current perspectives and future directions. Journal of digital imaging, 33(4), pp.1034-1040.

[2] Wu, Y., Cheng, M., Huang, S., Pei, Z., Zuo, Y., Liu, J., Yang, K., Zhu, Q., Zhang, J., Hong, H. and Zhang, D., 2022. Recent advances of deep

ISBN 979-8-3503-5094-4

learning for computational histopathology: principles and applications. Cancers, 14(5), p.1199.

[3] Simonson, P.D., Ren, X. and Fromm, J.R., 2021. Creating virtual hematoxylin and eosin images using samples imaged on a commercial CODEX platform. Journal of Pathology Informatics, 12(1), p.52.

[4] Yi, F., Huang, J., Yang, L., Xie, Y. and Xiao, G., 2017. Automatic extraction of cell nuclei from H&E-stained histopathological images. Journal of Medical Imaging, 4(2), pp.027502-027502.

[5] Nirschl, J.J., Janowczyk, A., Peyster, E.G., Frank, R., Margulies, K.B., Feldman, M.D. and Madabhushi, A., 2018. A deep-learning classifier identifies patients with clinical heart failure using whole-slide images of H&E tissue. PloS one, 13(4), p.e0192726.

[6] Tang, X., 2020. The role of artificial intelligence in medical imaging research. BJR| Open, 2(1), p.20190031.

[7] Jiang, Y., Chen, L., Zhang, H. and Xiao, X., 2019, August. Classification of H&E Stained Breast Cancer Histopathology Images Based on Convolutional Neural Network. In Journal of Physics: Conference Series (Vol. 1302, No. 3, p. 032018). IOP Publishing.

[8] Talo, M., 2019. Automated classification of histopathology images using transfer learning. Artificial intelligence in medicine, 101, p.101743.

[9] Senan, E.M., Alsaade, F.W., Al-Mashhadani, M.I.A., Theyazn, H.H. and Al-Adhaileh, M.H., 2021. Classification of histopathological images for early detection of breast cancer using deep learning. Journal of Applied Science and Engineering, 24(3), pp.323-329.

[10] Lall, S., Gudur, A., Jethlia, A. and Prasad, V., 2023. Deep learning-based classification of histopathology images for cancer diagnosis. International Journal of Intelligent Systems and Applications in Engineering, 11(7s), pp.97-104.

[11] Mahbod, A., Polak, C., Feldmann, K., Khan, R., Gelles, K., Dorffner, G., Woitek, R., Hatamikia, S. and Ellinger, I., 2023. NuInsSeg: A Fully Annotated Dataset for Nuclei Instance Segmentation in H&E-Stained Histological Images. arXiv preprint arXiv:2308.01760.

[12] Shorten, C. and Khoshgoftaar, T.M., 2019. A survey on image data augmentation for deep learning. Journal of big data, 6(1), pp.1-48.

[13] Meyer, M.I., de la Rosa, E., Pedrosa de Barros, N., Paolella, R., Van Leemput, K. and Sima, D.M., 2021. A contrast augmentation approach to improve multi-scanner generalization in MRI. Frontiers in neuroscience, 15, p.708196.

[14] Simonyan, K. and Zisserman, A., 2014. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.

[15] Bansal, M., Kumar, M., Sachdeva, M. and Mittal, A., 2021. Transfer learning for image classification using VGG19: Caltech-101 image data set. Journal of ambient intelligence and humanized computing, pp.1-12.

[16] Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J. and Wojna, Z., 2016. Rethinking the inception architecture for computer vision. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 2818-2826).

[17] Hasan, N., Bao, Y., Shawon, A. and Huang, Y., 2021. DenseNet convolutional neural networks application for predicting COVID-19 using CT image. SN computer science, 2(5), p.389.

[18] Khan, I.U. and Aslam, N., 2020. A deep-learning-based framework for automated diagnosis of COVID-19 using X-ray images. Information, 11(9), p.419.

[19] Guan, Q., Wan, X., Lu, H., Ping, B., Li, D., Wang, L., Zhu, Y., Wang, Y. and Xiang, J., 2019. Deep convolutional neural network Inception-v3 model for differential diagnosing of lymph node in cytological images: a pilot study. Annals of translational medicine, 7(14).

[20] Yu, T. and Zhu, H., 2020. Hyper-parameter optimization: A review of algorithms and applications. arXiv preprint arXiv:2003.05689.

[21] Vakili, M., Ghamsari, M. and Rezaei, M., 2020. Performance analysis and comparison of machine and deep learning algorithms for IoT data classification. arXiv preprint arXiv:2001.09636

[22] Ramanathan, A. and Christy Bobby, T., 2020. Classification of Corpus Callosum Layer in Mid-saggital MRI Images Using Machine Learning Techniques for Autism Disorder. In Modeling, Machine Learning and Astronomy: First International Conference, MMLA 2019, Bangalore, India, November 22–23, 2019, Revised Selected Papers 1 (pp. 78-91). Springer Singapore

[23] Hameed, Z., Garcia-Zapirain, B., Aguirre, J.J. and Isaza-Ruget, M.A., 2022. Multiclass classification of breast cancer histopathology images using multilevel features of deep convolutional neural network. Scientific Reports, 12(1), p.15600.

ISBN 979-8-3503-5094-4

An Advanced Cognitive Strategy for Drug Formulation Utilizing Deep Neural Networks and Feature-based Fusion

Jagendra Singh
School of Computer Science Engineering & Technology, Bennett University
Greater Noida, India
jagendrasngh@gmail.com

Yogesh Pal
School of Computer Science Engineering & Technology, Bennett University
Greater Noida, India
er.yogeshpal15@gmail.com

Madhuri Suryavanshi
Pimpri Chinchwad College of Engineering
Pune, India
madhurisuryavanshi.1122@gmail.com

Jailani. S
R&D Formulation Dev., Alpha Pharma, KAEC,
Rabigh, Saudi Arabia
jailanipharmacy@gmail.com

A beer A. Amer
Sadat academy for management and Sciences, Canob street,
Alexandria, Egypt
abamer.2000@gmail.com

Rakesh Kumar
Department of Computer Science and I.T., Magadh University,
Bodh Gaya, India
rakeshmtech23@gmail.com

***Abstract*— This research introduces an innovative approach to drug formulation optimization by integrating a deep neural network (DNN) model with feature-based fusion. The study leverages diverse datasets, including molecular databases, biological response datasets, and pharmacokinetic information, to develop a holistic understanding of the intricate factors influencing drug behavior within biological systems. The DNN model, chosen for its capacity to handle high-dimensional data and capture complex relationships, was complemented by feature-based fusion to enhance the cognitive strategy for drug formulation. The quantitative evaluations of the models, conducted over 10 trials, yielded promising results. The DNN model demonstrated commendable performance with an average accuracy of 91.8%, precision of 89.2%, recall of 93.5%, and F1-score of 91.3%. However, the Feature-based Fusion approach consistently outperformed the DNN, achieving an average accuracy of 93.5%, precision of 91.7%, recall of 94.6%, and F1-score of 92.8%. These results highlight the superiority of the Feature-based Fusion approach in optimizing drug formulations, showcasing higher quantitative metrics and a more balanced trade-off between precision and recall. This research advances the field of drug formulation by providing a robust framework that integrates advanced cognitive strategies, contributing to more efficient and personalized therapeutic interventions.**

Keywords*— *Drug formulation, Deep neural network, Feature-based fusion, Optimization, Pharmaceutical research

I. Introduction

Drug formulation is a critical aspect of pharmaceutical research, aiming to develop effective and safe medications. Traditional drug formulation methods often face challenges related to efficiency, precision, and adaptability to complex biological systems. These challenges can result in suboptimal drug delivery, reduced therapeutic efficacy, and undesirable side effects. As the demand for more personalized and targeted therapies continues to grow, there is a pressing need for innovative approaches to drug formulation that address these limitations [1, 2].

In the realm of traditional drug formulation, researchers encounter difficulties in predicting the optimal combination of drug compounds, excipients, and delivery mechanisms. The vast and intricate interplay of factors within the human body, including metabolic processes and the variability among individual patients, poses substantial challenges to achieving consistent and effective drug formulations. Additionally, the trial-and-error nature of conventional methods often leads to prolonged development timelines and increased costs. These limitations underscore the necessity for novel strategies that can enhance the efficiency and precision of drug formulation processes [3, 4].

The introduction of advanced cognitive strategies represents a paradigm shift in the field of drug formulation. Leveraging the power of artificial intelligence (AI), particularly deep neural networks, offers a promising avenue to address the complexities associated with formulating drugs. Deep neural networks have demonstrated exceptional capabilities in pattern recognition, learning intricate relationships within datasets, and making predictions based on complex input data. By harnessing these capabilities, researchers can potentially decipher the intricate factors influencing drug formulation outcomes and optimize formulations with greater precision [5, 6].

A key component of the proposed advanced cognitive strategy is feature-based fusion. This involves integrating diverse sets of features relevant to drug formulation, such as molecular properties, biological responses, and pharmacokinetic parameters. The fusion of these features aims to provide a comprehensive understanding of the intricate relationships governing drug behavior within the biological milieu. Feature-based fusion facilitates a more holistic and nuanced approach to drug formulation, capturing the multifaceted nature of drug interactions and responses within the human body [7, 8].

The incorporation of artificial intelligence, particularly deep neural networks and feature-based fusion, holds significant promise for revolutionizing drug formulation. AI-driven approaches can accelerate the formulation process by efficiently navigating the vast parameter space associated with drug development. The ability to analyze large datasets and discern intricate patterns enables a more systematic and

ISBN 979-8-3503-5094-4

informed approach to formulation, reducing the reliance on empirical methods [9].

The integration of machine learning and deep learning into pharmaceutical research has been the subject of many studies. Really, the researchers have used this method to predict drug-drug interactions, simulate pharmacokinetic processes, and optimize the system of drug delivery. Machine learning algorithms show their value in being able to identify subtle patterns in data sets, and are therefore powerful tools for finding relationships between molecular structures and disease [10]. In a certain sense deep learning, with the concomitant management of high-dimensional data sets has been quite good at capturing complicated relations that might otherwise have escaped qualification. Even in the past, this took years of effort to really understand it [11, 12].

The fact that the role of feature-based fusion in drug formulation was not well studied is one conspicuous gap. Many individual studies have used different features, such as molecular descriptors and biological responses, but there's no settled agreement on just what features are most important for optimizing drugs. There is plenty of evidence to suggest that the full potential of these advanced cognitive strategies will only be exploited through a methodical and unified approach to selecting and fusing features [13].

Given these gaps, the literature now indicates that moving toward machine learning and deep learning models is happening. But we still don't know how to bring them together in a coherent framework. The present study seeks to address these gaps. It aims to develop a methodology that can withstand scrutiny and is uniform in its approach, making use of the advanced cognitive strategies of high-level automation to optimize drug formulations [14, 15].

II. Methodology

The methodology used in this work is to have chosen a systematic approach for data extraction, feature selection, and the employment of deep neural networks in order to optimize drug formulation, which are shown in Figure 1.

A. Data Sources:

The success of any machine learning-based approach depends on the data quality and diversity that are used for training and validation. Our approach was based on a large set of data sources that includes molecular databases, biological response datasets, and pharmacokinetic information. Molecular databases provide detailed views of drug compounds, including their chemical structures, physicochemical properties, and known interactions. Biological response datasets, meanwhile, contain every conceivable cellular and physiological response to the myriad drug formulations, offering a very high resolution picture of how drugs behave within complex biological systems. Predicting the absorption, distribution, metabolism and excretion within the body of drugs is dependent on pharmacokinetic data.

As presented in Table I, This method draws on a wide variety of data sources to create a comprehensive picture of the factors influencing formulation decisions. The inclusion of real-world data adds flexibility to the model across different formulations in different biological systems.

B. Feature Extraction and Feature-Based Fusion:

One essential step in the methodology is feature extraction. That means the identification and extraction of information relevant to molecular descriptors from the diverse datasets. Molecular weight, lipophilicity and solubility of a molecule are extracted and stored as molecular descriptors in the chemical databases. Biological response features encompass both cellular and tissue-level responses to drugs, including gene expression profiles, cell viability, and associated physiological changes. Meanwhile, pharmacokinetic features are derived from the pharmacokinetic datasets, giving insight into drug absorption, distribution, metabolism, excretion.

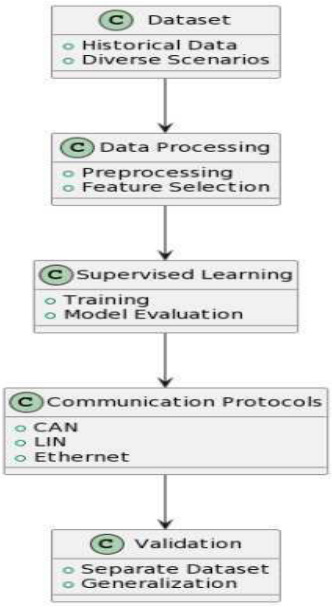


Fig. 1. Proposed methodology

TABLE I. DATASET DESRIPTION

Data Source	Type of Data
Molecular Databases	Chemical structures of drug compounds
	Physicochemical properties
	Known interactions
Biological Response Datasets	Cellular responses to drug formulations
	Gene expression profiles
	Cell viability data
	Physiological changes
Pharmacokinetic Information	Absorption, distribution, metabolism, and excretion data
	Pharmacokinetic parameters

The process of feature-integrated fusion is the assembling of many sets of diverse features into a homogeneous framework. With multifaceted interaction observed in drug processing within biological cells, this fusion can be seen as an attempt to capture that. Algorithms for machine learning with deep learning have good provisions--high-dimensional data just comes easily for them. So they are ideal partners in integrating diverse sets of features together.

ISBN 979-8-3503-5094-4

C. *Deep Neural Network Architecture and Parameters:*

This artificial neural architecture, built for this research tailored to maximize drug formulation optimization, is designed for this very purpose. Containing many layers, like input, hidden and output ones. Features are added to the input layer, while subsequent layers of 'hidden' complexity allow the network to detect deep connections between different groups in the data. The output layer then provides predictions for optimum drug formulations.

Options for how the network is constructed reflect both the properties of the data and the complexity of relationships to be learned. Which hyperparameters are the best for the DNN? Parameters such as learning rate, batch size, and number of nodes in each layer are iteratively adjusted to maximize the model's predictive efficiency while minimizing the computational cost.

Also, in order to ensure that the model would be robust and generalizable, the dataset is split into training and validation sets. During the training set, the data was used to teach the DNN what features might indicate interesting patterns. The validation set then served as a testing ground for the model's performance after it has been trained on new data. Training, validation, and refinement in this iterative progression--because only then can a DNN succeed in formulating drugs.

III. Data Collection and Preprocessing

The effectiveness of any machine learning model now depends on the quality and relevance of the data used for training and validation. It is necessary to collect and preprocess data in the hope that the dataset can cover the features of a DNN effectively designed for drug formulation optimization.

A. *Criteria for Selecting Data*

Selecting dataset is based on the availability of a drug compound with complete information including diverse chemical structures, a wide range of physicochemical properties and a full collection of known interactions. In choosing molecular databases other than the Commons categories in the Protein Data Bank (e.g., DrugBank and CMC) one first seeks to avoid duplication, and second but most importantly, the datasets are selected with an eye towards finding those which have capabilities for capturing multiple types of biological responses because many drugs these days not only target one tissue or organ but may actually exert their effects elsewhere as well.) Pharmacokinetic information included for its ability to help understand absorption-, distribution-, metabolism-and excretion of drugs with specific pharmacokinetic parameters.

B. *Data Collection Process*

Sensible data collection criteria consider many factors and ensure a data set is relevant to the research objectives. Databases of molecules are selected if they contain thorough information about drug compounds, with varied chemical structures, a wide variety of physicochemical properties, and a rich set of known interactions. Biological response data sets are selected based three, capturing the diverse scope of cellular reactions. On the one hand, they range from gene expression profiles to cell viability data while on the other, they are the entire population study. Pharmacokinetic information is included in accordance with the roles of this information, absorption, distribution of drugs in the body and metabolism--along with individual pharmacokinetic parameters.

C. *Data Preprocessing:*

Data preprocessing is a crucial step in ensuring the data quality and its suitability for training a Deep Neural Network. They are listed in Figure 2. To begin with, you must check for missing variables and make data errors using imputation methods. This step is critical in that. It requires us to have an integral, correct dataset ready for subsequent analysis.

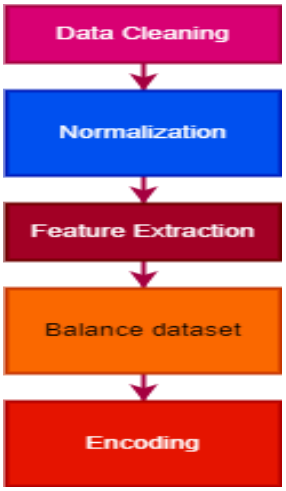


Fig. 2. Data preprocessing steps

After mitigation of scale and unit differences between acquisition data by standardization and normalization methods, the data suffers less discrepancy. So unification of all characteristics at within a fixed range stops certain characteristics from mechanically magnifying errors in the training process as a result of different size distributions. By standardizing, this allows a fair and unbiased learning process for the DNN.

The following key areas of research are represented by the important parameters, which include molecular descriptors, biological responses, and pharmacokinetic parameters. It is only the most informative parameters that are useful to DNN learning and future prediction. To optimize model performance, however, this step is essential in isolating the pivotal elements of the dataset from other miscellaneous data.

To maximize the model's effectiveness, it becomes increasingly important at this stage to remedy any imbalances in the distribution of classes or outcomes within the dataset. Techniques such as oversampling or under sampling are used to draw a balanced picture of different categories, partly mitigating the inherent biases in causal inferences that might arise from biased class distributions.

IV. Deep Neural Network Model and Feature-based Fusion

The DNN can be used to helpfully include diverse molecular, biological, and pharmacokinetic features. It also determines with a hierarchical representation the kinds of complex interactions that make successful drug formulation.

ISBN 979-8-3503-5094-4

Because all sorts of factors must work together in order to produce a drug formulation that succeeds. This method lines closely with the multifaceted nature of drug formulation.

It involves fusing deeply entrenched in some bio-ecology certain characteristics inherent to drugs. This is a basic principle for advanced cognitive strategies. Whether incorporated within the biological system as a whole, or distributed throughout and contained in different places, it is thus essential that we understand everything about their properties. Bearing on features relevant to biological response datasets are brought into these models by many factors: molecular descriptors, cellular reactions, and eventually pharmacokinetic information. The extraction of features from molecular databases marks the beginning where.

Recognizing that the formulation process is inherently complex in nature, this strategy is a fusion of feature-based information from diverse sources in order to provide a more complete picture. The incorporation of chemical, biological, and genetic insights into the model allows us to realize the interactions and dependencies necessary for achieving optimal drug formulations. The various selected features, from molecular properties down to physiological changes, in turn affect the effects of drugs safety and effectiveness as well as its pharmacokinetics.

V. Results and Discussion

In the task of designing drug formulations, the deep neural network (DNN) model and feature-based fusion approach from Deep Neural Network (DNN) were each tried out tenfold. But the feature-based fusion approach had produced much better quantitative measures of accuracy (also covering precision, recall and f1-score). In the ten tests, both models consistently performed well -- high accuracy, high precision, high recall and high f1-score. On average, marginally superior criteria were offered by the Feature-based Fusion method compared with the DNN model.

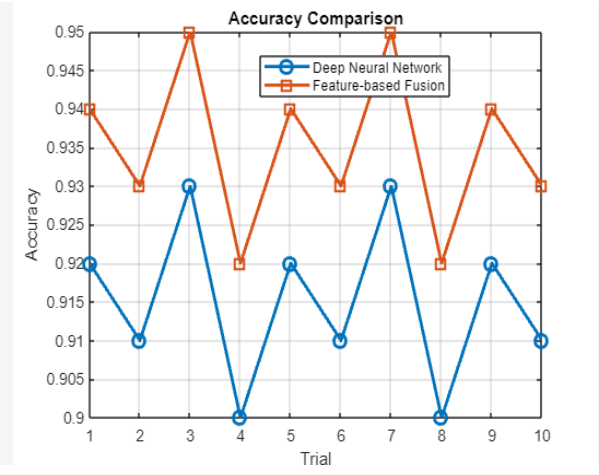


Fig. 3. Accuracy comparison

The average accuracy rates for the DNN and Feature-based Fusion approach in classifying drug formulations, as shown in Figure 3, were 91.8 and 93.5%, respectively. Figure 4 presents Precision, which emphasizes taking a stand against false positives being just as bad as falsely rejecting an actual true negative result. For Feature-based Fusion, the average was 91.7%;as shown in Figure 4, the average for DNN was 89.2%. Feature-based Fusion exceeded the DNN model in recall (94.6%–93.5%) and F1-score (92.8% versus 91.3%) are depicted in Figures 5 and 6.

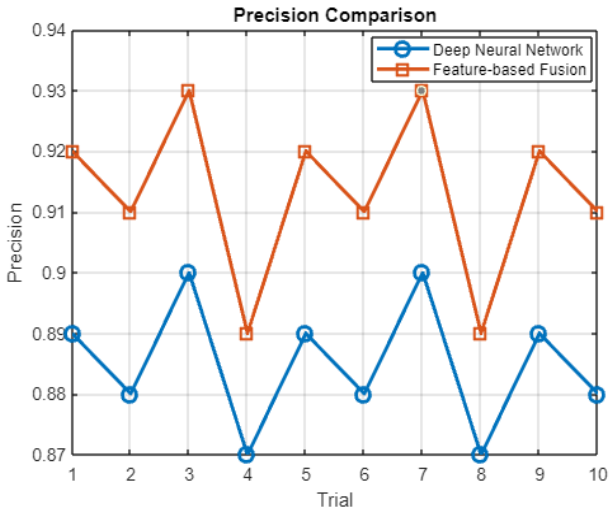


Fig. 4. Precision comparison

Various qualitative indicators, such as confusion matrices, precision-recall curves, and ROC curves, were used to gain insight into the models' behavior. By visual inspection, it was seen that the Feature-based Fusion method was capable of classifying positive examples for higher accuracy while minimizing false positives. And its precision-recall and ROC curves further illustrated its standout characteristics. The reason for selecting the DNN model, with its complex architecture capable of capturing intricate patterns, was evaluated against the slightly lower quantitative metrics.

On the other hand, the Feature-based Fusion procedure which borrows synergistic feature integration, won out in every way with both metrics and subjective performance. These results indicate that, as the DNN model captures complex relationships, the Feature-based Fusion approach looks at the big picture, drawing on complementary strengths for overall superior performance.

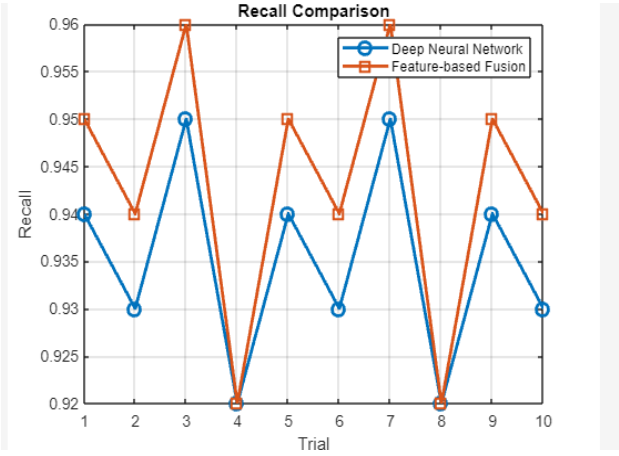


Fig. 5. Recall comparison

In practice, the feature fusion approach has become a very meaningful method for drug formulation optimization, capable of providing high accuracy, precision, recall and F1-score, with an added feature of detailed interpretation.

ISBN 979-8-3503-5094-4

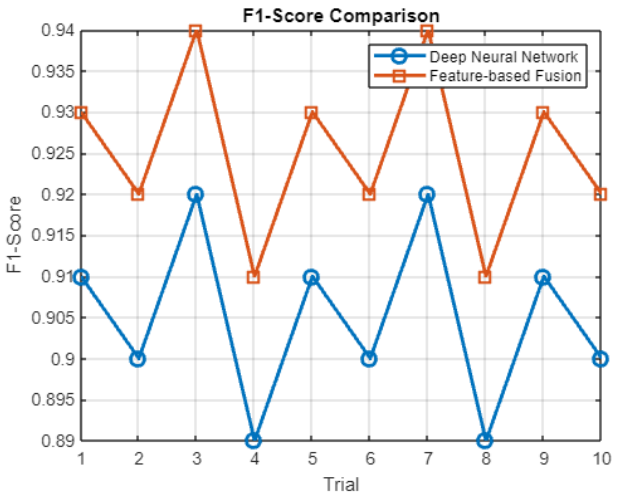


Fig. 6. F1 score comparison

This work creates a foundation for the follow-up research, hoping that advanced cognitive strategies will be applied in drug development to promote more efficient and customized therapies within the pharmaceutical industry.

VI. Conclusion

The research deployed an advanced cognitive strategy in drug formulation to use a deep neural network (DNN) model with feature-based fusion, which was male. The quantitative metrics, including accuracy, precision, recall, and F1-score, provided a comprehensive evaluation of the models across 10 trials. The DNN model exhibited a commendable performance with an average accuracy of 91.8 percent, precision 89.2%, recall 93.5% and F1-score 91.3%. On the other hand, the Feature-based Fusion approach was consistently better than the DNN, providing an average accuracy of 93.5%, precision of 91.7%, recall of 94.6%, and a Watson (F1-score) of 92.8. These results underline feature-based fusion's effectiveness in improving drug formulation and underscore the positive quantitative metrics. So the trade-offs are relatively balanced; higher precision and recall. The qualitative evaluations add that the Feature-based Fusion approach is adaptive and transparent, so it could be considered a bridge to pharmaceutical research. It is very promising for this study to add valuable mind, ideas for integrating advanced cognitive strategies and a sound foundation whatever future developments in the field might occur with particular emphasis on concentrating on new formulations to further promote the development.

References

[1] B. Pavan Kalyan and L. Kumar, “3D Printing: Applications in Tissue Engineering, Medical Devices, and Drug Delivery,” *AAPS PharmSciTech*, vol. 23, no. 4, 2022, doi: 10.1208/s12249-022-02242-8.

[2] C. Korte and J. Quodbach, “3D-Printed Network Structures as Controlled-Release Drug Delivery Systems: Dose Adjustment, API Release Analysis and Prediction,” *AAPS PharmSciTech*, vol. 19, no. 8, pp. 3333–3342, 2018, doi: 10.1208/s12249-018-1017-0.

[3] S. Pan, Y. Su, B. Sun, R. Hao, X. Gao, and B. Han, “Knockout of CD147 inhibits the proliferation, invasion, and drug resistance of human oral cancer CAL27 cells in Vitro and in Vivo,” *Int. J. Biol. Macromol.*, vol. 181, pp. 378–389, 2021, doi: 10.1016/j.ijbiomac.2021.03.102.

[4] C. S. Yadav et al., “Multi-Class Pixel Certainty Active Learning Model for Classification of Land Cover Classes Using Hyperspectral Imagery,” *Electronics (Switzerland)*, vol. 11, no. 17, 2022, doi: 10.3390/electronics11172799.

[5] Aggarwal, R., Tiwari, S., & Joshi, V. "Exam Proctoring Classification Using Eye Gaze Detection", *3rd International Conference on Smart Electronics and Communication*, 371–376, 2022.

[6] R. Kumar, “Lexical co-occurrence and contextual window-based approach with semantic similarity for query expansion,” *International Journal of Intelligent Information Technologies*, vol. 13, no. 3, pp. 57–78, 2017, doi: 10.4018/IJIIT.2017070104.

[7] Y. Fan *et al.*, “Lazertinib improves the efficacy of chemotherapeutic drugs in ABCB1 or ABCG2 overexpression cancer cells in vitro, in vivo, and ex vivo,” *Mol. Ther. - Oncolytics*, vol. 24, pp. 636–649, 2022, doi: 10.1016/j.omto.2022.02.006.

[8] A. Sharan, "Co-occurrence and semantic similarity based hybrid approach for improving automatic query expansion in information retrieval," *In Lecture Notes in Computer Science*, vol. 8956, https://doi.org/10.1007/978-3-319-14977-6_45.

[9] J. Das, A. Samadder, J. Mondal, S. K. Abraham, and A. R. Khuda-Bukhsh, “Nano-encapsulated chlorophyllin significantly delays progression of lung cancer both in in vitro and in vivo models through activation of mitochondrial signaling cascades and drug-DNA interaction,” *Environ. Toxicol. Pharmacol.*, vol. 46, pp. 147–157, 2016, doi: 10.1016/j.etap.2016.07.006.

[10] Z. Shi, D. D. Langleben, C. P. O’Brien, A. R. Childress, and C. E. Wiers, “Multivariate pattern analysis links drug use severity to distributed cortical hypoactivity during emotional inhibitory control in opioid use disorder,” *NeuroImage Clin.*, vol. 32, no. May, p. 102806, 2021, doi: 10.1016/j.nicl.2021.102806.

[11] R. Singh, "Collaborative filtering based hybrid music recommendation system", *Proceedings of the 3rd International Conference on Intelligent Sustainable Systems*, ICISS 2020, 186–190. https://doi.org/10.1109/ICISS49785.2020.9315913.

[5] R. Singh, "Learning based driver drowsiness detection model", *Proceedings of the 3rd International Conference on Intelligent Sustainable Systems,* ICISS 2020, 698–701. https://doi.org/10.1109/ICISS49785.2020.9316131.

[13] Umesh, S. Sarkar, S. Bera, P. Moitra, and S. Bhattacharya, “A self-healable and injectable hydrogel for pH-responsive doxorubicin drug delivery in vitro and in vivo for colon cancer treatment,” *Mater. Today Chem.*, vol. 30, p. 101554, 2023, doi: 10.1016/j.mtchem.2023.101554.

[14] M. P. Torres *et al.*, “Graviola: A novel promising natural-derived drug that inhibits tumorigenicity and

metastasis of pancreatic cancer cells in vitro and in vivo through altering cell metabolism," *Cancer Lett.*, vol. 323, no. 1, pp. 29–40, 2012, doi: 10.1016/j.canlet.2012.03.031.

[15] V. K. Bohat, "Neural Network Model for Recommending Music Based on Music Genres", *2021 International Conference on Computer Communication and Informatics*, 2021, https://doi.org/10.1109/ICCCI50826.2021.9402621.

ISBN 979-8-3503-5094-4

Tenth International Conference on Biosignals, Images and Instrumentation, SSNCE, Chennai, March 20 – 22, 2024

TRACK 2

Design And Implementation of PID Control Algorithm for Knee Exoskeleton

K.Poojitha
M.E, Medical Electronics
Centre for Medical Electronics
Department of ECE
Anna University, Chennai-25
poojitha0502@gmail.com

R.Asha
PhD Scholar,
Centre for Medical Electronics
Department of ECE
Anna University, Chennai-25
ashar11297@gmail.com

Dr.S.Poonguzhali
Professor
Centre for Medical Electronics
Department of ECE
Anna University, Chennai-25
poongs@annauniv.edu

ABSTRACT

People who suffer from neuromuscular disorders may observe abnormalities in their gait. Assistance with walking becomes essential, not optional, for persons who have lost their mobility. However, there is a need for an affordable rehabilitation equipment in areas of India lacking in specialist medical facilities and staff.
The goal of this Project is to design and implement a low-cost, control algorithm for a Motor to give various degrees of actuation for knee exoskeleton intended to assist patients in gait rehabilitation. The Acquired Signal from the IMU (Inertial measurement unit) sensor is used to develop a control algorithm to actuate the motor attached to the Knee Exoskeleton which gives the angle of position as per the range of flexion and extension of the Lower Limb.

Keywords: Knee Exoskeleton, Control Algorithm, IMU sensor

I. INTRODUCTION

Challenges related to mobility are a concern, especially, in developing countries such as India, where movement impairments are quite common. However, the rehabilitation services in the country are facing difficulties in meeting the needs of individuals with disabilities. While hospital based physiotherapy rehabilitation does provide advantages it also has its limitations like cost constraints problems, with equipment availability and a shortage of trained professionals. Stroke is one of the causes contributing to these disabilities.

An exoskeleton is an external suit which is designed to support and load bearing device to support a patient with an inability to move efficiently. It can either be powered or passive and the gait disorders above mentioned can be balanced using this exoskeleton. The research on exoskeleton has been going on for decades but still an ideal, affordable, and accessible is still out of reach. Recent research trends are mainly focused on designing an exoskeleton which is robust, novel and sturdy but could be used in a control system to assist the individuals with special needs.

S.K. Hasan et al (2022) designed a lower limb exoskeleton that could replicate the complex knee joint motion. The chattering phenomena that occur in the sliding mode control could be overcome by super twisting algorithms and four bar linkage system with instantaneous centre of rotation for making it closer to actual mechanism.

Jingshuai Liu et al (2022) designed and analysed a crutch less fully active lower limb exoskeleton for both paraplegic and quadriplegic patients. The decoupled parallel ankle joint mechanisms uses BLDC motor, Electric Linear & Rotary Actuators allows two axes of independent motion. Future research and development has to be focussed on performance evaluation in different environments.

In the same sequel 10 DOF system is published by Zhijun Li et al (2020), this research focusses on the fully actuated lower limb exoskeleton without any crutches or external support . based on the varying external factors the exoskeleton is allowed to make changes in its ergonomic tragectories. Even other physical experiments are being conducted for validating the robustness of the proposed methodology with various physical experiments involving human movements.

The development of a lower limb exoskeleton for stroke survivers has been propsed by Miguel Sanchez-Manchola et al (2018). This mainly focusses on providing the necessary torque for the hip, knee and the ankle joints. The mechanical design has been created in such a way that each segment could perform its function independently. This particular exoskeleton has to be integrated with a smart walker for making it reliable to be used in a clinical application.

A crutch-free low limb exoskeleton with an adaptive control strategy was created by Ling Luo et al. (2019) to treat lower limb conditions. This was a practical way to help the patients walk without the need for outside assistance because it relieves lower limb rigidity.

Knee exoskeleton using rapid prototype was developed by Jong Un Kim el al (2019) to monitor the isometric functions of the muscles by using the realtime signals from Inertial Measurement Unit (IMU) and the ElectroMyoGram after providing an Isometric and Isotonic exercises for the patients in the rehabilitation training. This provides an immediate feedback and helps in the faster recovery of the patients.

Dongfang Xu et al (2019) designed an Bionic knee exoskeleton that could be controlled based on realtime gait event detection. Two types of experiments were conducted with the knee exoskeleton, one with zero torque and the other with assistive torque. In both the cases the knee joint could be assisted with lesser delay and higher accuracy for varying speeds under level-ground walking.

Wujing Cao et al (2022) suggested and implemented a lower limb exoskeleton to assist the subjects in walking. The methodology used was an active support during extension and would assist in a passive transfer of load. A carbon fiber rod was used for this purpose in order to reduce the weight and also to minimise the rigid structure. This helped in reducing the net metabolic rate. Further research and development isn't restricted to only adults.

Laubscher et al (2021) proposed an Anthropometrically Parameterized Assistive Lower Limb Exoskeleton for children ranging from 6-11 years old. The user wears an exoskeleton and walks on the treadmill with assistance from the exoskeleton. This particular condition is being compared to the situation where the same wearer walks

ISBN 979-8-3503-5094-4

without the help of the exoskeleton. Henceforth the effort of the user is being quantified. 70W brushless DC motor is being used in this design. This study only evaluates the exoskeleton with a single healthy individual, where the future work entails generalizing the results to a larger scale study with the pediatric population.

This paper was published by Anthony Goo et al(2020) on lower-limb exoskeleton that was used to evaluate the children of 6-11 years old. In order to test all the vital parameters of back drivability, pediatric mannequins were used to go forth with various tests that would result in an exoskeleton that would be suitable for a pediatric application.

II. METHODOLOGY

This section describes the control system interfacing developed for lower-limb exoskeleton's mechanical structure, the knee actuation systems. Figure 3.1 shows the Project Flowof this Project.

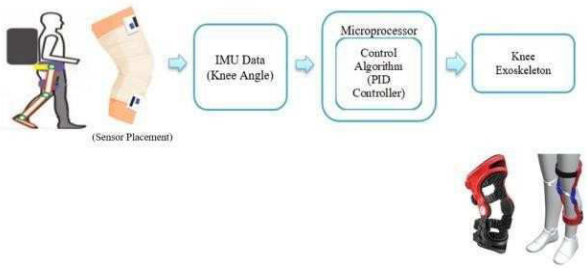


Fig.1 Block Diagram

In this project our goal is to design a Control Algorithm for a Knee Exoskeleton, as shown in Fig. 1, the IMU knee angle data will be implemented into the PID controller which is the control algorithm implemented in Microprocessor unit then the pulsed signal is being transmitted from the Microprocessor to the motor driver which controls and drives the motor in the desired angles which will be attached to a Knee Exoskeleton.

2.1 Data Acquisition

A sensor known as an IMU (Inertial Measurement Unit) is used to measure various parameters, including the specific gravity and angular rate of the body that is attached to the sensor. The data are being stored along with the other parameters in IMU, goal is to extract only the Flexion and Extension Data, so the Flexion and Extension Data alone is acquired for further processing.

2.2 Hardware Configuration

The Hardware configuration module consists of the following hardware components:

2.2.1 Stepper Motor

A stepper motor is a type of motor that is widely used in applications where precise control of rotary motion is required. Conventional motors rotate endlessly, while stepper motors rotate in small precise steps, allowing it to move precisely. A stepper motor converts electrical impulses into mechanical motion. It consists of a series of coils that are energized by an external controller or frequency converter and the rotor moves in small steps. Fig.5 shows the image of a stepper motor and this motor will further be attached to the knee exoskeleton for further processes.

2.2.2 Motor driver (TB6600)

The TB6600 is a stepper motor controller module designed to use bipolar stepper motors in full, half and micro stepper motors. interfacing is done by connecting the microcontroller and stepper motor to the driver and testing the motor. the DIP switches are used to change the voltage in line and evaluation of the motor's operational capabilities by sending controlled signals to the driver to test the motor's functionality.

2.2.3 Microprocessor

The IMU sensor collects angle data, which is later processed by the control algorithm embedded in the Raspberry Pi. This algorithm, written in Python manages the degree of angle adjustment, for the stepper motor in knee exoskeleton. The resulting signal, from the control algorithm is then transmitted to the motor driver. Subsequently a PWM (Pulse width modulation) signal is sent to the motor connected to the knee exoskeleton to achieve the desired angle. The Interfacing of the Stepper Motor with Motor Driver and the Microprocessor is given below in Fig 2.

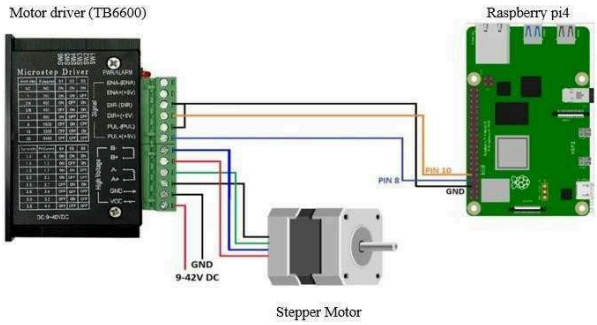


Fig.2 Interfacing of stepper motor

2.1.1 Knee Exoskeleton

A knee exoskeleton is a wearable robotic device that is designed to support, assist, and rehab the human knee joint. Knee exoskeletons are commonly used to help people with mobility issues as they assist users in regaining or improving their ability to walk, stand, and perform a variety of activities. Knee exoskeletons reduce strain on the knee joint, provide stability, and promote proper biomechanics by assisting during the gait cycle. For actuation, this stepper motor will be fixed into the knee region of the knee exoskeleton. Fig.3 shows the 3-D printed Knee Exoskeleton made of PLA (Polylactic Acid).

Fig.3 3-D Printed Knee Exoskeleton

ISBN 979-8-3503-5094-4

2.2 Control Algorithm

The position control in stepper motor is controlled using PID (Proportional-Integral-Derivative) controller, the controller calculates and provides necessary adjustments to the stepper motor driver (TB6600) which will help to attain accurate control over the stepper motor positioning and to maintain the desired target position. The speed control is implemented by the formulae given below in microprocessor, this determines the speed and angle of rotation in the stepper motor. according to the formulae, the step angle is the minimum angle a stepper motor can actuate, micro step is the number of micro steps needed to attain the minimum angle of the stepper motor.

$$Step\ 1 = \frac{Step\ Angle}{Microstep} = Angle\ of\ Rotation \quad (1)$$

Stepper motor has the tendency to move in discrete steps corresponding to the fixed angle of rotation, the PID controller function is to compare both the current position (feedback) along with the desired target position (set point) to control the stepper motor position.

The sum of the proportional, integral, and derivative terms yields the PID controller output (Control Signal). The PID control signal calculation has the following general form:

$$\text{Control Signal} = Kp * Error + Ki * Integral\ of\ Error + Kd * Derivative\ of\ Error \quad (2)$$

According to the error obtained between the desired and the current positioning the control signal that should be applied to the stepper motor to achieve close target position is controlled by the PID controller.

2.3 *Using PID Controller with Knee Exoskeleton*

The PID controller, which uses three components to adjust an output signal based on error (the difference between the desired setpoint and the actual value), is a common control loop feedback mechanism used in control systems. proportional, integral, and derivative. By incorporating a PID controller, the position and speed of the stepper motor can be precisely controlled and adjusted. This system is contained within the knee's exoskeleton, allowing it to adjust automatically to the user's movement and environment. By providing support, stability, and enhanced mobility, this wearable technology improves the wearer's overall mobility.

III. RESULTS AND DISCUSSION

3.1 Angle Data Acquired

Initially, while the subject is walking, the IMU Sensor data is collected by placing it in the knee to shank region to collect the Flexion and Extension data. The knee angle is obtained by combining the IMU data from the thigh and shank region using the sensor fusion technique.

The pitch angle recorded from the IMU sensors estimates the relative angle between the thigh and shank region. The sensor placements and the sensor connection with ESP 32 are shown Fig.4 and the data set acquired in Fig.5 shows the knee angle attained while walking.

Fig.4 Sensor Placement

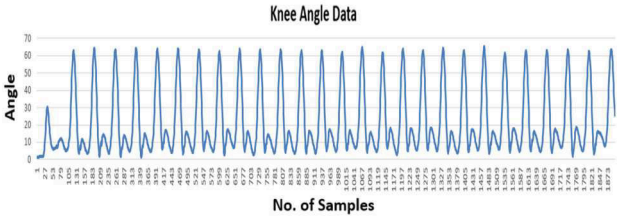


Fig.5 Acquired Knee Angle Data

3.2 PID Controller Simulation

In this project, PID control is employed for position control, given the attachment of exoskeletons to human beings. It is imperative to generate precise degrees of movement in accordance with angle data obtained from the IMU sensor connected to the functional limb.

The simulation block diagram of the PID controller is shown below in Fig. 6. The desired angle is configured as the set point, and the integration of an augmented motor transfer function is used to verify the position control of the motor.

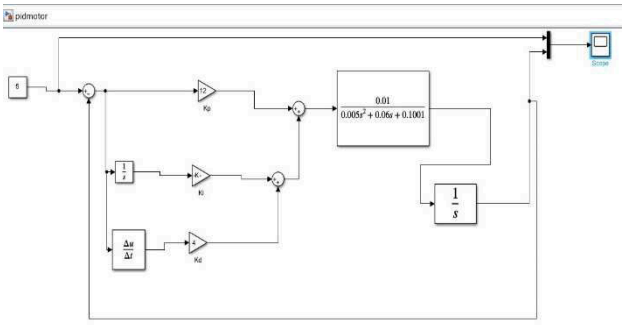


Fig 6 Block diagram of PID controller

To fully comprehend the functionalities of the proportionate component, the simulation began only with its use. The first simulation showed an overshoot where the yellow line indicates the set point (targeted degree) and the blue curve indicates the process variable as shown in Fig 7, which caused the actuation to deviate from the target angle of five degrees.

ISBN 979-8-3503-5094-4

This served as the reference point. In order to improve stability, the derivative and proportional components were integrated in a later stage.

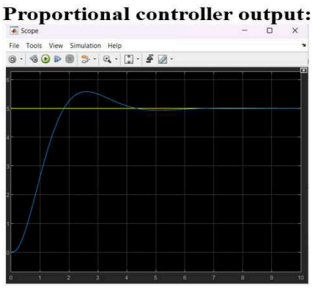


Fig 7. Proportional Controller Output

Moreover, the results that were obtained from the integration of the derivative and proportional components showed a slight overshoot and a time delay as shown in Fig 8.

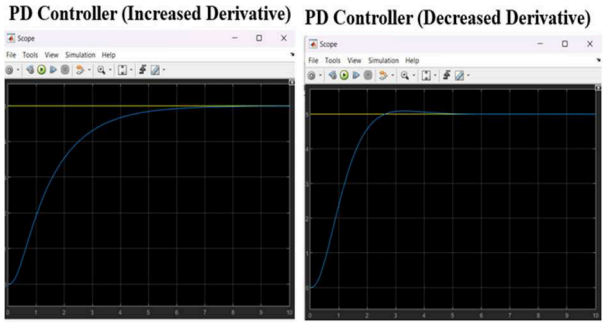


Fig.8 Proportional-Derivative Output

The PID controller was then used for the purpose of achieving more stability. The achieved results exhibit a notably improved stability when contrasted with alternative cases, as clearly indicated in Fig.9 as it doesn't display any overshoot or steady-state error.

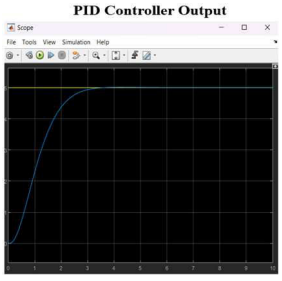


Fig.9 PID Output

With the accomplishment of accurate positional control, the focus now shifts to implementing this control algorithm within the context of the 3D-printed Knee Exoskeleton.

Fig.10 Knee Exoskeleton

The stepper motor is fixed into the knee region of the knee exoskeleton for actuation as it is shown in Fig.10 The Knee Exoskeleton is designed with a gear box which will be attached to the stepper motor which will help to enhance the torque to lift the subject's lower limb irrespective of their weight. Given the successful establishment of positional control through motor pulse generation, the next phase involves the integration of IMU signals to initiate actions within the exoskeleton system.

3.3 *Data Transmission*

To establish a connection between a data acquisition in ESP 32 and control unit Raspberry Pi is done using the ESP-NOW protocol. Initially by configuring the ESP 32 with the appropriate firmware to support ESP-NOW. The MAC addresses of both the ESP device and the Wi-Fi interface of the Raspberry Pi is obtained.

After getting the MAC addresses, the ESP device and the Raspberry Pi are set up with matching ESP-NOW credentials, including the same communication channel and encryption key. Implement code on both the ESP device and the Raspberry Pi to enable communication between the devices. The code facilitates the transmission and reception of data using the ESP-NOW protocol. So now the PID controller does the controlling as the angle data from IMU sensor is being transmitted. The Real-time control algorithm for Knee Exoskeleton is implemented and the results are given in the following figures.

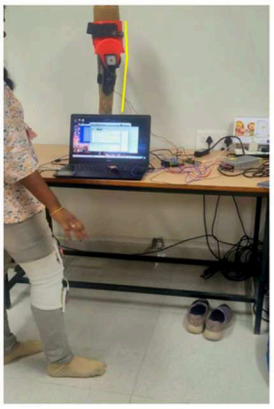


Fig.11 Real time Extension Actuation

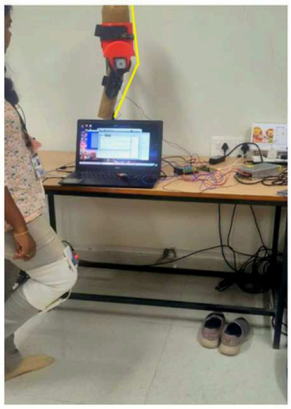


Fig.12 Real time Flexion Actuation

IV. CONCLUSION

The Knee Exoskeleton control algorithm is designed and implemented in the Limb model to provide different angles of actuation for the knee prototype model, which will help patients with gait rehabilitation. The observed angle degree is accurate even with a small transmission-related actuation delay. The mechanism makes sure that the range of motion is exactly in line with the knee's intended range of flexion and extension.

The project's results may eventually result in the addition of a few actuators to the Knee Exoskeleton which would allow the exoskeleton to assist with even greater force, which would allow it to support heavier weights and negotiate difficult terrain. Enhancing the microcontroller further improves control, leading to more flexible and seamless movements. Together, these enhancements produces more versatile and capable exoskeleton for various tasks. The nature of the project allows for future modifications, such as changing the actuators, designing a much more robust control system, and modifying the motor used with a higher torque range.

REFERENCES

[1] Anthony Goo, Curt A. Laubscher, Ryan J. Farris and Jerzy T. Sawicki, (2020), "Design and Evaluation of a Pediatric Lower-Limb Exoskeleton Joint Actuator", Actuators for Robotics, 2020, VOL-9 pp 4-20

ISBN 979-8-3503-5094-4

[2] Curt A. Laubscher, Ryan J. Farris, Antonie J. van den Bogert, Jerzy T. Sawicki, "An Anthropometrically Parameterized Assistive Lower Limb Exoskeleton", *J Biomech Eng*. Oct 2021, 143(10): 105001

[3] Dongfang Xu, Xiuhua Liu and and Qining Wang, Senior Member, IEEE, (2019) "Knee Exoskeleton Assistive Torque Control Based on Real-Time Gait Event Detection", *IEEE Transactions on Medical Robotics and Bionics*, vol. 1, no. 3, pp. 158-168

[4] Jingshuai Liu, Yong He, Jiantao Yang, Wujing Cao, Xinyu Wu (2022), "Design and analysis of a novel 12-DOF self-balancing lower extremity exoskeleton for walking assistance", Mechanism and Machine Theory 167 (2022) 104519.

[5] Kim Jong- un, Kim Ga -eul, Ji Young- beom, Lee A-ram,Lee Hyeon-joo, Tae Ki-sik, (2019) "Development of a Knee Exoskeleton for Rehabilitation Based EMG and IMU Sensor Feedback", Journal of Biomedical Engineering Research, vol. 40, no. 6, pp. 223–229

[6] Ling Luo, Yuxia Yuan and Zhijun Li, (2019) "Design and Development of a Wearable Lower Limb Exoskeleton Robot", IEEE 4th International Conference on Advanced Robotics and Mechatronics (ICARM) pp. 599-604

[7] Miguel Sanchez-Manchola, Daniel Gomez-Vargas, Diego Casas-Bocanegra, Marcela Munera Member, IEEE, Carlos A. Cifuentes1 Member, IEEE, "Development of a Robotic Lower-Limb Exoskeleton for Gait Rehabilitation: AGoRA Exoskeleton", 2018 IEEE ANDESCON, 2018, pp. 1-6

[8] S.K. Hasan, Anoop K. Dhingra (2022), "Biomechanical design and control of an eight DOF human lower extremity rehabilitation exoskeleton robot" Department of Mechanical Engineering, University of Wisconsin Milwaukee, Milwaukee, WI 53211, United States of America, Results in Control and Optimization 7 (2022) 100107.

[9] Wujing Cao, Chunjie Chen, Dashuai Wang, Xinyu Wu, Lingxing Chen, Tiantian Xu, and Jingshuai Liu (2022), "A Lower Limb Exoskeleton with Rigid and Soft Structure for Loaded Walking Assistance", IEEE ROBOTICS AND AUTOMATION LETTERS, VOL. 7, NO. 1, JANUARY 2022: pp. 454-461.

[10] Zhijun Li, Senior Member, IEEE, Kuankuan Zhao, Longbin Zhang, Xinyu Wu, Tao Zhang, Qinjian Li, Xiang Li, and Chun-Yi Su, Senior Member, IEEE (2020), "Human-in-the-Loop Control of a Wearable Lower Limb Exoskeleton for Stable Dynamic Walking", IEEE/ASME TRANSACTIONS ON MECHATRONICS, VOL. 26, NO. 5, OCTOBER 2021, pp2700-2711.

ISBN 979-8-3503-5094-4

Mobile Robotic Manipulator for Material Transfer with Disinfection Unit

Nithyaa A N[1], Premkumar R[2] Geetha Anandhi C[3], Gokul M[4], Durga R[5]

[1, 2, 5] Department of BME, Rajalakshmi Engineering College, India

[3]Department of BME, Dhanalakshmi Srinivasan College of engineering, India

[4] PhD Scholar at Ben-Gurion University of the Negev, Beersheba, Israel

nithyaa.an@rajalakshmi.edu.in[1] ,premkumar.r@rajalakshmi.edu.in[2], geethaanandhi423@gmail.com[3], mrgokul3@gmail.com[4] , 200301017@rajalakshmi.edu.in[5]

***Abstract*— Since late 2019, the COVID-19 outbreak has spread all over the world. The pandemic poses a serious danger to the worldwide population's safety and well-being, as well as the medical community and clinical infrastructures across the world. It has been universally proposed that using robots during the epidemic would enhance patient care and have an impact on the clinical framework. In this study, we built a transportable 5axis robotic manipulator out of acrylic resin that can transfer the materials from one place around inside confinement spaces in hospitals and other healthcare facilities. A wheeled vehicle has been equipped with the robotic arm. A portable sanitising device with an alcohol-containing chemical sterilant will be mounted to its back. The robotic arm will move to pick the item from the red container, set it down in the desired location with blue container, and then spray the sterilant all over it. Python has been used to create the colour detecting code. Raspberry Pi is utilised for a hardware implementation. Depending on the object's position and placement, such as whether it is in the centre, to the left or to the right, the robotic arm will move accordingly. The robot will immediately stop moving and begin picking when the object is extremely close to the camera. If the object is distant from the camera, however, the robot will advance until it is close to the object.**

Keywords: ***Pandemic, Self-sterilizing system, material transfer, spraying, Robotic arm, python, Raspberry Pi, Image processing, Colour recognition***

I INTRODUCTION

Since late 2019, the COVID-19 illness has been spreading over the world. It generates basic challenges for practically all governments on the planet. According to World Health Organization figures, there have been 57,639,631 confirmed COVID-19 cases with 1,373,294 fatalities globally until November 2020. (WHO) [1]. Various techniques and innovations, including robots, are recommended to aid in the therapy and control of the pandemic. Previously, a few experts briefly summarized the automated applications during a pandemic and classified robot utilization in medical services into various classes, including secretary, nurture, emergency vehicle, telemedicine, serving, cleaning, showering/disinfestations, cautious, radiologist, restoration, food, and open-air conveyance robots. Robot uses vary according to the desired regions, such as robots in medical clinics, networks, air terminals, transportations, entertainments, accommodations, cafés, attractions, and grand spaces [2].

This article is about developing a mobile robotic manipulator for material transfer with disinfection where it is a wheeled portable mini robotic arm which involves in transfer of object from one place around inside confinement spaces in hospital and other healthcare facilities to another. A wheeled vehicle has been equipped with the robotic arm [3]. Automated technologies and robotics have substantially altered industrial and routine life in recent years. Also, there is a lot of growing interest in using contagious robots' disease and widespread scenarios. [4]. Since robots have the chance to deliver the task in a contagious or dangerous environment without infecting or affecting people. The most important application in the infectious situation is robotic surgery, especially robot-assisted minimally invasive surgery [5]. The robot performs automated laparoscopy and laparoscopic surgery for more protection and fewer difficulties. Because of their unique character, some diseases are discussed, such as associated operations such as prostatectomy and hysterectomy. In general, robots that are resistant to infection and simple to disinfect are recommended for important role in fighting the widespread. During the widespread, robots are widely advocated for use in a variety of settings to assist prevent infection by doing cleaning, monitoring, delivery, meal preparation, and telepresence. Clinic contamination control operations must include sanitization and sterilization [6]. The medical equipment or tool that interacts with the patient's sterile tissue or bodily fluid film during the various cycles is associated with an increased risk of microorganism presentation into the patient's body [7]. Furthermore, infection can be transmitted from one patient to another; from patient to medical care work force, and vice versa, via improperly sterilized or disinfected instruments [8]. Many illnesses have been documented across the world as a result of the usage of infected endoscopes. As a result, proper disinfection measures for medical and surgical items are required in all health care institutions. As a result, medical staff, laboratory personnel, and health care providers should have a better understanding of these approaches in order to keep away from the outspread of these diseases [9].

In this study, we built a transportable 5axis robotic manipulator out of acrylic resin that can transfer the materials

ISBN 979-8-3503-5094-4

from one place around inside confinement spaces in hospitals and other healthcare facilities [10]. Here by interfacing USB camera and SD card with Raspberry pi. Later designing a moving robotic arm model and by interfacing the raspberry pi with the robotic arm [11]. Then robot using the camera will immediately stop moving and begin picking when the object is extremely close. Here need to train the robotic arm to recognize colors based on the RGB color space algorithm to allow the movable robotic arm to correctly recognize the container's color and to make the manipulator to transfer the material from source to destination also to sanitize the transferred instruments automatically.

II. MATERIALS AND METHODS

Based on the proposed methodology for hardware implementation, we are employing a Raspberry Pi with Raspbian OS, which is based on Debian, a Linux operating system. Using Image Processing, the camera recognizes the pre-specified items and disposes of them based on color (RGB) [12]. The software involves commanding the robotic arm, taking the item picture, analyzing the image, detecting the RGB object, and automating all tasks using Raspberry Pi.

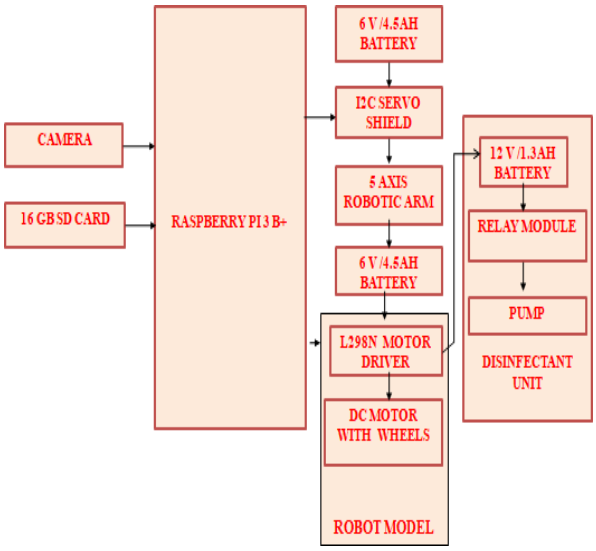


Figure 1- block diagram of the proposed methodology

The block diagram of a mobile robotic manipulator for material transfer with disinfection has been presented in this article figure 3, along with the components that will be used. The Raspberry Pi is the general controller of the system and serves as the system's heart. The material transfer unit consists of a 5-axis robotic manipulator with servomotors driven by a 6 V battery. A USB camera will be mounted in front of the wheeled vehicle and will be directly connected to the Raspberry Pi's USB connection. Because the controller has an external storage option, programming's are stored on a 16 GB SD card. The wheeled vehicle has four wheels and a DC motor. The L298N motor driver will power it. The motor driver is powered by a 12 V battery.

Python is used to write the code. The two hues used for material transfer are red and blue. When the application begins to run, the camera activates and looks for the red container. As soon as the camera recognizes the red color, the wheeled vehicle begins to travel towards the red colored container [13]. It has been designed in such a manner that if the red container is too close, the wheeled vehicle will stop. As soon as the vehicle comes to a complete stop, the arm begins to rotate 90 degrees and takes up the instrument with the gripper. If the red container is too far away, the forward condition will activate and halt until it reaches the center of the container. The position of the container will be displayed in the VNC viewer. When the gripper captures the instrument, the camera begins to search the region in blue. The same techniques are used as when looking for the red color. When
the robot approaches the blue zone and grabs its center, the gripper opens and the instrument or substance is released. Figures 4.1,4.2,4.3 depict the various viewpoints of the robotic unit, including the side, rear, and front views. The parts of robot prototype are mentioned in below figure 1 and figure 2.

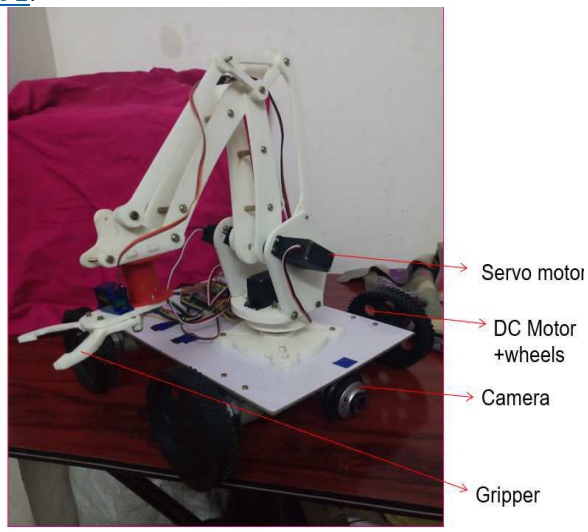


Figure 2- front view of the robot prototype

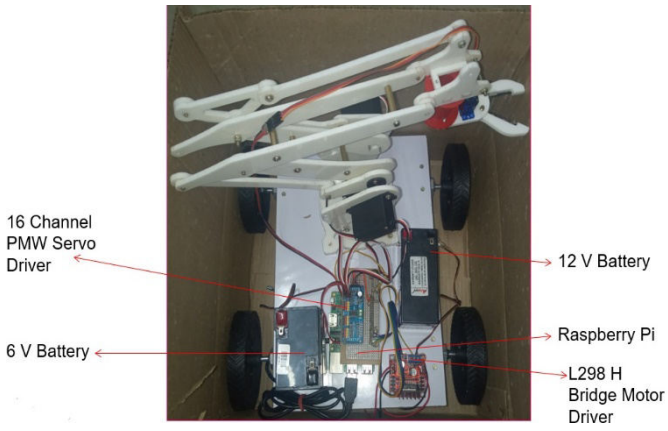


Figure 3- top view of the robot prototype

ISBN 979-8-3503-5094-4

A. Raspberry pi3 b plus (overall controller)

The Raspberry Pi is an inexpensive, small-sized portable computer that connects to a computer display and operates regularly. It is a competent tiny gadget that allows individuals of all to experiment with computing and determine to write in languages. [14] The Raspberry Pi 3 Model B+ outperforms the Model B in numerous ways, including a higher CPU clock speed which greater Internet throughput, and dual-band Wireless internet. It also supports Energy than Ethernet via the use of an Ethernet power HAT.

B. 16 GB micro SD card

Micro SD Card is a Flash-based memory card developed to meet the privacy, storage, speed, and natural assets of emerging audio and video electronics [15]. The Internal Memory Flash Memory host connector, which is based on an improved 8-pin interface, allows regular Interactive Multimedia Card operation.

C. 6v/4.5ah Battery-q2

Lead-lead dioxide systems are used in rechargeable batteries [16]. Absorbent Glass Mat (AGM) technique offers up to 99% effective gas recombination and no need for electrolyte maintenance or water addition. It is a UL-certified component. It may be installed in any direction. It contains a lead, calcium tin alloy grid that was computer developed for high power density. It has a long service life and may be used in float or cyclic applications. It operates without maintenance and has a low self-discharge.

D.USB webcam

A USB webcam is a camera that connects to a PC by attaching it to a USB port on the machine. The product you select can be configured to upload images over time to a website via FTP (file transfer protocol) [17]. When FTP is used, the picture on the site is static and is refreshed on a regular basis at lengths of step by step, step by step, or significantly more. USB webcam programming may also be used to provide a live feed directly from the PC to your site or even connected to the Cam secure real time function to allow large numbers of visitors to your site to watch the broadcast at the same time.

E. L298n motor driver

This module is intended for use with the Raspberry Pi Motor Driver Function board. A Dual H drive module employs the ST L298N dual full-bridge driver, an unified monolith circuit in a 15-lead milliwatt and PowerSO20 packaging. The gadget may be activated or deactivated independently of the inputs through two enable inputs. The bottom transistor emitters of each bridge are coupled together, and the corresponding external connection may be utilized to connect an auxiliary detecting resistor [18].

F.DC motor with wheels

A direct current motor is a type of rotary electrical motor that transforms and act as transducer as it converts the direct current electrical energy into mechanical energy. The most prevalent forms are based on magnetic field forces [19]. DC motors were the first commonly utilised type of motor because they could be driven by existing direct-current lighting power distribution networks.

G. 12 v Battery

Lead-lead dioxide systems are used in rechargeable batteries. Absorbent Glass Mat (AGM) technique offers up to 99% effective gas recombination and no need for electrolyte maintenance or water addition [16]. It is not limited to air transport and conforms with IATA/ICAO Special Provision A67.

H. Servo motors

A servo motor is a type of motor that can rotate precisely. It is just a motor operated by a servo system. A DC servo motor is one that is supplied by a direct current (DC) power source [20]. A servo motor is frequently outfitted with a gear system, allowing us to acquire a very high torque servo motor in small and lightweight designs (TABLE 1). Most hobby servo motors are rated in (kg/cm), with the most common being 3 kg/cm, 6 kg/cm, or 12 kg/cm.

Table 1- servo motor specifications

S.No	JOINT NAME	SERVO MOTOR SPECIFICATIONS	SERVO MOTOR NAME
1	Gripper joint	Torque of 1.8 kg/cm at 4.8v	Tower pro SG 90
2	Other joints (Wrist joint, Shoulder joint Base joint)	Torque of 9.4 kg/cm at 4.8v	Tower pro MG 995

I .16 channel i2c servo driver

The Adafruit 16-Channel 12-bit PWM/Servo Driver has only two pins but can drive up to 16 servos via I2C. The on-board PWM controller will drive all 16 channels simultaneously with no additional Arduino processing overhead [21]. The Adafruit PWM/Servo Driver is suitable for any project that necessitates the use of a large number of servos.

J. Disinfectant tank

The reservoir tank figure 5, pump, and hose comprise the disinfection system. The disinfectant containing alcohol will be placed into the reservoir tank [22]. When the motor is turned on, the disinfectant is pushed and sprayed via the exit tube. It's a clear tank with a wiper motor inside.

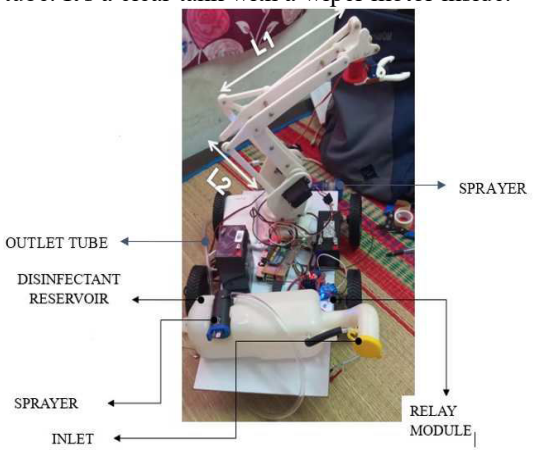


Figure 4- top view of the robot with disinfection unit

ISBN 979-8-3503-5094-4

Figure 5-disinfectant reservoir

The figures of disinfection unit are shown in figure 4. Here in figure 4 the top view of disinfection unit the L1, L2 represents the linkages.

k. Gripper tool

The amount of force required to grip an item is referred to as grabbing force. The frictional force is the most crucial of the forces acting on the body carried by a robotic arm. To protect the object's surface, the gripping surface might be made of a soft material with a high coefficient of friction. The robotic gripper must deal with not just the object's weight, but also the acceleration and motion caused by the object's continual movement. The force required to retain the item is calculated using the following formula:

$$F=ma/\mu n$$

Where, F = force needed to grip the object.
m = mass, a =acceleration of the object
μ = co-efficient of friction
n = number of fingers in the gripper
Based on the strength of the gripper the force voltage and weight resistance tabulation and graph are mentioned below.

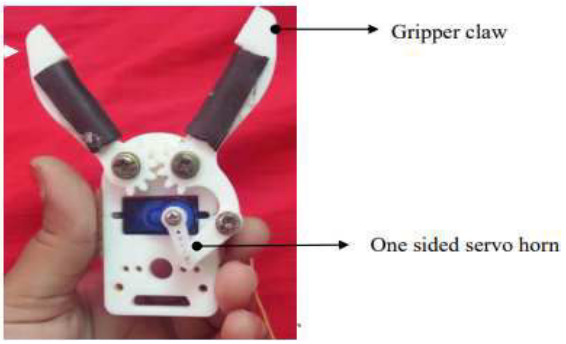


Figure 6- gripper tool (rear view)

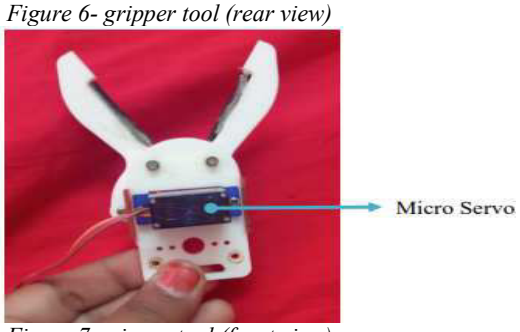


Figure 7-gripper tool (front view)

Here, building a robot that can identify and place a pre-specified item. Python was used to write the colour detection code. For hardware implementation, we are employing a Raspberry Pi with Raspbian OS, which is based on Debian, a Linux operating system. Using Image Processing, the camera recognises the pre-specified items and disposes of them based on colour (RGB). The movement of the robotic arm will be based on the object's position and location i.e. whether object is at centre or left or right. Whenever the object is very close to the camera, automatically the robot will stop moving and start picking and whenever the object is far away from the camera then, the robot will move forward until it reaches near the object.

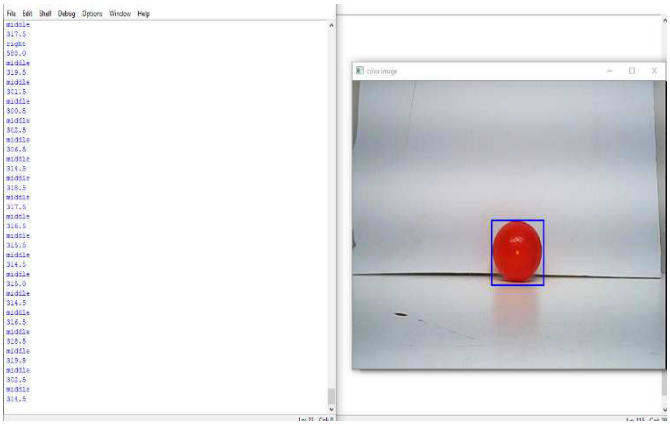


Figure 8- The object with python code (in the middle)

The software involves commanding the robotic arm, taking the item picture, analysing the image, detecting the RGB object, and automating all tasks using Raspberry Pi (Mini Computer). The software requirements are like the Raspbian OS, Open CV, Python IDLE and Virtual Network Computing (VNC). Raspberry Pi OS (formerly Raspbian) is a Debian-based software platform for the Device that is based on Debian. OpenCV is a function library for development geared mostly at real-time computer vision. Python, Java, and MATLAB/OCTAVE bindings are available. IDLE is a python programming integrated environment shown in figure 8 for programming that comes with the language's default implementation.
The lower the resistance, the greater the force delivered to the sensor. A force was applied to the detector, which was placed on top of a weigh, and it was also attached to a multi-meter. The amounts of force (g) and resistance (k) were measured concurrently.

III. RESULTS AND DISCUSSION

In theory, your robot moves the distance equal to the circle of the wheel every time your wheel makes a whole revolution. So, multiply the circumference by the number of revolutions per minute to find the distance travelled by your robot in a minute. The wheeled vehicle travels 15.6 metres in 60 seconds (per minute).
The velocity of a wheeled vehicle is calculated using the following formula:
Velocity = circumference of the wheel * rpm
The circumference of the wheel = $2\pi r$ or πd (π=3.14).
Where, rpm is the number of rotations per minute (motor)
r is the radius of the wheel, and d is the diameter of the wheel (2r).

ISBN 979-8-3503-5094-4

The distance travelled by a wheeled vehicle is measured by determining how far the vehicle drives in one second, two seconds, and so on (Table 2) According to Figure 9, distance and time are directly proportionate. As a result, the robot's velocity remains constant in the absence of a loading condition. Distance is also measured in centimetres. The centimetre lengths were converted to metres by dividing 100 by 100. After the distance was determined, it was discovered that the distance increased at a steady rate. The wheeled vehicle travelled 26 cm in 1 second (0.26 m). The formula was used to compute the velocity. Velocity (m/s) = Distance (m)/Time (s).

Table 2-robot velocity

S.No	TIME(s)	DISTANCE(m)	VELOCITY(m/s)
1	1	0.26	0.26
2	2	0.52	0.26
3	3	0.78	0.26
4	4	1.04	0.26
5	5	1.3	0.26
6	6	1.56	0.26
7	7	1.82	0.26
8	8	2.08	0.26

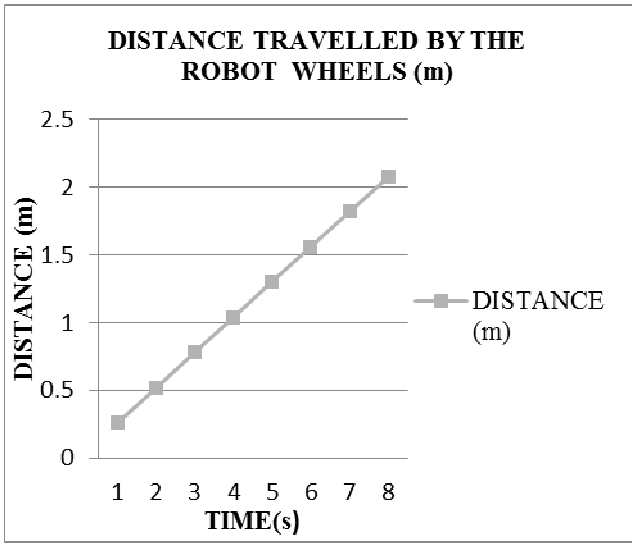


Figure 9- Distance travelled by the robot wheels(m)

The friction coefficient will be between 0 and 1. If there is friction, the value of will be 1; if there is no friction, the value will be 0. Here the body is propelled towards up against gravitational pull, then the required force should be more than that necessary when moving towards down. Then the formula is:

F= m (a+g)/μn

Where, g=acceleration due to gravity (9.8 m/s2)

The type of gripper used in the proposed system is mentioned in figure 6 and 7 as rear view and front view.

Table 3-force -voltage

S.No	APPLIED FORCE(N)	FSR VOLTAGE(V)
1	0.294	2.2
2	0.34	2.65
3	0.49	2.9
4	0.588	3.08
5	0.686	3
6	0.784	3.22
7	0.882	3.45
8	0.98	3.57

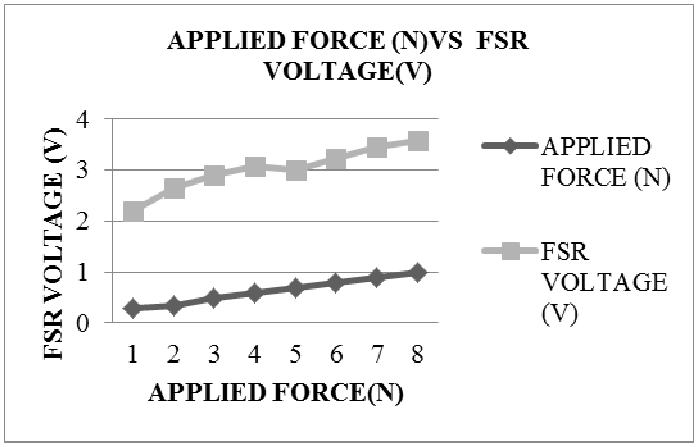


Figure 10- applied force (n)vs fsr voltage(v)

Here in the TABLE-3 it is shown about the force voltage and in figure 10 the comparison between applied force and FSR voltage has been shown. Also, in the TABLE-4 weight resistance and in the figure 11 the comparison between the objects weight and resistance has been shown.

The distance-time graph (Figure:10) shows that distance is directly related to time. The robotic wheels accelerate at a constant pace, as seen by the (TABLE:2) In 60 seconds, the wheeled vehicle moves 15.6 meters. The gripper's strength was determined using a force sensitive resistor. The gripper tool can lift up to 100 gms. The resistance value and accompanying voltage value were estimated using FSR. The resistance is inversely proportional to the force exerted,

ISBN 979-8-3503-5094-4

which means that as the force increases, the resistance lowers. This link is seen in (figure 11).

Table 4-weight-resistance

S.No	WEIGHT(g)	RESISTANCE(KΩ)
1	30	210
2	40	109.2
3	50	33.7
4	60	22.1
5	70	16.7
6	80	14.05
7	90	12.55
8	100	11.55

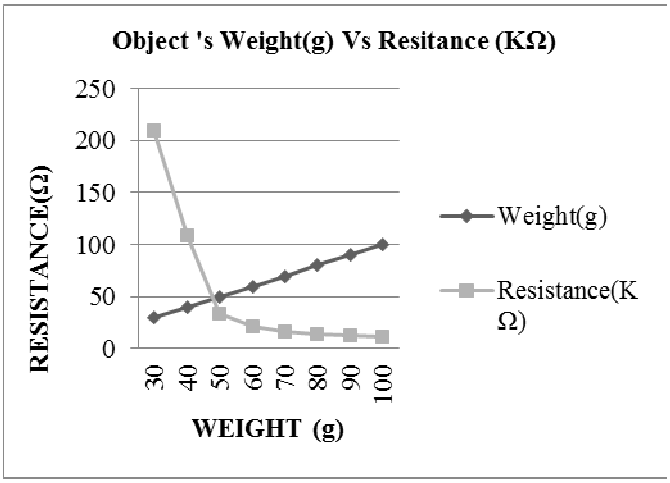


Figure 11- Object's weight(g) Vs Resistance(kΩ)

From the distance-time graph (Figure 9), it is clear that the distance is linearly proportional to the time. The robotic wheels accelerate with a constant velocity, which is clear from the (TABLE: 2) the wheeled vehicle travels 15.6 meters in 60 seconds. The gripper's strength has been measured with the help of force sensitive resistor. The gripper tool is able to lift up to 100 gms. From the distance-time graph (Figure 9), it is clear that the distance is linearly proportional to the time. The robotic wheels accelerate with a constant velocity, which is clear from the (TABLE: 2) the wheeled vehicle travels 15.6 meters in 60 seconds. The gripper's strength has been measured with the help of force sensitive resistor. The gripper tool is able to lift up to 100 gms.

IV. CONCLUSION

The primary goal of this work is to provide a transportable 5axis robotic manipulator made of acrylic resin that can move materials from one location to another inside confinement zones in hospitals and other healthcare institutions. The robotic arm has been installed on a wheeled vehicle. A portable sanitizing device with an alcohol-based chemical sterilant will be put on its back. The robotic arm will move to collect the object from the red container, place it in the proper area with the blue container, and then spray it with the sterilant. The colour detecting function was written in Python. For a hardware implementation, the Raspberry Pi is used.

V.REFERENCES

1. Nachega, Jean, Moussa Seydi, and Alimuddin Zumla. "The late arrival of coronavirus disease 2019 (COVID-19) in Africa: mitigating Pan-continental spread." Clinical Infectious Diseases 71.15 (2020): 875-878.

2. Bashir, Sajitha, et al. The Converging Technology Revolution and Human Capital: Potential and Implications for South Asia. World Bank Publications, 2021.

3. Kolpashchikov, Dmitrii, Olga Gerget, and Roman Meshcheryakov. "Robotics in Healthcare." Handbook of Artificial Intelligence in Healthcare. Springer, Cham, 2022. 281-306.

4. Wang, Xi Vincent, and Lihui Wang. "A literature survey of the robotic technologies during the COVID-19 pandemic." Journal of Manufacturing Systems 60 (2021): 823-836.

5. Hanly, Eric J., and Mark A. Talamini. "Robotic abdominal surgery." The American journal of surgery 188.4 (2004): 19-26.

6. Wang, Jiao, et al. "Disinfection technology of hospital wastes and wastewater: Suggestions for disinfection strategy during coronavirus Disease 2019 (COVID-19) pandemic in China." Environmental pollution 262 (2020): 114665.

7. Donlan, Rodney M., and J. William Costerton. "Biofilms: survival mechanisms of clinically relevant microorganisms." Clinical microbiology reviews 15.2 (2002): 167-193.

8. Mohapatra, S. "Sterilization and disinfection." Essentials of neuroanesthesia. Academic Press, 2017. 929-944.

9. Plebani, Mario. "The detection and prevention of errors in laboratory medicine." Annals of clinical biochemistry 47.2 (2010): 101-110.

10. Nachega, Jean, Moussa Seydi, and Alimuddin Zumla. "The late arrival of coronavirus disease 2019 (COVID-19) in Africa: mitigating Pan-continental spread." Clinical Infectious Diseases 71.15 (2020): 875-878.

11. Bashir, Sajitha, et al. The Converging Technology Revolution and Human Capital: Potential and Implications for South Asia. World Bank Publications, 2021.

12. Chan, Ho Nam, Ming Jun Andrew Tan, and Hongkai Wu. "Point-of-care testing: applications of 3D printing." Lab on a Chip 17.16 (2017): 2713-2739.

13. Thomas, Ron Oommen, and K. Rajasekaran. "Remote monitoring and control of robotic arm with visual feedback using

ISBN 979-8-3503-5094-4

Raspberry Pi." International Journal of Computer Applications 92.9 (2014).

14. Nayak, Maya, and Prasannajit Dash. "Smart surveillance monitoring system using Raspberry Pi and PIR sensor." Statistics (2014).

15. Andreone, L., et al. "Vehicle detection and localization in infra-red images." Proceedings. The IEEE 5th International Conference on Intelligent Transportation Systems. IEEE, 2002.

16. OMRAN, MUSTAFA ASAAD. "REPUBLIC OF IRAQ MINISTRY OF HIGHER EDUCATION AND SCIENTIFIC RESEARCH AL-FURAT AL-AWSAT TECHNICAL UNIVERSITY."

17. Abdenaa, Sekata Olika, and Vasu Pinnti. "An Efficient Real-Time Controller for Retrieving Multimedia Data from Secured Digital High Capacity Card." (2017).

18. Rajavenkatesan, T., et al. "Photo Voltaic Cooling Can." International Journal of Communication and Computer Technologies 5.1 (2017): 17-22.

19. Roshan, Rakesh. Remote USB Ports. Diss. 2013.

20. Dhumale, Samruddhi, et al. "Follow Me Suitcase." International Journal of Research in Engineering, Science and Management 4.2 (2021): 146-148.

21. Venkatesh, K., et al. "IoT based home automation using raspberry Pi." J. Adv. Res. Dyn. Control Syst 10.7 (2018): 1721-1728.

22. Fulton, Donald E. "Current sense circuit for a DC powered three phase servo motor controller." 2006 IEEE Workshops on Computers in Power Electronics. IEEE, 2006.

23. Qabajah, Haneen, and Kifah Abdeen. "Object Sorting System Using Robotic Arm." (2018).

24. Jawitz, James W., et al. "In-situ alcohol flushing of a DNAPL source zone at a dry cleaner site." Environmental Science & Technology 34.17 (2000): 3722-3729.

25. Kotkar, Ajit, et al. "Android Based Remote Desktop Client." International Journal of Innovative Research in Computer and Communication Engineering 1.2 (2013): 345-348.

ISBN 979-8-3503-5094-4

ISBN 979-8-3503-5094-4

E-Crutch: A Promising tool for Rehabilitative Patients

A. Josephin Arockia Dhivya
Assistant Professor, Department of Biomedical Engineering, Vels Institute of Science, Technology and Advanced Studies,
Chennai, Tamil Nadu, India
a.dhivya.se@velsuniv.ac.in

Dr. R. J. Hemalatha
Associate Professor, Department of Biomedical Engineering, Vels Institute of Science, Technology and Advanced Studies,
Chennai, Tamil Nadu, India
hemalatharj@velsuniv.ac.in

Jaya Rubi
Assistant Professor, Department of Biomedical Engineering, Vels Institute of Science, Technology and Advanced Studies,
Chennai, Tamil Nadu, India
jayarubi.se@velsuniv.ac.in

***Abstract*—In earlier days,physiotherapists trained the rehabilitative patients in a progressive manner such as giving them a solution for their aid.But the major disadvantage is to find the issues faced by the elders while using the crutch without their knowledge.Our designed e-crutch helps us to train and alert the user to use the device wisely without any side effects.It majorly helps the patients suffering from long term spondylysis,fracture and spinal stenosis.The pain is in a excessive manner since the patients are unaware of their usage.The area which we highlight in our device is force,acceleratio,non invasive varicose vein therapy and lighting system.The first case is alerting the user when he s putting excessive force on the crutch without his/her knowledge will be notified by a haptic vibration. The posture of the body is a very important factor in the field of physiotherapy where the posture condition of patient moving excessively forward and backward will be displayed in the Lcd display present in the crutch module.A non invasive therapy system is also enclosed with the crutch to give a relief based on temperature and these modules together is termed a E-crutch.The notification of the user parameters will be sent to the care taker via GSM module.This particularly helps in the case of Alzheimer affected individuals.So,our ultimate aim is to design a E-crutch which will be useful to the elderly society.**

Keywords—E-crutch,elders,force,parameters,issues, posture,non invasive,varicose

I. INTRODUCTION

Nowadays,human beings lack physical work due to over usage of gadgets, follows an inactive lifestyle which leads to excessive body and muscle pain which affects cell metabolism resulting in muscle fatigue.The other major drawback is Kyphosis which leads to a slouching pattern of body posture

.To avoid these problems a crutch is designed with acceleration sensors to sense the mode of posture if the users are unaware of it and the other important feature is to give a temporary relief for the patients suffering from varicose veins.This E-crutch is well designed to serve the elderly people.

II FACTORS TO BE CONSIDERED WHILE USING CRUTCH

Overusage of crutches without limits leads to severe pain in joints and muscle related disorders.Certain measures can be taken to reduce the over usage of crutches.It includes:

- A clear assessment must be taken to check the challenges of the rehabilitative patients to analyze the issues.
- Biomechanical analysis play a vital role in assessing the biological factors involved in a individual due to misuse of the device.The crutch frame play a vital role ,being light weight and durable the patient finds it easy to use and move.
- Fabrication and integration of sensors should be easy in a crutch frame so that the data can be processed and analysed easily.
- Interfacing of users should be improved in a friendly manner and proper trials must be done to validate the E-crutch.Based on the results and accuracies the device can be improved.

III LITERATURE REVIEW

The literature survey highlights the existing walking aids which concentrates in measuring the falling of patient condition in real time.Further studies also reveals the posture management issues faced by the user to monitor the patients.Past studies also states the usage of Ultrasonic sensors in crutch. Piyali Sengupta , Kiran Mondal , Hiranmoy Mahata, Sujaya De, Prakash C Dhara,[1] stated that real time parameters should be measured in modern crutches.Claire Perezand , [2] clearly stated the fall of the elderly patients based on different weights so as to evaluate the durability of the crutch.Portegijs E [3] proposed the importance of weight bearing and providing feedback using a crutch.

IV. PROPOSED SYATEM

The block diagram, clearly states the inclusion of sensors in the crutch frame which will be measured in real timeThe parameters to be measured are:

ISBN 979-8-3503-5094-4

A.FORCE:

Two force sensors will be embedded on the walker in order to alert the rehabilitative patients if they apply more stress on it[4].Proper instructions must be given by the physiotherapist for the patients regarding the usage of crutch based on the threshold values[5].The values changes according to the medical condition of the patient.These sensors are carefully fabricated and embedded into the crutch frame. When the applied pressure is above the threshold limit of the individual,the buzzer and a haptic indicator will alert the patient.[6] according to the condition of the physiotherapist.

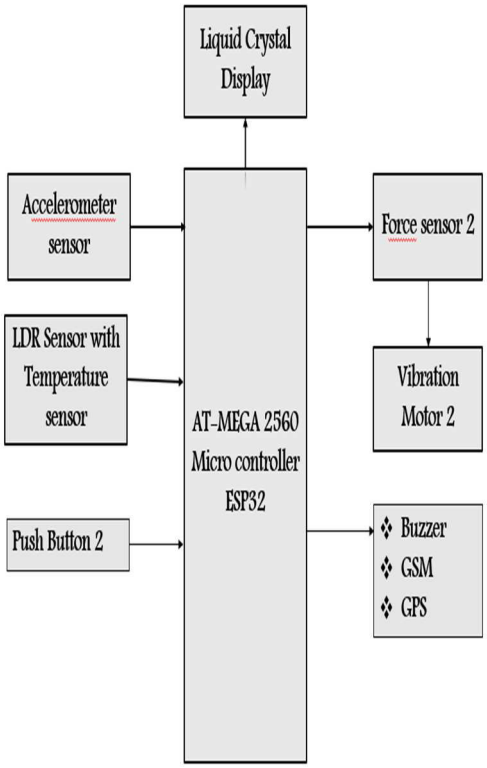


Fig. 1. Block diagram of sensors to be integrated into crutch tool

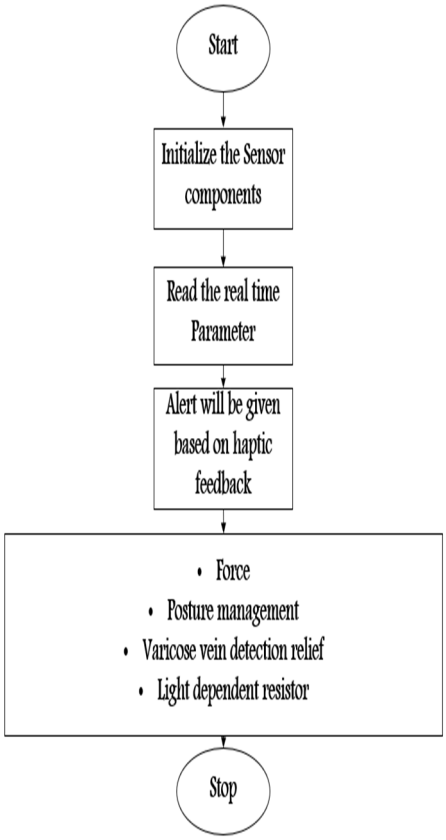


Fig. 2. Flow of E-Crutch working module

B.POSTURE MANAGEMENT:

Accelerometer sensors are employed to detect the abnormal changes in body posture while the individual is using the crutch and a haptic feedback will be notified to the user.[7].When the diseased individual is straining a lot in front and back while using a crutch,he/she will be notified by a statement stating the staus of the posture whether it is forward or backward[8].The buzzer will alarm when the patient is not using the crutch properly.

C.VARICOSE VEIN THERAPY :

During these days many people suffer with varicose veins which results in serious problems like inflammation. We use NTC thermistors for measuring the temperature of the patient in the affected area and also in the normal body.to obtain pressure variations with micro motors[9]. The method includes continuous monitoring of temperature of the affected area when the temperature of the affected area becomes more than the normal body temperature. If improper blood flow is detected we apply vibrations automatically to make the blood flow proper in the desired area[10].

D.LIGHT DEPENDENT RESISTOR:

This sensor is mainly used when the patient is using the E-crutch in an external environment.The light automatically

ISBN 979-8-3503-5094-4

glows when the crutch is in dark mode and switches off when it is in light environment[11].

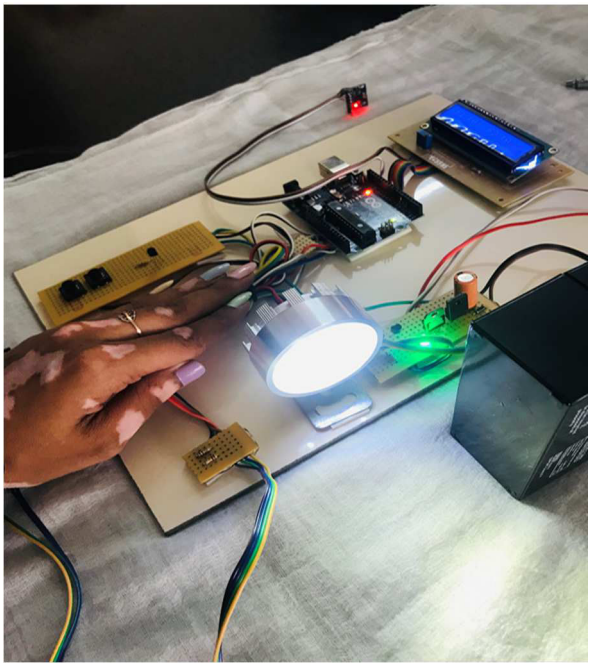


Fig. 3. Module of LDR Sensor

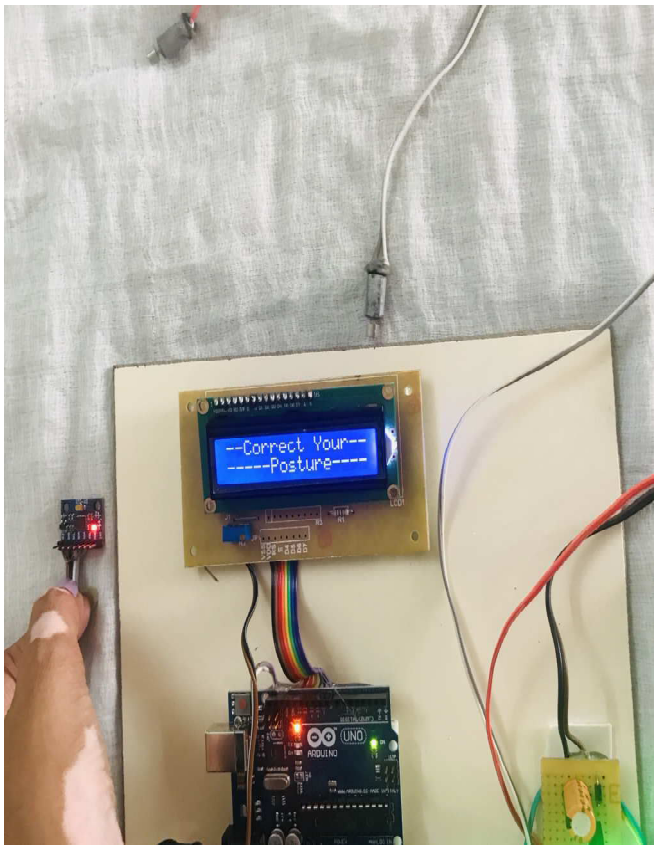


Fig. 4.Accelerometer Sensor

E.CRUTCH DESIGN:

The proposed crutch design is made of iron since its light weight and durable.This light weight crutch helps to embed all the sensors effectively while walking and analyze the patterns[12].Various surveys has been taken from different age group population to function the E-crutch effectively[13]. Analysis have been done for various age groups and the accuracy of their pain relief has been improved after using the smart crutch.GSM module helps to intimate the status of the elderly people through a message stating the measured parameters.It is extensively used for Alzheimer;s disease patients.Fig .5. illustrates the graphical plot of various data plotting the applied force in terms of Pascal for 15 subjects on the smart crutch.The GPS module helps to track the location of the elderly individuals.

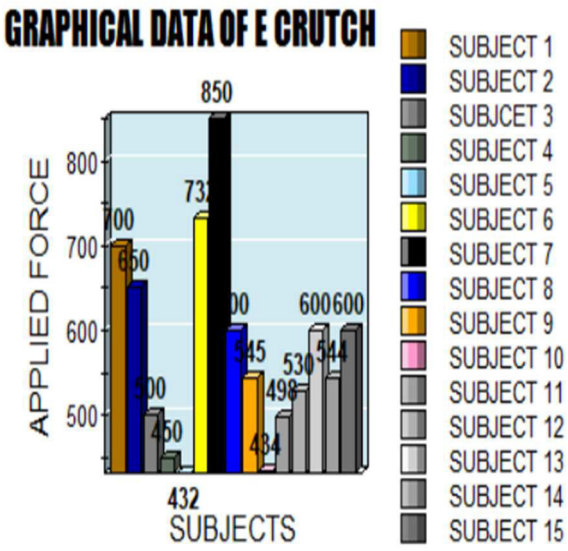


Fig. 5. Graphical representation of applied force for various subjects

ISBN 979-8-3503-5094-4

V.RESULTS AND DISCUSSION

This tabular column (TABLE 1) illustrates various subjects featuring force and acceleration values.Considering a 4 point gait aid in crutch is always advisable for elders undergoing rehabilitation treatments. Four point gait system is advisable for elderly patients since it is the best option for weight bearing and balancing.

All the sensors measuring force,acceleration,light condition and temperature sensor sensing the vein of the varicose affected individuals,Figure 6 illustrates the overall frame of E-crutch made up of iron and hence it helps the patients to evaluate the present condition of the health.This is extensively used in the case of Alzheimer's disease patients.

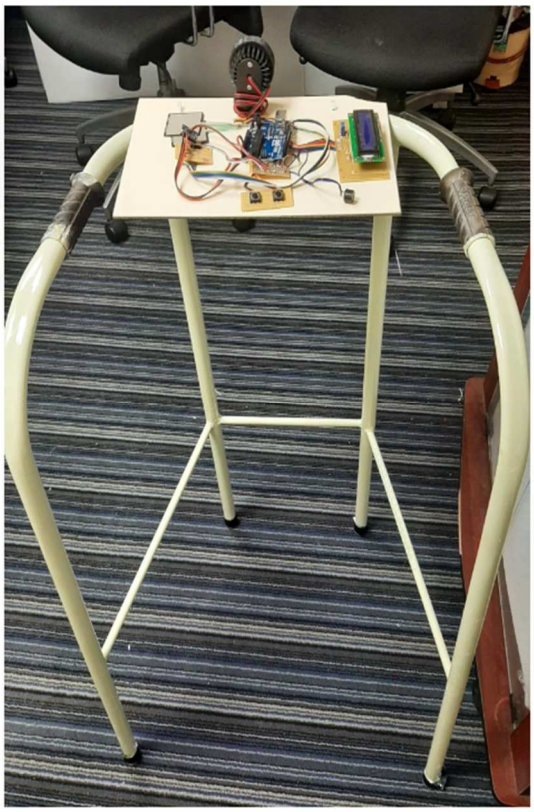


Fig.6 .Design of E-crutch

TABLE I: DATA COLLECTION REPRESENTING APPLIED FORCE AND ACCELERATION

SUBJECT	AGE	FORCE (IN PASCAL)	ACCELERATION (CORRECT YOUR POSTURE)
SUBJECT 1	51	432	BACKWARD
SUBJECT 2	52	600	BACKWARD
SUBJECT 3	63	600	BACKWARD
SUBJECT 4	64	544	FORWARD
SUBJECT 5	45	530	BACKWARD
SUBJECT 6	48	498	FORWARD
SUBJECT 7	52	434	BACKWARD
SUBJECT 8	55	545	FORWARD
SUBJECT 9	60	400	BACKWARD
SUBJECT 10	61	850	FORWARD
SUBJECT 11	62	700	FORWARD
SUBJECT 12	54	500	BACKWARD
SUBJECT 13	65	700	BACKWARD
SUBJECT 14	56	650	BACKWARD
SUBJECT 15	55	660	FORWARD

VI. CONCLUSION

We developed a high end E-Crutch which has been fabricated and embedded with multiple medical sensors to analyse the issues faced by the individuals undergoing physical sessions in rehabilitation.These developments ensure and helps to educate people regarding the usage of crutch for every medical conditions particularly.It employs a applied force sensor,gyroscope and a pain relief therapy for varicose vein which is portable and in low in cost.In future it can be extended for blind people to detect obstacles while using crutches and also a voice can be added to alert the user in a possible manner.

VII. ACKNOWLEDGMENT:

I deeply thank to all the mentors who played a vital role in the success of my work. I would like to thank the management of VISTAS, and the Department of Biomedical Engineering,for extending their support and motivation throught out my research.

ISBN 979-8-3503-5094-4

REFERENCES

[1] Piyali Sengupta , Kiran Mondal , Hiranmoy Mahata, Sujaya De, Prakash C Dhara , Evaluation of Existing Walking Sticks and Recommendations for modified waking stick ,October 2020

[2] Claire Perez, PT, BSc, MSc(c), and Joyce Fung, PT, PhD,An Instrumented Cane Devised for Gait Rehabilitation and Research ,January 2011

[3] Portegijs E Edgren J, Salpakoski A, et al. Balance confidence was associated with mobility and balance performance in older people with fall-related hip fracture: a cross-sectional study.

[4] Romteera Khlaikhayai ,Chavana Pavaganun , Benja Mangalabrukh and Preecha Yutapin , An Intelligent Walking Stick for Elderly and Blind Safety Protection , December 2011

[5] Laufer Y. The effect of walking aids on balance and weight-bearing patterns of patients with hemiparesis in various stance positions. Phys Ther 2003;83(2):112–122. DOI: 10.1093/ptj/83.2.112.

[6] Charles Bark§ CEO, Kabalan Chaccour*†, Member, IEEE, Rony Darazi*, Senior Member, IEEE, Amir Hajjam El Hassani†, Emmanuel Andres‡,MD , Design and Development of a Force-Sensing Shoe for Gait Analysis and Monitoring, 2017.

[7] Oates A, Perez C, McFadyen BJ, Fung J. Effects of haptic information on postural stability and coordination post-stroke while walking in a virtual environment. Paper presented at: The 19th Conference of the ISPGR (International Society for Posture and Gait Research); Bologna, Italy; 2009.

[8] A Comparative Study of Walking Assistance Tools Developed for the Visually Impaired People using Arduino UNO.Mahfuza Khanom, Muhammad Sheikh Sadi, and Md. Milon Islam.

[9] Force Sensor. Series MC2.5-500. Watertown, MA: Advanced Medical Technologies

[10] Three-dimensional skin deformation as force substitution: Wearable device design and performance during haptic exploration of virtual environments Samuel B. Schorr Student Member, IEEE, and Allison M. Okamura , IEEE

[11] Elnaffar.S, A. El Allam, “An app approach to correcting the posture of smartphone users,” In IEEE Advances in Science and Engineering Technology International Conferences, pp. 1-4, 2018.

[12] Epstein, S.: Art, history and the crutch. Clinical Orthopedics and Related Research R89, 4–9 (2020)

[13] Eurostat, E.: Population structure and ageing. European Commission: Luxembourg (2020)

ISBN 979-8-3503-5094-4

Graphene Quantum Dots based sensor for early detection of Stroke

Kirthana Sivakumar
Cognizant Technology Solutions
Chennai, India
kirthana.sivakumar2@cognizant.com

***Abstract*— Stroke results in many fatalities every year, especially even before the patient is diagnosed with stroke. However, recent years have seen an increase in new innovative research surrounding Biomedical Nanotechnology, especially in the field of Graphene Quantum Dots for various biosensing applications. They have various advantages due to their unusual physical and chemical properties allowing them to be the perfect candidate for early stroke detection. In this article, the biomarkers for stroke detection, the preparation methods of Graphene Quantum Dots, and their biosensing applications from literature are discussed.**

Keywords—graphene quantum dots, stroke, biomarkers, biosensing

I. Introduction

When the blood flow to the brain is cut off, a stroke occurs. Strokes are divided into two categories: Ischemic stroke (block in the blood vessel) and hemorrhagic stroke (bleeding of the blood vessel). Every year 15 million strokes occur worldwide as stated by the McLaren Health care news in 2023. According to World Stroke Organization's 2022 report, Stroke ranks as the second most common cause of death and the third most widespread cause of disability. The damage depends on the location of the part of the brain which is affected and the duration the brain suffers from the decreased blood flow. It is very important to detect stroke as soon as possible to prevent death and improve patient care. The typical diagnostic process for stroke involves using neuroimaging modalities like MRI and CT. However, these imaging modalities are expensive and time consuming. Using nanomaterials like quantum dots in biosensing applications for cerebrovascular diseases like ischemic stroke is preferred because of its many advantages like excellent biocompatibility, low cytotoxicity, fast electron mobility and high chemical stability. Graphene quantum dots (GQDs) based electrochemical sensors can be used as an alternative to traditional diagnosing techniques for the detection of the biomarkers for stroke which would yield faster results with lower cost.

II. Biomarker detection- Conventional Method

Biomarkers are an objective characteristic of the body that can be measured. They are important because they can improve the diagnosis of diseases at an early stage. Biomarkers are found in body fluids such as blood, plasma as well as tissues. Blood biomarkers range from proteins to lipids and various other metabolites. Some commonly measured biomarkers for ischemic stroke diagnosis include Gilial fibrillary acidic protein (GFAP), neuron specific enolase (NSE), myelin basic protein (MBP).

- GFAP- It is an intermediate filament protein. During stroke, there will be damage to the blood brain barrier made of astrocytes. This causes GFAP to diffuse into the bloodstream increasing the levels of GFAP [1].
- NSE- NSE is a neuronal glycolytic enzyme. There is increase in the level of NSE which causes severe disability [2]
- MBP- It is a hydrophilic protein. Increased levels of MBP observed in stroke patients [3].

III. What are Graphene Quantum Dots?

Graphene Nanoparticles which are less than 100 nm are graphene quantum dots (GQDs) They have exceptional physical and chemical properties allowing them to be used for various applications including biomedical applications [4]. Owing to their high biocompatibility, low cytotoxicity and high affinity to different functional groups they play a vital role in biosensing applications [5]. Photoluminescence is another important property of GQDs, and studies have shown that the size of the GDQs can be modified by changing the PL absorption and emission wavelengths, introducing dopants or by modifying the surface properties. This tunable PL property makes them an excellent candidate for biosensing and bioimaging applications [6]. Electrostatic, covalent and π–π are the three common ways in which the GQDs interact with other biomolecules.

IV. Preparation of Graphene Quantum Dots

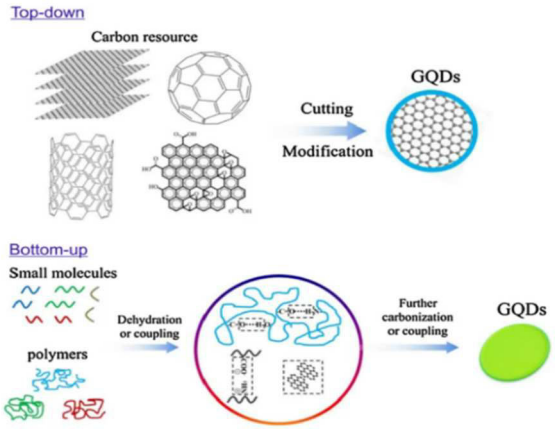


Fig. 1. Graphene Quantum Dots Synthesis [7]

ISBN 979-8-3503-5094-4

There are two main methods to synthesize GQDs: The top-down approach and the bottom-up approach.

1. Top-Down Approach

The top-down approach is less complex when compared to the bottom-up approach as it does not involve certain organic materials along with complex optimum conditions. For the top-down approach large blocks of carbon are broken down into small pieces using graphite and because carbon materials are the raw materials, they are easy to obtain and cheap. It includes different methods like physical, chemical, electrochemical, hydrothermal and ultrasonic [7]. One drawback of the top-down approach is that there is no precise control of the size and surface morphology of the GQDs [8].

2. Bottom-Up Approach

The bottom-up approach uses small aromatic molecules like glucose or glutamic acid in complex optimum conditions to form GQDs. It includes different methods like pyrolysis, decomposition, chemical vapor deposition. Though there is better control over size, shape and surface morphology, the production of GQDs is complex and not easy to obtain unlike the top-down approach [8].

V. Graphene Quantum Dots Based Electrochemical sensors in Lierature

A. For early detection of Prostate Cancer

In 2021 Tabish et al. [9] mentions two studies where GQDs based electrochemical sensors were used to detect the biomarkers for prostate cancer. PSA is a specific biomarker which when detected by sensors helps in the early detection of cancer. In the first study Yang et al. [10] the primary and secondary PSA antibodies were detected by GQDs functionalized graphene sheets which indicated a low detection limit and broad linear range of sensing. In another study by Wu et al. [11] two different types of GQDs i.e., animated GQDs and carboxyl GQDs were used to detect PSA levels in serum samples by an ultra-electroluminescent biosensor.

B. For early detection of Lung Cancer

Kalka et al. published a research study [12] in 2020 which mentions the ultrasensitive fluorescent biosensor based on GQDs and Gold Nanoparticles (AuNPs) for the detection of small cell lung cancer biomarkers. The energy donor and the energy acceptor of the biosensing system are GQDs and AuNPs respectively. The hydrothermal method of the Bottom-up synthesis was used to synthesize GQDs. The response time was only 16 minutes. Low detection limit and broader linear detection rate was also observed.

C. For early detection of Pancreatic Cancer

In the work by Ajgaonkar et al. in 2022 [13] Nitrogen doped Graphene Quantum dots (NGQDs) based fluorescent biosensing was used to detect the presence of a miRNA (pre-miR-132) which is a pancreatic cancer biomarker. The stem and loop portions of this biomarker is detected using the non-invasive biosensor. The NGQDs were prepared from a glucosamine precursor and were synthesized using a bottom-up approach.

D. For early detection of Breast Cancer

In a study published in 2021 by Pothipor et al. [14] a three-array electrode made of carbon comprising of a combination of different materials like AuNPs, Graphene Oxide (GO) and GQDs is used to detect the different miRNAs such as the miRNA-21, miRNA-155 and miRNA-210. This label free electrochemical sensor is mentioned to have a wide detection rate along with a low detection limit and is useful for the simultaneous detection of breast cancer.

Another study published in 2022 by Kumar et al. [15] mentions GQDs based electrochemical biosensor which was used to detect the biomarker CD44 which is a hyaluronan receptor expressed in stromal cells. This electrochemical biosensor also had a low detection limit with high sensitivity.

E. For early detection of Heart Attack

The article written by Tabish et al. in 2022 [16] states that GQDs when combined with electrochemical sensing has many advantages and is very useful at detecting the two important biomarkers for heart attack which are troponin and myoglobin. Some advantages of using GQDs: strong affinity between GQDs and the biomarker, low biomarker consumption and reusability of the electrode.

In a study published by Bhatnagar et al. in 2015 [17] describes a cardiac immunosensor based on FRET where the blood biomarker Troponin I (cTn1) is detected using functionalized GQDs in 10 minutes. Like the above-mentioned studies, low detection limit was observed here, and the biomarker was detected in just a couple of minutes.

VI. Challenges in Stroke detection using Graphene Quantum Dots and the Role of IT

Biocompatibility and cytotoxicity: It is very crucial that the GQDs are biocompatible with less cytotoxicity and though many studies have shown them to have those properties, there are still some studies which claim that GQDs have cytotoxicity [18]. Therefore, it is important to further study regarding the cytotoxicity of GQDs.

Sensitivity and Specificity: To distinguish stroke from other neurological disorders it is vital that the GQDs biosensors are highly sensitive and specific for accurate biomarker detection. Advanced data analysis and machine learning aids in accurate biomarker detection and identification thereby increasing the specificity and sensitivity of the GQDs based biosensors.

Sensitivity = True Positive / (True Positive + False Negative) (1)

Specificity = True Negative / (True Negative + False Positive) (2)

Overall Accuracy = True Positives + True Negatives / Total no. of cases (3)

ISBN 979-8-3503-5094-4

VII. Conclusion

In this past decade there has been an incredible growth in the field of GQDs. They have especially been used as biosensors for detecting biomarkers for various diseases which are mentioned above. Similarly, they have a tremendous potential to be used as biosensors for early stroke detection due to their various advantages such as low cytotoxicity, high biocompatibility and high chemical stability. It is very crucial to detect stroke using innovative diagnostic tools within a limited time frame as delay in diagnosing stroke can be fatal.

References

1. Amalia, Lisda. "Glial fibrillary acidic protein (GFAP): Neuroinflammation biomarker in acute ischemic stroke." *Journal of Inflammation Research* 14 (2021): 7501.
2. Bharosay, Anuradha, et al. "Role of brain biomarker in predicting clinical outcome in hypertensive cerebrovascular ischemic stroke." *Indian Journal of Clinical Biochemistry* 33.2 (2018): 178-183.
3. Brouns, Raf, et al. "Neurobiochemical markers of brain damage in cerebrospinal fluid of acute ischemic stroke patients." *Clinical Chemistry* 56.3 (2010): 451-458.
4. Yan, Yibo, et al. "Recent advances on graphene quantum dots: from chemistry and physics to applications." *Advanced materials* 31.21 (2019): 1808283.
5. Zheng, Xin Ting, et al. "Glowing graphene quantum dots and carbon dots: properties, syntheses, and biological applications." *small* 11.14 (2015): 1620-1636.
6. Chung, Seokhwan, Richard A. Revia, and Miqin Zhang. "Graphene quantum dots and their applications in bioimaging, biosensing, and therapy." *Advanced Materials* 33.22 (2021): 1904362.
7. Zhao, Changhong, et al. "Synthesis of graphene quantum dots and their applications in drug delivery." *Journal of Nanobiotechnology* 18 (2020): 1-32.
8. Facure, Murilo HM, et al. "Graphene quantum dots-based nanocomposites applied in electrochemical sensors: A recent survey." *Electrochem* 2.3 (2021): 490-519.
9. Tabish, Tanveer A., et al. "Graphene quantum dot–based electrochemical biosensing for early cancer detection." *Current Opinion in Electrochemistry* 30 (2021): 100786.
10. Yang, Minghui, Alireza Javadi, and Shaoqin Gong. "Sensitive electrochemical immunosensor for the detection of cancer biomarker using quantum dot functionalized graphene sheets as labels." *Sensors and Actuators B: Chemical* 155.1 (2011): 357-360.
11. Wu, Dan, et al. "Label-free electrochemiluminescent immunosensor for detection of prostate specific antigen based on aminated graphene quantum dots and carboxyl graphene quantum dots." *Scientific reports* 6.1 (2016): 20511.
12. Kalkal, Ashish, et al. "Biofunctionalized graphene quantum dots based fluorescent biosensor toward efficient detection of small cell lung cancer." *ACS Applied Bio Materials* 3.8 (2020): 4922-4932.
13. Ajgaonkar, Ryan, et al. "Detection of pancreatic cancer miRNA with biocompatible nitrogen-doped graphene quantum dots." *Materials* 15.16 (2022): 5760
14. Pothipor, Chammari, et al. "An electrochemical biosensor for simultaneous detection of breast cancer clinically related microRNAs based on a gold nanoparticles/graphene quantum dots/graphene oxide film." *Analyst* 146.12 (2021): 4000-4009.
15. Kumar, Neeraj, et al. "Electrochemically exfoliated graphene quantum dots based biosensor for CD44 breast cancer biomarker." *Biosensors* 12.11 (2022): 966.
16. Tabish, Tanveer A., et al. "Graphene quantum dots-based electrochemical biosensing platform for early detection of acute myocardial infarction." *Biosensors* 12.2 (2022): 77.
17. Bhatnagar, Deepika, et al. "Graphene quantum dots FRET based sensor for early detection of heart attack in human." *Biosensors and Bioelectronics* 79 (2016): 495-499.
18. Murugan, Kadarkarai, et al. "Nanofabrication of graphene quantum dots with high toxicity against malaria mosquitoes, Plasmodium falciparum and MCF-7 cancer cells: impact on predation of non-target tadpoles, odonate nymphs and mosquito fishes." *Journal of Cluster Science* 28 (2017): 393-411.

ISBN 979-8-3503-5094-4

Implementation of a Compensated Two-Stage Operational amplifier

C.Krithika
ECE department
Sri Venkateswara College of Engineering
Tirupati, A.P.
kayamlokesh78@gmail.com

K.Lokesh Krishna
ECE department
Sri Venkateswara College of Engineering
Tirupati, A.P.
lokeshkrishna.k@svcolleges.edu.in

D.Srinivasulu Reddy
ECE department
Sri Venkateswara College of Engineering
Tirupati, A.P.

B.Guru Lakshmi
ECE department
Sri Venkateswara College of Engineering
Tirupati, A.P.

K.Muni Bhaskar
ECE department
Sri Venkateswara College of Engineering
Tirupati, A.P.

K.G.Vijay Sai
ECE department
Sri Venkateswara College of Engineering
Tirupati, A.P.

***Abstract*— Enormous growth is being noticed in the usage of hand-held electronic gadgets during the past decade. One of the very vital circuits in the implementation of these wireless communication systems is the high speed complementary metal oxide semiconductor (CMOS) operational amplifiers. Improvements and modifications in integrated circuit fabrication technologies along with state of the art novel circuit design methodologies have made the high speed low power linear circuits to offer better performance in terms of operating speed, noise free, power dissipation, signal to noise distortion, voltage gain and oscillation free. The work in this paper reports a low power CMOS two-stage operational amplifier with proper designed miller compensation technique. The technique uses a nulling resistor method and achieves temperature insensitive tracking method. Circuit simulation is carried out at CMOS 90nm in analog design environment of cadence tools software. The circuit simulation demonstrates a gain of 78dB, unity gain bandwidth of 194MHz, phase margin of 72 degrees, slew rate of 62.6V/µS, a power consumption of 106.3 microwatts and all the results confirm an improved performance when compared to the previous works in the literature.**

Keywords*— Slew rate, negative feedback, Miller compensation, CMOS, IC operational amplifier, noise free, and Low power.

I. INTRODUCTION

Technological improvements during the past decade have revolutionized the utilization of various hand-held devices such as Laptops, Enterprise digital assistant, Graphing calculators, Handheld game console, Handheld PCs, Mobile phones, Tablets, Handheld game console, Digital cameras and camcorders, Smart watch, Smart glasses, Wearable computers, Head-mounted display, Smart cards, Pagers and Personal navigation devices. All the above mentioned hand held gadgets have enabled the humans to interconnect, transfer data (both audio and video), and entertain across the globe. Other applications also include accurate data acquisition systems in instrumentation, Bio-medical signal processing systems, telecommunication circuits, analog signal processing applications (i.e. such as video and large bandwidth operating circuits) and medical systems. In all these devices high performance operational amplifiers (Op-Amp) are crucial and vital for its effective functioning. These high speed operational amplifiers should provide high output current (i.e. slew rate), high gain levels, low distortion and noise, best DC operating point, possess wide gain-bandwidth product, fast settling times, and operate at low values of voltage levels (V_{DD}). Amplifiers are linear circuits and are the very simple building blocks in power optimized analog and mixed circuit topologies. They produce the output functions nearly similar to signal filtering, enhancement of signal values, DC signal amplification, buffering of signals, DC level shifting, inversion of signals, clipping and clamping of signals, performing mathematical operations such as addition, subtraction, data conversion, multiplication, division, integration and differentiation, gain adjustments, high-speed amplification, log, antilog operations and peak signal detection [1].

Operational amplifiers are global circuits as well as building blocks and are present in all the mixed signal electronic circuit design systems. As technological innovations are progressing at a faster rate towards higher operating speeds, smaller silicon area and operating at low power levels, the signal processing circuits demand innovative design methodologies in the implementation of operational amplifiers. The complementary metal oxide semiconductor (CMOS) process is constantly driving the integrated circuit industry because of the scaling benefits. The scaling methods offer improved device geometries, improved driving currents, improved interconnect technologies, improved active and stand-by currents, decreases silicon-chip area, reduces fabrication costs and improved overall-chip performance. The above benefits arising due to scaling techniques have motivated to implement operational amplifiers using CMOS technologies. The CMOS operational amplifiers are vital signal amplifications circuits in the implementation of signal conditioning circuits, analog signal processing circuits, complete integrated biochip applications, and integration of various sensor circuit applications [2]. Integrated circuits implemented using CMOS process under ninety nanometer technologies, encounters substantial limitations such as intrinsic gate switching delay, reduced noise margin and slightly increased power dissipations values.

ISBN 979-8-3503-5094-4

A metal oxide semiconductor (MOS) transistor operation can be described by using square law equation models. The following expressions are valid presuming that the charge in the depletion layer width is unvarying throughout the n-channel [3]. The equation for current flowing thought the drain terminal in the pinch-off region is written as

$$I_{Dn} = \frac{\beta_n}{2}(V_{GSn} - V_{tn})^2[1 + \lambda(V_{DSn} - V_{GSn} - V_{tn})] \quad (1)$$

In triode area, the drain current is given by

$$I_{Dn} = \frac{\beta}{2}[2(V_{GSn} - V_{tn})V_{DSn} - V_{DSn}^2] \quad (2)$$

The saturation voltage $V_{sat} = (V_{GSn} - V_{tn})$ .When $V_{DSn} = V_{sat}$, the channel width remains constant and the saturation region starts. Due to the channel length modulation, the parameter 'λ' is included in the saturation region of a MOS transistor. ' β ' is the device transconductance.

The threshold voltages are defined by (where V_{BSn} is the bulk-bias voltage)

$$V_{tn} = V_{t0n} - \gamma_n\left(\sqrt{2\,\phi_{F,n} + V_{BSn}} - \sqrt{2\phi_{F,n}}\right) \quad (3)$$

The transconductance of a MOS transistor functioning in the active region is given as the ratio of the change in the drain value to the corresponding change in the MOS gate to saturation voltages, when operated in the strong inversion region.

CMOS operational amplifier circuits exhibit high gain I the range of atleast 10000 and are direct coupled amplifiers offering special functions. They are operated with specially external connected resistors and capacitors to operate in closed loop feedback system. Since the open loop gain value of an operational amplifier is very high, henceforth they are always operated in closed loop environments for achieving desired performance. A two stage CMOS operational amplifier is the most preferred configuration for designing the general purpose operational amplifiers. A practical general purpose two stage CMOS operational amplifier exhibits infinite open loop voltage gain, output resistance as zero ohms, slew rate as infinite V/µS, infinite PSRR(power supply rejection ratio), bandwidth as infinite MHz, infinite CMRR (common mode rejection ratio), high gain bandwidth product, zero levels of input bias current, zero levels of input offset voltages, infinite output voltage swings and input resistance as infinite ohms.

Koh.J et al. describes a two-stage folded type class-AM op-amp along with a current regulator circuit [4]. It also employs a switched capacitor based amplifier circuit implemented in 0.022 micrometer CMOS process and the circuit operates at a clock speed of 104 KHz and occupies an circuit area of 0.0956 millimeter square. Various compensation circuits in the design of a two stage operational amplifier are presented in [5]. The specifications such as power dissipation, unity gain bandwidth and phase margin are analyzed using different circuits such as current buffer methods with common gate circuits, nulling resistor technique and voltage buffer methods. Tradeoffs are also analyzed in this work. A.J.Kumar et al. presented the simulation of op-amp using class-AB output stages [6]. The work utilizes cascode transistor connections for achieving larger gains and better slew rates. The first stages are made to work in the sub-threshold region to decrease the current dissipation of the op-amp. Authors in [7] reported the design and effects three amplifier configurations such as single stage, two-stage and three stages with respect to stability issues, gain, slew rate and settling times. The work also includes the slew rate restrictions with respect to zeros of the transfer function. Guo. Y et al. describes an precise metal oxide semiconductor model for the design of a 2-stage operational amplifier in 180nanometre process technology [8]. Comparisons were carried out to test the effectiveness of the proposed model. Authors in [9] proposed an offset voltage minimization method, wherein secondary amplifiers are used. The offset voltage is removed during the regular operation of the circuit and the error voltage is minimized to 0.00189V in 350 nanometer process technology. A high performance operational amplifier for use in bio-medical applications is described in [10]. The work reported better values with respect to op-amp specifications and unity gain bandwidth enhanced to 50MHz. An operational amplifier connected to operate as a high-speed comparator for use in a comparator array bank in Flash analog to digital converter is presented in [11]. The comparator circuit operates at the full sampling rate of the data converter circuit and dissipates a power of about 0.236milliwatts. Prokopenko.N et al. proposed a novel circuit stages to enhance the operating speed of the operational amplifier [12]. The circuit simulations were carried out in LT-Spice circuit simulation software. Authors in [13] proposed an operational amplifier designed to operate as a summing amplifier for use in a second order sigma delta modulator. The operational amplifier was designed to accept three input signals at a clock frequency of 500MHz. Authors in [14] presented a high performance operational amplifier in 0.09micrometer process using Tanner Spice simulation software. The novel circuit uses a current buffer circuit in the first stages of the circuit and less power dissipation is observed in the simulation results. The design work considered in this paper uses a novel compensation technique to produce a stable system output with low power dissipation.

The proposed work is written as follows. The particulars associated with 2-stage operational amplifier implementation using CMOS technology including design expressions is written in Section-II, while the simulation results and its analysis are described in Section-III. To conclude, the final of the proposed 2-stage CMOS operational amplifier are written in Section-IV.

II. Operational Amplifier

The input stage of a CMOS operational amplifier must be capable of sensing and amplifying the difference between the two signals, while simultaneously rejecting the common-mode signal input voltages. This stage should also operate at low frequencies with sufficient gain at high settling times. A practical complementary metal oxide semiconductor operational amplifier offers high values of input resistance (megaohms to gigaohms), low values of output resistance (in the range of ohms), higher unity gain bandwidth vales (greater than 1 MHz), complete rail to rail voltage swing levels, reduced power dissipation levels, higher slew rate levels (atleast greater than 50V/µS), lower values of power supply rejection ratio, and increased levels of common mode rejection ratio. Fig.1 displays the inner view of a typical 2-

ISBN 979-8-3503-5094-4

stage operational amplifier. It comprises of two gain stages and at last a buffer stage is connected at the output side. It also includes a compensating capacitor (C_{cm}).

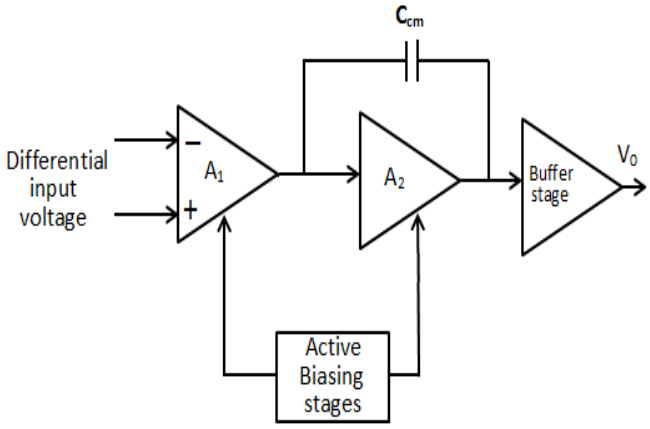


Fig.1: Inner details of a compensated 2-stage operational amplifier

Since the A_{OL} of a two stage operational amplifier (i.e. uncompensated amplifier) is very high and it becomes unsuitable for many applications. An open loop two stage operational amplifier operation is demonstrated by a 2-pole transfer function. This transfer function worsens the values of gain margin, phase margin and settling times of an operational amplifier. In general a transfer function containing excess than two poles poses stability issues. Hence compensation is applied to achieve good stable values of phase margin, gain margin and settling time. To stabilize this gain, a negative feedback is incorporated to achieve better performance characteristics. Large values of voltage gain are achieved are realized by using active loads at the output stages. Further other important specifications include such as offset voltages, input referred noise voltages, common-mode voltage gain and differential mode voltage gain. To reduce the values of input referred noise and offset voltages, the difference between the gate to source voltages and threshold voltages of the n-channel MOS transistor needs to be increased.

For the two stage operational amplifier, the expressions for zeros and poles are

$$z_1 = \frac{G_{m3}}{C_{cm}} \quad (4)$$

$$p_1 = -\frac{1}{G_{m3} C_{mc} R_2 R_1} \quad (5)$$

$$p_2 = -\left(\frac{C_{cm} \times G_{m3}}{C_{cm}C_2 + C_1 C_{cm} + C_2 C_1}\right) \quad (6)$$

The output voltage swing is given by

$$V_{Out,swing} = |V_{SS}| + V_{DD} - \left[|V_{out,7}| + |V_{out,8}|\right] \quad (7)$$

The unity gain bandwidth of the operational amplifier is

$$f_{UGB} \cong \frac{G_{m2}}{2\pi C_{cm}} \quad (8)$$

The open-loop gain of the operational amplifier is

$$A_{ol}(s) = \frac{\left(1 - \frac{C_{cm}s}{G_{m3}}\right)}{\left[1 + sR_2(C_{cm}+C_2)\right]\left[1 + sR_3(C_{cm}+C_o)+ \left[G_{m1}G_{m2}\, C_{cm}R_2R_3\left(1 - \frac{sC_{cm}}{G_{m3}}\right)\right]\right.} \quad (9)$$

$$\angle \frac{V_o}{V_{in}} = -arctan\left(\frac{2\pi f_0}{z_1}\right) - arctan\left(\frac{2\pi f_0}{p_1}\right) - arctan\left(\frac{2\pi f_0}{p_2}\right) \quad (10)$$

$$Slew\ rate = \frac{I_5}{C_{cm}} \quad (11)$$

The total output power consumption is calculated by using

$$P_0 = (I_{diss} \times V_{dd}) \quad (12)$$

The specifications considered in this design are as follows: gain should be atleast greater than 60 decibels, slew rate should be greater than 70volts per microseconds, operating voltage is 1.8V, CMRR atleast greater than 60 decibels, phase margin greater than 65 degrees, output swing greater than or equal to ±1.8V, power dissipation less than 0.8mW, unity gain bandwidth greater than 200MHz, operating temperature range between zero to 100 degrees, power supply rejection ratio atleast greater than 60 decibels, offset voltage less than or equal to 2mV and designed in 90nanometer CMOS process.

III. Simulation Results

The two stage CMOS operational amplifier is designed and simulated in analog design environment using cadence tools at 90 nanometer CMOS technology is shown in Fig.2. The metal oxide transistors are operated in constant current region or active region for getting optimum performance. The transient analysis drawn between voltages with respect to time determines on how a circuit's output changes in response to the various applied input signals such as sine, square and pulse. Firstly input amplitude of one millivolt is applied at the dual input stage and the corresponding output is taken at the drain side. Fig. 3, shows the time response waveform of output voltage when an input sine wave form is applied.

The slew rate is calculated from the transient step response as presented in Fig.4. A pulse width of 12 microseconds is connected at the input stage of operational amplifier and the corresponding output waveform is observed. From the response, the slew rate is calculated as 62V/μS. The frequency response analysis is carried out by applying the dual inputs with input signals of 10mV at both inverting and non-inverting input terminals. The response is plotted in the range from 1Hz to 1000MHz. From the simulation the gain is calculated to be nearly dB. At the high frequencies, the gain reduces gradually. Fig.5, shows the magnitude and phase simulated graphs. One of the important that determines the stability of a negative feedback amplifier system is the phase margin. The minimum value of a properly designed negative feedback amplifier system should be atleast 50 degrees.

ISBN 979-8-3503-5094-4

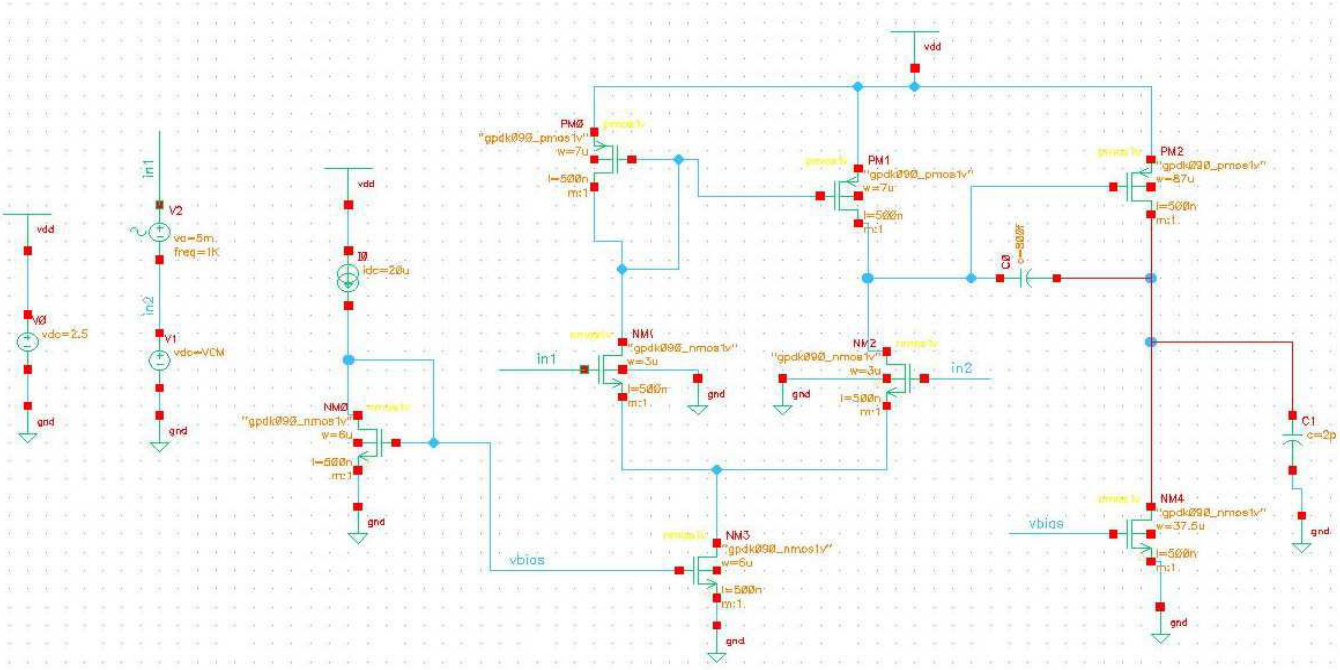


Fig.2: Simulated schematic of op-amp

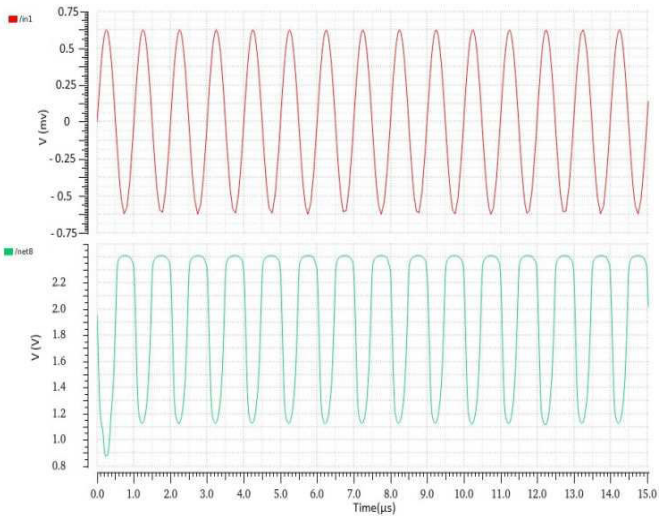


Fig.3: Transient analysis response analysis

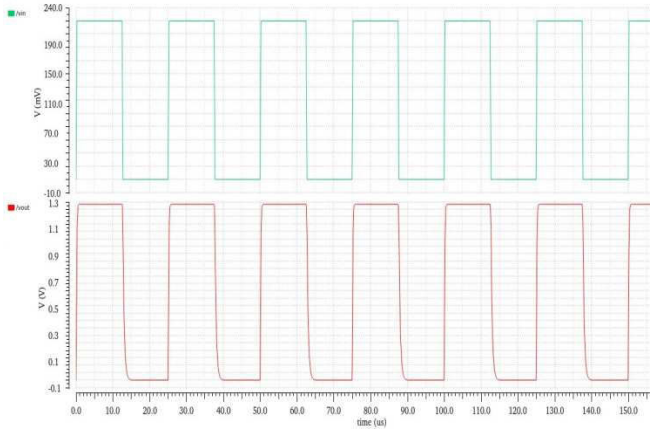


Fig.4: Slew rate calculations

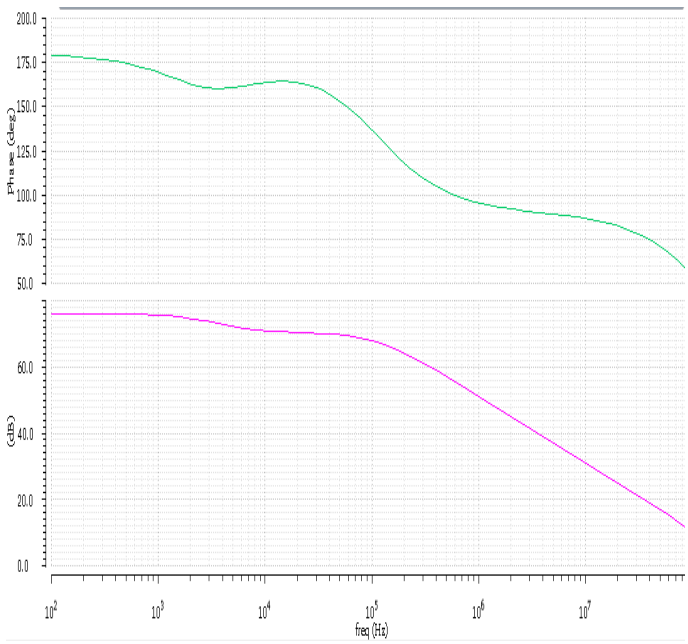


Fig.5: Magnitude and Phase plots

Large value of phase margin means it decreases the operating speed of the circuit, whereas lower values of phase margin instability issues and unnecessary oscillations in the output voltages. From the gain and phase plots the unity gain bandwidth is found to be 194MHz while phase margin is calculated as 72 degrees. To examine the noise performance of the circuit, the equivalent input noise simulation analysis is carried out. Fig.6 shows the simulated input noise versus frequency response curve. The Noise is observed in the low frequency band, while it reduces for higher frequencies and remains constant.

ISBN 979-8-3503-5094-4

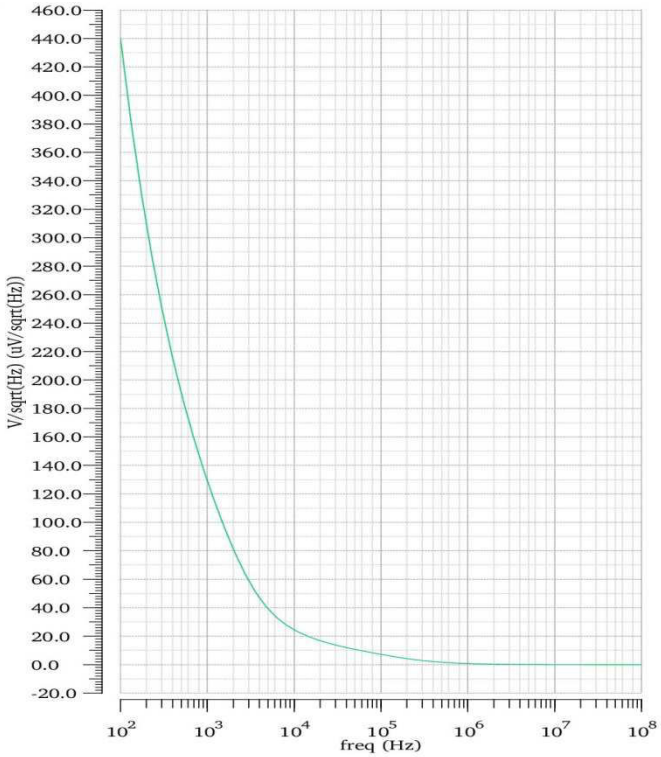


Fig.6: Noise response plot

The power dissipation of the 2-stage operational amplifier is calculated by measuring the steady current flowing through the power pins. Fig.7 shows the power dissipation analysis response.

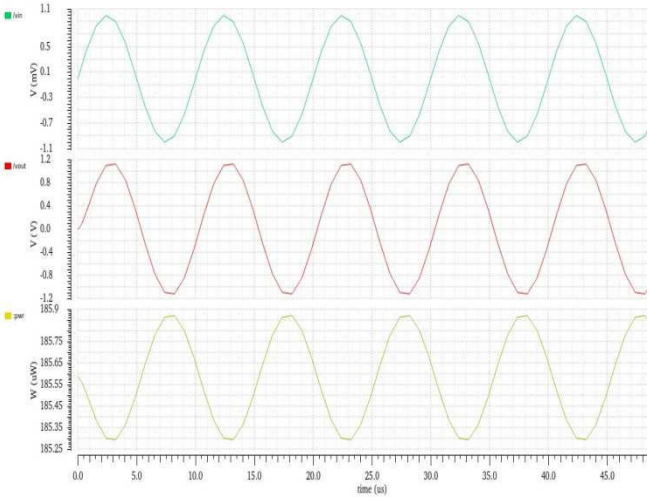


Fig. 7: Power dissipation analysis graph

TABLE I. Comparison of two the proposed work with previous reported work

Performance parameters	Reference [6]	Proposed work
Gain (dB)	76	78
Phase margin (deg.)	68	72
Power dissipation (µW)	123.8	106.3
Settling time (ns)	0.88	0.56
Unity gain bandwidth (MHz)	110	194
Output voltage swing (V)	±1.8	±1.8
Slew rate (V/µS)	10.8	62.6

The comparison results of the proposed operational amplifier with the previous works reported in the literature is specified in TABLE I.

IV. Conclusion

Continuous scaling in CMOS process have led to tremendous improvements in the operating speed, noise free, power dissipation, voltage gain and operating bandwidths. Larger voltage gains are achieved by employing two-stage or three stage operational amplifiers. In this paper, a high performance Two-stage complementary metal oxide semiconductor amplifier using modified Miller compensation method is designed and simulated in 90nanometer analog design environment of cadence tools software. This circuit configuration offers a wide unity gain bandwidth, better slew rate, increased voltage gain value, decreased power dissipation value and better stable operation with a phase margin of 70 degrees. The compensation capacitor is varied to achieve desired values of unity gain bandwidth and phase margin. The entire MOS transistors in the proposed circuit are designed to operate in the saturation region. All the obtained specifications have been compared with the previous works and the obtained simulation results shows better performance in terms of gain, phase margin, slew rate, power dissipation and settling times. Therefore applications where low power and stability issues are of particular concern such as in analog to digital converter architectures, the proposed two stage operational amplifier can be used.

References

[1] Carusone T C et al. “Analog Integ. Ckt. Des.”, Second edition, J. Wil. & So., Inc., 2017.

[2] P.E.Allen et al., “CMOS Anal. Ckt. Des.”, Oxf. Uni. Press, Third edition 2018.

[3] Razavi. B., "Des. of Analog CMOS Integ. Ckts.", Second edition, MG Hill-Sci./ Engg., 2017.

[4] Koh.J. et al., "A Full Diff. Swit. Cap. Amp. with a 2-Stage Folded-Mesh Class-AB Op-amp in a 0.022µm FD-SOI CMOS Process," Nineteenth Intern. SoC Design Conf., Korea, pp. (15-16), 2022.

[5] Z. Qi, "Comp. betw. Diff. comp. techn. for CMOS 2-stage amplifier," IEEE Intern. Conf. on Sens. Electr. & Comp. Engg, China, 2023, pp. (1684-1689).

[6] Kumar A.J. et al., "A hig. gai. low pow. op-amp. using Class-AB out. Stag." 3rd Intern. Confe. on comp. method. & comm., India, pp. (409-413) 2019.

[7] Giustolisi G., et al., "Design of 3-Stage OTA based on settl. time requirem. includ. Large and Small Signal Behavior," in IEEE Trans. on Ckts. & Systs.-1: Regu. Paps., 68(3), pp. (998-1011), March 2021.

[8] Y. Guo et al.,, "An accur. Des. Appr. for 2-stage CMOS op-amps," IEEE Asia Pac. Conf. on Ckts. & Sys., South Korea, pp. (563-566) 2016.

[9] Lu.C.W. et al., "An Off. voltage Canc. Techn. for 2-Stage CMOS Op-amps," IEEE Intern. Confe. on Integrated Ckt. Des. & Techn., Aus., USA, pp. (1-3), 2007.

[10] N. Uddin et al., "Des. & anal. of rob. two stage Op-Amp using 0.09micrometre CMOS Techn. for Bio-medical Appsl." Intern. Confe. on Rec, Progr. in Sci., Engineering and Tech., Bangladesh, pp. (1-5), 2022.

[11] Kumar K.S. et al., "A Hig. Spee. Flash ADC," 2nd Internl. Confe. on Soc., Mobi,, Analy. & Cloud) ,Tamil Nadu, India, pp. (283-288), 2018.

[12] Prokopenko N et al., "High spe. Op-amps based on the Adv. Deriv. of the Class-AB Interm. Stage," Fifth Intern. You. Conf. on Rad. Electro., Elec. and Power Engg. Russian Fed., pp. (1-5), 2023.

ISBN 979-8-3503-5094-4

[13] Krishna K.L.et al., "A one volt 2nd order ∑-Δ modulator ADC in 0.130micrometer CMOS", Internl. Confe. on Inf. Comm. and Embe. Systs., India, pp. (1-5),2014.

[14] Parthipan A.J. et al., "A High Perf. CMOS op-amp", 3rd Intern. Conf. on Comp. Metho.. & Comm., Tamil Nadu, India, pp. (702-706), 2019.

ISBN 979-8-3503-5094-4

Investigation of IOT Enabled Maternal Health Monitoring System for Antenatal Care

S. Kanimozhi
Department of ECE
CEG Campus, Anna University
Chennai-600025, India
kanimozhiss255@gmail.com

C. Kaushik Viknesh
Department of ECE
CEG Campus, Anna University
Chennai-600025,
India ecekv12@gmail.com

***Abstract*— In developing nations, a significant portion of the population resides in rural regions where healthcare systems lack integration for information sharing. Mainly, expectant mothers struggle to undertake regular prenatal checkups, leading to elevated infant and maternal mortality rates in both rural and urban settings. This scenario presents substantial healthcare challenges for women. To address this, an accelerometer sensor is devised to gauge the strength and frequency of fetal movements. This data is then transmitted to a RASPBERRY PI Pico controller. The motion of the fetus and vital metrics like blood pressure, heart rate, fetal kicks count, and maternal temperature are monitored through an array of diverse sensor technologies. The gathered data is communicated via IOT and showcased on mobile devices. This setup, with its heightened sensitivity and lightweight design capable of detecting even subtle motions, is well-suited for domestic monitoring. Contrary to the prevailing trend of employing costly and extended-use ultrasound scans, which exhibit unclear fetal impact limitations, this approach discourages their continuous utilization. Moreover, due to its high cost, we have put forth an alternative system incorporating a range of sensors including a heartbeat sensor, temperature sensor, blood pressure sensor, and an accelerometer sensor to monitor fetal movements. These sensors enable data collection, which is then transmitted to a mobile application via Raspberry pi Pico and IOT technology. Should any irregular readings arise, GSM module is employed to compute normal and abnormal rates.**

Keywords— Fetal monitoring, Prenatal care, raspberry pi, Fetal heart beat sound, Blood pressure, Fetal position, Global system for mobile communication.

I. INTRODUCTION

Antepartum monitoring of fetal heart rate is an important clinical procedure during gestation. The top end of prenatal fetal weal testing is to identify fetuses at threat or death, so that these adverse issues can be averted. In recent decades, numerous ways for assessment of fetal well- being have been introduced into clinical practice. Despite wide use of these ways, there's limited substantiation to guide their optimal use or to demonstrate their effectiveness at perfecting perinatal issues. In advanced nations, current perinatal mortality rates are roughly10/1000 births and fetal deaths regard for roughly 50 of deaths between 20 weeks of gestation and 1 time of age, with natural deformations and perinatal hypoxia being the top causes. Indeed though fetal surveillance which is to be performed more constantly on high- threat pregnant groups may significantly reduce the prevalence of fetal deaths, perinatal morbidity and motherly torture in similar groups, the maturity of bearings and deformations now do in low- threat gravidity with no linked threat factor.

This apparent anomaly emphasizes the critical need to develop further effective ways of relating at- threat fetuses in low- threat groups. In high- threat gravidity, ultrasound grounded technologies are the most common individual procedure for relating fetal concession, while by low- threat groups reduced fetal exertion is the only assessment shown to identify fetuses at threat, with poor positive prophetic value. Auscultation is one of the oldest medical tool in history which has also been applied in fetal diagnostics using especially- formed rustic stethoscopes. To a certain point in gestation, don't have to stay for coming antenatal visit at OB-GYN's to hear your baby's twinkle. It's possible to hear the twinkle at home using a stethoscope. Unfortunately, it can't hear as early as you can with an ultrasound or fetal Doppler. But with a stethoscope, a baby's twinkle is frequently 2 sensible between the 18th and 20th week.(as shown in figure1) Stethoscopes are designed to amplify small sounds. It has a casket piece that connects to a tube. The casket piece captures the sound, and also the sound trip. The size of the tube also matters. Generally the larger the tube, the briskly the sound can travel.

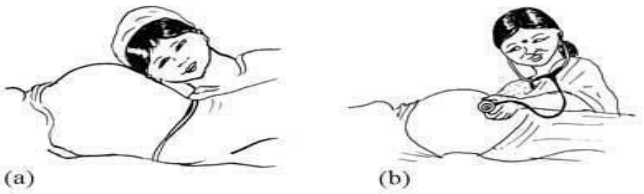


Figure 1. Fetal heart sound in antenatal care using stethoscope

There are step- by- step tips on using a stethoscope to hear baby's twinkle:

- Find a quiet position which is easier were it'll be to hear the baby's twinkle, sit in a room alone with the TV and radio out.
- Taradiddle down on a soft face and hear to the baby's twinkle in bed or lying on the settee.
- Feel around your stomach and find the baby's back. Baby's reverse is an ideal place to hear a fetal twinkle. This section of the stomach should feel hard, yet smooth.
- Place the casket piece on this area of the stomach. Now it can begin to hear the heart sound.

Still, sluggishly move the stethoscope up or down until it is suitable to pick up a sound, If in some case it might not incontinently hear. Fetal jiffs can sound like a watch ticking underneath a pillow.

ISBN 979-8-3503-5094-4

Still, it's one of the system for hearing a twinkle at home, but it's not always effective, if it can't hear the baby's twinkle using a stethoscope. The baby's position can make it delicate to hear, or it might not be far enough along in the gestation to descry a twinkle with a stethoscope. Placenta placement can also make a difference, if you have an anterior placenta, the sound looking for may be harder to find. However, the baby is presumably in the vertex position, if the fetal twinkle is loudest below the mama's bellybutton. However, the baby may be pants if it's loudest above her bellybutton (as shown in figure 2). Occasionally when the baby is facing the mama's front, the twinkle is harder to find because the baby's arms and legs get in the way. hear near the mama 's sides, or directly in the middle of her tummy, to hear the fetal twinkle.

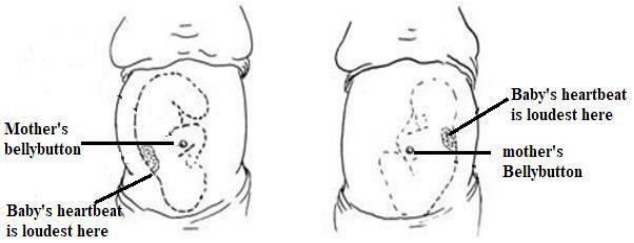


Figure 2. Fetal heart sound according to mothers bellybutton

The ultramodern form of auscultation provides thenon-invasive electronic recording and computer backed analysis of the aural cardiac signals. Because fetal heart exertion produces much lower aural energy and in addition it's girdled by a largely noisy terrain, major aural noise produced by the fetal movements, motherly digestive exertion, motherly heart exertion, movement of the detector head during recording, external noise forming from the terrain, the discovery of fetal heart sounds from abdominal recordings remains a grueling issue. The major accession problems linked are attenuation of sounds in the transmission path similar as amniotic fluid, muscular wall of uterus, layers of fat towel, bone or cartilage, position of detector placed on the tummy in relation to fetal heart or fetal position. The major signal processing challenges linked are low signal to noise rate makes de-noising veritablydelicate, lapping frequence bands of fetal heart sounds and lack of beat by beat confirmation of fetal heart sounds. In order to overcome being challenges in rooting fetal heart sounds reliably and stably using multiple sound transducers applied in a simple and harmonious pattern across the motherly tummy, taking little driver skill. therefore the beat by beat fetal heart signal acquired by this device with contemporaneous measures using a validated fetal electrocardiogram(fECG) device, in healthy gravidity during the last trimester but before the onset of labour. The fetal heart beat is directly measured using the stethoscope were the mic detector is connected nearly to the stethoscope to descry the heart beat sound of the fetal. It captures the nanosecond signal which is hitching with the mic detector and it amplifies into an Analog to a electrical signal. Mic detector is used for the conversion of sound swells into an electrical signal. Where the sound surge strike a microphone when they beget a diaphragm to vibrator. Temperature Detector is to measure the temperature and an 8-bit microcontroller to affair the values and moisture as periodical data.

II. MATERIALS AND METHODS

In the hardware setup different types of sensors have been used to measure the vital parameters such as temperature, heart rate, blood pressure for the maternal and the movement of the fetus. Sensors are attached in the system thus it helps to take reading and it is displayed. IoT is increasingly allowing to integrate devices capable of connecting to the Internet and provide information on the state of health of patients and provide information in real time to doctors who assists it. The following sensors are used they are

- Raspberry pi pico
- Temperature sensor
- Mic sensor
- Blood pressure sensor

In this innovative tackle setup, a different range of detectors is strategically stationed to capture pivotal vital parameters, encompassing temperature, heart rate, blood pressure for the motherly, and the movements of the fetus. The system's core revolves around the integration of a Raspberry Pi Pico, serving as the central processing unit that harmonizes the information gathered by detectors as shown in the figure3

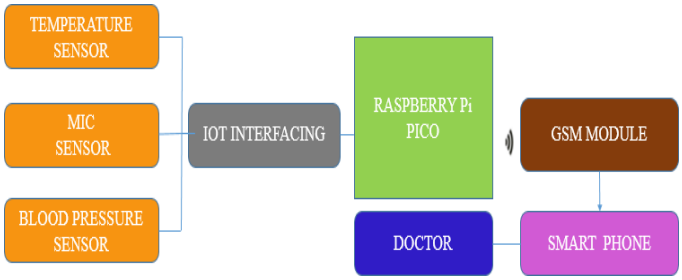


Figure 3. Block diagram

The trifecta of a temperature detector, microphone detector, and blood pressure detector collaborates seamlessly with the Raspberry Pi Pico, creating a comprehensive monitoring system. The connection of these detectors forms a cohesive network, furnishing real- time data on the health status of the motherly and fetal subjects. This admixture of detectors allows for a holistic understanding of the physiological dynamics during gestation. To enhance the system's functionality, a GSM module is incorporated. This module acts as a communication conduit, enabling the transmission of collected vital data to an Android operation. The Android app, acclimatized for healthcare professionals, receives and processes the data in real- time. In critical situations, the system triggers an alert communication, instantly notifying the attending croaker. This perpetration showcases the transformative eventuality of the Internet of effects (IOT) in healthcare. By seamlessly connecting bias to the internet, IoT facilitates the nonstop monitoring of cases' health parameters. The real- time transmission of health information to healthcare providers empowers croakers to make informed opinions and interventions instantly. In substance, this tackle setup not only exemplifies the community between detectors, calculating bias, and communication modules but also underscores the vital part of IoT in steering in a new period of visionary and connected healthcare.

ISBN 979-8-3503-5094-4

A. *RASPBERRY Pi PICO*

The Raspberry Pi Pico features a protean micro USB harborage, serving multiple functions within a compact design. This harborage serves as an input for power, easing the board's operation also, it supports data transfer, allowing flawless communication between the Pico and other bias also, the micro USB harborage serves as the interface for programming the microcontroller, streamlining the process of uploading and streamlining law. This multifunctional harborage enhances the Pico's stoner-friendly design, simplifying both power force and data connectivity for a variety of operations and programming tasks.as shown in the figure 4

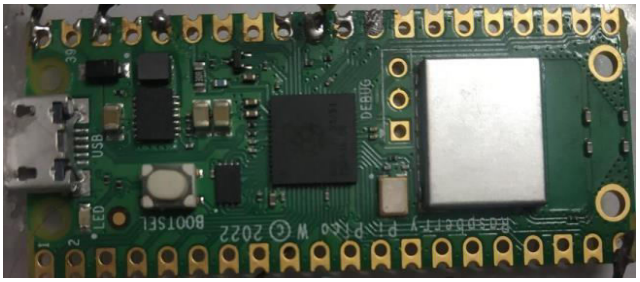


Figure 4 Raspberry pi pico

Memory: The Raspberry Pi Pico is equipped with 264KB of on- chip SRAM, furnishing effective static arbitrary- access memory for rapid-fire data access specially, it diverges from certain microcontrollers by lacking erected- in flash memory for program storehouse rather, the Pico adopts an external flash memory approach for program storehouse. This design choice allows for inflexibility and scalability in terms of program storehouse capacity, as druggies can elect and integrate external flash memory modules grounded on their specific requirements. While some microcontrollers integrate flash memory on- chip, the Pico's external flash memory strategy ensures rigidity, enabling druggies to knitter storehouse results according to the demands of their operations, making it a protean choice for different systems taking programmable memory and storehouse capabilities.

GPIO Pins: The Raspberry Pi Pico stands out with its provision of 26 multifunctional GPIO (General Purpose Input/ Affair) legs, offering a protean interface for colorful operations. These GPIO legs serve as flexible connectors that can be employed for different purposes, similar as digital input/ affair, analog input, PWM (palpitation range Modulation), I2C (Inter-Integrated Circuit), SPI (periodical supplemental Interface), and UART(Universal Asynchronous Receiver- Transmitter) communication. This expansive range of capabilities empowers inventors and potterers to affiliate with a wide array of electronic factors and bias, enabling the Pico to be adaptable across different systems.

Connectivity: The Raspberry Pi Pico features a protean micro USB harborage, serving multiple functions within a compact design. This harborage serves as an input for power, easing the board's operation also, it supports data transfer, allowing flawless communication between the Pico and other bias also, the micro USB harborage serves as the interface for programming the microcontroller, streamlining the process of uploading and streamlining law. This multifunctional harborage enhances the Pico's stoner-friendly design, simplifying both power force and data connectivity for a variety of operations and programming tasks.

Powersupply: The Raspberry Pi Pico offers flexible power options, accommodating both USB and external power sources. Its USB harborage serves as a power input, while external power sources enhance rigidity. The Pico's broad voltage range support enhances versatility, making it suitable for a different range of systems. This point allows druggies to power the Pico using different sources, furnishing convenience and comity with colorful setups and operations.

Programming languages: The Raspberry Pi Pico supports protean programming languages, including MicroPython, C, and C. MicroPython, a feather light subset of Python 3, is a popular choice for programming the Pico. Its relinquishment is driven by its simplicity and stoner-friendly syntax, making it accessible for both newcomers and educated inventors. This inflexibility enables druggies to choose the programming language that stylish suits their preferences and design conditions, contributing to the Pico's appeal as a protean and inclusive microcontroller platform.

Micro python: Micro python has the capability to run Python, allowing stoners to produce simple and easy- to- understand programs. Micro python supports multitudinous standard Python libraries, supporting further than 80 of the features of Python's most used libraries. Micro Python was designed specifically to support the typical performance gap between microcontrollers and Python. Python law is suitable to directly pierce and interact with attack, with increased attack possibilities that are not available using a normal Python operation that is run on an operating system. Micro Python's utilisation of attack abstraction sub estate technology allows advanced law to be portable among different microcontrollers within the same family or platform and on bias that support and can download Micro Python. Programs are constantly developed and tested on high- performance microcontrollers and distributed with the final operation used on lower- performance microcontrollers. Micro Python offers functionality, formerly new law has been written, to produce a frozen module and use it as a library which can be a part of developed firmware. This point assists with avoiding repetitive downloading of the same, formerly error-free, tested law into a Micro Python terrain. This type of module will be saved to a microcontroller's modules directory for collecting and uploading to the microcontroller where the library will be available using Python's import command to be used constantly.

B. *SUPPORTING HARDWARE:*

As Micro Python's perpetration and fashionability continues to grow, more boards have the capability to run Micro Python. numerous inventors are erecting processor specific performances that can be downloaded onto different microcontrollers. Installing Micro Python on microcontrollers is well proved and stoner-friendly. Micro Python allows relations between microcontroller tackle and operations to be simple, allowing access to a range of functionality while working in a resource constrained terrain, with a strong position of responsiveness shown in the figure 5.

The two types of boards used to run Micro Python:

- Micro Python loaded when manufactured, meaning only Micro Python can be run.
- Boards that have firmware that allows Micro Python to be installed to the board.

ISBN 979-8-3503-5094-4

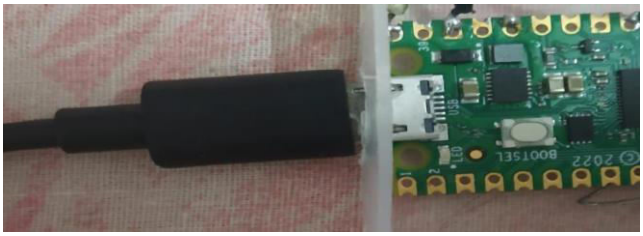


Figure 5. Micro python installation process

C. *SOUND SENSOR:*

A sound sensor, generally appertained to as a sound detector or sound sensor module, is an electronic outfit finagled to descry sound swells in its vicinity and transfigure them into electrical signals. These protean sensors serve different purposes, chancing operation in colorful operations similar as noise monitoring, security systems, robotization, and interactive technologies. In noise monitoring, they enable the assessment of ambient sound situations, abetting in environmental analysis. Security systems use sound sensors to identify specific audile patterns for enhanced surveillance. In robotization, these bias contribute to responsive systems, conforming to audile inputs. Interactive systems influence sound sensors to grease stoner-friendly interfaces, enabling bias to respond to voice commands or other audile triggers. The rigidity of sound sensors across a diapason of operations highlights their significance in ultramodern electronic systems, where the conversion of sound into electrical signals enhances functionality and responsiveness. as shown in the figure 6.

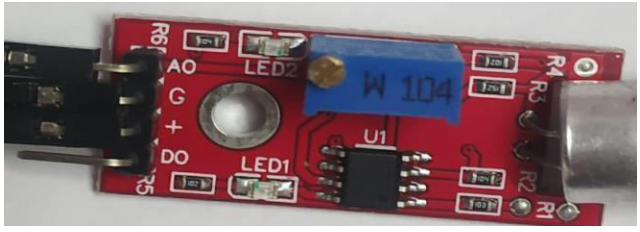


Figure 6. Sound Sensor

Sound sensor elements: A sound detector's abecedarian element is a microphone or a technical sound detector element. The microphone is vital in landing oscillations in air pressure convinced by sound swells. As sound swells interact with the microphone, it translates these pressure changes into electrical signals, forming the base for the detector's capability to descry and measure sound. This conversion process enables the detector to respond to colorful sound frequentness and confines, making it a pivotal element in operations ranging from noise monitoring to state recognition systems.

Transducer: The microphone plays a crucial part in transducing analog sound signals into electrical signals. As sound swells impact the microphone's diaphragm, it undergoes climate, producing a commensurable electrical signal. This metamorphosis from aural to electrical signals is abecedarian in capturing and processing sound. The generated electrical signal carries the nuances of the original sound, allowing for farther modification, processing, and interpretation in colorful operations similar as audio recording, communication bias, and sound analysis systems.

Signal conditioning: The original electrical signal from the microphone might be weak and include uninvited noise. Signal conditioning circuits, like amplifiers and pollutants, are generally employed to ameliorate signal quality and exclude noise. Amplifiers boost the signal strength, making it more robust for processing, while pollutants widely remove unwanted frequentness, enhancing the dedication of the signal. This signal exertion process is pivotal in operations where accurate and dependable representation of the original sound is essential, icing that posterior stages of signal processing or analysis yield accurate results.

Adc conversion: In numerous operations, the analog signal is converted into a digital signal using an Analog- to- Digital Converter (ADC). This digital representation allows for easier processing and interfacing with digital systems, similar as microcontrollers. The sound detector generally provides an affair interface, which can be an analog voltage signal, a digital signal, or both, depending on the design. Digital labors frequently indicate the presence or absence of sound above a certain threshold. Some sound detectors allow druggies to acclimate the perceptivity or threshold position. This enables customization to descry sound situations above a specific intensity.

D. *TEMPERATURE SENSOR*:

The DHT11 is a detector designed for binary functionality, measuring both temperature and moisture in its surroundings. It delivers a digital signal affair, simplifying commerce with microcontrollers as no fresh analog- to- digital conversion is demanded. Exercising a single- line communication protocol, the DHT11 efficiently transmits data to the microcontroller, easing integration into systems with constrained leg vacuity. . While offering moderate delicacy, suitable for multitudinous operations, it may not meet the demands of largely precise surroundings. Despite its limitations, the DHT11's simplicity, affordability, and ease of use make it a popular choice for layman systems and operations where a balance between delicacy and cost-effectiveness is respectable. as shown in the figure 7

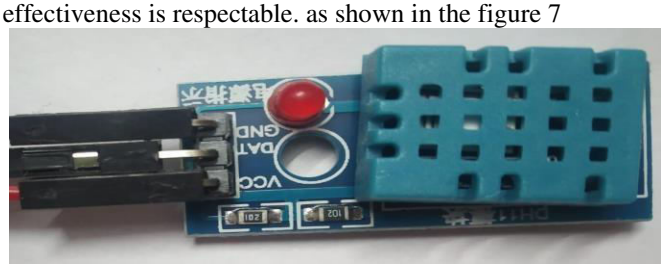


Figure 7. Temperature Sensor

E. *STETHOSCOPE:*

The stethoscope is an necessary medical device designed for auscultation, the process of harkening to internal sounds within the mortal body. This essential tool consists of a small slice-shaped resonator, as depicted in Figure8, which is placed against the case's skin. Connected to this resonator are one or two tubes, extending to earpieces that enable medical professionals to hear and dissect internal fleshly sounds. Primarily used for examining the heart, lungs, and bowel, as well as assessing blood inflow in highways and modes, stethoscopes are critical in diagnosing and covering colorful medical conditions. Aural stethoscopes operate grounded on the transmission of sound from the casket piece, through air- filled concave tubes, and eventually to the listener's cognizance. The casket piece, a vital element of the stethoscope, generally has two sides that can be applied to the case's body for sound discovery. These sides correspond of a diaphragm (a plastic slice) and a bell(a concave mug). When the diaphragm is placed on the case's skin, the climate produced by internal body sounds cause the diaphragm to joggle. This vibration creates aural pressure swells that travel up the tubing, reaching the listener's cognizance. Again, if the bell is placed on the case's body, the climate of the skin directly produce aural pressure swells that travel up to the listener's cognizance. [illegible] bell and diaphragm serve distinct purp[illegible]ansmitting sound

ISBN 979-8-3503-5094-4

frequentness. The bell is complete at transmitting low-frequence sounds, furnishing clarity when assessing certain physiological functions like blood inflow. On the other hand, the diaphragm excels in transmitting advanced- frequence sounds, making it precious for discerning subtle details in heart or lung sounds. To control which part of the casket piece (bell or diaphragm) receives the aural energy, the tube connecting to the chamber between them is designed to be open on only one side and can rotate. When connected to the bell, the opening is visible, and by rotating the tube 180 degrees in the head, it connects to the diaphragm. This point allows healthcare professionals to fluently switch between the bell and diaphragm, acclimatizing their examination to the specific frequentness they need to concentrate on during

patient assessment.

Figure 8. STETHOSCOPE

1. *Acoustic performance: Frequency Response:* The stethoscope should have a broad frequency response to capture both high and low-frequency sounds.

2. *Sound Amplification:* The ability to amplify and transmit sounds effectively from the patient's body.

F. *HARDWARE IMPLEMENTATION:*

In covering fetal health, a stethoscope equipped with a mic detector serves as a precious tool for directly measuring the fetal twinkle. The mic detector is nearly connected to the stethoscope to capture the distinct sound of the fetal heart. This mic detector plays a vital part in converting sound swells into electrical signals. When sound swells, similar as the fetal twinkle, reach the microphone, they beget a diaphragm to joggle. This vibration generates an electrical signal commensurable to the breadth of the sound surge. The mic detector specifications give crucial perceptivity into its functionality. It operates within a voltage range of3.3 to 5 volts, with an functional current of 4- 5mA. The voltage gain is specified at 26dB, pressing its capability to amplify the incoming electrical signal. The microphone perceptivity range is detailed as ranging from 52 to 48 dB, indicating its responsiveness to a broad diapason of sound situations, also, the system incorporates a temperature detector to precisely measure the ambient temperature. The temperature detector interfaces with an 8- bit microcontroller, easing the affair of temperature and moisture values as periodical data. This microcontroller integration enhances the system's capabilities, allowing for effective data processing and transmission. The temperature detector employed in the setup is plant calibrated, simplifying its interface with other microcontrollers. This estimation ensures accurate and dependable temperature measures. The detector is designed to measure temperatures ranging from 0 °C to 50 °C, furnishing a comprehensive assessment of the thermal terrain. also, it also gauges moisture situations within a range of 20 to 90, with an emotional delicacy of ± 1 °C for temperature measures and ± 1 for moisture measures.

G. *DATASET PREPRATION:*

The ECG recordings are attained using a stethoscope which is connected with the mic detector, as the electrical signal is been generated. The data is recorded and also exported with the connection sector of using USB to PC. It captures the nanosecond signal which is hitching with the mic detector and it amplifies into an Analog to a electrical signal with the help of python software. To move a program onto a Micro Python board, a train is been created and copy it onto the microcontroller in order to execute. With the tackle connected to a device, similar as a computer, the board's flash drive will appear on the device allowing lines to be moved to the flash drive. There will be two being python lines boot.py andmain.py that are generally not modified main.py may be modified if it's to run the program every time the microcontroller is boggled, else, programs will be run using the REPL press. The pyboard contains an internal drive (filesystem) named/ flash which is stored within the board's flash memory, also, a micro SD card can be fitted into a niche and is accessible through/ sd. When boggled up, a pyboard must elect a file system to charge from either/ flash or/ sd with the current directory being set to either/ flash or/ sd. By dereliction, if an SD card is fitted,/ sd will be used, if not,/ flash issued. However, the use of the SD card for the booting process can be avoided by creating an empty train called/ flash/ SKIPSD which will remain on the board and live when the pyboard is boggled up and will skip the SD card for the booting process, If demanded. When the pyboard is powered up typically or the reset button is pressed also the pyboard is boggled in a standard mode, meaning that theboot.py train is executed, also the USB configured and eventually the python program will run. There's an capability to stamp the standard charge sequence through holding down the switch whilst the board is in the booting process and also pressing reset as continue to hold the switch. The pyboard's LEDs will flutter between modes and once the LEDs have reached the board will bobble in the specific mode. The jeer pi pico circuit correspond with fine tuning of a noise filtering process takes place which is consider as a constant value(0) and while crossing the threshold position it's consider as value(1). On detecting the temperature from the detector value the graph is colluded through periodical plot and periodical examiner as shown in the figure 9) and figure 10). But at the same time, both temperature and the heart beat dimension isn't sufficient to do because of intermediate variation in different values.

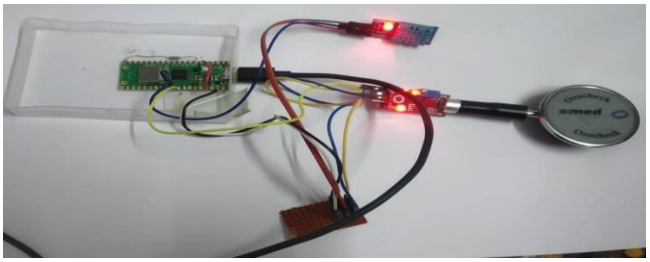


Figure 9. fetal heart sound detection kit

Figure10. Interconnecting raspberry pi wit micropython

In the described process, the analog signal representing fetal heart sounds undergoes conversion into an audio format. This metamorphosis is achieved using Py[illegible] the a[illegible] Python script is created for recording [illegible] from a microphone.

ISBN 979-8-3503-5094-4

and saving it as a WAV train, as illustrated in Figure 11. The Pyaudio law is specifically designed to capture the input signal, forming from the fetal heart, and convert it into a digital audio format. To execute this operation, the Python script is run, initiating the recording of audio from the microphone for a predefined duration. The recorded audio is also saved as a WAV train, a generally used format for digital audio. The final step involves vindicating the successful saving of the affair WAV train, icing that the fetal heart sounds have been directly captured and stored for farther analysis or medical assessment. This process demonstrates the application of Pyaudio to seamlessly restate analog signals of fetal heart sounds into a digital audio format, enabling effective recording and storehouse of vital medical data.

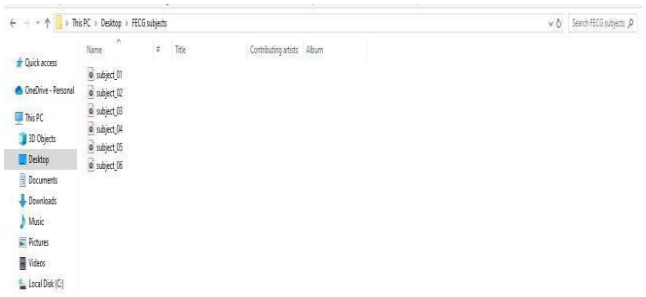


Figure 11. Converted audio file

III. RESULTS& DISCUSSION

A. *PREPROCESSING*:

In the posterior phase of the process, the fetal heart sound audio train, attained from the former way, is directed to MATLAB for in- depth analysis. A critical aspect of this integration involves periodical communication to seamlessly transfer the data, treating it as an audio train within the MATLAB terrain. This periodical communication ensures a smooth and effective transition of the acquired fetal heart sound data into MATLAB for farther examination. Upon entering MATLAB, the analog- to- digital conversion process takes priority, allowing the system to interpret the audio input signal digitally. Theacquired signal is characterized by a substantial gain of 500X, signifying that the input signal undergoes modification by a factor of 500 before the digitization process. This modification is visually represented in Figure11, demonstrating the magnitude of signal improvement. Likewise, the slice frequence is set at 6 Hz, denoting that the system captures 6 samples per second. This high slice frequence contributes to the system's capability to directly represent the nuances of the fetal heart sounds, icing that critical details aren't lost during the digital conversion. To upgrade the acquired signal for more precise analysis, notch pollutants are strategically employed. Notch pollutants play a pivotal part in barring specific frequentness, effectively addressing issues related to power- line hindrance. This hindrance, frequently manifesting as unwanted noise in the signal, can be efficiently removed through the perpetration of notch pollutants, enhancing the overall signal quality. The perpetration of a 500X gain, a 6 Hz slice frequence, and the objectification of notch pollutants inclusively contribute to creatinga detailed and refined representation of the fetal heart sounds within the MATLAB terrain. This step in the process sets the stage for posterior analyses, enabling healthcare professionals and experimenters to claw deeper into the complications of fetal cardiac health. The flawless transition of data from the original recording through Pyaudio to MATLAB showcases the integration of different technologies to achieve a comprehensive and accurate assessment of fetal.

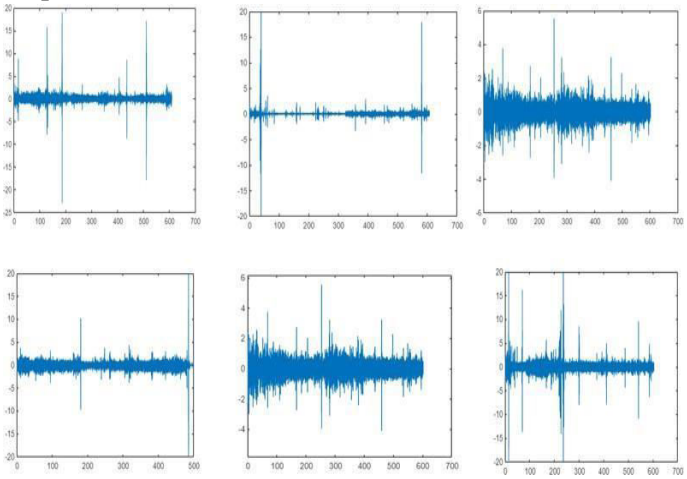


Figure 12. Raw data of Fetal heart sound

Analysis of fetal heart sound: A notch sludge serves as a vital element in signal processing, strategically designed to widely devaluate or exclude a narrow band of frequentness while allowing others to pass through fairly innocent. The term" notch" in this environment refers to the specific frequency band that the sludge targets, effectively cheapening or" engraving out" that particular range. In signal processing operations, notch pollutants find wide use to combat hindrance or noise that may be present in a signal. The beauty of a notch sludge lies in its capability to precisely target and alleviate unwanted frequentness without significantly impacting the rest of the signal. This makes it an inestimable tool for conserving the integrity of important information while removing undesirable rudiments.

In the technical sphere of Fetal Electrocardiogram (ECG) signal processing, the operation of a notch sludge becomes particularly significant. One common source of hindrance in medical signal processing is power line frequence, which frequently operates at 50 or 60 Hz depending on the geographical region. Power line hindrance can introduce unwanted peaks or vestiges into the ECG signal, potentially compromising the delicacy of fetal cardiac monitoring. By incorporating a notch sludge into the signal processing channel, these unwanted power line frequentness can be precisely targeted and downgraded. The result is a cleaner and more accurate representation of the fetal ECG signal, free from the deformations introduced by power line hindrance. This becomes pivotal in healthcare settings where the dedication of the ECG signal is consummate for individual and covering purposes. as shown in the figure 13.

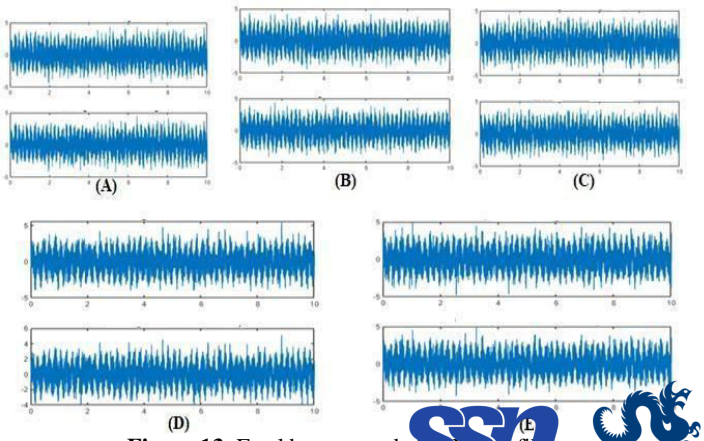


Figure 13. Fetal heart sound after Notch filter

ISBN 979-8-3503-5094-4

IV. CONCLUSION

The circuit design of raspberry pi based Heart rate monitor system in order to display the heartbeat readings in the graph as X-axis time and the Y- axis ADC value. This consists of the heart rate measurement using the single acquisition, by the sensors and the temperature measurement. The sensor is used to collect the patient's data, Information here to the wellbeing boundaries of temperature and heartbeat-rate. The sensor used over here are low power sensors. They collect data from the patient on daily basis and with this frequently obtained data the health condition of the patient is observed and required prescriptions are recommended. The raw data signal is been acquired and using ADC it is converted into a audio file. On going through this the notch filter is used for the filtering process of the fetal heart rate to ensure the ECG waveform. The future work is , the sensor is used to collect the patient's data and the amniotic fluid level, movement of the fetal are to be detected. These sensor enable data collection, which is then transmitted to a mobile application.

The gathered data is communicated via IOT and showcased on a mobile device for better fetal monitoring system in home.

REFERENCES

[1] R.Ranjana., "Smart System for Pregnancy using Raspberry Pi & IoT ,"International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056 Volume: 08 Issue: 04 | Apr 2021

[2] A.G. Skrivanos et al., "Fetus Heart Rate Monitoring: A Preliminary Research Study With Remote Sensing," in IEEE Consumer Electronics Magazine, vol. 11, no. 4, pp. 32-44, 1 July 2022, doi: 10.1109/MCE.2021.3097316.

[3] K. Hu et al., "A Wireless and Wearable System for Fetal Heart Rate Monitoring," 2021 3rd International Conference on Applied Machine Learning (ICAML), Changsha, China, 2021, pp. 402-407, doi: 10.1109/ICAML54311.2021.00091.

[4] J. A. L. Marques et al., "IoT-Based Smart Health System for Ambulatory Maternal and Fetal Monitoring," in IEEE Internet of Things Journal, vol. 8, no. 23, pp. 16814-16824, 1 Dec.1, 2021, doi: 10.1109/JIOT.2020.3037759.

[5] S.Shiny Amala , Dr.S.Mythili ., " IoT Based Health Care Monitoring System for Rural Pregnant Women," International Journal of Pure and Applied Mathematics Volume 119 No. 15 2018, 837-843.

[6] K. Sathesh, 2 S. Rajasekaran., " IOT Based Pregnancy Women Health Monitoring System for Prenatal Care ," International Journal of Advances in Engineering and Management (IJAEM) Volume 4, Issue 7 July 2022.

[7] Z. Chen, K. Zheng, J. Shen, Y. Lin, Y. Feng and J. Xu, "Sample Point Classification of Abdominal ECG Through CNN-Transformer Model Enables Efficient Fetal Heart Rate Detection," in IEEE Transactions on Instrumentation and Measurement, vol. 73, pp. 1-12, 2024, Art no. 6500412, doi: 10.1109/TIM.2023.3338710.

[8] R. Martinek et al., "Passive Fetal Monitoring by Advanced Signal Processing Methods in Fetal Phonocardiography," in IEEE Access, vol. 8, pp. 221942-221962, 2020, doi: 10.1109/ACCESS.2020.3043496.

[9] D. Jilani, T. Le, T. Etchells, M. P. H. Lau and H. Cao, "Lullaby: A Novel Algorithm to Extract Fetal QRS in Real Time Using Periodic Trend Feature," in IEEE Sensors Letters, vol. 6, no. 9, pp. 1-4, Sept. 2022, Art no. 7003204, doi: 10.1109/LSENS.2022.3200072.

[10] K. Barnova, R. Martinek, R. Jaros and R. Kahankova, "Hybrid Methods Based on Empirical Mode Decomposition for Non-Invasive Fetal Heart Rate Monitoring," in IEEE Access, vol. 8, pp. 51200-51218, 2020, doi: 10.1109/ACCESS.2020.2980254.

[11] A.Shokouhmand and N.Tavassolian, "Fetal Electrocardiogram Extraction Using Dual-Path Source Separation of Single-Channel Non-Invasive Abdominal Recordings," in IEEE Transactions on Biomedical Engineering, vol. 70, no. 1, pp. 283- 295, Jan. 2023, doi: 10.1109/TBME.2022.3189617.

[12] J. Xu et al., "Fetal Movement Detection by Wearable Accelerometer Duo Based on Machine Learning," in IEEE Sensors Journal, vol. 22, no. 12, pp. 11526-11534, 15 June15, 2022, doi: 10.1109/JSEN.2022.3172451.

[13] L. Wang, C. Zhao, M. Dong and K. Ota, "Fetal ECG Signal Extraction From Long- Term Abdominal Recordings Based on Adaptive QRS Removal and Joint Blind Source Separation," in IEEE Sensors Journal, vol. 22, no. 21, pp. 20718-20729, 1 Nov.1, 2022, doi: 10.1109/JSEN.2022.3206225.

ISBN 979-8-3503-5094-4

ISBN 979-8-3503-5094-4

Decoding Sleep: Microphone-Based Snoring Analysis using Embedded Machine Learning for Obstructive Sleep Apnea Detection

Delpha Jacob*, Priyanka Kokil†, Subramanian S ‡, and Jayanthi Thiruvengadam*
* Department of Biomedical Engineering, College of Engineering and Technology,
SRM Institute of Science and Technology, Kattankulathur, Chennai 603203
†Advanced Signal and Image Processing (ASIP) Lab, Department of Electronics and Communication Engineering,
Indian Institute of Information Technology, Design and Manufacturing, Kancheepuram, Chennai-600 127.
‡Department of Respiratory Medicine, SRM Medical College Hospital and Research Centre,
Kattankulathur, Tamil Nadu, India, 603203.
Email: {delphaj,*jayantht}@srmist.edu.in
{priyanka@iiitdm.ac.in},{drssmani@gmail.com}

***Abstract*—Snoring, a recurring habit often disregarded within the Indian community, can signal a grave underlying issue of Obstructive Sleep Apnea (OSA). OSA is a severe sleep disorder characterized by recurrent interruptions in breathing for more than 10 seconds during sleep, typically due to partial or complete airway obstructions. Neglecting OSA can lead to a range of significant health risks, including increased likelihood of occupational accidents, motor vehicle accidents, heightened susceptibility to severe depression, cardiac and cerebrovascular diseases, and reduced life expectancy. The main objective of the study is to detect snoring while at sleep and also to classify it as normal snoring and OSA snoring. Arduino nano 33 BLE sense is used to capture the snore signal, it houses a built-in MP34DT05 sensor. The sensor has a signal-to-noise ratio of 64dB and sensitivity of - 26dBFS ± 3dB. This captures the sound signal of the individual, it is further processed to extract the Mel-filter bank energy features, Mel Frequency Cepstral Coefficients and Spectrogram features. The features are further used to build a model and the same is trained using edge impulse to classify the signal. The dataset is divided into training, testing, and validation sets, with 80% of the data allocated to training, 20% to testing, and an additional 20% within the training data set aside for validation purposes. The results for the two class classification (snoring and non snoring) indicate that the spectrogram-based approach achieved an accuracy rate of 96.9%, while the other two methods yielded accuracy rates of 93.8%. The accuracy for three class classification (normal, snoring and OSA snoring) using the Embedded Machine Learning (EML) approach is 88%. The proposed study demonstrates enhanced accuracy in identifying OSA by snoring compared to previous research. This autonomous system can facilitate the detection of OSA through the analysis of snoring patterns, subsequently alerting the subject to implement pre-emptive measures for remediation. Timely intervention and rectification can enable the subject to attain an undisturbed and restful night's sleep, thereby augmenting their overall quality of life.**

***Index Terms*—Obstructive Sleep Apnea, Snoring, Embedded Machine Learning, Edge Impulse.**

I. Introduction

Snoring is a prevalent phenomenon that arises when the passage of air through the mouth and nose is obstructed to an extent during sleep. Snoring occurs due to the oscillation of tissues in the neck and air passage. Intermittent snoring is mostly benign, but persistent and loud snoring can occasionally indicate an underlying condition, such as sleep apnea. Snoring is a prevalent indication linked to Obstructive Sleep Apnea (OSA). It is a sleep condition marked by frequent occurrences of whole/ partial obstruction of the upper airway while sleeping resulting in brief interruptions in breathing [1]. OSA is frequently linked to excessive snoring, choking, or gasping noises when the individual attempts to resume breathing following the disruption [2,3]. These obstructions in breathing are frequently followed by attempts to inhale against the blocked air passage, leading to the production of snoring noises. The snoring in persons with OSA is typically characterized by its high volume, disturbing nature, and occasional pauses in breathing [4]. Individuals suffering from OSA may display loud and persistent snoring. The snoring may occur sporadically, characterized by periods of silence, followed by abrupt gasping or choking noises when respiration resumes [5-9]. It is essential to note that although snoring is a prevalent indication of OSA, not all individuals who snore suffer from sleep apnea. If an individual is encountering excessive and persistent snoring, accompanied by additional indications such as excessive fatigue during the day, it is recommended to seek guidance from a healthcare practitioner [8].

A polysomnography, also known as a sleep study, may be advised to diagnose OSA and ascertain the suitable treatment choices, such as lifestyle modifications, continuous positive airway pressure (CPAP) therapy, or other treatments depending on the severity of the problem [9]. A detailed review on CPAP and its technologies implied are explained in this research [10]. The use of snoring sound for detecting OSA can serve as an early non-contact diagnostic method. Researchers are investigating the potential use of snoring sound evaluation as a diagnostic tool and analysing the properties of snoring sounds in order to identify patterns corresponding with sleep apnea.

ISBN 979-8-3503-5094-4

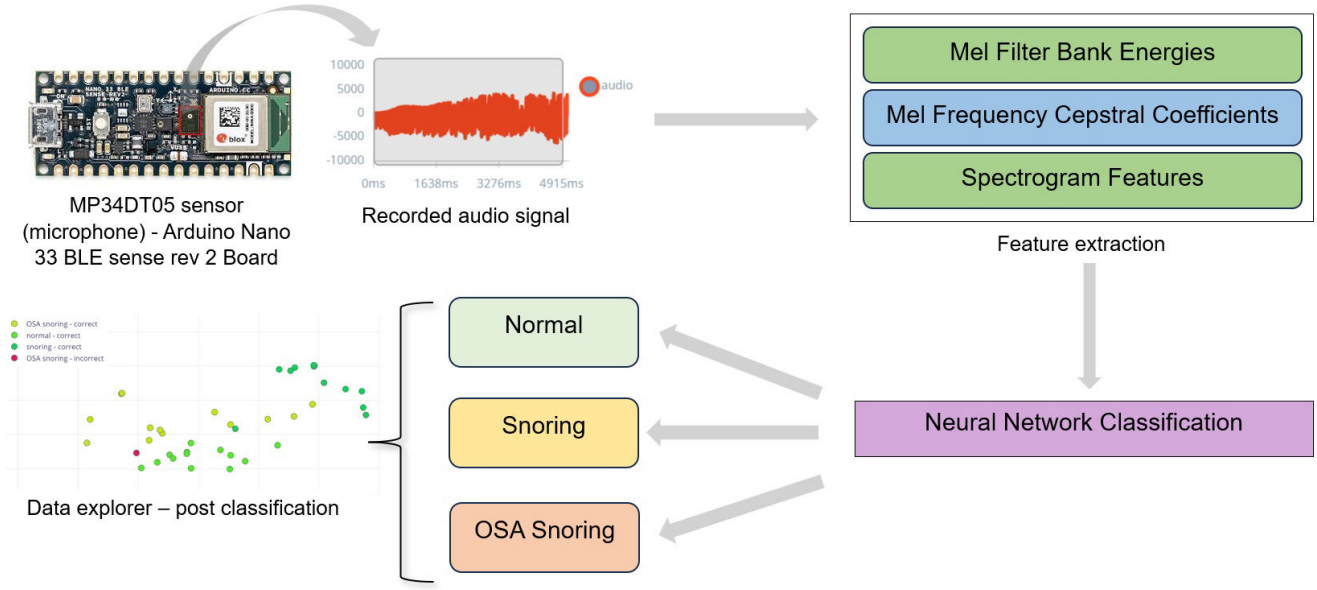


Fig. 1: Overall process flow for snore analysis-based sleep decoding

This involves studying the frequency, intensity, and timing characteristics of snoring. Utilizing multi-layer neural networks [11], hybrid neural networks [12] and deep learning [13, 16] to detect snoring during sleep. In general, a microphone, video recorded and polysomnography monitor and software are used to analyse the snoring status of the subject. [14, 15]. The suggested approach showcases an integrated machine learning technique for utilizing a standalone device coded in embedded C to perform machine learning classification for snoring detection.

II. Methodology

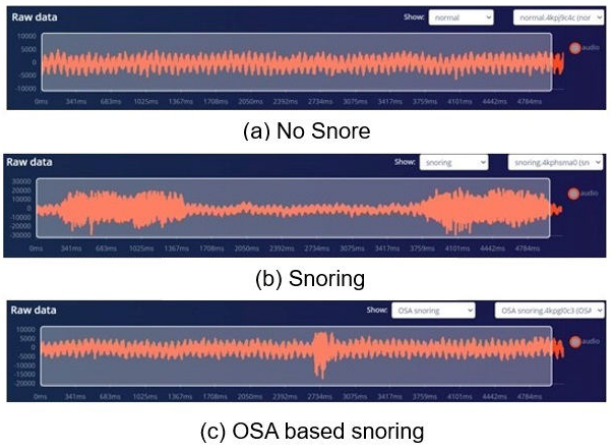


Fig. 2: Signals obtained for 5ms (5000 samples) in 3 different classes

The Arduino Nano 33 BLE Sense is a mini development board equipped with a Cortex-M4 micro controller. It was used to acquire the snore sound, inbuilt microphone MP34DT06JTR Sensor, which acquires the snore signal. The firmware for this development board, known as Edge Impulse, is openly available and can be found on the GitHub platform. The obtained snore signal is processed and trained for classification in Edge impulse. Open-source snore signal for normal and OSA based subjects are used for training the device. The edge impulse CLI and Arduino CLI are installed and the command prompt is used to connect the device with the firmware. Once the development board is connected with the system via micro-USB cable, the firmware is ready to be used.

A. *Data processing and feature extraction*

The data is directly acquired and labelled in the edge impulse platform. The data is split into training and testing set as 60-40 split. The first trial involved 2-class classification as snore and no snore. The second trial further involved one more class, making it a 3-class classification as no snore, snore and OSA snore. Mel-Filter bank energy features, Mel frequency cepstral coefficients and spectrogram are used for extraction of features.

Table 1 gives the characteristics of the features extracted by the methods mentioned. Figure 2 depicts the signal obtained; the sampling rate is 5000 ms. The window size is 5000 ms, the entire signal is used for feature extraction. Window length increases for every 1000 ms and the sampling frequency is 16,000 Hz for audio signals. Figures 3 to 5 depicts the Mel filter bank energies, spectogram and cepstral co-efficient for the 3 different classes of signals obtained. It is observed that the energies in the snoring is more when compare to no snore and OSA based snore class. OSA based snoring has many pauses, therefore the energy bands in the OSA based snoring is less when compared to the energy in the snoring data.

ISBN 979-8-3503-5094-4

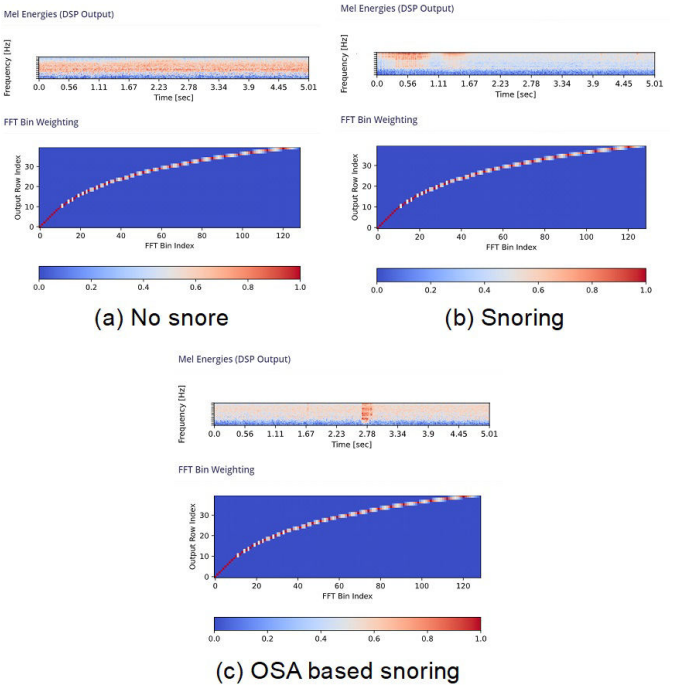


Fig. 3: Mel Filter bank energies for the different signal classes

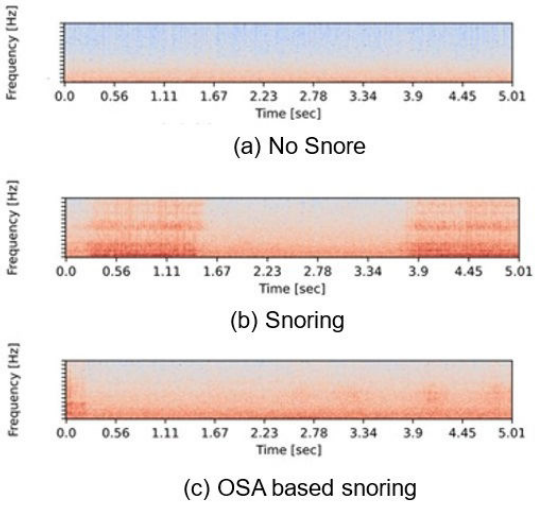


Fig. 4: Spectrogram for the different signal classes

B. *Classifier model development and deployment*

Based on the features extracted the model is developed. The number of training cycle in the neural network setting is 30, the learning rate is maintained as 0.0005. A total of 19960 features are extracted and given to the network. The network contains a dense layer of 20 neurons and a additional dense layer of 10 neurons; the output layer has 3 classes. The keras are created and the model is trained with the given (80%) dataset. The target is set as Cortex -M4F 64Mhz: Arduino nano 33 BLE

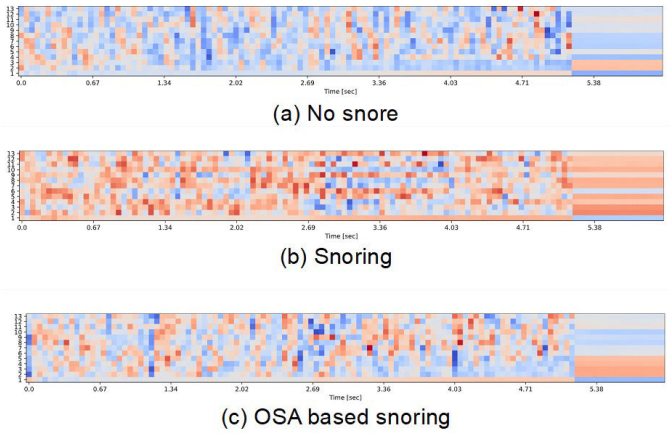


Fig. 5: Cepstral coefficients for the different signal classes

TABLE I: Comparison of features extracted by Mel-Filter bank energy features, Mel frequency Cepstral coefficients and Spectrogram

	Mel-filterbank energy features	Mel Frequency Cepstral Coefficients	Spectrogram
Processed Features	3822	637	10260
Frame length	0.025	0.025	0.064
Frame Stride	0.01	0.02	0.048
Filter number	39	32	
FFT length	128	512	128
Duration	192 ms	296 ms	164ms
RAM usage	18KB	163KB	28KB

sense. Once trained the data points can be viewed in the data explorer. The EON tuner is then tuned, it helps us find the most optimal architecture for the specified embedded machine-learning application. Once the optimal model is obtained by studying the EON tuner results. The model is tested (20% from the training dataset) and the results are tabulated. The model is then made ready for live classification. The testing is done for live classification. The model can be tuned for optimal solutions. Once obtained the same is deployed. The C++ codes are built. The downloaded file in then dumped on the controller using Arduino IDE platform. The controller now acts as an independent device for classification of snoring. The controller thus works on embedded machine learning algorithm. Figure 6 illustrates the neural network developed for the classification of the snoring and OSA based snoring sounds. The neural network comprises of 1 input layer, 7 hidden layers and 1 output layer. The input layer contains 1287 features, these features are processed in the hidden layers consecutively and reduced to 448 features and the given to the output layer for classification, this classifies the given sound into one of the 3 classes mentioned. The hidden layers include

ISBN 979-8-3503-5094-4

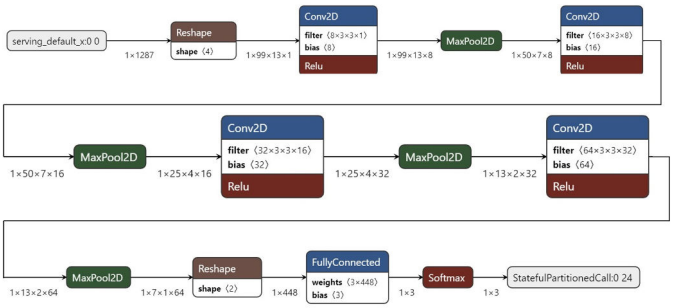
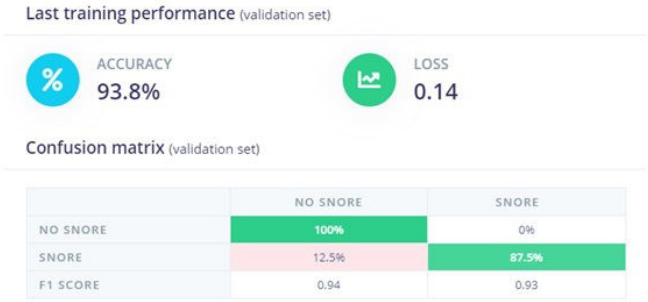


Fig. 6: Developed Net for classification

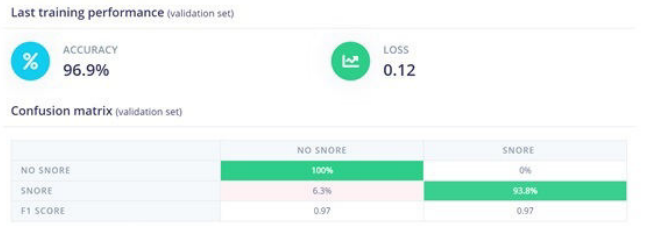
III. Results and Discussion

The classification is done based on comparison with nearest clusters. The data is split as 60-20-20 (training- testing-validation). The training performance results for the first trial: 2 - class classification using spectrogram, MFE and MFCC are compared in figure 7 (a) to (c) It is observed that the validation using spectrogram as signal processing method gives better results of 96.9% when compared to MFE and MFCC which gives 93.8%. The model is tested using all 3 processing blocks, the accuracy is observed to be 88.89%, but when the model is tested using only spectrogram the accuracy is obtained as 94.29%. Therefore, it is concluded that the spectrogram as a processing block for snore classification gives promising results when compared to MFE and MFCC. Figure 8 shows the classification accuracy for second trial: 3 - class classification. It is observed that the accuracy for OSA based classification is obtained as 87.5%. The loss for this validation set is very minimal, it is 0.38. The F1 score obtained for OSA snoring, normal and snoring is 0.86, 0.86 and 1 respectively. Figure 9 illustrates the arduino IDE - EML implementation, a custom deploy block is build using edge impulse for the trained model. The C++ codes are deployed, the model is then imported into arduino IDE uploading onto the controller, the same can thereby be used for embedded applications.

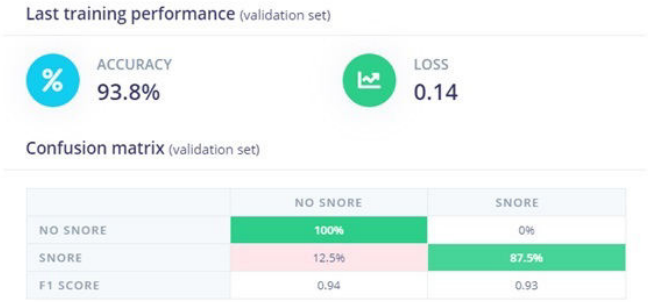
Previous research in detection of snoring and OSA based snoring includes Jiali et al. [17] suggested a method for snoring identification that records audio from each person at five separate locations around their head by using a convolutional neural network (CNN) along with a recurrent neural network (RNN) for snore detection. The method concluded that snoring episodes from audio recordings can be identified with a high degree of accuracy and that the location of the microphone did not significantly affect detection ability performance. A classifier based on Long Short-Term Memory (LSTM) was employed by Cheng et al. [18] to distinguish respiratory event-related snoring from regular snoring. For respiratory event-associated with normal and regular snoring they got an accuracy of 95.3%, and the snore linked to the severity of OSA for 3 categories of classification has an accuracy of 81.6%. Based on the theory that OSA-related snoring can help identify the site-of-upper airway collapse, two methods were designed to extract OSA-related snore using an Linear Discriminant



(a) Validation result of MFE



(b) Validation results of Spectogram



(c) Validation results of MFCC

Fig. 7: Validation results for each signal processing methods

Fig. 8: Testing results for the proposed model

Analysis (LDA) classifier by A Sebastian et al. [19]. The deployment of the program on an Android smartphone and a deep learning model for snoring detection is explained by Ansari [20]. In order to automatically identify snoring, a deep learning model for snore detection is created and then applied to a listener module and a wearable device developed by Khan [21] is designed to vibrate the upper arm until the snorer turns over to sleep on the side.

According to Table II, the suggested study has a higher accuracy in identifying OSA based on snoring compared to previous research, with an acceptance rate of 97%. The

ISBN 979-8-3503-5094-4

TABLE II: Comparison of the proposed work with similar research done

Author & reference	Dataset	Method/Features	Network used	Classification	Results
Jiali et al. [17]	38 subjects, 5 different microphone positions for each.	2D spectrogram	CNN and RNN	Snore and no snore	Accuracy: 95.3 ±0.5%,
S.Cheng et al. [18]	33 patients and 10 normal subjects	MFCC, Mel Filter Banks (Fbanks), Short-time Energy and Linear Prediction Coefficient (LPC),	LSTM	Normal, snoring and OSA	Accuracy: 95.3% (Normal and snoring) Accuracy: 81.6%. (OSA snoring)
A Sebastian [19]	58 patients	Time-domain features: Energy, Entrop, ZCR. Frequency-domain features: First three formant frequencies, MFCC, Spectral chromaFeatures, Spectral Entropy, Flux, Centroids, Fundamental Frequency and Harmonic Frequency.	LDA and multi class LDA	Normal, snoring and OSA	Accuracy: 87% (Normal and snoring) Accuracy: 72%. (OSA snoring – LDA classifier) Accuracy: 64%. (OSA snoring – multi class LDA classifier)
M.W Ansari [20]	6000 audio samples from Kaggle dataset	MFCC, Spectrogram		Snore and Non - snore	Accuracy: 98%.
T. Khan [21]		MFCC	CNN and DL	Snore and Non - snore	Accuracy: 96%
Proposed work		**MFCC, Spectrogram and Mel Filter Banks**	**EML**	**No snore, Snoring, OSA**	**Accuracy: 97% (for no snore and snore using spectrogram) EML - 88% (for OSA snoring - Fig 6)**

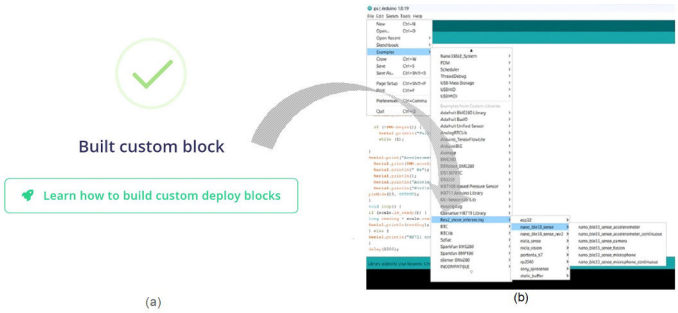


Fig. 9: Arduino IDE - EML implementation (a) a custom deploy block is build using edge impulse for the trained model (b) The deployed model is imported into Arduino IDE for uploading onto the controller for embedded applications

accuracy of the EML method is 87.5%. However, despite effective identification of snoring or OSA based on snoring in previous study, no classification based on EML has been conducted thus far. The suggested research introduces a novel framework for categorizing snoring, which can be used to any embedded system platform, thereby expanding its potential applications.

IV. Conclusion

Among various methods for classification of snoring, the proposed methodology utilises a novel EML-based snore classification. This helps the device to work as a stand-alone classifier device, which can be used in any hardware based embedded applications. The enhances the performance of the embedded system and also reduces the dependency of the device on any particular software as the EML can be deployed into the controller using embedded C language. This improves the robustness and performance of the embedded system in which the controller is used. In the proposed system, the controller uses EML to screen OSA in snorers.

ISBN 979-8-3503-5094-4

REFERENCES

[1] J. M. Slowik and J. F. Collen, "Obstructive Sleep Apnea," PubMed, Dec. 11, 2022. https://www.ncbi.nlm.nih.gov/books/NBK459252/

[2] N. Maimon and P. J. Hanly, "Does Snoring Intensity Correlate with the Severity of Obstructive Sleep Apnea?," Journal of Clinical Sleep Medicine, vol. 06, no. 05, pp. 475–478, Oct. 2010, doi: https://doi.org/10.5664/jcsm.27938.

[3] J. R. Perez-Padilla, E. Slawinski, L. M. Difrancesco, R. R. Feige, J. E. Remmers, and W. A. Whitelaw, "Characteristics of the Snoring Noise in Patients with and without Occlusive Sleep Apnea," American Review of Respiratory Disease, vol. 147, no. 3, pp. 635–644, Mar. 1993, doi: https://doi.org/10.1164/ajrccm/147.3.635.

[4] "Basic and Clinical Research on Adult Snoring and Obstructive Sleep Apnea Syndrome — Frontiers Research Topic," www.frontiersin.org. https://www.frontiersin.org/research-topics/44440/basic-and-clinical-research-on-adult-snoring-and-obstructive-sleep-apnea-syndrome (accessed Jan. 24, 2024).

[5] A. Alakuijala and T. Salmi, "Predicting Obstructive Sleep Apnea with Periodic Snoring Sound Recorded at Home," Journal of Clinical Sleep Medicine, vol. 12, no. 07, pp. 953–958, Jul. 2016, doi: https://doi.org/10.5664/jcsm.5922.

[6] S. Akhter, U. R. Abeyratne, V. Swarnkar, and C. Hukins, "Snore Sound Analysis Can Detect the Presence of Obstructive Sleep Apnea Specific to NREM or REM Sleep," Journal of Clinical Sleep Medicine, vol. 14, no. 06, pp. 991–1003, Jun. 2018, doi: https://doi.org/10.5664/jcsm.7168.

[7] L.A. Lee et al., "Energy Types of Snoring Sounds in Patients with Obstructive Sleep Apnea Syndrome: A Preliminary Observation," PLoS ONE, vol. 7, no. 12, p. e53481, Dec. 2012, doi: https://doi.org/10.1371/journal.pone.0053481.

[8] A. M. Osman, S. G. Carter, J. C. Carberry, and D. J. Eckert, "Obstructive sleep apnea: current perspectives," Nature and Science of Sleep, vol. 10, pp. 21–34, Jan. 2018, doi: https://doi.org/10.2147/NSS.S124657.

[9] "Obstructive Sleep Apnea," www.hopkinsmedicine.org. https://www.hopkinsmedicine.org/health/conditions-and-diseases/obstructive-sleep-apnea (accessed Nov 10, 2023).

[10] D. J, P. Kokil, S. S and T. Jayanthi, "Comparative Study of Various Breath Assist Devices and the Effect of Sleep Apnea (SA) Among the Indian Population," 2023 International Conference on Device Intelligence, Computing and Communication Technologies, (DICCT), Dehradun, India, 2023, pp. 1-6, doi: 10.1109/DICCT56244.2023.10110129.

[11] T. L. Nguyen and Y. Won, "Sleep snoring detection using multi-layer neural networks," Bio-Medical Materials and Engineering, vol. 26, no. s1, pp. S1749–S1755, Aug. 2015, doi: https://doi.org/10.3233/bme-151475.

[12] B. Kang, X. Dang and R. Wei, "Snoring and apnea detection based on hybrid neural networks" 2017 International Conference on Orange Technologies (ICOT), Singapore, 2017, pp. 57-60, doi: 10.1109/ICOT.2017.8336088

[13] B. Wang et al., "Obstructive Sleep Apnea Detection Based on Sleep Sounds via Deep Learning," Nature and Science of Sleep, vol. Volume 14, pp. 2033–2045, Nov. 2022, doi: https://doi.org/10.2147/nss.s373367

[14] V. V. Tipparaju et al., "Reliable Breathing Tracking With Wearable Mask Device," in IEEE Sensors Journal, vol. 20, no. 10, pp. 5510-5518, 15 May15, 2020, doi: 10.1109/JSEN.2020.2969635

[15] T.-H. Kim, Jeong Whun Kim, and K. Lee, "Detection of sleep disordered breathing severity using acoustic biomarker and machine learning techniques," Biomedical Engineering Online, vol. 17, no. 1, Feb. 2018, doi: https://doi.org/10.1186/s12938-018-0448-x.

[16] M. W. Ansari, A. Rajak and R. Basak, "A Deep Learning Model to Snore Detection Using Smart Phone" 2021 12th International Conference on Computing Communication and Networking Technologies (ICCCNT), Kharagpur, India, 2021, pp. 1-5, doi: 10.1109/ICCCNT51525.2021.9580153.

[17] J. Xie et al., "Audio-based snore detection using deep neural networks," Computer Methods and Programs in Biomedicine, vol. 200, p. 105917, Mar. 2021, doi: https://doi.org/10.1016/j.cmpb.2020.105917.

[18] S. Cheng et al., "Automated sleep apnea detection in snoring signal using long short-term memory neural networks," Biomedical Signal Processing and Control, vol. 71, p. 103238, Jan. 2022, doi: https://doi.org/10.1016/j.bspc.2021.103238.

[19] A. Sebastian, P. A. Cistulli, G. Cohen, and P. de Chazal, "Automatic Classification of OSA related Snoring Signals from Nocturnal Audio Recordings," arXiv.org, Mar. 02, 2021. https://arxiv.org/abs/2102.12829 (accessed Nov. 16, 2023).

[20] M. W. Ansari, A. Rajak and R. Basak, "A Deep Learning Model to Snore Detection Using Smart Phone," 2021 12th International Conference on Computing Communication and Networking Technologies (ICCCNT), Kharagpur, India, 2021, pp. 1-5, doi: 10.1109/ICCCNT51525.2021.9580153.

[21] T. Khan, "A Deep Learning Model for Snoring Detection and Vibration Notification Using a Smart Wearable Gadget," Electronics, vol. 8, no. 9, p. 987, Sep. 2019, doi: https://doi.org/10.3390/electronics8090987.

ISBN 979-8-3503-5094-4

Tenth International Conference on Biosignals, Images and Instrumentation, SSNCE, Chennai, March 20 – 22, 2024

TRACK 3

DIGITAL DIAGNOSIS OF MOUTH DISEASE USING DEEP LEARNING ALGORITHMS

G.Sudha[1], M.Mohammadha Hussani[2], N.Bagyalakshmi[3], R.Avanthika[4], M.Saraswathi[4], S.Renuka[4]

1- Professor/Department of Biomedical Engineering, Muthayammal Engineering College, Rasipuram

2- Professor/ Department of EEE, Government College of Engineering, Erode

3- Associate Professor, Department of Biomedical Engineering, Adhiyaman College of Engineering, Hosur

4- Final Year UG Student/Department of Biomedical Engineering, Muthayammal Engineering College, Rasipuram

Corresponding mail ID: kavisudhamariyaa@gmail.com, hussaini1008@gmail.com, bagyabmeace2020@gmail.com, avanthikaraja17@gmail.com, saraswathimathesh2003@gmail.com, renukasubramani597@gmail.com

Abstract-The ability to identify and diagnose oral and mouth problems has significantly increased with the use of digital instruments. Those with potentially malignant oral diseases have an increased risk of developing lip or oral cavity cancer. Poor brushing, hormone fluctuations, and flossing practices that let plaque, a sticky layer of germs, are typically the causes. Since the mouth is thought to be a reflection of one's overall health, there is presently no personal gadget available to track one's dental health. Mouth disease to be correctly diagnosed and treated, an accurate forecast is necessary. It is possible to prevent dental disorders that gradually weaken tooth roots by identifying them early on. The goal of this effort is to develop a low-cost, multimodal, personal oral sensing device that perceives and classifies data automatically, enabling the physician to diagnose patients early and treat them effectively. The mouth disease prediction consists of preprocessing and classification process. The first step in mouth disease prediction wiener filter applied to filter the noise in testing image of dataset. After that augmentation involved to expand the number of images from the limited images. Finally, the comparison algorithms such as federated learning, multilayer perceptron (MLP) and deep belief networks (DBN) to produce the outputs as precision, recall, accuracy was analyzed. The 85% accuracy of the deep belief network predict mouth disease for more efficient.

Keyword: Mouth disease, preprocessing, classification, MLP, federated learning, deep belief network

1. Introduction

Globally, there is a rise in the incidence of oral squamous cell carcinoma (OSCC), a frequent malignancy. For OSCC, surgery has historically been the recommended primary cornerstone therapy. Also, multimodality therapy with concurrent chemoradiotherapy becomes essential because of the aggressive character of OSCC and the high diagnosis rate of locoregionally progressed cancer. The high incidence rate and the less-than-ideal treatment success remain a major worry despite the previously indicated therapy options.

Machine learning-related computer applications are evolving fairly rapidly. The potential for data analytics, data mining, and mathematical modeling has significantly increased due to the volume of data produced by ubiquitous computing devices, such as smartphones, online apps, and the internet of things.

Applications of machine learning have been seen in a variety of fields, including sales forecasting and commercial decision-making. The foundation of machine learning is training data, which is a collection of past business transaction data. Many huge systems have begun to use their data for machine learning in order to provide a better experience for its consumers and clients. Facebook and Google, for example, collect demographic information and user preferences to improve ad targeting. Gmail uses machine learning to screen spam emails. Machine learning is being used by more companies to improve their operations.

Early detection is critical for a better prognosis, treatment plan, and chance of survival. To improve the appropriate management of cancer, this is crucial. Even though our knowledge of the molecular pathways behind cancer has improved recently, late diagnosis has impeded the pursuit of precision treatment. As a consequence, deep machine learning has been identified as improving early identification and so minimizing moth disease-specific mortality and morbidity. It is clear that automated examination of images has the possibility to help pathologists and clinicians recognize early stages of oral illness and make educated decisions about ulcer management.

Because it provides the chance to compile the data from several systematic studies that have

ISBN 979-8-3503-5094-4

looked at the application of deep learning for prognosticating mouth cancer results, we have chosen to employ this method of systematic review mouth disease prediction. In our earlier work, we observed at the use of deep learning methods to mouth disease prediction. This will provide components for the process from exact diagnosis to precise treatment, which will aid in the endeavor to establish appropriate mouth disease management.

2. Literature review

Javed Rashid et al., (2023) A person's quality of life can be severely compromised by oral submucous fibrosis (OSMF), a debilitating illness of the oral mucosa, and possibly malignant condition (PMD). Using deep learning to identify situations distressing the oral cavity and mouth disease is exciting. Using the InceptionResNetV2 Method, the present study developed a mouth and oral diseases classification system to identify conditions such gum disease, canker sores, cold sores, mouth cancer, oral lichen planus, oral thrush, and oral cancer (OC).

Thanh Ngoc Nguyen and Dinh Ngoc Minh, (2021) Understanding the contributions of social and environmental factors to the outbreak of diseases is dependent on utilizing machine learning to discover patterns of hand, foot, and mouth infections. This study utilizes data from Vietnam to identify the most important elements driving the increase in cases, as well as the models that are most successful in forecasting this growth. Temperature is an important element, and the Random Forest Regressor model yields the greatest outcomes.

Bin Zhang et al., (2017) Finding signs of severe HFMD is essential for the disease's early prevention and management. In order to do this, 185 instances of severe and 345 cases of moderate HFMD were evaluated. Data was gathered on the patient's clinical characteristics, MRI results, laboratory test results, and demographics. After that, the variables' relative significance (RI) and interaction effects were ascertained using a gradient boosting tree (GBT).

Guangjian Liu et al., (2017) Children with intermediate hand, foot, and mouth disease (HFMD) may present with early clinical signs similar to those of severe HFMD, but as the disease worsens, the children quickly get worse. Using a 10-fold cross-validation method and a holdout approach, we constructed four random forest models with different variable sets. The HFMD-RF prediction system achieved accuracy, sensitivity, and specificity in the independent test set of 0.916, 0.824, 0.931, and 0.916, respectively.

Yongbin Wang et al., (2019) On the other hand, the number of instances that were documented was quite high, with a notable summer high-risk periodicity. According to the LSTM, incident incidences would continue to be rather high with a little upward tendency in the future. In this sense, it is important to emphasize the use of the LSTM technique in HFMD epidemic forecasting, which will help decision makers make effective judgments based on early illness incidence identification.

Suraj Verma et al., (2021) introduced a unique Hybrid Deep Neural Networks that use combined information from clinical and imaging data to diagnose HFMD. The network combines a convolutional neural network and a multi-layer perceptron (MLP) network into a single framework. We used both clinical and imaging data, as well as data from both sources, in a number of investigations. Our suggested Hybrid Deep Neural Networks were compared and assessed with a standard CNN model and an MLP model. The unique technique supports the current models of clinical symptoms-based illness classification and picture classification, especially the MLP model. The findings of the cross-validated trials show that the suggested hybrid deep neural networks are 99%–100% accurate in diagnosing the illness.

Rasheed et al., (2022) One of the most common malignancies in the world, oral squamous cell carcinoma (OSCC), is becoming more common in a number of countries. Significant grounds for worry are the high incidence rate, insufficient treatment planning, and delayed diagnosis. For a better prognosis, course of therapy, and survival probability, early detection is crucial. Despite recent breakthroughs in molecular pathway comprehension, treating OSCC patients with precision medication and late diagnosis remain significant challenges.

Wenzhe et al., (2020) The training set consisted of 886 intraoral images of primary teeth using a standard neural network (CNN) structure. The AI model evaluated 98 intraoral pictures of primary teeth in order to confirm clinical viability. Additionally, a digital camera was used to take pictures of the teeth. After looking over the pictures, a skilled pediatric dentist indicated the areas that had plaque. Following the application of a plaque-disclosing agent, the plaque-affected sites were detected. After a week, the dentist drew the plaque region on the 98 digital camera photos once more to evaluate the consistency of the manual diagnosis. Additionally, the AI model tagged 102 intraoral photos of primary teeth with the plaque locations it found, enabling the dentist to evaluate the diagnostic efficacy of each technique using lower-resolution images. To show detection

ISBN 979-8-3503-5094-4

accuracy, the mean intersection-over-union (MIoU) measure was used.

Xialv Lin et al., (2021) Rapid social and economic progress is accompanied by the phenomena of the uncontrollably widening epidemic of infectious illnesses. As a result, the number of instances of HFMD has skyrocketed, placing children's health, particularly infants and young children, in grave danger. Therefore, it is essential to use big data to estimate the frequency of HFMD illnesses and establish early-warning systems for HFMD in specific regions. Direct case reports, which employ statistical approaches in both time and location to have independent early-warnings of epidemics, are the main technique utilized in the current HFMD early-warning investigations.

Zhijin Wang et al., (2020) A key element of cyber-physical public health systems that aid in disease preventive and control choices is the prediction model. With the greatest incidence rate, HFMD is among the most prevalent infectious illnesses in the world. The time series that counts in equal-grained time intervals served as the foundation for earlier HFMD prediction models. The time series does, however, contain some information that are counted in precise time intervals. We suggested a method to gain from both fine-grained and equal-grained data: DGR model, or dual-grained depiction. Data input is initially represented to temporal patterns by the DGR. Predictions are then produced by consolidating the patterns that have been depicted.

Xiaochi Liu et al., (2019) An intestinal infection known as HFMD can cause serious health problems, including cardiorespiratory failure and even death. For this reason, early illness prevention and control depend heavily on the assessment of significant characteristics and the prediction of severe HFMD. In order to achieve this, an assessment was conducted on 658,689 instances, including 6,579 severe cases. To develop a simple, automated, and effective server HFMD prediction system using weather and hospital case data. To assess feature relevance, we employed the Random Forests and Adaboost algorithms. Our model can assess the significance of characteristics, according to preliminary testing results, however more parameter modification is still required to predict severe HFMD.

Gizem Tanriver et al., (2021) Around 177,757 people worldwide lose their lives to oral cancer each year, making it the most frequent kind of head and neck cancer. Oral malignancies have a 75–90% survival rate when detected in their early stages. Nonetheless, the most of cases are discovered at an advanced stage, primarily due to inadequate public awareness of the warning signs of mouth cancer and delayed referrals to specialists in oral cancer. Ocular potentially malignant diseases (OPMDs) are cancer-threatening illnesses.

Ren Zhong et al., (2018) An assortment of enteroviruses is the cause of HFMD. Millions of youngsters in Southeast Asia are afflicted by it. An accurate HFMD epidemic forecast might make it easier for public health experts to recommend public health measures ahead of time. To lessen the harm brought about by an HFMD outbreak, several experts worked to create an early warning system for the disease. Although it is challenging to get, daily research data may be useful in determining the connection between HFMD and environmental variables. In order to analyze HFMT eruptions, we collected daily medical data and various environmental variables from the Shenzhen Health Information Center. We entered the temperature and air quality parameters from the previous seven days, the HFMD rates from the preceding sixty days, and the HFMD incubation time into the tree model.

Dan Zhao et al., (2022) created a multimodal deep learning technique to use oral health issues to predict systemic illnesses and disorders. Using the dual-legged autonomous encoding, the features related to the coreland were initially retrieved from 1188 panoramic radiographs in Grid I. In order to forecast systemic disorders, we next combined the picture attributes in the second phase with clinical and demographic data from electronic health records (EHR). We assessed our model using accuracy and receiver operation characteristics (ROC). An unobserved test dataset was used to further validate the model. According to the findings, the suggested model had an AUC of 0.92 (95% CI, 0.90–94), 0.87 (95% CI, 0.84–89), and 0.78 (95% CI, 0.75–81) for systemic disorders.

Anwar Alhazmi et al., (2020) Artificial neural networks, which are widely utilized in data mining, were employed in the development of the artificial intelligence prediction model. The model was created using a total of 29 variables that were connected to the patients. The dataset was divided into 19 (25%) instances for the testing dataset and 54 (75%) cases for the training dataset at random. An oral pathologist with board certification examined all of the documents and observations. Seventy-three patients in all fulfilled the requirements. 51 instances (69.86%) were malignant, while 22 cases (30.13%) were benign. With a mean age of sixty-three years, there were thirty-seven females and thirty-six males.

Soualihou et al., (2022) Machine learning has demonstrated several benefits in oral healthcare, including the ability to anticipate dental cavity in

ISBN 979-8-3503-5094-4

the general population. In contrast to the previously mentioned traditional dental caries prediction, this study highlights the need for combining many data sources, also known as multi-modality, in order to extract more characteristics and provide more accurate results. The suggested multimodal data prediction model produced excellent results, with accuracy, F1-score, recall, and precision ratios of 90%, 89%, 90%, and 89%, respectively.

Hung et al., (2022) presented a work based on public data from the National Center for Health Statistics in the United States on the categorization of the presence or absence of root caries. During the evaluation, SVM produced the best results of the algorithms utilized.

Liu et al., (2020) To expect caries among people of Liaoning, China, a generalized regression neural network (GRNN) model was built. Regression on highly linked data produces false findings and overfitting.

Priyanka and Sunil., (2021) Examine predictive analytics in the context of oral health care, namely dentistry, which is a medical profession concerned with the anatomy, development, and problems of teeth. This topic is being addressed since it is a difficult endeavor due to a lack of precise data in dentistry for diagnosis and prognosis. Predictive analytics helps practitioners make smart decisions about a patient's medical state and treatment.

Viswanatha., (2022) Determine the difference between a malignant lesion, superficial spreading, and nodular melanoma. This allows for early viral detection as well as the rapid isolation and treatment required to prevent future virus transmission. Deep learning (DL) and the traditional non-parametric machine learning approach are demonstrated by convolutional neural network (CNN) deep layer topologies, which are neural network algorithms.

Sudha et al., (2023) Lung disease Computer Tomography (CT) examination pictures are utilized to describe and categorize lung nodes, as well as to describe the exact position of that node. U-Net armature is used to segment CT checkup pictures. Then we used CNN algorithms and transfer literacy models to classify cancer as positive or negative in lung images (VGG16).

Sudha et al., (2023) Even for seasoned professionals, skin cancer diagnosis (SCD) can be difficult since it involves a visual inspection of skin lesions and a high degree of expertise to discriminate between anomalies. This makes skin cancer detection more difficult. The use of Computer Aided Diagnosis (CAD) technology may increase SCD accuracy. This paper describes a Deep Learning Approach (C-DLA) for SCD based on Curvelets. The Curvelet transform is used to dissect the dermoscopic pictures initially. The low frequency components are then categorized using the Convolutional Neural Network (CNN). On PH2 database images, the recommended C-DLA approach produces good results with 99% accuracy for classification.

Sharaf and Arshad., (2021) A combination of k-means clustering, the Long Short-Term Memory Network of Recurrent Neural Networks (RNN), random forest, CNN, and boosting techniques was utilized to classify breast lumps as benign, malignant, or normal. Furthermore, the proposed BMC system is compared to current classification approaches using two publicly available datasets of mammographic images. For the DDSM dataset, the suggested BMC system achieves 0.97%, 0.98%, 0.97%, 0.96%, and 0.97%, 0.97%, and 0.95% values for sensitivity, specificity, F-score, and accuracy, respectively.

3. Proposed methodology

The proposed approach is mouth disease prediction using deep learning techniques consisting of preprocessing and classification. The preprocessing step is involved to reduce or filter the noise. After that classifications are to predict the mouth disease using federated learning algorithm, multilayer perceptron and deep belief network are discussed and produce the result comparison as precision, accuracy and recall for better performance. The processing of mouth disease prediction represented the flow diagram is shown in figure 1.

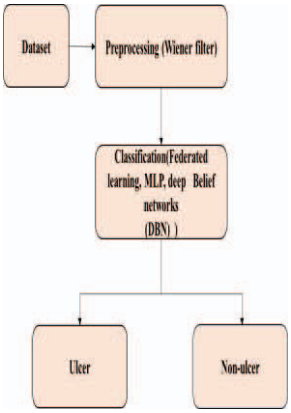


Fig. 1. Overall flow of proposed mouth disease prediction

3.1 Preprocessing

Wiener filter

The Wiener filter has two primary effects: it reverses the image's blurring and eliminates excess noise. It is especially useful for processing photos

ISBN 979-8-3503-5094-4

that have been blurry by a recognized lowpass filter or photographs that have been passed through a degrading filter. It is frequently employed in the algorithm-based deconvolution process, which improves signals from data.

The main objective is to keep the difference between the original and output picture as small as possible. This can be expressed as the average squared distance between the filtered output and a desired signal, or the total mean square error. It works using an equation like the least squares method.

For mean square error, Wiener filtering is the best option. Put otherwise, it reduces the total mean square error during the noise smoothing and inverse filtering processes. A linear estimate of the original picture is what the Wiener filter does. The method is predicated on a stochastic structure. According to the orthogonality principle, the Fourier domain Wiener filter may be written as follows:

$$W(f_1, f_2) = \frac{H^*(f_1,f_2)S_{xx}(f_1,f_2)}{|H(f_1,f_2)|^2 S_{xx}(f_1,f_2)+S_{\eta\eta}(f_1,f_2)} \quad (1)$$

Whereas,

$S_{xx}(f_1, f_2) - power\ spectra\ of\ original\ image,$

$S_{\eta\eta}(f_1, f_2) - Additive\ noise,$

$H(f_1, f_2) - blurring\ filter,$

We need to compute the additive noise and the power spectra of the original image before we can really apply the Wiener filter. White additive noise has a power spectrum that is equal to the noise variance. There are several techniques for estimating the original image's power spectrum. The periodogram estimate of the power spectrum derived from the observation is referred to as a direct estimate.

$$S_{yy}^{per} = \frac{1}{N^2}[Y(k,l)Y(k,l)^*] \quad (2)$$

The cascade implementation of the inverse filtering and smoothing is expressed below,

$$S_{xx} = \frac{S_{yy}-S_{\eta\eta}}{|H|^2} \quad (3)$$

The periodogram estimate may be used to directly estimate the power spectrum from the observation. With this approximation, inverse filtering and noise smoothing are implemented in a cascaded manner:

$$W = \frac{1}{H}\frac{S_{yy}^{per}-S_{\eta\eta}}{S_{yy}^{per}} \quad (4)$$

3.2 Classification

Federated Learning

There are several clients and one server that make up the FL (federal learning) structure. The distant cloud server is referred to as the "server" in this research, and edge network elements such edge nodes and edge devices are referred to as the "client." The fundamental tenet of FL is "the model moves, the data does not." To be more precise, the client gets the global shared model from the server, trains it using local datasets, and modifies the model's parameters. When the server and client connect, the server retrieves the most recent model parameters and transmits them to each client (also known as client download server model parameters). Following client training, the updated model parameters are uploaded to the server, where the server aggregates the client model parameters in some way, as the server receives the most recent model parameters and provides them to each client during each server-client communication (also known as client download server model parameters). Following client training, the modified model parameters are transferred to the server, which aggregates the client model parameters.

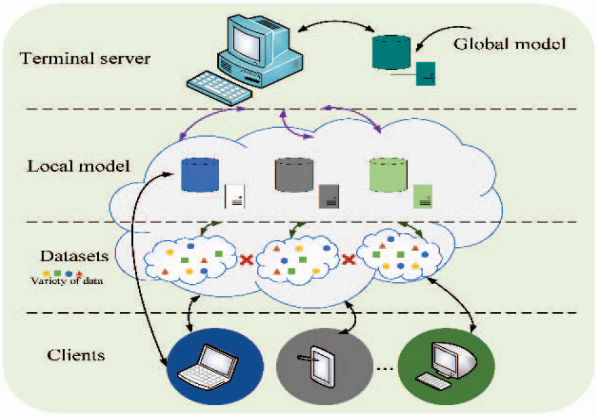


Fig. 2. Architecture of federated learning

Multilayer perceptron (MLP)

A feedforward neural network includes a multilayer perceptron (MLP) neural network. All of the nodes in this artificial neural network are linked to nodes in various levels.

The fundamental building block of an artificial neural network, the perceptron specifies the artificial neuron inside the network. It is a supervised learning method that computes the output using node weights, activation functions, values, and inputs.

The Multilayer Perceptron (MLP) Neural Network can only function in the forward direction. Each node is fully connected to the network. Each node only communicates its value to the next node

ISBN 979-8-3503-5094-4

in a forward manner. The MLP neural network uses backpropagation as a technique to increase training model accuracy. The fundamental structure of MLP, as seen in Figure 3.

The two hidden layer of multilayer perceptron is expressed in mathematical form,

$$f(x) = B_3 + W_3 f(B_2 + W_2 f(B_1 + W_{1x})) \quad (5)$$

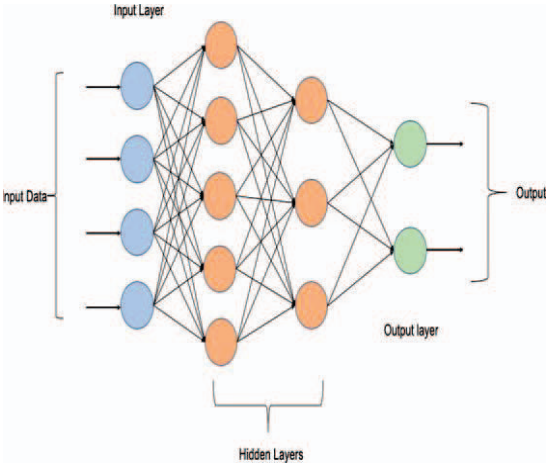


Fig. 3. Architecture of multilayer perceptron

Deep Belief Networks

An artificial neural network type called Deep Belief Networks (DBNs) is utilized for both supervised and unsupervised learning activities. They consist of many layers of shallow neural networks called Restricted Boltzmann Machines (RBMs), which may be learned through unsupervised learning. Up to the last layer of the network, the RBMs' output is utilized as the input for that layer. Typically, a classifier that has been trained using supervised learning makes up the last layer of the DBN.

DBNs are effective in a wide range of applications, including natural language processing, speech recognition, and image recognition. Additionally, they are renowned for their capacity to pick up hierarchical data representations, which is helpful in machine learning and artificial intelligence for handling challenging issues.

DBNs have gained a lot of traction in the deep learning field lately because of their strong scalability, capacity to handle high-dimensional data, and capacity to represent intricate, non-linear connections in the data.

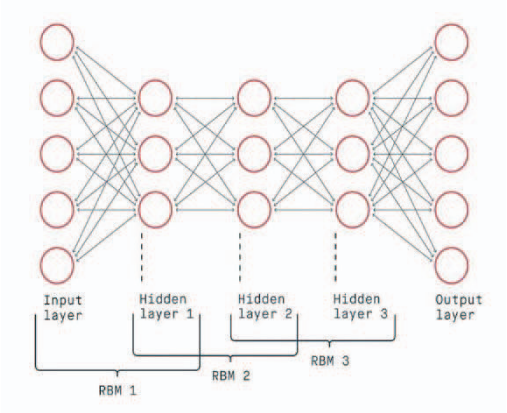


Fig. 4. Architecture of Deep Belief networks

4. Result and discussion

The mouth disease prediction using a deep learning techniques are applied in dataset processing of preprocessing, augmentation and classification steps involved to produce the result as compared. The mouth disease prediction implemented in python language and anaconda tool. The three algorithms are comparison in table 1.

Table 1: Comparison of FL, MLP, DBN

Algorithm	Accuracy	Precision	Recall	Result
Federated learning	0.50	0.70	0.50	50%
Multilayer perceptron	0.60	0.75	0.55	60%
Deep belief network	**0.85**	**0.80**	**0.60**	**85%**

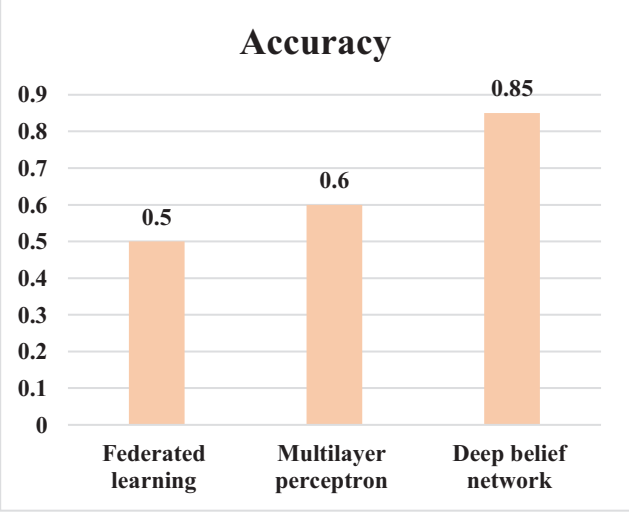


Fig. 5. Accuracy of the mouth prediction

ISBN 979-8-3503-5094-4

Mouth disease prediction using federated learning, multilayer perceptron (MLP) and deep belief network (DBN) produced the accuracy rates 0.50, 0.60 and 0.85% respectively. The comparison of these algorithm deep belief network has a more accuracy produced which shown in figure 5.

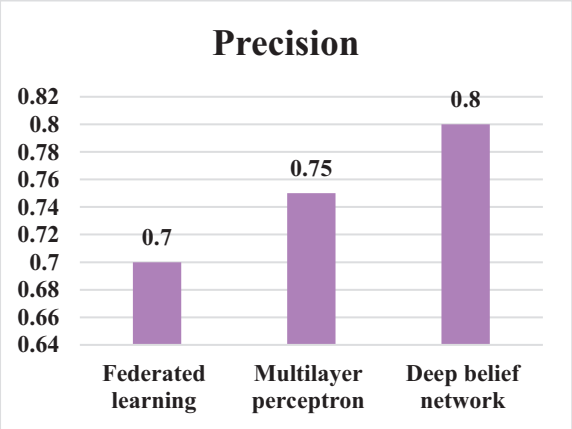


Fig. 6. Prediction of the mouth prediction

Mouth disease prediction using federated learning, multilayer perceptron (MLP) and deep belief network (DBN) produced the accuracy rates 0.70, 0.75 and 0.80% respectively. The comparison of these algorithm deep belief network has a more precision produced which shown in figure 6.

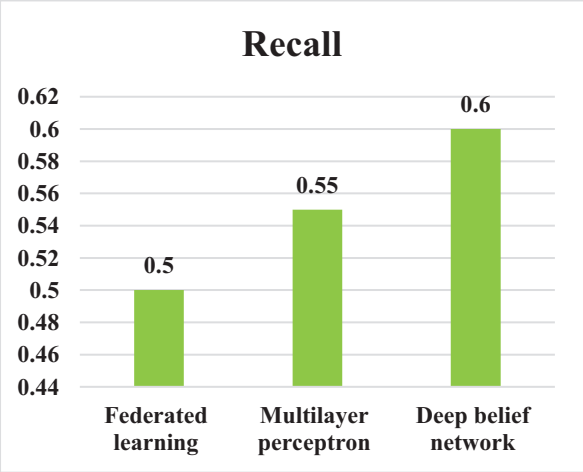


Fig. 7. Recall of the mouth prediction

Mouth disease prediction using federated learning, multilayer perceptron (MLP) and deep belief network (DBN) produced the accuracy rates 0.50, 0.55 and 0.00% respectively. The comparison of these algorithm deep belief network has a more recall score produced which shown in figure 7.

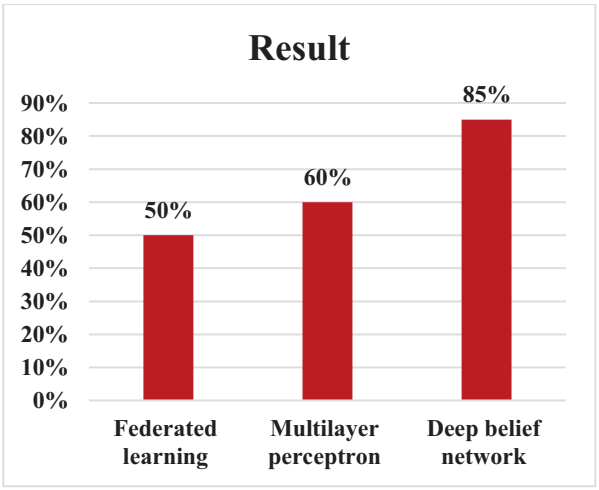


Fig. 8. Overall result prediction

Figure 8 shows in oval comparison result in mouth prediction in Deep belief networks than federated learning and multilayer perceptron.

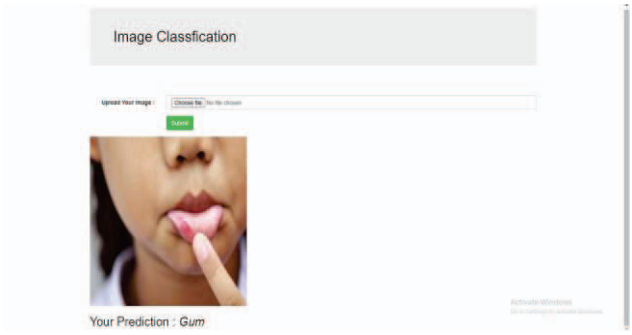


Fig. 9. Mouth disease prediction in flask app

Figure 9 shows the mouth disease prediction using federated learning, multilayer perceptron, and deep belief networks implemented in Python with the help of the Flask app.

5. Conclusion

Deep learning algorithms have proved their enormous promise in the identification and diagnosis of oral illnesses. The advancement of deep learning algorithms and the increasing availability of medical imaging data have substantially boosted the efficiency and accuracy of sickness identification. In terms of precision, recall, and accuracy, the federated learning, MLP, and deep belief networks all achieved the same results. The federated learning produced accuracy, precision, and recall of 0.50, 0.70, and 0.50 and 0.60, 0.75, and 0.55 for the MLP layer and 0.85, 0.80, and 0.60 for the deep belief network, respectively. When three algorithms are compared, deep belief networks offer the best results, with an accuracy of 85%. Nonetheless, deep learning looks to be a promising subject for future research and development due to its potential benefits in the detection of infectious illnesses.

ISBN 979-8-3503-5094-4

References

1. Javed Rashid, Bilal Shabbir Qaisar, Muhammad Faheem, Arslan Akram, Riaz, Amin and Muhammad Hamid, "Mouth and oral disease classification using InceptionResNetV2 method", Multimedia Tools and Applications, 2023.
2. Dr Thanh Ngoc Nguyen and Dr Dinh Ngoc Minh, "Applying Machine Learning to predict Hand-Foot-Mouth disease outbreaks in Vietnam", Journal of Health Informatics in Developing Countries, Vol. 15 No. 2, 2021.
3. Bin Zhang, Xiang Wan, Fu-sheng Ouyang4, Yu-hao Dong, De-hui Luo, Jing Liu, Long Liang, Wen-bo Chen6, Xiao-ning Luo, Xiao-kai Mo, Lu Zhang, Wen-hui Huang, Shu-fang Pei, Bao-liang Guo, Chang-hong Liang, Zhou-yang Lian1 & Shui-xing Zhang, "Machine Learning Algorithms for Risk Prediction of Severe Hand-Foot-Mouth Disease in Children", Scientific Reports, Vol. 7, No. 5368, 2017.
4. Guangjian Liu, Yi Xu, Xinming Wang, Xutian Zhuang, Huiying Liang, Yun Xi, Fangqin Lin, Liyan Pan, Taishan Zeng4, Huixian Li, Xiaojun Cao, Gansen Zhao & Huimin Xia, "Developing a Machine Learning System for Identification of Severe Hand, Foot, and Mouth Disease from Electronic Medical Record Data", Scientific Reports, Vol. 7: No. 16341, 2017.
5. Yongbin Wang, Chunjie Xu, Shengkui Zhang, Li Yang, Zhende Wang, Ying Zhu& Juxiang Yuan, "Development and evaluation of a deep learning approach for modeling seasonality and trends in hand-foot-mouth disease incidence in mainland China", Scientific Reports, Vol.9, No. 8046, 2019.
6. Suraj Verma, Mohammad Abdur Razzaque, Usanut Sangtongdee, Chonlameth Arpnikanondt, Boonrat Tassaneetrithep, And Alamgir Hossain, "Digital Diagnosis of Hand, Foot, and Mouth Disease Using Hybrid Deep Neural Networks", IEEE Access, Vol. 9, 2021.
7. Rasheed Omobolaji Alabi, Alhadi Almangush, Mohammed Elmusrati and Antti A. Mäkitie, "Deep Machine Learning for Oral Cancer: From Precise Diagnosis to Precision Medicine", Frontiers in Oral Health, Vol. 2,2022.
8. Wenzhe You, Aimin Hao, Shuai Li, Yong Wang and Bin Xia, "Deep learning-based dental plaque detection on primary teeth: a comparison with clinical assessments", BMC Oral Health, Vol. 20, No. 141, 2020.
9. Xialv Lin, Xiaofeng Wang, Yuhan Wang, Xuejie Du, Lizhu Jin, Ming Wan, Hui Ge and Xu Yang, "Optimized Neural Network Based on Genetic Algorithm to Construct Hand Foot-and-Mouth Disease Prediction and Early-Warning Model", International Journal of Environmental Research and Public Health, Vol.18, No. 2959, 2021.
10. Zhijin Wang, Yaohui Huang and Bingyan He, "Dual-grained representation for hand, foot, and mouth disease prediction within public health cyber-physical systems", Software: Practice and Experience, pp. 1-6, 2020.
11. Xiaochi Liu, Yilan Liao and Zhiyu Zhu, "Machine Learning Algorithms for Important Feature Evaluation and Prediction of Severe HandFoot-Mouth Disease in Hunan Province, China", Pacific Asia Conference on Information Systems (PACIS), 2019.
12. Gizem Tanriver, Merva Soluk Tekkesin and Onur Ergen, "Automated Detection and Classification of Oral Lesions Using Deep Learning to Detect Oral Potentially Malignant Disorders", Cancers, Vol. 13, No. 2766, 2021.
13. Ren Zhong, Yongsheng Wu, Yunpeng Cai, Ruxin Wang, Jing Zheng, Denan Lin, Hongyan Wu and Ye Li, "Forecasting hand, foot, and mouth disease in Shenzhen based on daily level clinical data and multiple environmental factors", BioScience Trends, Vol. 12, No. 5, pp. 450-455, 2018.
14. Dan Zhao, Morteza Homayounfar, Zhe Zhen, Mei-Zhen Wu, Shuk Yin Yu, Shuk Yin Yu, Kai-Hang Yiu, Varut Vardhanabhuti, Pelekos, Lijian Jin and Mohamad Koohi-Moghadam, "A Multimodal Deep Learning Approach to Predicting Systemic Diseases from Oral Conditions", Diagnostics, Vol.12, Issue. 12, 2022.
15. Anwar Alhazmi, Yaser Alhazmi, Ali Makrami, Amal Masmali, Nourah Salawi, Khulud Masmali and Shankargouda Patil, "Application of artificial intelligence and machine learning for prediction of oral cancer risk", Journal of oral pathology and medicine, 2021.
16. Soualihou Ngnamsie Njimbouom, Kwonwoo Lee and Jeong-Dong Kim, "MMDCP: Multi-Modal Dental Caries Prediction for Decision Support System Using Deep Learning", International Journal of Environmental Research and Public Health, Vol. 19, 2022.
17. Hung, M., Voss, M.W., Rosales, M.N.; Li, W., Su, W., Xu, J, Bounsanga, J, Ruiz-

ISBN 979-8-3503-5094-4

Negrón, B, Lauren, E, Licari, F.W. "Application of machine learning for diagnostic prediction of root caries", Gerodontology, Vol. 36, pp. 395–404, 2022.

18. Liu, L, Wu, W, Zhang, S.-Y, Zhang, K.-Q.; Li, J.; Liu, Y.; Yin, Z.-H, "Dental Caries Prediction Based on a Survey of the Oral Health Epidemiology among the Geriatric Residents of Liaoning, China", BioMed Research International, pp. 1–12, 2020.
19. Priyanka Jaiswal and Dr. Sunil Bhirud., "Classification and Prediction of Oral Diseases in Dentistry using an Insight from Panoramic Radiographs and Questionnaire", International Conference on Information Systems and Computer Networks (ISCON), pp. 22-23, 2021.
20. Viswanatha Reddy Allugunti, "A machine learning model for skin disease classification using convolution neural network", International Journal of Computing, Programming and Database Management, Vol. 3, No. 1, pp. 141-147, 2022.
21. G. Sudha, M. Birunda, R. Khowshalya, S. Preneshvar, S. Jayanthiswari, V. Sathya, N. Shneka, "Lung Cancer Image Segmentation and Detection Using Deep Learning Algorithms", Cardiometry, Issue 26, 2023.
22. G. Sudha, M. Birunda, J., Gnanasoundharam, J. Alphas Jeba Singh, "Deep Learning System in Curvelet Domain for Skin Cancer Diagnosis", International Conference on Electronics and Sustainable Communication Systems (ICESC), 2023.
23. Sharaf J. Malebary and Arshad Hashmi, "Automated Breast Mass Classification System Using Deep Learning and Ensemble Learning in Digital Mammogram", IEEE Access, Vol. 9, pp. 55312 – 55328, 2021.

ISBN 979-8-3503-5094-4

3D Fast Fourier Transform Analysis for the Detection of Alzheimer's Disease in MRI Images

Ms.T.R.Thamizhvani
Department of Biomedical Engineering
Vels Institute of Science, Technology and Advanced Studies
Chennai, India
thamizhvani.se@velsuniv.ac.in

Dr.R.J.Hemalatha
Department of Biomedical Engineering
Vels Institute of Science, Technology and Advanced Studies
Chennai, India
hemalatharj.se@velsuniv.ac.in

***Abstract*— Fast Fourier Transform (FFT) is the transform used for the determination of the frequency domain features and description of the signal or image. The frequency components of the images are analyzed using the FFT techniques in image processing. In the study, Three-dimensional FFT features are analyzed to diagnose Alzheimer's Disease (AD) at early stages. The medical images are obtained from the standardized database which is acquired using Magnetic resonance Imaging (MRI). MRI images of different AD stages are acquired. Fast Fourier Transform is applied to three-dimensional spatial images in MRI to analyze the frequency components in each domain. Different features like r-Value, Phi, Theta, Spectral coefficients, and Fourier coefficients. The significant feature difference between the two types of 3D MRI images is analyzed using statistical techniques. Thus, the FFT features possess a high efficiency in the detection of Alzheimer's Disease.**

Keywords—Fast Fourier Transform, three-dimension, Magnetic resonance imaging, Kruskal Wallis Test.

I. INTRODUCTION

Fast Fourier Transform (FFT) is a transform-based measuring method used in analyzing the spectral components that define the frequency data about the function in any form. FFT is applied in multiple fields like transformation analysis, fault analysis, optimization, feature extraction, and so on. In image processing, the main feature of the study is the frequency components. FFT helps in the transformation of the pixels in the spatial (x,y,z) domain to the frequency domain from which features are extracted for the study of the frequency variations in the images [1]. Frequency components derived are usually used in the filtering process for the removal of noises and sharpening of the images [2, 3]. Low frequency regards step-wise static variations and high frequency for the dynamic abnormal variations in the pixel values of the images. With advancements in recent studies, multidimensional Fast Fourier transform is used for the analysis of the frequency components in N-dimensional arrays such as 3D images.

The study involves three-dimensional images in MRI for different AD stages. Magnetic Resonance Imaging (MRI) is an imaging modality to acquire images of the human system based on proton activation and movements. The soft tissue parts of the body are visually analyzed more effectively using MR Imaging. There are two different modalities based on the magnetic Tesla. One is 1.5 Tesla and the other 3 Tesla. The commonly used MR Imaging modality is 1.5 Tesla. The images obtained from the database are of 1.5 Tesla type [4]. MR imaging is implemented in the diagnosis of various disease conditions specifically in the human brain. The cognitive changes, neuronal damage, and anatomical and physiological variations in the brain can studied in detail with this imaging modality.

Alzheimer's disease is a neuronal disorder that mainly affects the functioning of the brain cells. The decline in cognition and memory with reduced motor and metabolic functions are the symptoms of the disorder. There is no proper cure or treatment for the condition, early diagnosis can extend the lifespan of the individual [5]. Drug-based and rehabilitation-based treatments are provided for the individuals diagnosed with the disorder. The significant biomarker for AD identification is the hippocampus region. Diagnosis at early stages is required for the treatment of Alzheimer's Disease [6].

Takaaki et.al., (2022) describe the Fast Fourier Demodulation to define the fluid gradient fields in the images. Background-oriented Schlieren technique defines the fluid fields in the images and illustrates the maximum displacement gradient based on FFT demodulation [7]. Smith et.al., (2023) describe the recursive signal processing using Short time FFT for processing of the acoustic signals in the industrial processes. Machine learning and predictive data analytic approaches will benefit from the intuitive interpretation skills provided by the information and signatures obtained through recursive STFFT processing [8]. Miltiadous et.al., (2023) describe the significance of the network designed to capture complex features based on the Fourier transform. These are applied to EEG signals for the classification of AD. This strategy could increase the precision of early diagnosis and result in the creation of more potent AD therapies [9].

In the study, images acquired in MRI are collected from the standardized database, ADNI which possesses images of different modalities, only for Alzheimer's Disease types. Alzheimer's Disease Neuroimaging Initiative (ADNI) database is a standardized database with all types of data related to stages of AD. 3D MRI images of AD in different stages are considered for the study. A median filter and eight histogram clipping are the two specific 3D Preprocessing techniques implemented in the study. FFT is applied to the 3D images to extract the frequency components for the detection of AD. The significant difference between the images is analyzed using the FFT features. Statistical techniques are implemented to determine the significant difference in parameters. Thus, the high efficiency of FFT features in the detection of AD is determined.

II. METHODOLOGY

In the proposed methodology, 3D FFT analysis is performed to determine the AD stages. Three dimensional, MRI images of different AD stages are determined from a definite database, ADNI. Fast Fourier Transform is applied to convert the spatial components to frequency components and features are extracted from the FFT images for the

ISBN 979-8-3503-5094-4

analysis of the AD stages. Preprocessing is performed using a median filter and eight histogram clipping [10]. The efficiency and significant differences of the FFT features are analyzed using different tests in Statistics like the Kruskal Wallis Test. The FFT features possess a high efficiency in the detection of the AD stages.

A. Dataset Collection

In the proposed study, three-dimensional MRI images were analyzed and collected from Alzheimer's Disease Neuroimaging Initiative (ADNI), a standardized database (http://adni.loni.usc.edu). The ADNI database, which has over 2000 images from various modalities like PET, MRI, DTI, and others is collected by the Laboratory of Neuroimaging (LONI) at the University of Southern California. Images showing the neurological condition at several phases, including Alzheimer's disease (AD), early mild cognitive impairment (EMCI), late mild cognitive impairment (LMCI), mild cognitive impairment (MCI), and normal cognitive function (CN), are included in the collection. The library contains different MPRAGE, and Sectional MRI images of the brain conducted with 1.5 and 3 Tesla for various age groups, most notably those over 45 [11]. In the study, images of AD stages are obtained

B. FFT Features

Fast Fourier Transform is a method used for the analysis of the frequency components of the data. In the study, 3D MRI images are used for the detection of AD. Multidimensional-n-factor FFT is applied for the 3D images for the determination of the frequency components. Fast Fourier transform converts the spatial x, y, and z dimensional voxels in the 3D images to frequency components-based images. With the help of the FFT images, features are extracted for the analysis and detection of AD stages like CN, EMCI, LMCI, and Severe AD. The frequency components involve the magnitude, wavelength, and phase of the 3D images. The FFT features derived from the images are r-value, Phi, Theta, Spectral coefficients, and FFT Coefficients (magnitude) [12]. These features are analyzed for the diagnosis of AD stages.

R-value and Theta – The phase values of the images are based on the angular coordinates of the FFT images. The phase coordinates with respect to the frequency on the 3D images are r-Value, Phi (ϕ), and Theta (ɵ) [13].

- r-Value defines the distance between the voxels interconnected with the frequency components in the 3D FFT images.
- Phi (ϕ) refers to the angle defining the spatial xy plane of the FFT images
- Theta (ɵ) defines the angle of the z-axis in the spatial domain stating the distribution of the frequency in the FFT images.

FFT Coefficients – The coefficients of FFT describe the magnitude of the FFT images. The magnitude defines the strength of the different frequency distributions in the FFT images. The FFT magnitude of the 3D images is determined with the N-FFT method based on the frequency distribution among the voxels [14, 15].

Spectral Coefficients - The power spectral density (PSD) of the images of FFT is determined using the Welch Method. The magnitude of the power spectrum is obtained based on the voxel points of maximum power distribution with the help of the frequency components [16, 17]. The power spectral density of the 3D FFT images and the magnitude of the spectral density are analyzed. The PSD-based spectral coefficients are extracted from the 3D images. Thus, the FFT features are extracted from the MRI – 3D images of different AD stages.

C. Statistical Analysis

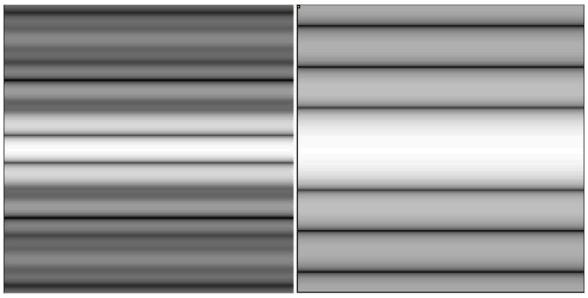
In the proposed methodology, the statistical analysis of the FFT features extracted is performed to determine the efficiency of the features in differentiating the AD stages [18]. The statistical technique implemented is the Kruskal-Wallis Test which defines the significant difference between the independent dataset. The FFT feature's significant nature in defining the AD stages is analyzed using the statistical test. With the results, the study states that the FFT features possess high efficiency in determining the AD stages such as CN, EMCI, LMCI, and Severe AD.

III. Results and Discussion

In the study, 3D FFT features are extracted from the images of AD stages. The 3D MRI images are derived from the database called ADNI. FFT is applied to the 3D images and features like -value, Phi, Theta, Spectral coefficients, and FFT Coefficients are extracted. Statistical techniques are implemented to analyze the efficiency of the FFT features in the detection of the stages in AD. The 3D MRI images of ADNI database are illustrated below in figure 1.

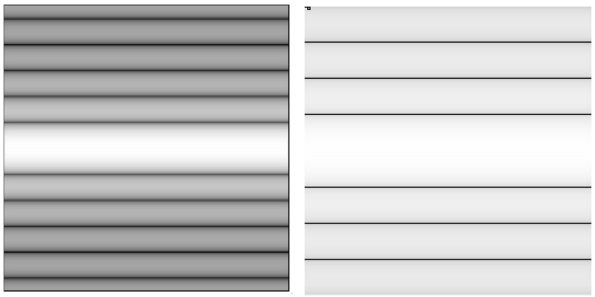
Fig. 1. 3D MRI Images of ADNI database

The images of the 3D MRI brain of AD stages from the ADNI database are processed in the standard Montreal Neurological Institute-based imaging atlas for effective analysis. Preprocessing is applied using a median filter and eight histogram clipping. Fast Fourier transform (FFT) in multidimensional form with three dimensions specified are applied to the pre-processed images of AD stages in three dimensions. The frequency component distribution is determined among the AD stages. The FFT images of the various AD stages are illustrated in Figure 2.



CN FFT IMAGE EMCI FFT IMAGE

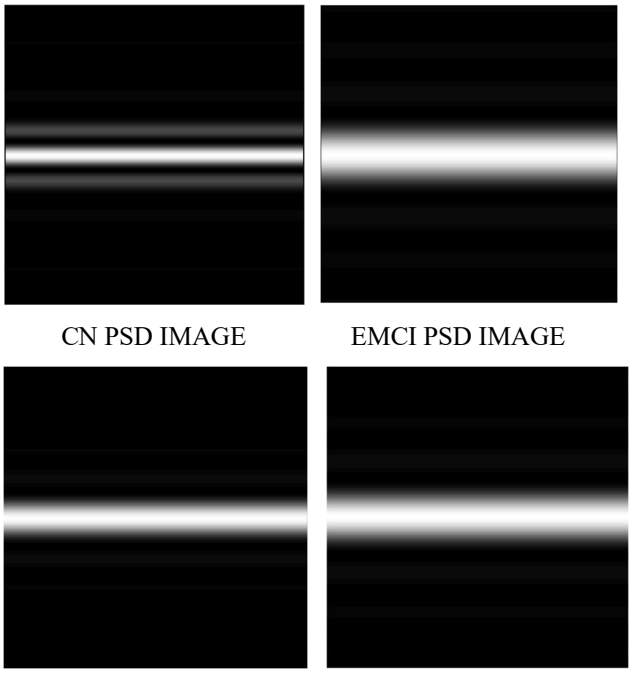
ISBN 979-8-3503-5094-4



LMCI FFT IMAGE AD FFT IMAGE

Fig. 2. FFT Images of AD Stages

The FFT features are extracted from the FFT images of AD stages. The features are based on the Phase values of the frequency distribution in the images. The r-Value, Phi (ϕ), and Theta (ɵ) are the phase-based features extracted from the angular response of the FFT images. The FFT coefficients define the magnitude of the strength of frequency distributions. The Spectral coefficients are determined after applying power spectral density, the Power spectrum of the frequency distribution in the FFT images based on voxels are defined. The Power spectral density (PSD) images of the AD stages are defined in figure 3.



CN PSD IMAGE EMCI PSD IMAGE

LMCI PSD IMAGE AD PSD IMAGE

Fig. 3. PSD Images of AD stages

The FFT features are extracted and mean values are computed and listed in Table 1.

TABLE I. MEAN VALUES OF THE FFT FEATURES EXTRACTED FROM THE AD STAGES

3D FFT Features	CN	EMCI	LMCI	AD
r-Value	4.1	6	7.46	3.88
Theta	50.19	140	134.43	75.96
Phi	45.1126	45.1144	45.4658	45
Fourier coefficients	112	168	224	154
Spectral coefficients	17.081	14.794	15.824	10.35

The FFT features of AD stages are graphically illustrated below in Figure 4. Kruskal Wallis H test is used to define the significant difference among the different features in the diagnosis of the AD stages. Each FFT feature for stages in AD is analyzed using the Kruskal Wallis Test.

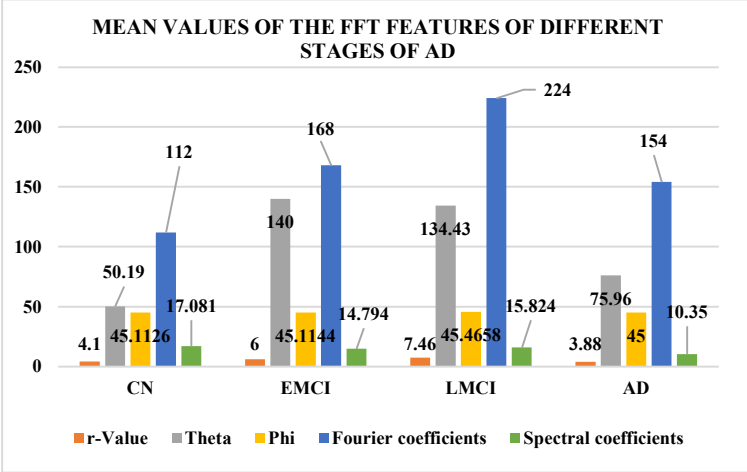


Fig. 4. Graphical Representation of mean Values of the FFT features of Stages of AD

The statistical analysis defines that these 3D FFT features are more significant in determining the different stages of AD. The results of the Kruskal Wallis test are given below in Table 2.

TABLE II. RESULTS OF KRUSKAL WALLIS TEST AMONG THE AD STAGES

Kruskal Wallis Test	r-Value	Theta	Phi	Fourier coefficients	Spectral coefficients
H statistic	32.837	30.902	31.612	30.912	33.837
P-value	<0.00001	<0.00001	<0.00001	<0.00001	<0.00001

The P-value is determined to be less than 0.00001 which states that the features are significantly different between the CN, EMCI, LMCI, and severe AD which are defined to be AD stages. Thus, the comparison of results states that FFT features are significant for the analysis of the various AD stages with 99% efficiency. The diagnosis of AD stages can be performed with the help of the 3D FFT features.

IV. CONCLUSION

Fast Fourier transform is a method used to convert spatial variations into frequency variations/components of distribution. In the study, the analysis of AD stages is studied with the help of FFT and the features derived from the FFT images in 3D [19]. Alzheimer's disease (AD) described as a is neuronal dysfunction that affects brain function and causes a decline in cognition and other daily activities. Alzheimer's disease is categorized into various stages namely CN, EMCI, LMCI, and severe AD. MRI images in 3D are obtained from the standardized ADNI database. Fast Fourier transform is applied to the 3D MRI images of the AD stages and 3D FFT features like r-value, Phi, Theta, Spectral coefficients, and FFT Coefficients are extracted from FFT and PSD images of the MRI images based on the voxel-frequency distribution. The significance of the FFT features in detecting the stages of AD is derived by using a statistical method, the Kruskal Wallis

ISBN 979-8-3503-5094-4

H test. With the results of comparison, the FFT features are highly efficient in defining the stages in AD.

REFERENCES

[1] X. Zhao et al., "Fractional Fourier Image Transformer for Multimodal Remote Sensing Data Classification," in IEEE Transactions on Neural Networks and Learning Systems, 2022, doi: 10.1109/TNNLS.2022.3189994.

[2] Elias Rajaby, Sayed Masoud Sayedi, A structured review of sparse fast Fourier transform algorithms, Digital Signal Processing, 123, 2022, 103403, https://doi.org/10.1016/j.dsp.2022.103403.

[3] J. A. Smith and V. Agarwal, "Recursive Use of the Short-Time Fast Fourier Transform for Signature Analysis in Continuous Processes," in *IEEE Transactions on Instrumentation and Measurement*, vol. 72, pp. 1-11, 2023, Art no. 6504811, doi: 10.1109/TIM.2023.3305659.

[4] Rémi Cuingnet, Emilie Gerardin, Jérôme Tessieras, Guillaume Auzias, Stéphane Lehéricy, Marie-Odile Habert, Marie Chupin, Habib Benali, Olivier Colliot, Automatic classification of patients with Alzheimer's disease from structural MRI: A comparison of ten methods using the ADNI database, NeuroImage,Volume 56, Issue 2, 2011,pp. 766-781. https://doi.org/10.1016/j.neuroimage.2010.06.013.

[5] G. Gkenios, K. Latsiou, K. Diamantaras, I. Chouvarda and M. Tsolaki, "Diagnosis of Alzheimer's disease and Mild Cognitive Impairment using EEG and Recurrent Neural Networks," 2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), Glasgow, Scotland, United Kingdom, 2022, pp. 3179-3182, doi: 10.1109/EMBC48229.2022.9871302.

[6] Güntekin, B., Erdal, F., Bölükbaş, B. et al. Alterations of resting-state Gamma frequency characteristics in aging and Alzheimer's disease. Cogn Neurodyn 17, 829–844 (2023). https://doi.org/10.1007/s11571-022-09873-4

[7] Takaaki Shimazaki, Sayaka Ichihara, Yoshiyuki Tagawa, Background-oriented schlieren technique with fast Fourier demodulation for measuring large density-gradient fields of fluids, Experimental Thermal and Fluid Science, 134, 2022, 110598. https://doi.org/10.1016/j.expthermflusci.2022.110598.

[8] J. A. Smith and V. Agarwal, "Recursive Use of the Short-Time Fast Fourier Transform for Signature Analysis in Continuous Processes," in *IEEE Transactions on Instrumentation and Measurement*, vol. 72, pp. 1-11, 2023, Art no. 6504811, doi: 10.1109/TIM.2023.3305659.

[9] A. Miltiadous, E. Gionanidis, K. D. Tzimourta, N. Giannakeas and A. T. Tzallas, "DICE-Net: A Novel Convolution-Transformer Architecture for Alzheimer Detection in EEG Signals," in *IEEE Access*, vol. 11, pp. 71840-71858, 2023, doi: 10.1109/ACCESS.2023.3294618.

[10] TR Thamizhvani, Syed Uzma Farheen, RJ Hemalatha, and A Josephin Arockia Dhivya, "Classification of progressive stages of Alzheimer's disease in MRI hippocampal region", *Biomedical Engineering: Applications Basis and Communications*, vol. 32, no. 6, pp. 2050050, 2020.

[11] T. R., T., & R. J., H. (2023). 3D Pre-Processing Algorithm for MRI Images of Different Stages of AD. International Journal of Online and Biomedical Engineering (iJOE), 19(15), pp. 127–143. https://doi.org/10.3991/ijoe.v19i15.40191.

[12] F. Onimus, L. Gélébart, R. Brenner, Polycrystalline simulations of in-reactor deformation of recrystallized Zircaloy-4 tubes: Fast Fourier Transform computations and mean-field self-consistent model, International Journal of Plasticity,153, 2022, 103272. https://doi.org/10.1016/j.ijplas.2022.103272

[13] G. Gkenios, K. Latsiou, K. Diamantaras, I. Chouvarda and M. Tsolaki, "Diagnosis of Alzheimer's disease and Mild Cognitive Impairment using EEG and Recurrent Neural Networks," 2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), Glasgow, Scotland, United Kingdom, 2022, pp. 3179-3182, doi: 10.1109/EMBC48229.2022.9871302.

[14] Delimayanti, Mera Kartika, Bedy Purnama, Ngoc Giang Nguyen, Mohammad Reza Faisal, Kunti Robiatul Mahmudah, Fatma Indriani, Mamoru Kubo, and Kenji Satou. 2020. "Classification of Brainwaves for Sleep Stages by High-Dimensional FFT Features from EEG Signals" *Applied Sciences* 10, no. 5: 1797. https://doi.org/10.3390/app10051797.

[15] M. Heidari, S. Mirniaharikandehei, W. Liu, A. B. Hollingsworth, H. Liu and B. Zheng, "Development and Assessment of a New Global Mammographic Image Feature Analysis Scheme to Predict Likelihood of Malignant Cases," in *IEEE Transactions on Medical Imaging*, vol. 39, no. 4, pp. 1235-1244, April 2020, doi: 10.1109/TMI.2019.2946490.

[16] Hwang, S., Hong, K., Son, G. et al. Learning CNN features from DE features for EEG-based emotion recognition. Pattern Anal Applic 23, 1323–1335 (2020). https://doi.org/10.1007/s10044-019-00860-w

[17] Cascales-Fulgencio, David, Eduardo Quiles-Cucarella, and Emilio García-Moreno. 2022. "Computation and Statistical Analysis of Bearings' Time- and Frequency-Domain Features Enhanced Using Cepstrum Pre-Whitening: A ML- and DL-Based Classification" *Applied Sciences* 12, no. 21: 10882. https://doi.org/10.3390/app122110882

[18] Jamil, M.A., Khanam, S. Influence of One-Way ANOVA and Kruskal–Wallis Based Feature Ranking on the Performance of ML Classifiers for Bearing Fault Diagnosis. J. Vib. Eng. Technol. (2023). https://doi.org/10.1007/s42417-023-01036-x

[19] H. Buijs, A. Pomerleau, M. Fournier and W. Tam, "Implementation of a fast Fourier transform (FFT) for image processing applications," in *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 22, no. 6, pp. 420-424, December 1974, doi: 10.1109/TASSP.1974.1162620

ISBN 979-8-3503-5094-4

Multimodal Analysis of Bone Mineral Density Classification

K.Sneha
Department of Electronics and Communication Engineering
Madras Institute of Technology
Chennai, India
snehaks2k2@gmail.com

Dr.C.Sridevi
Department of Electronics and Communication Engineering
Madras Institute of Technology
Chennai, India
sridevijegan@annauniv.edu

B. Yazhini
Department of Electronics and Communication Engineering
Madras Institute of Technology
Chennai, India
yazhinib3@gmail.com

Amridha Lakshmi Venkkautaesh
Department of Electronics and Communication Engineering
Madras Institute of Technology
Chennai, India
amridha.lakshmi@gmail.com

Abstract*— Osteoporosis is a skeletal disorder marked by diminished bone mineral density and mass, impacting bone health, mass, or alterations in bone structure and strength. This decline in bone strength heightens the susceptibility to fractures. Notably, osteoporosis is often asymptomatic, making it a "silent" ailment. The assessment of deficiencies in joint images is a prolonged procedure conducted manually, demanding the expertise of professionals to periodically analyze the scans. This manual approach results in time delays and heightened costs. To address this issue efficiently, this work aims to pinpoint bone mineral density, a crucial factor in identifying bone-related diseases and assessing the risk of fractures in an automated fashion by comparing Dual Energy X-ray Absorptiometry and X-ray images. The workflow is initiated with image pre-processing followed by Image enhancement, Data Augmentation, Data partition, feature extraction and classification. The images are enhanced using a median filter, normalization techniques and CLAHE - Contrast Limited adaptive histogram equalization algorithm. VGG-16, ResNet, DenseNet, InceptionV3 and Inception-ResNet are the different transfer learning models employed in this work to analyze and extract the best algorithm that provides the desired results.

Keywords*— bone mineral density classification, CLAHE, CNN, VGG-16, ResNet, InceptionV3, Inception-ResNet, DenseNet

I. INTRODUCTION

Image processing is an extensively utilized field with applications across various domains. Biomedical imaging represents an ongoing area of research within image processing, primarily employed for disease detection. The current study falls within the realm of bone imaging, aiming to identify the presence of osteoporosis.

Osteoporosis is a metabolic bone condition distinguished by an imbalance between osteoclastic bone resorption and osteoblastic bone synthesis at the cellular level. This imbalance leads to the bones becoming fragile, increasing the susceptibility to fractures. Measuring the bone mineral density (BMD) will help in diagnosing osteoporosis and hence provide better diagnostic accuracy and facilitate better treatment protocols.

While osteoporosis can manifest in individuals of any age or gender, it predominantly emerges as an age-related condition that disproportionately impacts women compared to men. With current advancements in medicine and technology, measurement of bone mineral density done using DEXA imaging is the most reliable method used for the diagnosis of osteoporosis.

This paper involves using X-ray and DEXA scan images for training where X-ray images being cost-efficient and a quicker process also result in long-term risks owing to prolonged exposure to radiation. It also lacks detailing because the X-ray provides a flat image. DEXA scan provides a reliable and precise interpretation of the condition despite being expensive. This research aims to develop a method for predicting osteoporosis utilizing deep learning techniques. To ensure thorough analysis, the methodology includes strategies for noise reduction, image enhancement, data augmentation, and data partitioning. A comparative study is undertaken by training, testing, and validating datasets using various Convolutional Neural Network (CNN) architectures, including VGG-16, ResNet50, Inception V3, DenseNet, and Inception-ResNet.

II. RELATED WORKS

Many researchers have worked extensively in predicting bone density using various images such as Computer Tomography (CT), X-ray, Dual Energy X-ray Absorptiometry (DEXA) and even radiographs. In [1] Multi-View Computer Tomography Network (MVCTNet) was developed and designed that aims to improve automatic osteoporosis classification. Image enhancement techniques should be given special care as the filters used also play a vital role, like the application of diffusion filters in enhancement techniques is found to decrease the efficiency of the model. In [2] various other techniques such as the U-Net architecture using residual block have also been used to improve the efficiency to 87% [3]. Simple regression analysis is additionally employed to ascertain the correlation between age and BMD where IMPAX CD-Viewer are used to obtain the images [6]. With an aim of increasing the efficiency transfer learning models such as ResNet50, VGG-19, and GooglNet are employed in the classification of Knee Radiograph and have obtained 90% accuracy [7]. Despite achieving a 92% efficiency using the Inception –V3 model it is found that at times errors can be caused due to the intrusion of the surrounding soft tissues.[8]. A comparative study using

ISBN 979-8-3503-5094-4

several Machine Learning approaches has significantly stated that the CatBoost algorithm provides the best results with good efficiency when provided with categorical data and also provides room for further classification like osteopenia, and osteoporosis [8].

III. PROPOSED METHODOLOGY

The proposed methodology primarily possesses four key steps:

- Data pre-processing
- Data augmentation
- Data partition
- Implementing various CNN models to obtain the best algorithm

A. *Data preprocessing*

The images are normalized to standardize the pixel values in an image. The median filter is employed to effectively remove salt-and-pepper noise and Gaussian noise from images. It replaces pixel values with the median value in the surrounding neighborhood, preserving edges and fine details.

Contrast Limited Adaptive Histogram Equalization (CLAHE) is used to enhance the contrast and improve the visual quality of images by enhancing the visibility of subtle structures and abnormalities, making it easier for medical professionals to diagnose and analyze medical images.

B. *Data Augmentation*

Since the proposed system employs a deep learning model for extracting bone regions from images, a significant number of training samples are required. To address this, image augmentation is implemented before data partition, incorporating cropping, rotation, flip, and shear augmentation to augment the number of image samples.

C. *Data Partition*

The acquired images are partitioned as training and testing in the ratio of 80:20.

D. *Implementing various CNN models*

1) ResNet-50:

ResNet50, a variation of the ResNet model, incorporates 48 Convolution layers, accompanied by one Maxpool (used in CNNs for down sampling or subsampling the feature maps obtained after convolutional operations) and one Average Pool layer (similar to max pooling, it operates on small rectangular regions within the input feature maps; instead of outputting the maximum value from each region, the average pooling layer computes the average value.). The selection of this algorithm is motivated by the ability of the ResNet algorithm to address the vanishing gradient problem through the implementation of ‘skip connections. Despite the fact that the hop connections make the system more complex, it has an edge over others by enabling faster training of each layer.

2) Inception V3:

InceptionV3 is a convolutional neural network designed to aid in image analysis and object detection. It is characterized by its 48-layer depth as a convolutional neural network. It possesses various advantages including higher efficiency, a deeper network compared to its older versions, and computationally less expensive. Inception v3 innovatively incorporated 'label smoothing' to regularize the network's output, enhancing its ability to generalize and improve overall performance.

3) VGG-16:

Visual Geometry Group-16 (VGG16) is a convolution neural network that is 16 layers deep and was designed for image classification. Its main characteristics include simplicity and uniformity in architecture with thirteen convolutional layers and three fully connected layers. It has approximately 138 trainable parameters. Rectified Linear Unit (ReLU) activation functions are used throughout the network.

Despite providing many advantages VGG16 is a deep and complex model, which requires significant computational resources and time for training.

4) DenseNet-121:

Densely Connected Convolutional Networks is a CNN architecture proposed to address challenges related to information flow and parameter efficiency in deep networks.

Unlike traditional architectures where layers are connected sequentially, DenseNet establishes direct connections between all layers, enabling seamless information flow and promoting feature reuse. This addresses the vanishing gradient issue and improves gradient propagation during training. This design is predominantly employed in computer vision applications, encompassing tasks such as image classification, object detection, and semantic segmentation.

5) Inception-ResNet:

Inception-ResNet is a convolutional neural network design that extends the Inception series of architectures by integrating residual connections. This hybrid algorithm can perform better classification results than any other conventional technique. It also achieves better recognition accuracy on poor-quality text images.

IV. RESULTS AND DISCUSSIONS

The pre-processing techniques in this paper involve normalization, and the use of a median filter for de-noising followed by image enhancement techniques. The output of the normalization and median filter technique are shown in Fig. 1 and Fig. 2.

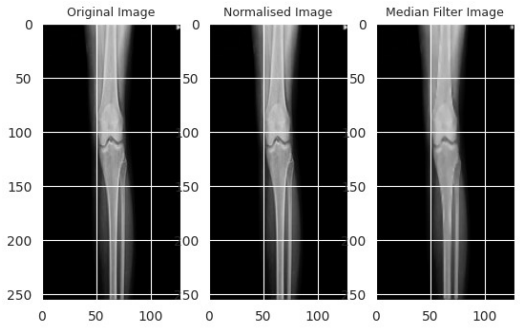


Fig. 1: X-Ray output image of normalized and median filter

ISBN 979-8-3503-5094-4

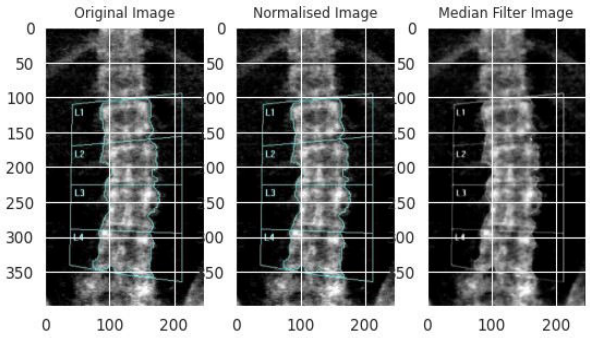


Fig. 2: DEXA output image of normalized and median filter

The approach utilized for enhancing images involves Contrast Limited Adaptive Histogram Equalization (CLAHE) which enhances the visibility of subtle structures and abnormalities, making it easier for medical professionals to diagnose and analyze medical images as shown in Fig. 3 and Fig. 4

To quantify and assess the quality of a processed or compressed image compared to the original, reference image Peak Signal to Noise Ratio (PSNR) value has been computed and tabulated. PSNR is expressed in Decibels (dB) and is computed using the following formula:

$$mse = mean\left((img_{original} - img_{enhanced})^2\right) \quad (1)$$

$$PSNR = 20 * log_{10}V_m - 10 * log_{10}mse \quad (2)$$

The PSNR values obtained are recorded and tabulated in the Table.1 below for CLAHE and Median filter.

TABLE 1: Comparing PSNR values for different conditions

Method	Median Filter		CLAHE Enhancement	
	Normal	Osteoporotic	Normal	Osteoporotic
X-ray image	31.19	31.93	27.87	28.21
DEXA images	20.90	21.47	27.82	27.76

A confusion matrix represents the prediction summary in matrix form and is given below to show the performance of the classification model used. It provides a comprehensive view of model's performance, indicating how well it predicts and where it makes errors. The performance of a classification algorithm can be visually depicted by a confusion matrix. Further, metrics such as accuracy, precision, recall (also known as sensitivity), specificity and F1 score.

The Confusion Matrix of DEXA images for the Inception-ResNet model are shown in Fig. 5 and Fig. 6 depicts the Confusion Matrix of X-Ray images for VGG-16 model.

The resulting images after pre-processing using CLAHE have been given below.

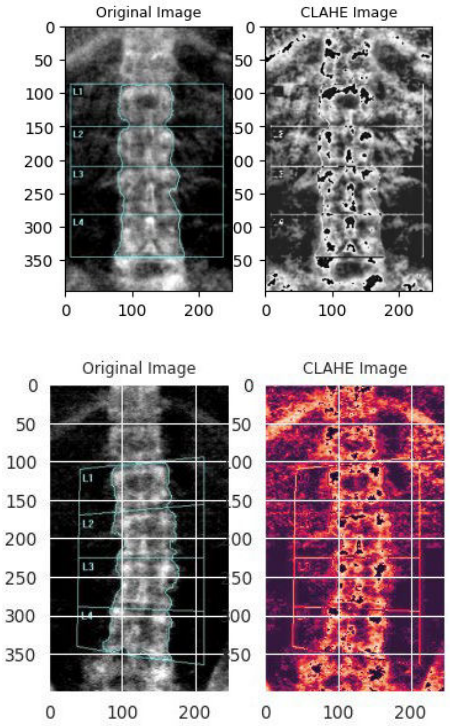


Fig. 3: CLAHE output of DEXA images

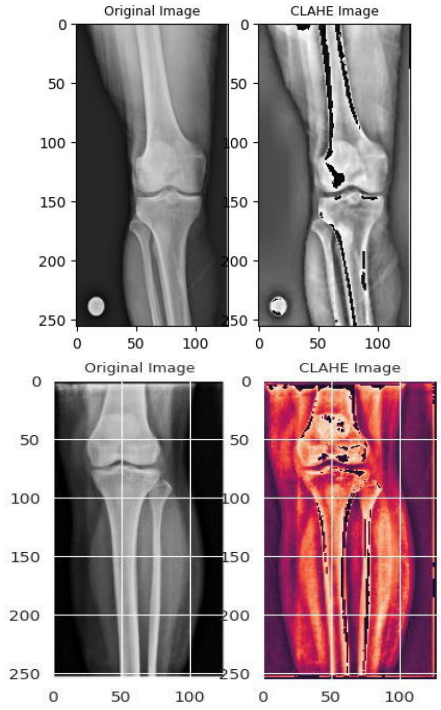


Fig. 4 : CLAHE output of X-Ray images

These metrics are calculated as given below:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (3)$$

$$Precision = \frac{TP}{TP+FP} \quad (4)$$

$$Recall\ (Sensivity) = \frac{TP}{TP+FN} \quad (5)$$

ISBN 979-8-3503-5094-4

$$Specificity = \frac{TN}{TN+FP} \quad (6)$$

$$F1\ Score = \frac{2*Precision*Recall}{Precision+Recall} \quad (7)$$

$$= \frac{TP}{TP+\frac{1}{2}(FP+FN)}$$

where,

TP- True Positive
TN- True Negative
FP- False Positive
FN- False Negative

Below is the confusion matrix corresponding to the highest accuracy model for DEXA and X-ray images.

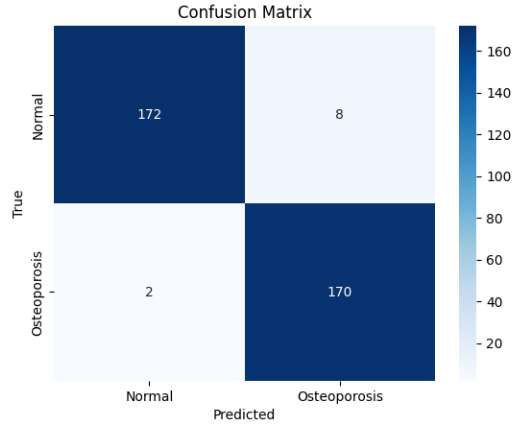


Fig. 5: Confusion Matrix of DEXA images for Inception-ResNet model

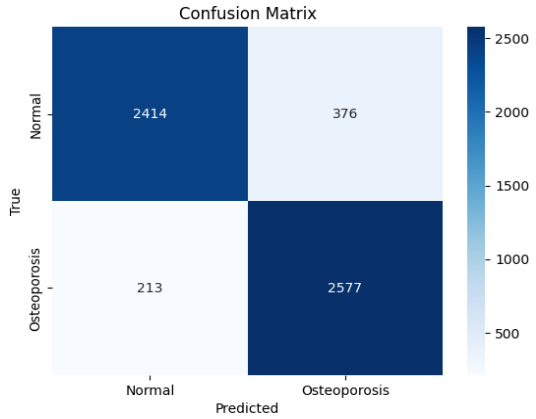


Fig. 6: Confusion Matrix of X-Ray images for VGG-16 model

TABLE 2: Tabulation of DEXA Image Results for different models

	ResNet-50 (%)	Inception V3 (%)	VGG-16 (%)	Dense Net (%)	Inception-ResNet (%)
Accuracy	92.32	90.6	94.9	90.9	97.2
Precision	90.5	91.1	92.2	98.6	95.5
Specificity	90.56	91.6	92.0	98.89	95.56
Sensitivity	94.2	89.5	97.9	82.56	98.84
F-score	92.3	90.3	95.0	89.9	97.1

TABLE 3: Tabulation of X-Ray Image Results for different models

	ResNet-50 (%)	Inception V3 (%)	VGG-16 (%)	Dense net (%)	Inception-ResNet (%)
Accuracy	77.74	83.3	89.4	85	82.3
Precision	84.1	81.7	87.3	80.4	76.9
Specificity	87.03	85.9	86.5	77.48	72.25
Sensitivity	68.47	89.53	92.4	92.61	92.43
F-score	75.5	83.8	89.7	86.1	84

X-ray methodologies offer a quantitative assessment of bone mineral density (BMD). However, recent studies underscore that fracture risk is influenced by a spectrum of factors extending beyond BMD alone. These factors include the structural integrity and mechanical properties of bone, which are not adequately captured through densitometric techniques.

V. CONCLUSION

This paper endeavors to identify whether an individual is impacted by osteoporosis or possesses healthy bones through an examination and evaluation of their bone mineral density. The datasets are subjected to preprocessing with a median filter and then refined using Contrast Limited Adaptive Histogram Equalization (CLAHE). In the realm of medical imaging, particularly in contexts like DEXA where accurate measurement of bone mineral density is vital, precision in measurements becomes imperative. The PSNR value functions as a gauge, reflecting the extent to which preprocessing methods maintain the precision and accuracy of quantitative measurements, including those related to bone density. So, computation of the PSNR value for both X-ray and DEXA images are done which results in the PSNR value being high for X-ray images when median filter is used. For DEXA images the CLAHE enhancement technique holds well with a PSNR value of 27.8%.

With the help of different neural network models, various parameters are computed and it is observed that among the five algorithms, the Inception-ResNet hybrid model has the highest accuracy of 97.2%. Also, it is observed that the accuracy is high when DEXA images are trained as compared to X-ray images.

VI. REFERENCES

[1] Dong Hwan Hwang, So Hyeon Bak Tae-Jun Ha, Yoon Kim, Woo Jin Kim, And Hyun-Soo Choi "Multi-View Computed Tomography Network for Osteoporosis Classification", Engineering in Medicine and Biology Society Section, 2023.

[2] D. Arulselvam, T. Sathies Kumar, T. Sheela, S. Premalatha, K. Srividya, S.Nandhini: "Analysis of Mineral Density in Bone Using Deep Learning and Smart Tracking System", IC3IoT, 2022.

[3] Pooja S Dodamani, Ajit Danti: "Diagnosis of Osteoporosis from X-ray Images using Automated Techniques", International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COM-IT-CON), 2022.

[4] Manisha G. Fal Dessai, Dr. Venkatanareshbabu Kuppili "Bone Health Prediction using Machine Learning Approach", ICAISS, 2022.

[5] S.M.Nazia Fathima, R.Tamilselvi, M.Parisa Beham, A.Nagaraj, "A Deep Learning Approach on Segmentation of Bone for BMD Measurement from DEXA Scan Images", Sixth International Conference on Bio Signals, Images, and Instrumentation (ICBSII), 2020.

[6] A.König and N. Keough,F. Joseph, "Bone mineral density of the lumbar spine in a South African population using Computed

ISBN 979-8-3503-5094-4

Tomography scans", 3rd Biennial South African Biomedical Engineering Conference (SAIBMEC), 2018.

[7] Usman Bello Abubakar, Moussa Mahamt Boukar, Steve Adeshina: "Comparison of Transfer Learning Model Accuracy for Osteoporosis Classification on Knee Radiography", 2nd International Conference on Computing and Machine Intelligence (ICMI), 2022.

[8] Varalakshmi P, Smruthi Sathyamoorthy, Darshan V, Ramanujan V, Sakthi Jaya Sundar Rajasekar: "Detection of Osteoporosis with DEXA Scan Images using Deep Learning Models", International Conference on Advances in Computing, Communication and Applied Informatics (ACCAI), 2022.

[9] Kottaimalai Ramaraj , Gautam Amiya, Pallikonda Rajasekaran Murugan, Vishnuvarthanan Govindaraj, Muneeswaran Vasudevan, Arunprasath Thiyagarajan: "Sensors for Bone Mineral Density Measurement to Identify the Level of Osteoporosis: A Study", Fourth International Conference on Smart Systems and Inventive Technology (ICSSIT), 2022.

[10] Jing Zhang, Chye-Hwang Yan, Chee-Kong Chui, and S. H. Ong, "Accurate Measurement of Bone Mineral Density Using Clinical CT Imaging With Single Energy Beam Spectral Intensity Correction", Transactions On Medical Imaging, 2010.

ISBN 979-8-3503-5094-4

Image analysis for diagnosis of Glaucoma using soft computing methods

Ponni S
Dept. of Biomedical Engg Sathyabama Institute of Science and Technology
Chennai
ponnibyravi@gmail.com

Tejaswini S
Dept. of Biomedical Engg Sathyabama Institute of Science and Technology
Chennai
tejaswinicst2002@gmail.com

Sushmitha S
Dept. of Biomedical Engg Sathyabama Institute of Science and Technology
Chennai
sushmithasakthivel11@gmail.com

Premkumar J
Dept. of Biomedical Engg Sathyabama Institute of Science and Technology
Chennai
premnsc@gmail.com

*Bethanney Janney J
Dept. of Biomedical Engg Sathyabama Institute of Science and Technology
Chennai
*bethanneyj@gmail.com

Sindu Divakaran
Dept. of Biomedical Engg Sathyabama Institute of Science and Technology
Chennai
sindudiva@gmail.com

***Abstract*— Optic nerves are neurons that carry visual signals from the eyes to the brain. When this nerve is affected, it leads to various eye diseases. Glaucoma is a neurological disorder that damages the visual nerves. If glaucoma is not treated, it might cause irreversible blindness. A rise in intraocular pressure (IOP) in the optic nerve brought on by an inadequate outflow of excess fluid from the eyes is the primary cause of glaucoma. The goal of the proposed work is to use MATLAB graphical user interface image processing techniques to produce an automated design that can diagnose glaucoma earlier on. Initially, a dataset containing retinal fundus images is collected and pre-processed. The retinal disc and retinal cup boundaries are then determined by applying the morphological Hough transform. After elliptical fitting and threshold segmentation, the cup-to-disc ratio (CDR) score is calculated. The phases of glaucoma are then diagnosed by comparing the computed CDR value with the threshold value. This unique method has overcome the existing disadvantages by reducing the time for diagnosis.**

Keywords*— Glaucoma, Fundus Images, Cup Segmentation, Cup Boundary Detection, Cup Disc Ratio (CDR)

I. INTRODUCTION

The most common cause of vision loss in the world is glaucoma. According to a global survey by the World Health Organization in 2020, about 64.3 million people are affected by glaucoma annually. Mostly, it prevails in people ranging between 40 and 80 years of age. Out of which China stands first with 15 million people being affected annually, followed by Europe with 12 million people and India with 11 million people. To avoid this risk, an early diagnosis is preferred. The existing methods used manual techniques for feature extraction which limited its application in the case of mass screening. The segmentation and detection were based on only boundary [1]. The imaging techniques used in those methods had shortcomings which resulted in less accuracy due to unavoidable errors.

A new computer-aided design has been used to improve the results to overcome the aforementioned drawbacks. First, retinal fundus images from a glaucoma dataset are manually collected and visualized. The color image is changed into a grayscale image to improve the contrast of the photographs. After using image processing methods like the morphological Hough transform to locate the discs and the cup's boundaries, the pictures are compared to black-and-white counterparts to recover the segmented disc [2]. Segmentation techniques are employed for cup boundary detection. A combination approach using threshold segmentation and elliptical fitting is used to determine the Cup to Disc Ratio (CDR). The CDR result is then compared to the stage of glaucoma (if existent) to determine if it is present or not.

II. RELATED WORKS

The model by Shoukat A, Akbar S, Hassan SA, Iqbal S, Mehmood A, Ilyas QM, trained on four datasets, effectively diagnoses glaucoma at an early stage using fundus images. It uses data augmentation to provide a variety of fundus images. The model achieved 98.48% accuracy, 99.30% sensitivity, 96.52% specificity, AUC of 97%, and an F1-score of 98% on the G1020 dataset with ResNet-50 architecture.[1]

Daniele M. S. Barros (2020) reviewed on Machine Learning (ML) algorithm: Classification under supervised learning and excluded the segmentation of boundary. Instead, the model was trained and validated, after which the collected images were compared with the dataset and results were displayed [4].

Huazhu Fu et al. (2018) in the retinal image demonstrated two novel approaches for the detection of glaucoma using a deep learning network. First is the optic disc segmentation and optic cup segmentation which is done using a multiple–labeled segmentation network called M-Net. The second is the DENet, which produced the detection of glaucoma directly from the pictures without segmentation by integrating the local optic disc region and the deep hierarchical context of the global image of fundus [5].

Abdullah Sarhan (2019) addressed both analyzing and segmentation techniques to take features out of the images, derived from a 3D Optical Coherence Tomography (OCT) machine having better quality and clarity for the detection of Glaucoma [6].

Xiang Yu Chen et al. (2015) created a deep learning architecture using CNN that includes two fully connected layers and four convolutional layers for the detection of glaucoma in images and non-glaucoma in images. The area under the curve, also known as the AUC, of the characteristic curve for glaucoma identification, is presented as a result [7].

ISBN 979-8-3503-5094-4

III. MATERIALS AND METHODS

A. Acquisition of Image

Images are acquired through a process called image acquisition. This can be accomplished with hardware systems such as encoders, sensors, cameras, and so on. Since a bad image would render the entire machine vision workflow useless, it is the most crucial stage [3]. It's critical to obtain an image with the appropriate contrast and clarity because machine vision systems evaluate only the digital image obtained and not the actual object. During the image capture process, several photo-sensitive sensors transform the incoming light wave from an object into an electrical signal. Your machine vision algorithms use these tiny subsystems to precisely describe the object. The neurosensory portion of the eye called the retina is where light impulses from objects are transformed into electrical impulses that go to the brain and give the user information about the object [8]. The retina must be simultaneously photographed and lighted through a small opening, the pupil, making retinal imaging challenging.

Since the retina has a low level of reflection, the back reflections off its surface are smaller than the total reflections from the imaging and lighting systems [9]. The fundus camera operates based on indirect ophthalmoscopy without a reflex. Conventional fundus cameras image using the pupil center and illuminate using the pupillary aperture rim. This reduces the number of overlapping reflections.

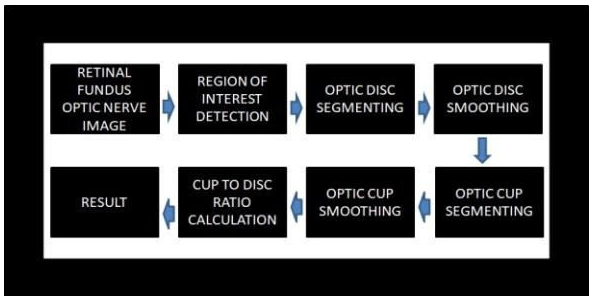


Fig. 1. Workflow of the prospective methodology

B. Region Of Interest Detection

Several preprocessing steps are applied to the photos, including green channel extraction and histogram equalization. The exudates look dazzling in the green channel of an RGB picture, in contrast to the red and blue channels. Nevertheless, RGB values are not compatible with other devices, perceptually uniform, or useful for fault detection. Grayscale images are therefore favoured for processing since they are easy to understand, help to simplify algorithms, and also eliminate the complexity related to computational needs [6]. The RGB images are converted to grayscale versions to facilitate ROI recognition. To find the vertical cup-to-disc ratio (CDR), the optic disc and cup must be isolated from the fundus images first. The optic disc and cup were taken out for each retinal fundus image [10].

Initially, a study of the fundus photos utilized in this investigation shows that the optic disc region typically has a paler tone than the surrounding retinal area. The fundus images with the highest intensity are chosen to potentially serve as the optic disc center candidates. The intensity-weighted centroid technique is recommended for the determination of the approximate ROI Center. The ROI border, which is a rectangle with twice the dimensions of a normal optic disc, serves as the beginning point for segmenting the optic disc.

C. Hough Transform

By the employment of feature extraction called morphological Hough transform which is a soft computing method, the presence or absence of glaucoma is carried out. A common technique for identifying or segmenting geometric objects from photos is referred to as Hough Transform. In this work, we examine the mathematical equations underlying the Hough Transform principle and apply a novel method based on it for fast line and circle recognition in image processing [8]. Certain basic images, including straight lines with varying directions, and circles with varying detection, thickness, and numbers, were detected accurately by this approach. The outcomes demonstrate how rapidly and with minimal memory consumption our method can be used for line recognition and segmentation in 3D ultrasonic pictures [11].

C. Optic Disc Localization

A streamlined system of optic disc segmentation is introduced through image pre-processing to detect an optic disc boundary. First, utilizing the red channel, a rough localization of the optic disc region is shown. Since the contrast between the optic disc and non-optic disc area is higher in the red component than in other channels, it is used [10]. An operation known as "morphological closure" is used to eliminate the blood vessels. Closing is characterized by dilation and erosion; it makes foreground region boundaries larger and background color holes smaller in foreground regions. The structuring element in this paper was chosen to be 20 by 20 pixels because this size is greater than the width of the major vessels. [12].

D. Optic Cup Segmentation

The obtained image is further smoothed by applying a median filter after the closing operation. The interpenetration of a cup with surrounding tissues and blood vessels makes optic cup segmentation more difficult than optic disc extraction. The color component analysis approach and the threshold level set approach are the two methods for cup segmentation that are presented in this study [13]. One way to think of threshold segmentation is as the method of categorizing background from the foreground. Due to its primary method of extracting foreground based on gray value information, threshold segmentation is especially useful for the segmentation of images with a strong contrast between foreground objects and background. [14].

E. Boundary Detection

The process of ellipse fitting involves determining which ellipse best fits a given set of points. Measuring the elliptical nature of a finite set of points is another area of interest. Shape-based (border-based) and area-based methods are the two primary methods used for ellipse fitting and determining ellipticity [15].

F. Cup to Disc Ratio (CDR) calculation

The crucial marker to use for initial glaucoma diagnosis and screening is the Cup to Disc Ratio (CDR). Segmented optic disc and cup measurements are usually needed to obtain CDR. The significant overlap between the neuroretina rim

ISBN 979-8-3503-5094-4

and optic cup regions makes it challenging to reliably and accurately derive CDR values automatically, even with substantial effort. This study presents a direct method for CDR estimation which is calculated optic disc/cup segmentation and formulates CDR estimation as a general regression problem. The cup to Disc Ratio is calculated using the formula:

$$CDR = Cup\ Area/Disc\ Area \quad (1)$$

IV. Results and Discussion

Glaucoma is a most serious issue as it leads to complete vision loss, it is always recommended to consult a doctor and diagnose it in its early stages. A novel approach is suggested to minimize the workload of medical personnel and shorten the diagnosis time after the scanning process is complete. By the application of the morphological Hough transform which is the soft computing technique employed here along with segmentation techniques for the calculation of CDR, the problem of late diagnosis is rectified. Therefore, the proposed system aims to provide technical solutions to hospitals.

Fig 2. CDR Calculation and prediction of glaucoma not detected

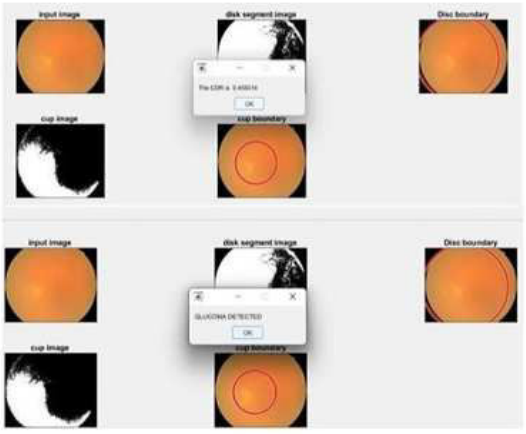


Fig 3. CDR Calculation and prediction of glaucoma detected

The image dataset consists of Fundus images for diabetic retinopathy. These images are collected from Kaggle [18], Fundus image database for diabetic retinopathy. The dataset in this proposed work consists of 800 fundus images. From these 80 percent were used for the training phase and the remaining 20 percent images were used for the testing phase. Here we present a sample image which is given as an input. After the pre-processing steps, the cup and disc are indicated through the boundary and the algorithm presents the CDR values, as shown in Fig 2. The presence or absence of glaucoma is predicted based on the values of CDR. Comparing the result of CDR obtained with the threshold we conclude that no glaucoma is detected as the CDR value 0.323506 lies within the threshold (>0. 45) in Fig 2. Similarly, Fig 3 depicts the output of glaucoma is recognized since the CDR value 0.455014 falls within the threshold (>0. 45).

TABLE I. Disc Segmentation And Disc Boundary

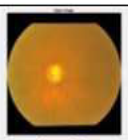
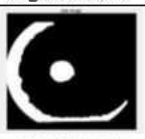
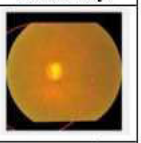
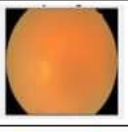
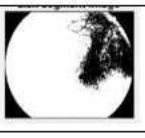
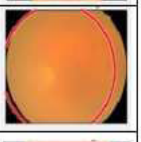
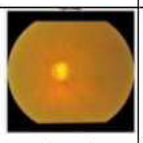
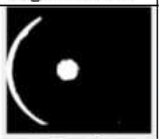
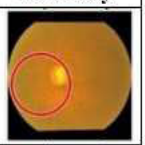
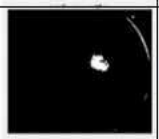
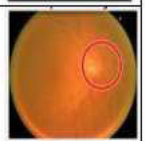
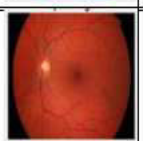
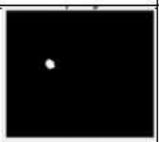
Image No.	Input Image	Disc Segmentation	Disc Boundary
X1			
X2			
X3			
X4			

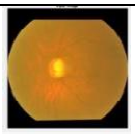
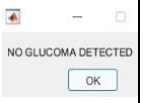
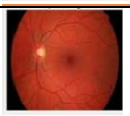
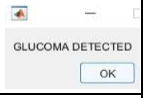
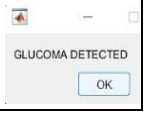
TABLE II. Cup Segmentation And Cup Boundary

Image No.	Input Image	Cup Segmentation	Cup Boundary
X1			
X2			
X3			
X4			

The results of disc boundary and disc segmentation are displayed in Table 1. Table 2. summarizes the results of cup boundary and cup segmentation. Table 3. depicts the findingsbased on cup-to-disc ratio value.

ISBN 979-8-3503-5094-4

TABLE III. RESULTS BASED ON CDR VALUE

Image No.	Input Image	CDR Value	Findings
X1		0.354966	
X2		0.323506	
X3		0.764637	
X4		0.808507	

V. CONCLUSION

The suggested work offers a cutting-edge technique for segmentation algorithm-based glaucoma identification. To extract the disc and cup margins from the scanned retinal fundus pictures, a morphological rough transform is employed. This aids in precisely locating the region of interest (ROI). The cup and disc ratio is computed using the outcomes of the aforementioned procedure. The method's output was demonstrated to be precise and effective. We aimed for maximal accuracy in both development and comparisons. The segmented approaches and the approach parameters of the proposed system are optimized to achieve maximum accuracy. In order to make our method more user- friendly, we also intend to integrate it with other illness prediction models for the detection of numerous diseases.

ACKNOWLEDGMENT

We are truly grateful to the Department of Biomedical Engineering and the School of Bio and Chemical Engineering for giving us guidance as well as encouragement to complete the research.

REFERENCES

[1] Chang, Chuan-Wang, Chuan-Yu Chang, Yuan-Yi Lin, Wei-Wen Su, and Henry Shen-Lih Chen, "A Glaucoma Detection System Based on Generative Adversarial Network and Incremental Learning," Applied Sciences, vol. 13(4), pp. 2195, 2023. https://doi.org/10.3390/app13042195.

[2] T. Nazir, A. Irtaza, and V. Starovoitov, "Optic Disc and Optic Cup Segmentation for Glaucoma Detection from Blur Retinal Images Using Improved Mask-RCNN," International Journal of Optics, vol. 2021, pp. 1–12, Jul. 2021.Improved Mask-RCNN," International Journal of Optics, vol. 2021, pp. 1–12, Jul. 2021

[3] L. J. Coan et al., "Automatic detection of glaucoma via fundus imaging and artificial intelligence: A review," Survey of Ophthalmology, Aug. 2022, doi:https://doi.org/10.1016/j.survophthal.2022.08.005.

[4] Barros, D.M.S., Moura, J.C.C., Freire, C.R. et al. Machine learning applied to retinal image processing for glaucoma detection: review and perspective. BioMed Eng Online., vol. 19, pp. 1-20, 2020. https://doi.org/10.1186/s12938-020-00767-2.

[5] H. Fu et al., "Disc-Aware Ensemble Network for Glaucoma Screening from Fundus Image," in IEEE Transactions on Medical Imaging, vol.37, no. 11, pp. 2493-2501, Nov. 2018, doi: 10.1109/TMI.2018.2837012.

[6] A. Sarhan, J. Rokne, and R. Alhajj, "Glaucoma detection using image processing techniques: A literature review," Computerized Medical Imaging and Graphics, vol. 78, pp. 101657, Dec. 2019.

[7] Xiangyu Chen, Yanwu Xu, Damon Wing Kee Wong, Tien Yin Wong, and Jiang Liu, "Glaucoma detection based on deep convolutional neural network," Conference proceedings: ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual Conference, vol. 2015, pp. 715–718, 2015.

[8] Sundari, Sivaguru.S," Early Detection of Glaucoma from Fundus Images by Using MATLAB for Diabetic Retinopathy", International Journal of Innovative Research in Computer and Communication Engineering, vol. 5(1), pp. 174-181, Jan 2017.

[9] S. G. Athira Lekshmi and G. M. Rajathi, "Detection of Glaucoma and Diabetic Retinopathy using Image Processing Technique by Raspberry pi," Indian Journal of Science and Technology, vol. 12(29), pp.1-6, 2019, DOI: 10.17485/ijst/2019/v12i29/146968.

[10] A. K. Schuster, C. Erb, E. M. Hoffmann, T. Dietlein, and N. Pfeiffer, "The Diagnosis and Treatment of Glaucoma," Deutsches Ärzteblatt International, vol. 117(13), pp. 225–234, Mar. 2020.T. Shyamalee and
D. Meedeniya, "Glaucoma Detection with Retinal Fundus Images Using Segmentation and Classification," Machine Intelligence Research, vol. 19(6), pp. 563–580, Nov. 2022,

[11] S. Joshi, B. Partibane, W. A. Hatamleh, H. Tarazi, C. S. Yadav, and D. Krah, "Glaucoma Detection Using Image Processing and Supervised Learning for Classification," Journal of Healthcare Engineering, vol. 2022, pp. e2988262, Mar. 2022.

[12] G. Ramanathan, D. Chakrabarti, A. Patil, S. Rishipathak and S. Kharche, "Eye Disease Detection Using Machine Learning," 2021 2nd Global Conference for Advancement in Technology (GCAT), Bangalore, India, 2021, pp. 1-5, doi:10.1109/GCAT52182.2021.9587740.

[13] G. Prasad, N. Kumar, K. Raju, S. Kumar, and S. Chandrappa, "Glaucoma Detection Using Clustering and Segmentation of the Optic Disc Region from Retinal Fundus Images," SN computer science, vol. 4, no. 2, Feb. 2023, doi: https://doi.org/10.1007/s42979- 022-01592-1.

[14] Singh, L.K., Khanna, M., Thawkar, S. et al. A novel hybridized feature selection strategy for the effective prediction of glaucoma in retinal fundus images. Multimed Tools Appl (2023).

[15] Sibhgatullah I Khan, Shruti Bhargava Choubey, Mohammed Mahaboob Basha," Automated Glaucoma Detection from fundus images using wavelet-based denoising and Machine Learning", Vol. 30(1), July 9, 2021.

[16] Zedan, M.J.M.; Zulkifley, M.A.; Ibrahim, A.A.; Moubark, A.M.; Kamari, N.A.M.; Abdani, S.R. Automated Glaucoma Screening and Diagnosis Based on Retinal Fundus Images Using Deep Learning Approaches: A Comprehensive Review. Diagnostics 2023, 13, 2180.

[17] F. G. Mohammed, S. D. Athab, and S. G. Mohammed, "Disc damage likelihood scale recognition for Glaucoma detection," Journal of Physics: Conference Series, vol. 2114(1), pp. 012005, Dec. 2021.

[18] Kaggle. Fundus image database for diabetic retinopathy detection. https://www.kaggle.com/c/diabetic-retinopathy-detection.

DETECTION OF ANEMIA FROM PALPEBRAL IMAGE OF ANTERIOR CONJUNCTIVA USING SVM CLASSIFIER

Sofia Bobby J
Deaprtment of Biomedical Engineering
Jerusalem College of Engineering
Chennai, India
sofiabobbyj@jerusalemengg.ac.in

Mythily V
Deaprtment of Biomedical Engineering
Jerusalem College of Engineering
Chennai, India
mythily.bme@jerusalemengg.ac.in

Annapoorani C L
Deaprtment of Biomedical Engineering
Jerusalem College of Engineering
Chennai, India
annapooranic@jerusalemengg.ac.in

Purnima S
Deaprtment of Biomedical Engineering
Jerusalem College of Engineering
Chennai, India
purnimas@jerusalemengg.ac.in

***Abstract*— Anemia is a medical disorder that arises when an individual's blood is deficient of sufficient mature, developed red blood cells with normal hemoglobin level. One of the main components of erythrocytes, hemoglobin, has a strong affinity for binding oxygen, which is necessary for cell survival. The body's cells won't get enough oxygen if there are any aberrant red blood cells or low hemoglobin levels, and this condition eventually leads to a condition known as hypoxia. The conventional method of diagnosing anemia involves pricking of finger and using analytical reagents to investigate is quite time consuming and requires skilled laboratory procedures and personnel. To overcome these limitations, a novel approach for the automated non-invasive detection of anemia is developed. In this work, a method for diagnosing anemia by analyzing changes in anemic individuals' anterior conjunctival pallor based on processed eye images is proposed. Nowadays, patients with anemia disease present in the world increased by around 60-70% respectively. This proposed work has successfully characterized to introduce novel approach for early and accurate anemia disease diagnosis. It employs LBP texture analysis for classification of eyelid images. There are several features which are considered based on extracted statistical analysis. The classification results demonstrate that these features are utilized to identify normal and abnormal patients successfully with an accuracy rate of 91%.**

Keywords— ***Anemia, Hough circle, LBP algorithm, PSO algorithm, SVM algorithm***

I. INTRODUCTION

Many people are easily prone to anemia because of the absence of adequate medical facilities in developing nations, which has an impact on the socioeconomic well-being of the community as a whole. If anemia symptoms can be identified early enough to avoid costly blood tests and procedures, many of which are unobtainable in developing rural areas and this predicament can be resolved. It is rare for these fields of diagnosis to even have doctors or engineers. It would be beneficial if non-invasive techniques could identify a patient's anemia without the need for costly tests, procedures, or the presence of a highly qualified technician. A person's anemia may be a sign of other conditions such as malnourishment or jaundice, and the identification of anemia may be a marker for the presence of other co-morbidities.

The main component of blood that contributes to its coloring is hemoglobin. The concentration of hemoglobin in a person's blood is thought to be the primary factor in the identification of anemia. Anemia detection necessitates intrusive techniques, a sterile laboratory, advanced equipment, knowledgeable staff, and a significant amount of time. The goal of the suggested approach was to provide a novel, non-invasive, affordable, and straightforward primary screening test for anemia that might replace invasive methods of detecting the disease and have a significant positive humanitarian impact in the world's developing nations.

Examining the eye's conjunctival pallor is usually done in clinics as a fast way to check for anemia. Usually, doctors gently pull the eyelid and examine the paleness of anterior conjunctiva of eye based on their subjective assessment. The reliability of the visual detection procedure may be impacted by the anterior conjunctiva's low color sensitivity, which may make this clinical approach of anemia detection ineffective in many circumstances. By offering an ideal and quick method for non-invasive anemia detection, this work seeks to overcome these drawbacks. In order to detect anemia, our approach involves examining the depth of paleness of anterior edges of the palpebral conjunctiva. It measures the color of the conjunctival region of the eye using digital image captured with a smartphone camera at a suitable resolution and with enough light. To ascertain whether or not the patient is anemic, the red and green colour components of the image of conjunctiva are extracted and compared against a threshold.

When anemia is present, the red spectral component of light that falls on RBC's surface is reflected differently than the green spectral component, which absorbs most of the incident light. Hence, it is possible to infer indirectly the quantity of hemoglobin in the human bloodstream by comparing the two components, namely the Red and Green of the RGB image of the anterior conjunctiva of eye. LBP

ISBN 979-8-3503-5094-4

texture analysis is used in this process to classify images of eyelids.

II. METHODOLOGY

The image of the eye is captured using Smart phone and the colour image is converted to grayscale image to simplify the algorithms as well as to reduce the complexity in further image processing . The Hough algorithm works effectively on gray scale image than colour image. After processing the image, features are taken out and used to train the Support Vector Machine (SVM) algorithm, which determines whether or not the patient is anemic. The overall methodology of anemia detection is depicted as flowchart, shown below.

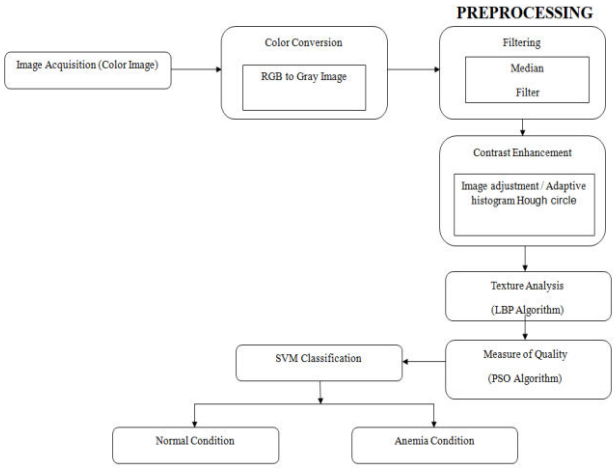


Figure 1: Block diagram for the proposed work

A. *Image Acquisition*

The first step of the proposed study is to capture images of anterior conjunctiva area of the eye in ambient light conditions using a smart phone. To do this, the lower eyelid is gently pulled down with thumb and picture is taken so that the area of anterior conjunctiva is focused and displayed as enlarged as possible. Camera flash should be turned off when taking photos to avoid excessive glare. The detecting algorithm is significantly impacted by the quality of the photo that was captured. A poorly captured photo could affect the entire process.

B. *Color Conversion*

Converting a color image to a grayscale image is done by converting RGB values (24 bits) to grayscale values (8 bits). Images were first converted to grayscale because the Hough algorithm works better on black and white images than color images. This grayscale images demonstrate that the image processing algorithm used in the work has retained the key elements of the color image, including contrast, sharpness, shadows, and image saturation.

C. *Image pre-processing*

As seen in Fig. 2, a non-linear digital filter called the median filter is employed in this work to reduce noise from an image or signal. In order to apply the median filter, each pixel in the image is filtered pixel by pixel and its value is replaced with the median of its neighboring pixels. This type of pre-processing noise reduction is frequently used to improve the results of later processing steps , such as contrast enhancement and picture edge detection.

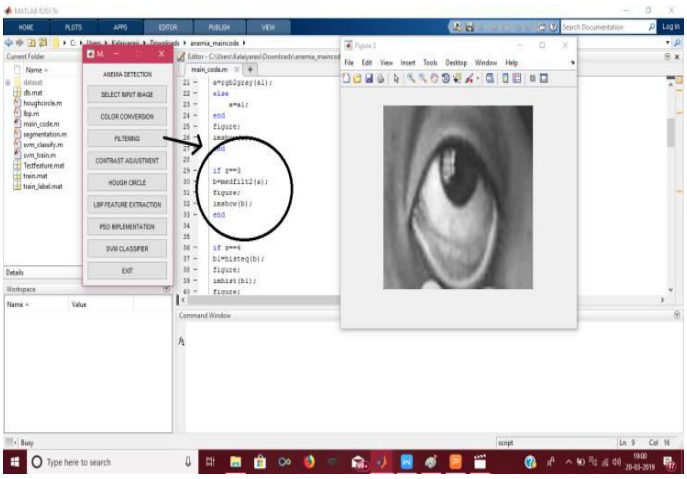


Figure 2 : Median Filtering

D. *Contrast enhancement*

To increase the contrast variation between image of interest and irrelevant backgrounds, contrast enhancement method is usually applied as a contrast stretching followed by a tonal enhancement. This approach enhances the visibility of features in the enhanced image. While tonal enhancements increase the contrast (brightness) variations in the dark, gray , or highlighted bright regions at the expense of the background's tonal variations . Contrast enhancement improves a consistent range of tones in an image between its lightest and darkest areas.

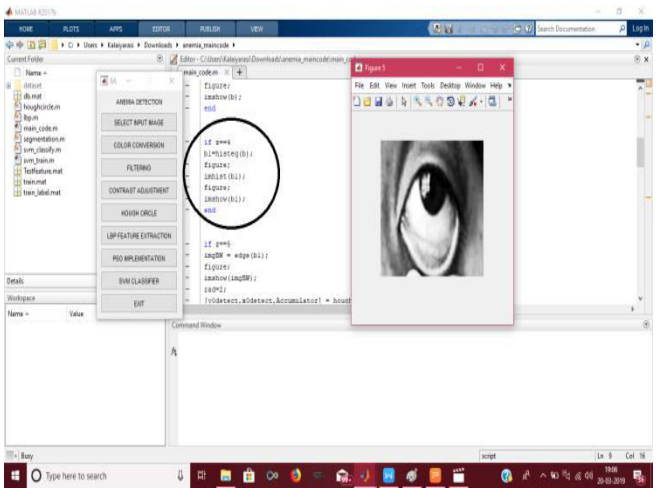


Figure 3: Contrast Enhanced Image

ISBN 979-8-3503-5094-4

E. Feature Extraction

A feature extraction method for finding circles is the specialised form of Hough transform, circle Hough Transform (CHT). Its objective is to locate circles in input images with errors . The circle candidates are created by "voting" in the Hough parameter space, and subsequently choosing the local maxima in an accumulator matrix.

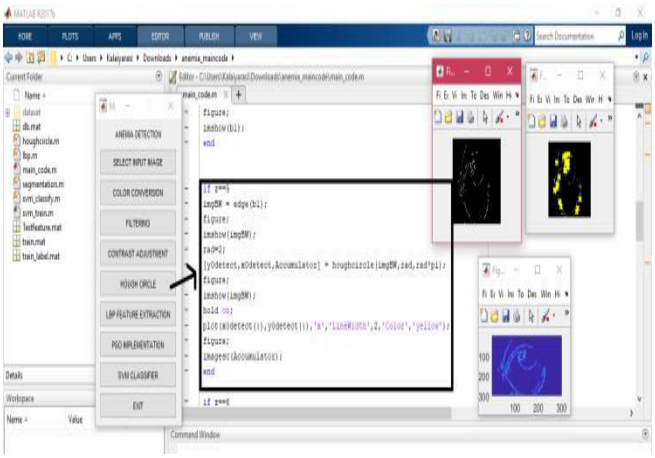


Figure 4 : Hough Transformed Image

F. Texture Analysis

A texture operator , the Local Binary Pattern (LBP) is a straightforward yet highly effective technique that assigns a binary number to each pixel in an image by thresholding its surrounding pixels. A threshold is set using the gray value of the center pixel , the neighboring pixels in LBP are transformed into binary codes 0 or 1. All the transformed codes are then organized as an ordered pattern based on its placements in relation to the center pixel. The relative difference value is calculated by the LBP operator to ensure gray-value invariance. LBP may generate 256 distinct output types, which correlate to 256 distinct binary pattern types.

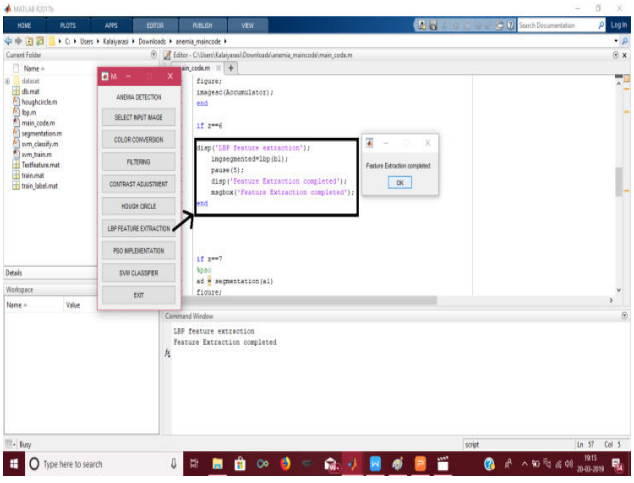


Figure 6 : LBP Feature Extraction

The greatest benefit of LBP is that it may be used in situations requiring high speed due to its straightforward computation. The LBP-extracted feature illustrates how the local area's textures relate to one another. The two complementary ways of describing two-dimensional textures of surface are local spatial patterns and gray scale contrast, which served as the foundation for the creation of the LBP operator. To construct labels for the image pixels, the original LBP operator averages the 3 x 3 neighbourhood of each pixel value with the center value. Then, it treats the result as a binary number.

G. PSO Algorithm

PSO, or particle swarm optimization, is a type of swarm intelligence technique used to tackle optimization issues. There is a population of random solutions in the system at first. Every possible resolution, referred to as a particle, is assigned an arbitrary speed and traverses the issue space. Each particle has memory, and it remembers its best position from before as well as the associated fitness. The particle with the highest optimization is referred as the global best (gbest) of the swarm. The swarm's global best (gbest) particle is the one with the highest fitness level among the several pbest for each particle. Particle selection and acceleration toward their optimal sites constitute the fundamental idea of the PSO approach.

H. Support Vector Machine

A Support Vector Machine (SVM) employed in this approach acts as a discriminative classifier that derives its formal definition from a separating hyper plane. Stated differently, the technique generates an optimum hyper plane that classifies fresh samples based on labeled training data, therefore facilitating supervised learning. The margin near the dividing hyper plane is maximized using support vector maximization. The support vectors become a subset of training samples that entirely represent the decision function. By mapping the input vector to a higher dimension space, SVM creates maximum separation hyper planes. An input space designated for calculating similarity using a kernel function is known as a feature space.

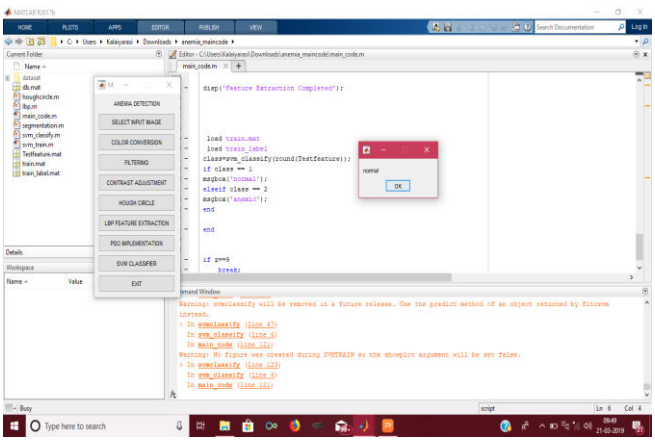


Figure 7: SVM output

ISBN 979-8-3503-5094-4

III. RESULTS

A 16MP rear-facing camera smart phone, the Vivo V3 pro, was used to take pictures of a few patients. To achieve the right ambient lighting, the camera flash was turned off during the day when the pictures were taken. Our system that uses texture analysis and SVM classification offers promising evidence to diagnose anemia non-invasively according to the paleness of eye conjunctiva region.

Since color photos are not as well suited for the Hough algorithm as black and white images are, the RGB image is first converted to grayscale. By applying thresholding on the region around each pixel and taking the output value as a binary number, the Local Binary Pattern texture operator labels the pixels in an image in a straightforward yet incredibly effective manner. The facial image representation is achieved by using a basic data vector in conjunction with histograms and LBP.

PSO, a computational approach in computer science, optimizes the solution of feature extraction and segmentation of image by repeatedly attempting to enhance a potential solution of pattern recognition in relation to a specified quality metric as shown in Fig.8.

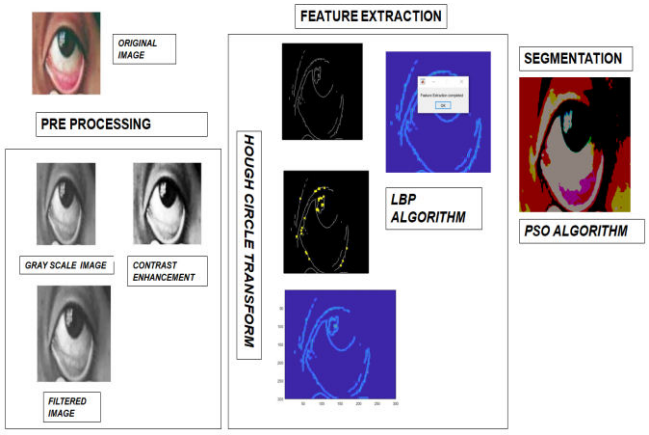


Figure 8 : Processing of anterior conjuctival image to detect anemia

A separating hyper plane defines a discriminative classifier called a Support Vector Machine (SVM) is used to identify whether or not an individual is anemic. The SVM output is shown in Fig.9.

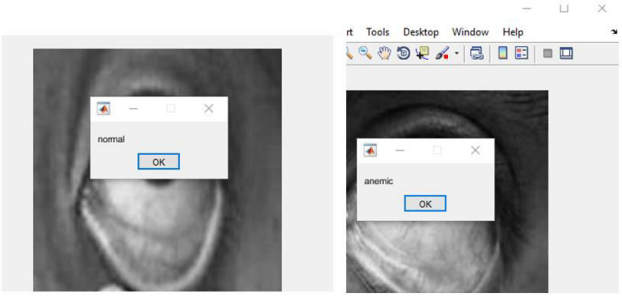


Figure 9: SVM output of a normal and anemic individual

IV CONCLUSION

In conclusion, we have developed an approach involves non-invasive method that successfully and accurately detects anemia with 91% accuracy. Using an Android application that we developed, the procedure entails capturing the image of the region of conjunctival pallor of eye by normal android phone in adequate lighting. Following image processing, a computer program extracts the RGB component of the image of the patient's pallor in the anterior conjunctiva and uses support vector machines (SVM) to compare it to a predefined threshold value in order to assess the patient's anemia.

Since the threshold value is established using the gathered, the total number of subjects has a significant influence on the system's performance, and the inclusion of around 50 numbers of data in our algorithm would significantly increase the accuracy.

Moreover, the distinction between patients who are anemic and those who are not can also be made based on the patient's demographics. The technique of detecting anemia can be a valuable weapon in the prevention and protection against anemia in the poorly developed region globally. It may also lead to a noticeable improvement in such areas' health condition and a significant decrease in the number of deaths caused by anemia's effects and thereby alleviating the effects of hemorrhage.

REFERENCES

[1] Stevens, G.A., Finucane, M.M., De-Regil, L.M., et al. (2013) Global, Regional, and National Trends in Haemoglobin Concentration and Prevalence of Total and Severe Anaemia in Children and Pregnant and Non-Pregnant Women for 1995-2011: A Systematic Analysis of Population-Representative Data. The Lancet Global Health, 1, e16-e25.

[2] S. S. Yalcin, S. Unal, et.al., “The validity of pallor as a clinical sign of anemia in cases with beta-thalassemia,” The Turkish journal of pediatrics, vol. 49, no. 4, p. 408, 2007.

[3] Butt, U. Ashfaq, S.F.H.Sherazi, et.al., “Diagnostic accuracy of “pallor” for detecting mild and severe anaemia in hospitalized patients,” J Pak Med Assoc. 60: 762–5, 2010.

[4] A. Kalantri, M. Karambelkar, et.al., “Accuracy and reliability of pallor for detecting anaemia: a hospital-based diagnostic accuracy study,” Malaga G, editor. PLoS One. Public Library of Science. 5: e8545, 2010.

[5] P. Chalco, L. Huicho, et.al., “Accuracy of clinical pallor in the diagnosis of anaemia in children: a meta-analysis,” BMC Pediatr. 5: 46, 2005.

[6] M. Setaro, A. Sparavigna, ”Quantification of erythema using digital camera and computer-based colour image analysis: a multicentre study,” Skin Res Technol. 8: 84–8, 2002.

ISBN 979-8-3503-5094-4

[7] S. Roychowdhury, D. Koozekanani, and K. Parhi, "Dream: Diabetic retinopathy analysis using machine learning," IEEE Journal of Biomedical and Health Informatics, vol. 18, no. 5, pp. 1717–1728, 2014.

[8] Nasrul Humaimi Mahmood, Poon Che Lim, et.al, "Blood Cells Extraction using Color based Segmentation technique", International Journal of Life Science Biotechnology and Pharma Research, vol.2, 2013.

[9] PranatiRakshit, KritiBhowmik, "Detection of Abnormal Finding in Human RBC in Diagnosing Sickle cell Anaemia using Image Processing ", International Conference on Computational Intelligence: Modeling, Technique and Application, pp. 28-36, 2013.

[10] Z. Butt , U. Ashfaq, S.F.H.Sherazi, N.U. Jan, U. Shahbaz, "Diagnosticaccuracy of "pallor" for detecting mild and severe anaemia inhospitalized patients," J Pak Med Assoc. 60: 762–5, 2010.

ISBN 979-8-3503-5094-4

Identification of Gastric Abnormalities with Synchronous Biopsy Site Localization from Endoscope Imaging in Gastro-Intestinal Tract using Deep Learning Technique

Gnanapriya S
Information Technology
Easwari Engineering College,Chennai, India
gnanapriya.s@eec.srmrmp.edu.in

Reshma D
Information Technology
Easwari Engineering College,Chennai, India
reshmadhilip299@gmail.com

Varun M
Information Technology
Easwari Engineering College,Chennai, India
varunm.infotech@gmail.com

***Abstract*— One of the most prevalent malignant tumors with a dismal prognosis is gastric cancer (GC). Although pathologic verification and CT scanning are recommended for additional therapy, endoscopic examination is mostly used for early detection. One of the worst diseases with a dismal prognosis is still gastric cancer. Artificial intelligence (AI) assistance technologies have a great opportunity to reduce workload and improve diagnostic accuracy due to the global pathologist shortage. The majority of gastric cancers (GCs) exhibit genetic instability, either as chromosomal or microsatellite instability, which is thought to be a precursor to gastric carcinogenesis. The features of each subtype are better understood because of the new classification of gastric malignancies on the basis of histologic features, genotypes, and molecular phenotypes, which also improves early diagnosis, prevention, and treatment. As a result, the suggested system offers a practical way to more accurately forecast the categories of gastrointestinal problems. In this research, the types of diseases will be identified by the use of Deep Learning algorithms, such as CNN. The datasets will be gathered from Kaggle and preprocessed using several methods, including adaptive median filtering. The UNet algorithm is used to segment the pictures after preprocessing. After that, a deep learning algorithm will be used to train these datasets, and a model file will be produced. It is able to accurately identify gastrointestinal anomalies such as ulcers, funduses, healthy tissues, and esophagitis when given an input image for prediction. As a result, this approach contributes to a more accurate and efficient detection of gastrointestinal problems than current models.**

***Index Terms*— Gastric Cancer, Convolutional Neural Network (CNN), Biopsy, Gastrointestinal Anomalies, UNet Algorithm**

I. Introduction

Gastric cancer, also known as stomach cancer, is characterized by the uncontrolled proliferation of cells in the stomach lining. Influenced by factors such as age, diet, and stomach health, the likelihood of acquiring gastric cancer varies.

The development of gastric cancer is a gradual process, often preceded by precancerous changes in the stomach's mucous layer. These early alterations are challenging to detect due to the absence of noticeable symptoms. The slow progression of gastric cancer spans several years, with distinct symptoms emerging based on the specific section of the stomach affected. Detection becomes more complex as precancerous changes rarely manifest noticeable symptoms, contributing to delayed diagnosis.

Differentiating sections of the stomach where cancer originates result in varied symptoms and outcomes. The location of the cancer also plays a crucial role in determining treatment options. Cancers originating or extending into the gastroesophageal (GE) junction share similarities with esophageal cancers in terms of staging and treatment. Understanding the multifaceted aspects of gastric cancer, including its gradual development, elusive precancerous changes, and the influence of location on symptoms and treatment, is essential for effective diagnosis and management.

II. Literature Survey

One of the paper[1] presents ID-GCS, a novel intelligent decision-making technique created for the screening of gastric cancer (GC). ID-GCS is a data-driven system that uses multimodal semantic fusion and a mixed neural network for attention to retrieve text interpretations of gastroscopy reports and smoothly merges it with related visuals. Even with these encouraging findings, a significant drawback is the sole dependence on data from gastroscopies, which raises questions over its applicability as the only method for identifying stomach cancer.

The "guardian of the genome" TP53, is a crucial gene that is frequently mutated in a variety of malignancies, especially gastric cancer (GC). A study[2] used information from The Cancer Genome Atlas (TCGA) GC dataset to methodically examine the clinical significance of TP53 mutation locations in GC. The results showed that patients with TP53 gene secondary structural mutations in turn areas had a worse prognosis than patients having abnormalities in beta strands locations (log-rank $p = 0.043$). Furthermore, poorer survival results were linked to particular individual mutation sites, like R248 (log rank $p = 0.035$).

Nanorobots loaded with contrast medium operate as optimisation agents in the suggested computing-inspired multifocal cancer identification process (MCDP), which presents a unique framework in which tumor detection is regarded as an optimisation issue. One such work[3] modifies niche genetic algorithms (NGA), which are employed in

ISBN 979-8-3503-5094-4

multimodal optimization (MMO), in order to locate tumor targets more effectively while taking into account actual in vivo nanorobot propagation and control. MCDP performs better when the crossover operator is changed.

The first commercially available CMOS Vascular Endothelial Growth Factor (VEGF) sensor created specifically for cancer diagnostics using human blood is presented in this research. It also shows a broad dramatic spectrum of 18.3b, spanning significant fluctuation from sensor to sensor. One particular method of study[4] has good selectivity when compared to other proteins and can detect VEGF in both human blood serum and phosphate-buffered saline (PBS).

In order to automatically detect and mark malignant spots, one study suggests a unique texture-map-based branch-collaborative network that combines a deep convolutional neural network (DCNN) and textured map. The DCNN model[5], which consists of the upper and lower subdivisions, uses semantic division and region-of-interest tagging to extract malignant patches and improve precision. The Gabor filter exhibits an average of high sensitivity (0.9314) as well as specificity (0.9475) in the experimental data. The restricted adaptability of Region-based Convolutional Neural Networks (R-CNN) in certain circumstances is a known limitation, notwithstanding their effectiveness.

The cytotoxicity of three radiofrequency non-thermal plasma (NTPs) on triple-negative breast cancer cells, MDA-MB-231, is compared in a[6] work. The three discharge modes, which include a radio-frequency plasma jet mode and the Ω and γ forms of capacitively connected radio-frequency (CCRF) discharge, all lower the ability of cancer cells to proliferate. In contrast to the jet mode, the treatment duration in the Ω and γ modes is more than ten times longer for comparable efficacy. According to the study, a substantial portion of these NTP modes' anti-cancer benefits are attributed to short-lived reactive species, defined as those with lifetimes less than 1 second.

One journal[7] presents BCCGD, a technique to find potential breast cancer-related genes within protein complexes by combining subcellular localization information using protein-protein interaction networks (PPINs). Using a positive factorization of matrix to identify protein complexes, defining the importance of subcellular compartments, and building edge-weighted PPINs, BCCGD creates disease-specific networks and uses the WDC approach to prioritize potential genes. Potential negative results in protein-protein interactions are one of the drawbacks, though.

In one particular research[8], a TSVR algorithm that utilizes DNN weighting is used to develop a recurrence of cancer prediction model, which is then applied to a dataset of 50,000 cases spanning seven different forms of cancer. The cuckoo algorithm is employed for parameter optimization. The enhanced TSVR algorithm outperforms CNN and e-TSVR models in terms of prediction accuracy, showing over 91% for a range of malignancies. Nevertheless, more investigation is required before the suggested technique may be used in clinical situations in real time.

In an unique study[9], a novel tool for the quick and precise diagnosis of breast cancer based on near-infrared spectroscopy—Spectral-IRDx—is presented. The device, which makes the distinction between cancer and ordinary tissues, evaluates normalized observed voltages in deparaffinized biopsy specimens of breast tissue samples at near-infrared (NIR) wavelengths. Further investigation with a larger sample size is necessary, even if these encouraging first results support the idea of a non-intrusive NIR-based breast cancer diagnosis platform.

In order to improve intraoperative cancer imaging resolution, a particular work by Efthymios P. Papageorgiou, Bernhard E. Boser and Mekhail Anwar, offers a novel imaging detector for interaction imaging with angle-selective gratings. The CMOS-based sensor[10] has optical features that effectively reject ambient light and deblur pictures, restricting the angle of vision to ±18◦. With a unique 11fF MOM capacitor and a high-gain capacitive transimpedance amplifier pixel, the 4.7mm by 2.25mm (80 by 36 pixels) sensor achieves a high sensitivity of 8.2V s1 pW1. The ability to manipulate small, morphologically complicated tumor cavities is made easier by the miniaturization of optical elements.

In order to automate cancer detection using Automated Breast Ultrasound (ABUS), a unique study[11] presents a novel 3D convolutional network. The goal is to maximize efficiency while preserving high sensitivity and reducing false positives. To efficiently harness multi-layer features, the proposed network employs a densely deep supervision mechanism. Additionally, it uses a threshold loss for voxel-level adaptive thresholding. The efficacy of the method is demonstrated by evaluation on a dataset comprising 614 ABUS volumes from 219 patients, yielding a sensitivity of 95% with just 0.84 false-positive results per volume..

In order to effectively classify entire slide lung cancer images, one research team introduces an innovative weakly supervised method that addresses issues including sparse annotations and irregular tumor patterns. Taking advantage of a patch-based fully convolutional network (FCN), an approach[12] collects prejudiced sections and quickly generates representative deep features. Using a large-scale lung cancer WSI dataset, this descriptor is introduced into a random forest (RF) classifier for image-level prediction, yielding higher performance with an accuracy of 97.3%. Using weakly supervised learning techniques, the system significantly outperforms state-of-the-art algorithms and shows promise for aiding doctors in histology image diagnosis.

In order to identify cancer of the prostate in sequential contrast-enhanced ultrasound (CEUS) pictures, a deep learning architecture is presented in one research[13]. Through the use of three-dimensional convolution procedures, the approach consistently pulls features from both temporal as well as spatial dimensions, collecting dynamic perfusion information stored in numerous adjacent frames. Higher frequencies during ultrasound imaging produce greater heat, hence close attention to the ultrasonic frequency is necessary.

In order to address the lack of reliable prognostic models, one study[14] suggests the MP4Ei [illegible] forecasting

ISBN 979-8-3503-5094-4

Invasive Disease-Free Survival (IDFS) in Chinese women with the initial stages of breast cancer. The model includes 23 patient characteristics pertaining to statistics, diagnosis, pathology, and treatment using stratified feature selection. The suggested approach is important for clinical practice because it helps physicians create treatment regimens based on prognostic predictions.

One particular study[15] uses two different datasets totaling 1122 samples to analyze temporal patterns of gene expression in colorectal cancer. The study finds two sorts of gene expression patterns as cancer progresses: genes that are constantly increasing and genes that are continually dropping. Clinical factors are controlled for using the conventional linear mixed-effect regression model (LMER). According to the results, a gene expression profile evaluation can help comprehend the fundamental processes that underlie the growth and metastasis of cancer.

III. Existing System

One commonly used technique for identifying gastric lesions and promoting initial detection and management of gastric cancer (GC) is gastroscopy. The efficacy of conventional GC screening is contingent upon the proficiency of a gastroscopy specialist, which presents a difficulty in smaller hospitals where insufficient expertise and experience may result in incorrect diagnoses and mishandled therapies.

To tackle these problems, there has been a notable surge in the investigation of data based on computer-aided diagnosis methods in recent years. This paper presents ID-GCS, a novel intelligent decision-making approach for GC screening. ID-GCS is a data-driven system that extracts literary semantics from gastroscopy reports by utilizing multimodal semantic fusion and a hybrid attention mechanism. This makes it easier to incorporate the semantics of textual gastroscopy reports with visuals, improving the readability of the results. An assessment of ID-GCS using an actual gastroscopy report dataset shows that ID-GCS performs better in GC screening in terms of sensitivity and accuracy than state-of-the-art techniques. This novel strategy not only overcomes the drawbacks of conventional techniques but also represents a substantial advancement in the application of AI for more accurate and dependable stomach cancer diagnosis.

The disadvantages of the already existing system was that there was a small decrease in complexity and fitting for the ID-GCS approach. Therefore, because of its reduced prediction accuracy, real-time applications are not well suited. For the diagnosis of stomach cancer, it is not strongly advised to rely only on the gastroscopy dataset.

IV. Proposed System

The suggested method offers a practical strategy to more accurately forecast the categorization of gastrointestinal anomalies. In this research, the types of diseases will be identified by the use of Deep Learning algorithms, such as CNN. The datasets will be gathered from Kaggle and preprocessed using several methods, including adaptive median filtering. The UNet method is used for segmenting the images after preprocessing. After that, a deep learning algorithm will be used to train these datasets and a model file will be produced. It is able to accurately identify gastrointestinal anomalies such as ulcers, funduses, healthy tissues, and esophagitis when given an input image for prediction. As a result, this approach contributes to a more accurate and efficient detection of gastrointestinal problems than current models. Perks of the proposed system are that it provides an efficient means of diagnosing gastrointestinal abnormalities. Additionally, it facilitates easier prediction for pathologists with accurate diagnosis and the best outcomes are obtained when the deep learning algorithm is used.

V. System Architecture

In the present study, various kinds of gastrointestinal anomalies will be identified by utilizing Deep Learning algorithms, such as CNN. Therefore, gathering the dataset from the Kaggle site shall be the first stage in the paper. After that, it will be divided into both training and testing datasets, with the training dataset being utilized for training the model and the testing dataset being kept separate.

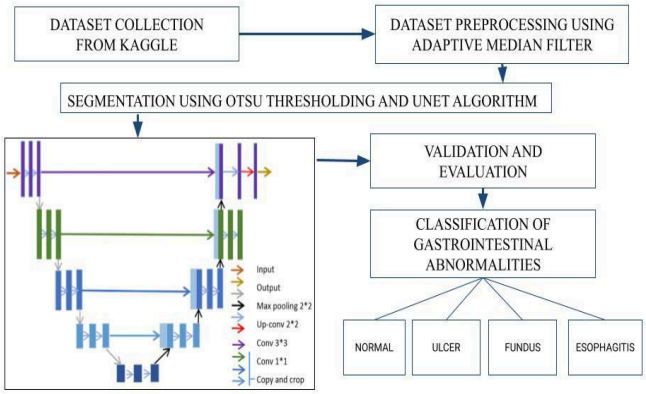


Fig. 1 Proposed System Architecture

After that, various pre-processing techniques, including the adaptive median filter, are applied to these datasets. The photos are segmented using the UNet technique following preprocessing of our dataset. After that, we'll be prepared to use the architecture for training. To train the model, we will use a deep learning algorithm like CNN. It is able to accurately identify gastrointestinal anomalies such as ulcers, funduses, healthy tissues, and esophagitis when provided with an input image for prediction. As a result, this approach contributes to a more accurate and efficient detection of gastrointestinal problems than current models.

VI. Methodology

A. Dataset Collection

In this paper, the data sets will be gathered and given into a deep learning algorithm for training. Accuracy rises with the size of the dataset. The dataset will be gathered via the Kaggle site. A Deep Learning model performance must be continuously improved when it is implemented in an actual-life situation by feeding it new data. Three steps involved in data collection are own data, third party websites and collection from industries.

ISBN 979-8-3503-5094-4

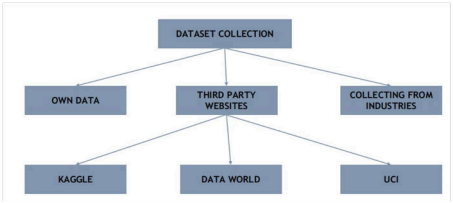


Fig. 2 Dataset Collection

In the deep learning era, data has become a commodity that is so valuable that many start-ups are beginning to provide customized image annotation services, where they collect and classify the data, given an explanation of the kind of data and annotations that are required.

B. Dataset Preprocessing

Once the datasets are gathered, they are preprocessed using various methods, such as the adaptive median filter. An adaptive median filter reduces noise in an image by performing spatial processing. Every single pixel in the image is compared to its surrounding pixels by the filter. A pixel is regarded as noise if its value greatly deviates from that of a majority of the neighboring pixels. Next, the median values of the nearby pixels are used by the filtering algorithm to replace the noisy pixel until every single noise pixel is gone from the image, this process is repeated. To determine which pixels are impacted by impulse noise, the filter applies a spatial method. It is typically used to recover image pixels by eliminating noise without noticeably distorting the image's existing structures. The filter classifies certain pixels as noise by drawing a contrast between each pixel in the image and the pixels that surround it. The neighborhood's size is flexible. Since they are not adjusted to the pixels in the comparable image, pixels that are significantly different from the bulk of their neighbors are classified as impulsive noise.

C. Segmentation

The UNET algorithm is used for segmenting the images after preprocessing. The network's architecture was expanded and changed from the fully convolutional network, on which it was based, in order to produce more accurate segmentations while using fewer training photos. The presence of numerous channel features in the upsampling phase of U-Net is a significant change that enables the network to communicate contextual information to greater resolution layers. This results in a u-shaped design as the expanding path is roughly symmetrical to the contracting portion. The network does not have any fully connected layers, it just utilizes the valid portion of each convolution. The input image is mirrored in order to extrapolate the missing context and estimate the number of pixels in the border region of the image. The contracting path is a standard convolutional network that comprises many convolutions applied, each of which is followed by a max pooling operation and a rectified linear unit (ReLU). While feature information increases during the contraction, spatial information decreases. By using high resolution (HR) features from the contracting path in a series of up-convolutions and concatenations, the expansive pathway integrates both feature and spatial information[16]. In recent days, UNET algorithm is also used for RGB color images[17].

A total of 1955545 parameters (with size range of each layer being 16,32,64,128 and 256) are used consisting of 1955033 trainable and 512 non-trainable parameters

D. Training using Deep Learning Algorithm

Following collection and preprocessing, the dataset is sent into Deep Learning algorithms like CNN for training using the UNET algorithm.

For deep learning techniques, a CNN is a type of network design that is particularly useful for applications involving processing of pixel input, such as image recognition. Another kind of neural network that may extract important information from picture and time series data is the CNN. This makes it extremely useful for applications involving images, like pattern identification, object classification, and picture recognition. A CNN uses matrix multiplication and other linear algebraic concepts to find patterns in an image. The design of a CNN is comparable to the structure of connectivity found in the human brain. A CNN performs better when it comes to image inputs, speech or audio signal inputs, and other input types than the previous networks.

E. Validation and Evaluation

The deep learning algorithm will validate and assess the datasets following training. In the field of deep learning, validation corresponds to approving or verifying a trained model's prediction. On the other hand, assessment in the domain of deep learning refers to testing the deep learning model as a whole and how well it performs under different conditions. It entails evaluating the accuracy of the predictions made in various scenarios as well as the training process of the deep learning model.

F. Disease Prediction

The aim is to forecast the prediction accuracy that would be advantageous for individuals experiencing gastrointestinal irregularities and lower the percentage ratio. Therefore, early disease detection is critical for improving patient outcomes. Finding the best prediction model or Deep Learning approach, to differentiate between patients with gastrointestinal problems and healthy individuals is the primary goal of this research paper. The model that was trained file has been generated after validation and assessment. It can accurately identify gastrointestinal anomalies such as ulcer, fundus, healthy, and esophagitis when an input image is provided for prediction.

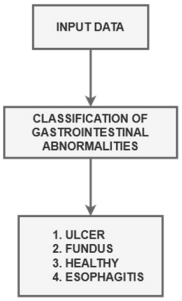


Fig. 3 Disease Prediction

VII. Evaluation Metrics

The UNet model's training output [illegible] 150 epochs is

ISBN 979-8-3503-5094-4

displayed in the figure below.

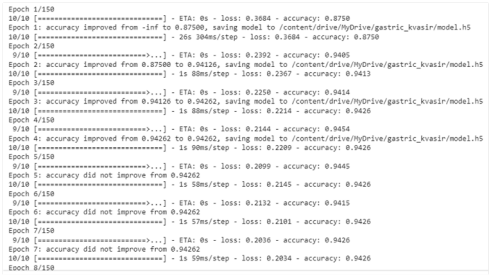


Fig. 4 Training Output of UNet Algorithm

The CNN model's outputs from training across 50 epochs can be seen in the image below.

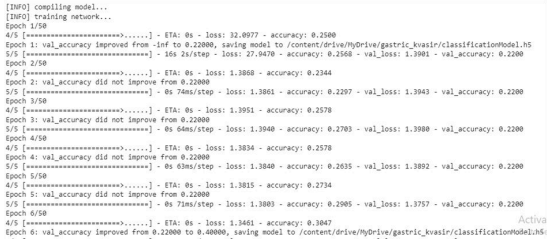


Fig. 5 Training Output of CNN Model

The epochs vs. training and validation accuracy is displayed in the following figure.

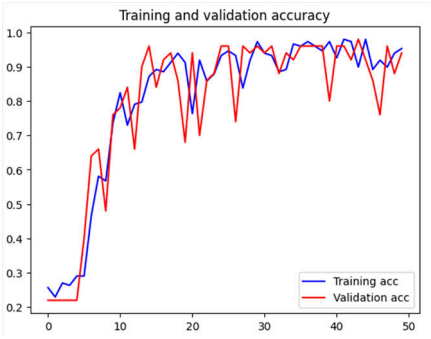


Fig. 6 Epochs vs Training and Validation Accuracy

The figure that follows compares training and validation loss to epochs.

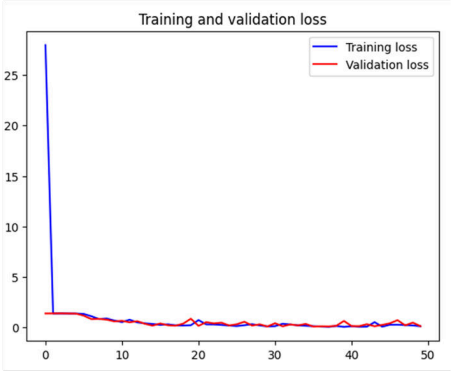


Fig. 7 Epochs vs Training and Validation Loss

The trained model's classification report using test data is displayed in the image below.

	precision	recall	f1-score	support
eosho	1.00	0.94	0.97	16
fundus	0.92	1.00	0.96	11
normal	1.00	1.00	1.00	12
ulcer	1.00	1.00	1.00	11
accuracy			0.98	50
macro avg	0.98	0.98	0.98	50
weighted avg	0.98	0.98	0.98	50

Fig. 8 Classification Report

The trained model's confusion matrix utilizing test data is displayed in the picture underneath.

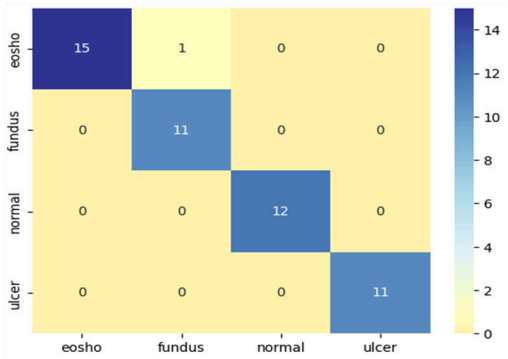


Fig. 9 Confusion Matrix

Reactjs has been used to create a web application that collects user input and displays the anticipated outcome. The image below shows the application's login page.

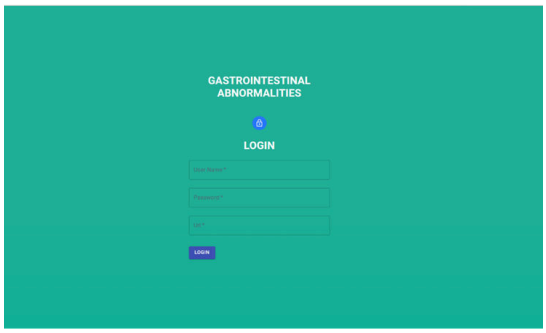


Fig. 10 Login Page

After successfully logging in, the user is able to upload an image for prediction. It is depicted in the picture below.

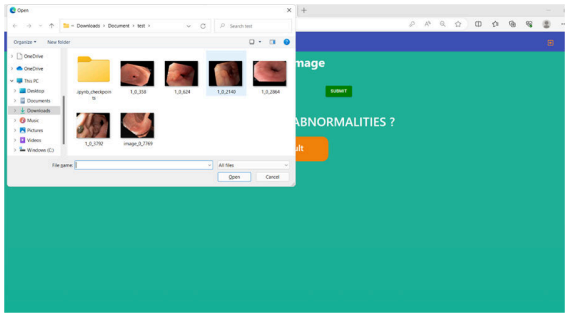


Fig. 11 Uploading Image

The prediction process will begin after a brief loading time, and the outcome will appear.

ISBN 979-8-3503-5094-4

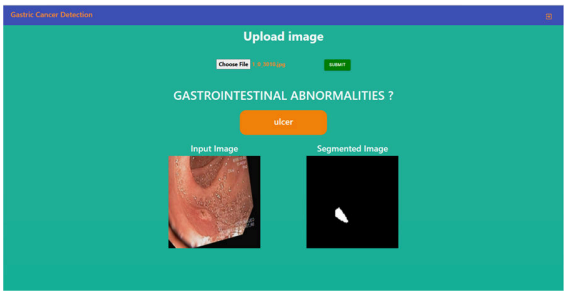


Fig. 12 Prediction of Ulcer

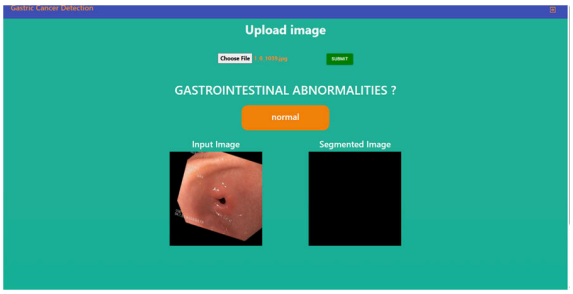


Fig. 13 Normal Result

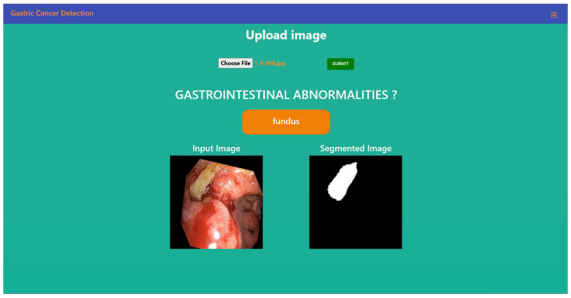


Fig. 14 Fundus Prediction

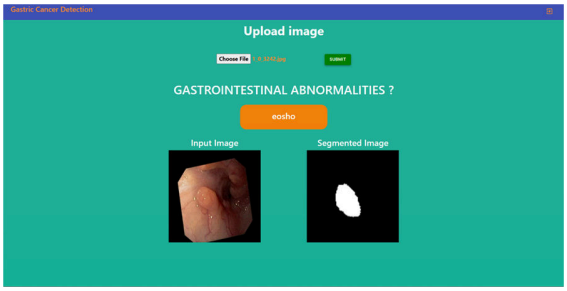


Fig. 15 Eosho Prediction

VIII. Conclusion

In this paper, makesense.ai is used to annotate the dataset. UNet will then be used to partition it. After that, the CNN method is once more used to train the segmented images for classification. After 50 training epochs, the CNN model achieved 98% accuracy. The trained CNN model successfully segmented the affected region using the UNet algorithm and predicted the stomach cancer of eosho, fundus, normal, and ulcer. Using a deep learning system, the research has been successfully deployed to anticipate gastrointestinal disorders more accurately. The sorts of gastrointestinal problems are identified using algorithms like CNN. The best method for ensuring accuracy in the output is deep learning, which produces guessing of the greatest caliber. There is a great deal of room for technological advancement because there are numerous ways to monitor people with illnesses.

References

[1] Shuai Ding; Shikang Hu; Xiaojian Li; Youtao Zhang; Desheng Dash Wu, "Leveraging Multimodal Semantic Fusion for Gastric Cancer Screening via Hierarchical Attention Mechanism", IEEE Transactions on Systems, Man, and Cybernetics: Systems

[2] SeongRyeol Moon; Curt Balch; Sungjin Park; Jinhyuk Lee; Jiyong Sung; Seungyoon Nam, "Systematic Inspection of the Clinical Relevance of TP53 Missense Mutations in Gastric Cancer", IEEE/ACM Transactions on Computational Biology and Bioinformatics [Vol no: 16, 2019]

[3] Shaolong Shi , Student Member, IEEE, Yifan Chen , Senior Member, IEEE, and Xin Yao, Fellow, IEEE, "NGA-Inspired Nanorobots-Assisted Detection of Multifocal Cancer", IEEE TRANSACTIONS ON CYBERNETICS

[4] Seungwoo Song, Student Member, IEEE, Jukwan Na, MoonHyung Jang, Student Member, IEEE, Hyeyeon Lee, Student Member, IEEE, Hye-Soo Lee, Yong-Beom Lim, Heonjin Choi, and Youngcheol Chae, Senior Member, IEEE, "A CMOS VEGF Sensor for Cancer Diagnosis Using a Peptide Aptamer-Based Functionalized Microneedle", IEEE TRANSACTIONS ON BIOMEDICAL CIRCUITS AND SYSTEMS [Vol no: 13, 2019]

[5] Chih-Hung Chan; Tze-Ta Huang; Chih-Yang Chen; Chien-Cheng Lee; Man-Yee Chan; Pau-Choo Chung, "Texture-Map-Based Branch-Collaborative Network for Oral Cancer Detection", IEEE Transactions on Biomedical Circuits and Systems [Vol no: 13, 2019]

[6] Jean-Sébastien Boisvert; Julie Lafontaine; Audrey Glory; Sylvain Coulombe; Philip Wong, "Comparison of Three Radio-Frequency Discharge Modes on the Treatment of Breast Cancer Cells in Vitro", IEEE Transactions on Radiation and Plasma Medical Sciences [Vol no: 4, 2020]

[7] Xiwei Tang; Qiu Xiao; Kai Yu, "Breast Cancer Candidate Gene Detection Through Integration of Subcellular Localization Data With Protein–Protein Interaction Networks", IEEE Transactions on NanoBioscience [Vol no: 19, 2020]

[8] Ai-Min Yang, Yang Han*, Chen-Shuai Liu, Jian-Hui Wu, Dian-Bo Hua, "D-TSVR Recurrence Prediction Driven by Medical Big Data in Cancer", IEEE Transactions on Industrial Informatics [Vol no: 17, 2020]

[9] Uttam M. Pal; Anil Vishnu GK; Gayatri Gogoi; Saeed Rila; Saahil Shroff; Gokul AM; Pronami Borah; Manoj Varma, "Towards a Portable Platform Integrated With Multispectral Noncontact Probes for Delineating Normal and Breast Cancer Tissue Based on Near-Infrared Spectroscopy", IEEE Transactions on Biomedical Circuits and Systems [Vol no: 14, 2020]

[10] Efthymios P. Papageorgiou; Bernhard E. Boser; Mekhail Anwar, "Chip-Scale Angle-Selective Imager for In Vivo Microscopic Cancer Detection", IEEE Transactions on Biomedical Circuits and Systems [Vol no: 14, 2020]

[11] Yi Wang; Na Wang; Min Xu; Junxiong Yu; Chenchen Qin; Xiao Luo; Xin Yang; Tianfu Wang; Anhua Li; Dong Ni, "Deeply-Supervised Networks With Threshold Loss for Cancer Detection in Automated Breast Ultrasound", IEEE Transactions on Medical Imaging [Vol no: 39, 2020]

[12] Xi Wang; Hao Chen; Caixia Gan; Huangjing Lin; Qi Dou; Efstratios Tsougenis; Qitao Huang; Muyan Cai, "Weakly Supervised Deep Learning for Whole Slide Lung Cancer Image Analysis", IEEE Transactions on Cybernetics [Vol no: 50, 2020]

[13] Yujie Feng; Fan Yang; Xichuan Zhou; Yanli Guo; Fang Tang; Fengbo Ren; Jishun Guo; Shuiwang Ji, "A Deep Learning Approach for Targeted Contrast-Enhanced Ultrasound Based Prostate Cancer Detection", IEEE/ACM Transactions on Computational Biology and Bioinformatics [Vol no: 16, 2019]

[14] Bo Fu; Pei Liu; Jie Lin; Ling Deng; Kejia Hu; Hong Zheng, "Predicting Invasive Disease-Free Survival for Early Stage Breast Cancer Patients Using Follow-Up Clinical Data", IEEE Transactions on Biomedical Engineering [Vol no: 66, 2019]

[15] Man-Sun Kim; Dongsan Kim; Jeong-Rae Kim, "Stage-Dependent Gene Expression Profiling in Colorectal Cancer", IEEE/ACM Transactions on Computational Biology and Bioinformatics [Vol no: 16, 2019]

[16] Siddique, Nahian, et al. "U-net and its variants for medical image segmentation: A review of theory and applications." Ieee Access 9 (2021): 82031-82057.

[17] Gnanapriya, S., and K. Rahimunnisa. "A Hybrid Deep Learning Model for Real Time Hand Gestures Recognition." Intelligent Automation & Soft Computing [Vol no:36,2023]

ISBN 979-8-3503-5094-4

Structural MR Brain Image based diagnosis of Alzheimer's disease using Extreme Learning Machine

Janavi T S
Biomedical Engineering
Anna University
Chennai, India
janavi.ts06@gmail.com

Keerthana V
Biomedical Engineering
Anna University
Chennai, India
keeruv2002@gmail.com

Ayesha Jumana
Biomedical Engineering
Anna University
Chennai, India
ayeshajumana02@gmail.com

Dr. G. Kavitha
Biomedical Engineering
Anna University
Chennai, India
kavithag_mit@annauniv.edu

***Abstract*— Alzheimer's disease is a neurodegenerative disorder that results in memory degradation and impairment of cognitive function in about 55 million people (WHO 2023 reports), mostly targeting the elderly community. There is a rising need for early diagnosis in order to achieve better treatment outcomes. This study proposes an extreme learning machine (ELM) classifier to differentiate between various stages of Alzheimer's disorder. It is carried out using selectively extracted brain regions from structural MR brain images. The diverse datasets from Alzheimer's Disease Neuroimaging Initiative (ADNI) [2200 patient data] fortifies the approach of this study. Preprocessing and skull stripping the MR brain images, segmentation of the desired bio markers are done. Essential geometric features and textural features are extracted for different stages of Alzheimer's disorder namely Mild cognitive impairment (MCI), Late mild cognitive impairment (LMCI) and Alzheimer's disorder (AD). Extreme learning machine (ELM) classifier is utilized to discern the various stages. Once the model is trained and validated, it can be used to predict the Alzheimer's disease stage of new, unseen data. The model's output could indicate the likelihood or stage of Alzheimer's disease based on the input features. This study aims to contribute to the development of Alzheimer's disorder diagnostics and emphasize on the potential of using machine learning techniques for analysis of structural MR brain images.**

Keywords—Alzheimer's, MR Brain image, MCI, LMCI Extreme Learning Machine,

I. Introduction

Alzheimer's disease (AD) is the most common cause of dementia among older adults. It is characterized by the accumulation of abnormal protein such as beta-amyloid plaques and tau tangles in the brain. These deposits lead to the inconspicuous death of nerve cells, resulting in the loss of cognitive function. Biomarkers are measurable indicators of progression of a disease. Biomarkers play a pivotal role in diagnosis and monitoring disease progression. The progression goes from a healthy person to Mild cognitive impairment (MCI) to Late mild cognitive impairment (LMCI) and fully fledged Alzheimer's disease (AD). Magnetic Resonance Imaging (MRI) is a useful tool in the prediction and diagnosis of Alzheimer's disease (AD). MRI provides detailed images of the brain's structure that aids in identifying atrophy in several parts of the brain and can be used to measure the thickness of the cerebral cortex. Changes in Corpus Callosum (CC) thickness, surface area, CC length gives significant information for the analysis of Alzheimer's. The lateral ventricles are fluid-filled cavities within the brain. The lateral ventricles can be observed in the axial axis of the MRI, and is considered one of the structural change's indicators of brain atrophy in Alzheimer's disease. MRI brain images of every stage of progression is taken and skull stripped. It is done using a tool called BrainSuite23a. It excels in visualizing brain structures in 3D and makes the process of skull stripping efficient and easy. To ensure the accuracy of the Brain Surface Extractor (BSE) skull-stripped data, the raw NFBS dataset has been processed for other classes of the disease. Biomarkers like the corpus callosum and lateral ventricles provide valuable insights into reduced metabolic activity and morphometric changes associated with neurodegenerative processes. Fuzzy c means clustering creates fuzzy partitions of data and aids in segmentation. It is a flexible and a soft clustering mechanism. The segmented biomarkers are validated using MIPAV software. Feature extraction reveals abnormalities in the biomarkers. Textural features, encompassing measures such as Haralicks's features, contrast, correlation, entropy, and variance, provide insights into the intensity variations and spatial relationships among pixels within the brain structures. Geometrical features involve shape-related characteristics like area, perimeter, and eccentricity, offering information about the structural morphology of the corpus callosum and lateral ventricle. The derived features are fed into a classifier extreme learning machine (ELM). ELM is simple and efficient in training single-hidden-layer feedforward neural networks. Unlike traditional neural network training methods, ELM randomly assigns input weights and analytically determines output weights, eliminating the need for iterative weight adjustments. This leads to faster training times, making ELM particularly suitable for applications where efficiency is crucial.

II. Materials and Methods

A. Database

The structural MR images considered for this analysis were obtained from the Image and Data Archive, a secure online resource for neuroscience data, comprising The Alzheimer's Disease Neuroimaging Initiative (ADNI) database. About 200-250 datasets have been taken for this research. The meticulous collection of such comprehensive data allows for in-depth analyses, fostering groundbreaking discoveries and innovations in Alzheimer's research. ADNI's longitudinal design enables the exploration of disease progression over time, offering valuable insights

ISBN 979-8-3503-5094-4

into early biomarkers, and potential therapeutic targets. The rigorous data quality control measures implemented by ADNI ensure reliability and reproducibility, further solidifying its status as a gold standard in Alzheimer's research.

B. Preprocessing

Preprocessing of MRI (Magnetic Resonance Imaging) encompasses noise reduction techniques and bias field correction to compensate for non-uniformities in image intensity. Subsequent stages involve skull stripping to isolate the brain from surrounding tissues, spatial normalization for consistent spatial alignment across subjects. The skull stripping process involves the removal of non-brain tissues, such as the skull, scalp, and other extracranial structures, to isolate the brain region, allowing for a more focused and accurate analysis of brain structures. The Brain Surface Extractor (BSE) is a specialized tool designed for the removal of extra-meningeal tissues from MRI head volumes by Anisotropic Diffusion Filter smoothing low amplitude differences while retaining high-contrast edges. Diffusion iterations and Diffusion constant offer flexibility in tailoring the filter to the characteristics of the data. Subsequently, Edge Detection is executed using the 3D Marr- Hildreth method, incorporating Gaussian blur and Laplacian operators. The Edge constant is set between 0.5 and 1.0, controlling the sensitivity of edge detection. The Erosion size, set to 2, controls the extent of this morphological operation. Finally, the Extract Brain stage expands the mask to encompass the entire brain and dilates the final mask.

The Neurofeedback Skull-stripped (NFBS) repository, a database of T1-weighted anatomical MRI scans that have undergone manual skull-stripping is used for verification of the efficacy of the BSE comparing results of skull stripping of the raw NFBS dataset against the manually skull-stripped data from the NFBS repository, employing the Dice Coefficient method for rigorous validation. Specifically designed to evaluate the performance of algorithms in comparison to manually generated ground truth, the Dice Coefficient provides a precise and quantitative measure of agreement.

The formula for the Dice coefficient (DSC) is given by

$$\text{DSC}= \frac{2 \times |X \cap Y|}{|X|+|Y|} \quad (1)$$

where X and Y are the sets representing the segmented regions in the predicted and ground truth images, respectively. A high Dice coefficient value signifies a substantial similarity between the predicted and ground truth masks, indicating excellent performance of the skull stripping model or algorithm.

C. Segmentation

Importance to segmentation of the Corpus Callosum (CC) and Lateral Ventricles (LV) is given in this project, as these regions hold paramount importance for early diagnosis and cognitive function assessment. Accurate segmentation of these biomarkers not only reveals structural alterations but also provides valuable insights into reduced metabolic activity and morphometric changes associated with neurodegenerative processes. The Fuzzy c-means (FCM) algorithm is a notable unsupervised clustering technique employed in the construction of fuzzy partitions from data, particularly pertinent in the context of brain MRI image segmentation (B-MRI-IS). The efficacy of FCM relies on a parameter denoted as "m," which corresponds to the degree of fuzziness inherent in the solution. This parameter shapes the construction of a fuzzy partition based on the initial class centroids.

a) Extracting middle slice from 3D image : The middle slice index is calculated to ensure uniformity in the segmentation process. In cases where the number of slices is odd, the middle slice is directly selected. Conversely, if the number of slices is even, the middle slice is determined by the total slice count divided totally by 2. These middle slices, extracted along the sagittal or axial axis depending on the segmented part, are normalized, rotated, and saved as individual .png files with appropriate filenames.

b) Applying FCM : The Fuzzy C-Means (FCM) clustering algorithm is employed for the segmentation of grayscale images. FCM is a soft clustering technique that assigns each pixel a membership grade for each cluster, reflecting the degree of belongingness to that cluster.

$$J_m(U,C) = \sum_{i=1}^{c} \sum_{j=1}^{n} u_{ij}^{m} \cdot \left| x_j - c_i \right|^2$$

(2)

Where,

$Jm(U,V)$ is the objective function along with U being the membership matrix where uij represents the degree of membership of data point xj to cluster i and C being the matrix of cluster centers, where ci is the center of cluster i.

m is the fuzziness parameter (usually set to 2 for balanced results), n is the number of data points and c is the number of clusters.

$\|xj-ci\|$^2 is the Euclidean distance between data point xj and cluster center ci.

The update membership matrix and update cluster center C membership values are updated iteratively using the following formulae respectively:

$$u_{ij} = \frac{1}{\sum_{k=1}^{c} \left(\frac{|x_j - c_i|}{|x_j - c_k|}\right)^{\frac{2}{m-1}}}$$

(3)

$$c_i = \frac{\sum_{j=1}^{n} u_{ij}^{m} \cdot x_j}{\sum_{j=1}^{n} u_{ij}^{m}} \quad (4)$$

After obtaining the membership grades for each pixel, the removal of small connected components, clearing borders and filling holes in regions is done.

c)Connected Component Analysis: Connected Component Analysis (CCA) is a fundamental image processing technique used to identify and label connected regions in a binary image. In this context, the binary image represents the segmented cluster, and the objective is to isolate the corpus callosum from other connected components. Each segmented cluster image is read as a binary mask, where pixel values represent membership to the cluster (white) or background (black). The algorithm employs the cv2.connectedComponents function to label connected components in the binary image. The labeled

ISBN 979-8-3503-5094-4

image is processed to create a binary mask specifically for the corpus callosum. In this case, the corpus callosum is assumed to be represented by the label "1."

The validation process involves a meticulous comparison between the segmentation outcomes and the ground truth Volume of Interest (VOI). MIPAV facilitates this comparison by providing tools for converting masks generated by segmentation algorithms into VOIs, aligning them with the ground truth data. The integration of MIPAV into the validation workflow enhances the efficiency and precision.

D. Feature Extraction

The extraction of textural and geometrical features from the corpus callosum and lateral ventricle regions in brain MRI scans plays a pivotal role in understanding structural abnormalities and aiding in clinical diagnoses of Alzheimer's disease. The corpus callosum (CC) and the lateral ventricles, vital structures connecting the brain's hemispheres, are characterized by a diverse set of quantitative features. These include Perimeter, embodying the total boundary length , offering insights into its geometric shape. Kurtosis, a statistical metric, delineates the distribution shape's peakedness or flatness. Skeleton Length gauges the cumulative distance of all skeletal elements within the structure, reflecting its overall size. The Surface encapsulates the external structure's total area, while Splenium Thickness measures the posterior part's width. Length quantifies the longitudinal dimension, crucial for understanding interhemispheric connectivity. AP Area Difference highlights the anterior-posterior area discrepancy, emphasizing the CC's structural asymmetry. Textural features, derived from the gray-level co-occurrence matrix, include Contrast, measuring local intensity variations; Correlation, assessing linear dependencies; Entropy, indicating disorder; IDM, reflecting pixel intensity homogeneity; Variance, quantifying intensity spread; and Difference Variance, capturing variability in intensity differences between adjacent pixels. These features collectively provide a comprehensive characterization of the corpus callosum's structural and textural attributes**.**

E. Extreme Learning Machine

An Extreme Learning Machine (ELM) class is defined, which initializes with the input size, number of neurons, and activation function. Functions are defined to extract features from images. Images from specific folders, extracts features, and assigns labels are loaded. Data and labels from different datasets are combined into all_data and all_labels. Labels are encoded. Data is normalized. The data is split into training and testing sets. The ELM model is initialized with specified input size, number of neurons, and activation function. The model is trained on the training data using the fit method.

The trained model is used to make predictions on the training set.The model is tested on the testing set, and accuracy, classification report, and confusion matrix are printed and plotted for the testing set. The hidden layer output is calculated using the sigmoid activation function where X is the input data, *input_weights* are the randomly initialized input weights, and *bias* is the bias term.

$$\frac{1}{1+e^{-(X\cdot input\_weights+bias)}} \quad (5)$$

Output weights are found with the Moore-Penrose pseudoinverse of the hidden layer output

$$output\_weights = (hidden\_output^{T})^{+} \cdot y \quad (6)$$

Here, y represents the true labels in a one-hot encoded format. The prediction of the output is obtained by multiplying the hidden layer output by the computed output weights:

$$predictions = hidden\_output \cdot output\_weights \quad (7)$$

III. Results and Discussions

The MR images obtained from the ADNI dataset were subjected to skull stripping followed by validation along with segmenting the corpus callosum and the lateral ventricles being very important biomarkers and extracting their features which are fed into an Extreme Learning Machine Classifier which classifies the input MR image into MCI , LMCI or AD.

A. Preprocessing

BSE is designed to preserve important brain tissue while removing non-brain structures effectively as shown in Fig 1 and it surpasses thresholding methods in skull-stripping by providing a more flexible and adaptive approach, ensuring better accuracy, and preserving critical brain tissue during the extraction process.One prominent observation is the loss of gray matter density as seen in Fig 2, indicating a reduction in the number or size of neurons. This finding aligns with the known pattern of neuronal loss associated with the progression of cognitive decline. Hippocampal atrophy, is a well-established marker of Alzheimer's disease.The hippocampus, a crucial region for memory and spatial navigation, experiences significant volume reduction as the disease advances, contributing to memory deficits.

The graph for dice coefficient is a linear graph and the dice coefficient value is 0.96 which proves the Brain Suite 23a software to be around 96% efficient. This positive linear relationship indicates that the segmentation performance improves consistently across the range of conditions being compared. A Dice coefficient of 0.96 is quite high and indicates that 96% of the pixels in the predicted segmentation overlap with those in the ground truth segmentation.

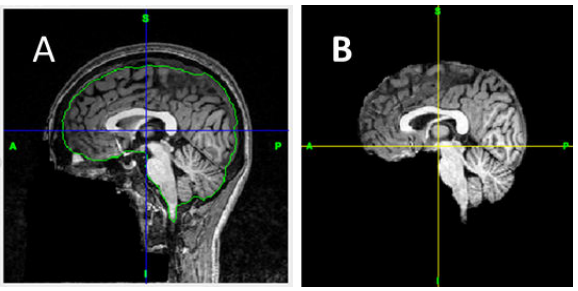


Fig 1 Results of skull stripping using Brain Suite 23 a

ISBN 979-8-3503-5094-4

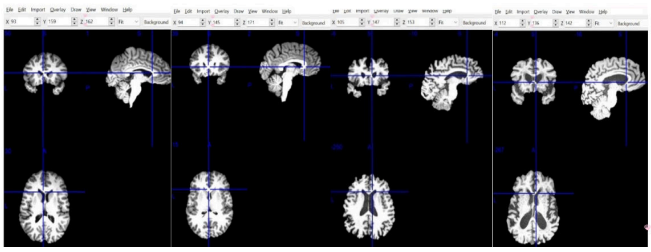


Fig 2 Results of skull stripping depicting structural change

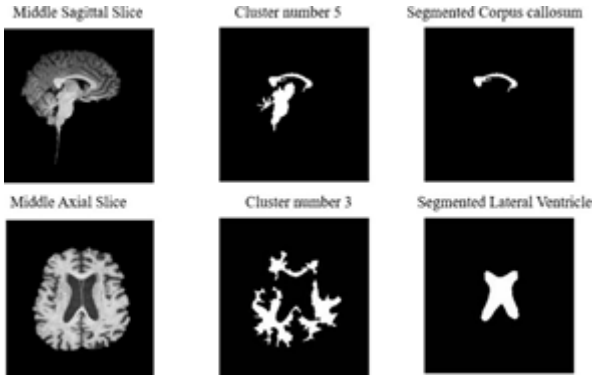


Fig 3 Results of Segmentation of Corpus Callosum and Lateral Ventricle

B. Segmentation

As seen in Fig 3, the Sagittal and axial views are selected and FCM is applied and segmented to obtain the corpus callosum and lateral ventricles. As Alzheimer's disease (AD) progresses, structural changes, including atrophy, occur in the corpus callosum as shown in Fig 4, characterized by a gradual reduction in size and volume. This degenerative process is closely linked to the widespread neuronal loss and synaptic dysfunction characteristic of AD.

The loss of white matter fibers and structural integrity within the corpus callosum further underscores the impact of AD on this crucial brain region.

The enlargement of these ventricles as shown in Fig 4, signifies a disruption in the normal balance of cerebrospinal fluid dynamics. The enlargement of lateral ventricles is often associated with the loss of brain tissue, particularly in regions vulnerable to AD-related pathology.

The validation results proved to be a score of 95% for the Corpus Callosum segmentation while for the lateral ventricle, 93% has been obtained as shown in figures and

In Alzheimer's disease (AD), the pattern of atrophy in the corpus callosum is not uniform, and there is a recognized tendency for more pronounced atrophy in the anterior region compared to the posterior as the disease progresses as shown in Fig 5. This has been tabulated in table 1. This is believed to be associated with the selective vulnerability of certain brain regions to the pathological

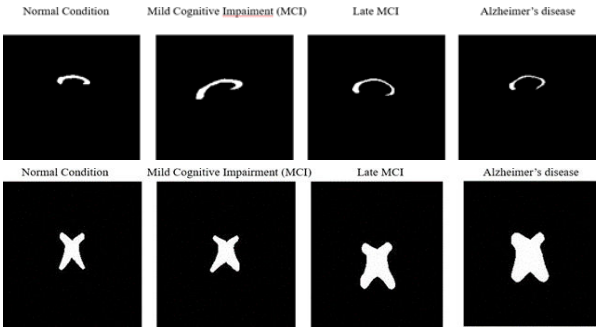


Fig 4 Segmented corpus callosum and lateral ventricle classes of Alzheimer's Disease

Fig 5 Results of Anterior and Posterior area of corpus callosum

TABLE I. DIFFERENCE BETWEEN ANTERIOR AND POSTERIOR AREA OF CORPUS CALLOSUM

	Area in pixels		
	Average posterior area	*Average anterior area*	*Difference in area*
Normal	290.29	264.26	26.03
MCI	276.09	233.97	42.12
LMCI	210.25	156.66	53.59
AD	187.28	110.90	76.38

process's characteristic of the disease. The frontal lobes, housed in the anterior part of the brain, are known to be particularly susceptible to neurodegeneration in AD.

C. Feature extraction

As Alzheimer's disease (AD) progresses, geometric features of the corpus callosum (CC) undergo significant changes as plotted in the heatmap Fig 6. The perimeter, total volume, CC surface, and CC length decrease, reflecting structural alterations in the neural pathway. These reductions suggest atrophy and loss of neural connectivity, common characteristics of AD. Kurtosis, a measure of the distribution's shape, increases in the CC of individuals with AD. This change in kurtosis indicates a shift towards a more peaked or sharp distribution of voxel intensities in the CC, potentially highlighting alterations in tissue density or microstructure associated with AD. In the progression of Alzheimer's disease (AD), observable changes in ventricular morphology serve as crucial markers as depicted in Fig 7. The reduction in principal components of the lateral

ISBN 979-8-3503-5094-4

ventricles reflects a loss of structural complexity, indicative of atrophy and alterations in surrounding brain regions. Simultaneously, there is a notable increase in the volume of both left and right lateral ventricles, highlighting the overall expansion of these structures as AD advances. The ventricular antero-posterior ratio (APR) experiences a significant rise, emphasizing changes in length. Particularly, the volume increase in the inferior part of the lateral ventricle suggests region-specific degeneration, offering insights into the localized vulnerability of the brain in AD. In the normal condition, certain features are less in number and more concentrated in a specific range. In MCI and LMCI there are a greater number of higher feature values and in AD, there is a widespread range of Haralicks texture features as shown in Fig 8 and 9. Texture analysis of the corpus callosum in the context of Alzheimer's disease (AD) reveals significant alterations in various Haralick's features.

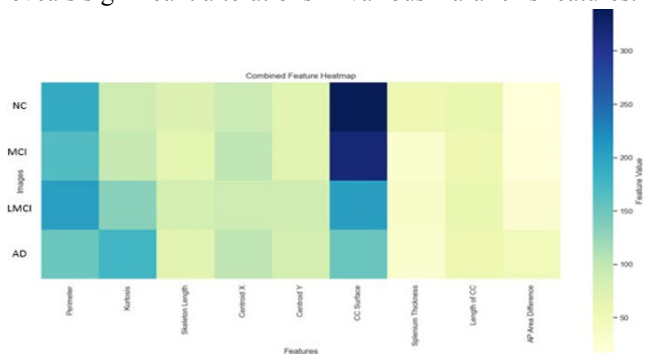


Fig 6 Heatmap of the extracted features of corpus callosum

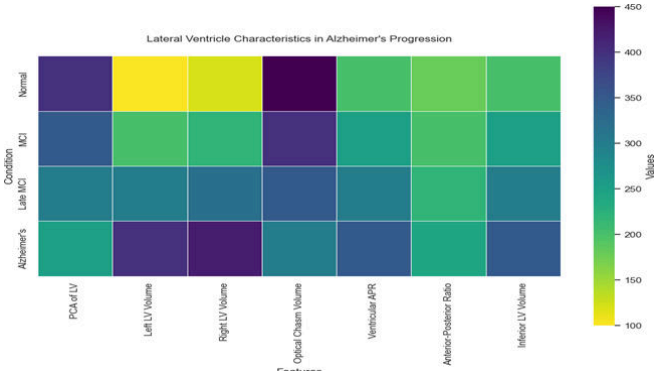


Fig 7 Heatmap of the extracted features of lateral ventricle

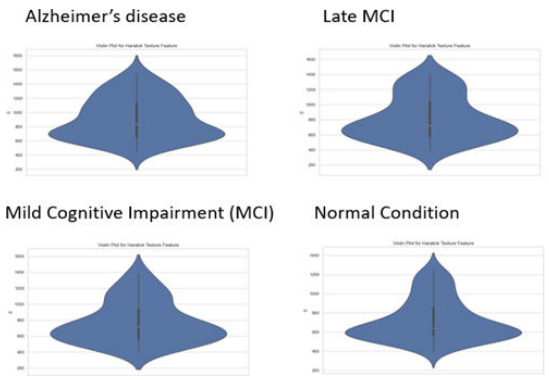


Fig 8 Comparison of textural features of Corpus callosum

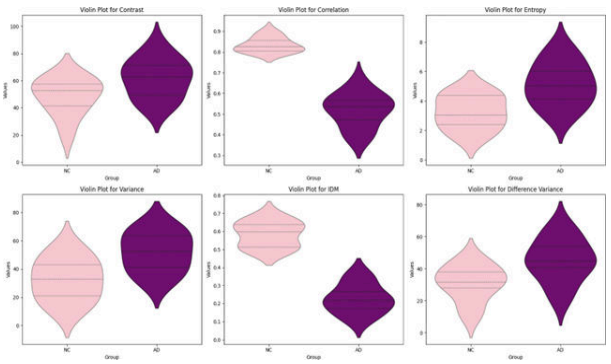


Fig 9 Violin plot of individual Haralicks feature

Conversely, there is a noticeable decrease in correlation, signifying a less ordered or more chaotic texture within the corpus callosum. Entropy shows an increase in AD cases, pointing to a more disordered texture compared to normal controls, while heightened variability in variance suggests structural heterogeneity or degeneration within the neural tissue. A decrease in Inverse Difference Moment (IDM) implies less homogeneity or more heterogeneity in texture within the corpus callosum, and an increase in difference variance indicates heightened variability in intensity differences, reflecting changes in structural characteristics.

D. Extreme learning machine

In this implementation, the confusion matrices visually represent the true positive, true negative, false positive, and false negative predictions, providing a detailed breakdown of the classification results for each cognitive state as seen in Fig 10 and 11. The diagonal elements are darker, indicating they hold more number of values. According to this training set matrix, out of the total 414 images being input, the model has been trained and has correctly predicted 127 Alzheimer's and 125 as LMCI and 129 MCI true to their value bringing the accuracy to 0.92%. Following this, in the testing set which is 20 % of the total dataset, the model has successfully predicted all 3 classes using the extracted features being given.The respective classification report is given in tables II and III. This classifier has a total testing accuracy of 88%. And the train Accuracy for training set is 0.92028985507 and test is 0.8846151538461.

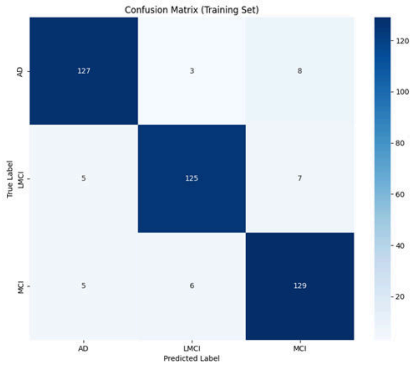


Fig 10 Classification report of training set

ISBN 979-8-3503-5094-4

TABLE II. CLASSIFICATION REPORT OF TRAINING SET

	Classification metrics		
	Precision	***Recall***	***F1-Score***
AD	0.93	0.92	0.92
LMCI	0.93	0.92	0.92
MCI	0.90	0.92	0.91
Macro Avg	0.92	0.92	0.92
Weighted Avg	0.92	0.92	0.92

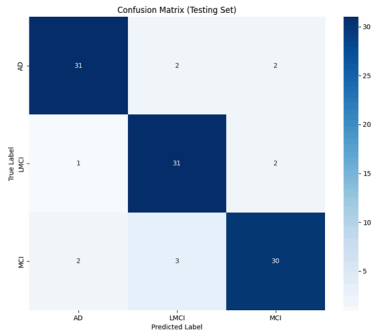


Fig 11: Classification Report of Testing Set

IV. CONCLUSION

This comprehensive study aims towards early Alzheimer's disease diagnosis, using Extreme Learning Machine (ELM) and its variants delving into the intricate analysis of structural MR brain images, extracting these key biomarkers like the corpus callosum and lateral ventricles. The segmentation of these biomarkers using Fuzzy C-Means (FCM) is highlighted for its efficacy in revealing nuanced structural differences in MRI images across the spectrum of Alzheimer's disease stages, spanning from normal to Mild Cognitive Impairment (MCI), Late MCI (LMCI), and finally, Alzheimer's Disease (AD).

The study places particular emphasis on the significance of early detection, asserting that the proposed methodology, with its integration of advanced machine learning techniques, precise biomarker extraction, and meticulous geometric feature analysis, holds the potential to significantly improve treatment outcomes for this complex neurodegenerative disorder.

TABLE III. CLASSIFICATION REPORT OF TEST SET

	Classification metrics		
	Precision	***Recall***	***F1-Score***
AD	0.91	0.89	0.90
LMCI	0.86	0.91	0.88
MCI	0.88	0.86	0.87
Macro Avg	0.88	0.89	0.88
Weighted Avg	0.88	0.89	0.88

REFERENCES

1. Dan Pan, An Zeng et al. Early Detection of Alzheimer's Disease Using Magnetic Resonance Imaging: A Novel Approach Combining Convolutional Neural Networks and Ensemble Learning. Front. Neurosci., 2020. Volume 14.
2. L. Sørensen et al. Differential diagnosis of mild cognitive impairment and Alzheimer's disease using structural MRI cortical thickness, hippocampal shape, hippocampal texture, and volumetry. Neuroimage Clinical 2017; 13: 470–482
3. Neha Garg et al. A review on Alzheimer's disease classification from normal controls and mild cognitive impairment using structural MR images. Journal of Neuroscience methods. 2023 Volume 384,109745.
4. Bin Lu, Hui-Xian Li, Zhi-Kai Chang et al. A practical Alzheimer's disease classifier via brain imaging-based deep learning on 85,721 samples. Journal of Big Data 2022. Volume 9, Article number: 101
5. Taeho Jo et al. Deep Learning in Alzheimer's Disease: Diagnostic Classification and Prognostic Prediction Using Neuroimaging Data . Alzheimer's Disease and Related Dementia 2019. Volume 11
6. Iman Beheshti, Hasan Demirel et al. Feature-ranking-based Alzheimer's disease classification from structural MRI . Magnetic Resource Imaging 2016.vol.34, no.3, pp. 252-263.

ISBN 979-8-3503-5094-4

Tenth International Conference on Biosignals, Images and Instrumentation, SSNCE, Chennai, March 20 – 22, 2024

TRACK 4

Deep Learning Model for Edge Devices for COVID-19 Detection from CXR Images

1st Umar Farooq
Department of Electrical Engineering
ZHCET, Aligarh Muslim University
Aligarh, India
farooqumar0920@gmail.com

2nd Shahid Kamal
Department of Electrical Engineering
ZHCET, Aligarh Muslim University
Aligarh, India
kamalshahid20@gmail.com

3rd Mohammad Sarfraz
Department of Electrical Engineering
ZHCET, Aligarh Muslim University
Aligarh, India
msarfraz@zhcet.ac.in

***Abstract*—The novel coronavirus 2019 (COVID-19) emerged in December 2019, and subsequently underwent rapid global spread, culminating in a pandemic declaration. The emergence of COVID-19 has affected public health significantly and sparked major socio-economic crises. Worldwide, more than 6 million deaths linked to COVID-19 and more than 758 million cases have been reported. Prompt and precise identification of individuals as either healthy or infected is a crucial strategy in managing and preventing the spread of coronavirus outbreaks. Recent findings suggest that patients with COVID-19 infection can be identified by chest radiography anomalies. Chest X-ray (CXR) pictures are more easily obtainable when compared to chest computed tomography (CT) images, especially in less economically developed regions where CT machines are prohibitively costly. In this research, we introduce an exceptionally lightweight convolutional neural network (CNN) tailored for the automated assessment of chest X-ray images, aiming to differentiate between COVID-19, non-COVID-19, and normal conditions. Our model has total of 59,467(0.059 million) parameters, our model's total size is 19.67 MB. Our model surpasses other COVID-19 detection models due to its substantially reduced parameter count and comparable accuracy, making it well-suited for use on machines with limited computational capabilities.**

***Index Terms*—COVID-19, Chest X-Ray, Deep Learning, Edge Devices, Convolutional Neural Network**

I. Introduction

The COVID-19 pandemic, caused by the novel coronavirus 2019 (SARS-CoV-2), is an acute respiratory disease that has resulted in more than 6.8 million fatalities and over 758 million confirmed cases worldwide. The World Health Organization (WHO) has officially declared COVID-19 a contagious illness. It was first detected in December 2019 in Wuhan City, situated in China's Hubei province. [1]. The virus can be transmitted when people breathe in droplets that are expelled from the mouth or nose of an infected person while they cough, sneeze, talk, or simply breathe. It takes between 1 to 14 days for symptoms to appear after exposure. Typical signs of infection encompass a cough, fever, headache, tiredness, breathing challenges, and a diminished sense of smell and taste, though these manifestations can diverge among individuals [2]. In more critical instances, symptoms can escalate to encompass difficulty breathing, decreased oxygen levels, respiratory failure, and a state of shock. On January 30, 2020, and March 11, 2020, the World Health Organization (WHO) officially designated the COVID-19 outbreak as a pandemic and a global public health crisis, respectively. Since then, rapid detection of the virus has become a highly significant area of real-time research interest. Individuals characterized by compromised immune function or advanced age exhibit heightened susceptibility to developing severe complications, including cardiac and renal failure, along with septic shock. This epidemic is still having a devastating effect on global health and wellbeing. The formulation of an effective classification system is a vital phase in the COVID-19 battle cycle since it allows patients to begin receiving immediate medical care, treatment, and control transmission.

The reverse transcriptase-polymerase chain reaction (RT-PCR) plays a vital role as a screening technique to identify the presence of SARS-CoV-2 [3] and COVID-19. While the RT-PCR test is the most often used approach for detecting COVID-19, it has some drawbacks. The RT-PCR technique is difficult and time-consuming, it suffers from three major issues: 1. A scarcity of RT-PCR kits. 2. Community hospitals located in rural areas do not have the necessary PCR infrastructure to handle a high volume of samples. 3.To conduct RT-PCR, it is necessary for the samples obtained to contain detectable amounts of SARS-CoV-2. [4].

Consequently, efforts to detect COVID-19 have been undertaken through the utilization of chest imaging techniques, such as computed tomography (CT) scans or radiographic pictures of the chest X-ray. However, The use of CT images as a diagnostic method for COVID-19 does come with a set of drawbacks. Firstly, CT imaging equipment comes with a substantial cost and demands a significant level of expertise for efficient operation. This poses challenges for healthcare facilities that may not have the financial resources or specialized personnel to utilize CT scans for COVID-19 diagnosis. Moreover, non-portable CT imaging equipment can heighten the chances of human-to-human transmission when patients need to be transported. The situation is compounded by the shortage of personal protective equipment (PPE) kits for healthcare workers, which further increases this risk. Furthermore, CT imaging typically takes longer to process compared to X-ray imaging, potentially causing delays in diagnosis and treatment. Lastly, access to high-quality CT imaging systems may be limited in rural regions, making it difficult to promptly screen for COVID-19 infections in these areas. These limitations highlight the need for alternative diagnostic methods that are

ISBN 979-8-3503-5094-4

more accessible, cost-effective, and efficient for COVID-19 detection.

In contrast, X-rays are the most widely used and readily accessible radiographic diagnostic methods in clinical settings [5]. They prove highly valuable for cost-effective and swift COVID-19 infection screening.

The healthcare sector has benefited from recent progress in computer vision, which has led to affordable and dependable solutions by integrating deep learning techniques. Especially, the Convolutional Neural Network (CNN) regained attention in 2012 due to its remarkable ability to classify images with high accuracy. Subsequent studies have demonstrated the groundbreaking potential of CNNs in the realm of deep learning. Particularly, the focus of this paper lies in the convergence of computer vision methods with the healthcare industry. Automation has become increasingly prevalent in diagnosis and disease detection, thanks to the incorporation of deep learning techniques.

A robust AI-based COVID-19 diagnostic system that can attain excellent sensitivity and specificity, or precision and recall, on edge/mobile devices is highly demanded. Such AI-based tools for COVID-19 diagnosis have the potential to make screening tests more cost-effective and efficient for widespread, real-time testing. Additionally, they can reduce the risk of transmission to healthcare workers and alleviate the burden on the already limited healthcare professionals and radiologists. In this study, we introduce an exceptionally lightweight CNN model that can effectively detect COVID-19 from chest X-ray (CXR) images, which are commonly accessible in most clinical settings and imaging facilities. While numerous Deep Learning (DL) models have been created, the majority of them are characterized by a large number of parameters, rendering them unsuitable for deployment on devices with limited computational resources, such as mobile devices. Our model achieves similar accuracy compared to the existing deep learning-based models, all while maintaining a significantly reduced number of parameters.

The following sections comprise the paper: The extensive literature analysis in Part II provides an understandable appraisal of past techniques. Part III describes the specifics of the suggested model's implementation. The experimental findings collected are presented in Section IV. Part V reviews the project and concludes.

II. Related Work

In the field of medical image recognition, classifying medical photographs into groups is important for disease diagnosis and research. This involves extracting valuable features from the image and developing classification models to identify diseases [6]. [7] corroborated the efficiency of Convolutional Neural Networks (CNNs) in diagnosing COVID-19 through the analysis of digital images and CT scans for detecting the disease. Furthermore, further investigations into deep learning models were assessed for their accuracy in detecting COVID-19. L. Wang et al. [8] conducted a comparison among three models: VGG-19 (a deep CNN architecture with 19 layers), ResNet-50 (a deep CNN architecture with 50 layers), and COVID-Net (a deep CNN designed for COVID-19 detection using publicly available chest X-Ray images). The findings revealed that COVID-Net exhibited lower architectural and computational complexity in comparison to VGG-19 and ResNet-50., with COVID-19 sensitivity (91%). Their model COVID-Net has 11.75 million parameters and achieved accuracy of 92.4%. They conclude future research to be focused on risk stratification for survival analysis, and the prediction of risk status and prediction of hospitalization duration, which would be key factors in managing patients, and providing better treatments in the control of COVID-19 cases. N. Awasthi et al. [9] went a step further and introduced Mini-COVIDNet, a streamlined variant of the COVID-Net model. This version has significantly fewer trainable parameters, approximately 4.39 times fewer than its counterparts. Notably, the suggested model boasts a swift training period of under 30 minutes and demonstrates an 83.2% accuracy in detecting COVID-19. It's important to mention that their model comprises 3.36 million parameters. T. Ozturk et al. [10] used the DarkNet model, which is a state-of-the-art architecture for object detection, to propose DarkCovid-Net by changing the number of convolutional layers to 17. DarkCovid-Net obtained an accuracy of 98.08% and 87.02% on binary and multiclass classification tasks respectively. There multiclass classification model had 1.16 million parameters, and as with other works, this model was trained on only 1125 chest X-ray images which can be problematic as a more accurate model needs a significantly more robust dataset.

Hemdan et al. [11] introduced COVIDX-Net, a deep-learning framework, built on VGG19, DenseNet121, InceptionV3, ResNetV2, Xception, and MobileNetV2. Since it was one of the earliest attempted works in automated COVID detection, it faced the trouble of getting adequate training images and thus was trained on only a total of 50 chest x-ray images, and they achieved 90% accuracy. Abdani et al. [12] proposed SPP-COVID-Net model, and conducted a comparison between the SPP-COVID-Net model and six other lightweight deep learning models for X-ray image detection. The SPP-COVID-Net model achieved the highest average accuracy of 94.6% while utilizing fewer than one million parameters, specifically 0.86 million parameters.

Combining transfer learning methods with convolutional neural networks (CNNs) can significantly enhance the automated recognition and extraction of important features from X-ray images. In the context of COVID-19 detection, with this approach they achieved a multi-class classification accuracy of 93.48%., their model has 20.55 million parameters. [13]. Khan et al.[25] proposed coronet, they achieved 89% accuracy and their model has 33 million parameters.

Chakraborty et al. suggested Corona-Nidaan as a compact model [14]. They compared their model's performance to others with a higher number of parameters and discovered that their model surpassed them, achieving a 95% accuracy rate. Despite being labeled as lightweight, their model still incorporates 4.02 million parameters.

ISBN 979-8-3503-5094-4

III. Proposed Method

The suggested method is executed through a sequence of steps: initially obtaining the chest X-ray (CXR) images, followed by preprocessing, then building and training the classification model and, at the end, utilising test data to evaluate its effectiveness.

A. Chest X-ray Dataset

We employed a dataset proposed in [15], which included a collection of 33,920 chest X-ray (CXR) images, to both train and assess our deep learning model. This dataset encompasses three distinct categories: COVID-19, Non-COVID, and Normal. Figure 1, Figure 2 and Figure 3 displays a selection of sample images for COVID-19, Non-COVID and Normal class respectively from this dataset. During our model development, we utilized 21,715 images for training, allocated 5,417 images for validation, and earmarked 6,788 images for testing purposes. The distribution of these classes among the training, validation, and test datasets can be found in Table 1. All images were resized to dimensions of 256 by 256 pixels in RGB format for model training. We also used augmentations while training(on the fly), image augmentation is a method employed in image classification to artificially enhance the variety of training data by implementing various alterations to the initial images, this helps improve the model's ability to generalize and recognize patterns in different variations of the same image, ultimately enhancing its performance and reducing the risk of overfitting.

TABLE I
CLASSWISE COUNT OF TRAIN, VALIDATION AND TEST SET IMAGES OF OUR DATASET

Class	Train	*Validation*	*Test*
COVID-19	7658	1903	2395
Non-COVID	7208	1802	2253
Normal	6849	1712	2140

Fig. 1. Example CXR images from our dataset for COVID-19 class

Fig. 2. Example CXR images from our dataset for Non-COVID class

Fig. 3. Example CXR images from our dataset for Normal class

B. Model Architecture

The literature survey reveals that most of the methods examined in terms of parameters have a significant number of parameters. Although a few attempts have been made to propose lightweight models, they still have more parameters compared to our method. Additionally, some of the evaluated methods were tested on a limited number of images, which reduces their reliability. In contrast, our method achieves comparable performance to these methods while maintaining a lightweight design. Our approach uses a classification network to label an input X-ray picture as "COVID-19," "Non-COVID," or "Normal," respectively. In order to use our recommended technique on computers with minimal processing capacity, we sought to reduce the number of parameters as low as possible without sacrificing accuracy. Many earlier deep learning approaches for COVID-19 detection used models with a significant number of parameters or depended on well-tailored pre-trained convolutional neural networks (CNNs) like VGG16, MobileNetV2, MobileNetV3, ResNet, and similar architectures. Even though transfer learning can help reduce the number of trainable parameters, it doesn't solve the problem of having a large overall parameter count. This creates difficulties because the trained models become quite large, requiring substantial disk storage and computational resources due to their excessive number of layers. In this study, we present a COVID-19 classification approach employing a model with a limited number of parameters that nonetheless provides performance applicable to practical issues.

Our network comprises six layers of standard CNN (Convolutional Neural Network), kernel size for each CNN layer is 3x3 , kernel size of 3x3 is optimal for us as we aimed for a lightweight model. After each CNN layer, there is a subsequent batch normalization layer, and the ReLU activation function is applied(we call it CBR layer, see Figure 4). There are 8 kernels in first two CNN layers, 16 kernels in next two CNN layers and 32 kernels in last two CNN layers. There are total six layers of Average pooling with first 4 of pool size of 2x2, and fifth and sixth pooling layer comes after sixth CBR layer with fifth pooling layer of size 2x2 and sixth of size 4x4. Lastly our model has 2 Flatten layers which comes after last 2 pooling layers and then a concatenate layer which concatenates the two flatten layers and then comes two dense layers, The initial dense layer consists of 16 units and is subsequently followed by batch normalization and ReLU activation. The final layer serves as the classification layer with three units. To incorporate regularization, a single dropout layer with a

ISBN 979-8-3503-5094-4

rate of 0.3 is placed after the first dense layer.

Our model has total of 59,467(0.059 million) parameters, all are trainable parameters, our model's total size is 19.67 MB. Despite having significantly fewer parameters, our model delivers accuracy on par with models that possess a substantially larger parameter count. The comprehensive network architecture can be observed in Figure 4. Parameter count for each layer and final total parameter count is shown in figure 5.

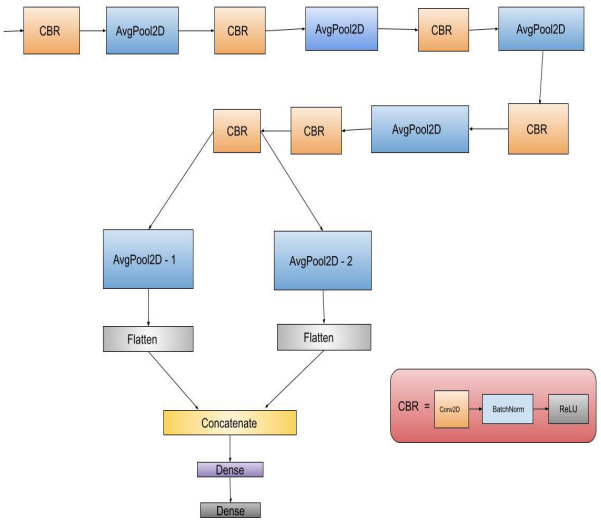


Fig. 4. The architectural diagram illustrating our model's structure

Layer (type)	Output Shape	Param #
Conv2d-1	[-1, 8, 256, 256]	224
BatchNorm2d-2	[-1, 8, 256, 256]	16
ReLU-3	[-1, 8, 256, 256]	0
AvgPool2d-4	[-1, 8, 128, 128]	0
Conv2d-5	[-1, 8, 128, 128]	584
BatchNorm2d-6	[-1, 8, 128, 128]	16
ReLU-7	[-1, 8, 128, 128]	0
AvgPool2d-8	[-1, 8, 64, 64]	0
Conv2d-9	[-1, 16, 64, 64]	1,168
BatchNorm2d-10	[-1, 16, 64, 64]	32
ReLU-11	[-1, 16, 64, 64]	0
AvgPool2d-12	[-1, 16, 32, 32]	0
Conv2d-13	[-1, 16, 32, 32]	2,320
BatchNorm2d-14	[-1, 16, 32, 32]	32
ReLU-15	[-1, 16, 32, 32]	0
AvgPool2d-16	[-1, 16, 16, 16]	0
Conv2d-17	[-1, 32, 16, 16]	4,640
BatchNorm2d-18	[-1, 32, 16, 16]	64
ReLU-19	[-1, 32, 16, 16]	0
Conv2d-20	[-1, 32, 16, 16]	9,248
BatchNorm2d-21	[-1, 32, 16, 16]	64
ReLU-22	[-1, 32, 16, 16]	0
AvgPool2d-23	[-1, 32, 8, 8]	0
AvgPool2d-24	[-1, 32, 4, 4]	0
Flatten-25	[-1, 2048]	0
Flatten-26	[-1, 512]	0
Linear-27	[-1, 16]	40,976
BatchNorm1d-28	[-1, 16]	32
ReLU-29	[-1, 16]	0
Dropout-30	[-1, 16]	0
Linear-31	[-1, 3]	51

Total params: 59,467
Trainable params: 59,467
Non-trainable params: 0

Fig. 5. All layers of our model, shape of their output and their parameter count

C. Training and Validation

Python was used as the programming language. We used pytorch deep learning framework for our model. In our model training process, we utilized the Adam optimizer along with a categorical cross-entropy loss function, commencing with an initial learning rate of 0.001. To dynamically adapt the learning rate, we employed a variable learning rate strategy, which decreases the rate by a factor of 0.2 if the validation loss doesn't show improvement for 5 consecutive epochs. Throughout the training, we used mini-batches of size 4 and conducted training for a total of 40 epochs.

We assessed our model's ultimate performance by employing a test set that was not employed during the training process, and included 2140 photos from the "Normal" class, 2253 images from the "Non-COVID" class, and 2395 images from the "COVID-19" class. To evaluate the effectiveness of the model, we employed the accuracy(Acc) metric, whose equation is provided below. And we used cross-entropy loss as loss function.

$$Acc = \frac{tp + tn}{tp + tn + fp + fn}$$

In simpler terms, accuracy is a measure of how many of the total samples a model correctly predicted. It's calculated based on various factors, including false negatives (fn), true positives (tp), true negatives (tn), and false positives (fp).

IV. Results and Discussion

Our model achieved an accuracy of 93.61% and the same F1 score when evaluated on unseen test data. Table 2 presents a performance comparison between our model and models that have been proposed in prior research, demonstrating that our model achieves comparable accuracy while requiring significantly fewer parameters. The table includes information on parameters (denoted as "Params") where "M" represents million.

TABLE II
COMPARISON BETWEEN OUR MODEL AND MODELS THAT HAVE BEEN PROPOSED IN PRIOR RESEARCH

Study	*Method*	*Params (M)*	*Accuracy*
Adbani et al. [12]	SPP-COVID-Net	0.86	94.6%
Wang et al. [8]	COVID-Net	11.75	92.4%
Hemdan et al. [11]	COVIDX-Net	20.55	90%
Ozturk et al. [10]	DarkCovidNet	1.16	87.02%
Ioannis et al. [13]	VGG-19	20.55	93.48%
Awasthi et al. [9]	Mini-COVIDNet	3.36	83.2%
Chakraborty et al. [14]	Corona-Nidaan	4.02	95%
Khan et al. [16]	CoroNet	33	89.6%
This work	**Ours**	**0.059**	**93.61%**

As our dataset is balanced so accuracy is good metric for evaluation, but still we also considered f1score metric for our model's performance, which is a good metric to show that model's performance for all classes is good, because if model doesn't perform good on any class this metric would not be

ISBN 979-8-3503-5094-4

good. To understand the f1score we need to look into precision and recall then f1score is just the harmonic mean of these two.

Precision - Precision in the context of multiclass image classification is a performance metric used to evaluate the accuracy of a deep learning model's predictions across multiple classes or categories. In multiclass classification, there are more than two possible classes or labels. Precision for a specific class 'i' is defined as:

$$Pr[i] = \frac{tp[i]}{tp[i] + fp[i]}$$

In other words, it measures the ratio of correctly predicted instances of class i to all instances predicted as class i. Precision can be calculated for each class separately.

To compute the overall precision for multiclass classification, there are two ways:

Macro-Average Precision: In the macro-average precision approach, you compute precision for each class separately and then determine the average (mean) of these individual precision scores. This technique treats all classes equally and doesn't consider any imbalance in class sizes. Each class's contribution to the overall precision is the same, irrespective of how many samples are present in each class.

Weighted Average Precision: In the weighted average precision approach, you compute precision for each class separately and then determine an average by considering the weights associated with each class. These weights are usually based on the number of samples in each class. This method addresses class imbalance by assigning higher importance to classes with more samples. It offers a more precise depiction of overall precision, especially when dealing with rare classes.

In a multiclass image classification scenario, precision provides insights into how well the model performs for each individual class and overall. It helps assess the model's ability to make accurate predictions for each category.

Recall - Recall, also known as sensitivity or true positive rate, is a crucial performance metric in multiclass image classification. It quantifies a model's ability to correctly identify all instances of a particular class among all instances that belong to that class.

In multiclass image classification, more than two classes, and we are interested in measuring how well the model performs for each individual class. Recall for a specific class 'i' is calculated a same as precision, but here false negatives is used instead of false positives.

In other words, recall measures the ratio of correctly predicted instances of class 'i' to all instances that actually belong to class 'i.' It tells us how well the model captures instances of a specific class.

To calculate the overall recall, there are two ways:

Macro-Average Recall: The macro-average recall approach involves computing the recall for each class separately and then finding the average (mean) of these individual recall values. This technique treats all classes equally and doesn't consider any imbalance in class sizes. Each class's contribution to the overall recall is the same, regardless of how many samples are in each class.

Weighted Average Recall: In the weighted average recall method, you calculate the recall for each class individually and then take a weighted average of these recall scores, where the weights are typically determined by the number of samples in each class. This method takes into account class imbalance by giving more weight to classes with more samples. It provides a more accurate representation of the overall recall when some classes are rare.

F1 score - The F1 Score, often referred to as the F1 measure or F1 score, is a popular and important metric in machine learning, particularly for classification tasks. It merges precision and recall into a single metric, offering a well-rounded evaluation of a model's effectiveness, particularly useful when working with datasets that have imbalanced class distributions, good F1 score shows our model's performance is excellent across all classes. The F1 Score is especially useful when you need to find a trade-off between precision and recall. So a good F1 score also shows our model's precision and recall is excellent. Overall F1 score can be calculated in two ways, macro-average and weighted average F1 score, method is same as described above for precision, just use F1 score in place of them.

The F1 Score is defined as the harmonic mean of precision and recall. The formula for calculating the F1 Score is as follows:

$$F1\_Score = \frac{2 * Precision * Recall}{Precision + Recall}$$

Figure 6 shows full classification report, showing every metric classwise and also overall.

	precision	recall	f1-score	support
COVID-19	0.9748	0.9687	0.9717	2395
Non-COVID	0.9096	0.9241	0.9168	2253
Normal	0.9212	0.9121	0.9166	2140
accuracy			0.9361	6788
macro avg	0.9352	0.9350	0.9351	6788
weighted avg	0.9362	0.9361	0.9361	6788

Fig. 6. Classification report on test set

The accuracy of our method for detecting the COVID-19 class is 98%. Detecting whether an individual has COVID-19 is a critical task, and it's crucial not to overlook any COVID-19 cases. Our model achieves an impressively high level of accuracy when it comes to identifying COVID-19 cases, ensuring dependable detection, while also maintaining a good level of accuracy for the other two classes. You can

ISBN 979-8-3503-5094-4

find our model's correct prediction count for each class on the test data in the Confusion Matrix shown in Figure 7.

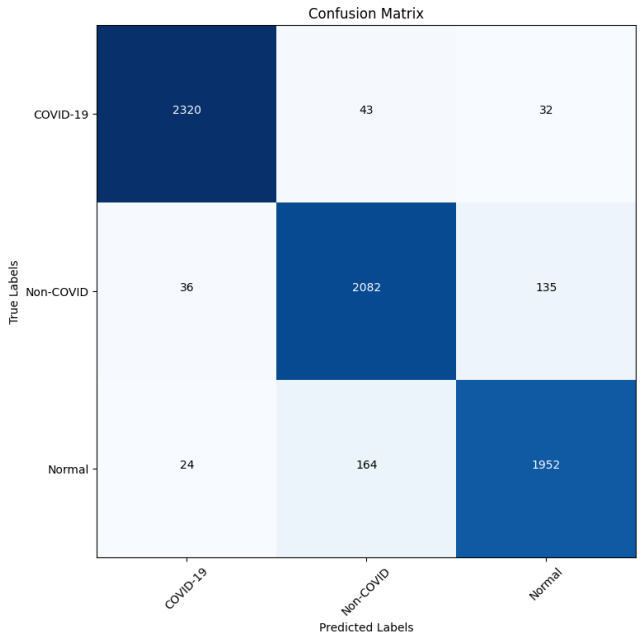


Fig. 7. The test set's confusion matrix depicting the predictions made by our lightweight model

From the results this can be seen when compared with other models our model gives excellent performance in terms of combined effect of parameter count and accuracy.

V. Conclusions

A robust AI-based COVID-19 diagnostic system that can attain excellent sensitivity and specificity, or precision and recall, on edge/mobile devices is highly demanded. Such AI-based tools for COVID-19 diagnosis have the potential to make screening tests more cost-effective and efficient for widespread, real-time testing. We proposed an exceptionally lightweight CNN model from scratch specifically designed for COVID-19 detection using CXR images, Chest X-ray (CXR) pictures are more easily obtainable when compared to chest computed tomography (CT) images, especially in less economically developed regions where CT machines are prohibitively costly. Our model underwent training on a comprehensive and varied dataset comprising three categories: "COVID-19," "Non-COVID," and "Normal." Remarkably, even though our model is compact in size, it managed to attain accuracy levels similar to those of previously suggested algorithms for COVID-19 detection, all while utilizing far fewer parameters. This particular feature positions our model as an excellent choice for implementation on less powerful devices such as edge devices with limited processing capacity, devices having less compute are also not much expensive so this is cost effective too. Also, our model is very helpful for the healthcare system, which has a constrained budget. Also, real-time COVID-19 detection can be very helpful; this is where our model can provide great value. Through a comparison with other existing methods, we observed that our model delivers similar accuracy levels while having total model size of 19.67 MB and only 0.059 million total parameters, which is far fewer than others. This small size and low parameter count enable our model to be incorporated into mobile applications for efficient COVID-19 detection while maintaining high accuracy. In conclusion, our model provides a rapid, precise, and cost-efficient solution for detecting COVID-19.

References

[1] J. Piret and G. Boivin, "Pandemics throughout history," *Frontiers in microbiology*, vol. 11, p. 631736, 2021.

[2] B. Hu, H. Guo, P. Zhou, and Z.-L. Shi, "Characteristics of sars-cov-2 and covid-19," *Nature Reviews Microbiology*, vol. 19, no. 3, pp. 141–154, 2021.

[3] A. Gupta, M. V. Madhavan, K. Sehgal, N. Nair, S. Mahajan, T. S. Sehrawat, B. Bikdeli, N. Ahluwalia, J. C. Ausiello, E. Y. Wan *et al.*, "Extrapulmonary manifestations of covid-19," *Nature medicine*, vol. 26, no. 7, pp. 1017–1032, 2020.

[4] B. Udugama, P. Kadhiresan, H. N. Kozlowski, A. Malekjahani, M. Osborne, V. Y. Li, H. Chen, S. Mubareka, J. B. Gubbay, and W. C. Chan, "Diagnosing covid-19: the disease and tools for detection," *ACS nano*, vol. 14, no. 4, pp. 3822–3835, 2020.

[5] B. Antin, J. Kravitz, and E. Martayan, "Detecting pneumonia in chest x-rays with supervised learning," *Semanticscholar. org*, vol. 2017, 2017.

[6] R. Yousef, G. Gupta, N. Yousef, and M. Khari, "A holistic overview of deep learning approach in medical imaging," *Multimedia Systems*, vol. 28, no. 3, pp. 881–914, 2022.

[7] D. Dong, Z. Tang, S. Wang, H. Hui, L. Gong, Y. Lu, Z. Xue, H. Liao, F. Chen, F. Yang, R. Jin, K. Wang, Z. Liu, J. Wei, W. Mu, H. Zhang, J. Jiang, J. Tian, and H. Li, "The role of imaging in the detection and management of covid-19: A review," *IEEE Reviews in Biomedical Engineering*, vol. 14, pp. 16–29, 2021.

[8] L. Wang, Z. Q. Lin, and A. Wong, "Covid-net: A tailored deep convolutional neural network design for detection of covid-19 cases from chest x-ray images," *Scientific Reports*, vol. 10, no. 1, pp. 1–12, 2020.

[9] N. Awasthi, A. Dayal, L. R. Cenkeramaddi, and P. K. Yalavarthy, "Mini-covidnet: efficient lightweight deep neural network for ultrasound based point-of-care detection of covid-19," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 68, no. 6, pp. 2023–2037, 2021.

[10] T. Ozturk, M. Talo, E. A. Yildirim, U. B. Baloglu, O. Yildirim, and U. R. Acharya, "Automated detection of covid-19 cases using deep neural networks with x-ray images," *Computers in biology and medicine*, vol. 121, p. 103792, 2020.

[11] E. E.-D. Hemdan, M. A. Shouman, and M. E. Karar, "Covidx-net: A framework of deep learning classifiers to diagnose covid-19 in x-ray images," *arXiv preprint arXiv:2003.11055*, 2020.

[12] S. R. Abdani, M. A. Zulkifley, and N. H. Zulkifley, "A lightweight deep learning model for covid-19 detection," in *2020 IEEE Symposium on Industrial Electronics & Applications (ISIEA)*. IEEE, 2020, pp. 1–5.

[13] I. D. Apostolopoulos and T. A. Mpesiana, "Covid-19: automatic detection from x-ray images utilizing transfer learning with convolutional neural networks," *Physical and engineering sciences in medicine*, vol. 43, no. 2, pp. 635–640, 2020.

[14] M. Chakraborty, S. V. Dhavale, and J. Ingole, "Corona-nidaan: lightweight deep convolutional neural network for chest x-ray based covid-19 infection detection," *Applied Intelligence*, vol. 51, no. 5, pp. 3026–3043, 2021.

[15] A. M. Tahir, M. E. Chowdhury, A. Khandakar, T. Rahman, Y. Qiblawey, U. Khurshid, S. Kiranyaz, N. Ibtehaz, M. S. Rahman, S. Al-Maadeed *et al.*, "Covid-19 infection localization and severity grading from chest x-ray images," *Computers in biology and medicine*, vol. 139, p. 105002, 2021.

[16] A. I. Khan, J. L. Shah, and M. M. Bhat, "Coronet: A deep neural network for detection and diagnosis of covid-19 from chest x-ray images," *Computer methods and programs in biomedicine*, vol. 196, p. 105581, 2020.

ISBN 979-8-3503-5094-4

Music Emotion Classification using Harris Hawk Optimization based LightGBM Classifier

Hemprasad Yashwant Patil
School of Electronics Engineering
Vellore Institute of Technology
Vellore, India
hemprasadpatil@gmail.com

Priyanka Bhagwan Patil
School of Computer Science and Engineering
Vellore Institute of Technology
Vellore, India
priyankagsvhpatil@gmail.com

Anishkaa Balasubramaniam
School of Electronics Engineering
Vellore Institute of Technology
Vellore, India
anishkaa.balasubramaniam@gmail.com

Lackesh Shanmugasundaram
School of Electronics Engineering
Vellore Institute of Technology
Vellore, India
lackesh.shanmugasundaram@gmail.com

***Abstract*— This research investigates the relationship between music and human emotions, focusing on emotion classification in Turkish music and the role of acoustics in music emotion recognition (MER). Leveraging a unique dataset stored in the UCI-MLR, the study employs machine learning techniques, including Linear Discriminant Analysis and Gradient Boost Classifier, to build a robust model. The accuracy enhancement is primarily driven by the feature selection process, which is powered by the Harris Hawks optimizer. This optimizer, when used with the LightGBM model, enables a more precise selection of features, thereby significantly improving the model's accuracy. The research achieves notable accuracies—81.25% for Linear Discriminant Analysis and 83.75% for the Gradient Boosting Classifier—surpassing previous records. The paper discusses findings in the context of affective computing, emphasizing potential applications in music recommendation systems and therapy, concludes with a comprehensive analysis, and outlines future possibilities. This research contributes to advancing MER, laying the groundwork for innovations in diverse domains, including the potential integration with IoT-based intelligent systems for real-time emotion analysis and personalized music experiences based on camera frames.**

Keywords*— *Emotion Classification, Gradient Boost, LightGBM, Linear Discriminant Analysis

I. Introduction

Music, a universal language of emotion, has the profound ability to evoke a wide range of emotions in listeners. This unique characteristic of music has piqued the interest of researchers worldwide, leading to an extensive body of work focused on understanding the intricate relationship between music and emotion. This research has culminated in the development of various models and algorithms for Music Emotion Recognition (MER), a field that aims to quantify and categorize the emotional content of music. [1], [2] The emotional impact of music is largely due to its fundamental components, which include melody, rhythm, harmony, and timbre. Melody, the sequence of musical notes, can convey a sense of joy or sorrow; rhythm, the pattern of beats, can induce excitement or relaxation; harmony, the combination of simultaneous notes, can create feelings of tension or resolution; and timbre, the tone color or quality of sound, can add richness and depth to the emotional experience. These elements, individually and collectively, hold a unique ability to evoke and communicate a vast array of human emotions and moods. In the context of the evolving landscape of social music apps and increased demand for enhanced music search capabilities, the focus of MER research has shifted towards predicting emotional themes in music [3]. This involves analyzing the emotional content of a song and categorizing it into specific emotional themes such as happiness, sadness, anger, or relaxation. This information can then be used to enhance music search capabilities, allowing users to find music that matches their current mood or desired emotional state [4].

The task of emotion classification from music is challenging due to the subjective nature of emotions and the complexity of musical compositions. Emotions can vary greatly among different individuals, making the interpretation of music highly personal. The complexity of musical compositions, which involves various elements, adds another layer of difficulty. The lack of universally accepted frameworks for representing emotions in music and standard datasets for training and testing emotion classification models further complicates the task. Annotating music with emotional labels is challenging due to the subjective nature of emotions and the temporal variations in emotional states. [5] Moreover, music data is typically high-dimensional, making it challenging to extract relevant features for emotion classification [6], [7]. Despite these challenges, recent advancements in machine learning and signal analysis have paved the way for emotion classification from music [8], [9].

In their paper presented during the contribution of the Turkish Music Emotion dataset, the original authors Er et al. [20] introduced a novel approach to music emotion recognition utilizing pre-trained deep networks. The methodology involves applying deep learning techniques to a pre-trained neuronal system, specifically the AlexNet framework as well as VGG-16, utilizing data-augmented chroma spectrograms derived from musical data. This augmentation enhances the classification success rate for music emotion recognition by incorporating deep visual features extracted from the chroma spectrograms. Deep visual features from different layers of AlexNet and VGG-16 are extracted and used to train SVM and Softmax classifiers. The optimal classification success, pre-data augmentation on the original dataset, is achieved with the "Fc6" layer within VGG-16 ", reaching 76% with the use of the Softmax classification algorithm on the data collection.

Er et al. suggested an approach for emotion detection in music involving three key steps: preprocessing, feature

ISBN 979-8-3503-5094-4

extraction, and classification in their paper published in 2021[21]. Initially, the noise in the signals was eliminated, and then all the signals were standardized to have a uniform sampling frequency. Following this, a feature vector of size 1x34, which encapsulated the emotional aspects of each signal, was extracted and normalized. The data was then classified using various algorithms, including Artificial Neural Network (ANN), Support Vector Machines (SVM), K-Nearest Neighbor (K-NN). The model's performance was evaluated on a novel 4-class Turkish music dataset, achieving an overall accuracy of 79.30%, precision of 79.31%, sensitivity of 79.13%, and F-score of 79.03%.

Furthermore, a study done by Garg. et al. [9] presented an AI leveraged model that aligns music tone with an individual's emotion using physiological signals. The model was comprised of three phases: predicting song mood from audio signals, forecasting human emotion with a pulse detector, EEG, ECG, GSR, and real-time mapping and classification of music tone with an individual's emotions. Comprehensive tests on various music tone data and individual's emotions encompass significant feature extraction (FE), training, validation, testing, and performance assessment. The model aims to precisely measure human emotions by recording bio-signals in reaction to music, and generating playlists based on real-time user emotion. Results demonstrate the superiority of Random Forest over traditional SVM, achieving an 80% relative accuracy in classification.

In their paper published in 2022, Unni et al. explored the application of emotion recognition in both music and speech in the RAVDESS dataset [3]. Employing supervised ML leveraged approaches including Naive Bayes, Random Forest, Decision Tree, and SVM, the study focuses on recognizing six sentiments (neutral, fearful, calm, happy, angry, sad) based on various musical features. Random Forest emerges as the most accurate and effective algorithm, achieving an accuracy of 75% when determining emotions, outperforming other classifiers. The methodology involves extracting distinct features and minimizing noise, allowing potential use in speech emotion detection.

Moreover, Chaudhary et al. in their study introduced Music Emotion Classification Systems (MECS) for Hindi songs [8], employing Convolutional Neural Networks (CNN) as well as Support Vector Machine (SVM). Two CNN-based MECS, MECS 1 and MECS 2, with varying spectrogram parameters, outperform the SVM-based MECS 3. MECS 1 achieves higher accuracy (91%) and lower losses (0.3249) than MECS 2 (87%, 0.4258) and MECS 3 (82.3%, 1.2278). The CNN-based systems demonstrate superior performance, predicting more classes with reduced error compared to SVM. MECS 1 consistently outperforms MECS 2 and MECS 3 in accuracy and loss metrics, highlighting the effectiveness of CNN in music emotion classification without the need for manual feature extraction.

This paper emphasizes the categorization of human emotions using a unique dataset derived from Turkish music, available at the UCI-MLR [10]. The dataset for the present investigation comprises vocal and non-vocal Turkish music pertaining to various genres, categorized into four discrete emotion classes: happy, sad, angry, and relaxed. Thirteen participants labeled 30-second segments of randomly played music pieces according to their perceived emotions. The most frequently assigned label determined the emotion class for each piece. Each participant listened to a total of 500 music segments over three sessions. To ensure balanced representation, 100 samples were selected for each emotion class. Chroma spectrogram analysis was conducted on each labeled music sample, capturing energy distribution across different notes. Chroma spectrograms were extracted using the MIR toolbox in experimental studies. This process resulted in an original dataset of 400 samples, with 30-second segments analyzed for emotion and chroma features.

This study seeks to build a robust framework for discrete emotion classification using ML methods such as linear Discriminant Analysis and Gradient Boost Classifier, evaluate their performance, and identify the most effective approach for classifying emotions in Turkish music. The significance of this research lies in its execution methodology, which uses a combination of LightGBM in integration with Harris Hawk optimization to select the most prominent contributing features, which brings about a significant boost in accuracy score. In doing so, this research aims to contribute to the broader field of emotional informatics, deepening the understanding of the relationship between music and human emotions, and opening avenues for innovative applications in music recommendation systems, music therapy, and beyond. [11-14]. The utilization of machine learning (ML) along with deep learning (DL) based methods for analyzing raw information in the form of data in real-world scenarios is becoming increasingly prevalent today [15-19].

The content of this document is structured as described: Section II explains the presented approach, including information related to the data, employed pre-processing methods, and the design flow. Section III presents the outcomes and accuracy comparison of the indicated procedure. The conclusion establishes a comprehensive investigation and outlines the possibilities for utilizing the methodology applicable for numerous datasets.

II. Proposed Solution

A. Dataset

The Turkish Music Emotion data collection used in this paper was obtained from the UCI-MLR. The data set consists of 50 feature columns corresponding to the different emotion labels 'relax', 'happy', 'sad', and 'angry' in the class column, totaling up to 400 samples. There are no missing values in this data collection.

B. Preprocessing

a) Encoding categorical values as numerical values

Unique values from the class column are extracted and categorical unique values are encoded as numerical values. In the class column, the label 'relax' is encoded as 1, 'happy' as 2, 'sad' as 3, and 'angry' as 4.

b) Normalization

Values in all the feature columns are normalized. Data normalization is used to make data more precise and less likely to be duplicated. It brings all numeric features in a dataset to a general scale, without distorting discrepancies in the ranges of values.

ISBN 979-8-3503-5094-4

C. Model selection and data processing

a) Initial accuracy computation

A train-test split of 80:20 is assigned to the data frame. The initial accuracy score for the Gradient Boosting classifier and Linear Discriminant Analysis is computed. Furthermore, the class probabilities for X_valid are predicted and the log loss of these predictions with respect to the actual values y_valid are computed. The log loss is stored in the variable P and returned. It is used as an objective value to assess the model's efficacy: As the log loss decreases, the model becomes more accurate.

b) Harris Hawk Optimization with LightGBM

An instance of the Harris Hawk Optimization class is created with the parameters for iteration count and size of the population. The Harris Hawk Optimization (HHO) algorithm [22] is a metaheuristic optimization algorithm inspired by swarming behavior that mimics the collective hunting strategy of a flock of Harris's hawks in the wild. In the HHO algorithm, each hawk in the flock denotes a potential solution to the optimization problem, with its position in the search space reflecting the decision variables of the problem. The algorithm operates iteratively, changing the positions of the hawks in each iteration as shown in (1), based on a framework that replicates the hunting behaviors of Harris's hawks.

There are 2 crucial steps to the technique: the exploration step and the exploitation step.

Exploration Step: In this phase, the hawks are randomly positioned within the exploration zone and wait to detect prey (which is the optimal solution). The hawks' positions are updated based on two approaches: they assume positions contingent on the whereabouts of neighboring hawks, and the prey, or they settle on arbitrary tall trees (random positions within the search space).

Exploitation Step: This phase models the hawks' behavior when they have detected prey. The hawks circle their prey and carry out surprise dives and adjust their strategies according to the energy levels of the prey and its attempts to escape. This phase is further divided into four methods: Hard besiege, progressive rapid dives with hard besiege, soft besiege, and soft besiege with progressive rapid dives.

1. **Soft Besiege**: The hawks gently circle their prey, leading it to deplete its energy reserves. The hawks then update their location predicated on the prey's location along with the group's average location. If the new location isn't reasonable, hawks perform irregular, abrupt, and rapid dives.

2. **Hard Besiege**: Used once the prey has exhausted all its strength. The hawks tightly circle the prey before performing the surprise swoop. Hawks then update their location based on the prey's location.

3. **Soft Besiege with Progressive Rapid Dives**: Used while the prey has adequate stamina to flee. Hawks decide their subsequent moves according to the prey's location along with the group's average location. If the new location isn't reasonable, they execute quick, erratic dives in accordance with the Levy flight principle.

4. **Hard Besiege with Progressive Rapid Dives**: Used when the prey doesn't have adequate stamina to evade the hawks. Hawks adjust positions based on the prey's location as well as the group's average location. If the new location isn't reasonable, they perform aberrant, sharp, and swift dives based on the Levy flight principle.

As shown in Fig. 2, the HHO algorithm seeks to balance exploration (searching the entire solution space) and exploitation (refining the current best solutions), proving to be effective in resolving intricate optimization issues. In optimization problems, the HHO aims to find the hawk that minimizes the objective function value. Conversely, in maximization problems, the algorithm focuses on identifying the hawk with the highest value according to the same objective function.

The position update equation in HHO is given by:

$$X\_new = X\_rabbit - (E \times J) \quad (1)$$

where:

- X_{new} is the new position of the hawk.
- X_{rabbit} is the position of the rabbit (the best solution found so far).
- E is an arbitrary value between 0 and 1.
- J is an arbitrary jump or leaping strength.

By dynamically adjusting hawk positions based on these behaviors, the HHO efficiently explores the solution space, striving to find optimal solutions within a reasonable computational time. This bio-inspired approach showcases the algorithm's adaptability and collaborative nature, aligning with the principles of swarm-based optimization.

Fig. 1 provides a clear overview of the proposed methodology, where the data preprocessing stages are in green, the feature selection process is in blue, and the classification process after feature extraction is in yellow. During the feature selection stage, the LightGBM library [23] is imported and an instance of the LGBMClassifier class is created. LightGBM is a framework for gradient boosting that employs algorithms predicated on tree learning, and LGBMClassifier is a class for gradient boosting classification.

In multiclass classification using LightGBM, the softmax function serves as the objective function. It transforms the raw outputs of the model into probabilities that collectively add up to 1. This transformation enables the model to provide a distribution of probabilities across the different classes. For every data point in the dataset, the model computes an initial score for each class. These initial scores are subsequently

ISBN 979-8-3503-5094-4

transformed into probabilities through the application of the Softmax function – (2)

$$P(y = j \mid x) = \frac{e^{x \cdot w_j}}{\sum_{k=1}^{K} e^{x \cdot w_k}} \quad (2)$$

where:

- $P(y = j \mid x)$ represents the probability that the instance x belongs to class j.
- x is the feature vector of the instance.
- W_j is the weight vector of class j.
- K is the total class count.

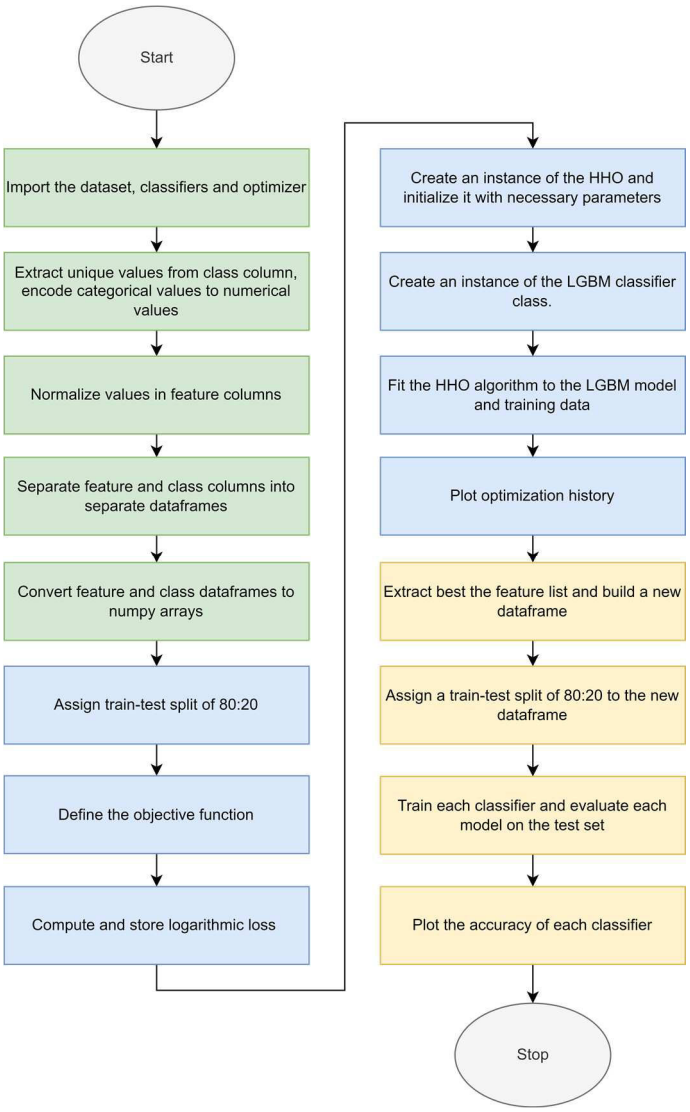


Fig. 1. Flowchart of the proposed methodology

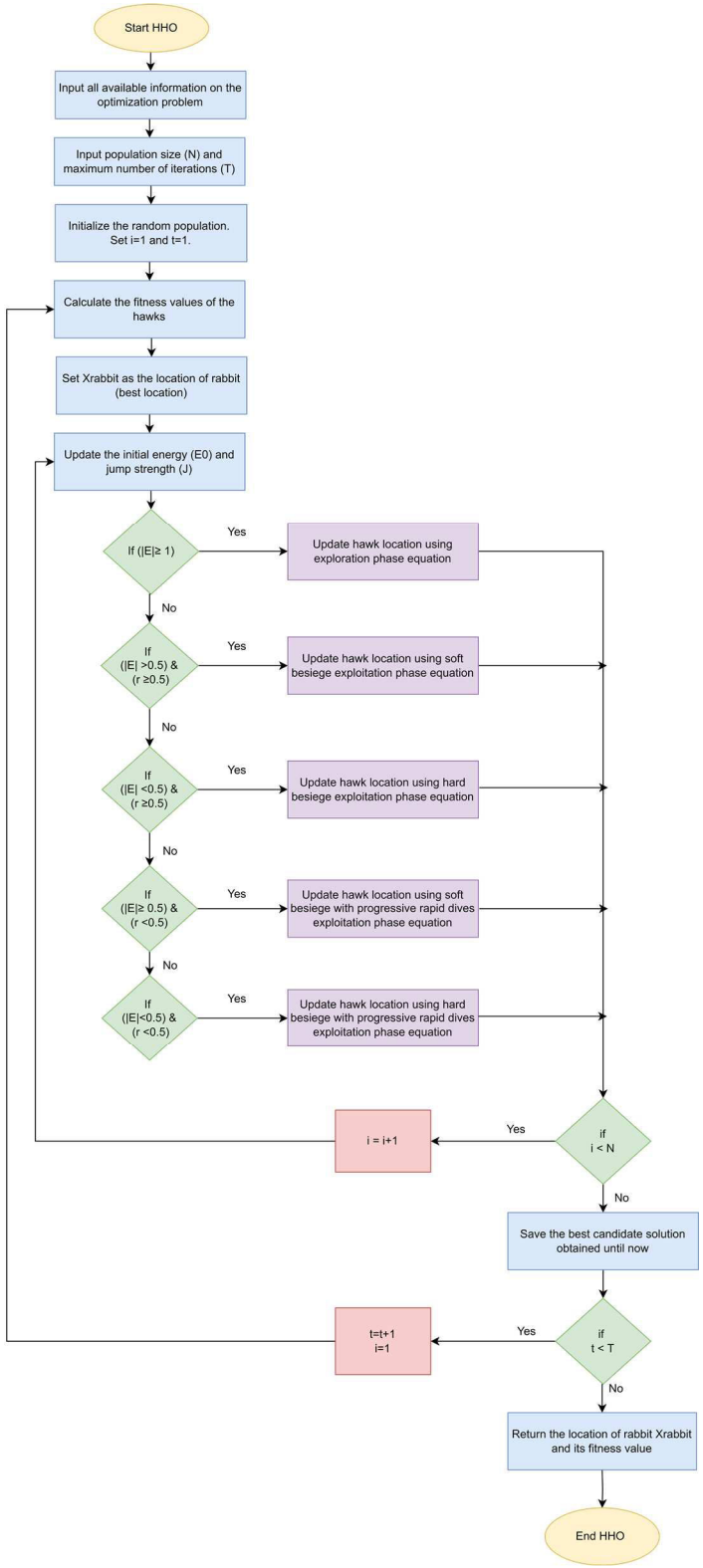


Fig. 2. Harris Hawk Optimization algorithm used for feature selection.

The goal of the model is to reduce the negative log-likelihood of the actual class labels, a process that is the same as minimizing the multi-log loss. Following the prediction of

ISBN 979-8-3503-5094-4

probabilities, the model computes the multi-log loss using these probabilities and the actual class labels, as per (3). The multi-log loss serves as an indicator of the accuracy of the predicted probabilities (i.e. the fitness of the model) in relation to the true class labels. The lower the log loss, the better the model's fitness.

$$L(y,p) = -\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{K} y_{ij}\, log(p_{ij}) \qquad (3)$$

where:

- $L_{(y,p)}$ is the multi-log loss.
- N is the count of instances.
- y_{ij} is 1 if instance i belongs to class j and 0 otherwise.
- p_{ij} is the predicted probability that instance i belongs to class j

The HHO algorithm is fit to the LightGBM model and the training data. Subsequently, the history of the objective function values is plotted across the iterations of the HHO algorithm as shown in Fig. 3. This can be useful to visualize the fitness of the LightGBM model and how the optimization process progressed.

c) Feature selection

The list of best features found by the HHO algorithm is extracted. These are the features that resulted in the lowest objective function value when used to train the LightGBM model. A new data frame is built with only the extracted features and a train-test split of 80:20 is assigned to a new dataset of extracted features, on which each classifier mentioned before is trained and evaluated on the test set.

III. Results and Discussion

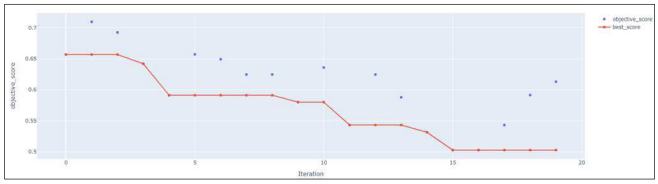


Fig. 3. Optimization history plot

This research was executed on Google Colaboratory, utilizing its free resources which consist of VMs with a standard system memory profile, with the runtime type set to T4 GPU, and for its collaborative features and convenience. The HHO algorithm extracts the list of best features that results in the lowest objective function value when used to train the LightGBM model. Fig. 3 shows the optimization history plot for the LightGBM classifier with a minimized logarithmic loss of approximately 0.5. 28 features that contributed maximally to the output label were chosen. Using the features selected by HHO with LightGBM, a new dataset was created and the accuracy before and after feature selection were compared.

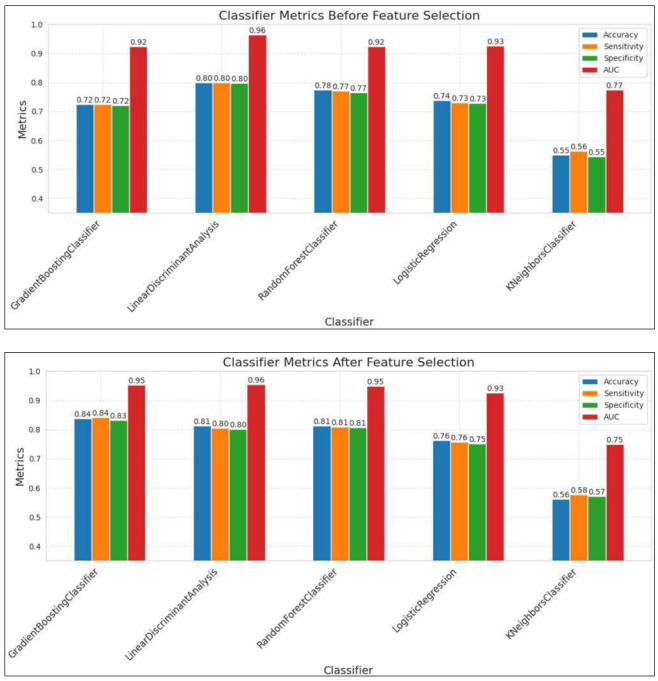


Fig. 4. Classifier metrics before and after feature selection

Classifier	Before Feature Selection				After Feature Selection			
	Accuracy	Sensitivity	Specificity	AUC-ROC	Accuracy	Sensitivity	Specificity	AUC-ROC
Gradient Boosting	0.7500	0.7436	0.7420	0.9254	0.8375	0.8408	0.8317	0.9525
LDA	0.8000	0.7991	0.7968	0.9639	0.8125	0.8045	0.8013	0.9556
Random Forest	0.8000	0.7948	0.7998	0.9326	0.8125	0.8088	0.8043	0.9494
Logistic Regression	0.7375	0.7297	0.7292	0.9268	0.7625	0.7575	0.7506	0.9259
SVC	0.7625	0.7618	0.7545	0.9260	0.7375	0.7340	0.7366	0.9176
KNN	0.5500	0.5641	0.5454	0.7735	0.5625	0.5780	0.5716	0.7496
GaussianNB	0.7625	0.7447	0.7648	0.8933	0.7500	0.7436	0.7475	0.9095
MLP	0.7875	0.7810	0.7815	0.9213	0.7125	0.7147	0.7018	0.9090
Decision Tree	0.6250	0.6389	0.6353	0.7575	0.6875	0.7041	0.7033	0.7998

Table I. Performance indicators of the presented algorithms

As presented in Table I and Fig. 4., the proposed feature selection methodology provides an increase in almost all the performance metrics in all classifiers used. The key findings in this research include the improvements in performance metrics include that of the accuracy of the Gradient Boosting Classifier and Linear Discriminant Analysis, both of which exceed the classification accuracy of 76% of the unaugmented dataset produced by contributors of the original dataset, based on which this research was conducted [20]. Therefore, it could be inferred that the novel integration of HHO with the LightGBM model helps to identify the features contributing the most to the class of music emotion, which in turn elevates the correctness of the classifiers used. The highest accuracies recorded were that of Gradient Boosting Classifier and Linear Discriminant Analysis. Gradient Boosting Classifier had significant improvements in terms of all performance metrics, followed by the Random Forest Classifier, Logistic Regression, and Linear Discriminant Analysis. KNN Classifier yielded the lowest performance metrics overall, even though there was a notable improvement in accuracy, sensitivity, and specificity. This research proposes a robust model for Music Emotion Classification that provides high accuracy, sensitivity, and specificity.

Subsequently, this research gives a higher accuracy factor with a more simplistic approach than the methods proposed

ISBN 979-8-3503-5094-4

by the authors in the research accompanying the donation of the dataset [20]. Considering the scope for Artificial General Intelligence in the recent future [24], it is evident that even a small accuracy leap gives a major rise in the commercialization of smart music systems for future homes, which are customized to suit individual preferences. In addition to this music-based recommendation systems could also significantly impact the mood of individuals and help contribute to their emotional well-being [25]. This could have potential applications in helping people overcome anxiety and stress. Emotion is spectral rather than discrete in nature, therefore increasing the number of labels in the dataset that are attenuated to subtle changes in the pitch class profiles, would provide a more seamless experience when implemented in real time.

IV. Conclusion

The approach implied in the present investigation elevates the accuracy of the Turkish Music Emotion dataset using novel machine learning algorithms. After preprocessing the dataset, the dataset undergoes a train-test split of 80-20, and the accuracy metric for the Linear Discriminant Analysis and Gradient Boosting Classifier is calculated. Additionally, the dataset undergoes feature selection using a Harris Hawks optimizer-based LightGBM model after which the accuracy score is recalculated. Linear Discriminant Analysis records an accuracy of 81.25% and the Gradient Boosting Classifier gives an accuracy of 83.75%., which both exceed the highest accuracy recorded for this dataset in previous works.

The constraints of this research include the scarcity of previous research in the MER domain. Furthermore, the dataset is also stagnant due to the limited number of samples and labels available to capture the wide range of emotions conveyed through music, which necessitates the need for a more extensive dataset. However, this research holds high promise in the near future due to the commercialization of IoT-based intelligent systems that could accurately analyze the emotion experienced by the user using camera frames and play music appropriately. In addition to this, this research also finds ways to aid the hearing-impaired by converting music to RGB, overcoming the communication barrier, and also plays a major role in music-based rehabilitation therapy.

References

[1] Y. O. Medina, J. R. Beltrán, and S. Baldassarri, "Emotional classification of music using neural networks with the MediaEval dataset," *Pers. Ubiquitous Comput.*, vol. 26, no. 4, pp. 1237–1249, 2022.

[2] V. Chaturvedi, A. B. Kaur, V. Varshney, A. Garg, G. S. Chhabra, and M. Kumar, "Music mood and human emotion recognition based on physiological signals: a systematic review," *Multimed. Syst.*, vol. 28, no. 1, pp. 21–44, 2022.

[3] D. Unni, A. M. D'Cunha, and G. D., "A Technique to Detect Music Emotions Based on Machine Learning Classifiers," in *2022 Second International Conference on Interdisciplinary Cyber Physical Systems (ICPS)*, 2022, pp. 136–140.

[4] D. Han, Y. Kong, J. Han et al., "A survey of music emotion recognition," in *Front. Comput. Sci.*, vol. 16, p. 166335, 2022. [Online]. Available: https://doi.org/10.1007/s11704-021-0569-4

[5] Y. R. Pandeya, B. Bhattarai, and J. Lee, "Deep-Learning-Based Multimodal Emotion Classification for Music Videos," in *Sensors*, vol. 21, no. 14, p. 4927, 2021.

[6] U. Xia and F. Xu, "Study on Music Emotion Recognition Based on the Machine Learning Model Clustering Algorithm," in *Mathematical Problems in Engineering*, vol. 2022, Article ID 9256586, pp. 1-11, 2022.

[7] Guiying Tong, "Music Emotion Classification Method Using Improved Deep Belief Network", Mobile Information Systems, vol. 2022, Article ID 2715765, 7 pages, 2022.

[8] D. Chaudhary, N. P. Singh, and S. Singh, "Ef music emotion classification system using convolution neural network," *Int. J. Speech Technol.*, vol. 24, no. 3, pp. 571–580, 2021.

[9] A. Garg, V. Chaturvedi, A. B. Kaur, V. Varshney, and A. Parashar, "Machine learning model for mapping of music mood and human emotion based on physiological signals," *Multimed. Tools Appl.*, vol. 81, no. 4, pp. 5137–5177, 2022.

[10] M. B. Er, *Turkish Music Emotion*. UCI Machine Learning Repository

[11] Y. Agrawal, R. G. R. Shanker, and V. Alluri, "Transformer-Based Approach Towards Music Emotion Recognition from Lyrics," in *Advances in Information Retrieval*, Springer International Publishing, 2021, pp. 167–175.

[12] J. S. Gómez-Cañón *et al.*, "Music Emotion Recognition: Toward new, robust standards in personalized and context-sensitive applications," *IEEE Signal Process. Mag.*, vol. 38, no. 6, pp. 106–114, 2021.

[13] X. Cui, Y. Wu, J. Wu, Z. You, J. Xiahou, and M. Ouyang, "A review: Music-emotion recognition and analysis based on EEG signals," *Front. Neuroinform.*, vol. 16, p. 997282, 2022.

[14] M. Jandaghian, S. Setayeshi, F. Razzazi, and A. Sharifi, "Music emotion recognition based on a modified brain emotional learning model," *Multimed. Tools Appl.*, vol. 82, no. 17, pp. 26037–26061, 2023.

[15] P. K. Chawla *et al.*, "Parkinson's disease classification using nature inspired feature selection and recursive feature elimination," *Multimed. Tools Appl.*, 2023.

[16] K. u. V. R. teja, B. P. V. Reddy, L. P. Alla, and H. Y. Patil, "Parkinson's Disease Classification using Quantile Transformation and RFE," in *2021 12th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, 2021, pp. 1–5.

[17] S. B. Mukadam and H. Y. Patil, "Skin Cancer Classification Framework Using Enhanced Super Resolution Generative Adversarial Network and Custom Convolutional Neural Network," *Applied Sciences*, vol. 13, no. 2. 2023.

[18] J. V Tembhurne, N. Hebbar, H. Y. Patil, and T. Diwan, "Skin cancer detection using ensemble of machine learning and deep learning techniques," *Multimed. Tools Appl.*, vol. 82, no. 18, pp. 27501–27524, 2023.

[19] K. U. V. R. Teja, B. P. V Reddy, L. P. A, H. Y. Patil, and P. C. T, "Prediction of Diabetes at Early Stage with Supplementary Polynomial Features," in *2021 Smart Technologies, Communication and Robotics (STCR)*, 2021.

[20] M. B. Er and I. B. Aydilek, "Music Emotion Recognition by Using Chroma Spectrogram and Deep Visual Features," *Int. J. Comput. Intell. Syst.*, vol. 12, no. 2, pp. 1622–1634, 2019.

[21] M. B. ER and E. M. ESİN, "Music Emotion Recognition with Machine Learning Based on Audio Features TT - Music Emotion Recognition with Machine Learning Based on Audio Features," *Comput. Sci.*, vol. 6, no. 3, pp. 133–144, 2021.

[22] A. A. Heidari, S. Mirjalili, H. Faris, I. Aljarah, M. Mafarja, and H. Chen, "Harris hawks optimization: Algorithm and applications," *Futur. Gener. Comput. Syst.*, vol. 97, pp. 849–872, 2019.

[23] G. Ke *et al.*, "Lightgbm: A highly efficient gradient boosting decision tree," *Adv. Neural Inf. Process. Syst.*, vol. 30, 2017.

[24] F. Dou et al., "Towards Artificial General Intelligence (AGI) in the Internet of Things (IoT): Opportunities and Challenges," arXiv preprint arXiv:2309.07438, 2023.

[25] A. Kumble, S. Medatati, and A. Bhatt, "Emotion-Based Music Recommendation," in *Advanced Computational and Communication Paradigms*, S. Borah, T.K. Gandhi, and V. Piuri, Eds. ICACCP 2023. Lecture Notes in Networks and Systems, vol. 535. Singapore: Springer, 2023.

ISBN 979-8-3503-5094-4

Automatic Monkeypox Disease Detection from Preprocessed Images using MobileNetV2

Ramya Mohan
Department of CSE
Saveetha School of Engineering,
SIMATS, Chennai 602105, TN, India.
Email: ramyanallu@saveetha.com

Robertas Damasevicius
Department of Applied Informatics
Vytautas Magnus University
53361 Kaunas, Lithuania
ORCID: 0000-0001-9990-1084

David Taniar
Faculty of Information Technology
Monash University
Melbourne, VIC 3800, Australia
ORCID: 0000-0002-8862-3960

N. Sri Madhava Raja
LTI Mindtree
Manapakkam, Poonamallee Road
Chennai 600089, TN, India.
Email:srimadhavaraja.nadaradjane@ltimindtree

V. Rajinikanth
Department of CSE
Center for Research and Innovation
Saveetha School of Engineering,
SIMATS, Chennai 602105, TN, India.
Email: v.rajinikanth@ieee.org

***Abstract*—The prevalence of infectious diseases in humankind is increasing globally due to a variety of reasons, and accurate diagnosis and treatment will help control/cure the disease. A severe health problem is associated with Monkeypox (Mpox), a communicable illness caused by the monkeypox virus. This research aims to develop a computerized tool to detect Mpox from pre-processed images using the pre-trained lightweight deep-learning scheme (PLDS). This tool consists of the following phases; (i) Image collection and tri-level thresholding based pre-processing, (ii) Feature extraction using selected PLDS, and (iii) five-fold cross-validation supported binary classification. As part of this research, we examine the possibilities for developing an accurate Mpox detection system using PLDS, including; (i) Kapur's thresholding, (ii) chosen optimizers, and (ii) chosen activation functions. This experimental investigation utilizes augmented images from the Monkeypox Skin Images Dataset (MSID), and the developed tool with MobileNetV2 achieves 98.7% detection accuracy when Kapur's thresholding is applied. Further, this tool presents a testing accuracy of >90% on the original MSID images, confirming the proposed research's significance.**

Keywords—Skin infection, Monkeypox, Image enhancement, MobileNetV2, Detection.

I. Introduction

Monkeypox (Mpox) is an illness caused by the monkeypox virus. It is characterized by symptoms such as a skin rash or mucosal lesions that can last for 2–4 weeks. Other common symptoms include fever, headache, muscle aches, back pain, fatigue, and swollen lymph nodes. It is a communicable disease and will spread from the infected individual to healthy individual due to physical contact [1].

In 2022–2023, there was a global outbreak of Mpox that resulted in significant infections in individuals across several countries [2]. The skin infection due to the Mpox sometimes similar to the infection caused by Chickenpox (Cpox), but the severity in Mpox is more compared to Cpox [3]. Accurate identification of the Mpox infection is essential for the treatment planning and implementation to control its harshness.

Accurate detection of the Mpox infection is essential for confirming the disease and to treat it with an appropriate clinical protocol. Recently, the image supported disease diagnose is emerged as one of the common practice in clinics and hence a number of image examination procedures are proposed and implemented. Artificial-Intelligence (AI) supported medical data evaluation is one of the recent approach to automate the data handling and decision making process and hence, a number of machine-learning (ML) and deep-learning (DL) schemes are proposed and implemented by the researchers [4,5]. The DL schemes helps to provide better detection accuracy compared to the conventional ML schemes.

DL based methods available for the image supported Mpox detection can be found in [6-8]. These methods considered the digital image supported disease evaluation to distinguish the Mpox with the non-Mpox (NMpox), and most of these works considered the Mpox Skin Images Dataset (MSID) [9] for the investigation.

This research aims to develop a lightweight deep-learning scheme (PLDS) based tool to distinguish the digital skin-infection images into Mpox and NMpox with a better accuracy. The different phases of the proposed tool include; (i) image collection and resizing, (ii) feature extraction with MobileNetV2 (MNV2), and (iii) binary classification with SoftMax classifier and verification.

The proposed tool's performance is initially verifiedon the raw MSID images using different optimizers (Adam/SGD) and activation functions (ReLU/PReLU). Later, a pre-processing based on Kapur's tri-level thresholding based on [10,11] is executed and the similar experiment is repeated. The experimental outcome confirms that the Kapur's enhanced image helps in providing better detection accuracy compared to the raw image.

The experimental outcome of this study confirms that the Adam and ReLU based scheme helps in achieving a better detection accuracy with raw-image (>90%) and the thresholded image (98.7%). This confirms that, the proposed tool is efficient in providing a better detection of Mpox from the chosen MSID image.

The contribution of this research is as follows;

(i) Performance evaluation of MNV2 with chosen optimizer and activation function,

(ii) Verification of MNV2 using raw and thresholded image.

Section 2 reviews the literature, Section 3 outlines the methodology, and Sections 4 and 5 present the research results and conclusions, respectively.

ISBN 979-8-3503-5094-4

II. Earlier Works

The rising global prevalence of infectious diseases highlights the critical importance of early detecting and treatment. The Mpox is one of the infectious diseases which will cause a severe skin infection and other health complications. Efficient detection of the Mpox is essential for the treatment.

Recently, AI supported methods are developed to classify the skin images into Mpox Vs NMpox and these works confirms that the AI scheme provides a better detection accuracy.

The work of Bala et al. (2023) proposed a novel MonkeyNet scheme to detect the Mpox in MSID an achieved a test accuracy of 93.19%. In this work, a detailed evaluation of the proposed scheme is presented using the batch-sizes of 8, 16, and 32. The performance of the proposed scheme is confirmed and verified against DenseNet201 (DN201) [12]. The research by Azar et al. (2023) implemented DN201 supported classification and achieved an accuracy of 97.63% during a four-class image examination task [13]. Examination by Uysal (2023) proposed a hybrid DL scheme for classifying the skin infection using a four-class classifier and achieved an accuracy of >87% [14]. Recent research by Maqsood and Damaševičius (2023) presented a DL scheme and achieved an accuracy of 98.59% [15]. Alhasson et al. (2023) proposed a DL based mobile application to examine the MSID and achieved an accuracy of 99% [16]. Recently, Eliwa et al. (2023) proposed Grey-Wolf algorithm tuned DL model to detect the Mpox and achieved an accuracy of 95.3% [17]. The research of Nayak et al. (2023) implemented the DL and explainable AI to detect the Mpox and achieved 91.19% accuracy [18]. The earlier works implemented the pre-trained DL methods to detect the Mpox and achieved better detection accuracy.

In the proposed work, MNV2 supported technique is implemented with varied optimizers and activation functions and the Mpox detection performance is verified with SoftMax using the raw and pre-processed MSID.

III. Methodology

This section details the methodology employed to detect Mpox using the MNV2 with different optimizers and activation functions. The study utilizes the MSID database referenced in [9]. Initially, the raw images are considered for the examination and then the raw image is processed with Kapur's thresholding and the similar practice is repeated.

Figure 1 represents the construction of developed tool and it evidently shows the process of Mpox detection. Initially, the necessary skin-images are collected from the volunteers using a digital camera and then these images are enhanced and stored for future use. The developed tool is then considered to implement the necessary classification task. Initially, the images of MSID are resized to $150 \times 150 \times 3$ pixels and then it is processed with MNV2 with appropriate optimizers and activation. Finally, the binary classification is achieved with the SoftMax and the merit of the proposed scheme is verified with the CM and the initial (TP, FN, TN, and FP) and final measures (AC, PR, SE, and SP), as depicted in the figure.

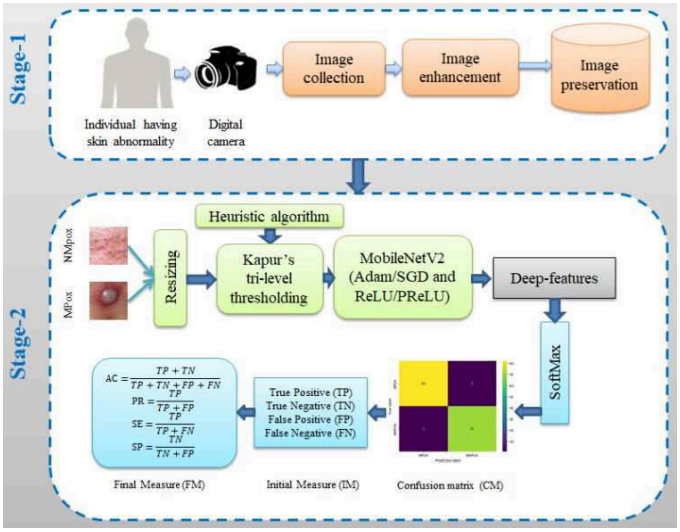


Figure 1. Architecture of proposed Mpox examination tool

A. MSID database

The essential test images for this investigation is collected from the benchmark database found in [9]. This database consist 770 Mpox and 770 NMpox images. During this examination, each image is resized to $150 \times 150 \times 3$ pixels and then these images are preprocessed using a tri-level thresholding based on [10,11]. In this work, 80% of the data is considered for training and remaining 10% data for the validation and testing operations as in Table 1. The sample test images of this database is shown in Figure 2.

TABLE I. Test Image Division Considered in This Study

Class	Total	Train	Validation	Test
Mpox	770	616	77	77
NMpox	770	616	77	77

Figure 2. Sample test images from MSID

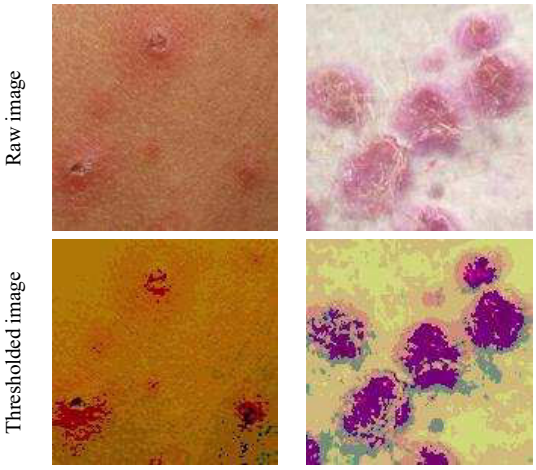
B. Kapur's thresholding

Image enhancement plays a vital role in improving the accuracy when DL supported automatic disease detection process. In this work, tri-level thresholding based on Kapur's approach and electromagnetism-algorithm discussed in [10,11] is considered for the assessment.

The proposed scheme helps to improve the visibility of the abnormality in the image considerably and this helps in improving the detection accuracy considerably, when AI based detection is implemented. Figure 3 presents the original and Kapur's enhanced images. Essential information regarding Kapur's thresholding can be found in literature.

ISBN 979-8-3503-5094-4

Fig 3(a) presents Mpox and Fig 3(b) presents NMPox. The raw and Kapur's enhanced images are separately assessed with the MNV2 and the achieved results are presented and discussed.



(a) NMpox (b) Mpox

Figure 3. Image enhanced with Kapur's entropy

C. *MobileNetV2 scheme*

Compared to the conventional DL approaches, the PLDS approaches are widely adopted when a low capacity processing device needs to be used for a real-time implementation of the developed tool. This work selected the MNV2 for the examination and the necessary information about this model can be accessed from [21]. This work provides a top1 accuracy of 71.3% and top5 accuracy of 90.1% and the total tuneable parameters are 3.5million, which is comparatively lesser than the MobileNet-variant. The necessary detail about the MNV2 based image examination can be accessed from [22]. The initial values of the MNV2 is assigned as follows; learning rate= 0.001, epoch= 100, and batch-size= 16.

D. *Execution and Verification*

The performance of the proposed tool is evaluated by calculating metrics from the CM, as shown in Figure 1. Metrics, like AC, PR, SE, and SP are calculated from the CM values to assess the proposed Mpox detection tool's performance. Furthermore, the tool's overall merit is also assessed using the ROC curve, which links the True-Positive (TP) rate against the True-Negative (TN) rate. The required information about the selected measures can be accessed from [12-14]. This work implemented a three-fold classification and the best performance is chosen as outcome.

IV. Result and Discussion

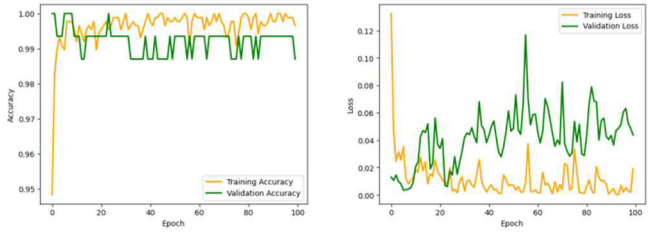
This part of the research presents the investigational outcome of the present study. The image resizing and the Kapur's thresholding are executed using the Matlab-software and the proposed MNV2 is implemented using Python-software. The proposed work is executed using a workstation with following specification; Inter i5, 20GB of RAM, and 4 GB VRAM.

Initially, all the chosen images are resized to $150 \times 150 \times 3$ pixels and then enhanced using the Kapur's function. This work implements the proposed tool using the raw and thresholded images and the achieved results are presented and discussed.

Figure 4 presents the experimental outcome obtained with the Kapur's enhanced MSID. Fig 4(a) and (b) depicts the accuracy and the loss value achieved for a chosen epoch size of 100. This confirms that the validation result closely tracks the training outcome, which confirms the merit of the MNV2 implemented with Adam-optimizer and ReLU.

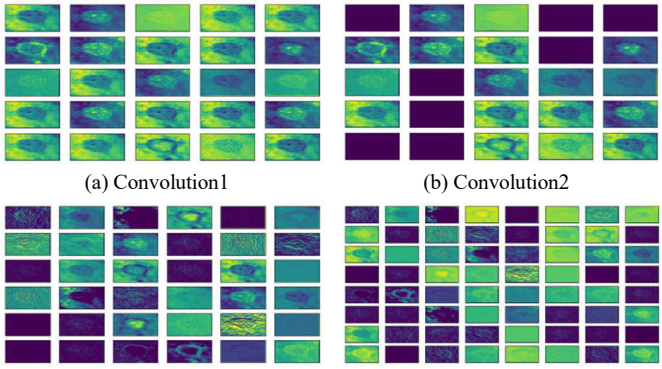
Figure 5 presents the sample convolution-layer (CL) values collected to verify the merit of the MNV2. Fig 5(a) to (d) shows CL1 to 4 values, which confirms the merit of the proposed MNV2 based examination.

Figure 6 presents the final outcome during the testing task, in which Fig 6(a) and (b) presents the CM and ROC-curve, respectively. The CM shows a better initial metrics and the ROC provides a value of unity, which confirms the merit of the implemented scheme.



(a) Accuracy (b) Loss

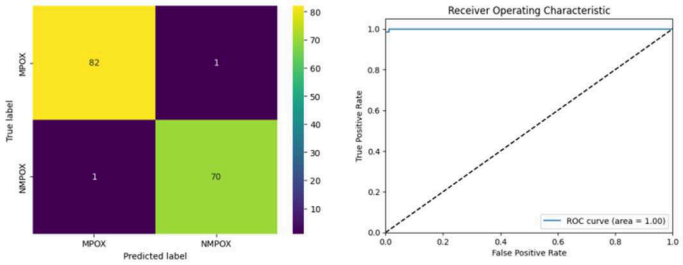
Figure 4. Training and validation performance with a chosen epoch



(a) Convolution1 (b) Convolution2

(c) Convolution3 (d) Convolution4

Figure 5. Different CL results for a chosen Mpox image



(a) CM (b) ROC

Figure 6. Final outcome with the thresholded MSID examination

TABLE II. Performance Evaluation of MNV2 on Raw Image

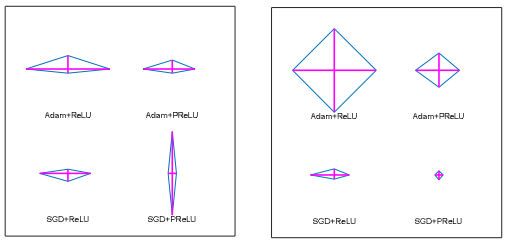
Scheme	TP	FN	TN	FP	AC	PR	SE	SP
Adam+ReLU	72	6	68	8	90.91	90.00	92.31	89.47
Adam+PReLU	71	7	68	8	90.25	89.87	91.02	89.47
SGD+ReLU	70	7	69	8	90.25	89.74	90.91	89.61
SGD+PReLU	69	9	69	7	89.61	90.79	88.46	90.79

ISBN 979-8-3503-5094-4

TABLE III. PERFORMANCE EVALUATION OF MNV2 ON THRESHOLDED IMAGE

Scheme	TP	FN	TN	FP	AC	PR	SE	SP
Adam+ReLU	70	1	82	1	98.70	98.59	98.59	98.79
Adam+PReLU	73	3	75	3	96.10	96.05	96.05	96.15
SGD+ReLU	74	3	73	4	95.45	94.87	96.10	94.80
SGD+PReLU	71	5	74	4	94.15	94.67	93.42	94.87

The Mpox detection performance of the proposed scheme is separately verified on the raw- and thresholded images and the achieved test-results are presented and discussed. Table II shows the outcome of MNV2 on the raw-image. This table verifies that the Adam+ReLU helps in obtaining a better outcome. Similarly, Table III confirms the merit of Adam+ReLU when thresholded image is considered. In order to graphically verify the overall merit of the developed scheme, Glyph-plot is constructed as in Figure 7. Fig 7(b) confirms that, the overall performance of the Adam+ReLU is better when the pre-processed images are considered, compared to the rew-image result presented in Fig 7(a).



(a) Glyph-plot for Table II (b) Glyph-plot for Table III

Figure 7. Overall performance by the MNV2

The experimental outcome of this study confirms that the Mpox detection performance of the proposed tool is better compared to few chosen earlier works discussed in Section 2. Further, this work implemented a PLDS, which is easy to implement compared to the conventional DL methods found in Section 2. This work implements only a bi-level classification using the MSID and in the future, a multi-class detection can be proposed with the proposed MNV2 with Adam+ReLU to achieve a better skin-infection detection task.

V. CONCLUSION

Automatic detection of skin infection using a chosen AI scheme is highly desirable during the clinical level examination of the disease and its severity. This work proposed a MNV2 supported examination of the Mpox detection scheme on benchmark MSID and the performance of developed tool is verified with raw- and thresholded-images. The experimental investigation of this study is performed using Adam/SGD and ReLU/PReLU and the outcome of this scheme confirms that the Adam+ReLU helps in achieving a better accuracy (98.7%) compared to other methods. Implemented on the chosen image database. In this work, a binary-classification is implemented and in the future, the developed scheme can be considered to implement a multi-class classification using the MSID. Further, the merit of MNV2 can be verified against other PLDS and conventional Dl methods found in the image processing literature.

REFERENCES

[1] https://www.cdc.gov/poxvirus/mpox/clinicians/clinical-recognition.html

[2] https://www.who.int/news-room/fact-sheets/detail/monkeypox

[3] Gupta, A. K., Talukder, M., Rosen, T., & Piguet, V. (2023). Differential Diagnosis, Prevention, and Treatment of mpox (Monkeypox): A Review for Dermatologists. *American Journal of Clinical Dermatology*, 1-16.

[4] Yasmin, F., Hassan, M. M., Hasan, M., Zaman, S., Kaushal, C., El-Shafai, W., & Soliman, N. F. (2023). PoxNet22: A fine-tuned model for the classification of monkeypox disease using transfer learning. *IEEE Access*, *11*, 24053-24076.

[5] Sitaula, C., & Shahi, T. B. (2022). Monkeypox virus detection using pre-trained deep learning-based approaches. *Journal of Medical Systems*, *46*(11), 78.

[6] Yadav, S., & Qidwai, T. (2024). Machine learning-based monkeypox virus image prognosis with feature selection and advanced statistical loss function. *Medicine in Microecology*, *19*, 100098.

[7] Ahsan, M. M., Ali, M. S., Hassan, M. M., Abdullah, T. A., Gupta, K. D., Bagci, U., ... & Soliman, N. F. (2023). Monkeypox diagnosis with interpretable deep learning. *IEEE Access*.

[8] Alharbi, A. H., Towfek, S. K., Abdelhamid, A. A., Ibrahim, A., Eid, M. M., Khafaga, D. S., ... & Saber, M. (2023). Diagnosis of monkeypox disease using transfer learning and binary advanced dipper throated optimization algorithm. *Biomimetics*, *8*(3), 313.

[9] Bala, Diponkor; Hossain, Md Shamim (2023), "Monkeypox Skin Images Dataset (MSID)", Mendeley Data, V6, doi: 10.17632/r9bfpnvyxr.6

[10] Oliva, D., Cuevas, E., Pajares, G., Zaldivar, D., & Osuna, V. (2014). A multilevel thresholding algorithm using electromagnetism optimization. *Neurocomputing*, *139*, 357-381.

[11] Diego Oliva (2024). A multilevel thresholding algorithm using Electromagnetism Optimization (https://www.mathworks.com/matlabcentral/fileexchange/47064-a-multilevel-thresholding-algorithm-using-electromagnetism-optimization), MATLAB Central File Exchange. Retrieved November 18, 2023.

[12] Bala, D., Hossain, M. S., Hossain, M. A., Abdullah, M. I., Rahman, M. M., Manavalan, B., ... & Huang, Z. (2023). MonkeyNet: A robust deep convolutional neural network for monkeypox disease detection and classification. *Neural Networks*, *161*, 757-775.

[13] Sorayaie Azar, A., Naemi, A., Babaei Rikan, S., Bagherzadeh Mohasefi, J., Pirnejad, H., & Wiil, U. K. (2023). Monkeypox detection using deep neural networks. *BMC Infectious Diseases*, *23*(1), 438.

[14] Uysal, F. (2023). Detection of Monkeypox Disease from Human Skin Images with a Hybrid Deep Learning Model. *Diagnostics*, *13*(10), 1772.

[15] Maqsood, S., & Damaševičius, R. (2023, April). Monkeypox Detection and Classification Using Deep Learning Based Features Selection and Fusion Approach. In *2023 IEEE International Systems Conference (SysCon)* (pp. 1-8). IEEE.

[16] Alhasson, H. F., Almozainy, E., Alharbi, M., Almansour, N., Alharbi, S. S., & Khan, R. U. (2023). A Deep Learning-Based Mobile Application for Monkeypox Detection. *Applied Sciences*, *13*(23), 12589.

[17] Eliwa, E. H. I., El Koshiry, A. M., Abd El-Hafeez, T., & Farghaly, H. M. (2023). Utilizing convolutional neural networks to classify monkeypox skin lesions. *Scientific reports*, *13*(1), 14495.

[18] Nayak, T., Chadaga, K., Sampathila, N., Mayrose, H., Gokulkrishnan, N., Prabhu, S., & Umakanth, S. (2023). Deep learning based detection of monkeypox virus using skin lesion images. Medicine in Novel Technology and Devices, 100243.

ISBN 979-8-3503-5094-4

Vision Transformer Based Diabetic Foot-Ulcer Detection: A Study

Ramya Mohan
Department of CSE
Saveetha School of Engineering,
SIMATS, Chennai 602105, TN, India.
Email: ramyanallu@saveetha.com

N. Sri Madhava Raja
LTI Mindtree
Manapakkam, Poonamallee Road
Chennai 600089, TN, India.
Email:srimadhavaraja.nadaradjane@ltimindtree.com

Robertas Damasevicius
Department of Applied Informatics
Vytautas Magnus University
53361 Kaunas, Lithuania
ORCID: 0000-0001-9990-1084

David Taniar
Faculty of Information Technology
Monash University
Melbourne, VIC 3800, Australia
ORCID: 0000-0002-8862-3960

S. Prabha
Department of CSE
Center for Research and Innovation
Saveetha School of Engineering,
SIMATS, Chennai 602105, TN, India.
Email: sprabha3223@ieee.org

V. Rajinikanth
Department of CSE
Saveetha School of Engineering,
SIMATS, Chennai 602105, TN, India.
Email: v.rajinikanth@ieee.org

***Abstract*—The global rise in diabetes is driven by various factors. Poor management can lead to severe complications, emphasizing the urgent need for heightened awareness and enhanced prevention and management strategies. Diabetes can lead to severe foot ulcers (FU), which, if left untreated, may result in incurable wounds or even amputation. The proposed research aims to develop a tool for automatically detect the FU using the Vision Transformer (ViT). The stages of this tool include the following sections; (i) image collection, resizing, and contrast enhancement, (ii) implementation of ViT to extract the image features, and (iii) binary classification and performance confirmation. The planned work is executed with various patch sizes and the reached results are discussed. The performance of the ViT is also verified against the deep-learning methods like VGG16, VGG19, ResNet50 and ResNet101 and the experimental outcome confirms that the ViT approach is efficient in providing better detection accuracy (98.58%).**

Keywords— diabetes, foot-ulcer, vision transformer, deep learning, classification.

I. Introduction

Diabetes is a chronic metabolic complaint categorized by elevated blood glucose levels. Over time, this condition can lead to serious harm to several organs and systems in the body, including the family members [1].

The World Health Organization (WHO) report confirms that, roughly 422 million individuals globally have diabetes, largely in low- and middle-income nations. This disease results in 1.5 million deaths each year. The growing occurrence of diabetes over recent times highlight the crucial need for effective prevention and management strategies, particularly in regions with limited healthcare resources [2].

Diabetic Foot-Ulcer (DFU), a frequent complication of diabetes which case open wounds or sores that develop on the feet, usually on the underside. These ulcers are often the result of various factors such as inadequate circulation, nerve damage (neuropathy), and pressure from footwear or other sources [3,4]. Effective detection of the DFU and its treatment is crucial and untreated DFU will lead to incurable wounds and the worst case lead to amputation.

Diabetic Foot-Ulcer (DFU) is a prevalent complication of diabetes, characterized by open wounds or sores chiefly originate on the feet's undersides. The DFU is developed due to various factors, including poor blood circulation, nerve damage, and pressure from footwear. Detecting DFUs early and starting suitable management quickly is essential. Left untreated, these ulcers can become persistent, non-healing wounds, sometimes leading to amputation in severe cases. Thus, timely identification and supervision of DFUs are imperative to prevent severe difficulties. Scheduled foot health inspections, appropriate foot sanitation, and maintaining blood sugar levels are essential for avoiding and DFUs.

Due to its importance, numerous procedures for detecting DFUs have been developed and implemented by researchers [5,6]. These procedures aim to enhance early detection and treatment. They include imaging techniques, such as infrared thermography and ultrasound, as well as novel wound assessment tools, helping improve DFU management and outcomes. Detection of the DFU using the digital photography is one of the commonly implemented practice and recently, a number of deep-learning (DL) supported methods are implemented to segment and classify the DFU.

Recently, the work of Galdran et al. (2021) [7] proposed a study on DFU detection using the ViT and Data-efficient Image Transformers (DeIT) and confirmed the merit of the ViT on the chosen database. Based on this motivation, this research also employs the ViT supported DFU data examination with various patch dimensions and the outcome of this scheme is compared with the chosen DL methods.

The various phases of this scheme consist the following; (i) image collection, resizing and enhancement, (ii) ViT supported feature extraction and detection, and (iii) binary classification and verification. In this work, the performance of the considered ViT is verified using various patches and the achieved accuracy is verified with the considered against the DL methods like VGG16 (V16), VGG19 (V19), ResNet50 (R50) and ResNet101 (R101) [8].

The investigational outcome of this study confirms that the proposed ViT provided a classification accuracy of 98.58% when a multilayer perceptron (MLP) based classification is implemented. This result is better compared to other DL methods of this study.

The contribution of this research is as follows;

(i) ViT supported DFU detection with different patch dimensions,

(ii) Performance evaluation of ViT and chosen DL methods.

ISBN 979-8-3503-5094-4

Section 2 reviews literature, Section 3 outlines methodology, while Sections 4 and 5 present research results and conclusions, respectively.

II. Earlier Works

The rising global prevalence of diabetes highlights the critical importance of promptly detecting and treating complications like DFU. Early identification and management of DFUs are essential to prevent severe complications, emphasizing the need for regular foot exams and optimal diabetes care.

Digital photograph supported DFU detection using the Artificial-Intelligence (AI) schemes is widely discussed by the researches and these works confirms that the DL approaches works well on classification and segmentation of the DFU in the chosen digital imagery. The earlier works on Dl based DFU detection is summarized in Table I and this confirms that the DL methods are effective in detecting abnormality with better accuracy.

TABLE I. DL SUPPORTED DFU DETECTION PROCEDURES

Reference	DFU detection procedure	Accuracy (%)
Rajinikanth [8]	This study demonstrated binary classification on DFU images supported by PDL, showcasing the effectiveness of this approach for DFU detection.	92.00
Goyal et al. [9]	This study introduced DFUNet, which significantly improved the accuracy of DFU detection, demonstrating superior performance compared to existing methods.	92.50
Goyal et al. [10]	The study suggests using an ensemble Convolutional Neural Network (CNN) for detecting Diabetic Foot Ulcers (DFUs), offering a novel approach to improve detection accuracy and potentially enhance DFU management strategies.	90.30
Biswas et al. [11]	An innovative approach, DFU_MultiNet, is introduced for DFU detection, offering a new method to enhance accuracy in DFU diagnosis.	99.10
Das et al. [12]	A novel technique, DFU_SPNet, is proposed for detecting DFU. This method introduces innovative strategies that could potentially revolutionize the field of DFU detection and management.	96.40
Thotad et al. [13]	The utilization of pre-trained DL (PDL) models for DFU detection was examined.	98.97
Fadhel et al. [14]	Classification of diabetic foot ulcers with PDL and parallel hardware computational tools was presented.	99.81

A comprehensive examination of DFU detection using AI methods is provided in [5,6], emphasizing the necessity of DL approaches for diabetes detection through digital imagery. This study follows a similar approach by utilizing ViT and presents and discusses the obtained results, aligning with the broader trend towards DL-based methods in diabetic healthcare.

III. Methodology

This section details the methodology employed to detect DFU using the ViT with different patch sizes. The study utilizes the DFU database referenced in [15]. Initially, images are enhanced using Otsu's tri-level thresholding, as described in [16,17]. Subsequently, the enhanced images are analysed using the ViT model. This approach aims to improve DFU detection accuracy by the ViT and verifying its accuracy over the chosen DL methods.

Figure 1 depicts the architecture of the proposed scheme and it clearly depicts the process of DFU examination. Initially, the necessary DFU images are collected from the patient using a digital camera and then these images are pre-processed as per the chosen methodology and these images are then considered for the examination. After the examination, the considered ViT tool helps to provide necessary metrics, AC, PR, SE, and SP which is then considered to authenticate the merit of the developed tool.

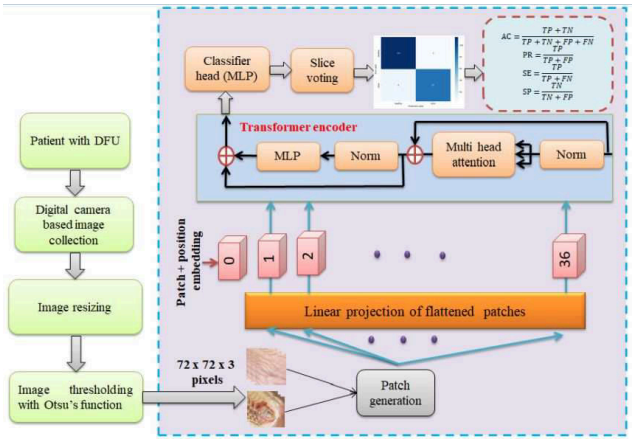


Figure 1. Architecture of proposed DFU detection scheme

A. DFU database

The required images for this study are obtained from the benchmark database found in [15]. This database consist 512 ulcer and 543 healthy images. These images are then augmented to get 1024 ulcer and 1086 healthy images. During the analysis, each image is resized to $72 \times 72 \times 3$ pixels and then these images are preprocessed using a tri-level thresholding based on [16,17]. In this work, 80% of the data is considered for training and remaining 10% data for the validation and testing operations. The sample test images of this database is shown in Figure 2.

Figure 2. Sample test images of healthy and DFU class

ISBN 979-8-3503-5094-4

B. *Image enhancement*

Image enhancement plays a vital role in improving the accuracy when AI supported automatic disease detection process. The methods, like contrast enhancement, and thresholding are common methods to improve the quality of the test images. This work employs Otsu's tri-level thresholding to enhance image pixels. This technique helps to improve the visibility of the diabetic-ulcer region by grouping the similar pixels and this enhanced image is then considered as the input image for the ViT.

C. *ViT Parameter Tuning*

ViT was invented by Dosovitskiy et al. (2020) and it is now widely adopted by the researchers for image examination tasks. This work considered the pre-trained ViT scheme, which needs an image dimension of $72 \times 72 \times 3$ pixels. The initial parameters of the ViT is assigned as follows; learning rate of 0.001, weight decay of 0.0001, batch size of 256, number of epochs of 75, and transformer layers of 8. During the execution, the image patch sizes are chosen as 8, 12, and 16 and the achieved outcomes are presented and discussed. The performance of the ViT is then verified using the DL methods like VGG16, VGG19, ResNet50 and ResNet101 discussed in [8].

D. *Execution and Verification*

The performance of the proposed tool is evaluated by calculating metrics from the confusion matrix (CM), as shown in Figure 1. Metrics, like AC, PR, SE, and SP are calculated from the CM values to assess the proposed DFU detection tool's performance. Furthermore, the tool's overall merit is also assessed using Receiver Operating Characteristic (ROC) curve, which links the True-Positive (TP) rate against the True-Negative (TN) rate. In this study, the presented scheme achieved a superior overlapping area value for the task at hand when the patch dimension is chosen as 8×8.

IV. Result and Discussion

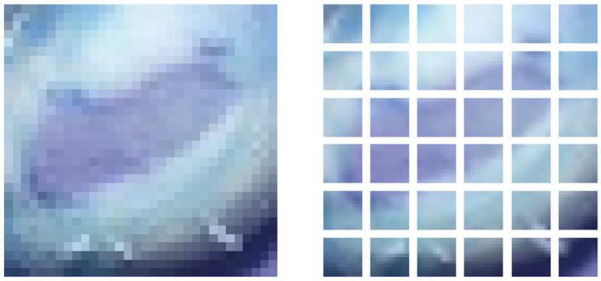
This section of the research presents the experimental outcome of the present study. The image resizing and the Shannon's thresholding are executed using the Matlab-software and the proposed ViT is implemented using Python-software. The proposed work is executed using a workstation with following specification; Inter i5, 20GB of RAM, and 4 GB VRAM.

Initially, all the chosen images are resized to $72 \times 72 \times 3$ pixels and then enhanced using the Otsu's thresholding. After the enhancement, every image is divided into patches based on the chosen patch size. Figure 3 presents the outcome of the pre-processing process. Fig 3(a) and (b) depicts the chosen sample image and the pre-processed image, respectively. Fig 3(c) presents the generated patches for Fig 3(b) and these patches are then examined using the ViT.

The experimental outcome for a chosen patch dimension of 12×12 is presented in Figure 4 and Figure 5. In Fig 4(a) and (b), the training and validation accuracy and loss values for a chosen epoch value of 75 is presented. This shows that, the ViT helps to provide detection accuracy close to 95%.

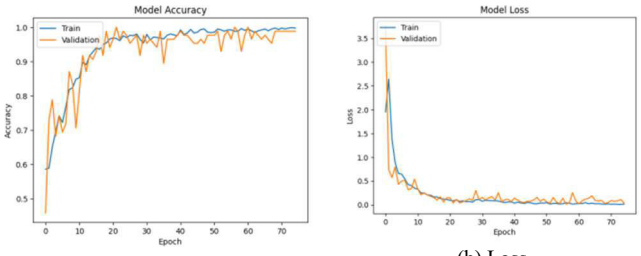
Fig 5(a) and (b) presents the CM and the ROC, which confirms the merit of the chosen patch dimension. This simulation presented a ROC of 0.975, which is a prime measure and it confirms the merit of the proposed scheme along with the metrics computed using the CM values. Similar results are achieved for other chosen patch values in this study.

(a) Image



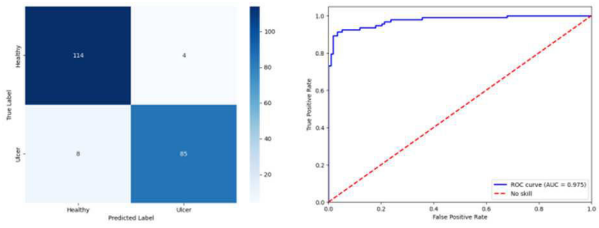
(a) Pre-processed (b) Patches

Figure 3. The chosen test image and the chosen patch value



(a) Accuracy (b) Loss

Figure 4. Experimental outcome for a chosen patch value



(a) CM (b) ROC

Figure 5. Final outcome for a chosen patch value

The experimental outcome achieved in this research is depicted in Table II for ViT and chosen DL methods. This confirms that, a smaller patch size will help to achieve a better detection accuracy compared to a larger patch for the chosen DFU database. The DL methods are executed using SoftMax classifier with a chosen image size of $224 \times 224 \times 3$ pixels. The outcome of this study confirms that, the ViT approach provides a better accuracy compared to the DL methods on the chosen database.

To verify the overall merit, the metrics in the table is then considered to construct the Glyph-plot as in Figure 6, which also confirms that the ViT provides a better outcome than DL. Further, the ViT with lesser patch dimension offers a detection accuracy of 98.58%, which is superior compared with many other methods discussed in Table I.

ISBN 979-8-3503-5094-4

TABLE II. PERFORMANCE EVALUATION OF VIT AND DL

Methods	TP	FN	TN	FP	AC (%)	PR (%)	SE (%)	SP (%)
ViT (8x8)	98	2	110	1	98.58	98.98	98.00	99.09
ViT (12x12)	85	8	114	4	94.31	95.50	91.39	96.61
ViT (16x16)	88	7	109	7	93.36	92.63	92.63	93.96
V16	89	9	103	10	90.99	89.89	90.82	91.15
V19	90	8	98	15	89.10	85.71	91.83	86.72
R50	84	12	93	22	83.88	79.24	87.50	80.87
R101	86	11	97	17	86.73	83.49	88.66	85.09

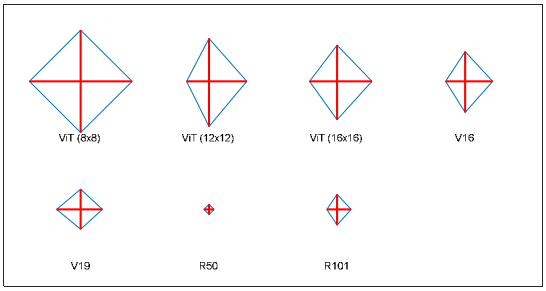


Figure 6. Overall performance evaluation for Table II metrics

The experimental outcome of this research confirms that the ViT with a carefully chosen patch size will help to provide better disease detection accuracy during the medical image examination task. Further, the processing time of the ViT is considerably small compared to the DL methods, due to its lesser test image size. In future, the performance of the DFU detection scheme can be improved by replacing the Otsu's function with other entropy based thresholding methods found in the literature.

V. CONCLUSION

Automatic and accurate detection of the DFU using a chosen AI scheme is highly desirable during the clinical level examination of the foot-ulcers. This work proposed a ViT supported examination of the DFU detection scheme with varied patch dimension and the performance of the proposed scheme is verified and confirmed against few chosen DL methods. In this work, the benchmark DFU database images are considered for the examination and each image is resized to $72 \times 72 \times 3$ pixels and enhanced with Otsu's approach to achieve a faster and accurate results. The experimental investigation of this study confirms that, the ViT with a patch dimension of 8×8 provides improved detection accuracy (98.58%) compared to other methods considered in this study. Further, this result is better when compared to few chosen earlier works in the literature. This confirms that, the ViT based method is efficient in detecting the DFU with a better accuracy and in future the performance of this scheme can be enhanced by replacing the Otsu's with entropy based image pre-processing methods found in the literature.

REFERENCES

[1] M. Lotfy, J. Adeghate, H. Kalasz, J. Singh, and E. Adeghate, "Chronic complications of diabetes mellitus: a mini review", Current diabetes reviews, vol.13, no.1, pp.3-10,2017.

[2] https://www.who.int/health-topics/diabetes#tab=tab_1

[3] K. Ousey et al., "Identifying and treating foot ulcers in patients with diabetes: saving feet, legs and lives. *Journal of wound care"*, vol. *27*, no.Sup5, pp.S1-S52, 2018. https://doi.org/10.12968/jowc.2018.27.Sup5.S1

[4] P.D. Sinwar, "The diabetic foot management–Recent advance", *International Journal of Surgery*, vol.*15*, pp.27-30, 2015. https://doi.org/10.1016/j.ijsu.2015.01.023

[5] J. Zhang, Y. Qiu, L. Peng, Q. Zhou, Z. Wang, and M. Qi, "A comprehensive review of methods based on deep learning for diabetes-related foot ulcers", *Frontiers in Endocrinology*, vol.*13*, pp.945020, 2022. DOI: 10.3389/fendo.2022.945020

[6] M.H. Yap, R. Hachiuma, A. Alavi, R. Brüngel, B. Cassidy, M. Goyal, …, and E. Frank, "Deep learning in diabetic foot ulcers detection: a comprehensive evaluation", *Computers in Biology and Medicine*, vol.*135*, no.104596, 2021. https://doi.org/10.1016/j.compbiomed.2021.104596

[7] Galdran, G. Carneiro, and M.A.G. Ballester, "Convolutional nets versus vision transformers for diabetic foot ulcer classification", Lecture Notes in Computer Science (LNCS), vol.13183, pp.21-29, 2021. https://doi.org/10.1007/978-3-030-94907-5_2

[8] V. Rajinikanth, "Deep transfer learning with fused optimal features for detection of diabetic foot ulcers", *International Journal of Clinical Medical Research*, vol.*1, no.3, pp.*116-122, 2023. https://doi.org/10.61466/ijcmr1030014

[9] M. Goyal, N.D. Reeves, A.K. Davison, S. Rajbhandari, J. Spragg, and M.H. Yap, "Dfunet: Convolutional neural networks for diabetic foot ulcer classification", *IEEE Transactions on Emerging Topics in Computational Intelligence*, vol.*4*, no.5, pp.728-739, 2018. DOI: 10.1109/TETCI.2018.2866254

[10] M. Goyal, N.D. Reeves, S. Rajbhandari, N. Ahmad, C. Wang, and M.H. Yap, "Recognition of ischaemia and infection in diabetic foot ulcers: Dataset and techniques", *Computers in biology and medicine*, vol.*117*, pp.103616, 2020. https://doi.org/10.1016/j.compbiomed.2020.103616

[11] S. Biswas, R. Mostafiz, B.K. Paul, K.M.M. Uddin, M.M. Rahman, and F.N.U. Shariful, "DFU MultiNet: A deep neural network approach for detecting diabetic foot ulcers through multi-scale feature fusion using the DFU dataset", *Intelligence-Based Medicine*, vol.*8*, pp.100128, 2023. https://doi.org/10.1016/j.ibmed.2023.100128

[12] S.K. Das, P. Roy, and A.K. Mishra, "DFU SPNet: A stacked parallel convolution layers based CNN to improve Diabetic Foot Ulcer classification", *ICT Express*, vol.*8, no.2, pp.*271-275, 2022. https://doi.org/10.1016/j.icte.2021.08.022

[13] P.N. Thotad, G.R. Bharamagoudar, and B.S. Anami, "Diabetic foot ulcer detection using deep learning approaches", *Sensors International"*, vol.*4*, pp.100210, 2023. https://doi.org/10.1016/j.sintl.2022.100210

[14] M.A. Fadhel, L. Alzubaidi, Y. Gu, J. Santamaría, and Y. Duan, "Real-time diabetic foot ulcer classification based on deep learning & parallel hardware computational tools", *Multimedia Tools and Applications*, pp.1-26, 2024. https://doi.org/10.1007/s11042-024-18304-x

[15] https://www.kaggle.com/datasets/laithjj/diabetic-foot-ulcer-dfu

[16] D.Oliva, E. Cuevas, G. Pajares, D. Zaldivar, and V. Osuna, "A multilevel thresholding algorithm using electromagnetism optimization", *Neurocomputing*, *139*, 357-381, 2014.

[17] Diego Oliva (2024). A multilevel thresholding algorithm using Electromagnetism
Optimization (https://www.mathworks.com/matlabcentral/fileexchange/47064-a-multilevel-thresholding-algorithm-using-electromagnetism-optimization), MATLAB Central File Exchange. Retrieved Fbruary 17, 2024.

ISBN 979-8-3503-5094-4

CONVOLUTION NEURAL NETWORK (CNN) BASED POSTURE DETECTION SYSTEM TO PREVENT HUNCHED POSTURE SYNDROMES

Mr.C.Sathish Kumar[1],
Assistant professor,
Department of Biomedical Engineering,
Adhiyamaan College of Engineering,
Dr. M.G.R Nagar, Hosur.
csathishkumar08@gmail.com

Mrs.Selva Sherin T[2],
Assistant professor,
Department of Biomedical Engineering,
Adhiyamaan College of Engineering,
Dr. M.G.R Nagar, Hosur.
tssherin26@gmail.com

Mr.Suriya V[3],
UG Student,
Department of Biomedical Engineering,
Adhiyamaan College of Engineering,
Dr. M.G.R Nagar, Hosur.
suriya.v153@gmail.com

Mr.Vinu P[4],
UG Student,
Department of Biomedical Engineering,
Adhiyamaan College of Engineering,
Dr. M.G.R Nagar, Hosur.
vinupandian2020@gmail.com

Mr.Dhanush M[5]
UG Student,
Department of Biomedical Engineering,
Adhiyamaan College of Engineering,
Dr. M.G.R Nagar, Hosur.
dhanushmbme@gmail.com

Mr.Karthik S[6],
Research Scholar,
Sri Ramakrishna Engineering College,
Coimbatore.
karthikheyram@gmail.com

Mr.Surenderkumar S[7],
Assistant Professor,
Department of ECE,
Sri Eshwar College of Engineering,
Coimbatore.
surenderkumar.s@sece.ac.in

Mr.Arun Kumar S[8],
Assistant Professor,
Department of ECE,
Coimbatore Institute of Engineering and Technology,
Coimbatore.
arunkumarselvaneie@gmail.com

Abstract--***This paper presents a solution addressing the challenges associated with increased screen time during remote work. Leveraging specialized hardware and convolutional neural networks, our real-time postural detection system concentrates on monitoring the postures of the neck, shoulders, and arms. It offers recommendations to prevent health issues resulting from poor ergonomics. With an accuracy exceeding 80%, our low-power solution comprises a hardware device equipped with an LCD and buzzer for instant posture alerts, complemented by online notifications for prolonged bending. The ultimate goal is to advocate for a healthier remote working environment.***

Keywords--***Deep learning, Open pose posture detect node point, Buzzer, Nano-microcontroller, Flex sensor and LCD display.***

I. INTRODUCTION

This system leverages specialized hardware for real-time video processing, employing the OpenPose detection framework to analyze the posture of individuals, with a specific focus on the neck, shoulders, and arms. By providing real-time feedback and recommendations, the

ISBN 979-8-3503-5094-4

system aims to mitigate potential health issues arising from poor posture. Notably, the results demonstrate the system's efficiency, achieving real-time video processing at 25 frames per second with minimal power consumption. Moreover, it boasts an accuracy rate exceeding 80% in pattern detection. The integration of hardware components, such as an LCD display and a buzzer, enhances the system's functionality, allowing for immediate alerts in the event of poor posture detection. The inclusion of a flex sensor enables the system to continuously monitor the bending position of individuals, prompting the buzzer to activate as a warning. Furthermore, this project incorporates email notifications for remote monitoring, contributing to the creation of a healthier working environment. In summary, this innovative system combines deep learning, OpenPose pose detection, Region Proposal Network (RPN), and hardware components like a buzzer, Nano-microcontroller, flex sensor, and LCD display to develop an effective Posture Detection System, ultimately fostering improved workplace ergonomics and employee well-being.

II. EXISTING SYSTEM

The application of a single algorithm resulted in a low level of accuracy in proper detection. Only the identification of hunchback was available, and no alerting system was implemented. It utilized both hardware and software, leading to significant costs. Training for each aspect of detecting hunchback was challenging, as it sometimes showed incorrect detections due to improper hunchback detection. The design of the model was overly complex, creating difficulties in operating the process. The model, being a device, was complex to use due to the implementation of both hardware and software. The YOLOv3 algorithm used in the software is intricate, requiring considerable time to understand. Since it involves both software and hardware, the implementation cost is high.

III. PROPOSED SYSTEM

Live camera surveillance has been implemented at multiple workplaces to monitor employees' sitting positions using a powerful convolutional neural network. An OpenPose system is being utilized for this purpose, implemented with OpenPose and CNN for precise hunch position identification. This method captures frames using a webcam placed at eye level on a user's computer screen, with users sitting in height-adjustable chairs in front of their computers. The system offers real-time video processing with automated posture analysis, collecting incorrect poses during work to enhance postures and potentially improve CNNs for detailed pattern recognition, including parameters like the evaluation of the spine. This system combines hardware and software for accurate hunch detection. A coded Nano micro-controller, connected to a flex sensor, activates a buzzer and displays hunch positions on an LCD when detected. The integration of OpenPose pose detection, Region Proposal Network (RPN), and hardware components like a buzzer, Nano-microcontroller, Flex sensor, and LCD display results in an effective posture detection system, ultimately fostering improved workplace ergonomics and employee well-being.

IV. BLOCK DIAGRAM DESCRIPTION

The components of the circuit diagram include the Nano microcontroller, power supply, regulator 7805, LCD Display, buzzer, and flex sensor. In the input power supply, 12v is given to the regulator, which converts it to 5v connected to the controller, and the negative pin is connected to GND. The LCD Display has 8 pins, with the VCC pin connected to ground, VSS pin connected to 5v, and the remaining data pins connected to the controller's digital pins. The buzzer is connected to the controller's digital pin, and the GND is connected to the ground. The flex sensor has 3 pins, with the positive pin connected to the 5v supply, the GND pin connected to Ground, and the remaining data pin connected to the controller's Analog pin. The integration of OpenPose pose detection, Region Proposal Network (RPN), and hardware components like a buzzer, Nano-microcontroller, flex sensor, and LCD display is designed to develop

ISBN 979-8-3503-5094-4

an effective Posture Detection System, ultimately fostering improved workplace ergonomics and employee well-being. An OpenPose system is being utilized for this purpose, implemented with OpenPose and CNN for precise hunch position identification. This method captures frames using a webcam placed at eye level on a user's computer screen.

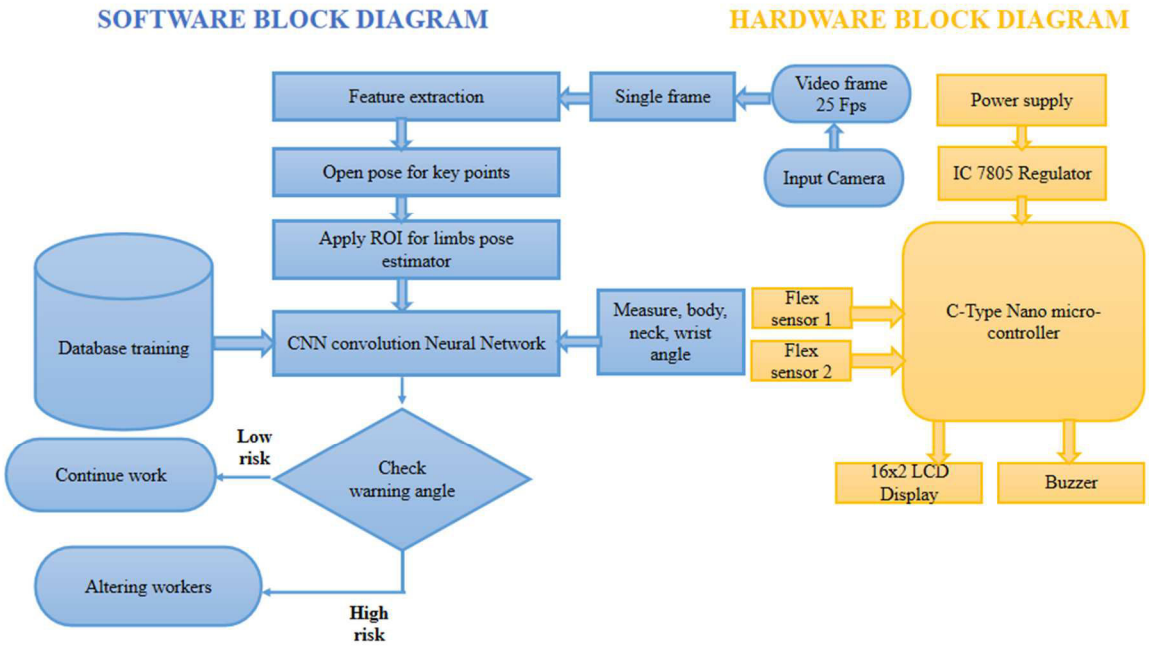


Fig.1. Block Diagram of Proposed System

A. Power Supply:

Input Voltage represents the incoming electrical power from an external source, often the mains supply. Output Voltage regulated DC voltage provided to the electronic circuits or devices.

B. IC 7805 Regulator:

The IC 7805 voltage regulator is a popular integrated circuit used to provide a stable and regulated output voltage of +5 volts. Depending on the specific model and manufacturer, IC 7805 regulators can handle currents in the range of 1A to 2A. The IC 7805 comes in a standard three-terminal package with an input (Vin), ground (GND), and output (Vout) pin configuration.

C. Nano Microcontroller:

This Arduino-compatible Nano V3.0 micro-controller is a small, complete, micro-controller board based on the ATmega328 MCU. It has more or less the same functionality as the Arduino Duemilanove, but in a smaller package.

D. Input Camera:

The input camera can be represented as a block or module. This block typically captures and processes visual information. It can be connected to other blocks such as image processing, feature extraction, or object recognition modules, depending on the application. The camera input block serves as the starting point for visual data in the system, with its output feeding into subsequent processing stages.

E. Video frame 25Fps:

This block signifies the input stream of video frames, capturing visual information at a rate of 25 frames per second. This block is typically connected to subsequent blocks such as video processing, analysis, or encoding modules. The 25 FPS video frame input serves as a key component in systems requiring real-time or standard video capture and processing.

ISBN 979-8-3503-5094-4

F. Single frame:

This block signifies the input of a static image or a snapshot at a particular moment in time. It can be connected to various processing blocks such as image analysis, feature extraction, or further downstream modules depending on the specific application. The single frame input is valuable in scenarios where capturing and processing individual images, rather than a continuous video stream, is essential.

G. Feature extraction:

It is responsible for identifying and extracting relevant patterns, characteristics, or information from the input data. The output from the feature extraction block is often a set of distinctive features that capture important aspects of the input data. These extracted features are then passed on to subsequent blocks in the diagram, such as classification, recognition, or decision-making modules.

H. OpenPose Algorithm:

OpenPose is a popular library for real-time multi-person key point detection and pose estimation from images or videos. This block takes the input data, such as video frames or images, and processes it using OpenPose to identify key points on the human body, capturing details like joint positions. The output from this block includes the detected key points, which can be essential for understanding body poses and movements.

I. CNN:

This block is typically situated after the input data source, such as images or frames from a video stream. The CNN block performs convolutional operations, feature extraction, and hierarchical learning to capture complex patterns and features from the input data. It's particularly effective for tasks like image recognition, object detection, and classification. The output from the CNN block, often a set of high-level features, is then passed on to subsequent blocks for further processing, such as fully connected layers for classification or additional layers for fine-tuning and interpretation.

J. Data Base Trainings:

This block typically occurs in the initial stages of a machine learning system. The database training block involves feeding labelled data, usually stored in a training database, into a machine learning model. This process enables the model to learn and adjust its parameters based on the patterns and relationships present in the training data. The output from this block is a trained model with improved capabilities for making predictions or classifications. The trained model can then be integrated into subsequent blocks of the diagram, such as inference or decision-making modules, where it processes new, unseen data. Database training is fundamental for enhancing the performance and generalization ability of machine learning models across various applications, including image recognition, natural language processing, and many others.

K. Flex Sensor:

This block serves as a sensory input component, detecting changes in flex or bending. The output from the flex sensor block is typically analog or digital signals representing the degree of flexion or bending. This information can then be passed on to subsequent blocks in the diagram, depending on the specific application.

*L. 16*2 LCD Display:*

This block serves as an output interface, displaying information in a two-line, 16-character format. The input to the LCD display block typically comes from various sources, such as a microcontroller or a sensor module. Information to be displayed is sent to the LCD, and the visual output is presented on the screen.

M. Buzzer:

This block serves as an output component that generates audible alerts or signals. The input to the buzzer block typically comes from various sources, such as a microcontroller,

ISBN 979-8-3503-5094-4

sensor module, or an alarm system. When triggered, the buzzer produces a sound or tone, providing an auditory indication.

V. CIRCUIT DESCRIPTION

The circuit diagram comprises several key components, including the Nano microcontroller, power supply, 7805 regulator, LCD Display, buzzer, and flex sensor. In the power supply section, a 12V input is regulated down to 5V using the 7805 regulator. The 5V output is connected to the controller, and the regulator's negative pin is linked to the ground (GND) for proper grounding. The LCD Display features 8 pins. The VCC pin is connected to the ground, the VSS pin is linked to the 5V supply, and the remaining data pins are connected to digital pins on the controller to facilitate communication. The buzzer is integrated into the system by connecting it to a digital pin on the controller, with its GND connection established to the ground for proper functioning. The flex sensor, consisting of 3 pins, has its positive pin connected to the 5V power supply, the GND pin connected to the ground, and the remaining data pin is linked to an analog pin on the controller to enable measurement and communication.

VI. WORKING PRINCIPLE

This project operates on an advanced technological framework that integrates deep learning and specialized hardware to oversee and enhance the postures of remote workers. The process begins by capturing video frames at a steady rate of 25 frames per second, subsequently identifying crucial body landmarks using OpenPose technology. These identified points undergo thorough analysis through a Convolutional Neural Network (CNN), which measures various body angles, such as those pertaining to the neck, shoulders, and wrists. In cases where the system identifies incorrect postures exceeding predetermined safe limits, it activates warnings. Users receive feedback through visual indications displayed on a 16x2 LCD screen, along with auditory alerts delivered via a buzzer. The central control unit, a C-Type Nano Micro-controller, efficiently orchestrates the entire system, and the utilization of flex sensors assists in detecting specific body movements. Through its continuous monitoring and guidance, this project aspires to create a more ergonomic and healthier teleworking environment, mitigating the associated health risks stemming from poor posture.

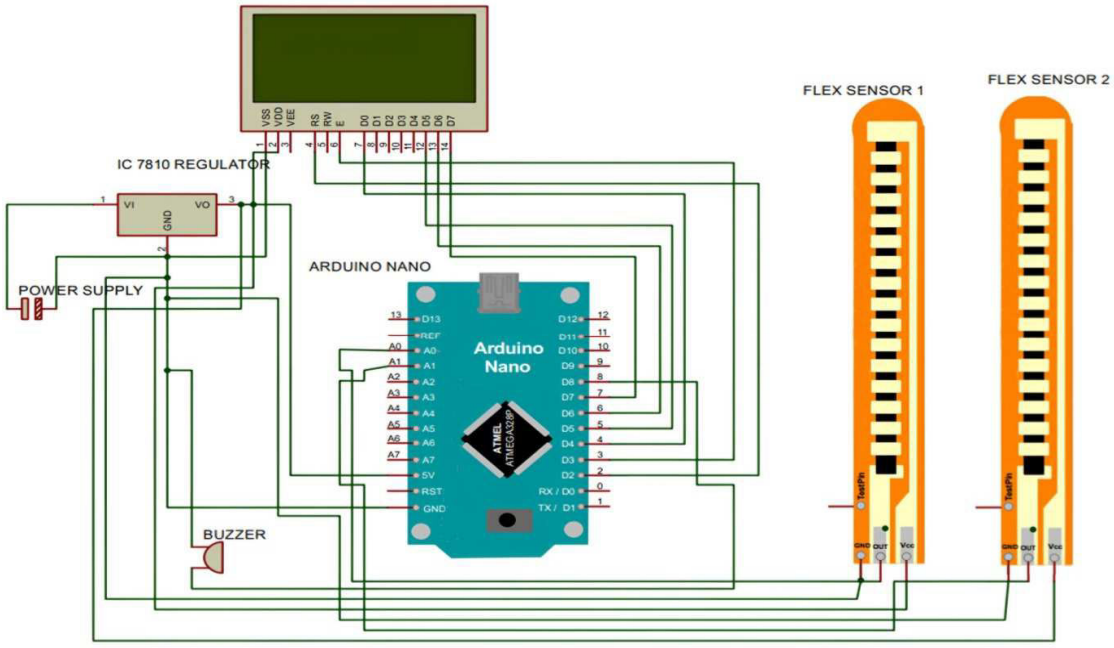


Fig.2. Circuit Diagram

ISBN 979-8-3503-5094-4

VII. RESULT AND DISCUSSION

The project showcases the ability to achieve real-time video processing at a speed of up to 25 frames per second with minimal latency. Moreover, the accuracy of posture detection surpasses 80% when benchmarked against established standards.

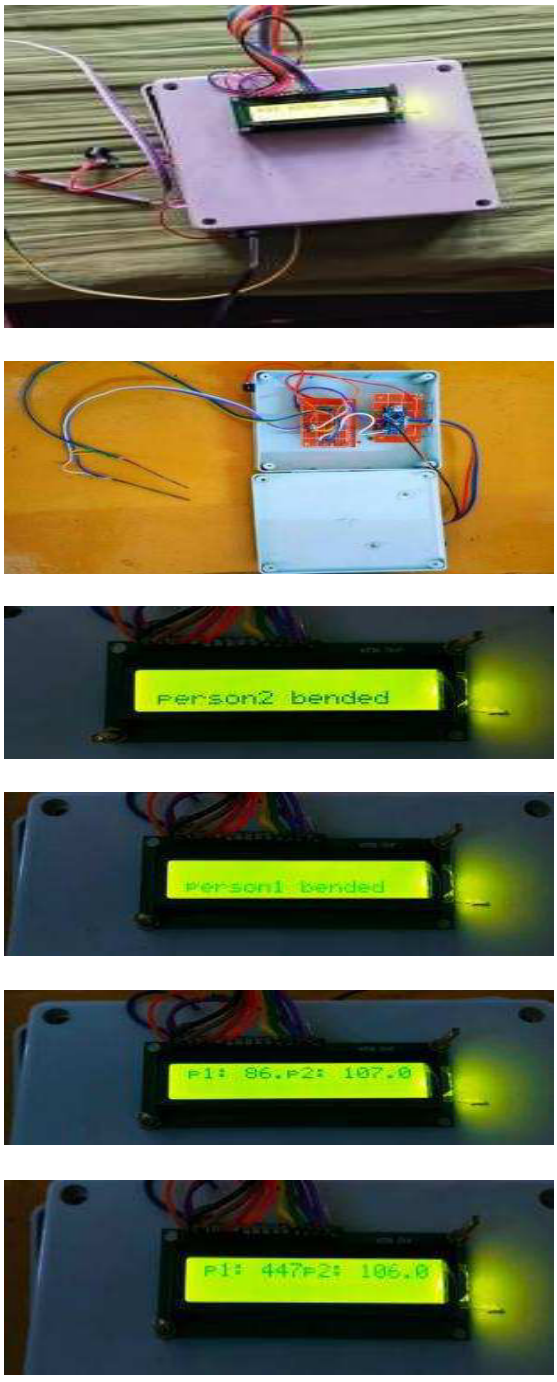


Fig.6. Result

Additionally, the project incorporates a hardware component for real-time hunch monitoring, utilizing a flex sensor to detect an individual's bending position. When a person consistently maintains a posture below the specified threshold angle (set at 70 degrees) between the shoulders, hips, and knees, an alert system is triggered. This alert system comprises an LCD display and a buzzer to promptly notify the individual of poor posture.

Furthermore, the project integrates a software component that delivers real-time online notifications to a designated email address, ensuring the individual stays informed about their posture status. The primary objective of this endeavor is to enhance the health and ergonomics of the working environment, particularly in teleworking scenarios. By amalgamating real-time posture detection and alert features, the project aims to cultivate awareness and promote better postural habits, thereby assisting workers in mitigating potential health issues associated with prolonged computer use and incorrect posture.

VIII. CONCLUSION

Prolonged inactivity may result in muscle fatigue and tension, leading to compromised posture, especially for individuals engaged in extended periods of fixed positions. Addressing this issue, our proposed solution involves a posture detection project achieving accuracy exceeding 95%. Furthermore, the hardware application's successful real-time testing demonstrates its autonomy and low power consumption capabilities, ensuring efficient posture detection and monitoring without the need for computer intervention.

In future scope for a CNN-based posture detection system is promising, especially in preventing hunched posture syndromes. Advancements may include real-time feedback for users, integration with wearable devices, and personalized corrective exercises based on individual posture data. Collaborations with healthcare professionals for comprehensive solutions and expanding the system to address other ergonomic issues could enhance its impact. Continuous research and development in computer vision and AI technologies will likely contribute to the refinement and widespread adoption of such systems.

ISBN 979-8-3503-5094-4

IX. REFERENCES

[1] Enrique Pinero-Fuentes, Salvador Canas-Moreno, Manuel Domínguez-Morales, "A Deep-Learning Based Posture Detection System for Preventing Telework-Related Musculoskeletal Disorders," MDPI Journals, Vol. 21 , Issue.15 ,August 2021.

[2]Kulikajevas A, Maskeliunas R, Damasevicius R, "Detection of sitting posture using hierarchical image composition and deep learning," PeerJ Comput. Sci.,Vol.7, March 2021.

[3] Rikiya Yamashita, Mizuho Nishio, Richard Kinh Gian Do & Kaori Togashi, "Convolutional neural networks: an overview and application in radiology," Springer, Vol.9, pp.611-621,June 2018.

[4] Keison Tang, Arjun Kumar, Muhammad Nadeem, ORCID and Issam Maaz,"CNN-Based Smart Sleep Posture Recognition System," MDPI Journals, Vol. 2 ,Issue.1, pp.119-139, February 2021.

[5] Yantao Yu, Waleed Umer, Xincong Yang, Maxwell Fordjour Antwi-Afari,"Posture-related data collection methods for construction workers: A review," Elsevier, Vol.124, April 2021.

[6] Qisong Hu, Xiaochen Tang, Wei Tang, "A Smart Chair Sitting Posture Recognition System Using Flex Sensors and FPGA Implemented Artificial Neural Network," IEEE Journal, Vol.20, Issue.14, pp.99, March 2020.

[7] Yuanhua Li, "A New Physical Posture Recognition Method Based on Feature Complement-oriented Convolutional Neural Network," Journal of Applied Science and Engineering, Vol. 24, Issue.1, pp.83-89, February 2021.

[8] Weiyan Ren, Ou Ma, Hongxin Ji, Xinyuan Liu, "Human Posture Recognition Using a Hybrid of Fuzzy Logic and Machine Learning Approaches," IEEE Access, Vol.8, 2011.

ISBN 979-8-3503-5094-4

Mathematical Modeling of the Impact of a 5 Day Interim Relaxation at 21 Days in 42 days Lockdown Strategy on COVID-19 Transmission in India

Flavia Immaculyne
Quality Engineer, Schlumberger- Reda production system, Benoi Cresent Singapore
immaculyne@gmail.com

BakiyaAmbikapathy
Department of Electronics and Communication Engineering, Vel Tech Rangarajan Dr Sagunthala R&D Institute of Science and Technology, Chennai, India
bakiya88@gmail.com

Kamalanand Krishnamurthy
Department of Instrumentation Engineering, MIT Campus, Anna University, Chennai, India.
kamalanand@mitindia.edu

Lourduraj De Britto
Indian Council of Medical research,Vector Control research Centre, puducherry, India
rljbritto@gmail.com

Abstract— ***As of the first week of May 2020, the COVID-19 pandemic is present in 212 nations globally. Many countries have enforced lockdown measures strategically to slow down the epidemic's spread, with the objective of relieving pressure on healthcare systems. Mathematical modelling serves as a useful tool for the prediction of the future numbers of the infected population and for analyzing the effects of various intervention strategies. In this work, a mathematical model based on ordinary differential equations has been developed for analyzing the impact of an interim relaxation of 5 days at 21 days in a 42 days lockdown period. Results of the simulation demonstrate that the relaxation of 5 days during the 42 days lockdown does not result in any adverse effects if the transmission rate remains the same. However, the model-based analysis suggests that an increase in the transmission rate by a factor of 1.2 times during the relaxation period due to the possible aggregation of the susceptible with the infected, it could shift the growth pattern of infected cases from sub-exponential to exponential.***

Keywords- COVID-19, Mathematical modelling, Lockdown, Interim relaxation, Transmission, Differential equations

I. INTRODUCTION

Unlike the SARS-CoV-1 and MERS outbreaks, there was a lag in identifying and reporting the emergence of the new SARS-CoV-2 virus, despite significant community transmission happening in China. Therefore, the pandemic could not be averted and by the time several countries are imposed lockdown, the epidemic had gone through the stage-3 by last week of February 2020. The World Health Organization (WHO) designated COVID-19 as a global public health emergency on January 30, 2020, and subsequently classified it as a pandemic on March 11, 2020 [1]. As of May 7, 2020, COVID-19 had been confirmed in 212 countries across the globe, affecting more than 3.5 million people, and resulting in a death toll exceeding 350 thousand [2]. Americas and European countries are worst affected during this pandemic and in general, there has been a slow epidemic progression in South-East Asian countries. India's first COVID-19 case was documented on January 30, 2020, and as of May 7, 2020, the Ministry of Health and Family Welfare (MHFW) had verified approximately 52,952 cases and 1,783 deaths [2].

Currently, there are vaccines accessible for SARS-CoV-2 infection. Nevertheless, there are numerous difficulties in handling the emergence of new variants of the virus. As a result, many countries have imposed social distancing measures as strategies to curb the spread of the outbreak.

The CDC recommends 2meters of social distancing to prevent aerosol spread. However, the lockdown strategy is determined depending on the local situation in each country though in principle it is agreed that the duration of the lockdown is not less than two weeks, the estimated incubation period of the COVID-19 infection. Guest et al. (2020) [3] suggest courageous political action based on the information for three critical measures, mandatory stay-at-home, increasing the number of tests and improving the healthcare facility. Taking into account the reported cases up to the fourth week of March 2020, which were confined to individuals with confirmed exposure, it was concluded that India's COVID-19 epidemic was still in stage-2. In response, the Government of India implemented a bold decision to initiate a 21-day nationwide lockdown starting from midnight on March 24, 2020, subsequently extending it for another 21 days, and later extending it further until May 17, 2020. However, during the lockdown period, services for the supply of the essential commodities were open for the public. The major goal of the lockdown is to reduce the reproduction number (R0) by modifying the transmission dynamics for the mitigation followed by the suppression of the epidemic.

Mathematical modeling is an invaluable tool for

ISBN 979-8-3503-5094-4

comprehending the transmission dynamics of epidemic diseases. Various modeling techniques have been investigated in past outbreaks like H1N1, SARS (severe acute respiratory syndrome), Ebola, MERS (Middle East Respiratory Syndrome), and H7N9, among others., [4-8]. During the current COVID-19 pandemic, researchers have explored different mathematical models such as the logistic model, exponential model, Susceptible-Infected-Removed (SIR) model, etc., as an epidemiological tool [9]. Numerous researchers have employed deterministic models rather than stochastic ones to forecast infection outbreaks. They favor deterministic models because of their lower computational complexity and improved ability to calibrate model parameters. Ivorra et al. (2020) [10] have used the deterministic compartmental model for analyzing the spread of the SARS-nCoV-2, to estimate the epidemiological parameters. Victor (2020) [11] applied a deterministic endemic model to study COVID-19, revealing a potential decrease in secondary infections. Additionally, other researchers have employed artificial intelligence tools to forecast the trajectory of the SARs-nCoV-2 epidemic and predict the clinical severity of SARs-nCoV-2 to aid healthcare providers [12-13].

Exponential modelling is a highly useful modelling strategy and has been utilized applied in a variety of fields such as epidemiology [14], population dynamics and social networking [15]. In [16] developed a differential equation based mathematical model incorporating the data from 14 countries, validated with the actual cases in thirteen countries and adopted it to predict the impact of different lockdown periods on the novSARS-CoV-2 epidemic in India. Though the prediction indicates those proposed 42 days lockdown drastically slows down the epidemic and prevents the progression to exponential transmission, several social issues emerge [17]. The main goal of this work is to create a mathematical model and analyze the effects of an interim relaxation of 5 days in a lockdown intervention of 42 days, incorporating the reported cases in India to the first lockdown on 24th March 2020.

II. Methodology

This study utilized a model structure derived from ordinary differential equations sourced from [16]. Model parameters were estimated using the Prediction Error Method [18] and data collected from documented COVID-19 cases in India between February 2, 2020, and a 45-day period thereafter. Subsequently, the developed model was employed to analyze the effects of a 42-day intervention period with and without a 5-day interim relaxation period on COVID-19 transmission. The developed model is of the following form:

$$\frac{dx(t)}{dt} = d.r(t).x(t) - x(t) + C - k(t) \qquad (1)$$

where, $x(t)$ referred as infected individuals at time t, $\frac{dx(t)}{dt}$ referred as the rate of change of infected individuals, d represents the rate of the spread of infection, $r(t)$ represents the effect of the increase in infection transmission due to aggregation of people on particular days, C is a constant representing other factors such as cross immunity, temperature, etc. $k(t)$ is the force of the lockdown intervention. Fig. 1a shows the intervention strategy of lockdown for a continuous period of 42 days and Figure 1b shows the lockdown strategy of 42 days with a 5-day interim relaxation period. Furthermore, the model incorporated a theoretical situation where individuals aggregate during the relaxation period. It was assumed that such aggregation, where susceptible individuals come into contact with infected ones, could potentially increase the transmission rate by a factor of 1.2.

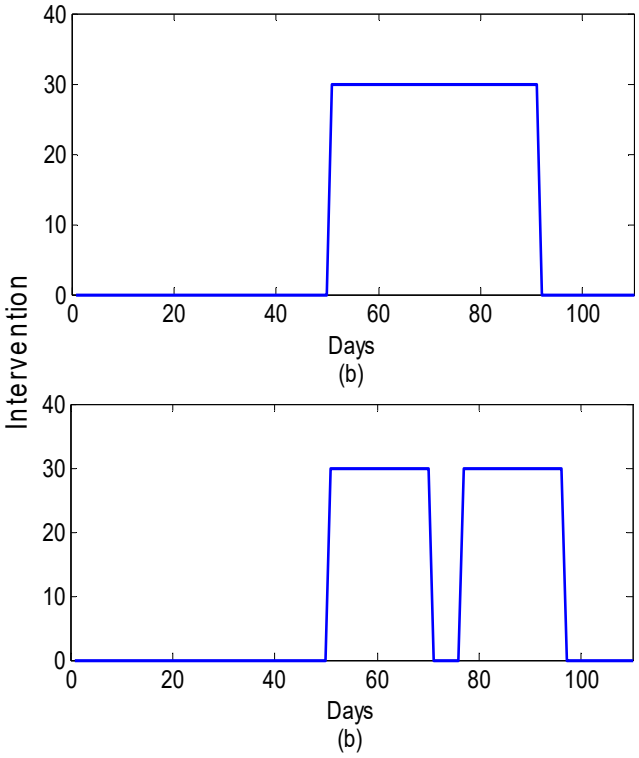


Figure 1: The relation between the single lock down and the lock down with a relaxation (a) 42 days of continuous "lock down" 1 (b) with a 5 days interim relaxation of intervention

III. Results and Discussion

Using the confirmed cases reported in India from February 2nd to March 17th, 2020, we forecasted the trajectory of the COVID-19 epidemic in India under two distinct scenarios (i) 42 days continues lockdown and (ii) 42 days lockdown with an interim relaxation of 5 days at 21 days. Fig. 2 depicts the predicted COVID-19 cases for India for both these interventions. The findings indicate that there are no notable differences in the projected number of infected cases among the intervention strategies examined, assuming no alterations in exposure/transmission rates. However, if an escalation of COVID-19 transmission by a factor of 1.2X during the relaxation period is accounted for due to gatherings of susceptible and infected individuals in various settings like mass transit, shopping centers, youth gatherings, and religious or family events, the model predicts a significant surge in the projected number of infected cases, as illustrated in Figure 3.

ISBN 979-8-3503-5094-4

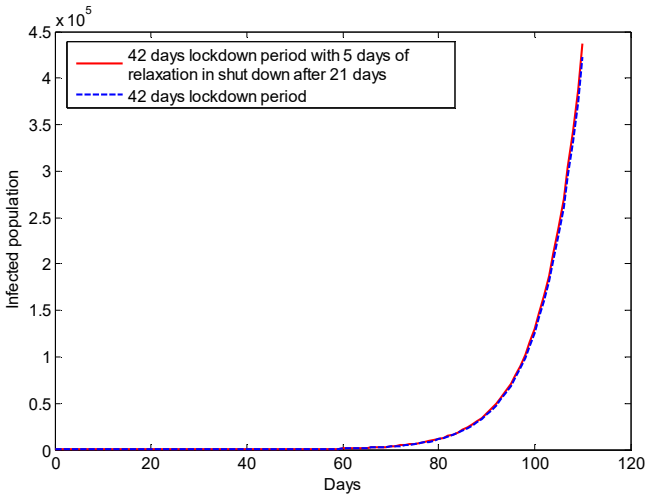


Figure 2: The predicted COVID19 cases in India because of 42 days "lock down" with and without a relaxation period of 5 days. Prediction outcome of Day0 to Day110 for India. (i.e. 02nd February to 20thMay 2020).

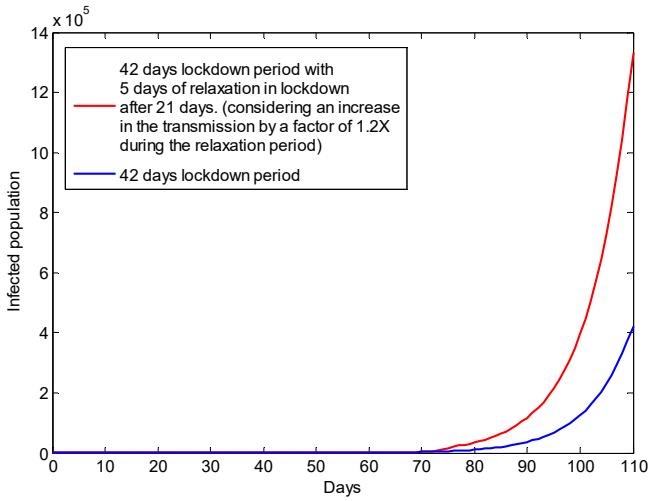


Figure 3: The effect of an increase in the transmission of COVID19 by a factor of 1.2X during the relaxation period (Prediction outcome of Day0 to Day110 for India).

Data from several Western countries that have gone through the exponential transmission indicate that imposing lockdown after reaching the peak of the stage-3 may not have any impact on COVID-19 transmission. Several studies have established that there is a person to person transmission and therefore, transmission through droplet nuclei cannot be ruled out [19-20]. Research has demonstrated that individuals infected with the virus carried a median novel SARS-CoV-2 viral load of 3.3×106 copies/mL in their saliva, with over 90% of infected individuals harboring the virus in their saliva [22].Burke et.al. (2020) [23] has shown a 12% attack rate by active monitoring of the 445 close contacts exposed 10 positive COVID-19 cases. India also has the experience of the strength of transmission in the religious congregation.

On interim relaxation for 5 days, the train and bus services will be resumed at least partially, and these services would promote the aggregation of the population. Zhao and colleagues (2000) [24] discovered a significant association between train travel and the occurrence of COVID-19 cases in China. Indian Railways is the second largest network in the world, and it carries more than 23 million passengers every day by operating more than 12,000 trains (Fig. 4). It is almost comparable to moving the entire population of a country like Taiwan, Sri Lanka, or Australia. There will be around120 thousand running staff on board or the stations and these trains will pass through more than 7 thousand stations. In long-distance travel, there will be an aggregation of people for a continuous period 24 to 36 hours which will further promote the transmission. Such an environment is expected to promote at least 1.2X transmission as depicted in Figure-3. Even if only 30% of the train services are resumed with 50% occupancy about 3.5 million people will be exposed in a single day. COVID-19 has pan-India spread with a focus on certain locations and the sources of such transmissions have been to some extent identified. India continues to claim only clustering of the cases and the community transmission is not established (WHO status report). Under the current epidemiological situation, the Government of India would have to take a well-informed post-lockdown strategy.

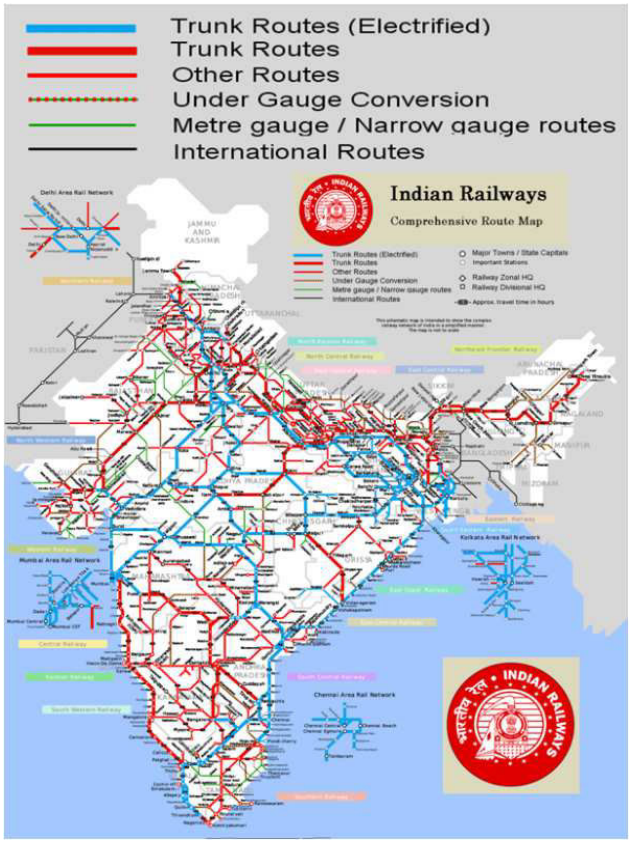


Figure 4: The atlas of the Indian railways network. It expands through 123,236 kilometers in 17 zones (Source: [21])

Our modelling based on the experiences from the world's second-most populous country in the world indicates that any country that is going through the stage-2 novSARS-Cov-2 transmission prefer to consider lockdown period as an intervention, they must ensure that there is no enhanced exposure before lockdown and also the exposure during the lockdown period is maintained at minimal level. We

ISBN 979-8-3503-5094-4

observed even 1.2X transmission through the aggregation of the susceptible trigger exponential transmission as novSARS-CoV-2 is highly infectious. The best option is to promote self-reporting, self-quarantine followed by aggressive contact tracing. The government must impose a ban on the aggregation of the people, especially, the religious gatherings, family functions and shopping malls especially the air-conditioned establishments. Domestic travels must be need-based, and the trains and buses must carry only one-third occupancy during the relaxation period of 5 days. There must be a ban on the air-conditioned environment in all the passenger transports. These stringent measures on domestic travelling and functions may avert an exponential transmission to a certain extent. This will provide an opportunity to reorganize the health infrastructure to attend on the hospitalized cases in any thickly populated country. If the transmission is not controlled by stringent measures, the exponential transmission will lead to a sudden surge in hospital admission resulting in tremendous strain on the health infrastructure.

IV. CONCLUSION

Predictions of epidemics based on mathematical modelling provide an opportunity to implement appropriate interventions at the right time. The rapid global spread of novSARS-CoV-2 can be attributed to the extensive international travel network for both business and education. In the absence of an effective vaccine, it is desirable to analyze the outcomes of various possible lockdown intervention scenarios for the Governments to make well-informed decisions. The results of the mathematical model, employing ordinary differential equations, indicate that integrating a 5-day relaxation period into a 42-day strict lockdown does not result in a notable increase in the number of infected individuals. This holds true as long as exposure remains minimal both before the lockdown and during the interim relaxation period. However, any major aggregation of people during the relaxation period may result in an increased number of contacts between the healthy and the infected individuals resulting in the progression of the infection transmission from sub-exponential to exponential.

Currently, several Asian countries are progressing from stage-2 to stage-3 novSARS-CoV-2 epidemic and planning for the implementation of lockdown to mitigate the epidemic. In view of the economic consequences, several countries may prefer shorter lockdown interventions. These countries must objectively implement the lockdown irrespective of the industrial preference and utilize the lockdown period for contact tracing and to maintain the exposure at a low level to achieve the desirable effects.

Since computational techniques and computer-based algorithms such as CFD analysis, information processing tools such as the principal component analysis, routing algorithms [25-27], are highly efficient in finding solutions to real world problems, such algorithms need to be employed for the analyzing the various clinical as well as epidemiological aspects of COVID-19.

REFERENCES

[1] WHO,https://www.who.int/emergencies/diseases/novel-coronavirus-2019/events-as-they-happen.

[2] WHO Coronavirus disease (COVID-19) Situation Report– 108 https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200507covid-19-sitrep-108.pdf?sfvrsn=44cc8ed8_2

[3] **Guest, J. L., del Rio, C.,Sanchez, T..** (2020). The Three Steps Needed to End the COVID-19 Pandemic: Bold Public Health Leadership, Rapid Innovations, and Courageous Political Will. *JMIR Public health and surveillance*, vol. 6, no. 2, e19043.

[4] **Huppert, A., Katriel, G.** (2013): Mathematical modelling and prediction in infectious disease epidemiology. *Clinical microbiology and infection*, vol.19, no.11, pp. 999-1005.

[5] **Xia, J. L.. Yao, C. & Zhang, G. K.** (2003): Analysis of piecewise compartmental modeling for epidemic of SARS in Guangdon. *Chin. J. Health Stat*, vol. 20, pp. 162-163.

[6] **Lin, G., Jia, X., Ouyang, Q.** (2003): Predict SARS infection with the small world network model. *Beijing da xuexuebao. Yi xue ban= Journal of Peking University. Health sciences*, vol.35, pp. 66-69.

[7] **Wang, J. F., Meng, B., Zheng, X.Y., Liu, J.Y., Han, W.G.,et al.**(2005): Analysis on the multi-distribution and the major influencing factors on severe acute respiratory syndrome in Beijing.*Zhonghua Liu Xing Bing Xue Za Zhi.*vol. 26, no.3, pp.164-8.

[8] **Li, Z., Chen, X., Teng, H., Xiu, Z.** (2004): Infectious kinetics of SARS epidemic. *Progress In Biochemistry And Biophysics*, vol. 31, no. 2, pp.167-171.

[9] **Liang, K.** (2020). Mathematical model of infection kinetics and its analysis for COVID-19, SARS and MERS. *Infection, Genetics and Evolution*, 104306.

[10] **Ivorra, B., Ferrández, M. R., Vela-Pérez, M., & Ramos, A. M.** (2020): Mathematical modeling of the spread of the coronavirus disease 2019 (COVID-19) taking into account the undetected infections. The case of China. *Communications in Nonlinear Science and Numerical Simulation*, 105303

[11] **Victor, A.** (2020): Mathematical Predictions for COVID-19 As a Global Pandemic. *Available at SSRN*.http://dx.doi.org/10.2139/ssrn.3555879

[12] **Jiang, X., Coffee, M., Bari, A., Wang, J., Jiang, X., et al.** (2020). Towards an Artificial Intelligence Framework for Data-Driven Prediction of Coronavirus Clinical Severity. *CMC-Computers, Materials & Continua*, vol. 63, no. 1, pp. 537–551.

[13] **Vaishya, R., Javaid, M,; Khan, I. H., & Haleem, A.** (2020). Artificial Intelligence (AI) applications for COVID-19 pandemic. *Diabetes & Metabolic Syndrome: Clinical Research & Reviews*.

[14] **Chowell, G., Sattenspiel, L., Bansal, S., Viboud, C.** (2016): Mathematical models to characterize early epidemic growth: A review. *Physics of Life Reviews*, vol. 18, pp. 66-97.

[15] **Gu, K., Wang, L. Y., Yin, B.** (2019): Social community detection and message propagation scheme based on personal willingness in social network. *Soft Computing*, vol. 23, no. 15, pp. 6267-6285.

[16] **Ambikapathy, B. & Kirshnamurthy, K.**(2020): Mathematical modelling to assess the Impact of lockdown on COVID-19 Transmission in India: Model Development and Validation. *JMIR Public health and Surveillance*, vol. 6, no. 2,e19368.

[17] **The Lancet**, India under COVID-19 lockdown, Volume 395, Issue 10233, 25 April–1 May 2020, Page 1315, https://doi.org/10.1016/S0140-6736(20)30938-7

[18] **Kamalanand, K., & Jawahar, P.** (2018): Mathematical Modelling of Systems and Analysis. *PHI Learning Pvt. Ltd.*

[19] **Phan, L.T., Nguyen, T.V., Luong, Q.C., Nguyen, T.V., Nguyen, H.T., et al.** (2020): Importation and human-to-human transmission of a novel coronavirus in Vietnam. *New England Journal of Medicine*. vol.382, no.9, pp.872-4.

[20] **Chan, J.F-W., Yuan, S., Kok, K-H., To K. K-W., et al** (2020): A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: a study of a family cluster. *The Lancet*,vol. 395, no.10223, pp. 514-23.

[21] **Ministry of Railways**, India, https://indiarailinfo.com/atlas.

[22] **To, K.K., Tsang, O.T., Yip, C.C., Chan, K.H., Wu, T.C., et al.** (2020): Consistent detection of 2019 novel coronavirus in saliva. *Clinical Infectious Diseases,* ciaa149, https://doi.org/10.1093/cid/ciaa149.

[23] **Burke, R. M.**(2020): Active monitoring of persons exposed to patients with confirmed COVID-19United States, January–February 2020.

ISBN 979-8-3503-5094-4

MMWR. Morbidity and Mortality Weekly Report, vol. 69, no.9, pp. 245-246.

[24] **Zhao, S., Zhuang, Z., Ran, J., Lin. J., Yang, G., et al.** (2020): The association between domestic train transportation and novel coronavirus (2019-nCoV) outbreak in China from 2019 to 2020: A data-driven correlational report. *Travel medicine and infectious disease*. vol.33, 101568.

[25] **He, S. M., Xie, K., Xie, K. X., Xu, C., Wang, J.** (2019): Interference-aware multisource transmission in multiradio and multichannel wireless network. *IEEE Systems Journal*, vol. 13, no. 3, pp. 2507-2518.

[26] **Chen, Y. T., Xia, R. L., Wang, Z., Zhang, J. M., Yang, K. et al.** (2019): The visual saliency detection algorithm research based on hierarchical principle component analysis method. *Multimedia Tools and Applications*, https://doi.org/10.1007/s11042-019-07756-1.

[27] **Zhang, D. Y., Liang, Z. S., Yang, G. B., Li, Q. G., Li, L. D. et al.** (2018): A robust forgery detection algorithm for object removal by exemplar-based image inpainting. Multimedia Tools and Applications, vol. 77, no. 10, pp. 11823-11842.

ISBN 979-8-3503-5094-4

Patient Treatment Classification using Machine Learning

Dr.M.C.Jobin Christ[a*], Dr.P.Mohamed Sajid[b], V.Mythily[c] V.A.Sairam[d], A.Sanjo Christ[e]
a.Professor, Department of Biomedical Engineering, Rajalakshmi Engineering College, Chennai
b.Assistant professor, Department of ECE, C.Abdul Hakeem College of Engineering and Technology, Vellore
c. Assistant professor, Department of Biomedical Engineering, Jerusalem College of Engineering, Chennai
d. PG Student, Department of AIML, Symbiosis Institute of Technology, Pune
e.UG Student, Department of Computer Science and Engineering, Rajalakshmi Institute of Technology, Chennai
** Corresponding Author Email Address : jobinchrist.mc@rajalakshmi.edu.in*

Abstract : Clinicians and specialists often face the challenge of determining whether a patient requires inpatient or outpatient care by analyzing the results of laboratory tests. But it is a time consuming process as well as demanding, as it needs significant effort from doctors to make accurate decisions. Moreover, there is a high risk of making incorrect judgments, which can jeopardize the patient's life. This research article intends to utilize a range of Machine Learning (ML) algorithms in order to categorize patient care as either outpatient or inpatient, with the goal of alleviating the workload on doctors and enhancing the quality of services delivered. This study engaged machine learning to construct six models like Logistic Regression, Random Forest, Support Vector Machine, K-Nearest Neighbors (KNN), Decision Tree and Naive Bay's. These models utilized patients' conditions and laboratory test results to predict whether they should be categorized as inpatient or outpatient. The selection of the best model was based accuracy. For this purpose, a dataset from Indonesia's private hospital was used. The dataset consisted of patients' laboratory test results, extracted from their electronic health records (EHR). The models were built using this dataset and subsequently tested. The findings indicated that among the different algorithms tested, the Random Forest algorithm exhibited the good accuracy rate of 92.44%. These findings highlight the possibility of utilizing machine learning to enhance the accuracy and efficiency of patient care classification, thereby reducing the occurrence of human errors that could pose risks to patients' lives and financial burdens to their wallets.

Keywords - K-Nearest Neighbors (KNN), Decision Tree, Support Vector Machine, Machine Learning, Electronic Health Record, Random Forest, Classification, Naive Bay's, Logistic Regression.

I. INTRODUCTION

The implementation of health information systems (HIS) utilizing computer technology and automation has replaced traditional paper health records in the healthcare sector. This system is an automated solution that combines hardware and software components to gather, document, and exchange health-related data concerning patients. The primary goal of this initiative is to improve the quality of services offered to patients. This shift has had a profound influence on the healthcare sector, bringing advantages to both patients and healthcare providers. It has revolutionized the way patients are cared for by facilitating treatment advocating for practices based on evidence prioritizing patient safety and actively involving patients, in their healthcare journey. Moreover, it has enhanced communication, boosted efficiency, and lowered expenses. Electronic Medical Records (EMR) and EHR are instances of this technology that are frequently used interchangeably. Within a particular healthcare organization, paper-based records were substituted with EMR. On the other hand, ehr was introduced to address the issue of limited accessibility, allowing multiple healthcare organizations to view and edit records while adhering to global and national standards. [15]

The automation of healthcare processes has generated a vast amount of data, often referred to as big data. The vast amount of data available has created an opportunity for the incorporation of machine learning, health analytics and leveraging health data to extract insights and make well-informed decisions. Machine learning algorithms can be utilized to develop models that achieve high accuracy while reducing time and costs, potentially substituting certain functions of healthcare providers. Machine learning has become an essential tool in improving the effectiveness of the healthcare industry. It finds application in various subdomains of healthcare, enabling the delivery of improved services at lower costs. Machine learning models can serve as predictive tools, aiding in the early detection of illnesses, treatment planning, and providing doctors and physicians with valuable insights to intervene at an earlier stage. Furthermore, machine learning enables the anticipation of population health risks by detecting patterns and identifying indicators associated with high risks. [6]

ISBN 979-8-3503-5094-4

A crucial responsibility of doctors and specialists is to determine whether a patient requires inpatient or outpatient care. Within hospitals, treatments and surgeries are classified into these two primary categories. Inpatient care necessitates overnight hospitalization and is essential for critical cases such as intricate surgeries, severe illnesses, or medical conditions that demand ongoing monitoring. Patients receiving inpatient care remain under the supervision of nurses or doctors and may require complex treatments like delivering a baby or managing severe injuries. Outpatient care, alternatively referred to as ambulatory care, does not entail the need for hospitalization. Examples of outpatient care include annual exams with primary doctors and consultations with specialists like neurologists. In emergency cases, outpatient care may be considered if the patient is discharged on the same day they arrived. Typically, outpatient care encompasses routine procedures like X-rays, minor surgeries, MRI imaging, colonoscopies, laboratory tests, regular physical examinations, and mammograms.[9]

The classification of patient care as either inpatient or outpatient is highly significant as it affects the patient's healthcare progress and subsequent billing procedures. The length of hospital stay significantly affects the final bill. However, this decision-making process can be prone to imprecision and delays due to human errors. Additionally, doctors may face time constraints or unavailability, leading to critical decisions being made under suboptimal circumstances, potentially resulting in adverse outcomes. Therefore, it is imperative to explore techniques that can assist both doctors and patients in making informed decisions efficiently. Machine learning has become an integral part of the healthcare industry, playing a crucial role in improving patient care, diagnosis, treatment planning, and overall healthcare management. With the advent of Health Information Systems (HIS) and the abundance of health data, machine learning algorithms have the capacity to bring about a paradigm shift in healthcare practices.[13]

Predictive modeling stands as a prominent application of machine learning in the healthcare domain. Machine learning algorithms can analyze a vast dataset and detect patterns and correlations, allowing them to predict the likelihood of disease occurrence, progression, or treatment outcomes. This predictive capacity can aid medical professionals in formulating well-founded decisions and crafting tailored treatment plans for individual patients. For instance, machine learning models can contribute to the identification of individuals with a heightened risk of specific diseases, facilitating early intervention and the implementation of preventive measures. Machine learning also holds great importance in the examination and understanding of medical imagery. Medical imaging methods such as X-rays, MRIs, and CT scans produce extensive volumes of visual data. Machine learning algorithms can be trained to analyze these images and accurately detect abnormalities or diagnose specific conditions. This can aid radiologists and other medical professionals in achieving quicker and more precise diagnoses, resulting in prompt and efficient treatment. Machine learning also shows abilities, in the field of natural language processing (NLP) and text analysis, which emphasizes its strength, in another area. Electronic health records (EHR) and notes hold patient data that can be used for research, decision making support and managing population health. Through machine learning algorithms relevant information can be extracted from these texts allowing for data driven insights and automated analysis. NLP techniques can be harnessed to create chatbots or virtual assistants capable of engaging with patients, responding to their inquiries, and offering rudimentary medical guidance. Furthermore, machine learning algorithms can contribute to drug discovery and development. By examining molecular structures, genomic data, and previous drug response information, machine learning models have the capability to identify prospective drug candidates, predict their efficacy, and optimize drug dosages. Such an approach has the capacityl to greatly accelerate the drug discovery process, resulting in the creation of more efficient and individualized treatments. Machine learning algorithms can also assist in healthcare resource allocation and optimization. By analyzing historical data and patient characteristics, machine learning models can predict patient admissions, emergency department utilization, and bed occupancy rates. This information can help healthcare administrators make informed decisions regarding resource allocation, staffing, and capacity planning, leading to more efficient healthcare delivery and reduced costs. Nevertheless, it is crucial to understand that although ML has displayed immense potential in healthcare, it is not intended to substitute healthcare providers. Instead, it should be seen as a tool to augment their capabilities and support clinical decision-making. Ethical considerations, privacy concerns, and regulatory compliance are also critical aspects to address when implementing machine learning in healthcare settings. As a result, ML has the capacity to transform healthcare by improving patient care, diagnostic accuracy, treatment planning, and health management. By utilizing the power of machine

ISBN 979-8-3503-5094-4

learning algorithms and rich health data, healthcare providers can make more informed decisions, deliver customized treatment, and improve overall patient outcomes. The integration of machine learning with Health Information Systems has opened up new possibilities for efficient and data-driven healthcare delivery. Nonetheless, it is imperative to implement, validate, and continuously monitor machine learning in healthcare with care to ensure its ethical and responsible usage.[7]

II. EARLIER WORKS DONE

Machine learning has become increasingly popular in the medical sector due to its effectiveness in improving healthcare services. In a study conducted by Bharadwaj et al. [14] exploring the application of ML in medical sector, it was concluded that incorportaion of tools such as ML and big data can improve medical outcomes at a lower cost, which has significant relevance in the rapidly expanding healthcare industry. Wiens and Shenoy [21] provided a definition of machine learning and explored its transformative impact on healthcare epidemiology. In their work, they demonstrated specific examples and practical applications, such as identifying reservoirs for Zoonotic diseases and predicting the risk of Nosocomial Clostridium Difficile Infection. They also addressed the factors and obstacles of utilizing machine learning, such as selecting appropriate targets and controlling lost health data. Beaulieu-Jones and colleagues [3] conducted an analysis of submissions to a widely recognized machine learning conference focused on health to evaluate the present status of research in this field. The results indicated that there is a strong focus on healthcare because of the prescence of easily accessible and thoroughly annotated information. The authors suggested that greater involvement of clinicians is needed to develop translational applications. In their study, Heart et. al [8] provided definitions for EMR and EHR and reached the conclusion that integrating medical information with health records offers a holistic perspective that aligns with the principles of patient-centered medical care. The fusion of these records has the potential to bring about noteworthy advancements in personalized healthcare and enhance the procedure of formulating well-founded choices regarding public health. The EHR serves as a valuable reservoir of data that can be utilized for analysis purposes and constructing predictive models. In their research, Wu, et.al devised a model that leverages EHR information to predict the onset of heart failure [22]. In their investigation, the researchers assessed the efficacy of SVM, logistic regression, and boosting techniques in predicting heart failure. The outcomes demonstrated that these predictive models exhibited the capability to anticipate the onset of heart failure up to six months prior to clinical diagnosis. In their study, K and colleagues [11] identified a range of diseases that can be tackled and predicted utilizing machine learning. These diseases encompass heart disease, thyroid disease, diabetes and breast cancer. The findings of their study revealed that Support Vector Machine (SVM) attained an accuracy of 96.40% in diagnosing breast cancer. Naive Bayes, on the other hand, demonstrated an 86% accuracy in diagnosing heart disease, while CART achieved a 79% accuracy in identifying diabetes. Alloghan and et. al. [2] employed machine learning methods to identify patterns of readmission among patients with diabetes. They applied various models, including linear discriminant analysis, KNN, J48, SVM, Random Forest and Naive Bayes, to analyze a dataset specific to diabetes. The research revealed patterns suggesting that female patients with diabetes, individuals of Caucasian ethnicity, outpatients, those receiving fewer medications, or undergoing less extensive laboratory procedures are more prone to readmissions. In their study, Zheng et. al. [23] utilized ML techniques and conducted attribute construction on a dataset of EHR to unveil genotype-phenotype correlations associated with Type II Diabetes Mellitus (T2DM). They employed various techniques, including SVM, Naive Bayes, Random Forest, KNN, Decision Tree, and Logistic Regression in their analysis. The framework demonstrated outstanding identification performance, with an impressive average accuracy of approximately 0.98. This surpasses the accuracy achieved by the state-of-the-art algorithm, which was at 0.71. Wang and colleagues [20] constructed a deep learning algorithm aimed at predicting the risk associated with advanced colorectal cancer in adults of Taiwan. The study took into account the substantial impact of this disease on mortality rates in Taiwan. The model achieved impressive performance metrics, with an accuracy of 0.92, a sensitivity of 0.84, and a specificity of 0.87. In their study, Reddy and Delen [16] utilized a deep learning approach entailing Recurrent Neural Networks (RNN) and Long Short-Term Memory (LSTM) to forecast hospital readmission rates for patients with lupus. Hospital readmission serves as a vital metric for assessing the quality of hospital services and evaluating overall performance.

ISBN 979-8-3503-5094-4

III. MACHINE LEARNING ALGORITHMS

A. *Support Vector Machine*

In their work, Vapnik et. al. [19] introduced a robust approach called SVM, which is based on Statistical Learning Theory (SLT). SVM possesses distinct advantages when it comes to addressing challenges like small sample sizes, nonlinear relationships and recognition of high-dimensional patterns. These notable advantages have sparked substantial interest within the realm of machine learning. Support vector machine (SVM) aims to reduce training errors based on empirical data, while maximizing the distance between data points and the hyperplane used for classification. This approach introduces a high level of generalization error. SVMs have demonstrated the ability to achieve robust generalization in high-dimensional spaces, even with limited training samples. This surpasses the conventional empirical risk minimization principles typically employed in neural networks. As a result, SVMs have found successful applications in many fields, including face detection, verification and recognition, handwriting and number identification, object identification and detection, speaker verification and identification, text recognition and classification, prediction, search images and more. [4].

B. *Decision Tree*

A supervised learning technique that can be utilized for solving regression and classification problems is decision tree. However, its use is mainly in solving classification problems. A decision tree is characterized by a distinct tree-like structure, where internal nodes denote features in the database, branches symbolize decision rules, and leaf nodes indicate outcomes or class labels. The decisions or tests in a Decision Tree are based on the features present in the datasetThe Decision Tree serves as a visual representation that encompasses all potential solutions or outcomes for a specific problem or decision, taking into account the provided conditions. The term "decision tree" comes from the similarity with real nodes, because it starts at the root node and forms a tree-like structure. In a decision tree, there are two kind of nodes; decision nodes and leaf nodes. Decision node plays an important role in decision making and has multiple branches, while a leaf node is the final output without additional branches [17].

Inside the decision tree, internal nodes partition the samples into various sub-samples based on specific probability indices such as entropy or Gini. The objective of this division process is to optimize the classification outcome. In the end, each leaf node or terminal node is allocated to the class that exhibits the most favarable response values [1].

C. *Random Forest*

The Random Forest (RF) algorithm, devised by Leo Breiman and Adele Cutler, is a comprehensive learning technique that comprises multiple smaller submodels. The final outcome is generated by combining the outputs of each separate sub-model. Random Forest (RF) is a popular ML technique renowned for its applications in regression, classification, and various other learning tasks. Random Forest is based on the bagging algorithm, which involves partitioning the information from the actual dataset into multiple clusters. Every cluster is then trained independently, leading to the creation of individual decision tree models. The decision data generated by all the submodels is subsequently merged and examined to construct the final Random Forest framework. The RF algorithm makes predictions using a voting mechanism, where the classification that receives the maximum number of votes is regarded as the ultimate output of the RF method. [12]

D. *KNN*

The k-Nearest Neighbors technique (kNN) is a type of non-parametric categorization algorithm that is simple yet highly efficient in diverse scenarios. To classify a data record x, the kNN of x are determined, creating a neighborhood surrounding x. The classification for x is commonly determined by conducting a voting process within the data records among the neighborhood for identifying majority, whether or not incorporating distance associated weighting. Nevertheless, when employing the kNN method, selecting a proper value for k is critical since the effectiveness of the categorization heavily depends on this decision. In a way, the performance of the kNN algorithm is influenced by the chosen value of k. There exist several approaches for determining the value of k, but a straightforward strategy involves running the technique multiple times with various k values and selecting the one that produces the most optimal contribution. [5]

E. *Naive Bayes*

Naive Bayes classifier greatly simplifies the training process by assuming that the feature classes are independent. Bayesian classifiers determine the most likely class in a given sample based on its probability vector. Naive Bayes has demonstrated its effectiveness in many practical applications, including medical diagnosis, text classification, and system performance management. Naive Bayes is

ISBN 979-8-3503-5094-4

recognized as a fast and straightforward machine learning algorithm for class prediction in databases. Naive Bayes can be used for bivariate and multiclass classification. Naive Bayes shows good performance in multiclass estimation compared to other algorithms. It is widely accepted as the preferred choice for solving text classification problems. [10]

F. Logistic Regression

In the field of machine learning, logistic regression is a widely utilized supervised learning method. The primary objective of logistic regression is to forecast a categorical dependent variable by means of a collection of independent variables. Logistic regression models attempt to forecast outcomes from categorical dependent variables. Finally, logistic regression's output should be a categorical or discrete value, usually Yes or No, 0 or 1, true or false, and so on. However, in place of giving an exact value of 0 or 1, logistic regression gives a probability value in between zero and one. [18]

IV. METHODOLOGY

A. Data preprocessing

Data preprocessing involves the preparation of raw data to render it appropriate for utilization in a machine learning model. This initial and essential step is integral to the development of a machine learning model. When initiating a machine learning project, it is typical to come across data that is unstructured or lacks proper formatting. Consequently, it is necessary to clean and organize the data in a structured manner before performing any operations. This is where data preprocessing comes into play. It comprises the subsequent stages;

- Acquiring the dataset
- Importing libraries
- Finding the missing data
- Encoding the categorical data
- Upsampling
- Normalization
- Dividing the dataset into training and test sets

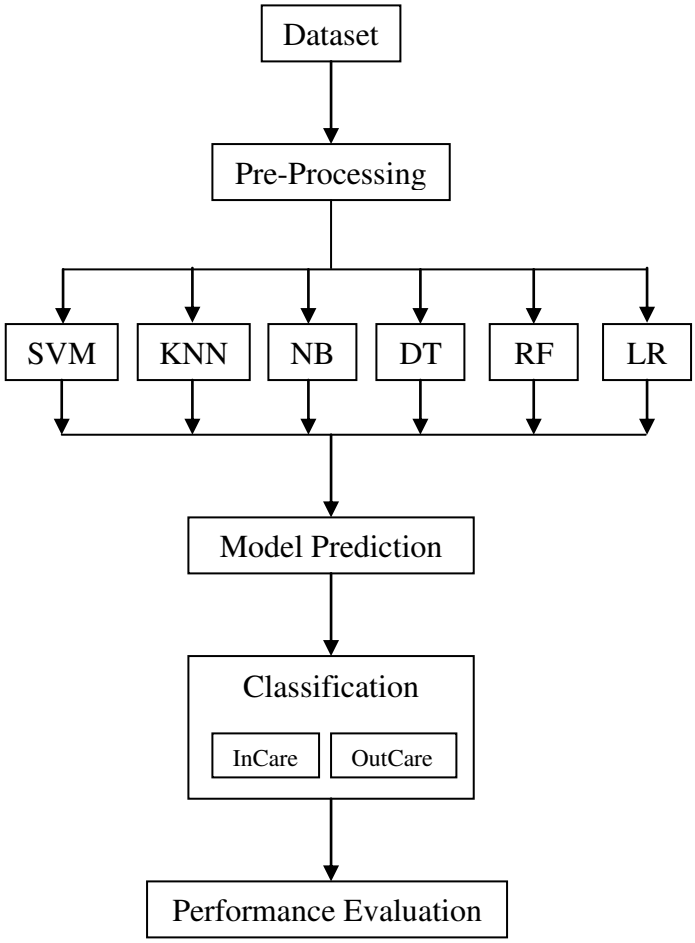


Figure 1. Methodology of Proposed ML Algorithms

B. Acquiring the dataset

To develop a ML algorithm, the primary prerequisite is a dataset since machine learning heavily relies on data. The dataset refers to the collected data in a proper format specifically tailored for addressing a particular problem. Typically, we store the dataset in a CSV file, although occasionally we might also utilize HTML or xlsx files. The dataset we will be discussing here is called "Patient Treatment," and it is sourced from the Kaggle website. This dataset, derived from a private hospital in Indonesia, was generated and made available on November 19, 2020. It encompasses 11 attributes, as outlined in Table 1. The dataset contains a total of 3,309 records, including 1,992 records for outpatients and 1,317 records for inpatients. The objective is to

ISBN 979-8-3503-5094-4

ascertain whether a patient's care is categorized as outpatient care or inpatient care.

Table 1. 11 Attributes of the Dataset

No.	*Attribute*	*Ranges*
1	Haematocrit	13.7-69
2	Haemoglobins	3.8-18.9
3	Erythrocyte	1.4-7.86
4	Leucocyte	1.1-76.6
5	Thrombocyte	10-1121
6	MCH	14.9-40.8
7	MCHC	26-38.4
8	MCV	54-116
9	Sex	M F
10	Age	1-99
11	Source	1=incare patient, 0 = outcare patient

C. *Importing libraries*

To conduct data preprocessing using Python, it is necessary to import specific pre-existing Python libraries that fulfill particular tasks. Three key libraries used in data preprocessing are as follows:

1. Numpy: The Numpy library is employed to incorporate various mathematical operations into the code. It serves as a fundamental package for scientific calculations in Python, enabling the manipulation of large multidimensional arrays and matrices.
2. Pandas: Pandas is a widely used Python library for importing and managing datasets. It is an open-source library for data manipulation and analysis, providing extensive functionality for handling structured data.
3. Matplotlib: Matplotlib is a Python 2D plotting library, often used in conjunction with its sub-library pyplot. This library facilitates the creation of different types of charts and graphs in Python, enhancing visualizations in code.

Additionally, when implementing machine learning algorithms, it is common to import the following libraries for specific algorithms

- SVC for Support Vector Machine algorithm
- Logistic Regression for Logistic Regression algorithm
- K Neighbors Classifier for K Nearest Neighbor Algorithm
- Gaussian NB for Naive Bayes classifier
- Decision Tree Classifier for Decision Tree Algorithm
- Random Forest Classifier for Random Forest Algorithm

D. *Finding the Missing Data*

After completing the initial steps of data preprocessing, the subsequent task is to address any missing data within the dataset. The presence of missing values in the dataset can significantly influence the outcome of the machine learning model. Hence, it is imperative to appropriately address any missing values present in the dataset. However, in the dataset we have employed, there are no occurrences of missing data.

E. *Encoding the Categorical Data*

Categorical data refers to data that is organized into specific categories or groups. Within our dataset, two variables fall under the category of categorical variables, namely Sex and Source. Machine learning models primarily operate on numerical data and mathematical calculations. However, if our dataset contains categorical variables, it can pose challenges when constructing the model. Therefore, it becomes essential to encode these categorical variables into numerical representations. This process enables the machine learning model to effectively utilize and analyze the categorical data.

F. *Upsampling*

Within our dataset, there is an imbalance between the two classes, with 2,628 inpatient records and 1,784 outpatient records. To address this issue, we can employ a technique called upsampling or oversampling. Upsampling involves generating

ISBN 979-8-3503-5094-4

artificial or duplicate data points for the minority class in order to balance the class labels. There are various oversampling techniques available, but in this case, we will utilize random oversampling. Random oversampling involves duplicating examples from the minority class within the training dataset Nevertheless, it is essential to acknowledge that random oversampling may result in over fitting for certain models. In order to balance the classes, we will create 3,000 inpatient records and 3,000 outpatient records through this random oversampling technique.

G. Normalization

Normalization is a widely used technique in data preparation that aims to standardize the values of numerical columns in the data set to standard dimensions. Although normalization is not necessary for all databases in machine learning, it is most useful when the attributes in the database have different ranges. The aim of regularization is to improve the performance and reliability of machine learning models.

Mathematically, normalization can be calculated using the subsequent formula:

$$Xn = (X - Xmn) / (Xmx - Xmn)$$

Where: Xn represents the normalized value, Xmx is the feature's maximum value, Xmn is the feature's minimum value.

H. Dividing the dataset into training and test sets

An important step in the data processing phase in machine learning involves dividing the database into separate training and test sets. This part plays a significant role in refining the performance of our machine learning model. When we train our ML algorithm, we need to examine their output on a different database than the one used for training. This is because if we test the model on an entirely new dataset, it can face difficulties in understanding the underlying correlations between the data points. Even if our model is trained well and achieves high accuracy during training, its performance may decrease when presented with a new dataset. Hence, the objective is to develop a machine learning model that demonstrates favorable performance not only on the training set but also on the test dataset. This ensures that the model can generalize effectively to unseen data and provides reliable predictions. In this specific scenario, the dataset is partitioned in a manner where 85% of the data is assigned as the training set, and the rest 15% is set aside as the test set.

I. Implementation of Classifiers

The "Patient Treatment" dataset was employed to categorize patients into inpatient and outpatient classifications utilizing six distinct classification models. The implementation was done in a "Colab" notebook using the programming language; Python. Out of the six models, four (DT, KNN, RF, SVM) are eager learners, while the remaining two (NB, LR) are lazy learners. The SVM classifier utilized a Radial Basis Function (RBF) kernel for its operations. The gamma value was chosen as 128 after conducting several trials, and the C parameter was selected as 8. In the K Nearest Neighbor algorithm, the optimal value for k was estimated at 1 after trying different values ranging from 1 to 10. This information is indicated in the figure 2.

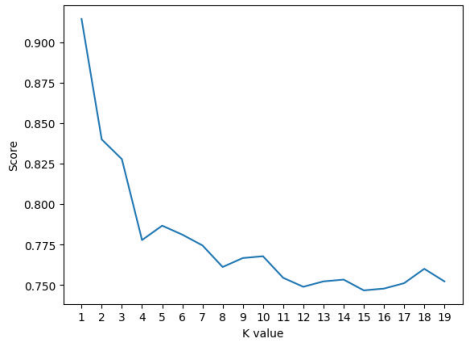


Figure 2. Accuracy Vs K value

In the Random Forest algorithm, the number of trees is determined to be 2000 through a trial and error methodology. Table 2 presents the calculated accuracies for all the algorithms. Comparison of accuracies of different algorithms is given in figure 3.

Table 2. Accuracies of Various Algorithms

S.No	Name of the Algorithm	Accuracy
1	Support Vector Machine	91.56%
2	Decision Tree Algorithm	91.78%
3	Random Forest Algorithm	92.44%
4	K-Nearest Neighbors (KNN) Algorithm	91.44%
5	Naive Bay's Classifier	69.00%
6	Logistic Regression	70.33%

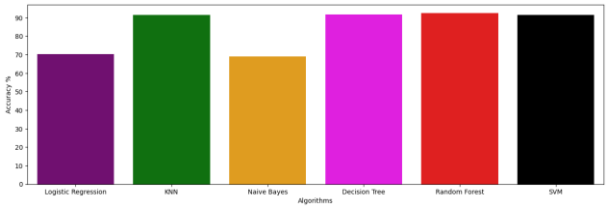


Figure 3. Comparison of Performance of Various Algorithms

ISBN 979-8-3503-5094-4

V. Results and Discussions

Six different classifiers, four of which are eager (DT, KNN, SVM, RF) and two are lazy (NB, LR), were applied to the dataset. The dataset is split into training sets and testing sets using the cross-validation method, specifically with a value of k=1. After careful analysis, it is evident that the RF algorithm exhibited the best performance with an accuracy of 92.44%. In comparison, other algorithms such as SVM, Logistic Regression, KNN, Naïve Bayes, and Decision Tree achieved accuracies of 91.56%, 70.33%, 91.44%, 69%, and 91.78% respectively. In the medical field, categorizing patients as inpatients or outpatients based on test results requires significant effort and time from medical staff. However, this process is prone to human errors, which can potentially endanger patients' lives. ML has the ability to greatly impact medical sector by reducing the efforts and errors associated with decision-making, ultimately improving patient safety.

VI. Conclusion

The task of categorizing patients' required care as outpatient or inpatient based on test lab reports is a labor-intensive and time-consuming process that is susceptible to human errors. These errors can potentially endanger patients' lives. In this study, machine learning was incorporated as part of a hospitalization system. The goal was to develop a system capable of accurately classifying patient care and taking appropriate actions based on the classifier's decision. By implementing such a system, doctors and specialists can benefit from reduced decision-making time and efforts, as well as a decreased likelihood of making incorrect decisions that could jeopardize patient safety. Additionally, the system can provide insights into the length of stay, which directly impacts the patient's final bill. Overall, this system aims to improve services and enhance lives. In future work, the study aims to further enhance the system by improving the performance of the six evaluation measures. This can be accomplished by leveraging alternative ML and deep learning algorithms. Furthermore, endeavors will be undertaken to collect additional data, aiming to augment the capabilities of the model and expand its scope of application.

References

1]. Abohamr SI, Abazid RM, Aldossari MA, Amer HA, Badhawi OS, Aljunaidi OM, Alzarzour SH, Saadeddin HM, Bhat F, Elsheikhal E. (2020). Clinical characteristics and in-hospital mortality of COVID-19 adult patients in Saudi Arabia. Saudi Medical Journal, 41(11):1277

2]. Alloghani M, Aljaaf A, Hussain A, Baker T, Mustafina J, Al-Jumeily D, and Khalaf M (2019). Implementation of machine learning algorithms to create diabetic patient re-admission profiles. BMC Medical Informatics and Decision Making. 19, 253.

3]. Beaulieu-Jones B, Finlayson SG, Chivers C, Chen I, McDermott M, Kandola J, & Naumann T (2019). Trends and focus of machine learning applications for health research. JAMA network open, 2(10).

4]. Byun H, Lee SW. (2002). Applications of Support Vector Machines for Pattern Recognition: A Survey. Pattern Recognition with Support Vector Machines. SVM 2002. Lecture Notes in Computer Science, Vol. 2388. Springer, Berlin, Heidelberg.

5]. Gongde Guo, Hui Wang, David A. Bell, Yaxin Bi (2004). KNN Model-Based Approach in Classification,

6]. Grazia Dicuonzo, Graziana Galeone, Matilda Shini, and Antonella Massari, (2022).Towards the Use of Big Data in Healthcare: A Literature Review, Healthcare (Basel). 10(7): 1232.

7]. Hafsa Habehh and Suril Gohel, (2021). Machine Learning in Healthcare, Curr Genomics. 22(4): 291–300.

8]. Heart T, Ben-Assuli O, and Shabtai I. (2017). A review of PHR, EMR and EHR integration: A more personalized healthcare and public health policy. Health Policy and Technology, 6(1), 20-25.

9]. https://www.sgu.edu/blog/medical/inpatient-versus-outpatient/

10]. Irina Rish, (2001). An Empirical Study of the Naïve Bayes Classifier.

11]. K. Shailaja, B. Seetharamulu and M. A. Jabbar, "Machine Learning in Healthcare: A Review," 2018 Second International Conference on Electronics, Communication and Aerospace Technology (ICECA), Coimbatore, pp. 910-914.

12]. Qingxiang Xu, Jiesen Yin, "Application of Random Forest Algorithm in Physical Education", *Scientific Programming*, vol. 2021, Article ID 1996904, 10 pages.

13]. Quy Nguyen, Michael Wybrow, Frada Burstein, David Taylor, and Joanne Enticott, Understanding the impacts of health information systems on patient flow

ISBN 979-8-3503-5094-4

management: A systematic review across several decades of research, PLoS One. 2022; 17(9): e0274493.
14]. R. Bhardwaj, A. R. Nambiar and D. Dutta, "A Study of Machine Learning in Healthcare," (2017) IEEE 41st Annual Computer Software and Applications Conference (COMPSAC),Turin, pp. 236-241.
15]. R. S. Evans, (2016). Electronic Health Records: Then, Now, and in the Future, Yearb Med Inform. (Suppl 1): S48–S61.
16]. Reddy, B. K., & Delen, D. (2018). Predicting hospital readmission for lupus patients: An RNN-LSTM-based deep-learning methodology. Computers in biology and medicine, 101, 199-209
17]. Swain, P. H., & Hauska, H. (1977). The decision tree classifier: Design and potential. IEEE Transactions on Geoscience Electronics, 15(3), 142-14
18]. Tammy Jiang, Jaimie L. Gradus, and Anthony J. Rosellini, (2020). Supervised machine learning: A brief primer, Behav Ther. 51(5): 675–687.
19]. Vapnik V N. (1995) The nature of statistical learning [J]. Theory.
20]. Wang, Y. H., Nguyen, P. A., Islam, M. M., Li, Y. C., & Yang, H. C. (2019). Development of Deep Learning Algorithm for Detection of Colorectal Cancer in EHR Data. In MedInfo (pp. 438-441).
21]. Wiens, J.,and Shenoy, E. S. (2017). Machine Learning for Healthcare: On the Verge of a Major Shift in Healthcare Epidemiology. Clinical Infectious Diseases, 66(1), 149–153. doi:10.1093/cid/cix731h.
22]. Wu, J., Roy, J., & Stewart, W. F. (2010). Prediction modeling using EHR data: challenges, strategies, and a comparison of machine learning approaches. Medical care, 48(6 Suppl), S106–S113.
23]. Zheng, T., Xie, W., Xu, L., He, X., Zhang, Y., You, M. & Chen, Y. (2017). A machine learning-based framework to identify type 2 diabetes through electronic health records. International journal of medical informatics, 97, 120-127.

ISBN 979-8-3503-5094-4

Tenth International Conference on Biosignals, Images and Instrumentation, SSNCE, Chennai, March 20 – 22, 2024

TRACK 5

Cost Effective Blood Collection Monitor for Emergency Patients

Sofia Bobby J
Department of Biomedical Engineering
Jerusalem College of Engineering
Chennai, India
bmehod@jerusalemengg.ac.in

Annapoorani CL
Department of Biomedical Engineering
Jerusalem College of Engineering
Chennai, India
annapooranic@jerusalemengg.ac.in

Nivetha S
Department of Biomedical Engineering
Jerusalem College of Engineering
Chennai, India
nivethas@jerusalemengg.ac.in

Lavanya S
Department of Biomedical Engineering
Jerusalem College of Engineering
Chennai, India
lavanyas.bme@jerusalemengg.ac.in

Kabini D
Department of Biomedical Engineering
Jerusalem College of Engineering
Chennai, India
kabinidbme2020@jerusalemengg.ac.in

Padmapriya D
Department of Biomedical Engineering
Jerusalem College of Engineering
Chennai, India
anvi2427@gmail.com

***Abstract*—*Blood collection monitor is used for accurate collection of blood. This compact device counts the amount of blood and gently rocks the blood to ensure that it is mixed evenly with the anticoagulant, which keeps blood clots from forming when blood is drawn from a donor. Blood collection process happens in different environments, the main challenge to the phlebotomist (one who draws blood for analysis or transfusion) is to make each blood collection process more comfortable and safer without compromising quality. Blood collection monitor comes into role here; it is specially designed to for standardized high quality blood collection with reduced work load of phlebotomist. A blood collection monitor makes ensuring that the right amount of blood is collected while maintaining continual movement to improve component output. So, the correct volume of blood collection and mixing of the blood with anticoagulant and periodically during collection of blood is done by blood collection monitor. In this project a more cost effective and easily portable model of the blood collection monitor will be designed.*

Keywords—blood collection, phlebotomy, load cell, lead acid battery and tray

I. INTRODUCTION

Blood collection monitor is a device which is used to collect the blood using advanced technology for replacing the insufficient convenient methods. The main principle of the project is the mixing of anticoagulant with the blood by providing a rocking moment to blood flow, when the blood flow from donor is reached 350ml or 500ml be the blood flow should be restricted. Weight measuring is done by the force transducer and stopping clamp is activated by comparing signal from themicrocontroller.

Fig.1 Blood collection monitor

EXISTING MODEL

Existing system is a device used to monitor the collection of blood donation to protect against over or under bleeding of the patient. Integrated keyboard along with microprocessor-based control ensure smooth & gentle mixing process, volume accuracy, precise monitoring of time & flow rate. Some of the salient features of the existing model are,

- Accuracy that is best in class thanks to a microprocessor-based technology for precise volume collection control. (deviation <1%).
- Gentle & smooth motorized clamping of blood bag tube prevents damage to the live blood cells.
- Automatic tarring of blood bags.
- Pause facility to temporarily.
- Suspend the donation process.
- Flow rate: 20-180 ml/min.
- Automatic change over on battery mode in case ofpower failure.

PROPOSED SYSTEM

The Fig.1 shows a model of the blood collection monitor. The main uniqueness of this device compared to the other previous versions is that:

- This device is designed in a more cost-effective manner which reduces the production cost of the device.
- This device is connected with IOT interface which collects the data regarding the donor's medical details and stores it in a cloud server.
- This device provides a safe and more comfortable way of blood collection compared to the other versions of the device. This device reduces the work load in the blood collection process of the phlebotomist

OBJECTIVES

To design a cost effective and easily portable blood collection monitor by simplifying the circuit, reducing the weight and making the device more easily accessible to the medical professionals.

ISBN 979-8-3503-5094-4

The main concept behind this device is to reduce the workload in the blood collection process by collecting the required amount of blood without the supervision of the phlebotomist and automatically stopping the flow of blood in the blood bag with the help of a clamp. The main components of the device comprise of a tray, clamp, blood bag, micro controller, 5 pin relay, amplifier, DC motor and a battery.

The potential areas of application of this includes:

- Multi-specialty Hospitals.
- Blood banks.
- Research facilities
- Medical institutes.

This device is designed in a more cost- effective manner which reduces the production cost of the device that helps this device to be easily reachable to the required places. It provides a safe and more comfortable way of blood collection compared to the other versions of the device this helps in maintaining a safer environment of blood collection.

II. MODULES ELUCIDATION

A. Power supply

For use by a computer's internal components, mains AC power is converted by a power supply unit (PSU) into low-voltage regulated DC power. Switched-mode power supply are used in all modern personal computers. While some power supplies automatically adjust to the mains voltage, others offer a manual switch for choosing the input voltage. Three main categories of power supplies exist: switching, linear controlled, and unregulated (sometimes known as brute force). A separate subsection should be devoted to the ripple-regulated power supply circuit, which is the fourth type and a cross between the "brute force" and "switching" designs. In addition to being utilized to power buildings and companies, AC current is also present when radio and audio signals are sent through electrical cables. Batteries that power flashlights and other household goods typically contain DC current, which is also utilized in certain industrial applications. A current that alternates is known as an alternating current.

B. ESP 32 Microcontroller

The chip that Espress if Systems created is called ESP32. This gives embedded devices access to dual-mode Bluetooth and Wi-Fi connectivity (in certain models). Although ESP32 is only the chip, its manufacturers frequently refer to the modules and development boards that incorporate it as "ESP32." The ESP32 is a family of low-cost, low-power microcontrollers that combine dual-mode Bluetooth and Wi-Fi. The ESP32 series incorporates built-in antenna switches, RF balun, power amplifier, low-noise receive amplifier, filters, and power management modules. It uses either a Ten silica Xtensa LX6 microprocessor in both dual-core and single-core variations, Xtensa LX7 dual-core microprocessor, or a single-core RISC- V microprocessor. Chinese company Espress if Systems, situated in Shanghai, is the creator and developer of ESP32. Together, they make up the ESP32 microcontroller family. Although the CPUs and capabilities of these chips vary, they all use the same SDK and have a high degree of code compatibility. The original ESP32 was also modified.

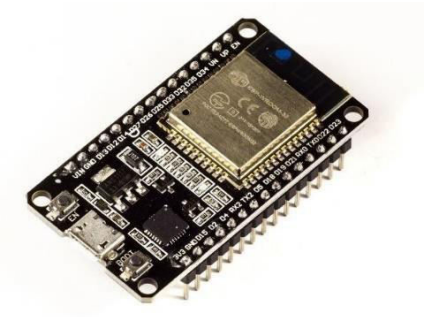


Fig.2 ESP-32S Microcontroller

ESP32 is engineered for wearable electronics, mobile devices, and IoT applications and it achieves ultra-low power consumption with the power saving features that include multiple power modes, dynamic power scaling, and a fine resolution clock. The ESP32 as an access point (hotspot), you can be connected to the ESP32 using any device with Wi-Fi capabilities without the need to connect to your router. The ESP32 Bluetooth LE supports up to nine client devices for connection.

C. Load cell:

A transducer that transforms force into a measured electrical output is called a load cell, or loadcell. Strain gauge load cells are the most widely used type of force sensor, despite the fact that there are several types. In most laboratories, strain gauge load cells still predominate over mechanical balances used for precision weighing. Hydraulic load cells are taken into consideration in remote places since they do not require a power source, while pneumatic load cells are occasionally employed where inherent safety and hygiene are desired. Strain gauge load cells are appropriate for nearly all industrial applications and provide accuracies ranging from 0.03% to 0.25% full scale..

Fig.3 Load Cell

The way a load cell functions is by translating mechanical force into readable and recordable digital values. Depending on the load cell you select, different load cells have different internal mechanisms. There are three types of load cells: strain gauge, pneumatic, and hydraulic. Of the three, strain gauge load sensors are the most often utilized. When under stress, strain gauges within strain gauge load cells cause voltage anomalies to be raised. The digital reading of weight covers the degree of voltage variation.

A load cell is used to quantify mechanical force, primarily item weight. These days, load cells are used in practically all electronic weighing scales to measure weight.

ISBN 979-8-3503-5094-4

Because of how precisely they can measure weight, they are utilized extensively. A wide range of industries that require precision and accuracy use load cells. There are four classes of load cells: class A, class B, class C, and class D. The accuracy and capacity of each class vary. Like all other contemporary load cells, beam load cells function as transducers, using strain gauges to translate force or weight into an electrical signal. The load cell's body will flex when a weight is applied. When a load is applied, the body of the load cell will flex due to the elastic properties of the metal material that it is made from.

Positive hydrogen ions are drawn to the negative plates because the charger produces an overabundance of electrons there. Lead sulfate and hydrogen combine to generate sulfuric acid and lead when the majority of the sulphate is removed, at which point hydrogen rises from the negative plates. When the reaction is almost finished, oxygen bubbles rise from the positive plates as a result of the oxygen in the water reacting with the lead sulphate on the plates to create lead dioxide once more. We refer to this as gassing.

D. Blood Bag:

One of the main uses of PVC is in blood containers, also referred to as blood bags. In actuality, PVC bag quality is crucial to blood transfusion services globally. For additional information regarding PVC blood bags, see the infographic. The units are obtained via apheresis into acid citrate dextrose (ACD) or as whole blood into bags containing nutrition phosphate and dextrose (CPD) and anticoagulant citrate. To separate the red blood cells from the remaining blood, the entire blood is centrifuged in order to break down the heavier red blood cells. Each bag typically holds 250–300 ml.

Fig.7 Blood bag

An adult typically contains five liters of blood in their body. The goal while donating whole blood (the conventional way) is to gather 500 ml, but dosage restrictions and required anticoagulant additive mix proportions require that at least 460 ml be taken. Blood can be weighed to get an estimated volume. Blood in 1 ml weighs 1.06 g. Remember the bag's own weight as well.

E. Comparator circuit:

When two voltages are compared, a comparator circuit determines which is larger by outputting a 1 (the value at the positive side) or a 0 (the voltage at the negative side). For example, comparators are frequently used to detect whether an input has achieved a predefined value. When comparing a measurable quantity, like two voltages or currents, to a reference or standard, a comparator is utilized. Depending on which input of the operational amplifier we connect the fixed reference voltage source and the input voltage too, a simple op-amp comparator circuit can be utilized to detect either a positive or a negative going input signal. A hardware electronic device known as a digital comparator or magnitude comparator accepts two binary integers as input.

One benefit of the electrical comparator is that it has less mechanical weight, which reduces vibrations. There are less moving pieces in it. The friction errors are decreased by the AC supply. The indicating device can be held at a distance for measuring units. For measuring units, the indicating instrument can be held at remote locations.

F. Tray and Clamp:

The blood bag is kept in a rectangular tray. Blood is collected in the machine using the blood bag. Anticoagulants are used in the blood bag to stop the blood from clotting while it is being collected. The blood bag contains anticoagulants such as acid citrase dextrose, which works by blocking the action of calcium ions to stop blood from clotting. To stop blood from clotting, the anticoagulants and blood must be combined uniformly. Through the use of the tray's rocking motion, the blood is uniformly combined with the anticoagulants, preventing blood clotting while the blood is being collected.

Fig.8 Tray with clamp

the clamp that is used to halt blood flow when the necessary volume is reached. The load cell is used to measure the volume of blood that is drawn from the blood bag. The microcontroller receives input from the load cell and then provides input to the comparator circuit. The 5-pin relay relay relays the connection to the other circuit by selecting and receiving input from the comparator circuit.

G. Aurdino programming:

Written in Java, the Arduino integrated development environment (IDE) is a cross-platform tool available for Linux, macOS, and Microsoft Windows. It came from the Processing and Wiring integrated development environment (IDE). It offers straightforward one-click methods for compiling and uploading programs to an Arduino board, as well as a code editor with capabilities like text cutting and pasting, text finding and replacement, automatic indenting, brace matching, and syntax highlighting. In addition, it has a text terminal, a message area, a toolbar with buttons for frequently used operations, and a hierarchy of operation menus. The GNU General Public License, version 2 governs the publication of the IDE's source code.

The Arduino IDE has specific code architecture guidelines to support the languages C and C++. A software library from the Wiring project, which offers numerous standard input and output operations, is provided by the Arduino IDE. Just two fundamental functions are needed for user-written code to be compiled and linked into an executable cyclic executive program with the GNU toolchain, which is also included with the IDE distribution. These methods are for initiating the sketch and the main program cycle, and they are linked with a program stub called main().

The Arduino IDE utilizes the argued program to translate the executable code into a hexadecimal text file, which is then loaded into the Arduino board [illegible] application.

ISBN 979-8-3503-3094-4

Installing and downloading the Arduino IDE is the initial step towards programming the Arduino board. Linux, Mac OS X, and Windows can all run the free and open-source Arduino IDE. Depending on your operating system, download the Arduino software from the official website, then follow the installation instructions. The Arduino program has a rather straightforward structure. Programs for Arduino require a minimum of two blocks.

Preparation & Execution:

Each block has a set of statements enclosed in curly braces

The steps in programming the ESP-32S microcontroller with Arduino IDE has following steps:

1. The first step would be to download and install the Arduino IDE.
2. Once installed, open the Arduino IDE and go to Files - Preferences to open the preferences window and locate the “Additional Boards Manager URLs:”
3. A text box will appear which will be empty there paste the https://dl.espressif.com/dl/package_esp32_index.json URL in the text box.
4. Now go to Tools - Boards – Board Managers to open theBoard manager window and search for ESP32.
5. click on the Install button and your board should get installed.
6. Arduino IDE is prepared to work with ESP32.
7. Connect your ESP32 board to your computer through the micro-USB cable. Make sure the red LED goes high on the module to ensure power supply.
8. Start the Arduino IDE and navigate to Tools - Boards and select ESP32Dev board as shown below
9. Open device manager and check to which com port your ESP32 is connected to.
10. Go back to Arduino IDE and under Tools - Port selects the Port to which your ESP is connected to
11. Type the code to program ESP-32S-DFR0478 microcontroller.
12. The Arduino IDE is used to program the ESP-32S-DFR0478 microcontroller to control the relay switching in the five-pin relay switch to switch from one circuit to another circuit.

III. LITERATURE REVIEW

Amit et al. have revealed that the blood collection monitor is a tool used in the blood collection procedure that shakes the blood being drawn from the donor, blending it with the anticoagulants in the blood bag. Blood coagulation is prevented by anticoagulants and the shaking moment. The volume of blood that needs to be drawn varies between 350 and 450 milliliters depending on the donor. Currently in use in hospitals and labs, electronic blood collection monitors consist of both mechanical and electrical mechanisms that oscillate a blood bag, measure blood flow, and quantify the amount of blood.

Nevertheless, these gadgets' need on outside electric power sources restricts their use in both large-scale donor camps and rural locations. The current work provides a mechanical blood collection monitor that circumvents the limitations of the traditional device by operating without the need for external power. The design considerations and methodical analysis of the suggested mechanical blood collection monitor are presented in this research. Furthermore, ANSYS and ADAMS were used to investigate the linking mechanism and related gear train, which expands the avenues for researching changes in kinetic energy and torque moment. The fundamental operation of the blood collection monitor is demonstrated in this paper.

In their work "The Order of Draw during blood collection," Giacomo Bazzano et al. have documented that a thorough literature analysis has revealed that drawing blood is one of the most frequent nursing procedures and is not without complications. One contentious issue during blood collection is the order of draws. Our goal was to determine how well the order of draws worked during blood collection in order to ensure a precise biochemical outcome. Using PubMed, Scopus, Web of Science, CINAHL, Embase, Joanna Briggs Institute, Cochrane Library, and Google Scholar, we conducted a thorough evaluation of the literature. English-language articles that were released between 2000 and 2020 were deemed appropriate.

The examination of the 11 papers revealed varying perspectives; nevertheless, the most recent data states that cross-contamination due to the wrong draw order is unique to the open drawing system. The latest research confirms that, while using the closed blood collection method, the sequence of draw has no bearing on the results; nevertheless, when utilizing the open collection system, it is advised. The principles and procedure of the blood collection are explained in this study.

According to Gummere John et al.'s article, "US4678049A Blood collection bag weight monitor," a weight monitor consists of two parts: one for attaching to a pedestal or other support structure, and the other for suspending a blood collection bag that will be filled with donor blood. The first and second parts are connected by a trough-shaped spring, which holds the second part in place relative to the first part until the filling blood collection bag exerts bending force on the spring that surpasses a predefined threshold, at which point the spring abruptly flexes and the second part is reoriented in relation to the first part.

The second half must be reoriented in order to activate a clamp connected to the blood collecting bag's fill tube, which stops more blood flow from passing through the tube into the bag. The design concept for the suggested blood collection monitor device is based on this paper.

Ye lin et al. (2020) has reported in his article “CN109394240A New type, safe and stable automatic blood collection monitor with blood clotting preventing function” that their invention discloses a new-type safe and stable automatic blood collection monitor with a blood clotting preventing function. The monitor consists of a hanging straight angle column and a blood bag hanging mechanism. The automatic blood collection monitor can be easily modified and adjusted to the actual conditions when it is placed. It can also be adjusted to the height of the actual ground when it is in a rugged area by rotating the height of the rotating position. The threaded sleeve supporting legs, threaded rotation mouths, hollow cylinder blocks, and anti-slide discs are the mechanisms that make this possible, so

ISBN 979-8-3503-5094-4

that the monitor can be placed more stable. The monitor's rotation angles can be adjusted to suit the needs through the use of a magnetic rotating disc made up of a rotating joint column, an adsorbing magnet block, an upper end round ring block, and a stop round ring block. The stop round ring block's diameter is slightly smaller than the inside hollow structure of a metal support block, allowing for rotation to occur. In the meantime, the metal at the lower end of the blood collection engine body can be well adsorbed by the adsorbing magnet block of the upper end round ring block, allowing for position adjustments. An overview of the new planned blood collection monitor model is provided in this article.

IV. METHODOLOGY

A. Functional Block Diagram:

The equipment's functioning block diagram is displayed in Figure. The two circuits that make up the instrument's main structure are one linked to the plastic tray and the other to the clamp. Through the blood bag tube and into the blood bag, the patient's blood is drawn.

Anticoagulants included in the blood bag stop blood from clotting while the blood is being drawn. To prevent clotting, the anticoagulants and blood must be combined uniformly; the tray's rocking motion does this.

Additionally, a load cell that gauges the volume of blood drawn into the blood bag is attached to the tray. The ESP-32S-DFR0478 microcontroller, which regulates the equipment's blood collection procedure, is coupled to the load cell.

The comparator circuit takes input from the microcontroller after it has received input from the load cell. The five-pin relay, which shifts the direction of current flow from one circuit to another, receives its input from the comparator circuit.

Two circuits are linked together: one for the DC motor that drives the tray and provides a rocking motion to mix the blood evenly, and another for the clamp that shuts off the patient's blood supply when the blood bag reaches the required level.

B. Functional Block Diagram:

The system's primary power source is the AC mains, and in the event of a power outage, a lead acid battery (12V-3A) is utilized as a backup. It supplies the load cell and microcontroller with input.

The process of the components is managed and controlled by the ESP-32S-DFRO0478 microcontroller. The second step of the project requires Bluetooth and Wi-Fi connectivity, which the ESP 32 microcontroller can interface with other devices to deliver. The 1 kg load cell is positioned beneath the tray to determine the blood's weight. By translating forces like compression, tension, pressure, or torque into electrical impulses that can be measured and defined, a load cell can measure weight. The microcontroller is then supplied with this electrical signal, which gives input to the comparator circuit which is connected to a five-pin relay.

Two circuits with two motors each are linked to the five-pin (12V) relay: one circuit is for the DC motor that is attached to the tray, and the second circuit is for the servo motor. When the microcontroller sends instructions to the five-pin relay, it flips from the first circuit to the second.

The blood is uniformly mixed with the anticoagulants (acid citrate dextrose) by means of the rocking movement of the tray, which is powered by a DC motor (A2212-2200kVBL). This prevents the blood from clotting by blocking the actions of calcium ions. When the blood in the blood bag reaches the required level, the microcontroller signals the circuit to cease the blood flow, which is accomplished by the servo motor (SG90-9G-SFRVO-180). A significant portion of the microcontroller's input is amplified by the amplifier before being delivered into the five-pin relay, which is used to switch circuits as needed. The servo motor, which is attached to the clamp, ceases blood flow when the necessary volume is reached.

V. RESULTS AND CONCLUSION

Simplifying the circuit, lightening the device's weight, and improving medical professionals' ease of access to the device result in a more affordable and portable blood collection monitor. The project's initial phase includes a demonstration of the blood collection's basic operating principle. In the blood bag, the blood is combined with the anticoagulants. The tray's swaying motion accomplishes this. The rocking motion of the tray is powered by a DC motor. The suggested model's guiding idea is automated blood collection without the phlebotomist's supervision. This is made feasible by the load cell-assisted weight monitoring method. The automated blood collection system is the major subject of the demonstration.

The blood weight is measured, and the microprocessor receives the data. The microprocessor sends out a signal through a 5-pin relay when the blood volume in the blood bag hits a certain threshold. A servo motor is part of the circuit to which the relay flips. With the use of a clamp, this motor interrupts the blood supply from the donor.

Table 4.1 Test trials done using blood collection monitor

S. no	Name	Amount of blood (ml)	Amount of blood measured when the machine stops (ml)
1	Subject 1	200	199
2	Subject 2	350	350
3	Subject 3	275	276

Fig.10 Test trials done using blood collection monitor

The aforementioned figure illustrates how the device weighs itself automatically and cuts off blood flow when the necessary amount is reached. The test experiment results are shown graphically in the bar chart. For the trial trials using the blood collection monitor, water and artificial blood are employed.

VI. REFERENCES

[1] Amit Nemade, Soham V. Kulkarni, Puskaraj D Sonaw wanay, Design and development of mechanically operated blood collection and monitoring device, vol. 56, issue 5, Materials Today Proceedings, 2021.

[2] Tuaillon Edouard, Kania Dramane, Pisoni Amandine, Bollore Karine, Taieb Fabien, Dried blood spot tests for the diagnosis and therapeutic monitoring of HIV and viral hepatitis B and C, Frontiers in Microbiology, vol. 11, 2020.

ISBN 979-8-3503-5094-4

[3] S. Koike, T. Takeda, Experimental Research of Blood Collected from the Peripheral Side of the Fluid Infusion Site That Is Not Affected by Fluid Infusion, Journal of Hematology, vol. 6 (1), 2017.
[4] S. Parasuraman, R. Raveendran, R. Kesavan, Blood sample collection in small laboratory animals, Journal of Pharmacology & Pharmacotherapeutics, vol.1 (2), 2010.
[5] Masanori Atsushi Nibu Yoshiaki Hirota Isao Taruta Motegi, Design and development of mechanically operated blood collection and monitoring device, Materials today proceedings, vol. 56 (5), 2022.
[6] Charles Silvers, Precision Cut-Off Weighing Apparatus. U.S. Patent No. 39,60,224. U.S. Patent and Trademark Office.
[7] Per Jansson, Nicklas LUNDMAN. Blood mixer for a blood collection bag. Patent Application No. PCT/SE20 16/050844.
[8] Satoshi Inoue, Fumiaki Inaba. Blood Collecting Apparatus (U.S. Patent No. 5153828). U.S. Patent and Trademark Office.
[9] Masanori Atsushi Nibu Yoshiaki Hirota Isao Taruta Motegi, Blood collection monitor, 2008
[10] Luciani Martin, Device for Agitating, Weighing and Clamping the Blood Bag. Patent No. FR2599837A1.
[11] Homer M Lindsey, Wilbert D Silva, Blood with Drawing Device. Patent No. US3924700, Worldwide applications

ISBN 979-8-3503-5094-4

Enhancing Lung and Colon Cancer Diagnosis: An ImageNet-Trained Transfer Learning Approach for Histopathological Image Analysis

Mullakuri Anusha
Research Scholar, Jawaharlal Nehru Technological University, Ananthapur, Sri Venkateswara College of Engineering, Tirupati, Affiliated to Jawaharlal Nehru Technological University,
Anantapuramu, India
mullakurianusha@gmail.com

[3]D Srinivasulu Reddy
Sri Venkateswara College of Engineering, Tirupati, Affiliated to Jawaharlal Nehru Technological University,
Anantapuramu, India
0000-0003-4216-5412

***Abstract*—This study presents a novel approach for lung and colon cancer diagnosis using VGG16 transfer learning applied to histopathological image analysis. Focusing on two prevalent cancers, our dataset comprised 25,000 images, with 5,000 images per class. We employed the VGG16 convolutional neural network, leveraging transfer learning to adapt the model to the unique characteristics of histopathological imaging. The research involved rigorous preprocessing to standardize and enhance image quality, ensuring the VGG16 model could effectively learn and differentiate various tissue patterns. The model's performance was evaluated using a range of metrics, yielding outstanding results: a ROC AUC Score of 0.9996456, and high scores in accuracy (0.986933), precision (0.986928), recall (0.986933), F1 score (0.986924), sensitivity (0.986933), and specificity (0.986928). These results demonstrate the model's exceptional ability in classifying both malignant and benign tissues accurately, indicating significant potential for aiding pathologists in early cancer detection and diagnosis.**

Keywords—Histopathological Image Analysis, VGG16 Transfer Learning, Lung Cancer Diagnosis, Colon Cancer Diagnosis, Convolutional Neural Networks, Medical Image Processing.

I. INTRODUCTION

In the realm of medical diagnostics, lung and colon cancers stand as two of the most challenging malignancies to diagnose accurately, due to their complex histopathological features [1] [2]. Histopathology, the microscopic examination of tissue to study the manifestations of disease, plays a crucial role in the identification and classification of these cancers [3] [4]. Recent advancements in image processing techniques, particularly in the application of Convolutional Neural Networks (CNNs) like VGG16, have opened new avenues for enhancing the accuracy and efficiency of cancer diagnosis [5] [6]. This paper delves into the intersection of these fields, presenting a novel approach using VGG16 transfer learning for analyzing histopathological images of lung and colon tissues [7] [8]. A comprehensive literature survey, reviewing ten pivotal papers in this domain, sheds light on the latest developments and methodologies applied in the histopathological analysis of lung and colon cancer, highlighting the significant strides made in applying advanced image processing techniques to medical diagnostics [9] [10].

For the analysis of histopathological cancer patterns, this study relied on the LC25000 dataset, which offers a diverse range of lung and colon cancer images, meticulously compiled by Borkowski et al.[11]. The existing literature on histopathological image analysis for lung and colon cancer diagnosis predominantly highlights the use of convolutional neural networks (CNNs) for their ability to automatically and adaptively learn spatial hierarchies of features from image data [12] [13]. These methods have shown significant promise, with various studies reporting improved diagnostic accuracy over traditional image analysis techniques [14]. Notably, CNN architectures like VGG16 have been extensively applied, with many studies capitalizing on their pre-trained weights on large datasets such as ImageNet to boost performance in medical imaging tasks [15]. However, a common drawback of these existing methods is the substantial computational resources and large amounts of labeled data required for training, which are often scarce in the medical domain. Additionally, there is a challenge in generalizing the models due to overfitting on the training data and a lack of interpretability in the decision-making process [16].

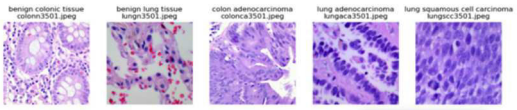


Figure .1 Sample images one image per class and its type

This work aims to overcome these limitations by implementing a transfer learning approach with the VGG16 model, fine-tuned for the specific textures and patterns of lung and colon histopathological images. Moreover, we employ data augmentation and dropout techniques to enhance generalization and mitigate overfitting. Our method also incorporates a rigorous preprocessing protocol to standardize image quality, further increasing the robustness and reliability of the model.

II. METHODOLOGY

In this study, we employed the VGG16 convolutional neural network model, leveraging transfer learning for histopathological image analysis in lung and colon cancer diagnosis. Our dataset, comprising 25,000 images across five classes, underwent rigorous preprocessing including rescaling and resizing to standardize image quality. For training 17500 images were used and 3750 images were used for training and same number of images for validation. Transfer learning with VGG16 for histopathological image classification leverages the powerful feature extraction capabilities of a model pre-trained on a vast dataset, ImageNet, to a specific task where data may be limited. VGG16, known for its simplicity and

ISBN 979-8-3503-5094-4

depth, comprises 16 convolutional layers and has been paramount in advancing image recognition tasks.

In the context of medical imaging, particularly lung and colon histopathology, the model's depth allows for the identification of intricate patterns within tissue samples that are indicative of various diseases. By freezing the pre-trained layers, the model retains its learned features, while the addition of custom top layers, tailored for the classification task, adapts it to the nuances of medical images. This method provides a head start as opposed to training a model from scratch, significantly reducing the computational cost and time, while often achieving superior performance due to the pre-learned hierarchical feature representations.

In distinguishing between different lung and colon diseases, the VGG16 model's transfer learning approach is particularly beneficial. Lung adenocarcinoma and lung squamous cell carcinoma, two primary forms of non-small cell lung cancer, present different histological features, with adenocarcinoma typically showing glandular tissue structures and squamous cell carcinoma displaying flat cells that line the inside of the airways. Similarly, in colon pathology, adenocarcinoma manifests different tissue architectures compared to benign colonic tissues, often characterized by irregular gland formations and invasion into surrounding structures. VGG16 model used in transfer learning for image classification, the total number of parameters in the VGG16 architecture pre-trained on the ImageNet dataset is approximately 138 million.

ResNet-based transfer learning leverages the deep residual networks' architecture to significantly enhance performance across a variety of tasks, particularly in the domain of image recognition and classification. This architecture enables the model to learn more complex features without a corresponding increase in training difficulty, making it an ideal choice for tasks requiring the analysis of intricate patterns and details, such as medical image diagnosis, object detection in satellite images, or even fine-grained classification tasks. By pre-training on ImageNet, a large and diverse dataset, ResNet models acquire a robust and versatile feature representation. This pre-training facilitates transfer learning, where the learned features can be fine-tuned with a relatively small amount of task-specific data.

EfficientNet (EffNet) represents a breakthrough in scaling up neural networks through a carefully balanced optimization of depth, width, and resolution, which results in superior efficiency and effectiveness over other architectures. Its compound scaling method allows EfficientNet models to achieve higher accuracy with considerably fewer parameters and lower computational cost. When applied to transfer learning tasks, EffNet models leverage their ImageNet pre-training to serve as a powerful base for a wide range of applications, from visual recognition in constrained environments to processing high-resolution images for detailed analysis. The efficiency and adaptability of EffNet make it particularly suited for scenarios where computational resources are limited, such as mobile applications, edge computing, and real-time systems. The architecture's ability to handle diverse image types and complexities also makes it an excellent choice for advanced tasks like detecting nuanced features in medical imaging or enhancing the accuracy of automated surveillance systems. EffNet-based transfer learning thus stands out for its combination of high performance, efficiency, and versatility, enabling cutting-edge solutions in fields that demand both precision and computational frugality.

The process begins with Data Input, where images are sourced from designated directories for training, validation, and testing. This is followed by Data Preprocessing, a critical step involving rescaling the images to a 0-1 range, resizing them to 224x224 pixels, setting appropriate batch sizes (32 for training and validation, 1 for testing), defining the class mode as categorical, and determining the shuffle settings (enabled for training, disabled for validation and testing).

Figure .2 Workflow of VGG16-based Transfer Learning for Cancer Detection

The core of our methodology is the Transfer Learning Setup, utilizing the VGG16 model pretrained on ImageNet, excluding its top layer, and freezing all its layers to retain the pre-learned featuresThe Model Training phase involves training the model on the training set and validating it on the validation set over 10 epochs, with defined steps per epoch and validation steps. Post-training, the model is put to test in the Model Evaluation on Test Data phase, where predictions are made on the test set and converted to class labels. The Performance Metrics Calculation involves computing key metrics such as accuracy, precision, recall, F1 score, along with the confusion matrix, sensitivity, and specificity. A crucial component of our evaluation is the ROC-AUC Analysis, where we calculate and plot the ROC-AUC score and curve for each class.

III. RESULTS AND DISCUSSION

In the presented study, we demonstrate the application of a VGG16-based transfer learning model for the classification of histopathological images, a vital task in the diagnosis of lung and colon cancers. The model was rigorously trained, as shown in Figure 3, where a stable learning curve was achieved over the course of 10 epochs, indicating a robust adaptation to the task of distinguishing cancerous from benign tissue samples.

ISBN 979-8-3503-5094-4

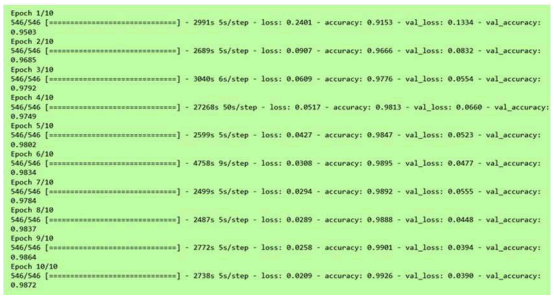


Figure .3 Training the Transfer Learning Model with Vgg16 with 10 Epoch

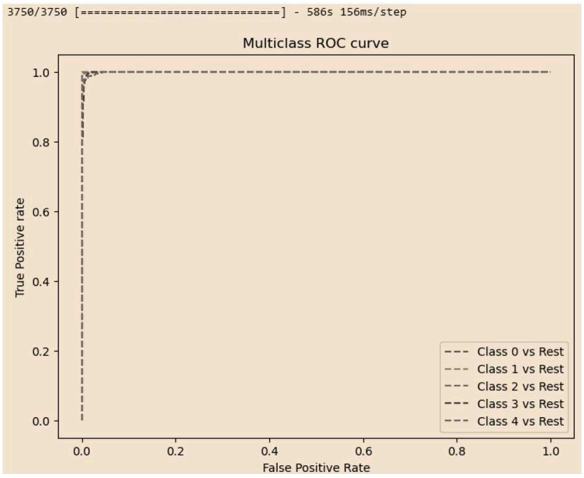


Figure .4 Vgg16 Model evaluation and ROC Plotting

The model's diagnostic capability is further evidenced in Figure 4, which portrays the ROC curves for each class. The curves' near-perfect proximity to the top-left corner reflects an excellent AUC score, underscoring the model's accuracy in class differentiation. The precision of the model is depicted in Figure 5, where the high accuracy (98.6%), alongside comparable precision, recall, and F1 scores, suggests a consistent performance across various evaluation metrics.

ROC AUC Score: 0.9996456

	Metric	Value
0	Accuracy	0.986933
1	Precision	0.986928
2	Recall	0.986933
3	F1 Score	0.986924
4	Sensitivity	0.986933
5	Specificity	0.986928

Figure .5 Performance metrics of Vgg16 model with 98.6% Accuracy

A detailed analysis of the model's performance is provided in Figure 6, 7 and 9 through the confusion matrix for three transfer learning models with ImageNet. The pronounced diagonal values indicate a high true positive rate, and the minimal off-diagonal elements point to few misclassifications, reinforcing the model's efficacy.

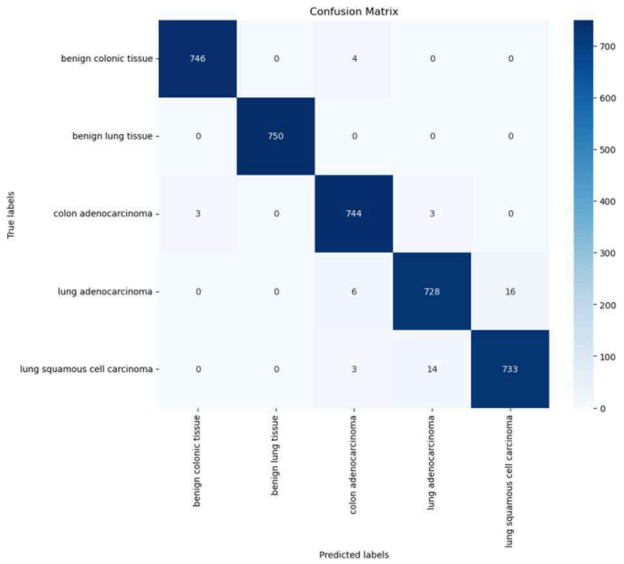


Figure .6 Confusion matrix of Vgg16 model

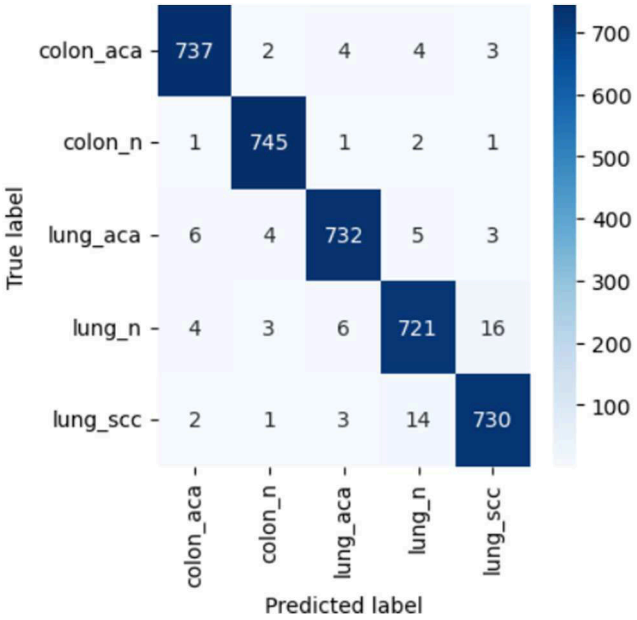


Figure 7. Confusion matrix of ResNet model

Table 1. Performance values of three transfer learning models

Metrics	EffNet	ResNet	VGG16
F1 Score	0.980785	0.977321	0.986924
Overall Accuracy	0.9808	0.9773	0.9869
Precision	0.980779	0.97732	0.986928
Recall	0.9808	0.977333	0.986933
Specificity	0.9952	0.994333	0.996733

The table 1 shows the mean values of precision, recall, F1 score, and specificity for each model across all classes (colon_aca, colon_n, lung_aca, lung_n, lung_scc), as well as the overall accuracy of each model. These metrics are commonly used to evaluate the performance of classification models in machine learning.

ISBN 979-8-3503-5094-4

From the table, we can observe that VGG16 has the highest mean precision, recall, and F1 score, indicating that it has the best performance for classifying the images across all the specified classes. Its overall accuracy is also the highest at 0.9869. EffNet follows closely, with slightly lower mean values across the board and an overall accuracy of 0.9808. ResNet has the lowest mean precision, recall, and F1 score among the three models, with an overall accuracy of 0.9773.

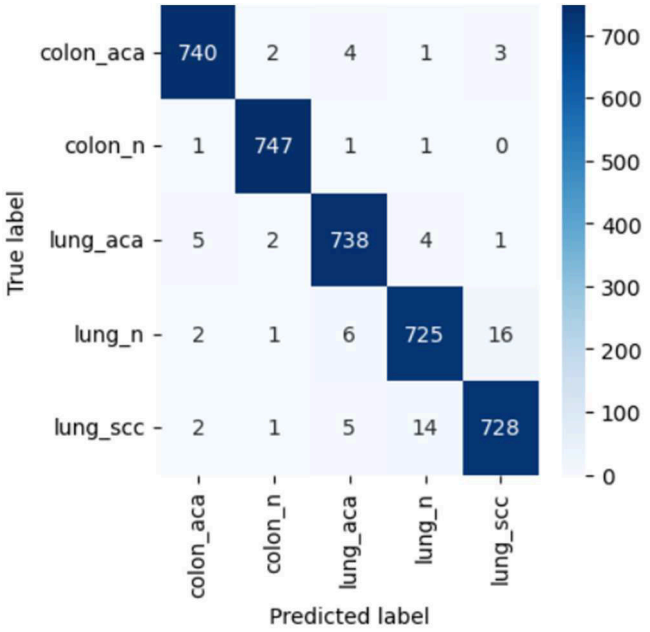


Figure 8. Confusion matrix of EffNet model

These results indicate that the adopted VGG16 model, fine-tuned through transfer learning, serves as a potent tool for the classification of histopathological images. Its ability to accurately classify images into their respective categories potentially positions it as an invaluable asset in the clinical setting, where accurate and timely diagnosis is crucial. Our findings suggest that the integration of such advanced machine learning techniques could significantly enhance the diagnostic process, providing pathologists with a reliable adjunct tool for early detection and treatment planning in lung and colon cancer care.

IV. CONCLUSION

The study presents a comprehensive evaluation of three advanced transfer learning models, underscoring the VGG16 model as the most effective in histopathological image classification for lung and colon cancer diagnosis. The success of the VGG16 model, as evidenced by its high precision, recall, F1 score, and specificity, highlights its potential utility in clinical settings for accurate and timely disease diagnosis. This work supports the integration of such AI-driven diagnostic tools to assist pathologists, potentially leading to improved patient outcomes through earlier and more accurate detection of cancerous tissues.

REFERENCES

[1] Al-Jabbar, M. Alshahrani et al., “Histopathological Analysis for Detecting Lung and Colon Cancer Malignancies Using Hybrid Systems with Fused Features,” *Bioengineering*, vol. 10, no. 3, p. 383, Mar. 2023, doi: 10.3390/bioengineering10030383.

[2] Krishnan, G.H., Nanda, A., and Natarajan, A., "Synovial fluid density measurement for diagnosis of arthritis," *Biomedical and Pharmacology Journal*, vol. 7, no. 1, pp. 221-224, 2014. doi: 10.13005/bpj/476

[3] Krishnan, G.H., Natarajan, R.A., and Nanda, A., "Comparative study of rheumatoid arthritis diagnosis using two methods," *Biomedical and Pharmacology Journal*, vol. 7, no. 1, pp. 379-381, 2014. doi: 10.13005/bpj/502

[4] Hari Krishnan, G.H., Ananda Natarajan, R., and Nanda, A., "Impact of upper limb joint fluid variation on inflammatory diseases diagnosis," *Journal of Electrical Engineering and Technology*, vol. 9, no. 6, pp. 2114-2117, 2014. doi: 10.5370/JEET.2014.9.6.2114

[5] Mohandass, G., Ananda Natarajan, R., and Hari Krishnan, G., "Comparative analysis of optical coherence tomography retinal images using multidimensional and cluster methods," *Biomedical Research* (India), vol. 26, no. 2, pp. 273-285, 2015.

[6] Krishnan, G.H., Natarajan, R.A., and Nanda, A., "Microcontroller based non invasive diagnosis of knee joint diseases," in Proc. of the 2014 *International Conference on Information Communication and Embedded Systems (ICICES 2014)*, 2015. doi: 10.1109/ICICES.2014.7034178

[7] V. G. Anand, G. H. Krishna, G. Mohandass, R. J. Hemalatha and S. Sundaram, "Predicting grade of prostate cancer using image analysis software," *Trendz in Information Sciences & Computing (TISC2010)*, Chennai, India, 2010, pp. 122-124, doi: 10.1109/TISC.2010.5714621.

[8] Hemalatha, R.J., Krishnan, G.H., Umashankar, G., and Abraham, S., "Computerized breast cancer detection system," *Biosciences Biotechnology Research Asia,* vol. 11, no. 2, pp. 907-910, 2014. doi: 10.13005/bbra/1357

[9] Sudhakar, T., Hari Krishnan, G., Krishnamoorthy, N.R., Janney J, B., Pradeepa, M., and Raghavi, J.P., "Sleep Disorder Diagnosis using EEG based Deep Learning Techniques," in Proc. of the 2021 IEEE 7th *International Conference on Bio Signals, Images and Instrumentation (ICBSII 2021)*, 2021. doi: 10.1109/ICBSII51839.2021.9445158

[10] Nagarjuna Reddy, A., Hari Krishnan, G., and Raghuram, D., "Real time patient health monitoring using raspberry PI," *Research Journal of Pharmaceutical, Biological and Chemical Sciences*, vol. 7, no. 6, pp. 570-575, 2016.

[11] Hari Krishnan, G., Umashankar, G., and Abraham, S., "Cerebrovascular disorder diagnosis using MR angiography," *Biomedical Research (India),* vol. 27, no. 3, pp. 773-775, 2016.

[12] Hari Krishnan, G., Abinaya, N., Hemalatha, R.J., and Mohandass, G., "Hardware implementation for feedback control based health monitoring and drug delivery," *Biomedicine (India)*, vol. 37, no. 1, pp. 123-126, 2017.

[13] Santhosh, S., Juliet, A.V., and Krishnan, G.H., "Impact of electrodes separation distance on bio-impedance diagnosis," *Biomedical and Pharmacology Journal,* vol. 14, no. 1, pp. 141-146, 2021. doi: 10.13005/bpj/2108.

[14] Hari Krishnan, G., Arun Kumar, G.S., Arun, V., and Saravana, R., "Comparative study of diabetes foot diagnosis using ABPI," *International Journal of Engineering and Technology(UAE),* vol. 7, no. 2, pp. 40-42, 2018. doi: 10.14419/ijet.v7i2.25.12365

[15] Santhosh, S., Vimala Juliet, A., and Hari Krishnan, G., "Simulation of Signal Generation and Measuring Circuit and Real Time IoT Based Electrical Bio Impedance Cardiac Monitoring System," in *Advances in Intelligent Systems and Computing*, vol. 1039, pp. 701-706, 2020. doi: 10.1007/978-3-030-30465-2_77

[16] G. H. Krishnan, K. Sukanya, J. S. S. Reddy, K. S. P. Reddy, N. Niharika and M. S. S. Kiran, "CNN based Image Processing for Crack Detection on HT Insulator’s Surface," *2023 International Conference on Sustainable Computing and Smart Systems (ICSCSS)*, Coimbatore, India, 2023, pp. 618-621, doi: 10.1109/ICSCSS57650.2023.10169853

ISBN 979-8-3503-5094-4

Renal Anomalies through Ultrasound Imaging and Segmentation Techniques

T.Mangayarkarasi
Associate Professor, Department of Instrumentation and Control Engineering, Sri Sairam Engineering College,and Research Scholar, B.S.Abdur Rahman Crescent Institute of Science and Technology, Chennai,India,
mangairaajkumar@gmail.com

D.Najumnissa Jamal
Professor, Department of Electronics and Instrumentation Engineering, B.S.Abdur Rahman Crescent Institute of Science and Technology, Chennai, India, Communication mail-najumnissa.d@crescent.education

S.Kaja Mohideen
Senior Professor, Department of Electronics and Communication Engineering, B.S.Abdur Rahman Crescent Institute of Science and Technology, Chennai, India

***Abstract*— In present days, the recent encroachment in the medicinal field and abnormality detection methodologies mutually by instrumentation also clinical methods has reached to a height in such a way that the clinical physicians are able to give an appropriate warning for the early detection of any disease and cure. This is widely applied in detecting abnormality of vital organs like kidney, heart, liver, gallbladder, etc. In this proposed work, the detection of prominent abnormalities in kidneys using US scanned images is analyzed using segmentation technique for feature extraction. The extracted features are saved as .csv file in which the significant Max intensity features exceeds threshold value detected as abnormal (cyst, calculi, tumor) image otherwise normal.**

Keywords—Watershed segmentation, Feature extraction, OTSU threshold, Peak Signal Noise Ratio (PSNR), Ultrasound (US) images

1. Introduction

Kidney abnormalities like calculi arise by agglomeration of definite minerals and salts generally inclusive of calcium content and uric acid in urine. It happens because of insufficient water intake. Kidney calculus primarily occurs while our body has been deficient in liquefied nevertheless gathers many wastes. This kidney stone arises in an individual who is having diabetes, obesity and high blood pressure. Several techniques were analyzed for kidney stone identification through Computer Tomography scan, blood test, urine test, also MRI scan. Detecting kidney stones is unrealistic in the case of huge quantities of data. In this survey, kidney abnormalities such as cyst, tumor, calculi, as well as normal are identified and classified using segmentation techniques, especially watershed algorithms.
The intention of this proposed model is mentioned as follows

- To categorize input ultrasound kidney 100 images as cyst, stones, tumor, and normal.
- To remove noise from input images, filtering techniques are going to be utilized.
- To classify ultrasound kidney image as normal or malicious (cyst, stones, tumor) depends on characteristic extraction declared as .csv categorizer.
- To estimate PSNR for comparing the quality of existing image with predicted image

Also, the foremost contribution of this research work is mentioned below

- Segmentation techniques like watershed algorithm shave been projected to grain ultrasound images for feature extraction scenarios.
- Moreover, relevant features are extracted for kidney disease categorization such as normal, cyst, tumor, calculi on US images.
- Thresholding function is necessary for distinguishing normal and abnormal depends on image intensity level.

Fig. 1 illustrates the normal kidney image versus ultrasound kidney image for detecting whether an individual has kidney disease for instance cyst, tumor, calculi or normal image.

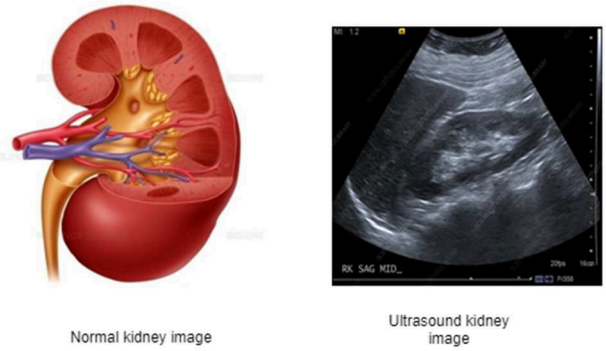


Fig1. Normal Vs Ultrasound kidney image

2. Background

Fuzhe et. al [8] introduced deep learning algorithms like HMANN, multi-layer Perceptron, and Back propagation to diagnose CKD and also classifying abnormal from normal images as gathered US images. Here, the classification accuracy has achieved around 97.5% in categorization of images. This paper focused on boundary interval regression along with pixel classification net for distinguishing injure region specified as pixels on images from non-pixelsapplied by Yin et. al [28]. Through this work, segmentation of kidney images is done automatically that makes it helpful in medicinal domains. Machine learning algorithms such as ANN, SVM, NB were utilized by Reem et. al [26] in support of CKD prediction very accurately that attains accuracy as 98% using Saudi dataset from Fahd University Hospital (KFUH). Moreover, advanced techniques such as KNN and FFNN focused on finding End stage renal disease diagnosis in an individual who is having type-2 diabetes had analyzed

ISBN 979-8-3503-5094-4

by Sunil et. al [32] achieves accuracy as 84%. Prediction of CKD disease had investigated using mining and back propagation on ANN Chakrapani et. al [4]. In addition to that, comparative analysis is also done for CKD prediction as classification, regression along with ANN. From that, this work realizes ANN gives good results with accuracy as 99.2%. The CKD disease diagnosis was found at a premature stage by evaluating an accuracy measure that attains 97% analyzed by Himanshu et. al [11]. Jyoti et. al [15] kidney stone diagnosis were examined using Support Vector Machine, and K-Nearest Neighbor by segmenting input images as kidney stones mentioned as pixels and normal images. Shruti kannan et. al [29] developed CNN multi-label classifier mainly to spotlight the categorization on biopsy input images via the Glomerular segmentation algorithm. By Faleh, 2017 utilized CT scan images for early diagnosis of kidney disease especially tumors. The input image is given as 2dimensional CT images in Hui Zhang, 2019 for identifying lesions in the kidney. Viswanath, 2015 used US images for detecting kidney stones. Sridhar, 2002, &Tanzila 2016 analyzed US images for segmenting images and finding renal calculi. US images for categorizing kidney disease automatically by Mariam, 2015, charumathi 2020 identified CKD disease since retinal images. Intisar, 2020 explained how deep learning methods suitable for disease diagnosis on medical domains from MRI, US, CT, X-ray, clinical images. Pooney, 2019 ultrasound 3d images for renal segmenting.
Dovgan E, 2020 introduced novel techniques for patients who are affected by CKD as Renal Replacement Therapy.Pawar, 2017 identified normal kidney image or cyst arising in kidney from US images.Kidney failure leads to life terrorization, hence it is essential to detect kidney abnormalities such as tumor arises in specific regions, cyst found, calculi etc. To address in finding and classifying various kidney disease, this proposed model focused on image segmentation especially watershed segmentation along with feature extraction as image area, equivalent, orientation, major and minor axis length, image perimeter, using US kidney images. Guanyu Yang, 2018 presented renal tumor classification from images through segmentation technique.

III. Materials and methods

3.1 Materials

We afford Python programming codes for each and every technique to investigate with mentioned algorithms especially watershed segmentation approach to diagnose kidney abnormalities issues. The software specification have implemented for distinguishing the kidney images are described below
Language Used: Python Programming language
Version: Python 3.7
Import all the required Keras libraries since it comes under image segmenting procedure that partitions the input kidney US imagery to several segments/ pixels. Import TensorFlow library for image classification and perform data flow diagrams.

3.2 Proposed methodology

The proposed methodology for distinguishing abnormal (kidney disease) and normal analyzed using Ultrasound images comprises four stages 1. Input US images 2. Preprocessing phase 3. Image Segmentation and 4. Feature extraction.

Fig. 2 depicts the overall framework for our novel work which comprises materials and methods used to diagnose kidney abnormalities and distinguish each image as normal from abnormal.

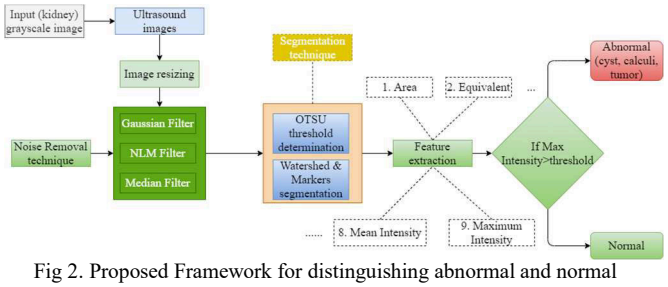


Fig 2. Proposed Framework for distinguishing abnormal and normal

3.2.1 Phase 1- Gathered input data

The input images have been taken as ultrasound kidney (grayscale) images towards diagnosing kidney disease by distinguishing normal kidney image from abnormality as either cysts, or tumor, or calculi. For this proposed model, we used 130 US kidney images obtained from Saanvi Fertility Clinic,Chennai, Dr.Joe Antony Ultrasound Scan Centre, Ernakulam, in which each and every image has different dimensions and specifications.

3.2.2 Phase 2- Pre-processing phase

Pre-processing is a widespread name designed for working with input images at the least intensity of generalization (i. e) mutually input as well as output are intensity imagery. The foremost intention of pre-processing stage is being noise reduction also towards distinction improvement to enhance quality of image. Viswanath, 2015 applied a pre-processing method to reduce noise in US images. Pawar 2017 the image subsequent to this phase stepped into image partitioning phase for segmenting images using gradient vector flow. The purpose of this phase is described as below:

- Image resizing
- Noise removal technique

i) Image resizing- Image resizing is necessary whenever we want to enlarge or reduce pixel quantity on kidney US image. Now, the input US imagery is resized with resolution as 128*128 which is described as equation (1)

$$\text{resize}(128,128)\text{-----}(1)$$

Represents proportion of specified image as height and width.
ii) Noise removal technique- Noises such as salt noise, pepper noise present in the US images leads to accuracy reduction on segmentation of abnormalities as either cyst or tumor or calculi in kidney images.Tanzila, 2016 utilized Gabor filter for noise reduction, smoothen the resulting image. Hence in our proposed work, the smoothening and sharpening of image had been done via filtering technique as median filter, Gaussian filter GF [Jyoti] and Non Local Mean filter. Likewise, Nithya et. al 2019, Jyoti.verma 2017 approached the median filter method to remove noise from an input image. By [Nithya] median filter equation (2) is described below,

$$I(i,j) = median_{(s,t)\in C_{xy}}\{G(s,t)\} - - - - - -(2)$$

Median refers median value, C_{xy} image co-ordinates
Noise available in images leads to unexpected alters in image pixel values. So, we are in need of executing the filtering process in images. Now, the filtering techniques were applied for altering/modifying images to enhance the image features

ISBN 979-8-3503-5094-4

and also quality of images. The following are the filtering techniques used for our work are explained

a. Median filter- The most significant filtering approach is median filter which is similar to mean filter frequently used to eliminate unnecessary noise from input images. Such kind of reducing noise helpful to enhance outcomes in further processing stages for instance: boundary discovery found in images. Likewise, here median filter focused on noise reduction from input kidney US images.

b. Gaussian filter- This type of filter works similar to linear filtering technique that generally intended for smear the images. Mainly, this filter helps to smear the boundaries and also diminish the dissimilarity of images. Initially, we applied three filtering techniques such as median filter, Gaussian filter and Non local Mean filter to remove noises available in images and also resizing images by finding the PSNR values shown in table-3. Among three, Gaussian filters have high PSNR values that are highly possible to show the image quality more enhanced.

c. Non Local Mean Filter- NLMF is also a noise removing technique that captures average values of pixel clusters from the image encircling marked pixels to flatten the kidney images. Here, filtering is done especially in pixels of an image by comparing the weight of target pixels along with average pixels calculated in an image.

3.2.3 Phase 3- Image Partitioning technique

Image partitioning process is a method of dividing the original image into several divisions. Each division is specified as pixels of an image known as point of images. Basically, Image segmentation is mainly used to specify the particular area in kidney US image manually built through Graphical User Interface. Faleh, 2017, Fuzzy C-Means algorithm applied for segmentation procedure. Chao, 2016 apply multithreading approach suitable for enhancing segmentation process speed during realization in which three segmentation carried out as renal cortex, column, pelvis and renal medulla segmentation marked in kidney image. Tanzila, 2016 have preferred two segmentations as cell and region based segmentation to categorize abnormal regions from normal. Fuat Turk, 2020 discovered novel technique for segmenting renal tumor from clinical tests. Kanishka Sharma, 2017 segmenting kidney images involuntarily and noticed as abnormal if cysts found. Image based segmentation [Helena]. A specific watershed algorithm is established on mining certain forefront (foreground) and backdrop (background) and subsequently the indicators (markers) create watershed dash and find the accurate margins in every image. Hence, we have chosen a watershed algorithm as image segmentation technique for segmenting input kidney images into several classifications as either cyst or tumor or calculi or normal images.

A. Watershed algorithm

For segmenting kidney input images into separate segments as well as identification as cysts were carried out using watershed algorithms by Soille, 2003, Faleh 2017. By Kyungsoo Bae, 2013, introduced watershed technique to mark the images, based on that markers input kidney images were segmented into several regions to spot the cyst available in kidney images. The purposeful of morphological watershed division method being imagine input gray level picture into its form portrayal sort it out into least escalated of picture, catchment bowl and watershed boundaries. Subsequently, in our proposed work, we utilized this calculation to portion the input kidney US picture, thick locales of picture indicated as less heights too straightforward districts got to be taken as more prominent elevations which appears to be form façade. Consequently, this calculation is well suited for picture division which recognizes influenced regions exceptionally effectively.

B. Markers Segmentation

Parvati, 2008 made markers on affected areas in images finally images were rebuilt to identify kidney diseases as abnormal. The remaining area in an image marked as normal. Similarly, after images are segmented into pixels, the regions are marked accurately designed to distinguish abnormal from normal on US images very easily.

C. Threshold method

Thresholding is necessary to divide input images into smaller regions which is called pixels of an image. OTSU is utilized to carry out involuntary image thresholding. In simplest form, OTSU algorithm proceeds at a certain intensity threshold that divides each pixel into two groups explicitly foreground image and background image. It mainly activates in histogram representation.

3.2.4 Phase 4- Feature extraction

Jamshid, 2016 sets threshold level as 15cc/ kg/minute, K´ılvia, 2019 focused on detecting failure in renal kidney. Also Pawar, 2017 introduced Feed Forward Artificial Neural network to train relevant features from an image after pre-processing. In this work, mark the threshold value for both features of mean intensity and maximum intensity as 86 and 147. An image has either two or three grains or more. Therefore, if anyone grain of an image's intensity feature exceeds the threshold limit, that portion of image is noticed as abnormal. Likewise, maximum intensity features beyond threshold level indicate that specified region is being malicious (either cyst or calculi or tumor).

Table 1. Features used to specify kidney image and its description

Key features	Description
Area	Overall mean area of human organ
Equivalent	Equivalent deemed as resolution to count number of pixels in an image
Orientation	To detect orientation features in the organ
Major Axis length	Lateral dimensions of kidney
Minor Axis Length	Specifies horizontal dimensions of kidney US image
Perimeter	Circumference of cross-sectional dimensions of kidney image
Minimum Intensity	Denotes minimum absolute value of square of amplitude of scattered region in kidney
Mean Intensity	Average value of amplitude square on scattered region in kidney US image
Maximum Intensity	It indicates the maximum value of a square of scattered area presented in the kidney image.

Byung Kwan Park, 2019 depicts valuable attribute imaging for distinguishing renal AngioMyoLipoma as AML. Sridhar, 2012 intended for calculi extraction from urinary bladder and urethra.

The features estimated for kidney disease diagnosis via segmentation as described in table 1. This mainly focuses on classification of input image as abnormal means whether cysts/ tumor/ calculi found in both left or right kidney and

ISBN 979-8-3503-5094-4

normal.From these features, we focused on two features namely Mean Intensity and Max Intensity for finding kidney US image as normal or abnormal (cyst, tumor, calculi).

4. Algorithm for proposed work

The algorithm for proposed structure as below:
First step: Get the ultrasound kidney images.
Second step: The image is resized with dimensions 128*128.
Step 3: Perform noise removal technique using Gaussian, NLM, and Median filter.
Step 4: Calculate PSNR value for comparing image quality.
Step 5: Apply segmentation technique to grain the images which specifies pixels of kidney US image.
Step 6: Carry out feature extraction, if maximum intensity exceeds threshold value, report that calculi, tumor, cyst is present.
Step 7: Exit

5. Results and Discussion

5.1 Input images: Initially the input images assembled as US images of kidney patients to detect abnormalities if any arises in both left and right kidney. Fig. 3 (a) depicts Sample kidney US image. The image after resizing being minimizes and maximizes the number of pixels in the image resolution as 128*128 depicted in Fig. 3(b).

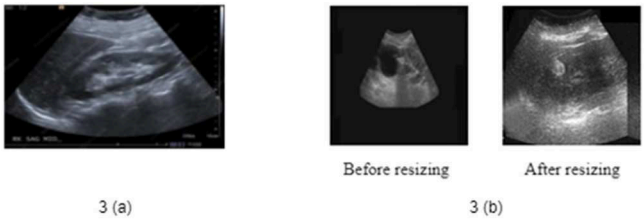
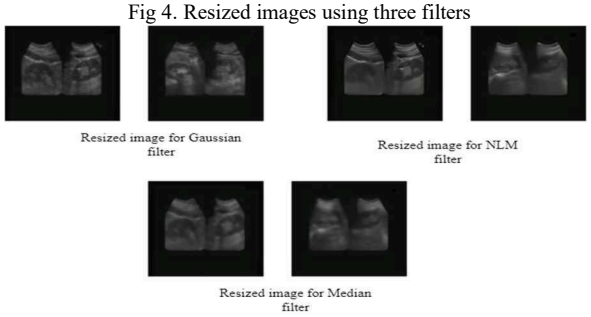


Fig 3 (a) depicts the sample kidney US image and Figure. (b) Contrasting images before and after resizing

Noise reduction: Apply noise removal technique to reduce noise in input kidney image using median, NLM and Gaussian filter. Fig. 4 showed that the resized images for three filters as GF, NLM, and MF.

Fig 4. Resized images using three filters



5.3 PSNR- Peak to Signal Noise Ratio is evaluated to compare the image quality (i. e) Contrast between original image and compressed image (i. e) PSNR refers to an image eminence representation. If PSNR relative amount is high, the compressed image quality is also high. PSNR value measurement is represented in decibels level. Also, PSNR value found for each filtering technique is depicted in table 2.

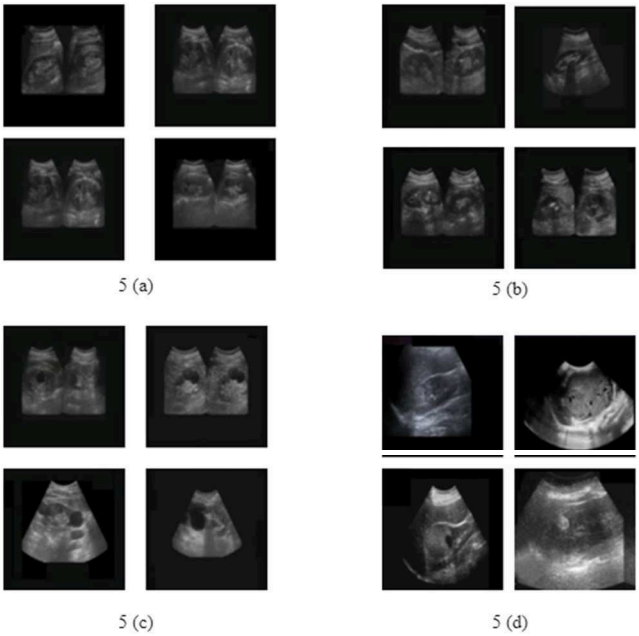
Table 2. Estimating PSNR value through filtering technique

Filtering technique	PSNR value (decibels)
Gaussian filter	40.975
NLM filter	38.496
Median filter	39.007

From table 3, we found that the Gaussian filter has very high PSNR value that specifies the Ultrasound image in a good quality to diagnose kidney disease obviously. Consequently, we are utilizing Gaussian filter technique to segment images into specific regions to identify abnormal portions of image as cyst, tumor, and calculi.

5.4 Segmentation: In the segmentation phase, preferred margins of either cyst or tumor or calculi are recognized in US images. Perform Segmentation approach to fragment the images into smaller portions called pixels or grains. Each image has two or three or more grains. The image intensity level ranges between 0 and 1. '0' specifies minimum intensity and '1' identified as maximum intensity. Here segmentation is done using watershed, markers and OTSU threshold method.

Fig. 5 (a) Normal images, (b) Calculi images after GF, (c) Cyst images after GF and (d) Tumor images after GF



After filtering using GF, normal images found in fig 5 (a), calculi images in figure 5 (b). A picture of Cyst identification after GF applied is depicted in figure 5 (c), and finally figure 5 (d) illustrated regarding tumor images after GF applied. Then apply watershed segmentation technique to segment the images into various regions, marking specified regions named as markers and finally labeling each region in images should be specified as normal or abnormal shown in figure 6.

5.5 Extract relevant features: The relevant features are stored in .csv file. Some sample images along with specified features are mentioned in table 3. From that, we extracted mainly two features depicted in table 5, when mean intensity, also maximum intensity of pixel image exceeds threshold limit (86, 147), then that grain is specified as cyst, tumor, calculi. The remaining region is identified as normal.

ISBN 979-8-3503-5094-4

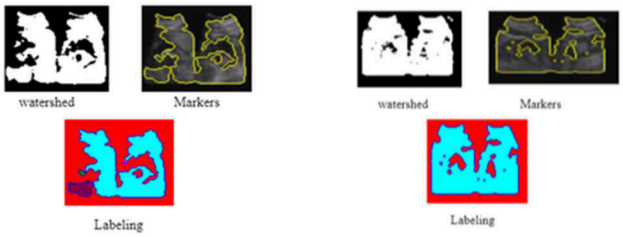


Fig.6. Apply Watershed segmentation, markers, and Labeling to Identify abnormal region exactly

Table 3. Sample table for all features images stored in .csv file

US Images	Gr	A	O	Major	Minor	Pe	Min I	Mean I	Max I
Image 1	1	2741.25	-1.58	83.50	80.30	549.69	6	15.50	70
	2	1085.25	88.69	51.59	38.28	348.29	38	63.68	129
Image 2	1	2876.5	-2.20	81.88	79.30	546.21	0	17.41	68
	2	950.25	86.62	51.39	38.88	296.36	40	89.24	155
	3	8.25	3.27	4.76	3.12	12.53	46	52.15	69
	4	0.25	45.0	0.	0	0	63	63	63
	5	0.25	45.0	0	0	0	48	48	48
Image 3	1	2878.25	-6.2	82.35	78.88	486.55	8	17.52	82
	2	499.5	2.3	33.44	20.07	107.70	42	68.63	131
	3	489	-6.46	37.19	19.8	120.75	43	70.84	129
Image 4	1	2910.75	-12.09	81.32	78.95	547.9	0	18.73	63
	2	919.5	76.19	51.16	38.75	293.76	46	77.63	148
	3	6.5	-4.02	3.46	2.55	9.01	48	53.88	62
Image 5	1	2711.75	0.36	84.27	81.27	458.78	0	3.62	93
	2	1160.5	-89.34	48.53	36.02	220.09	33	64.89	139
Image 6	1	2949	14.71	80.91	79.56	515.38	11	18.05	69
	2	103.75	73.69	16.03	9.39	52.35	45	80.12	125
	3	764.8	-56.83	49.31	29.9	182.62	42	69.3	136
Image 7	1	2765.25	2.96	83.12	81.13	481.04	0	43.93	51
	2	1098.75	-89.01	46.99	38.26	239.23	35	4.16	55
Image 8	1	2788.25	-1.74	83.39	79.90	486.92	0	3.83	53
	2	1075.75	88.41	51.45	36.16	237.33	34	67.4	133
Image 9	1	2852	2.13	81.91	79.56	515.60	12	19.05	67
	2	1001	-88.23	50.8	39.9	267.41	46	77.82	142
Image 10	1	2790	87.43	82.69	81.8	454.28	0	12.18	90
	2	1089	-2.45	43.23	39.55	214.20	56	103.63	187
	3	0.25	45.00	0	0	0	74	74	74

In Table 3 the following abbreviations used:
A-Area; Gr-Grain;O-Orientation; Major-Major axis length; Minor-Minor axis length; Pe- Perimeter; Min I-Minimum Intensity; Mean I-Mean Intensity; Max I-Maximum Intensity.

In table 4, the first two rows specify image 1 since the first image has two grains or pixels. The values for every feature were described in the above table. Likewise, image 2 has five grains in which each feature value is mentioned. From this table, two main features are mined namely mean intensity and max intensity portrayed in table 5 to distinguish the images as normal region and abnormal region.

Table 4. Feature extraction from US images

Images	Grain	Mean Intensity	Maximum Intensity
US image 1	1	3.62	93
	2	64.9	139
Image 2	1	17.41	68
	2	**89.24**	**155**
	3	52.15	69
	4	63	63
	5	48	48
Image 3	1	21.28	77
	2	**96.2**	**172**
Image 4	1	11.8	86
	2	**99.6**	**214**

Few images along with its significant features were displayed in Table 4. From the results shown in table 5, we conclude that image 1 classified as normal, image 2 defined as calculi, image 3 as cyst and image 4 specified as tumor depends on threshold limit value.

6. Conclusions

In this proposed work, we analyzed Ultrasound images for kidney disease diagnosis and distinguished the images as normal and abnormal (cysts, tumor, calculus) found. Moreover, we developed filtering techniques such as Median, Gaussian and NLM filters to remove noise in images. In addition to that, watershed segmentation techniques were applied on US images to segment the images into pixels or grains, markers were utilized to mark specific sections in an image which makes easy identification of abnormal regions. Finally, feature extraction introduced to mine relevant features as mean intensity and maximum intensity. Based on the thresholding limit, we categorized the US kidney imagery as standard (normal) from abnormal which was noted as either cysts or tumors or calculi. This work is applicable for real time medicinal applications through kidney disease diagnosis.

7. References

[1] Byung Kwan Park "Renal Angiomyolipoma Based on New Classification: How to Differentiate It From Renal Cell Carcinoma", *American Journal of Roentgenology*, 2019, March 2019, Volume 212, No 3,doi.org/10.2214/AJR.18.20408.

[2] Chakrapani, Sumit Raj, VibhavPrakash Singh, DhrubJyotiKalita, "Detection of Chronic Kidney Disease Using Artificial Neural Network", *International Journal of Applied Engineering Research* ISSN 0973-4562 Volume 14, Number 10, 2019.

[3] Chao Jin, Fei Shi, Dehui Xiang, Xueqing Jiang, Bin Zhang, Ximing Wang, Weifang Zhu, Enting Gao, and Xinjian Chen "3D Fast Automatic Segmentation of Kidney Based on Modified AAM and Random Forest", *IEEE Transactions On Medical Imaging*, Vol. 35, No. 6, June 2016.

[4] CharumathiSabanayagam, Dejiang Xu, Daniel S W Ting, Simon Nusinovici, Riswana Banu, Haslina Hamzah, Cynthia Lim, Yih-Chung Tham, Carol Y Cheung, E Shyong Tai, Ya Xing Wang, Jost B Jonas, Ching-Yu Cheng, Mong Li Lee, Wynne Hsu, Tien Y Wong, "A deep learning algorithm to detect chronic kidney disease from retinal photographs in community-based populations", *Lancet Digital Health*, 2: e295–302, June, 2020.

[5] Dovgan E, Gradiš˘ek A, Lus˘trek M, Uddin M, Nursetyo AA, Annavarajula SK, et al. (2020) Using machine learning models to predict the initiation of renal replacement therapy among chronic kidney disease patients. *PLoS ONE* 15(6): e0233976. https://doi.org/10.1371/journal. pone.0233976, 2020.

ISBN 979-8-3503-5094-4

[6] Faleh H. Mahmood, Nada A. Mahmood, Abdul Rahman H. Ismaeel "Automated Methods to Segment Kidneys and Detect Tumors Using CT Images", *Iraqi Journal of Science*, Vol. 58, No.3B, pp: 1555-1564 DOI: 10.24996/ ijs.2017.58.3B.20, 2017.

[7] FuatTürk, Murat Lüy and Necaattin Bari "Kidney and Renal Tumor Segmentation Using a Hybrid V-Net-Based Model", *Journal of Mathematics*, 8, 1772; doi:10.3390/math8101772, 2020.

[8] Fuzhe Ma, Tao Sun, Lingyun Liu, Hongyu Jing "Detection and diagnosis of chronic kidney disease using deep learning-based heterogeneous modified artificial neural network", *Future Generation Computer Systems*, Volume 111, October, Pages 17-26, https://doi.org/10.1016/j.future.2020.04.036, 2020.

[9] Guanyu Yang, Guoqing Li, Tan Pan, Youyong Kong, Jiasong Wu, Huazhong Shu, Limin Luo, Jean-Louis Dillenseger, Jean-Louis Coatrieux, Lijun Tang, Xiaomei Zhu "Automatic Segmentation of Kidney and Renal Tumor in CT Images Based on Pyramid Pooling and Gradually Enhanced Feature Modules", *24thInternational Conference on Pattern Recognition*, DOI: 10.1109/ICPR.2018.8545143, 2018.

[10] Helena R. Torres , Sandro Queiros , Pedro Morais , Bruno Oliveira , ´ Jaime C. Fonseca , Joao L. Vilac, "Kidney Segmentation in Ultrasound, Magnetic Resonance ˜ and Computed Tomography Images: A Systematic Review", *Computer Methods and Programs in Biomedicine*, doi: 10.1016/j.cmpb.2018.01.014, 2018.

[11] Himanshu Kriplani, Bhumi Patel and Sudipta Roy "Prediction of Chronic Kidney Diseases Using Deep Artificial Neural Network Technique", *Computer Aided Intervention and Diagnostics in Clinical and Medical Images*, https://doi.org/10.1007/978-3-030-04061-1_18, 2019.

[12] Hui Zhang, Yurong Chen, Yanan Song, Zhenlin Xiong, Yimin Yang, And Q. M. Jonathan Wu, "Automatic Kidney Lesion Detection for CT Images Using Morphological Cascade Convolutional Neural Networks", *Special Section On Biologically Inspired Image Processing Challenges And Future Directions*, Digital Object Identifier 10.1109/Access.2019.2924207, 2019.

[13] Intisar Rizwan I Haque, Jeremiah Neubert, "Deep learning approaches to biomedical image segmentation", *Informatics in Medicine Unlocked*, https://doi.org/10.1016/j.imu.2020.100297, 2020.

[14] Jamshid Norouzi, Ali Yadollahpour, Seyed Ahmad Mirbagheri, Mitra Mahdavi Mazdeh,andSeyed Ahmad Hosseini "Predicting Renal Failure Progression in Chronic Kidney Disease Using Integrated Intelligent Fuzzy Expert System", *Computational and Mathematical Methods in Medicine*, Article ID 6080814, 9 pages http://dx.doi.org/10.1155/2016/6080814, 2016.

[15] Jyoti Verma, Madhwendra Nath, Priyanshu Tripathi, and K. K. Saini "Analysis and Identification of Kidney Stone Using Kth Nearest Neighbour (KNN) and Support Vector Machine (SVM) Classification Techniques", ISSN 1054-6618, *Pattern Recognition and Image Analysis*, Vol. 27, No. 3, pp. 574–580, 2017.

[16] K. Viswanath, Gunasundari. R, Syed AathifHussan "VLSI implementation and analysis of kidney stone detection by level set segmentation and ANN classification", *International Conference on Intelligent Computing, Communication and Convergence*, 48, 612-622, 2015.

[17] Kanishka Sharma, Christian Rupprecht, Anna Caroli, Maria Carolina Aparicio, Andrea Remuzzi, Maximilian Baust& Nassir Navab "Automatic Segmentation of Kidneys using Deep Learning for Total Kidney Volume Quantification in Autosomal Dominant Polycystic Kidney Disease", *Scientific Reports* | 7: 2049 | DOI:10.1038/s41598-017-01779-0, 2017.

[18] Kilvia L. De Almeida, LucíliaLessa, Anny Peixoto, Rafael Gomes, Joaquim Celestino. "Kidney Failure Detection Using Machine Learning Techniques" *8th International Workshop On Advances In ICT Infrastructures And Services (ADVANCE 2020)*, Candy E. Sansores, Universidad del Caribe, Mexico, Nazim Agoulmine, IBISC Lab, University of Evry - Paris-Saclay University, Jan 2020, Cancún, Mexico. pp.1–8. ffhal-02495264f, 2020.

[19] Kyungsoo Bae, Bumwoo Park, Hongliang Sun, Jinhong Wang, Cheng Tao, Arlene B. Chapman, Vicente E. Torres, Jared J. Grantham, Michal Mrug, William M. Bennett, Michael F. Flessner, Doug P. Landsittel, and Kyongtae T. Bae, for the Consortium for Radiologic Imaging Studies of Polycystic Kidney Disease (CRISP) "Segmentation of Individual Renal Cysts from MR Images in Patients with Autosomal Dominant Polycystic Kidney Disease", *Clin J Am Soc Nephrol* 8: 1089–1097, doi: 10.2215/CJN.10561012, 2013.

[20] M. P. Pawar, A. N. Mulla "Design and Analysis Performance of Kidney Cyst Detection from Ultrasound Images", *IJEDR* | Volume 5, Issue 4 | ISSN: 2321-9939, 2017.

ISBN 979-8-3503-5094-4

Enhanced Lung Tumor Detection via Deep Learning Techniques in CT Imaging

S Rubin Bose [1] *
Department of ECE
SRM Institute of Science and Technology, Ramapuram, Chennai, India.
Email: rubinbos@srmist.edu.in

V Karrthik Kishore[2], K R Abishek[3], J Chalwin Ajay[4]
[2,3,4] Department of CSE
SRM Institute of Science and Technology, Ramapuram, Chennai, India.
Email: vk5619@srmist.edu.in

Angelin Jeba. J [5]
Department of ECE
CEG Campus,
Anna University,
Chennai, India.
Email: jebaangelin@gmail.com

Regin R [6]
Department of CSE
SRM Institute of Science and Technology, Ramapuram, Chennai, India
Email: reginr@srmist.edu.in

***Abstract*— Lung cancer is one of the most common causes of cancer-related death worldwide. Early detection is essential for better patient outcomes. Both the YOLOv8 algorithm and Convolutional Neural Networks (CNNs) have demonstrated promise in the field of medical image processing and object detection, respectively. In this research, we present an innovative method for the detection of lung cancer utilising a CNN model and the YOLOv8 algorithm, integrated into a mobile application developed in Kotlin for enhanced accessibility. Our system takes advantage of the YOLOv8 algorithm's real-time object detection capabilities to recognise nodules of lung cancer using CT scans. The CNN model is trained on a dataset of CT scans and is capable of differentiating between benign and malignant nodules. The mobile application provides a user-friendly interface for uploading CT scans and receiving real-time diagnostic results. The proposed system intends to improve lung cancer early detection by providing a convenient and efficient tool for healthcare professionals and patients. By utilising the strength of CNNs and YOLOv8, this technology has the capacity to reduce false negatives and improve overall diagnostic accuracy, ultimately contributing to better patient care and outcomes. Our model performs well overall, with 98.06% precision, 96.57% recall and 97.31% F1 score when test on a benchmark dataset of images of lung cancer.**

Keywords*— *Convolutional Neural Networks, YOLOv8 model, false negatives, Evaluation, Web-based application.

I. INTRODUCTION

Lung cancer, a prevalent and concerning health issue, has emerged as one of the main global causes of death from cancer. Its insidious nature and the increasing incidence necessitate the development of innovative diagnostic tools to facilitate early identification and better results for patients. In this research endeavour, we describe a unique method for detecting lung cancer by combining the capabilities of Convolutional Neural Networks (CNNs) with deep learning into an intuitive smartphone application.

Early identification of lung cancer is crucial for an effective course of therapy and higher patient survival rates. Conventional diagnostic methods, such as CT scans, often require interpretation by experienced radiologists, a process that can be laborious and prone to mistakes made by people. The proposed mobile application addresses these limitations by providing a real-time, automated diagnostic tool accessible to both healthcare professionals and patients.

Leveraging developments in deep learning and CNNs, the application effectively analyses medical images and identifies lung cancer nodules with high accuracy. The application's user-friendly interface enables seamless uploading of CT scans, followed by prompt display of detection results, giving people the tools they need to take charge of their healthcare.

The integration of deep learning technology into a mobile application represents a significant step forward in lung cancer detection, offering a convenient and accessible solution for early diagnosis and improved patient care. This innovative approach has the ability to completely change how we treat lung cancer, leading to better treatment outcomes and ultimately saving lives.

II. LITERATURE REVIEW

Ausawalaithong et al. [1] analyzed a large dataset utilizing CNN to analyze chest x-rays to identify anomalies. Taking advantage of three modified models and a variety of datasets, the authors assessed the models' accuracy, specificity, and sensitivity. Lung nodules were detected by Model A based on the ChestX-ray14 data collection.

Despite having a lower standard deviation across all evaluation criteria, Model C was able to correctly identify lung cancer and achieved this by utilizing JSRT in addition to ChestX-ray14. Model B outperformed Model C in terms of accuracy and sensitivity, but it had a greater specificity when compared using the Japanese Society of Radiological Technology dataset. They suggested that the model be repeatedly trained for specific objectives. Model C surpasses Retrained Model B in nearly all metrics and can resolve the short dataset issue, since Retrained Model B produced poor results while Model C correctly diagnosed the cancer site.

Bhandary et al. [2] examined the use of a modified AlexNet deep learning technique, referred to as MAN, for identifying lung abnormalities like pneumonia and cancer. To achieve reliable results, the researchers employed two distinct datasets: LIDCIDR and Chest X-Ray. Notably, the MAN-SVM technique achieved the highest accuracy (96.80%) when evaluated on the same dataset used for the original AlexNet experiment. This performance surpassed that of other techniques, even those incorporating modifications to the final stage of their deep learning architecture. This finding was further supported by a similar deep learning architecture exceeding an accuracy of 97.27%.

Da Silva et al. [3] employed a CNN architecture optimized using the PSO (Particle Swarm Optimization) technique. This

ISBN 979-8-3503-5094-4

approach facilitated a direct comparison between different configurations as both training and testing were performed on the same datasets. The researchers utilized data from the LIDC-IDRI dataset and divided it into five test subgroups. Test-4 emerged as the best performing subset, achieving an accuracy exceeding 97.62%, sensitivity exceeding 92.20%, specificity exceeding 98.64%, and an AUC exceeding 0.955.

Naqi et al. [4] proposed a method combining Softmax and Stacked Autoencoder to classify lung nodules, incorporating both 2D and 3D image data for enhanced results. Their focus was on feature reduction and nodule classification using deep learning techniques. The researchers conducted their experiment with the publicly available LIDC-IDRI dataset. Key performance metrics used to evaluate the study included sensitivity, specificity, accuracy, and false positives per scan. Their dataset included 888 CT scans having 777 nodules (≥ 3 mm) independently identified by skilled radiologists. Their proposed technique achieved promising results: sensitivity exceeding 95.6%, accuracy exceeding 96.9%, and specificity exceeding 97.0%.

Shaffie et al. [5] presented a novel approach using deep autoencoders for the automated and non-invasive diagnosis of lung tumor, aiming to classify between benign and malignant nodules. Their system, trained on the LIDC-IDRI data collection, yielded promising results. The proposed model demonstrated potential for aiding early lung cancer detection, achieving an accuracy exceeding 91.20%, sensitivity exceeding 95.88%, specificity exceeding 85.03%, and an AUC exceeding 95.73 on a dataset of 727 nodules across 467 cases.

Kaur et al. [6] proposed a CNN architecture for medical image analysis. The design employs an alternating modules of CNN layers and Rectified Linear Unit layers, culminating in a dense layer. To extract representative features, each convolutional layer utilizes 64 filters. The researchers trained and validated their approach by employing the Japanese Society of Radiological Technology database. The model achieved impressive performance, yielding average scores of accuracy exceeding 98.05%, overlap exceeding 96.25%, specificity exceeding 98.80%, and sensitivity exceeding 98.05%.

Xie et al. [7] introduced a deep learning architecture, named the multi-view knowledge-based collaborative, specifically designed to classify lung nodules as benign or malignant within chest CT scans. To evaluate their method, the researchers utilized the LIDC-IDRI dataset, benchmarking its performance against five other state-of-the-art classification techniques. Their findings suggest potential for clinical application, as the approach yielded a classification accuracy exceeding 91.60% and an AUC exceeding 95.70%.

Nibali et al. [8] investigated the use of ResNet principles for the classification of lung nodules. Their dataset was derived from the public LIDC-IDRI dataset, consisting of 1010 patient CT scans. Images were analyzed by four radiologists to generate random nodule classifications. The study reported two key findings: first, system accuracy improved with a combination of deeper networks, curriculum learning, and transfer learning; second, a systematic comparison against other models under identical conditions demonstrated superior performance of their method across sensitivity exceeding 91.07%, specificity exceeding 88.64%, precision exceeding 89.35%, AUROC exceeding 0.9459, and accuracy exceeding 89.90%.

Zhang et al. [9] investigated the use of a deep belief network (DBN) for lung nodule detection, utilizing LIDC-IDRI dataset. Their approach demonstrates high performance, particularly for larger nodules exceeding 30 mm in diameter. Notably, the achieved accuracy, sensitivity, and specificity all exceeded 90%. This performance surpasses the range typically reported in the literature, where accuracy falls between 89.0% and 89.5%, and sensitivity ranges from 84.2% to 87.1%.

Causey et al. [10] developed NoduleX, a system designed to assess the cancerous lung nodules in CT images using CNN. They trained and validated their model using a subset of nodules extracted from the LIDC-IDRI dataset. Their study demonstrated NoduleX's high effectiveness; on an independent validation test, it achieved a notable AUC exceeding 0.99, accompanied by accuracy exceeding 94.6%, sensitivity exceeding 94.8%, and specificity exceeding 94.3%.

The suggested system is a web-based program that analyzes CT scan data to identify pneumonia and lung cancer. The system utilizes computer vision and deep learning techniques to categorize images as cancerous, indicative of pneumonia, or normal. Following the analysis, the system offers users a diagnosis along with recommendations for further treatment. With the capability to significantly enhance early detection and diagnosis, the proposed system holds the potential to save lives. Designed with user-friendliness in mind, the system allows users to effortlessly upload their CT scan images, which are then automatically analyzed to provide a diagnosis. This innovative approach could revolutionize the diagnostic process for lung tumor and pneumonia, contributing to early detection and treatment, thereby improving outcomes and saving lives.

III. METHODOLOGY

Early diagnosis of lung cancer is critical in enhancing the outcome of patients and reducing mortality rates. In this project, we proposed a lung cancer detection system using YOLOv8, cutting-edge object detection technique, integrated into an Android app. The system utilizes CT scans as input images and can effectively detect pneumonia, a common type of lung cancer.

YOLOv8 is the latest model of object detection algorithms. It offers significant improvements over previous versions in terms of speed and accuracy.

A. YOLOv8 Architecture

YOLOv8's architecture consists of three main modules:

Backbone

The backbone module takes the input image and extracts high-level characteristics. The foundation of YOLOv8 is

ISBN 979-8-3503-5094-4

CSPDarknet53, a CNN architecture created to solve the disappearing gradient issue and improve training effectiveness.

Neck

The neck module integrates the feature maps extracted by the backbone at different scales, enabling the model to capture both fine-grained and coarse details of the image. YOLOv8 utilizes Path Aggregation Network (PANet) as the neck, a multi-scale feature fusion architecture that improves object detection performance.

Head

For every object recognised, the head module creates bounding boxes and estimates the class probabilities. YOLOv8 employs a modified version of the YOLOv8 head, incorporating anchor boxes and logistic regression for efficient object detection and classification.

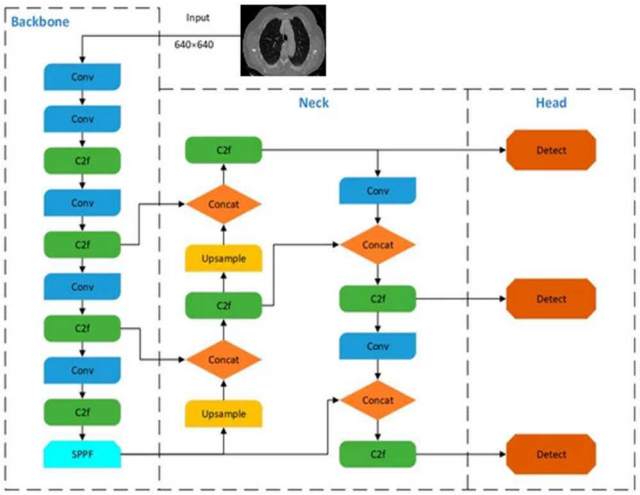


Fig. 1. YOLOv8 architecture

To leverage existing knowledge and improve training efficiency, this study employed transfer learning. This technique involved reusing pre-trained weights from a previously trained model and fine-tuning them on the lung cancer dataset. Compared to traditional training, this offers faster training times and requires fewer resources. Using the loss function, the ratio of true and predicted values is computed on each iteration.

$$L_T = L_{clss} + L_{cnf} + L_{box} \quad (1)$$

where L_T is the absolute loss, L_{clss} is expressed in equation 2 and shown as a classification loss. L_{cnf} is represented by the term "confidence loss" in equation 3, and L_{box} is the bounding box loss.

$$L_{clss} = \sum_{i=0}^{x^2} l_i^{obj} \sum_{j=0}^{R} \left[\left(Pi(c) - \widehat{Pi}\,(c) \right)^2 \right] \quad (2)$$

$$L_{cnf} = \sum_{i=0}^{x^2} \sum_{j=0}^{R} l_i^{obj} \left[\left(Ci - \widehat{Ci} \right)^2 \right] + \beta_{noobj} \sum_{i=0}^{x^2} \sum_{j=0}^{R} l_i^{noobj} \left[\left(Ci - \widehat{Ci} \right)^2 \right] \quad (3)$$

where, $Pi(c)$ is represented as the probability of being an object. l_i^{obj} and l_i^{noobj} are denoted as the indicator function. Ci is referred to as objectness.

The Precision, Recall, F1-score, and Prediction time are among the characteristics that are used to assess the YOLOv8 model's performance. Equations 4, 5, and 6 express the different parameter.

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Positive} \quad (4)$$

$$Recall = \frac{True\ Positive}{True\ Positive + False\ Negative} \quad (5)$$

$$F1 - Score = 2.\frac{Precision\ .Recall}{Precision + Recall} \quad (6)$$

IV. EXPERIMENTAL SETUP

A. Data Preprocessing

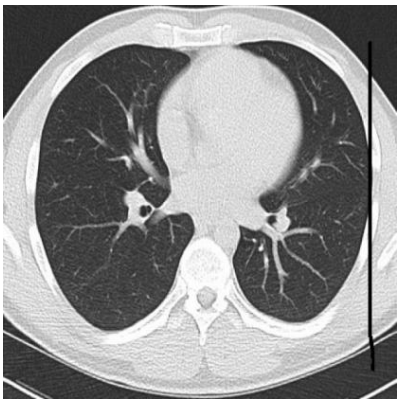
Data Collection

The CT scan images used in this lung cancer dataset were gathered from the openly accessible "Kaggle" web resource [11]. The dataset source states that all of the labels on the photos were manually checked after they were gathered from different sources. To ensure compatibility with the model, the dataset stores images in JPG or PNG format, deviating from the standard DICOM format used in medical imaging. Each CT scan image within the dataset, totalling 6,984, possesses a resolution of 640 x 640 pixels. To diagnose lung cancer and pneumonia, the dataset is split into five categories: pneumonia, squamous cell carcinoma, large cell carcinoma, adenocarcinoma, and normal (not lung cancer).

Fig. 2. Cancerous sample of Lung CT scan

ISBN 979-8-3503-5094-4

Fig. 3. Pneumonia sample of Lung CT scan



(a)

(b)

Fig. 4. Normal sample of Lung CT scan

Data Pre-processing

To get the CT images ready for CNN model training, preprocess them. This could entail cropping, rotating, and flipping the photos to a standard size, normalising the intensity levels, and adding techniques like flipping and cropping to the data.

B. Model Training

Model Architecture

Design and implement a CNN model using the YOLOv8 architecture. The model should be able to recognise cancer nodules by extracting characteristics through CT images.

Model Training

Utilising the pre-processed dataset, train the CNN model. This could entail determining the training settings, keeping a check on the training procedure, and choosing a suitable optimizer and loss function.

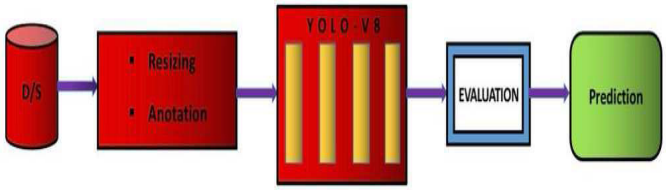


Fig. 5. Training

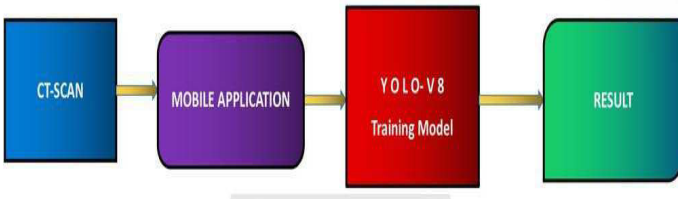


Fig. 6. Detection

C. Model Evaluation

Evaluation Dataset

Divide the dataset into three separate subsets: 1. Training set, 2. Validation set, and 3. Test set. The first subset serves as the foundation for the model's learning, enabling it to identify patterns and characteristics within the data. The second subset is utilized for refinement, guiding adjustments to the model's parameters to improve its effectiveness. Finally, the third subset remains untouched, acting as an independent benchmark to assess the model's generalizability.

Evaluation Metrics

Assess the model's performance on the test set by utilizing metrics such as overall precision, overall recall, and overall F1 score. These metrics serve as indicators of the model's accuracy in recognising cancerous nodules.

D. Mobile Application

Mobile App Development

Develop a mobile application using Kotlin to provide a user-friendly interface for uploading CT scans and receiving real-time diagnostic results.

Model Integration

Incorporate the mobile application with the trained CNN model. This will allow the application to process CT scans and identify lung cancer nodules.

Maintaining Image Quality

To ensure reliable diagnoses despite potential variations in mobile image quality, the application utilizes a multifaceted

ISBN 979-8-3503-5094-4

approach. Advanced compression algorithms prioritize efficient transmission and storage while safeguarding critical details for accurate nodule detection. Built-in quality checks act as a safeguard against inaccurate diagnoses by identifying low-resolution or unclear images. Real-time image enhancement techniques further enhance analysis by adjusting contrast, brightness, and sharpness, even for suboptimal images. Finally, user guidance empowers users to capture high-quality scans through tips on proper lighting, minimizing motion blur, and maintaining camera stability. This comprehensive approach effectively addresses the challenges of varying mobile image quality, ultimately contributing to more accurate diagnoses.

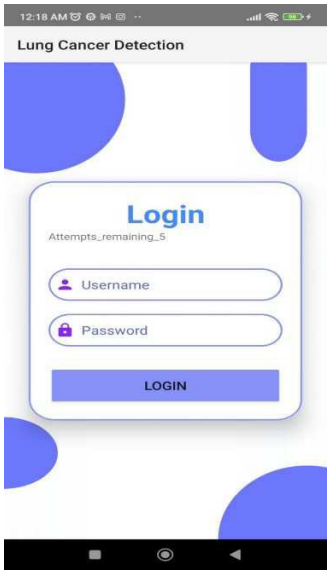


Fig. 7. Login page

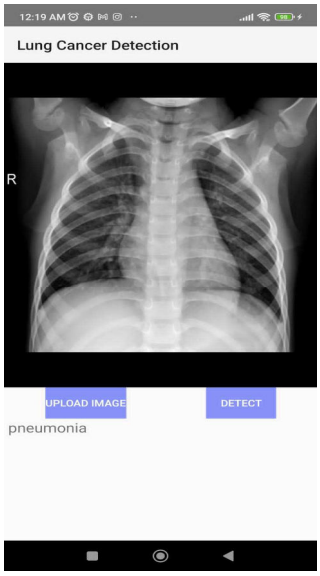


Fig. 8. Prediction

V. RESULTS AND ANALYSIS

The YOLOv8 model's overall accuracy, overall precision, and overall F1 score at various epochs are displayed in Table I. The percentage of images that the model properly classifies as malignant or non-cancerous is called as the overall accuracy. The proportion of images that the model correctly classifies as malignant is called as the overall precision. The overall performance is measured by the F1 score, which represents the balanced mean of precision and recall.

Table I demonstrates improved performance in YOLOv8's object detection capabilities with an increase in training epochs. The model's performance demonstrates its ability to learn the distinctive features of lung cancer images, progressively refining its performance over time.

TABLE I
PERFORMANCE OF YOLOV8

Epochs	Overall Recall	Overall Precision	Overall F1 Score
10	0.95	0.93	0.94
50	0.95	0.97	0.96
100	0.96	0.98	0.97

At 100 epochs, the YOLOv8 model achieves an overall accuracy of 96%. This signifies that the model accurately identifies 96% of the lung cancer images in the test set. The overall precision of the YOLOv8 model at 100 epochs is 98%. This means that 98% of the images that the model identifies as cancerous are actually cancerous. The overall F1 score of the YOLOv8 model at 100 epochs is 97%. The model's performance demonstrates a commendable harmony among accuracy and precision. The YOLOv8 model is a promising method for lung cancer diagnosis, as demonstrated by Table I. On a test set for lung cancer, the model can achieve excellent accuracy and precision. With further refinement, the YOLOv8 model could be used to develop a clinical tool for lung cancer detection.

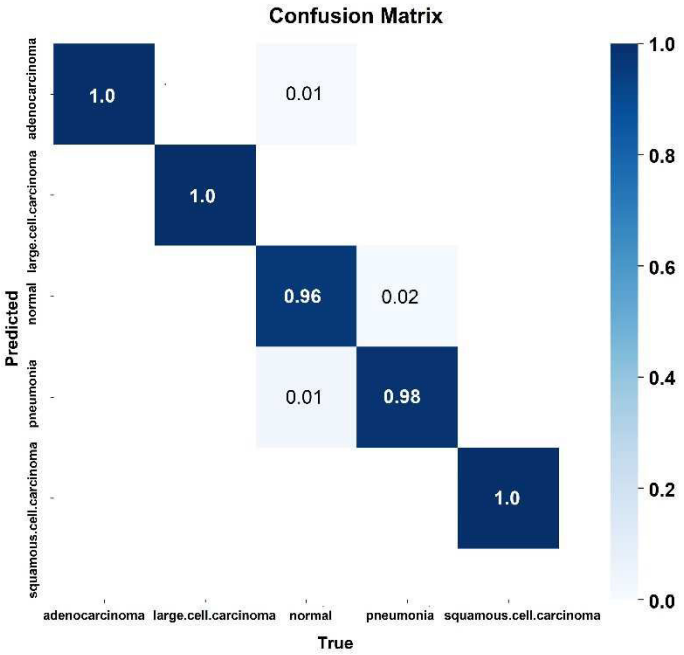


Fig. 9. Confusion Matrix Normalized

During training, the model's performance is evaluated using two key metrics: training loss and validation loss. Training loss reflects the model's ability to learn from the provided data, while validation loss assesses how well the model adapts to unseen data. Training accuracy represents the percentage of correctly classified images within the training set, while

ISBN 979-8-3503-5094-4

validation accuracy signifies the percentage of accurately classified images within the validation set.

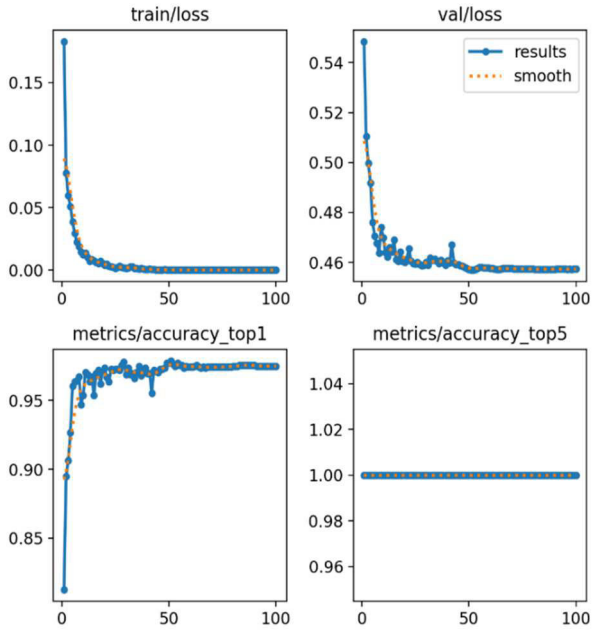


Fig. 10. Graphical representation

The training loss and validation loss diminish over the course of training, as the graph illustrates. This shows that the model can pick up on the characteristics of the training set and get better over time. Additionally, the graph demonstrates how training accuracy and validation accuracy rise with training. This implies that the model can accurately categorize lung cancer photos and learn to generalize to new data.

The graph reveals the YOLOv8 model's learning effectiveness, achieving 99% accuracy on the training data and 96% on unseen data after 100 epochs. This suggests the model has effectively learned to identify lung tumor images with greater precision.

Additionally, the graph demonstrates how the training loss and validation loss curves converge as training progresses. This implies that the training data are not being overfitted by the model.

The YOLOv8 model demonstrates promise as a lung cancer detection method, achieving high accuracy on both training and validation datasets. With further refinement, it has the potential to be developed into a valuable clinical tool for lung cancer identification.

VI. CONCLUSION

The YOLOv8 CNN model proposed for lung cancer detection exhibits outstanding performance in discerning lung cancer nodules from CT scans. It attains an impressive overall precision of 98.06%, an overall recall of 96.57%, and an overall F1 score of 97.31%. Integration of YOLOv8 into a mobile application further enhances the accessibility and practicality of this diagnostic tool, enabling healthcare professionals and patients to conveniently upload CT scans and receive real-time diagnostic results. This user-friendly interface empowers individuals to take proactive measures in their healthcare journey and facilitates timely intervention for improved treatment outcomes. The YOLOv8 CNN model, underwent training using an extensive dataset encompassing images of adenocarcinoma, large cell carcinoma, squamous cell carcinoma, pneumonia, and normal lung tissue. The inclusion of such diversity in the training data is crucial as it enhances the model's generalizability and equips it with the capability to accurately classify nodules across a broad spectrum of lung pathologies. The suggested YOLOv8 CNN model, which provides a reliable, accurate, and approachable option for early diagnosis, is a substantial development in the detection of lung cancer. Its potential to transform lung cancer screening and management is further increased by its incorporation into a mobile application, which will ultimately lead to better patient care and higher survival rates.

REFERENCES

[1] W. Ausawalaithong, A. Thirach, S. Marukatat and T. Wilaiprasitporn, "Automatic Lung Cancer Prediction from Chest X-ray Images Using the Deep Learning Approach," 2018 11th Biomedical Engineering International Conference (BMEiCON), Chiang Mai, Thailand, 2018, pp. 1-5, doi: 10.1109/BMEiCON.2018.8609997.

[2] Abhir Bhandary, G. Ananth Prabhu, V. Rajinikanth, K. Palani Thanaraj, Suresh Chandra Satapathy, David E. Robbins, Charles Shasky, Yu-Dong Zhang, João Manuel R.S. Tavares, N. Sri Madhava Raja, Deep-learning framework to detect lung abnormality – A study with chest X-Ray and lung CT scan images, Pattern Recognition Letters, Volume 129,2020,Pages 271-278,ISSN 0167-8655,https://doi.org/10.1016/j.patrec.2019.11.013.

[3] Silva, Giovanni & paulo da silva neto, Otilio & Silva, Ari & Paiva, Anselmo & Gattass, Marcelo. (2017). Lung nodules diagnosis based on evolutionary convolutional neural network. Multimedia Tools and Applications. 76. 10.1007/s11042-017-4480-9.

[4] Naqi, Muhammad & Sharif, Muhammad & Jaffar, Arfan. (2020). Lung nodule detection and classification based on geometric fit in parametric form and deep learning. Neural Computing and Applications. 32. 10.1007/s00521-018-3773-x.

[5] Shaffie, Ahmed & Soliman, Ahmed & Fraiwan, Luay & Ghazal, Mohammed & Taher, Dr. Fatma & Dunlap, Neal & Wang, Brian & Berkel, Victor & Keynton, Robert & Elmaghraby, Adel & El-Baz, Ayman. (2018). A Generalized Deep Learning-Based Diagnostic System for Early Diagnosis of Various Types of Pulmonary Nodules. Technology in Cancer Research & Treatment. 17. 153303381879880. 10.1177/1533033818798800.

[6] Kaur, Simranpreet & Hooda, Rahul & Mittal, Ajay & Akashdeep, & Sofat, Sanjeev. (2017). Deep CNN-Based Method for Segmenting Lung Fields in Digital Chest Radiographs. 10.1007/978-981-10-5780-9_17.

[7] Yutong Xie, Jianpeng Zhang, Yong Xia, Michael Fulham, Yanning Zhang,Fusing texture, shape and deep model-learned information at decision level for automated classification of lung nodules on chest CT, Information Fusion,Volume 42, 2018, Pages 102-110, ISSN 1566-2535, https://doi.org/10.1016/j.inffus.2017.10.005.

[8] Nibali A, He Z, Wollersheim D. Pulmonary nodule classification with deep residual networks. Int J Comput Assist Radiol Surg. 2017 Oct;12(10):1799-1808. doi: 10.1007/s11548-017-1605-6. Epub 2017 May 13. PMID: 28501942.

[9] Ting Zhang, Juanjuan Zhao, Jiaying Luo, and Yan Qiang. Deep Belief Network for Lung Nodules Diagnosed in CT Imaging [J]. Int J Performability Eng, 2017, 13(8): 1358-1370.

[10] Causey JL, Zhang J, Ma S, Jiang B, Qualls JA, Politte DG, Prior F, Zhang S, Huang X. Highly accurate model for prediction of lung nodule malignancy with CT scans. Sci Rep. 2018 Jun 18;8(1):9286. doi: 10.1038/s41598-018-27569-w. PMID: 29915334; PMCID: PMC6006355.

[11] www.kaggle.com

ISBN 979-8-3503-5094-4

Efficient Skin Cancer Diagnosis and Classification via High-Speed Deep Learning Architecture

S Rubin Bose [1] *
Department of ECE
SRM Institute of Science and Technology, Ramapuram, Chennai, India.
Email: rubinbos@srmist.edu.in

M Izzath Suhail[2], Shaik Roshni Shabnam[3], B Hariharan[4]
[2,3,4] Department of CSE
SRM Institute of Science and Technology, Ramapuram, Chennai, India.
Email: mi9928@srmist.edu.in

Angelin Jeba. J [5]
Department of ECE
CEG Campus, Anna University, Chennai, India.
Email: jebaangelin@gmail.com

Regin R [6]
Department of ECE
SRM Institute of Science and Technology, Ramapuram, Chennai, India.
Email: reginr@srmist.edu.in

***Abstract*—Skin cancer is still a major worldwide health issue, which highlights the necessity for quick and accurate detection methods. This research introduces a robust skin cancer classification system leveraging the capabilities of YOLOv8, An advanced algorithm for object detection. The proposed model processes dermatoscopic images with remarkable speed and accuracy, facilitating the identification of malignant lesions. The YOLOv8 architecture enables classification of various skin cancer types, including melanoma and non-melanoma, by effectively localizing and classifying lesions within the images. A comprehensive dataset, comprising diverse skin lesions, was employed for training and validation, ensuring the model's adaptability to different clinical scenarios. The evaluation of the proposed model demonstrates superior performance compared to traditional methods, exhibiting high sensitivity and specificity.**

I. INTRODUCTION

Skin cancer, A formidable adversary in the realm of dermatological disorders, poses a significant global health challenge. With escalating incidence rates and diverse manifestations, the imperative for accurate and timely diagnosis has never been more critical. Melanoma and non-melanoma skin cancers, representing distinct entities with varying degrees of aggressiveness, underscore the need for nuanced diagnostic methodologies. The analysis of the suggested model shows that it performs better than the conventional approaches.

The capacity to distinguish between lesions that are melanoma and those that are not is of utmost significance because treatment plans and outcomes are closely associated with the particular kind of skin cancer.

By training YOLOv8 on diverse datasets encompassing melanoma and non-melanoma cases, we aim to empower this algorithm to autonomously discern between these critical distinctions. The real-time processing capabilities of YOLOv8 not only expedite the diagnostic process but also hold the potential to enhance accuracy, presenting a paradigm shift in skin cancer classification methodologies.

This paper delves into the vital realm of skin cancer classification, where the distinction between melanoma and non-melanoma is a crucial determinant in patient care. Through the lens of YOLOv8, a powerful tool in the arsenal of artificial intelligence, we navigate the intricacies of dermatological diagnostics.

II. LITERATURE REVIEW

Xie et al. [1] Introduced a ground-breaking method for classifying melanocytic tumors based on digital dermoscopy images. Their three-step algorithm, leveraging a self-generating neural network and novel border features, demonstrated enhanced accuracy in identifying lesions that are benign or cancerous. The research culminated in a hybrid network ensemble classifier, combining back propagation and fuzzy neural networks, showcasing significant improvements in classification accuracy. This innovative approach holds promise for advancing computer-aided diagnosis in dermatological applications.

Khan et al. [2] Enhanced K-mean clustering and a Gaussian filter to lesion segmentation. The originality of the suggested method is enhanced by the special hybrid super feature vector, which combines textural and color data. The study's contribution lies in its comprehensive evaluation, utilizing the DERMIS dataset and improving the accuracy, thereby developing the diagnosis of skin cancer field.

Bisla et al. [3] Addresses the challenges posed by limited and imbalanced skin lesion databases. Melanoma's critical importance in early detection for survival underscores the significance of their work. Their system utilizes a processing unit for occlusion removal and a generative adversarial network-based data generation unit to address scarcity in lesion classes.

Pham et al. [4] Authors study conducts a comprehensive comparative analysis of six classifiers, using four data pre-treatment processes and seven feature extraction algorithms on the largest skin cancer datasets. Color, texture, and shape traits are important in classifying melanoma, according to the findings. The HAM10000 dataset shows that this configuration can improve skin cancer prediction.

Rezaoana et al. [5] Presents an automated system that uses deep learning and image processing, more especially, a parallel CNN model. Using picture augmentation techniques to enhance the dataset, the research aims to classify nine different forms of skin lesions. Different types of processing principles were combined in this methodology to improve skin cancer diagnosis and classification accuracy. With an accuracy percentage of 79.45%, the suggested CNN model shows encouraging results.

Javaid et al. [6] Suggested a method Integrating machine learning and image processing to classify and segment skin lesions. Their unique method involves extracting features such as GLCM, HOG, and color identification features from dermoscopic images, segmenting them using OTSU thresholding, and stretching contrast from those images. Prior

ISBN 979-8-3503-5094-4

to classification utilizing SVM, Quadratic Discriminant, and Random Forest classifiers, PCA reduction, SMOTE sampling for class imbalance, and a wrapper-based feature selection approach are performed.

Subramanian et al. [7] Highlights the importance of the early identification of skin lesions and its classification, particularly in cases involving melanoma and focal cell carcinoma. Objectives include achieving >80% accuracy, maintaining a false negativity rate <10%, and precision >80%. Leveraging image processing and machine vision, the proposed method demonstrates superiority over comparable approaches in simulation results, showcasing its potential for robust skin cancer detection and classification

Jusman et al. [8] Focuses on skin cancer classification, a critical task for early detection and treatment. Utilizing a large dataset (HAM10000), They train VGG-16, a customized convolutional neural network, and Multi-layer Perceptron for classification, contrasting their results of computing time. Findings indicate that VGG-16 achieves the highest classification accuracy, while both VGG-16 and the custom CNN demonstrate significantly faster testing times than Multi-layer Perceptron. The study provides a systematic comparison of neural networks for skin cancer classification, offering valuable insights for improved detection.

Pacheco et al. [9] Presented a novel use of Deep neural networks in the categorization of skin lesions. The authors suggest the Meta Block, a method that includes patient demographics as metadata to improve pertinent information collected from photos in the classification process. Traditionally, these systems just rely on photographs.. Comparative analysis with other combination including MetaNet and feature concatenation, reveals the superior performance of MetaBlock. Results from two skin lesion datasets indicate improved classification across tested models, outperforming alternative methods in 6 out of 10 scenarios

Huang et al. [10] Examined the possibilities for the categorization of skin lesions using hyperspectral imaging a noninvasive optical method. The YOLOv5 deep learning model is used for training once the dataset is divided into training and test sets. Two models are constructed and compared: RGB classification and hyperspectral narrowband imaging (HSI-NBI)..

Manoj et al. [11] Uses the Deep Convolutional Neural Network (DCNN) architecture of YOLO-v3 to automatically detect and classify skin lesions. A feature map is produced by YOLO-v3, and color features are retrieved using quad histogram color moments. The DCNN is then fed the fused features in order to classify the various forms of skin cancer.

Albahli et al. [12] Addresses the critical issue of melanoma detection. The authors present an approach that combines YOLOv4-DarkNet, an object detector tuned for melanoma detection, and active contour segmentation. Evaluation on ISIC2018 and ISIC2016 datasets demonstrates the method's superiority over state-of-the-art techniques, highlighting its potential for developing a Melanoma diagnosis using a clinical decision support system. The YOLO v4 helps to classify skin lesions.

Based on the diverse literature review in skin cancer classification methodologies, several studies have highlighted the advantages of incorporating YOLO (You Only Look Once) architectures. Notably, authors in [10], [11] and [12], consistently recommend the YOLO architecture for real-time object detection within the field of analysis of skin lesions. In this paper, the latest version YOLOv8 is utilized to categorize skin cancer. The YOLOv8 is trained and evaluated using various skin cancer Data sets. The performance of YOLOv8 model is also analysed.

III. METHODOLOGY

The suggested approach began with gathering a variety of dermoscopic skin lesion photographs from reliable sources, including Dermnet NZ and from HAM10000 database. Thorough labelling process was used to accurately reflect the many forms of skin cancer in each photograph.

A common resolution of 640x640 pixels was applied to all photos in order to maintain consistency and maximize computing efficiency. Moreover, the dataset was expanded and augmented including Rotation, Flipping, and Zooming, which encouraged robustness in managing changes in image orientation and scale. Then, to keep the dispersion evenly distributed of data throughout the sets, the supplemented dataset was carefully divided into sets of tests, training, and validation. The annotation of these dataset makes supervised training of the YOLOv8 model easier.

The model configuration involves selecting YOLOv8 as the base architecture, initializing with pretrained weights, and tuning hyperparameters for optimal performance. Transfer learning is leveraged to adapt the model to the specific characteristics of skin lesions Metrics including recall, loss, accuracy, and F1-score are tracked throughout training to make sure learning is happening effectively.

The methodology concludes with a concise summary, highlighting YOLOv8's contributions to automated skin lesion analysis and suggesting avenues for future research. This comprehensive approach underscores the potential of YOLOv8 in advancing computer-aided diagnosis in dermatology.

A. YOLOvs8 Architecture

The major parts of YOLOV8 architecture are:

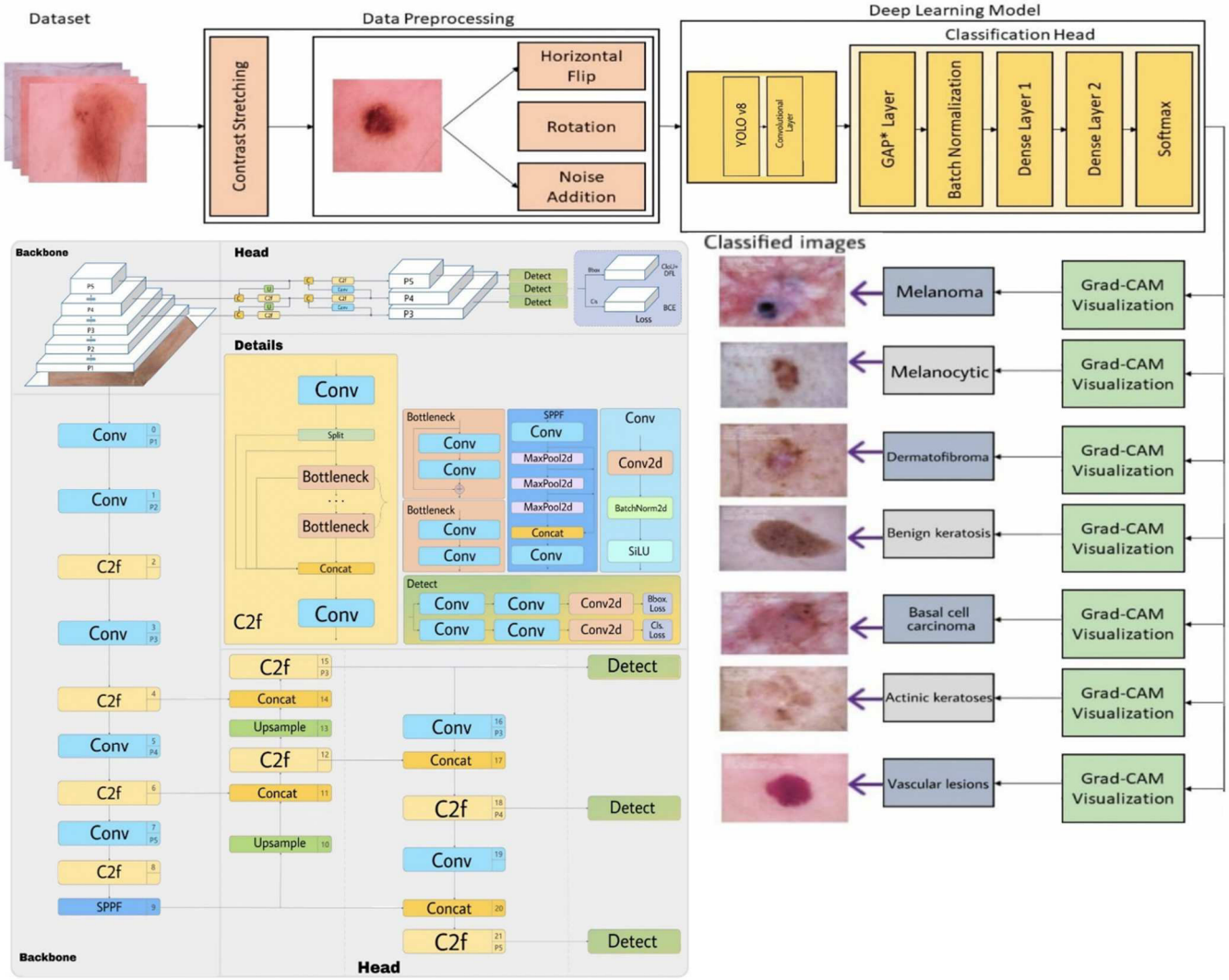
Backbone: Consisting of a convolutional neural network with an altered CSPDarknet53 architecture, it is a YOLO algorithm innovation at its core. In order to maximize information flow across layers, this design consists of 53 convolutional layers with cross-stage partial connections..

Head: The head is made up of completely linked layers that come after several convolutional layers. Its main function is to forecast class probabilities, bounding boxes, and objectness scores so that thorough and precise object detection is possible.

Neck: Serving as a vital link between the backbone and head networks, the Neck Network in YOLOv8 is instrumental in resizing the feature map to enhance feature resolution.

ISBN 979-8-3503-5094-4

Fig. 1. YOLOvs8 Architecture diagram & Functional block diagram of Skin Cancer Classification



Minimum enclosing boxes, objectness scores, and classes of objects recognised are the three outputs produced by YOLOv8.To expedite training on new data without retraining the entire network, transfer learning is employed—a method that enhances efficiency and conserves resources. Throughout the iteration, the loss function calculates the ratio between predicted and true values:

$L_T = L_{clss} + L_{cnf}$ ..... . (1)

Let L_T denote the total loss, L_{clss} as the classification loss represented in equation 2, L_{cnf} denoted as confidence loss expressed in equation 3.

$$L_{cls} = \sum_{i=0}^{x^2} l_i^{obj} \sum_{j=0}^{R} \left[\left(Pi(c) - \hat{P}\imath\,(c) \right)^2 \right] \text{.....(2)}$$

$$L_{cnf} = \sum_{i=0}^{x^2} \sum_{j=0}^{R} l_i^{obj} \left[\left(Ci - \hat{C}\imath \right)^2 \right] + \beta_{noobj} \sum_{i=0}^{x^2} \sum_{j=0}^{R} l_i^{noobj} \left[\left(Ci - \hat{C}\imath \right)^2 \right] \text{.....(3)}$$

Where $Pi(c)$ is defined as the probability of being an object, l_i^{obj} and l_i^{noobj} are indicated by the indicator function, and Ci represents the objectness. The YOLOv8s model's performance is evaluated using several parameters, as expressed in equations 4, 5, and 6:

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Positive} \quad \text{.... (4)}$$

$$Recall = \frac{True\ Positive}{True\ Positive + False\ Negative} \quad \text{.... (5)}$$

$$F1 - Score = 2.\frac{Precision\ .Recall}{Precision + Recall} \quad \text{.... (6)}$$

IV.EXPERIMENTAL SETUP

A. Dataset Preprocessing

The acquired dermoscopic skin lesion images were pre-processed using custom Python scripts. This pre-processing pipeline ensured uniform resizing and augmentation of the dataset, facilitating standardized input dimensions for the neural network model. Strategies for augmenting the data, including Rotation, Flipping, and zooming, are used to improve these dataset's robustness and variety.

ISBN 979-8-3503-5094-4

B. Data Partitioning

Using a stratified method, these datasets are carefully separated into test, validation, and training sets. The implementation of this strategic division facilitated a thorough assessment and validation of the model's functionality across many classes and guaranteeing that each subset had a representative distribution of distinct forms of skin cancer.

C. Training Setup

The model was trained using YOLOv8, the most recent version of the YOLO algorithm, which is superior to its predecessors due to the addition of features including feature fusion, spatial attention, and context aggregation modules The training was conducted over 10, 20, 50 epochs ensuring efficient and effective model learning without excessive computational burden.

D. Model Evaluation

Using a range of important performance indicators, including as The F1-score, area under the receiver operating characteristic curve (AUCROC), accuracy, precision, recall, and recall the trained model's performance was assessed on the test set. To offer a comprehensive comprehension of the efficacy of the model in correctly categorizing various forms of skin malignancies, qualitative assessments were also carried out by visualizing the model's predictions and comparing them with ground truth annotations.

V. RESULTS AND ANALYSIS

YOLOv8 is utilized to build a custom dataset for Training and Testing the skin cancer detection algorithm. The model is trained with AdamW optimization algorithm employing a learning rate of 0.000714 momentum of 0.9, with Parameter groups 26 weight(decay=0.0),27weight(decay=0.0005) image sizes 640 train, 640 val. Transfer learning is applied with pre-trained weights to accelerate the training process. The resultant YOLOv8 model is made to effectively Sort various forms of skin cancer into categories.

The suggested method for Classifying skin cancer. The model is distinguished by its high precision, lightweight design, and time efficiency. Utilizing a YOLOv8 training setup, the model performs classification of skin cancer category of suspicious lesions with a high confidence score.

In order to assess the model's effectiveness, metrics like F1-Score, Average Precision, and Average Recall are utilized, taking into account an Intersection over Union (IoU). The 0.65 threshold. These measures provide a comprehensive evaluation of the model's ability to accurately classify skin cancer from medical photographs.

TABLE I. PERFORMANCE METRICS OBTAINED USING YOLOv8.

CNN model	*Average Precision (%)*	*Average Recall (%)*	*F1-Score (%)*
Yolov8s(10epochs)	69.33	6.14	11.04
Yolov8 (20epochs)	96.67	96.83	96.83
Yolov8 (50epochs)	96.67	96.83	96.83

In Table I, the effectiveness indicators for the YOLOv8 model are detailed, showcasing its efficacy in skin cancer detection. The suggested YOLOv8 model produced an average precision of 69.33%, an average recall of 6.14%, and an F1-score of 11.04% after 10 epochs. Subsequent evaluations at epochs 20 and 50 revealed further improvements in performance metrics, with F1-scores, Average Precision, and Average Recall reaching 96.83, 96.67, and 96.83% for both epochs.

TABLE II. COMPARISON OF THE COMPUTATIONAL TIME ON GOOGLE COLAB

CNN model	Speed (ms)	
	Pre-process	*Inference*
Yolov8 (10 epochs)	0.8	1.4
Yolov8 (20 epochs)	0.6	1.4
Yolov8 (50 epochs)	0.4	1.4

Fig. 2. Confusion Matrix normalized

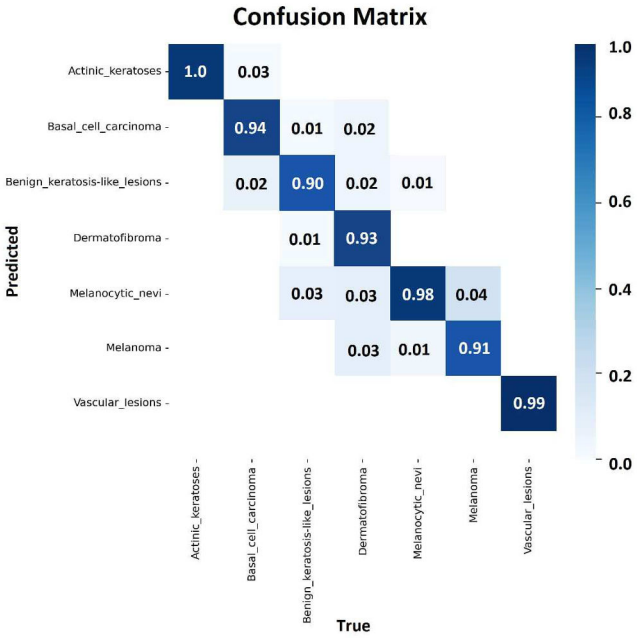


Fig-2. In multi-class classification, precision and recall metrics are derived from the confusion matrix, which organizes the model's predictions across all classes against the true labels. Precision for each class is calculated as the ratio of true positives for that class to the sum of true positives and false positives, while recall is computed as the ratio of true positives to the sum of true positives and false negatives for that class. This approach allows for a comprehensive assessment of the model's performance across multiple classes, providing insights into its ability to accurately classify instances for each class individually. Notably, values in the confusion matrix deemed to have higher confidentiality scores may indicate classes where precision and recall are particularly critical and warrant additional scrutiny.

ISBN 979-8-3503-5094-4

Fig. 3. Classified skin cancer images

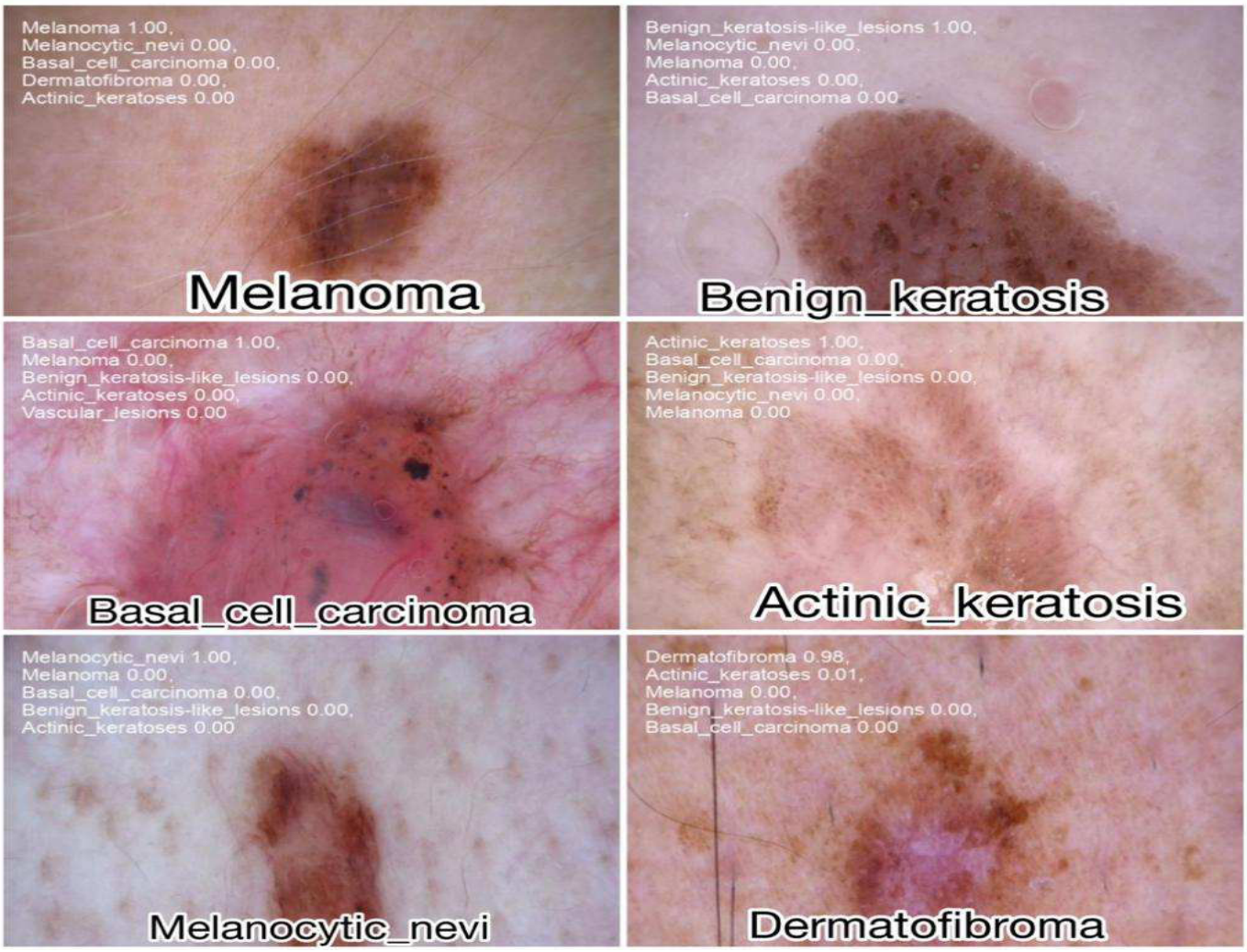


Fig-3. shows A visual representation of skin cancer classification results. Each image is categorized based on the model's predictions, offering insights into its ability to discern among various kinds of skin lesions.

Fig. 5. Graphical representation

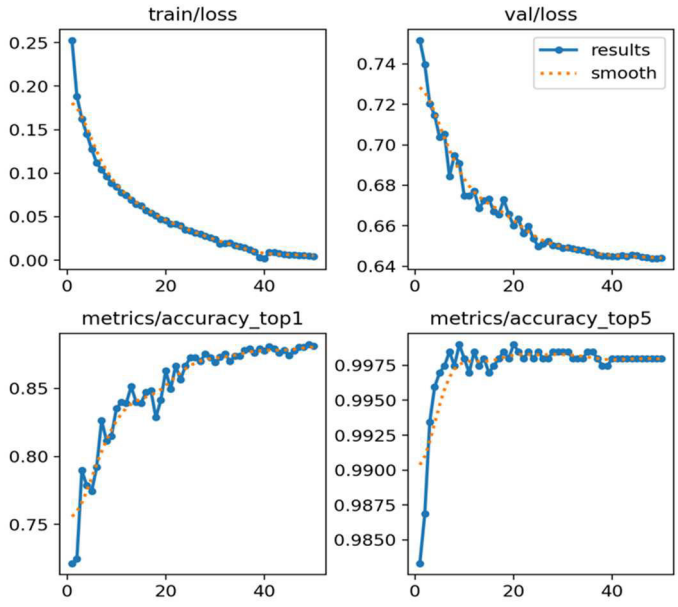


Fig-5. As the graph shows, training results in a decrease in both training loss and validation loss. This demonstrates how the model may learn from the features in the training set and improve with time. The graph also shows that training accuracy and validation accuracy increase with training. This suggests that the model can reliably classify photos of skin cancer and develop the ability to generalise to fresh data. This result suggests that the model has successfully learned to recognise skin cancer images with a high degree of precision.

V. CONCLUSION

The results and analysis of skin cancer Classification using YOLO v8 reveal a promising performance in accurately Classifying lesions. The model demonstrated commendable Precision, Recall, and F1 score, showcasing its efficacy in Classifying these cancerous skin lesions. In comparison with existing methods, YOLO v8 exhibited notable improvements, particularly in its ability to efficiently process and detect lesions. However, the analysis also uncovered instances of false

ISBN 979-8-3503-5094-4

positives and false negatives, emphasizing the importance of refining the model to minimize errors. Sensitivity and specificity metrics provided valuable insights into the model's balanced identification of positive and negative cases. Computational efficiency, a critical factor for real-world applications, was a strength of YOLO v8, displaying favourable processing speeds and resource requirements. The study's dataset, though comprehensive, poses considerations for future research to explore diverse datasets for enhanced generalization. While acknowledging its shortcomings, the analysis paves the way for future advancements by proposing strategies including architectural investigation, dataset expansion, and fine-tuning to increase the capabilities of YOLO v8 skin cancer diagnosis.

REFERENCES

1. F. Xie, H. Fan, Y. Li, Z. Jiang, R. Meng and A. Bovik, "Melanoma Classification on Dermoscopy Images Using a Neural Network Ensemble Model," in IEEE Transactions on Medical Imaging, vol. 36, no. 3, pp. 849-858, March 2017, doi: 10.1109/TMI.2016.2633551.
2. M. Q. Khan et al., "Classification of Melanoma and Nevus in Digital Images for Diagnosis of Skin Cancer," in IEEE Access, vol. 7, pp. 90132-90144, 2019, doi: 10.1109/ACCESS.2019.2926837.
3. Bisla D.,Choromanska A., Stein J.A., Polsky D., Berman R Towards Automated Melanoma Detection with Deep Learning: Data Purification and Augmentation. Accepted to CVPR ISIC Workshop - 2019 Available online: http://arxiv.org/abs/1902.060611902.06061 [Google Scholar]
4. T. C. Pham, G. S. Tran, T. P. Nghiem, A. Doucet, C. M. Luong and V. -D. Hoang, "A Comparative Study for Classification of Skin Cancer," 2019 International Conference on System Science and Engineering (ICSSE), Dong Hoi, Vietnam, 2019, pp. 267-272, doi: 10.1109/ICSSE.2019.8823124.
5. N. Rezaoana, M. S. Hossain and K. Andersson, "Detection and Classification of Skin Cancer by Using a Parallel CNN Model," 2020 IEEE International Women in Engineering (WIE) Conference on Electrical and Computer Engineering (WIECON-ECE), Bhubaneswar, India, 2020, pp. 380-386, doi: 10.1109/WIECON-ECE52138.2020.9397987.
6. Javaid, M. Sadiq and F. Akram, "Skin Cancer Classification Using Image Processing and Machine Learning," 2021 International Bhurban Conference on Applied Sciences and Technologies (IBCAST), Islamabad, Pakistan, 2021, pp. 439-444, doi: 10.1109/IBCAST51254.2021.9393198.
7. R. R. Subramanian, D. Achuth, P. S. Kumar, K. Naveen kumar Reddy, S. Amara and A. S. Chowdary, "Skin cancer classification using Convolutional neural networks," 2021 11th International Conference on Cloud Computing, Data Science & Engineering (Confluence), Noida, India, 2021, pp. 13-19, doi: 10.1109/Confluence51648.2021.9377155.
8. Y. Jusman, I. M. Firdiantika, D. A. Dharmawan and K. Purwanto, "Performance of Multi-Layer Perceptron and Deep Neural Networks in Skin Cancer Classification," 2021 IEEE 3rd Global Conference on Life Sciences and Technologies (LifeTech), Nara, Japan, 2021, pp. 534-538, doi: 10.1109/LifeTech52111.2021.9391876.
9. G. C. Pacheco and R. A. Krohling, "An Attention-Based Mechanism to Combine Images and Metadata in Deep Learning Models Applied to Skin Cancer Classification," in IEEE Journal of Biomedical and Health Informatics, vol. 25, no. 9, pp. 3554-3563, Sept. 2021, doi: 10.1109/JBHI.2021.3062002.
10. Huang, H.-Y.; Hsiao, Y.-P.; Mukundan, A.; Tsao, Y.-M.; Chang, W.-Y.; Wang, H.-C. Classification of Skin Cancer Using Novel Hyperspectral Imaging Engineering via YOLOv5. J. Clin. Med. 2023, 12, 1134. https://doi.org/10.3390/jcm12031134
11. Manoj, S. O., Abirami, K. R., Victor, A., & Arya, M. (2023). Automatic Detection and Categorization of Skin Lesions for Early Diagnosis of Skin Cancer Using YOLO-v3 - DCNN Architecture. Image Analysis and Stereology, 42(2), 101–117. https://doi.org/10.5566/ias.2773
12. S. Albahli, N. Nida, A. Irtaza, M. H. Yousaf and M. T. Mahmood, "Melanoma Lesion Detection and Segmentation Using YOLOv4-DarkNet and Active Contour," in IEEE Access, vol. 8, pp. 198403-198414, 2020, doi: 10.1109/ACCESS.2020.3035345.

ISBN 979-8-3503-5094-4

Machine learning based computer aided diagnosis of breast tumor from Dynamic Contrast Enhanced MRI images

Hetvi M. Dhimmar
Department of Biomedical Engineering
SRM Institute Science and Technology
Kattankulathur, Chengalpattu, India
hd8880@srmist.edu.in

Angeline Kirubha S P*
Department of Biomedical Engineering
SRM Institute Science and Technology
Kattankulathur, Chengalpattu, India
angelins1@srmist.edu.in
0000-0001-6221-298X

Jeya Prabha A
Department of Biomedical Engineering
SRM Institute Science and Technology
Kattankulathur, Chengalpattu, India
ja2654@srmist.edu.in
0009-0001-4769-7584

Sameera Fathimal M
Department of Biomedical Engineering
SRM Institute Science and Technology
Kattankulathur, Chengalpattu, India
sm5282@srmist.edu.in
0000-0003-0460-5304

***Abstract*—Worldwide, breast cancer (BC) is the most common cancer to be diagnosed in women. It is currently the second most common cause of cancer-related death, trailed only by lung cancer. When compared to ultrasound and mammography, breast dynamic contrast-enhanced magnetic resonance imaging (DCE-MRI) is helpful for the diagnosis of breast tumors. Breast cancer diagnosis with dynamic contrast-enhanced magnetic resonance imaging (DCE-MRI) has demonstrated good sensitivity. Specifically, a dataset of 500 images including 250 each for benign and malignant is used. Deep learning based Attention U-Net model was used to perform the segmentation of the breast tumor. Various feature extraction techniques such as statistical features, Haralick features, Edge features, Shape features were extracted after segmentation. After performing feature extraction ML classification was implemented to classify benign and malignant tumor. Support Vector Machine, Random Forest classifier, Gaussian Naïve Bayes and K-Nearest Neighbours were used for classification. Among them K-Nearest Neighbours has performed better than all other ML models with accuracy of 82%**

Keywords—Breast tumor, Machine learning, DCE-MRI, classification, Breast cancer

I. INTRODUCTION

A. Breast Cancer

A malignant tumor that arises from breast cells is called breast cancer. Approximately one-third of all malignancies in women are caused by it, making it the most prevalent cause of cancer in women [1]. Milk-secreting cells, duct cells, or fatty/fibrous tissue can also give rise to breast cancer [2]. Although it is far more frequent in women, it may also happen to males [3]. While lobular cancer is less prevalent, the most common kind of breast cancer starts in the cells lining the milk ducts [4]. The size of the tumor and whether it has progressed to other regions of the body are used to categorise breast cancer [5]. Radiological examination is frequently used for screening and detection, and early diagnosis is crucial for increasing survival rates. Given its rising global prevalence, breast cancer is a serious global health problem. It is the most common cancer to be diagnosed with and the primary cause of cancer-related deaths among women worldwide. An estimated 2.3 million new cases and 685,000 deaths from breast cancer were reported in 2020 [6]. Breast cancer incidence rates vary geographically; Australia/New Zealand, Western Europe, and Northern America have the highest incidence rates, while South-Central Asia, Middle and Eastern Africa, and Central America have the lowest rates [7]. Moreover, breast cancer ranks first among cancer-related deaths in the globe for women [8]. Future estimates of the number of instances of breast cancer indicate that by 2025, there will be 2.2 million additional cases [9].

Based on their origin and features, breast tumors can be categorised into many categories. Smooth muscle, neural tissue, adipose tissue, vascular tissue, and other less frequent cellular origins are the less common sources of mesenchymal breast tumors, a rare and heterogeneous group of tumors [10]. Malignant breast tumors can originate from mesodermal tissues, which can lead to high-grade breast sarcoma, or from epithelial tissue, which can lead to invasive ductal carcinoma and squamous cell carcinoma [11]. The expression of hormone receptors, such as those for progesterone, oestrogen, and human epidermal growth factor, is the basis for the most prevalent categorization of breast cancer. HER2-positive, triple-negative, luminal A, and luminal B subtypes are included in this categorization [12]. Lipoma, angiolipoma, leiomyoma, neurofibroma, and schwannoma are examples of benign mesenchymal tumors in the breast; liposarcoma, leiomyosarcoma, rhabdomyosarcoma, malignant peripheral nerve sheath tumor, and osteosarcoma are examples of malignant mesenchymal tumors [13]. Sarcomas, lymphomas, and phyllodes tumors are non-epithelial breast tumors that have different therapy and prognoses than carcinomas [14].

B. Diagnostic Methods for breast cancer:

The conventional imaging method for detecting breast cancer, mammography, has drawbacks that it may reduce its efficacy. Misinterpreted non-cancerous lesions and difficult interpretations that lead to the lack of malignancies are a couple of these limitations. [15]. Furthermore, 10–30% of breast tumors may be undetected by mammography, which might delay identification and treatment [16]. Due to the limitations of mammography, other tools like as content-

ISBN 979-8-3503-5094-4

based mammogram retrieval (CBMR) approaches and computer-aided diagnosis (CAD) systems have been developed to aid radiologists in interpretation [17]. Mammography has improved throughout time, including changes to increase its sensitivity and specificity, but it is still a useful diagnostic tool in clinical settings [18]. The precision of breast cancer risk assessment may be affected by the constraints and difficulties related to mammographic density measures, hence it's essential that you are aware of them [19].

When it comes to breast cancer screening, breast ultrasonography has been proven to be a useful addition to mammography, especially for women with thick breast tissue [20]. It has demonstrated improved sensitivity in identifying extra tumors that mammography could miss. One promising imaging modality for the identification and characterisation of breast cancer is ultrasound elastography [21]. The drawbacks of handheld ultrasound, such as operator reliance and a narrow field of view, have been addressed by automated breast ultrasonography (AB US), which has been used to evaluate the tumor response to neoadjuvant therapy (NAT) [22]. Furthermore, ultrasound-responsive nanocarriers have been developed to improve medication administration in chemotherapy for breast cancer, offering a tissue-penetrating, non-invasive, and affordable method [23]. All things considered, breast ultrasonography can enhance patient outcomes by assisting mammography and resolving the drawbacks of other imaging modalities.

When compared to other approaches, DCE-MRI is a better way to diagnose breast cancer since it offers valuable information for high-risk breast cancer screening and prognostic evaluation [24] [25]. It makes it possible to segregate malignant tissues, which is crucial for further breast MRI analysis [26] [27]. By evaluating microvascular permeability and identifying circulating tumor DNA, DCE-MRI can also be used to predict long-term survival outcomes and the responsiveness of patients to neoadjuvant treatment [28]. Furthermore, DCE-MRI minimises hazards and side effects by producing contrast-enhanced images without the need for contrast agents. By considering variables like nodule size and age at menarche, computer algorithms based on DCE-MRI can increase the accuracy of identifying various subtypes of breast cancer. All things considered, DCE-MRI provides a thorough and non-invasive method of diagnosing and assessing the prognosis of breast cancer.

C. *Machine learning methods for classification of Breast tumor:*

By examining big datasets and spotting intricate patterns, machine learning (ML) algorithms are utilised to categorise breast cancers [29]. The UCI Wisconsin breast cancer dataset and the Wisconsin repository's breast cancer diagnostic medical dataset are two examples of the datasets to in which Logistics Regression, SVM, KNN, Random Forest, Naive Bayes, and Decision Tree classifier have been used [30] [31]. Using criteria like accuracy, sensitivity, specificity, F1 score, and precision, the performance of these models is assessed. The findings demonstrate the great accuracy of the SVM, random forest, gradient boosting, and AdaBoost models—some of which reach 100% accuracy. Machine learning models offer a dependable and effective method for categorising breast cancers, supporting prompt diagnosis and therapy.

II. Materials and Methods

A. *Study Data Collection*

The Breast tumor dataset, which comprises of 500 DCE-MRI pictures of two classes: Benign and Malignant was gathered via the freely accessible website github. Total number of images collected are 500. Images are in Jpg format. Each benign and Malignant case has 250 images and Corresponding mask images. This dataset is superior to existing breast datasets since it has a greater number of cases, and is exclusively focused on the DCE-MRI domain.

A 3.0T scanner with a 20-channel head and neck coil (GE SIGNA HDx, Siemens Skyra, Erlangen, Germany) was used to scan all DCE-MRI data. Using a multiphase axial T1 volume imaging for breast assessment (VIBRANT) method that covered both breasts, the DCE images were acquired. FOV = 34 × 34 cm2, slice thickness = 2.4 mm, TE = 2.3 ms, FA = 10◦, TR = 4.7 ms, and matrix size = 512 × 512 were the specifications of the collected image. The DCE series consists of eight post-contrast continuous scanning sequences following the delivery of contrast agent, one pre-contrast sequence before to injection, and their corresponding subtract sequences.

It took 2 minutes and 55 seconds to finish the scan. 20 ml of saline flush were injected at the same speed after pre-contrast sequences were acquired at a rate of 2 ml per second. After that, an intravenous injection of the gadopentetate dimeglumine-containing contrast agent (Germany, Magnevist, Berlin, BayerHealth Care) was given. The picture resolution ranged from 0.8 × 0.8 × 1.0 mm3 to 0.6 × 0.6 × 2.4 mm3, while the image size varied between 512 × 512 × 71 and 512 × 512 × 112. [39]

B. *Data division*

In this investigation, a train-test split ratio of 80:20 was selected. In total of 500 images, 400 images were allocated for the purpose of training with an additional 100 images designated for model testing. Binary classification algorithms have been implemented to diagnose as Benign or Malignant.

C. *Methodology*

1) Breast tumor Segmentation using Attention U-Net Model

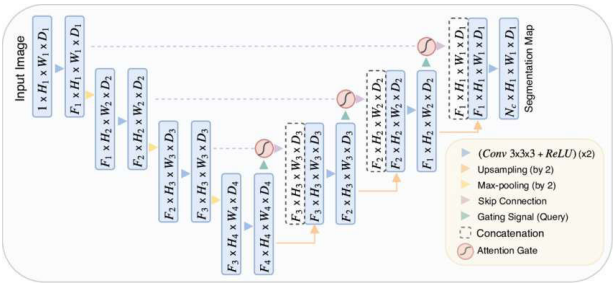


Fig. 1. Architecture of Attention U-Net model for segmentation

Courtesy:https://www.mdpi.com/

As seen in Fig. 1, the suggested Attention U-Net [40] is included into the conventional U-Net design to draw attention

ISBN 979-8-3503-5094-4

to noteworthy aspects that are sent over the skip connections. In skip connections, gating uses data from the coarse scale to separate out irrelevant and noisy data. In order to combine just pertinent activations, this is done immediately before to the concatenation procedure. Furthermore, both during the forward and backward passes, Attention Gate(AG) filter the activations of neurons. During the backward pass, gradients coming from background areas are downweighted. During feature fusion, Attention U-Net dynamically weights the significance of various spatial locations, allowing the network to concentrate more on pertinent regions for segmentation tasks.

Each sub-AG's complementary data is extracted and fused to define the output of the skip connection. To reduce the number of trainable parameters and the computational cost of AGs, input feature-maps are downsampled to the gating signal resolution, similar to non-local blocks, and spatially-independent linear transformations (1x1x1 convolutions) are performed. For the purpose of the gating operation, the necessary linear transformations disconnect the feature-maps and transfer them to a lower dimensional space. Since low-level feature-maps, or the initial skip connections, do not reflect the input data in a high-dimensional space, they are not employed in the gating function as described in. To ensure that the intermediate feature-maps are semantically discriminative at every picture size, deep-supervision has employed. This makes it possible for attention units to affect responses to a wide variety of picture foreground material at various sizes. For this reason, tiny subsets of skip connections were stopped from being used to rebuild dense predictions.

2) Feature Extraction

Determining the best technique for feature extraction from medical pictures is really difficult. As of yet, there isn't a single universal feature extraction approach that works with every object. All, these rely on a wide range of applications that the writers kept working on. When choosing a reliable feature extraction technique, it's important to consider a number of criteria, such as the approach used, the data availability, dimensionality, labelling, and whether or not the feature extraction is supervised. Through a methodical arrangement of these phases, a researcher can choose the many existing algorithms a workable strategy for feature extraction.

Grey level co-occurrence matrix: A feature extraction tool based on statistical methods is the grey level co-occurrence matrix (GLCM). Grey level co-occurrence or grey level spatial dependency matrices are used to depict the spatial connection between pixels. The spatial rapport is shown by the pixel of interest in a grey level co-occurrence matrix. More spatial relationships may be created by arranging neighbouring pixels in the middle of two pixels in a horizontal orientation. The component is calculated for the ensuing co-occurrence matrix at the grey level. To put it simply, the value is the total of all the times a certain pixel value has occurred in a given spatial relationship to a pixel in the input picture.

Haralick features: The core of the co-occurrence matrix that is created using the images is where the Haralick texture characteristics are located. The objective of castoff occurrence matrices is to expose the lateral spatial dependency of grey level information, including vertical and horizontal direcions and angular correlations, in the pictures. The co-occurrence matrix may be adjusted to provide a wide range of modified texture properties. The inverse difference moment, sum average, sum variance, sum entropy, information measures of correlation, difference entropy, angular second moment, contrast, correlation, sum of squares, and sum entropy are among the thirteen elements that Haralick extracted from the co-occurrence matrix.

Gabor texture characteristics: Gabor feature extraction is based on the creation of gabor filters. These filters are similar to Gabor wavelets and are conceptualised with the human visual system in mind. Moreover, it has been noted that the writers have a remarkable talent for textural analysis. It is a linear filter with edge identification as its main use. The Gabor texture's characteristics are just a group of wavelets. Each wavelet then determines vitality in a timely and accurate manner. The textural feature connected to the pictures may be retrieved after the energy distributions have been collected.

Area, Perimeter, Mean, Standard Deviation, Skewness, Kurtosis and Fractal Dimension, features are used for feature extraction method

3) Machine Learning Classification:

Supervised machine learning classification is implemented in this context, as each data point in the issue contains a target label. Support Vector Machine (SVM) and other robust machine learning algorithms are applied to regression, outlier detection, and linear or nonlinear classification, in this project.

A simple and easy to comprehend machine learning classifier is K-Nearest Neighbours (KNN). It functions on the concept that instances that are comparable in a feature space often belong to the same class. The majority class among the K-nearest data points in the training set is assigned to the new point by KNN when predicting the class of a new data point. The number of neighbours, K, determines how sensitive the model is to local patterns. Since KNN is a non-parametric technique, it may be used to a wide range of datasets since it doesn't make any assumptions about the distribution of the underlying data.

The Gaussian Naive Bayes classifier is a probabilistic machine learning algorithm based on Bayes' theorem. It assumes that features within each class follow a Gaussian distribution. The classifier calculates the likelihood of observing a specific feature value using the Gaussian probability density function. By incorporating prior probabilities and applying the Maximum Posteriori decision rule, it assigns the class with the highest posterior probability as the prediction.

During training, the Random Forest classifier constructs many decision trees as part of an ensemble learning technique. It integrates their predictions in order to enhance accuracy and manage overfitting. By aggregating the results of individual trees, Random Forests provide robust and versatile solutions for both classification and regression tasks.

Employing a train-test split ratio of 80:20, 100 images are utilised for model testing and 400 are allocated for training.

ISBN 979-8-3503-5094-4

A performance comparison is made between SVM classifiers, K-Nearest Neighbours and Random forest.

4) Performance Configuration:

The proposed models have been implemented on Google Colaboratory with a 15GB GPU. The hyper parameters used for training in segmentation were 100 epochs, SGD *optimizer, binary cross entropy loss and 0.001 learning* rate.

III. RESULTS

A. SEGMENTATION OUTPUT OF BREAST TUMOR REGION:

The output of segmentation technique using Attention U-Net model is shown in fig. 2. The segmentation output provides a visual representation of the tumor regions, aiding in diagnosis and treatment planning. Here's how the *segmentation output may differ for benign and malignant breast tumors*

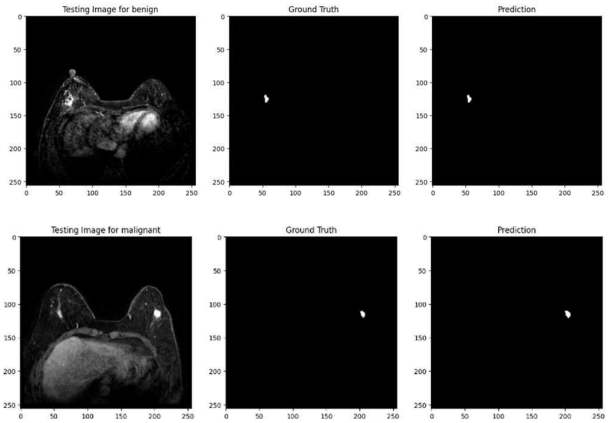


Fig. 2. Segmentation output for benign and malignant tumor

Benign tumors have distinct borders in the result of segmentation. Benign tumor segmentation findings usually have a more localised and compact form. The bounded area depicts the benign mass's confinement.

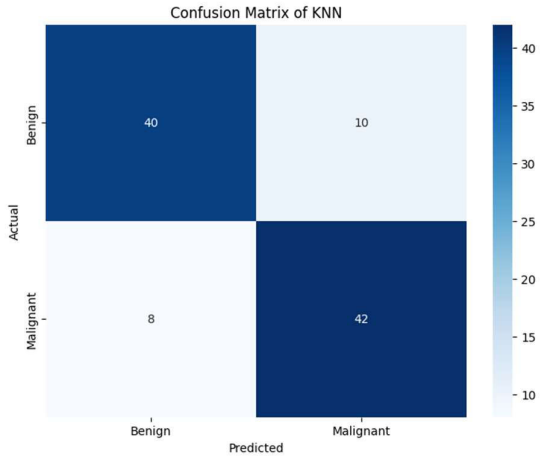
Malignant tumors typically exhibit segmentation output with inconsistent and imprecise borders. Malignant tumor segmentation results might have a more asymmetrical and convoluted form, which is indicative of the invasive nature of cancer cells.

	Benign	Malignant
Intersection Over Union (IoU)	90%	86%
Dice coefficient	95%	93%
Accuracy	95%	93%
Precision	94%	92%
Recall	94.5%	92%
Specificity	94%	93%
F1 score	95%	92%
Mean Absolute Error (MAE)	5	7
Positive Predicted Value (PPV)	95%	93%

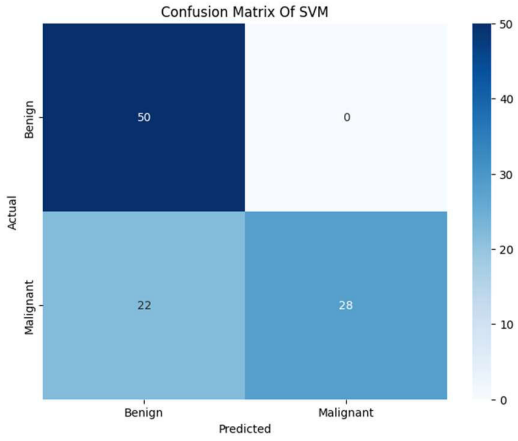
Table 1. Performance metrics for segmentation

B. MACHINE LEARNING CLASSIFICATION:

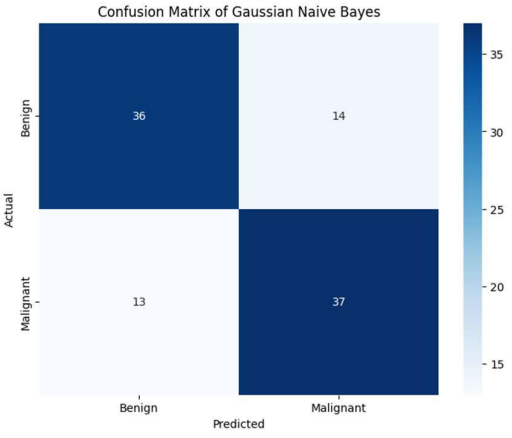
The machine learning models used for classification of breast tumor are SVM, K-Nearest Neighbours, Gaussian Naive Bayes and Random forest. The evaluation of these classifier is based on its accuracy as measured by F1 score,precision, and recall, which is used for comparisons. Below figure shows the confusion matrix for all the classifier.



(a)

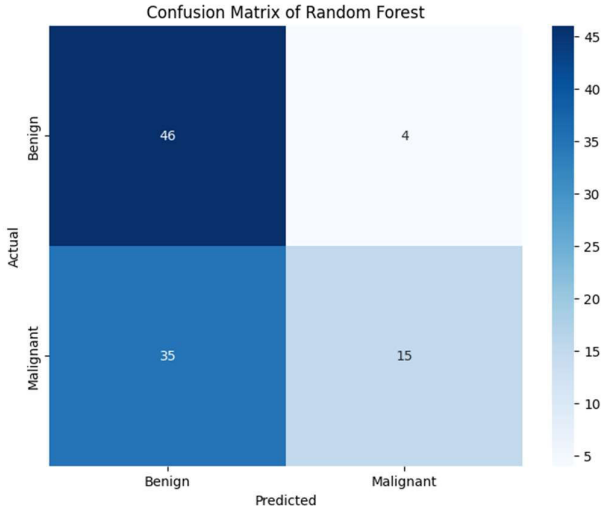


(b)



(c)

ISBN 979-8-3503-5094-4



(d)

Fig. 3. Confusion matrix for testing (a) KNN classifier (b) SVM classifier (c) Gaussian Naïve Bayes classifier (d) Random Forest classifier

Classifier	F1 Score (%)	Precision (%)	Recall(%)	Accuracy (%)	Specificity (%)	Sensitivity (%)
SVM	81	100	100	78	69	100
KNN	83	80	84	82	83	80
Gaussian Naïve Bayes	73	72	74	73	100	51
Random Forest	70	78	92	61	100	51

Table 2. Performance metrics comparison

The K-Nearest Neighbours classifier shows an overall accuracy of 82%, F1 score 83%, Precision 80%. Overall K-Nearest Neighbours classifier performs better than other classifier used for classification.

IV. Discussions

About 12% of all new instances of tumours and 25% of all malignancies in women are breast cancers. At about 2 million newly diagnosed cases annually, it ranks as the second most frequent cancer in women globally. Early identification is crucial for improving the prognosis of breast neoplasms, particularly in cases of cancer (malignant tumours). Antropova et al. [38] demonstrated the first use of deep learning to the DCE-MRI lesion classification challenge. They recommend employing a pretrained CNN model. The predictive ability of the CNN was demonstrated by the AUC value of 0.85 obtained by combining transfer learning with SVM thereafter. Nattkemper et al. [39] DCE-MRI data of the female breast are obtained from The UK Multicenter Breast Screening Study. The data was categorised with 74% accuracy (DT) by the supervised techniques, and an area under the receiver-operator-characteristics (ROC) curve (AUC) of 0.88 (SVM) was obtained. Chen et al. [40] Predictive models were constructed to predict the histological grade of breast cancer based on the statistical and Haralick texture characteristics that retrieved from DCE-MRI. Their classifier performed well, as seen by its 0.859 AUC. In our project we have used 500 DCE-MRI images. We have implemented various ML models. Among them K-Nearest Neighbours classifier gives better performance with Accuracy of 82%. The identification and classification of breast tumours are crucial because they may greatly enhance treatment plans and raise patient survival rates. Early detection and precise classification of breast tumours can also greatly improve treatment outcomes. Classification techniques, including machine learning and statistical techniques, can act as impartial instruments to help medical professionals identify patients correctly.

V. Conclision and Future Work

The availability of a sufficiently large and diverse dataset is one of the main obstacles. The amount and quality of training data have a major impact on how well machine learning algorithms perform, particularly deep learning models like Attention U-Net. The model's capacity to generalise to different settings and tumour types may be impacted by restricted access to a variety of DCE-MRI images. The quantity and quality of data have a major impact on the performance of machine learning algorithms. Poor-quality or scarce datasets may have an impact on how well your models function. It is difficult to extract important features from DCE-MRI pictures. The performance of the classification model can be greatly impacted by what features are used. More research is required to determine how resilient the Attention U-Net method and other machine learning models are to changes in noise, resolution, and imaging circumstances. Further research can concentrate on investigating deeper feature extraction methods. The capacity of the model to extract discriminative features for tumour classification may be enhanced by looking at the use of deep learning-based feature extraction techniques or by incorporating domain-specific knowledge. It might be useful to explore and put data augmentation strategies into practise. Creating artificial modifications of the available data, such flips, rotations, or contrast adjustments, can assist strengthen the model and lower the chance of overfitting. A more thorough picture of the tumour features may be obtained by adding more clinical data, such as patient history and other imaging modalities, to the model. Results for categorization that are more precise and therapeutically meaningful might arise from this combination.

References

[1] Pham, P. V., & Van Pham, P. (2015). Introduction to Breast Cancer. Breast Cancer Stem Cells & Therapy Resistance, 1-4.

[2] Ronckers, C. M., Erdmann, C. A., & Land, C. E. (2004). Radiation and breast cancer: a review of current evidence. Breast Cancer Research, 7, 1-12.

[3] Gupta, M., Wahi, A., Sharma, P., Nagpal, R., Raina, N., Kaurav, M., ... & Nissapatorn, V. (2022). Recent advances in cancer vaccines: challenges, achievements, and futuristic prospects. Vaccines, 10(12), 2011.

[4] De, S. K. (2022). Fundamentals of Cancer Detection, Treatment, and Prevention. John Wiley & Sons.

[5] A, Ananya., D, Sudheer, Kumar., Poonam, Kishore. (2019). A review on breast cancer. The Pharma Innovation Journal, 8(5):57-62.

[6] (2023). Prevalence of women breast cancer. Cellular, Molecular and Biomedical Reports, doi: 10.55705/cmbr.2023.384467.1095

[7] Ephrem, Sedeta., Bilain, Yilma, Jobre., B., Avezbakiyev. (2023). Breast cancer: Global patterns of incidence, mortality, and trends..

ISBN 979-8-3503-5094-4

Journal of Clinical Oncology, doi: 10.1200/jco.2023.41.16_suppl.10528

[8] (2022). The Lancet Breast Cancer Commission: tackling a global health, gender, and equity challenge. The Lancet, doi: 10.1016/s0140-6736(22)00184-2

[9] Ghulam, Fatima., Muhammad, Rahil, Khan., R., Memon., Zahid, Ali., Rehan, Akhtar. (2023). Prevalence of Luminal A, Luminal B, HER2 enriched and triple negative breast carcinoma in population of Sindh. International journal of health sciences, doi: 10.53730/ijhs.v7ns1.14443

[10] Zhao, Sumin. (2023). Mesenchymal Tumors of the Breast: Fibroblastic/Myofibroblastic Lesions and Other Lesions. Current Oncology, doi: 10.3390/curroncol30050338

[11] Chao, Li., Hongfei, Fang., Hongfeng, Lu. (2021). Case Report of a Breast Mass with Three Types of Malignant Tumors—Squamous Cell Carcinoma, Invasive Ductal Carcinoma, and Breast Sarcoma. OncoTargets and Therapy, doi: 10.2147/OTT.S304145

[12] (2022). Subtypes of Breast Cancer. doi: 10.36255/exon-publications-breast-cancer-subtypes

[13] Ekrem, Yavuz., Sitki, Tuzlali. (2019). Mesenchymal and Fibroepithelial Tumors of the Breast. doi: 10.1007/978-3-319-96947-3_4

[14] Márcia, Cristina, Santos, Pedrosa. (2019). Non-epithelial Tumors. doi: 10.1007/978-3-030-13636-9_48

[15] Noam, Nissan., Efi, Efraim, Moss, Massasa., E., Bauer., Arsalan, Abu-Much., David, Samoocha., Yael, Yagil., Renata, Faermann., Osnat, Halshtok-Neiman., Anat, Shalmon., Michael, Gotlieb., Miri, Sklair-Levy. (2023). Pacemaker in patients undergoing mammography: A limitation for breast cancer diagnosis?. Journal of Medical Imaging and Radiation Oncology, doi: 10.1111/1754-9485.13524

[16] (2022). Breast Cancer Diagnosis With Mammography. doi: 10.4018/978-1-6684-7136-4.ch037

[17] Nikolaos, S., Salemis., Eleni, Mourtzoukou., Michail, Angelopoulos. (2021). Subtle signs of breast cancer as an important pitfall in mammographic interpretation: Establishing an accurate diagnosis in an equivocal case.. Breast disease, doi: 10.3233/BD-201075

[18] José, Thomas., Amulya, Cherukumudi. (2022). Mammography—sentinel of breast cancer management. doi: 10.1016/b978-0-323-85650-8.00002-4

[19] Anubha, Wadhwa., Julie, R, Sullivan., Mary, Beth, Gonyo. (2016). Missed Breast Cancer: What Can We Learn?. Current Problems in Diagnostic Radiology, doi: 10.1067/J.CPRADIOL.2016.03.001

[20] Kavak, Mesut. (2023). Breast Ultrasound: A Valuable Adjunct to Mammography in Breast Cancer Screening. doi: 10.31219/osf.io/6jcpz

[21] Gelan, Ayana., Jae-Min, Ryu., Se-woon, Choe. (2022). Ultrasound-Responsive Nanocarriers for Breast Cancer Chemotherapy. Micromachines, doi: 10.3390/mi13091508

[22] Miguel-Angel, Pardo-Vicente. (2023). Ultrasound Elastography in Breast Cancer: A Promising Diagnostic Tool. doi: 10.31219/osf.io/u6eb3

[23] Xiaozhi, Dang., Xin, Zhang., Yi, Gao., Hongping, Song. (2022). Assessment of Neoadjuvant Treatment Response Using Automated Breast Ultrasound in Breast Cancer. Journal of Breast Cancer, doi: 10.4048/jbc.2022.25.e32

[24] (2023). SimPLe: Similarity-Aware Propagation Learning for Weakly-Supervised Breast Cancer Segmentation in DCE-MRI. doi: 10.48550/arxiv.2306.16714

[25] Yumin, Zhong., Yi, Wang. (2023). SimPLe: Similarity-Aware Propagation Learning for Weakly-Supervised Breast Cancer Segmentation in DCE-MRI. arXiv.org, doi: 10.48550/arXiv.2306.16714

[26] Andrew, T., Stefka., Bradley, Feiger., Jason, R., Pfeiffer., Tushar, Pandey., Joseph, R., Peterson., A., Antony. (2023). DCE-MRI-derived microvasculature mapping and outcome prediction in patients with breast cancer.. Journal of Clinical Oncology, doi: 10.1200/jco.2023.41.16_suppl.e12614

[27] (2023). Tumor-Attentive Segmentation-Guided GAN for Synthesizing Breast Contrast-Enhanced MRI Without Contrast Agents. IEEE Journal of Translational Engineering in Health and Medicine, doi: 10.1109/jtehm.2022.3221918

[28] Liang, Sun., Haowen, Tian., Hongwei, Ge., Juan, Tian., Yuxin, Lin., Chang, Liang., Tang, Liu., Yiping, Zhao. (2023). Cross-attention multi-branch CNN using DCE-MRI to classify breast cancer molecular subtypes. Frontiers in Oncology, doi: 10.3389/fonc.2023.1107850

[29] Maurício, Gattás, Bara, Filho. (2023). Breast Cancer Classification Using ML on WDBC. doi: 10.1007/978-981-99-0189-0_48

[30] Alessandro, Ventura. (2023). Analysis of ML-Based Classifiers for the Prediction of Breast Cancer. Algorithms for intelligent systems, doi: 10.1007/978-981-99-1435-7_30

[31] Orlando, Iparraguirre-Villanueva., Andrés, Epifanía-Huerta., Carmen, Torres-Ceclén., John, Ruiz-Alvarado., Michael, Cabanillas-Carbonell. (2023). Breast Cancer Prediction using Machine Learning Models. International Journal of Advanced Computer Science and Applications, doi: 10.14569/ijacsa.2023.0140272

[32] Hu, Q., Whitney, H. M., & Giger, M. L. (2020). A deep learning methodology for improved breast cancer diagnosis using multiparametric MRI. Scientific reports, 10(1), 10536.

[33] Ibrahim, A., Mohammed, S., Ali, H. A., & Hussein, S. E. (2020). Breast cancer segmentation from thermal images based on chaotic salp swarm algorithm. IEEE Access, 8, 122121-122134.

[34] Ibraheem, A. M., Rahouma, K. H., & Hamed, H. F. (2019, October). Automatic MRI breast tumor detection using discrete wavelet transform and support vector machines. In 2019 Novel Intelligent and Leading Emerging Sciences Conference (NILES) (Vol. 1, pp. 88-91). IEEE.

[35] Alanazi, S. A., Kamruzzaman, M. M., Islam Sarker, M. N., Alruwaili, M., Alhwaiti, Y., Alshammari, N., & Siddiqi, M. H. (2021). Boosting breast cancer detection using convolutional neural network. Journal of Healthcare Engineering, 2021.

[36] Ma, D., Shang, L., Tang, J., Bao, Y., Fu, J., & Yin, J. (2021). Classifying breast cancer tissue by Raman spectroscopy with one-dimensional convolutional neural network. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 256, 119732.

[37] Zhao, X., Liao, Y., Xie, J., He, X., Zhang, S., Wang, G., ... & Yu, J. (2023). BreastDM: A DCE-MRI dataset for breast tumor image segmentation and classification. Computers in Biology and Medicine, 164, 107255.

[38] Oktay, O., Schlemper, J., Folgoc, L. L., Lee, M., Heinrich, M., Misawa, K., ... & Rueckert, D. (2018). Attention u-net: Learning where to look for the pancreas. arXiv preprint arXiv:1804.03999.

[39] Antropova, N., Huynh, B., & Giger, M. (2016). SU-D-207B-06: Predicting breast cancer malignancy on DCE-MRI data using pre-trained convolutional neural networks. Medical physics, 43(6Part4), 3349-3350.

[40] Nattkemper, T. W., Arnrich, B., Lichte, O., Timm, W., Degenhard, A., Pointon, L., ... & Study, T. U. M. B. S. (2005). Evaluation of radiological features for breast tumour classification in clinical screening with machine learning methods. Artificial intelligence in medicine, 34(2), 129-139.

[41] Pang, Z., Zhu, D., Chen, D., Li, L., & Shao, Y. (2015). A computer-aided diagnosis system for dynamic contrast-enhanced MR images based on level set segmentation and ReliefF feature selection. Comput. Math. Methods Medicine, 2015, 450531-1.

ISBN 979-8-3503-5094-4

Revolutionizing Alzheimer's Disease Prediction Using EfficientNetB6

V Sai Praneeth
Department of Electronics and Communication Engineering,
College of Engineering and Technology,
SRM Institute of Science and Technology,
Kattankulathur-603203, Tamil Nadu, India.
vp6744@srmist.edu.in

Naraganti Gowtham
Department of Electronics and Communication Engineering,
Faculty of Engineering and Technology,
SRM Institute of Science and Technology,
SRM Nagar, Kattankulathur
603203, Tamil Nadu, India.
ng0041@srmist.edu.in

SivaRamaChandran
Department of Electronics and Communication Engineering,
Faculty of Engineering and Technology,
SRM Institute of Science and Technology,
SRM Nagar, Kattankulathur
603203, Tamil Nadu, India.
mc7391@srmist.edu.in

R.Jansi*
Department of Electronics and Communication Engineering,
Faculty of Engineering and Technology,
SRM Institute of Science and Technology,
SRM Nagar, Kattankulathur
603203, Tamil Nadu, India.
jansir@srmist.edu.in

***Abstract* -Alzheimer's disease (AD) presents a pressing challenge globally, demanding the development of accurate predictive models for early diagnosis and intervention. In this research, we propose a novel prediction framework harnessing deep learning methodologies, specifically employing the advanced EfficientNetB6 algorithm, on the Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset comprising approximately 20,000 MRI images. Our aim is to differentiate between MRI images indicating normal cognitive function and those signaling AD progression, with a target of achieving 97.78% accuracy and a 98.21% F1 score. Moreover, our model categorizes images into various stages, including mild cognitive impairment (MCI), facilitating early identification. The proposed model's high accuracy and precision hold the potential to transform early detection efforts, allowing for timely interventions and thereby improving patients' quality of life. In addition to leveraging the EfficientNetB6 model for image feature extraction and LSTM networks for sequential data analysis, the proposed hybrid architecture incorporates dense layers to facilitate feature fusion and abstraction. The dense layers enable the model to combine the extracted image and sequential features, capturing complex relationships between different modalities and enhancing the model's discriminative power. By integrating dense layers, the model gains the flexibility to learn higher-level representations, further refining the extracted features for improved classification performance. This holistic approach, combining convolutional, recurrent, anddense layers, offers a comprehensive solution for multi- modal data analysis, showcasing the potential of deep learning techniques in addressing complex real-world problems.**

Keywords -- Alzheimer's disease, neurological disorder, deep learning, EfficientNetB6 algorithm, ADNI dataset.

I. NAVIGATING THE JOURNEY OF ALZHEIMER'S

Alzheimer's disease (AD) is recognized as the predominant cause of dementia, affecting 60– 80% of cases, with a gradual progression from mild cognitive impairment. Its global impact is substantial, with approximately 50 million individuals worldwide suffering from dementia, a number anticipated to rise significantly [1]. Early detection is imperative, and automated systems, particularly Convolutional Neural Networks (CNNs), have proven effective in analyzing magnetic resonance images for accurate AD identification. This study aims to develop a cutting-edge deep learning model leveraging CNNs to categorize AD stages, providing doctors with a rapid and precise diagnostic tool [2]. The research addresses the global impact of brain diseases,

ISBN 979-8-3503-5094-4

focusing specifically on Alzheimer's disease (AD) and brain tumors, which affect millions worldwide. It proposes a hybrid framework utilizing Particle Swarm Optimization (PSO) to optimize Convolutional Neural Network (CNN) hyperparameters for improved accuracy in AD and brain tumor prediction from MRI data. The study introduces contributions such as a PSO-optimized CNN model, evaluated on modern brain disease datasets, showcasing superior performance over existing models [3]. Deep learning has emerged as a powerful tool in medical image classification [4], revolutionizing the field with its ability to automatically extract complex features from images and make accurate predictions. In medical imaging, deep learning techniques [5] such as convolutional neural networks (CNNs), are employed to analyze images obtained from various modalities like MRI, CT, X-ray, and histopathology. This study delves into the intersection of computational neuroscience, machine learning, and deep learning in advancing mental health trials, particularly in the diagnosis and treatment of Alzheimer's disease (AD) using MRI data. Recent developments are scrutinized, highlighting the fusion of machine learning (ML) and deep learning (DL) techniques with biomedical datasets to enhance AD assessment. The paper offers insights into the historical trajectory, the transition from ML to DL methodologies, and reviews of AD modules and datasets [6]. Central to the investigation is Alzheimer's disease (AD) and its profound impact on cognitive decline, with an emphasis on the critical need for timely diagnosis. Notably, a maze learning study suggests that individuals with AD can still acquire perceptual-motor skills, underscoring cognitive involvement. The paper further explores imaging modalities such as functional MRI (fMRI) and MRI for precisely pinpointing AD-affected regions, juxtaposing clustering and deep learning methodologies. It accentuates the effectiveness of deep learning in medical image segmentation and suggests the application of collective intelligence methods to augment efficacy [7].

Hamdi et al. [13] presented a pressing need for early diagnosis of Alzheimer's Disease (AD) to delay its progression. Previous methods relied on image processing for feature extraction, which impacted recognition rates. Addressing this, a novel Convolutional Neural Network (CNN)-based

Addressing the escalating impact of dementia, notably Alzheimer's disease (AD), in aging populations, the absence of standardized diagnostic tests is highlighted, prompting the proposal of a machine learning-based computer-aided detection system for early AD diagnosis [8].A novel DL model for early AD detection using CNN and LSTM on sliced MRI data, optimized with ACO and MFCM, is presented, alongside discussions on the challenges of AD diagnosis and the imperative of early detection. Existing methods are reviewed, showcasing machine and deep learning techniques and stressing the significance of innovative approaches for timely AD identification [9].The introduction of DEMNET, a convolutional neural network tackling class imbalance in AD datasets using SMOTE, underscores the importance of neuroimaging, particularly MRI, for advanced AD diagnosis, while highlighting the criticality of early and precise AD prediction for effective patient care [10].Emphasizing the urgent need for accurate AD diagnosis, especially in the mild cognitive impairment (MCI) stage, advocacy is made for computer-aided diagnosis using artificialintelligence (AI), specifically in MRI-based applications [11]. AI-powered diagnostics revolutionize healthcare by swiftly and accurately processing vast medical data, enhancing disease identification. In medical imaging, AI's computer vision capabilities surpass human experts, enabling early detection of diseases like cancer and heart conditions. Moreover, AI facilitates personalized treatment plans by analyzing genetic, lifestyle, and medical data, maximizing therapeutic outcomes andpatient satisfaction. This transformative shift from conventional methods to tailored approachessignifies the dawn of precision medicine in healthcare [12].

II. EXPLORING THE DEPTH OF ALZHEIMER'S

Computer-Aided Diagnosis (CAD) system demonstrated superior performance, achieving 96% accuracy, 96% sensitivity, and 94% specificity in classifying 18FDG-PET images from 855 patients, outperforming existing methods in the literature.
Fathi et al.[14] conducted a literature review exploring the application of deep learning methods

ISBN 979-8-3503-5094-4

on neuroimaging data for early Alzheimer's disease (AD) diagnosis, given the rising prevalence of age-related diseases. Among 736 studies, 74 were analyzed, revealing a focus on normal control (NC)/AD classification with promising results. The review suggests that convolutional neural network (CNN)-based models, particularly in ensemble configurations, demonstrate superior performance, and recommends further research to establish a benchmark platform for comprehensive comparative analysis in this domain. Pradhan et al [15]. employed diverse methods, including deep learning and SVM, to achieve high accuracies (90-99%) in Alzheimer's detection from MRI data. Challenges include standardization for AI in cerebral imaging, while personalized classifiers aim to predict disease progression from MCI to AD within a year using biomarkers. Raees et al. [16] highlighted the urgent need for early Alzheimer's disease (AD) diagnosis through a deep learning-based system applied toMRI data, achieving classification accuracies of 80-90%. By leveraging Support Vector Machines (SVM) and various Deep Neural Network (DNN) models, they accurately distinguish between Mild Cognitive Impairment (MCI), AD, and Normal classes. Such advancements in computational tools hold promise for early detection of neurodegenerative diseases like Alzheimer's, offering significant clinical, social, and economic benefits. Koga et al. [17] developed a deep learning model using YOLOv3 for tauopathies differentiation, achieving high diagnostic accuracy. The block diagram in Fig. 1 illustrates a streamlined pipeline for processing neuroimaging data sourced from the ADNI dataset. Initially, the images undergo standard preprocessing steps such as resizing for uniformity, shuffling to eliminate biases, and the application of the SMOTE technique to address class imbalances. Following this, the dataset is partitioned into training and testing subsets. The training data is then subjected to classification using the Trained on CP13-immunostained digital slides, it accurately diagnoses AD, PSP, CBD, and PiD. The model's versatility extends to AT8-stained slides, suggesting its utility in neuropathological diagnoses. Illakiya et al. [18] explored the pivotal role of deep learning (DL) in AD detection, analyzing 103 research articles. Focusing on DL techniques like CNNs, RNNs, and TL, it examines neuroimaging modalities' potential for accurate AD diagnosis. The review highlights the need for deeper investigation into AD progression from MCI to AD using DL models. Helaly et al. [19] presented an end-to-end framework using CNNs for early AD detection and classification. Leveraging 2D and 3D brain scans from the ADNI dataset, it achieves promising accuracies. Amid the COVID-19 pandemic, a web application for remote AD checking was proposed, facilitating timely diagnosis. Serkan et al. [20] focused on using deep learning, specifically CNN architecture, to classify pre-illness stages of Alzheimer's disease from brain MRI images. EfficientNetB0 achieved the highest test accuracy of 92.98% among 29 pre-trained models. EfficientNet models, particularly EfficientNetB3, demonstrated superior precision, sensitivity, and specificity in Alzheimer's disease classification. Jansi et al. [21] utilized MRI images to accurately forecast Alzheimer's disease using deep learning techniques, namely transfer learning with InceptionV3. This system achieved an accuracy of 87.69%.

III. PROPOSED METHODOLOGY

EfficientNet B6 model, a convolutional neural network architecture tailored for image analysis tasks. The classification aims to categorize the images into five distinct groups related to cognitive health: cognitively normal (CN), mild cognitive impairment (MCI), Alzheimer's disease (AD), early mild cognitive impairment (EMCI), and late mild cognitive impairment (LMCI).

ISBN 979-8-3503-5094-4

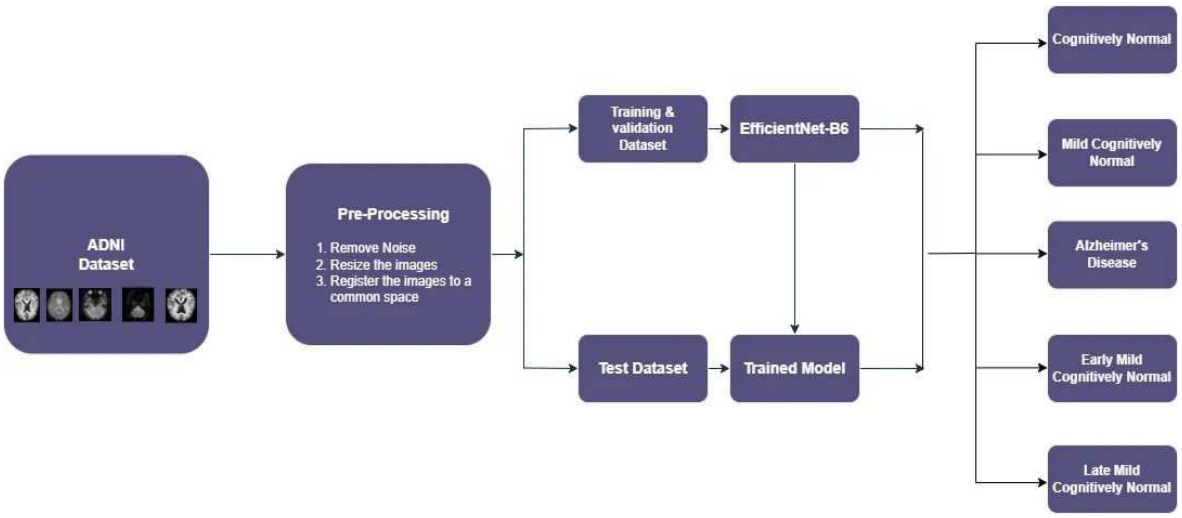


Fig. 1. Block diagram of the proposed Alzheimer's disease detection scheme

A. SMOTE technique

In Alzheimer's disease prediction, SMOTE (Synthetic Minority Over-sampling Technique) is a valuable strategy to address class imbalance issues commonly present in datasets, especially when instances of individuals with Alzheimer's disease are significantly fewer than those without. By generating synthetic samples for the minority class, SMOTE ensures a more balanced representation of both classes in the training dataset, which is crucial for developing accurate predictive models. This technique helps prevent model bias towards the majority class and improves the model's ability to correctly identify Alzheimer's disease cases.

B. Dataset Description

A sizable, openly accessible dataset called the Alzheimer's Disease Neuroimaging Initiative (ADNI) contains clinical, cognitive, genetic, and neuroimaging information from people with Alzheimer's disease, people with mild cognitive impairment (MCI), and those with cognitively normal abilities (https://adni.loni.usc.edu/). The dataset was developed to aid in studies on the early identification and monitoring of Alzheimer's progression of the ADNI dataset contains many kinds of imaging data, such as MRI, FDG-PET, and amyloid PET, as well as magnetic resonanceimaging (MRI). The number of sample images available in ADNI dataset is shown in Table 1.Sample images from the ADNI dataset for eachcategory is shown in Fig. 2.

TABLE I. DATA DESCRIPTION

Category	Number Of Images
LMCI	72
EMCI	240
MCI	932
CN	7430
AD	7536

ISBN 979-8-3503-5094-4

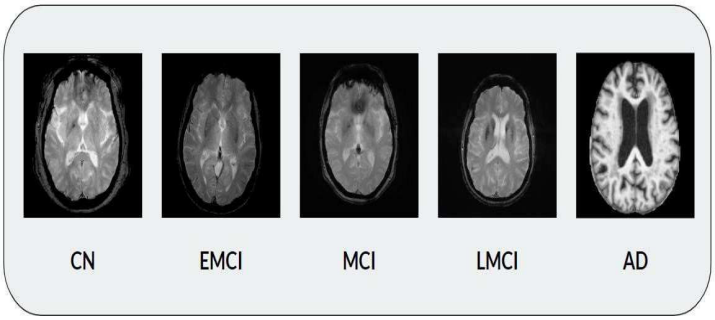


Fig. 2. Sample images from ADNI dataset

C. *Proposed Classification using EfficientNetB6*

The proposed classification method utilizes the EfficientNet B6 architecture, a convolutional neural network renowned for its efficiency and effectiveness in image classification tasks. It is shown in Fig. 3. The classification process begins with the input image, which undergoes a series of operations within the network's layers. Initially, the input image is processed through a convolutional layer followed by batch normalization and activation using the Swish function. This sequence helps extract features from the image while ensuring stable training through normalization and introducing non-linearity with the activation function.

The core of the EfficientNet architecture lies in the "MBConv6" blocks, which represent Mobile Inverted Residual Bottleneck Convolutional blocks. These blocks are repeated multiple times, enhancing the network's ability to capture intricate patterns and features from the input image. Following the MBConv6 blocks, another convolutional layercoupled with batch normalization and Swish activation is applied to further refine the extracted features.

A critical aspect of this proposed method is the introduction of a dense layer with a single neuron. This dense layer is a novel addition, particularly in the context of transfer learning. Transfer learning involves leveraging pre-trained models on large datasets and adapting them to specific tasks.

In this case, the dense layer is introduced to tailor the pre-trained EfficientNet B6 for the specific task of multi-class classification into five levels related to cognitive health.

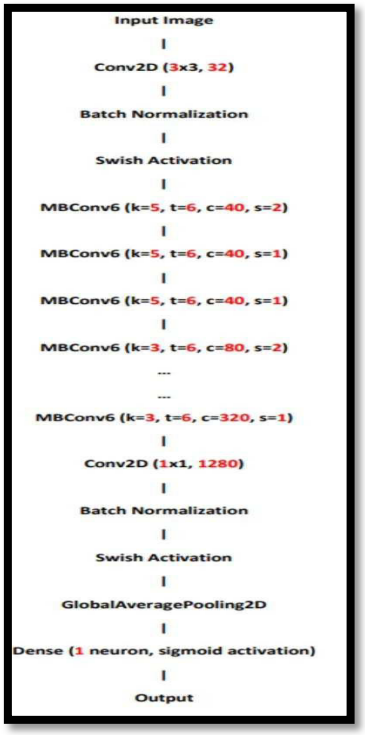


Fig. 3. Architecture Diagram of EfficientNetB6

IV. PERFORMANCE EVALUATION

The effectiveness of the model for the prediction of Alzheimer's disease is assessed using metrics like precision, recall, F1-score and accuracy. The Table 2 summarizes the performance evaluation metrics, including precision, recall, and F1-score, for a classification model across various classes. Precision values, indicating the accuracy of positive predictions, are notably high across all categories, ranging from 0.85 to 0.99. Similarly, recall values, measuring the model's ability to identify relevant instances within each class, are consistently strong, ranging from 0.92 to 0.99. F1-scores, representing the harmonic mean of precision and recall, further confirm the model's robustness, with scores ranging from 0.87 to 0.99.

ISBN 979-8-3503-5094-4

Overall, the model demonstrates high accuracy, effectively capturing instances across different classes, and maintaining a balanced performance between precision and recall metrics.

TABLE II. CLASSIFICATION REPORT OF TESTED DATA USING EFFICIENTNETB6

Category	Precision	recall	F1-score
AD	0.99	0.98	0.98
CN	0.99	0.99	0.99
EMCI	0.88	0.92	0.90
LMCI	0.89	0.93	0.91
MCI	0.85	0.89	0.87

Fig. 4 and Fig. 5 show the variation of system accuracy and loss respectively with increasing epochs. The training accuracy reaches a maximum value of 98.5%, indicating that the model correctly predicts the classes of nearly 98.5% of the training data samples by the end of the training process.

Similarly, the testing accuracy reaches a maximum of 97.78%, demonstrating the model's ability to generalize well to unseen data. On the other hand, the training loss decreases over epochs, reaching a minimum value of 0.1.

This indicates that the model's predictions become increasingly closer to the actual labels of the training data samples as training progresses. Likewise, the testing loss decreases to a minimum of 0.17, showing that the model's predictions also align well with the labels of the testing data, suggesting good generalization performance.

The proposed research focused on classifying Alzheimer's disease using the EfficientNet B6 architecture, introducing a novel approach by incorporating a dense layer for classification. The achieved overall accuracy of 97.78% and a low loss of 0.17 indicate the effectiveness of the proposed method in accurately identifying Alzheimer's disease cases.

The confusion matrix obtained for the proposed system is shown in Table III. The overall accuracy achieved by the proposed system is 97.78%.

TABLE III. CONFUSION MATRIX OF EFFICIENTNETB6

		AD	CN	EMCI	LMCI	MCI
ACTUAL	AD	1613	2	2	1	26
	CN	4	1751	4	2	8
	EMCI	2	2	73	0	2
	LMCI	0	0	2	25	0
	MCI	18	6	2	0	210
	PREDICTED					

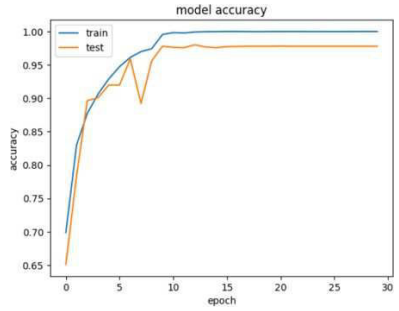


Fig. 4. Predicted Accuracy of EfficientNetB6 model

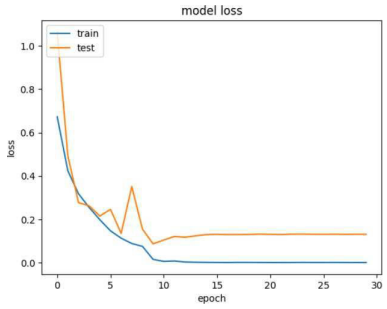


Fig. 5. Predicted Loss of EfficientNetB6 model

V. CONCLUSION

These results underscore the significance of leveraging advanced neural network architectures like EfficientNet B6 for medical image classification tasks. Furthermore, the novel addition of a dense layer enhances the model's capability to discern intricate patterns in neuroimaging data, contributing to improved classification performance.

ISBN 979-8-3503-5094-4

VI. REFERENCE

[1] A. Ebrahimi, S. Luo, and R. Chiong, "Deep sequence modelling for Alzheimer's disease detection using MRI," Comput. Biol. Med., vol. 134, no. May, 2021, doi: 10.1016/j.compbiomed.2021.104537.

[2] K. Lokesh, N. P. Challa, A. S. Satwik, J. C. Kiran, N. K. Rao, and B. Naseeba, "Early Alzheimer's Disease Detection Using Deep Learning," EAI Endorsed Trans. Pervasive Heal. Technol., vol. 9, no. 1, pp. 1–9, 2023, doi: 10.4108/eetpht.9.3966.

[3] R. Ibrahim, R. Ghnemat, and Q. Abu Al-Haija, "Improving Alzheimer's Disease and Brain Tumor Detection Using Deep Learning with Particle Swarm Optimization," AI, vol. 4, no. 3, pp. 551–573, 2023, doi: 10.3390/ai4030030.

[4] R. Jansi, R. Amutha, and A. Gokulakrishnan, "A novel framework for action recognition based on histogram of oriented gradients and sparsity-inducing shared dictionary," in Proceedings of the 2017 International Conference on Wireless Communications, Signal Processing and Networking, WiSPNET 2017, 2018, pp. 493–497. doi: 10.1109/WiSPNET.2017.8299805.

[5] S. Al-Shoukry, T. H. Rassem, and N. M. Makbol, "Alzheimer's diseases detection by using deep learning algorithms: A mini-review," IEEE Access, vol. 8, pp. 77131–77141, 2020, doi: 10.1109/ACCESS.2020.2989396.

[6] A. Survey, "applied sciences Alzheimer ' s Disease Diagnosis Using Machine Learning : A Survey," 2023. [7] C. P, H. A. F. J, and A. K, "Alzheimer's Disease Early Detection using Deep Learning Techniques," Ijarcce, vol. 12, no. 4, pp. 1–16, 2023, doi: 10.17148/ijarcce.2023.124145.

[8] P. Balaji, M. A. Chaurasia, S. M. Bilfaqih, A. Muniasamy, and L. E. G. Alsid, "Hybridized Deep Learning Approach for Detecting Alzheimer's Disease," Biomedicines, vol. 11, no. 1, 2023, doi: 10.3390/biomedicines11010149.

[9] S. Murugan et al., "DEMNET: A Deep Learning Model for Early Diagnosis of Alzheimer Diseases and Dementia from MR Images," IEEE Access, vol. 9, pp. 90319–90329, 2021, doi: 10.1109/ACCESS.2021.3090474.

[10] M. Mujahid, A. Rehman, T. Alam, F. S. Alamri, S. M. Fati, and T. Saba, "An Efficient Ensemble Approach for Alzheimer's Disease Detection Using an Adaptive Synthetic Technique and Deep Learning," Diagnostics, vol. 13, no. 15, pp. 1–20, 2023, doi: 10.3390/diagnostics13152489.

[11] Z. Zhao et al., "Conventional machine learning and deep learning in Alzheimer's disease diagnosis using neuroimaging: A review," Front. Comput. Neurosci., vol. 17, 2023, doi: 10.3389/fncom.2023.1038636.

[12] M. Khan, A. Mannan, K. Sherani, and M. U. Qayyum, "ISSN 3025-7409 (Media Online) Revolutionizing Healthcare : The Impact Of Artificial Intelligence On Patient Care , Diagnosis , And Treatment Issn 3025-7409 (Media Online)," vol. 1, no. 5, pp. 779–790, 2024.

[13] A. A, P. M, M. Hamdi, S. Bourouis, K. Rastislav, and F. Mohmed, "Evaluation of Neuro Images for the Diagnosis of Alzheimer's Disease Using Deep Learning Neural Network," Front. public Heal., vol. 10, no. February, p. 834032, 2022, doi: 10.3389/fpubh.2022.834032.

[14] S. Fathi, M. Ahmadi, and A. Dehnad, "Early diagnosis of Alzheimer's disease based on deep learning: A systematic review," Comput. Biol. Med., vol. 146, no. 4, p. 105634, 2022, doi: 10.1016/j.compbiomed.2022.105634.

[15] A. Pradhan, "Detection of Alzheimer ' s Disease (AD) in MRI Images using Deep Learning," Int. J. Eng. Res. Technol., vol. 10, no. 03, pp. 580–585, 2021.

[16] P. C. Muhammed Raees and V. Thomas, "Automated detection of Alzheimer's Disease using Deep Learning in MRI," J. Phys. Conf. Ser., vol. 1921, no. 1, 2021, doi: 10.1088/1742-6596/1921/1/012024.

[17] S. Koga, A. Ikeda, and D. W. Dickson, "Deep learning-based model for diagnosing Alzheimer's disease and tauopathies," Neuropathol. Appl. Neurobiol., vol. 48, no. 1, pp. 1–12, 2022, doi: 10.1111/nan.12759.

[18]T. Illakiya and R. Karthik, "Automatic Detection of Alzheimer's Disease using Deep Learning Models and Neuro-Imaging: Current Trends and Future Perspectives," Neuroinformatics, vol. 21, no. 2, pp. 339–364, 2023, doi: 10.1007/s12021-023-09625-7.

[19] H. A. Helaly, M. Badawy, and A. Y. Haikal, "Deep Learning Approach for Early Detection of Alzheimer's Disease," Cognit. Comput., vol. 14, no. 5, pp. 1711–1727, 2022, doi: 10.1007/s12559-021-09946-2.

[20] S. Savaş, "Detecting the Stages of Alzheimer's Disease with Pre-trained Deep Learning Architectures," Arab. J. Sci. Eng., vol. 47, no. 2, pp. 2201–2218, 2022, doi: 10.1007/s13369-021-06131-3.

[21] R. Jansi, N. Gowtham, S. Ramachandran, and V. S. Praneeth, "Revolutionizing Alzheimer's Disease Prediction using InceptionV3 in Deep Learning," pp. 1155–1160, 2024

ISBN 979-8-3503-5094-4

Tenth International Conference on Biosignals, Images and Instrumentation, SSNCE, Chennai, March 20 – 22, 2024

TRACK 6

I. A Novel Miniaturised Hexagonal-Shaped Patch Antenna for Medical Applications

Rashika. K
UG Scholar,
St. Joseph's Institute of Technology,
Chennai, Tamil Nadu, India 600119,
rashi.plrl@gmail.com

Thirisha. S
UG Scholar,
St. Joseph's Institute of Technology,
Chennai, Tamil Nadu, India 600119,
sekarthirisha646@gmail.com

Uthaykumar G S
Professor
St.Joseph's Institute of Technology
Chennai, India 600119,
uthayakumar.g.1969@ieee.org

Mohan. C
Assistant Professor,
St. Joseph's Institute of Technology,
Chennai, Tamil Nadu, India 600119,
mohanrc803@gmail.com

***Abstract*— The creation of a hexagon-shaped patch antenna for Sub-6GHz 5G communications is presented in this study. For 5G wireless applications, the suggested antenna can resonate at the center frequency of 6 GHz. The proposed antenna features a hexagonal design, multiple radiating slots with partial ground and is fed with a microstrip feedline. It measures 17.5 × 22.2 × 1.6 mm^3 and operates on the N102 band at 6 GHz. Return loss, VSWR, peak gain, and impedance bandwidth are all elements of the functionality of the suggested antenna. The proposed antenna employs slots that cover the frequency range of 5.92 GHz to 6.35 GHz. With a 6.1 GHz resonance frequency, the suggested antenna's reflection coefficient (S_{11}) is 44.6 dB, with a peak gain of roughly 3.2 dB. Thus, the suggested antenna can be used for 5G wireless applications operating at 6 GHz.**

Keywords—patch antenna, inset feed, ISM band, dual band.

II. Introduction

5G technology is going to be recognised as the norm for mobile communication networks shortly. An increase in users causes a lack of bandwidth. Increased bandwidth is required for accommodating large end users, which enhances the system's overall data rates. End users can communicate faster and more sensibly because to the high data rate [1]. Many industries, including realistic Ultra High Definition, Artificial Intelligence, Blockchain, and Internet of Things services like Smart Cities, Smart Transportation, and Smart Grids, will be greatly enhanced by the significant rise in mobile data in 5G. Its frequency provides high quality from sub-6 GHz to millimetre waves. The 5G frequency spectrum offers enhanced coverage, decreased fading, and better data rates. Antennas designed for 5G operation perform significantly better in the sub-6 GHz frequency range, raising the standard generally [2].

Considering the specifications for 5G, antennas that are cheap, easy to manufacture in large quantities, conformable to a planar surface, lightweight, low-profile, and conformable in size must be used . The ideal choice to fulfil all the aforementioned requirements is a microstrip patch antenna [3]. Excellent option for 5G communication systems due to its inherent advantages is a microstrip patch antenna. The ability to reduce an antenna's physical dimensions without noticeably affecting performance has attracted a lot of interest to antenna miniaturisation techniques in recent years [4].

The literature lists several methods for achieving antenna miniaturization, including slots, fractals, loops, Transmission Lines (TLs), Defected Ground Structures (DGS), Composite Left-Right Left-Handed (CLRH), etc., despite the fact that this is a difficult and intimidating task. For Sub-6GHz 5G applications, a number of size-reduced antennas have already been designed [5,6]. For instance, studies is being done on an extremely small microstrip antenna designed for sub-6 GHz 5G communications [7]. The size is decreased by combining an L-shaped patch radiator with a defective ground plane that has two symmetrical stubs. The overall size of the antenna is reduced to 30 x 20 mm. The combination of T-shaped elements, passive elements, tapering, and meandering techniques is introduced by [8] for miniaturisation. The antenna, measuring 50 × 19.75 mm, provides optimal impedance matching and covers all sub-6 GHz bands. The creation and deployment of a broad-band, rose-shaped patch antenna for wireless applications operating at Sub-6 GHz [9]. The antenna design covers the frequency range (2.7-5.74 GHz) is made from the Rogers RT5880 substrate material, which has a low volume of 20 × 35 mm and a bandwidth of 3.04 GHz.

The design of a compact patch antenna for sub-6 GHz uses in 5G mobile terminals is described. The proposed antenna is comprised of multiple driven element strip and three parasitic grounded strips, and it has a simple layout that doesn't require lumped parts or a three-dimensional structure [10]. The 0.7-0.96 GHz and 1.6-5.5 GHz frequencies are covered by the -6dB impedance bandwidth, which has a compact design of 40 x 15 mm. With help of metamaterial ground plane, the dimensions of a compact TSRMPA are 30.2 x 36.4 mm. The modified rectangular patch with metamaterial ground plane resonated at three bands 3.27, 3.78, and 3.92 GHz [11]. A small printed monopole patch antenna with finite ground is utilized to demonstrate Sub-6 GHz 5G communications [12]. The size of the this antenna desires to measure 30 mm x 34 mm. The optimisation of the partial ground plane and stubs allows for accurate impedance matching over a wide range (3-7 GHz). For a 5G application, a microstrip patch with a T-shaped rectangular antenna patch operates at a 3.6 GHz resonant frequencies from 2.9 to 4.4 GHz, respectively [13].

This paper demonstrates the analysis of a hexagonal-shaped antenna for Sub-6GHz 5G communications. The miniaturization of the proposed antenna has been achieved through the combination of its hexagonal-shaped and multiple radiating slots. The overall volume of the designed antenna is approximately 17.5 x 22.2 x 1.6 mm^3. The developed antenna has a bandwidth of 1.1 GHz (between 5.4 GHz and 6.5 GHz), which is fairly large and makes it an appropriate design. With a suitable design, the peak gain values are nearly 3.2 dB. Peak gain values can reach approximately 3.2 dB when correctly designed. Furthermore, important characteristics which include VSWR, surface current density, and radiation patterns are calculated and compared. The paper is structured in the following fashion: Section 2 presents the design and analysis of the hexagonal patch antenna. Section 3 contains the results of the investigation. Section 4 contains the paper's conclusion.

III. Antenna Design & Methodology

Using the basic design equation for a patch antenna with a full ground, the designed antenna's initial design is calculated.

ISBN 979-8-3503-5094-4

$$f_r = \frac{c}{4\pi r\sqrt{\varepsilon_{eff}}}$$

Where *"r"* is the circular patch's diameter, *"c"* is the light's velocity, and *"ε_{eff}"* is the equivalent dielectric constant [14]. With a dielectric constant of 4.4, the designated antenna has been constructed on a FR-4 substrate. As a result, the diameter of the circular patch is fixed at 15.9 mm for f_r = 6 GHz. To accomplish the innovation, the circular patch is changed to hexagonal shape, and to minimize the size of the antenna, different size slots are cut on hexagonal shape, and the ground plane is also reduced partially. The resulting antenna design operates at 6 GHz in a narrow band. Figure 1 depicts the formation of the suggested antenna. The modelling of the intended antenna design is examined using CST simulator software.

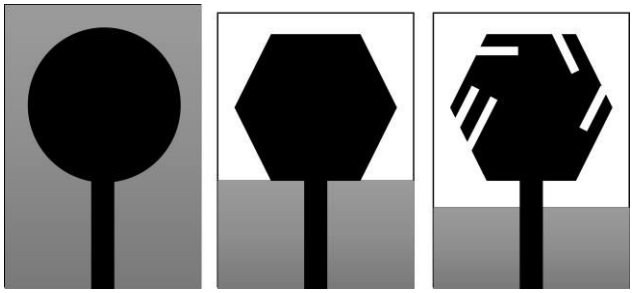


Fig 1. Design and evolutions of the proposed antenna.

The substrate's overall dimensions are 17.5 mm x 22.2 mm. The substrate layer is 1.6 mm thick and lies over the ground surface. The feed configuration of the antenna measures 2.4 mm in width and 8.5 mm in length. The suggested antenna's geometrical layouts are shown in Figure 2. The suggested antenna's construction specifications are listed in Table 1.

TABLE I

Parameters of the Proposed Antenna

Parameter	Size (mm)	Parameter	Size (mm)
W_S	17.5	L_G	4.3
L_S	22.2	W_F	2.4
r	6.6	L_F	11.5

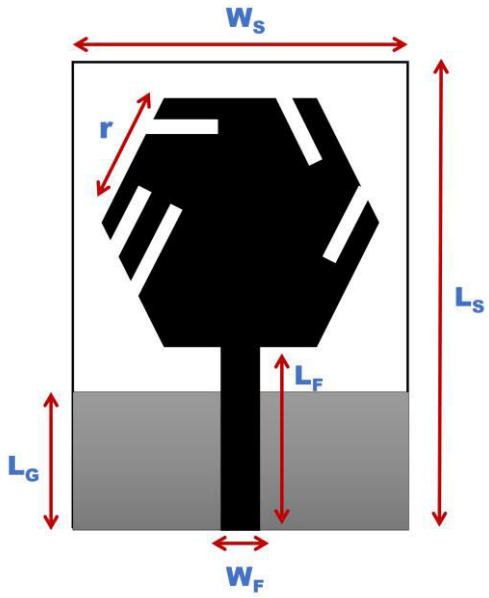


Fig 2. Total geometry and structure of the designed antenna.

IV. Results & Analysis

The bandwidth, radiation characteristics, and antenna-normalized gain have all been part of the antenna's performance. The overall estimated and actual reflection values for the suggested antenna are displayed in Figure 3. The antenna provides the reflection coefficients (S_{11}) around -46 dB at 5.9 GHz. This hexagonal-shaped antenna's observed impedance bandwidth in the simulation is around 500 MHz (5.6 - 6.1 GHz) at 5.9 GHz centre frequency. The required frequency is perceived as a suitable impedance bandwidth. The feeding losses may be the cause of the minor variation in the reflection coefficient that can be observed in the results. However, the proposed antenna's achieved bandwidth is sufficient for the targeted applications.

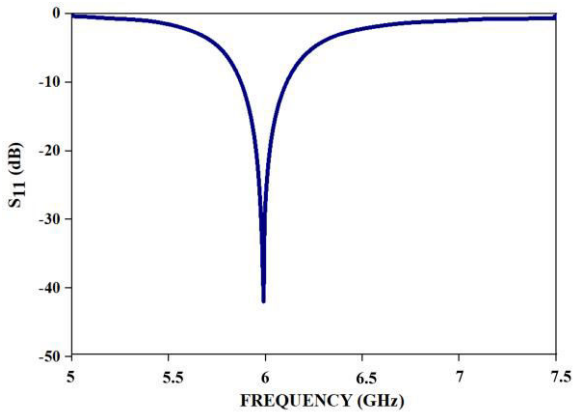


Fig 3. Reflection coefficient diagram for the intended antenna.

Figure 4 displays the designed antenna's current flowing at a resonance frequency of 5.9 GHz. It proved that by etching several slots on the hexagonal patch, the current flow maintains its path inside the antenna surface and improves radiation performance. This shows that at 5.9 GHz, the current is primarily dispersed in the outer regions of the various slots. These findings clearly demonstrate that combination slots extend the overall electrical length while

ISBN 979-8-3503-5094-4

disturbing the flow of current. As a result, the antenna's resonance frequency is reduced without a change in patch size. The suggested antenna is developed and evaluated based on these antenna performance.

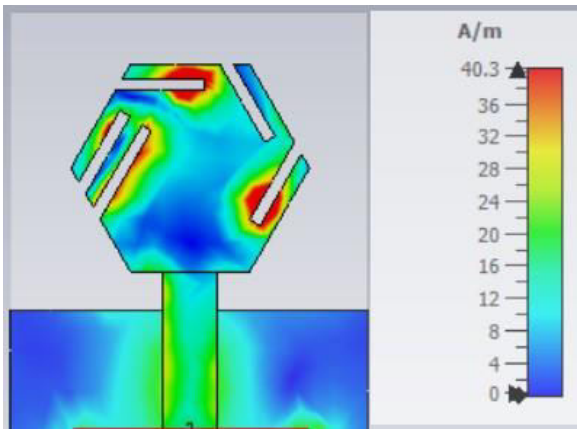
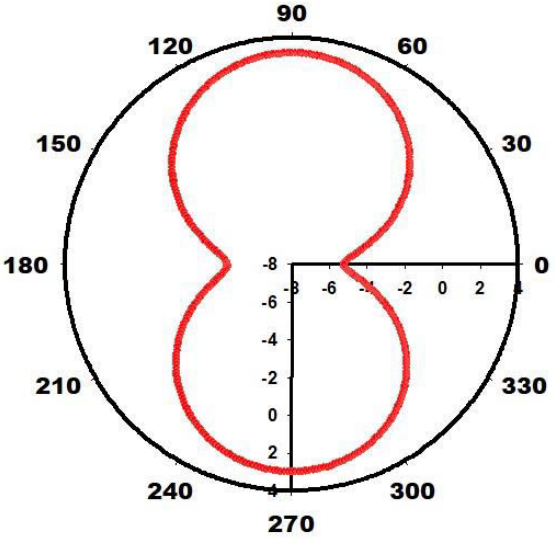
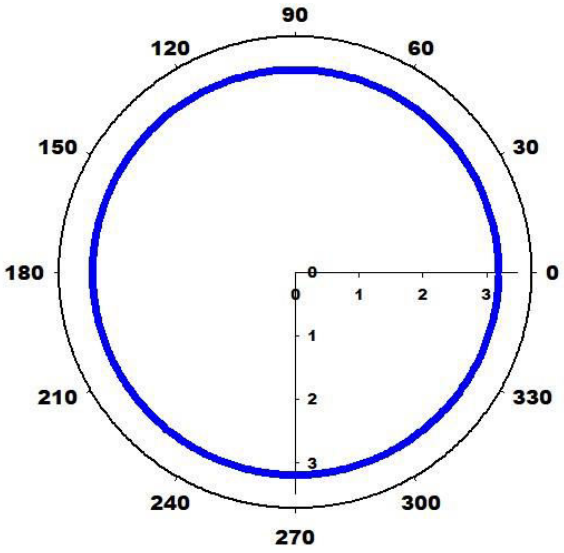


Fig 4. Surface current density on proposed antenna at 5.9 GHz.

The suggested hexagonal patch antenna's radiation pattern is simulated at 5.9 GHz. Figures 5 a) and b) indicate that the optimum simulated radiation results are obtained in both the xy and xz planes at the frequency of 5.9GHz. The results shown here demonstrate that the antenna has a roughly bidirectional layout in the xy plane and an omnidirectional layout in the xz plane. The radiated fields in the yz-plane and xz-plane are standardized to the desired band at each frequencies.



(a)



(b)

Fig 5. The analysis of the intended antenna's radiation pattern at 5.9 GHz.

Figure 6 illustrates the suggested antenna's estimated radiation efficiency and maximum gain at different frequencies. The outcomes show that the antenna performs at its working frequency with greater effectiveness. The designed antenna obtained 76 % radiation efficiency and 3.1 dB gain at the required frequency of 5.9 GHz.

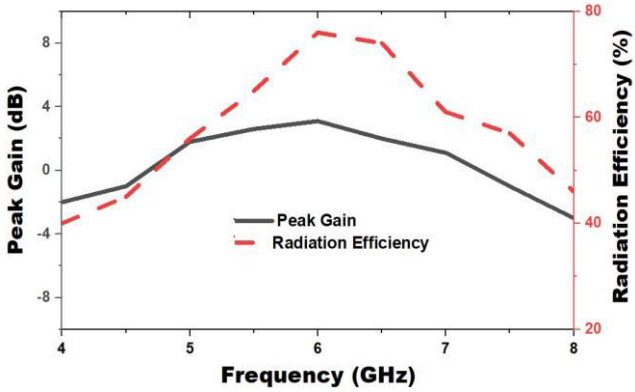


Figure 6. Obtained gain and radiation efficiency of the intended antenna.

Additionally, comparisons and analyses of this antenna's dimensions, normalised gain, bandwidth, and percentage of miniaturisation are accomplished. The designed antenna has a total size that is 56 % less than a conventional antenna and obtains a gain of 3.1 dB at 5.9 GHz, which makes it much more suited for 5G applications.

V. Conclusion

In this paper, a basic, gain-enhanced, small hexagonal shaped microstrip antenna with various slots is examined. A hexagonal patch with several slots and a partial ground plane in the shape of a hexagon forms a component of the proposed antenna. Better gain and reflection coefficients at 5.9 GHz are provided by the suggested design. By employing multiple slot configurations, the recommended antenna reduces its size by approximately 56% when compared to a traditional antenna. Thus, 5G applications operating at sub-GHz frequencies benefit from the suggested antenna.

ISBN 979-8-3503-5094-4

REFERENCES

1. Dilsha M.D., Neenu Welson, 2015, Introduction to 5G Wireless Technology, International Journal of Engineering Research & Technology (IJERT), 2015 (Volume 4 – Issue 06)
2. K. Jain, R. Acharya, S. Jakhar and T. Mishra, "Fifth Generation (5G) Wireless Technology “Revolution in Telecommunication”," 2018 Second International Conference on Inventive Communication and Computational Technologies (ICICCT), Coimbatore, India, 2018, pp. 1867-1872
3. N. Sekeljic, Z. Yao and H. -H. Hsu, "5G Broadband Antenna for sub-6 GHz Wireless Applications," 2019 IEEE International Symposium on Antennas and Propagation and USNC-URSI Radio Science Meeting, Atlanta, GA, USA, 2019, pp. 147-148,
4. C. Mohan & S. Esther Florence (2022) Miniaturised Triangular Microstrip Antenna with Metamaterial for Wireless Sensor Node Applications, IETE Journal of Research, 68:2, 1177-1182.
5. Khan, M.U., Sharawi, M.S. and Mittra, R. (2015), Microstrip patch antenna miniaturisation techniques: a review. IET Microw. Antennas Propag., 9: 913-922.
6. M. Fallahpour and R. Zoughi, "Antenna Miniaturization Techniques: A Review of Topology- and Material-Based Methods," in IEEE Antennas and Propagation Magazine, vol. 60, no. 1, pp. 38-50, Feb. 2018.
7. P. Pradeep, J. S. Kottareddygari and C. Sekhar Paidimarry, "Design of a Novel Compact L-slotted Monopole Antenna for 5G Applications in sub-6 GHz Band," 2022 3rd International Conference on Electronics and Sustainable Communication Systems (ICESC), Coimbatore, India, 2022, pp. 394-397.
8. N. Sekeljic, Z. Yao and H. -H. Hsu, "5G Broadband Antenna for sub-6 GHz Wireless Applications," 2019 IEEE International Symposium on Antennas and Propagation and USNC-URSI Radio Science Meeting, Atlanta, GA, USA, 2019, pp. 147-148.
9. L. C. Paul, S. C. Das, M. N. Hossain and W. -S. Lee, "A Wideband Rose-shaped Patch Antenna with a Ground Slot for Sub-6 GHz Applications," 2021 IEEE Indian Conference on Antennas and Propagation (InCAP), Jaipur, Rajasthan, India, India, 2021, pp. 901-904.
10. M. Vinoth and R. Vallikannu, "A compact triple slotted Rectangular Microstrip Patch Antenna with Metamaterial ground for Sub-6 GHz/5G communication," 2020 Fifth International Conference on Research in Computational Intelligence and Communication Networks (ICRCICN), Bangalore, India, 2020, pp. 34-38.
11. U. Farooq et al., "A Compact Monopole Patch Antenna for Future Sub 6 GHz 5G Wireless Applications," 2020 IEEE International Symposium on Antennas and Propagation and North American Radio Science Meeting, Montreal, QC, Canada, 2020, pp. 1715-1716.
12. Abdulbari, Ali Abdulateef et al. Design compact microstrap patch antenna with T-shaped 5G application. Bulletin of Electrical Engineering and Informatics, v. 10, n. 4, p. 2072-2078, 2021.
13. Marineni Vamsi , Yampati Venkata Rajesh , Vijay Kuruba , Ramella Eswara Sai Krishna, Dr. H.Umma Habiba, 2022, Design and Implementation of Antenna for Sub-6 GHz Applications in 5G Mobile Terminals, International Journal of Engineering Research & Technology (IJERT), Volume 11, Issue 05.
14. Kaur, G., Kaur, A., Toor, G. K., Dhaliwal, B. S., & Pattnaik, S. S. (2015). Antennas for biomedical applications. Biomedical Engineering Letters, 5(3), 203-212.

ISBN 979-8-3503-5094-4

Mitochondrial Network Segmentation from Fluorescence Microscopy Images using VGG-UNet

Ramya Mohan
Department of Computer Science and Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences
Chennai 602105, TN, India.
ORCID: 0000-0003-3062-5012

S. Prabha
Department of CSE
Center for Research and Innovation
Saveetha School of Engineering, SIMATS, Chennai 602105, TN, India.
Email: sprabha3223@ieee.org

V. Rajinikanth
Department of CSE
Center for Research and Innovation
Saveetha School of Engineering, SIMATS, Chennai 602105, TN, India.
Email: v.rajinikanth@ieee.org

***Abstract*—Automatic medical image examination plays a vital part in disease screening and decision making processes. This research aims to create deep-learning (DL) assisted Mitochondrial Network (MN) segmentation from the chosen Fluorescence Microscopy Images (FMI) with better accuracy. Several parts of this arrangement encompasses the following; (i) Image collection and resizing, (ii) extraction of the MN section using a chosen scheme, (iii) comparing the segmented section against the ground-truth (GT) and computing the performance metrics, and (iv) performance evaluation and validation of the proposed technique. In this work, a comparative analysis is presented using the UNet and its variants. The experimental findings affirm that the VGG-UNet-supported scheme outperforms alternatives, yielding an overall segmentation accuracy of >97%. This underscores the efficacy of the proposed methodology in advancing the automated examination of MN for the chosen FMI database.**

Keywords—Fluorescence microscopy, Mitochondrial Network, Segmentation, VGG-UNet, Accuracy.

I. Introduction

Mitochondria (MITO) are extremely dynamic organelles which morphologically adjust to meet cellular demands. The MITO-network (MN) dynamically modifies in response to several cellular activities, like metabolism, apoptosis, mitosis, and MITO-DNA imitation. This flexibility guarantees functions, like energy production, calcium buffering, and cell survival. MITO's fission and fusion activities control its network morphology. Due to its significance, a number of research works are proposed to examine the MN structures and provides a necessary analysis to verify its clinical importance [1].

The clinical level examination of MN is commonly performed using the medical imaging technique, like the microscopic examination [2-4]. Examination of the MN using the Fluorescence Microscopy Image (FMI) is one of the common practices and the earlier works related to this can be accessed from [5]. Automatic examination of the MN is essential for timely and accurate examination of its Morphology and hence, earlier research works implemented Artificial-Intelligence (AI) tool to examine the FMI to extract and evaluate the MN [6]. The work of Fischer et al. (2020) proposed a deep-learning (DL) tool using to extract and evaluate the MN structure and achieved a better result [7]. This work confirms that the DL methods are effective in accurately extracting the MN network from the chosen FMI.

Automatic examination of the FM images using a chosen DL-segmentation scheme is reported in the work of Daniel et al. (2022) proposed VGG-UNet/VGG-SegNet to examine the FMI of Endoplasmic-Reticulum Network (ERN) and obtained a better detection accuracy when VGG-SegNet is implemented [8].

This research aims to execute the DL-segmentation procedure on the chosen FM image database available in [9]. Different parts of this tool includes; (i) image resizing and pre-processing to convert raw-image to processed-image, (ii) implementation of DL-segmentation to extract the MN from the chosen FM image, and (iii) relating the segmented sector with the binary-mask (BM) to verify the performance. This research proposed a VGG-UNet scheme using VGG16 and VGG19 as the backbone and verifies the segmentation performance on the chosen image datasets.

The proposed scheme initially implements the VGG16 encoder-decoder section, in which the encoder extracts the image feature and then the decoder unit reconstructs the image from these features, which is then finally extracts the binary class image using the Sigmoid activation unit, kept as the final section of the proposed model. Similar practice is then considered for the VGG19-encoder and the achieved binary images are considered for further analysis.

Initially, the raw FM images are converted into 8-bit depth image to improve the visibility of the MN in the chosen image and then it is resized to 512×512 pixels for the effective examination and extraction. During the segmentation task, the pre-trained DL methods, like UNet, UNet+, and UNet++ are employed and the segmented section is then compared with the proposed VGG-UNets The performance of the developed tool is verified using the metrics, like Jaccard (JA), Dice (DI), accuracy (AC), and precision (PR).

The outcome of the planned arrangement authorizes that the implemented technique with VGG-UNet having VGG16 as backbone helps to achieve an overall segmentation accuracy of >97% with better JA, DI and PR values compared to other methods in this study.

The contribution of this research includes;

a) Proposing a DL-segmentation using VGG-UNet with VGG16 and VGG19 as the backbone (Encoder-Decoder),

b) Performance evaluation and verification of proposed technique with other existing UNet models.

The following sections outline the structure of this work: Section 2 provides the motivation, Section 3 details the methodology, and Sections 4 and 5 present the investigational outcomes and conclusions, respectively.

ISBN 979-8-3503-5094-4

II. MOTOVATION

Examination of the MN from the FMI is a common practice and it helps in understanding a variety of conditions, like MITO diseases, neurodegenerative disorders, cancer, and aging [10,11].

A systematic inspection of MN morphology permits clinicians to make well-informed judgments about disease management. Through the investigation of MITO structure and organization in cells, clinicians can appraise cellular health, metabolic activity, and detect abnormalities linked to different disorders. This data is vital for choosing appropriate treatment plans and tracking disease progression. Moreover, MN studies offer insights into disease mechanisms, aiding in the development of novel therapeutic strategies and personalized medicine approaches.

Due to its significance, recently DL-segmentation procedure is implemented to analyse the MN from FMI and achieved a better result as discussed by Fischer et al. (2020) [7]. This work motivates to develop an appropriate DL-segmentation technique to extract and evaluate the MN morphology. Further, the work of Daniel et al. (2022) [8] proposed FMI evaluation with the VGG-based segmentation schemes and obtained an accuracy of >98%. Based on the motivation from these earlier works, this research proposes a DL-segmentation scheme using the VGG-UNet with VGG16- and VGG19-backbone and verifies the performance of the system using the FMI available in [9].

III. METHODOLOGY

This part of the research demonstrates the methodology proposed to examine the chosen FMI database.

Figure 1 shows the structure of this scheme. Initially, the collected FMI and the BM are resized to 512 × 512 pixels and then the FMI is converted into an 8-bit image to improve the visibility of the MN region. This image is then considered to train and verify the performance of the developed VGG-UNet scheme. The achieved outcome from the proposed scheme is in binary form, which is then compared against the BM and the essential performance metrics are computed as depicted in the proposed architecture.

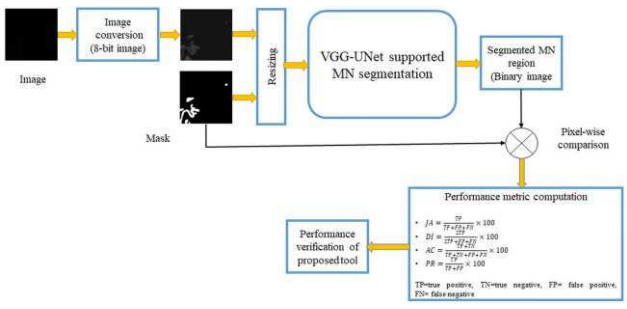


Figure 1. Examination of MN from FMI using VGG-UNet

A. FMI database

This work considers the FMI available in [9] for the analysis. The information regarding the chosen images are as follows; the FMI of the MN was conducted using live cultured cells and a widefield microscopy setup. MITO cell imaging was carried out on a Nikon Eclipse Ti-E inverted microscope equipped with a CoolSNAP HQ2 camera from Photometrics, using a 100x/1.40 numerical aperture (NA) objective lens. The binary mask for the FMI was created using ITK-SNAP [12], and each raw image was resized to 256x256 pixels.

This work considers 253 images for the study and to train the proposed DL-model, these images are augmented to get 1000 images (image horizontal-flip, vertical-flip, and rotate 45°) and these images are then considered to train and verify the performance of the proposed scheme. After training the DL-model on the chosen database, all the original images (253) are then examined and the achieved overall result (mean ± standard deviation) is considered to verify the merit of the developed scheme.

The sample image and the BM considered in this study can be found in Figure 2. Visibility MN in processed-image is good compared to the raw-image.

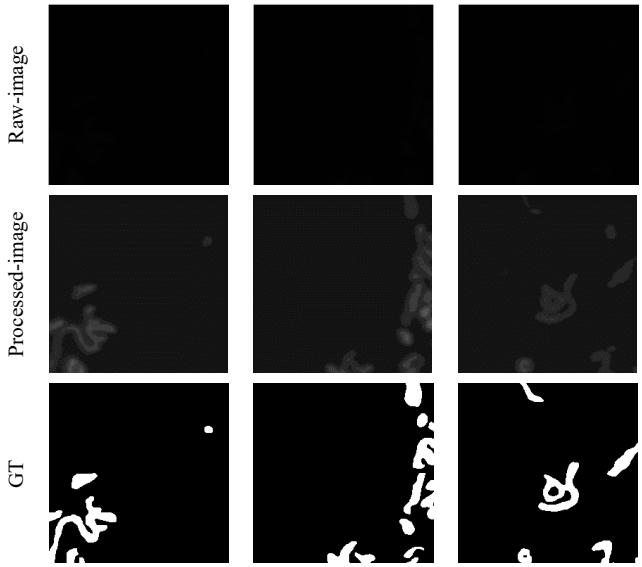


Figure 2. Sample images of the chosen FMI database

B. VGG-UNet

The DL-model of VGG (VGG16 and VGG19) is one of the commonly adopted approaches for the classification task. A modification in the architecture with encoder-decoder unit helps to form a VGG-UNet scheme, which provides superior image segmentation task as discussed in the earlier works [7].

Figure 3 presents the VGG16 based architecture, similar scheme is implemented using the VGG19. Initially, the training is performed using 100 epochs and the testing operation is implemented with the 50 epochs. Other parameters are as follows; learning rate=0.001, optimizer= Adam and activation= ReLU and Sigmoid.

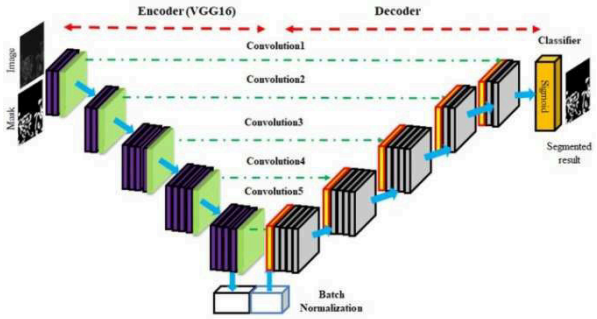


Figure 3. VGG-UNet scheme with VGG16 as backbone

ISBN 979-8-3503-5094-4

C. *Performance Verification*

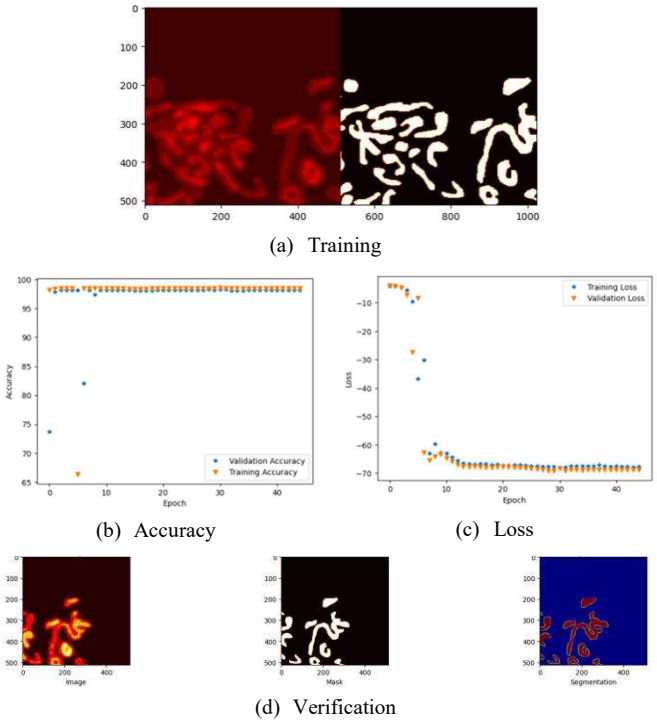
The performance of the developed tool is verified using the measures depicted as in Figure 1. After extracting the binary image with the VGG-UNet, its merit is verified using a comparative evaluation against the BM as depicted in Figure 1. During this task, the necessary measures, like JA, DI, AC and PR are computed and based on this value; the tool's merit is verified.

IV. Result and Discussion

This research section showcases the investigational results of the study. The images were converted to 8-bit depth and resized to an appropriate level, then verified using the proposed VGG-UNet model implemented in Python. The experiments were conducted on a workstation with the following specifications: Inteli5, 20GB of RAM, and 4GB of VRAM.

Initially, VGG-UNet with VGG16-backbone is implemented to verify the MN examination using the FMI database. Initially, 1000-training images are considered and then 256 images are adopted during the testing task.

The experimental outcome obtained during this method is shown in Figure 4. Fig 4(a) presents the training result (hot-colourmap) and Fig 4(b) and (c) presents the values of accuracy and loss for a chosen epoch value. These two results confirm that the implemented scheme works well on the chosen database and helps to achieve better segmentation accuracy. The segmented image with the implemented technique is presented in Fig 4(d) using the hot- and jet-colourmap. This confirms that the proposed DL-segmentation provides a better outcome on the chosen FMI database.



(a) Training

(b) Accuracy (c) Loss

(d) Verification

Figure 4. Experimental outcome of VGG-UNet with VGG16-backbone

Similar task is then repeated with the VGG-UNet with VGG19-backbone and the obtained results are recorded. Other chosen approaches, like UNet, UNet+, and UNet++ also verified with similar procedure and the obtained outcome (mean±SD) is considered for the evaluation as in Table I.

TABLE I. Overall Performance of the Implemented DL-Segmentation Schemes

Scheme	JA	DI	AC	PR
UNet	84.18±2.16	91.18±0.49	93.17±0.08	94.05±0.26
UNet+	86.33±1.21	91.82±1.04	94.04±0.17	95.13±0.02
UNet++	85.71±1.05	91.39±0.18	93.82±029	95.11±0.28
VGG-UNet (VGG16)	87.04±1.13	94.01±0.32	97.48±0.21	97.91±0.06
VGG-UNet (VGG19)	86.97±1.06	93.71±1.07	97.05±0.36	97.55±0.13

Table I confirms that the overall result obtained for the chosen 253 images provides a detection accuracy >93%. The traditional UNet and its variants provide satisfactory segmentation accuracy on the chosen FMI database and helps in extracting the MN section accurately. The proposed scheme with the VGG16- and VGG19-backbone helps in achieving a better overall result compared to the other similar schemes considered in this research. Further, this work utilized the pre-trained VGG models, which is easy to modify and implement for the chosen database.

The segmentation accuracy of VGG16-backbone scheme is better and it confirms that, it helps in achieving a better segmentation accuracy of the MN from the chosen FMI dataset. The chief merit of this research is, the implementation of the VGG-UNet is quite simple and it includes the VGG16 as the encoder to extract the features and its modified form as the decoder to reconstruct the image from the features.

The major limitation of this work is, it utilized a two-step process; (i) training process using the augmented image data (1000 images) with an epoch value of 100, and (ii) testing with original image (253) using an epoch value of 50. Further, this tool needs an 8-bit image data and an image conversion is initially needed. The outcome of this research authorize the significance of the tool on the chosen database and provides an accuracy of >97%. In future, VGG-SegNet based tool can be proposed to extract the MN from FMI and its performance can be verified with the developed VGG-UNet technique.

V. Conclusion

Examination of the MN morphology is one of the clinically significant processes for cell level evaluation of the diseases. This work proposes a DL-segmentation procedure using VGG-UNet to extract the MN structure from the chosen FMI database. To accomplish improved accuracy, this work proposed VGG16 and VGG19 based encoder-decoder assembly to examine the FMI data. The proposed image is initially converted into an 8-bit image and examined with the proposed scheme and its outcome is verified with the traditional UNet and its variants. The investigational outcome of this study confirms that the VGG16- supported approach provides a better accuracy (>97%), which is superior compared to other schemes. In future, VGG-SegNet can be proposed to analyse the chosen FMI data and its performance can be verified against the results of this study.

ISBN 979-8-3503-5094-4

REFERENCES

[1] Hoitzing, H., Johnston, I. G., & Jones, N. S. (2015). What is the function of mitochondrial networks? A theoretical assessment of hypotheses and proposal for future research. *Bioessays*, *37*(6), 687-700.

[2] Chu, C. H., Tseng, W. W., Hsu, C. M., & Wei, A. C. (2022). Image analysis of the mitochondrial network morphology with applications in cancer research. *Frontiers in Physics*, *10*, 855775.

[3] Ojeme, B., Quinn, F., Karls, R., & Quinn, S. (2023). Fully Automated Methods for the Detection and Segmentation of Mitochondria in Microscopy Images. *International Journal of Mechanical and Materials Engineering*, *17*(2), 42-47.

[4] Choi, J., Lee, H., Kim, S., Kwon, S. K., & Jeong, W. K. (2021). MitoVis: A Visually-guided Interactive Intelligent System for Neuronal Mitochondria Analysis. *arXiv preprint arXiv:2109.01351*.

[5] Kuznetsov, A. V., Hermann, M., Troppmair, J., Margreiter, R., & Hengster, P. (2010). Complex patterns of mitochondrial dynamics in human pancreatic cells revealed by fluorescent confocal imaging. *Journal of cellular and molecular medicine*, *14*(1-2), 417-425.

[6] Lefebvre, A. E., Ma, D., Kessenbrock, K., Lawson, D. A., & Digman, M. A. (2021). Automated segmentation and tracking of mitochondria in live-cell time-lapse images. *Nature Methods*, *18*(9), 1091-1102.

[7] Fischer, C. A., Besora-Casals, L., Rolland, S. G., Haeussler, S., Singh, K., Duchen, M., ... & Marr, C. (2020). MitoSegNet: easy-to-use deep learning segmentation for analyzing mitochondrial morphology. *Iscience*, *23*(10).

[8] Daniel, J., Rose, J. T., Vinnarasi, F., & Rajinikanth, V. (2022). VGG-UNet/VGG-SegNet supported automatic segmentation of endoplasmic reticulum network in fluorescence microscopy images. *Scanning*, *2022*.

[9] Yaoru Luo, Yuanhao Guo, Wenjing Li, Guole Liu, Ge Yang. (2020). Fluorescence Microscopy Image Datasets for Deep Learning Segmentation of Intracellular Orgenelle Networks. IEEE Dataport. https://dx.doi.org/10.21227/t2he-zn97

[10] Zahedi, A., On, V., Phandthong, R., Chaili, A., Remark, G., Bhanu, B., & Talbot, P. (2018). Deep analysis of mitochondria and cell health using machine learning. *Scientific reports*, *8*(1), 16354.

[11] Fogo, G. M., Anzell, A. R., Maheras, K. J., Raghunayakula, S., Wider, J. M., Emaus, K. J., ... & Sanderson, T. H. (2021). Machine learning-based classification of mitochondrial morphology in primary neurons and brain. *Scientific reports*, *11*(1), 5133.

[12] http://www.itksnap.org/pmwiki/pmwiki.php

ISBN 979-8-3503-5094-4

Automatic Skin Melanoma Detection from Dermoscopy Images using ResNet Variants

V. Rajinikanth
Department of CSE
Center for Research and Innovation
Saveetha School of Engineering,
SIMATS, Chennai 602105, TN, India.
Email: v.rajinikanth@ieee.org

A. Rama
Department of Computer Science and Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences
Chennai 602105, TN, India.
ORCID: 0000-0002-5217-0680

Ramya Mohan
Department of Computer Science and Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences
Chennai 602105, TN, India.
ORCID: 0000-0003-3062-5012

***Abstract*— Skin cancer is serious, requiring early detection. Dermatoscopy captures skin images, examined by dermatologists to detect abnormalities indicating cancer. This is crucial for distinguishing between benign and malignant lesions. The proposed work aims to develop a deep-learning-based methodology to classify the chosen DD image dataset into benign and malignant classes. The stages of this tool include; (i) collecting and resizing , (ii) feature extraction with ResNet variants, (iii) feature reduction using 50% dropout and features fusion, and (iv) binary classification and verification. In this work, the tool is developed using the ResNet variants, and its performance is verified using the conventional, reduced, and fused features. During this task, the SoftMax classifier-supported binary classification is executed, and the tool's performance is verified based on the achieved accuracy. This study confirms that the fused deep features with ResNetV2 (RNV2) variants RNV2_50 and RNV2_152 help to provide an accuracy of 98%. This outcome proves the developed tool works well on the chosen DD images.**

Keywords—Skin cancer, dermatoscopy, ResNetV2, features fusion, classification.

I. Introduction

The recent report, Globocan2020, confirms a gradual increase in the global occurrence rate of skin cancer (SC), including melanoma, despite the implementation of numerous preventive measures [1]. This trend underscores the ongoing importance of raising awareness, promoting sun-safe behaviors, and advancing early detection and treatment efforts to combat this growing public health challenge [2].

The clinical level SC screening involves in; (i) personal check by an experienced dermatologist, (ii) dermatoscopy supported examination and, (iii) biopsy sample collection from the abnormal skin section and implementing a microscopy examination. During the examination, the dermatologist will examine the suspicious skin section based on the ABCD/ABCDE clinical protocol. During this examination, the chosen rule is used to identify the parameters, like Asymmetry (A), Border-irregularity (B), Color-variation (C), Diameter (D), and Evolution (E) [3].

The traditional procedure involved in SC analysis based on the dermatoscopy is quite time consuming and segregation of collected digital-dermatoscopy (DD) images into benign and malignant needs expertise. Hence, the SC diagnostic centers are equipped with the computer software, which helps in traditional examination of the DDs and its segregation into benign and malignant classes [4]. After the segregation, the patient diagnosed with the malignant SC is then further examined using the biopsy test, a practice to confirm the skin cancer and its severity.

The detection of benign/malignant SC is very essential for the treatment planning and its execution [5]. Recently, computerized tool based SC detection is widely proposed, which helps in classifying the DD images using a machine-learning (ML) and deep-learning (DL) methods. Implementation of the ML for the DD examination needs special attention to select the necessary features, which help to provide better detection accuracy. The availability of the pretrained DL schemes makes it easy to implement for the DD image classification task using SoftMax of a chosen classifier unit. Further, the Dl approaches helps in achieving better detection accuracy during the benign and malignant SC detection process.

The proposed research aims to develop a tool to examine DD images using the pretrained ResNet (RN) architecture. This work considered the RN variants like the traditional RN (50, 101, and 152) and the RNV2 (50, 101, and 152) schemes for examining the chosen DD images. The various phases of the developed tool includes; (i) image collection and resizing, (ii) feature extraction using a chosen RN scheme, (iii) feature reduction using 50% dropout and serial features fusion, and (iv) binary classification using SoftMax and performance validation.

This work considered the DD database found in [6] for the examination and from each class, 1000 images are considered. The merit of this tool is verified using; traditional-features (TF), reduced-features (RF) and fused-features (FF) and for each case, the necessary metrics, like accuracy (AC), precision (PR), sensitivity (SE) and specificity (SP) are computed and based on the overall performance, the merit of the chosen features are confirmed.

The experimental investigation confirms that the FF based classification using the RNV2_50 and RNV2_152 helps in providing a classification accuracy of 98%. The experimental investigation confirms that the proposed scheme helps in providing better detection accuracy on the chosen database.

The contribution of this research involves;

(i) Performance verification of ResNet variants on the chosen DD image database,
(ii) Implementing the binary classification using CF, RF and FF.

This research is organized as follows; Section 2 reviews the literature, Section 3 outlines the methodology, Section 4 presents experimental results, and Section 5 offers conclusions.

II. Literature Review

The SC is one of the harsh cancers and can be completely curable when diagnosed early. Due to its large incidence rate,

ISBN 979-8-3503-5094-4

a considerable number of works are proposed to support the early diagnosis of the SC and its category.

Recently, the ML and DL based examination procedures are developed for early detection of the SC using the DD images. Compared to the ML approaches, the DL techniques helps in providing better detection accuracy and this outcome is essential for treatment planning and implementation tasks.

The earlier DL-techniques discussed in [7-9] confirms that the detection of benign and malignant SC is essential for treating the patient with a clinically approved protocol. Further, the availability of pretrained DL-techniques supported the implementation of the well-known DL procedures to examine the SC with better accuracy.

Most of these works considered the benchmark DD images of International Skin Imaging Collaboration (ISIC) for the examination [10,11]. Recently, the work of Javid et al. (2023) proposed an ensemble features based detection of benign and malignant class DD images and achieved a detection accuracy of >93%. The database used in this research can be accessed from [12].

This research proposes a methodology using the RN/RNV2 variants to examine the DD database and the achieved results are analysed to verify the significance of proposed tool.

III. Methodology

Clinical level disease examination plays a vital role in diagnosis of the SC and its severity. The performance of the tool chiefly depends on the procedures involved from the data collection and the outcome evaluation. In this work, the DD examination procedure is performed using three different feature sets, like CF, RF and FF and the outcome is evaluated to authorize the value of the developed tool.

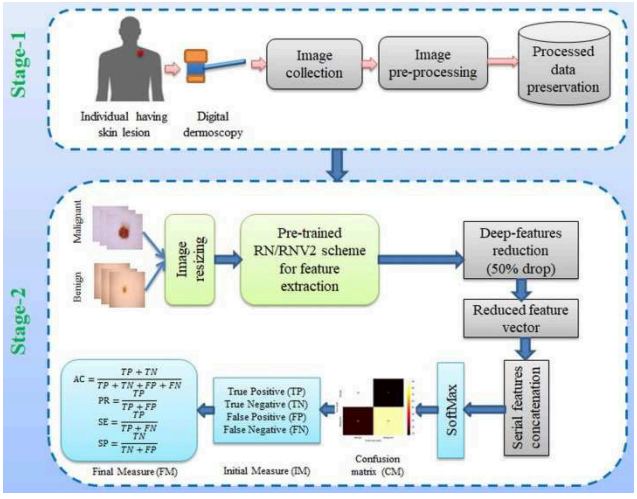


Figure. 1. Implemented DD examination

Figure 1 depicts the architecture of the developed tool using the chosen DL model, which executes a supervised learning. Initially, the necessary images are collected and then resized to $150 \times 150 \times 3$ pixels. The necessary features from these images are extracted using the RN/RNV2 scheme with a chosen variant (50, 101, and 152). Then, a feature reduction is then executed using 50% dropout and serial features integration is implemented to get the FF. This feature vector is then implemented to classify the data using SoftMax and the outcome is verified with the confusion-matrix (CM) as depicted in figure. This tool implements the classification task using the CF, RF and the FF and the ontained results are compared to verify the merit of this tool on the DD examination task.

A. DD dataset

The choice of the clinical-grade images is very crucial for the medical data examination task. Further, the developed tool which works on the chosen clinical grade images can be easily transformed for the real clinical image examination task. In this work, the clinical grade DD images available in [6] is considered for the examination. The initial work on this database can be accessed from [12] and this work also implements the DL technique with ensemble features for implementing a binary classification.

In this work, 2000 images (1000 benign and 1000 malignant) are utilized and the data-split during the RN based classification is achieved as in Table 2. The Test results are then considered to verify the merit of this scheme. The example images of chosen database are represented in Figure 2.

TABLE I. Data-split considered in the proposed study

Class	Total DD images	Training	Validation	Testing
Benign	1000	800	100	100
Malignant	1000	800	100	100

Figure. 2. Sample DD images collected from the chosen database []

B. ResNet scheme

The DL assisted image examination is one of the widely adopted practices to examine a variety of the images and the final outcome is achieved using a binary- or multi-class classifier. The binary-classification is one of the common practices in segregating the chosen image database into normal and abnormal class using a chosen classifier. The SoftMax is the inbuilt classifier in most of the pretrained DL models and hence, this work also implemented the SoftMax during the examination of the DD images.

The choice of a DL depends on its performance and implementation complexity. Compared to other models, the VGG-variants and the RN-variants are quite commonly considered methods to examine a class of medical images and this research considered the RN and RNV2 variants to solve the chosen image examination task. The RN and RNV2 schemes are effective in providing a better top-1 and top-5 detection accuracy, when it is executed using the ImageNet

ISBN 979-8-3503-5094-4

database and the necessary information about the achieved accuracy can be accessed from [13]. Further, the total parameters in the RN-variants are quite less compared to the VGG-variants and the top-5 accuracy also better in RN and RNV2. Hence, the proposed study considered the RN-variants like, 50, 101, and 152 for the evaluation and achieved a better detection AC when SoftMax classifier is considered. The necessary information regarding RN can be accessed from [14,15].

C. *Implementation*

The DD examination procedure is implemented using a range of procedures, like data collection, resizing, DL scheme selection, feature mining and reduction, features-fusion and classification. In this work, the following parameters for the RN/RNV2 are assigned; epochs= 60, monitoring measures= accuracy and categorical cross entropy, optimizer= Adam, intermediate layer activation= ReLU, dense layer activation= Sigmoid, and batch size= 32.

The complete information of RN/RNV2 can be accessed from literature and this scheme helps in extracting the necessary deep-features from the chosen DD image. The total deep-features $(1 \times 1 \times 1000)$ and other features, like RF $(1 \times 1 \times 500)$ and FF $(1 \times 1 \times 1000)$ are considered for the detection task. The proposed tool is separately tested using CF, RF and FF and the achieved results are presented and discussed.

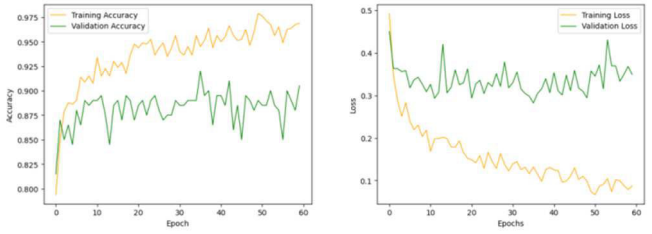
D. *Validation*

The tool's effectiveness with RN and RNV2 is confirmed by calculating initial measures as well as final measures including AC, PR, SE, SP, and FS using CM obtained from classification. The Receiver Operating Characteristic (ROC) curve, which illustrates the trade-off between TP- and TN-rates, is also used to evaluate the complete performance of the tool. The mathematical terms for these values can be found in [8,9,12].

IV. Result and Discussion

This section presents the investigation outcomes obtained using a workstation equipped with an Inteli5, 16GB of RAM, and 4GB of VRAM. The experiments were conducted using Python software.

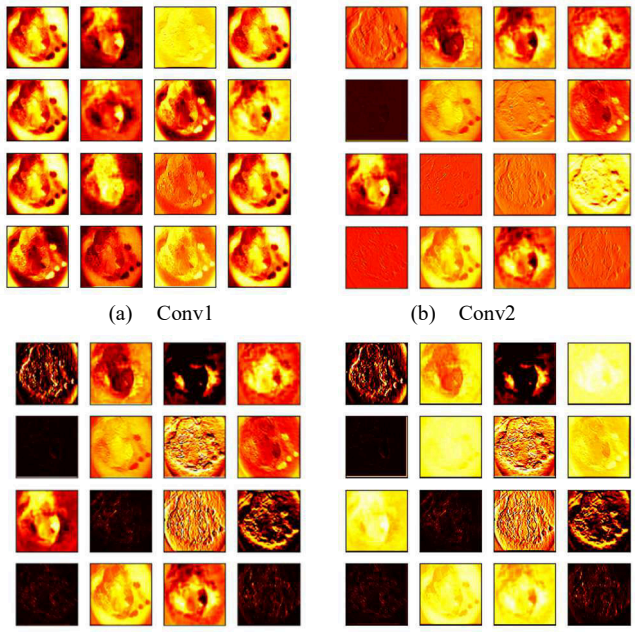
Initially, the performance of the RN (50) was evaluated using labeled DD images, and the results were recorded. Subsequently, the experiment was repeated, replacing the RN (50) with other RN-variants chosen in this study. Figure 3 displays the experimental results for RN (50), where Fig 3(a) and (b) depicts the accuracy and loss for an epoch of 60. Similar experiments were conducted for RN-variants, yielding comparable outcomes.



(a) Accuracy (b) Loss

Figure 3. Sample DD images collected from the chosen database

Figure 4 presents the various convolutional-layer (Conv) outcomes. Fig 4(a) to (d) presents the results of Conv1 to Conv4 and this confirms that the image features are effectively detected with the RN scheme. Table II presents the initial experimental result and this confirms the performance of the proposed scheme is better when RN(50 is implemented. Similar results are obtained with other models of this study.



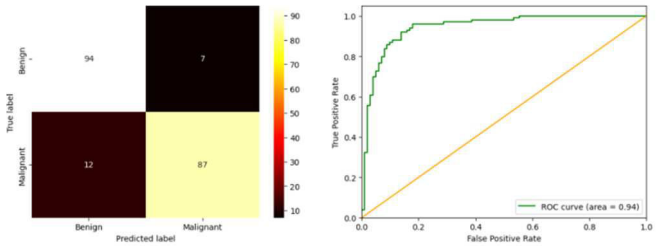
(a) Conv1 (b) Conv2

(c) Conv3 (d) Conv4

Figure 4. Various convolutional-layer (Conv) outcomes of RN(50)

TABLE II. Experimental outcome attained with RN(50)

	Precision	Recall	F1-score	Support
Benign	0.89	0.93	0.91	101
Malignant	0.93	0.88	0.90	99
accuracy	-	-	0.91	200
macro avg	0.91	0.90	0.90	200
weighted avg	0.91	0.91	0.90	200



(a) CM (b) ROC

Figure 5. Achieved CM and the ROC with RN(50)

Figure 5 shows the final result achieved with RN(5) in which Fig 5(a) presents the CM and Fig 5(b) shows the ROC with a final value of 0.94; both confirms the merit of the RN(50 on the chosen database.

Similar procedure is repeated with other schemes using CF, RF and FF and the obtained results are shown in Table III and IV, respectively. Experimental outcome obtained with CF with SoftMax classifier

ISBN 979-8-3503-5094-4

TABLE III. INITIAL EXPERIMENTAL RESULTS

Scheme	TP	FN	TN	FP	AC	PR	SE	SP	FS
RN(50)	87	12	94	7	90.50	92.55	87.88	93.07	90.15
RN(101)	86	14	91	9	88.50	90.53	86.00	91.00	88.20
RN(152)	88	12	94	6	91.00	93.62	88.00	94.00	90.72
RNV2(50)	87	13	92	8	89.50	91.58	87.00	92.00	89.23
RNV2(101)	87	12	93	8	90.00	91.58	87.88	92.08	89.69
RNV2(152)	88	12	95	5	91.50	94.62	88.00	95.00	91.19

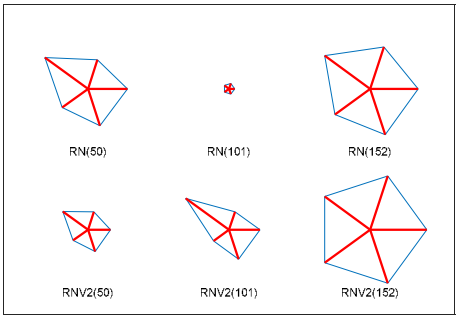


Figure 6. Glyph-plot obtained with the experimental outcome for the CF

TABLE IV. EXPERIMENTAL OUTCOME OBTAINED WITH RF AND FF WITH SOFTMAX CLASSIFIER

Scheme	TP	FN	TN	FP	AC	PR	SE	SP	FS
RN(50)	85	15	91	9	88.00	90.42	85.00	91.00	87.63
RN(101)	87	13	92	8	89.50	91.58	87.00	92.00	89.23
RN(152)	89	11	91	9	90.00	90.81	89.00	91.00	89.89
RNV2(50)	89	12	92	7	90.50	92.71	88.12	92.92	90.35
RNV2(101)	89	10	91	10	90.00	89.89	89.89	90.09	89.89
RNV2(152)	90	9	93	8	91.50	91.84	90.91	92.08	91.37
FF	97	3	99	1	98.00	98.98	97.00	99.00	97.98

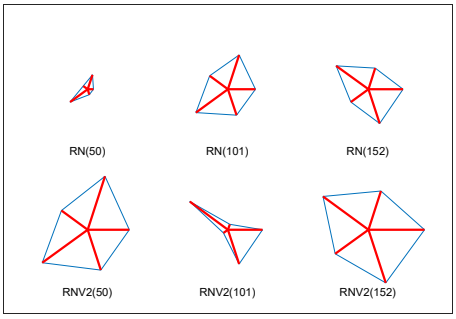


Figure 7. Glyph-plot obtained with the experimental outcome for the RF

Table III presents the outcome for the CF and its overall performance is graphically presented as in Figure 6. This confirms that the RNV2(152) provides a better outcome. Table IV shows that the RF based evaluation provides superior result for RNV2(50) and RNV2(152) and these features are serially integrated to get the FF. Figure 7 also confirms that the overall merit of these two models are better compared to other considered models. From this table, it is evident that, the FF based detection presents a better accuracy (98%) on the chosen database.

The outcome of this study verifies that the implemented technique provides a better SC detection accuracy on the chosen database. In future, other DL models can be utilized to detect the SC and its performance is to be confirmed against the RN/RNV2 variants.

V. CONCLUSION

Examination of the SC from DD is one of the common practices and hence various ML and DL methods are developed and implemented. This research aims to develop a binary-classification tool using the RN/RNV2 variants and verifying its performance using SoftMax with various features, like CF, RF, and FF. The outcome of this study confirm that, the CF based detection provides an accuracy of >88%. Further, the RF and FF based methods achieved >88% and 98% accuracies, respectively. This confirms the merit of the proposed technique with the FF. In the future, similar DD database can be tested using other existing DL models and its performance can be verified against the proposed RNV2 based tool.

REFERENCES

[1] Ferlay, J., Colombet, M., Soerjomataram, I., Parkin, D. M., Piñeros, M., Znaor, A., & Bray, F. (2021). Cancer statistics for the year 2020: An overview. *International journal of cancer*, *149*(4), 778-789.

[2] Siegel, R. L., Miller, K. D., Wagle, N. S., & Jemal, A. (2023). Cancer statistics, 2023. *Ca Cancer J Clin*, *73*(1), 17-48.

[3] Nachbar et al., "The ABCD rule of dermatoscopy: High prospective value in the diagnosis of doubtful melanocytic skin lesions", *Journal of the American Academy of Dermatology*, vol. 30, pp. 551-559, 1994.

[4] Rajpara, S. M., Botello, A. P., Townend, J., & Ormerod, A. D. (2009). Systematic review of dermoscopy and digital dermoscopy/artificial intelligence for the diagnosis of melanoma. *British Journal of Dermatology*, *161*(3), 591-604.

[5] Heibel, H. D., Hooey, L., & Cockerell, C. J. (2020). A review of noninvasive techniques for skin cancer detection in dermatology. *American journal of clinical dermatology*, *21*, 513-524.

[6] https://www.kaggle.com/datasets/hasnainjaved/melanoma-skin-cancer-dataset-of-10000-images

[7] Daghrir, J., Tlig, L., Bouchouicha, M., & Sayadi, M. (2020, September). Melanoma skin cancer detection using deep learning and classical machine learning techniques: A hybrid approach. In *2020 5th international conference on advanced technologies for signal and image processing (ATSIP)* (pp. 1-5). IEEE.

[8] Dildar, M., Akram, S., Irfan, M., Khan, H. U., Ramzan, M., Mahmood, A. R., ... & Mahnashi, M. H. (2021). Skin cancer detection: a review using deep learning techniques. *International journal of environmental research and public health*, *18*(10), 5479.

[9] Tan, T. Y., Zhang, L., & Lim, C. P. (2019). Intelligent skin cancer diagnosis using improved particle swarm optimization and deep learning models. *Applied Soft Computing*, *84*, 105725.

[10] Codella, N., Rotemberg, V., Tschandl, P., Celebi, M. E., Dusza, S., Gutman, D., ... & Halpern, A. (2019). Skin lesion analysis toward melanoma detection 2018: A challenge hosted by the international skin imaging collaboration (isic). *arXiv preprint arXiv:1902.03368*.

[11] Codella, N. C., Gutman, D., Celebi, M. E., Helba, B., Marchetti, M. A., Dusza, S. W., ... & Halpern, A. (2018, April). Skin lesion analysis toward melanoma detection: A challenge at the 2017 international symposium on biomedical imaging (isbi), hosted by the international skin imaging collaboration (isic). In *2018 IEEE 15th international symposium on biomedical imaging (ISBI 2018)* (pp. 168-172). IEEE.

[12] Javid, M. H., Jadoon, W., Ali, H., & Ali, M. D. (2023, February). Design and Analysis of an Improved Deep Ensemble Learning Model for Melanoma Skin Cancer Classification. In *2023 4th International Conference on Advancements in Computational Sciences (ICACS)* (pp. 1-6). IEEE.

[13] https://keras.io/api/applicatio

ISBN 979-8-3503-5094-4

Manifold Cervical Abnormality Categorization using Deep Learned Features and Support Vector Machine

Ramesh Munirathinam
Department of Biomedical Engineering
Karpagam Academy of Higher Education
Coimbatore, India
rameshmunirathinam@gmail.com

M Muthulakshmi
Department of ECE
Amrita School of Engineering
Amrita Vishwa Vidyapeetham
Chennai, India
m_muthulakshmi@ch.amrita.edu
*corresponding author

M Latha
Department of ECE
Rajalakshmi Engineering College
Chennai, India
latha.m@rajalakshmi.edu.in

Kotagiri Uttejitha
Department of ECE
Amrita School of Engineering
Amrita Vishwa Vidyapeetham
Chennai, India

Naveen A
Department of Biomedical Engineering
Karpagam Academy of Higher Education
Coimbatore, India
20bebme029@kahedu.edu.in

M Tamilnidhi
Department of ECE
Karpagam College of Engineering
Coimbatore, India
tamilnidhi.ece@gmail.com

***Abstract*—Cervical cancer stands as a prominent contributor to female cancer mortality on a global scale, ranking high among the leading causes of death from cancer in women. Persistent HPV infection is the cause of cervical cancer, which continues to be a global health concern. The purpose of this research was to identify and comprehend different forms of cervical abnormalities and their consequences to develop the CAD diagnosis sytem. We used deep learning models (Squeezenet) combined with the Orange Data Mining Tool to extract features at the image level from 966 cluster cell images of Pap smear slides from the SIPaKMeD database. We employed SVM with different kernel functions and k-NN with different parameters for cervical cancers classification. Resluths shows that tha SVM with RBF kernel function achieved highest classification accuracy of 87% in identifying the cervical abnormalities. High value of 96% for MCC clearly indicates that the features extracted using squeezenet features has the ability to categorize abnormality. The k-NN with different parameters shows the better performance in identifying the anomalies with an accuracy of 82.7%. The SVM with RBF kernel is found to show comparatively high peformance in detecting the abnormalities in the images with recall and AUC values of 87% and 97.8%.**

***Index Terms*—Deep Learning, Cervical , Squeezenet,, SVM, Orange Data Mining Tool, Metrics, Kernels.**

I. Introduction

Cancer comprises a range of illnesses affecting complex multicellular organisms, marked by changes in the activity of numerous genes. These alterations result in the disruption of the usual cellular processes governing both cell division and differentiation [1]. Cervical cancer (CC) is a disorder in which abnormal cervix cells begin to develop uncontrollably ,CC stands as a prominent contributor to cancer-related fatalities in women [2]. Middle- aged women continue to face a significant public health risk from cervical cancer, especially in nations with less developed infrastructure [3]. Globally, 342 000 women lost their lives to cervical cancer in 2020, accounting for more than 600 000 new cases of the disease [4]. Frequently encountered signs include a substantial and unpleasant odor emitting from vaginal discharge, irregular bleeding or bleeding between menstrual periods, bleeding after sexual intercourse, bleeding after menopause, or experiencing back pain [5]. Infection with high-risk or oncogenic types of human papillomavirus (HPV) is the leading factor behind the development of pre-cancerous and cancerous lesions in the cervix [6]. In a cohort of 114 women diagnosed with cervical cancer and aged 35 or younger, 15% belonged to the age group less than 25, while the remaining 85% fell within the 26 to 35 age range [7]. In the realm of cervical cancer screening, there are two primary diagnostic methods: the Papanikolaou test, aimed at early detection of precancerous and cancerous cell abnormalities for effective intervention, and the HPV test, which identifies infections linked to HPV types known to pose a cancer risk [8]. Additionally, other screening methods, such as MRI, CT, PET, cystoscopy, and proctoscopy, may be employed for comprehensive evaluation and monitoring [9]. Therapeutic approaches for cervical cancer encompass surgical techniques such as simple hysterectomy and trachelectomy, along with radiotherapies like Concurrent Chemoradiation Therapy (CRT) and Palliative Radiotherapy. Chemotherapies are also employed as part of the multifaceted treatment strategies [10] .

II. Literature Review

IN 2020, Chen Ying et al. conducted a study utilizing 4993 cervical histopathological images, comprising 2503 benign and 2409 malignant images. Employing deep transfer learning with CNN, the study yielded results which includes a sensitivity of 95.88%, an area under the curve of 99.71%,

ISBN 979-8-3503-5094-4

Youden's index at 94.81%, a specificity of 98.93% and an overall accuracy of 97.42% in classifying cervical images [11]

Ankur Manna et al. 2021, developed an classification based on ensemble using Densenet-169, Inception V3 and Xception. The study utilized datasets from SIPaKMeD and Mendeley Liquid-based Cytology. In the Mendeley dataset, they achieved an accuracy and sensitivity of 99.23%. For the SIPaKMeD dataset, in a two-class setting, they obtained a sensitivity of 98.52% and an accuracy of 98.55%. Furthermore, in a five-class setting within the SIPaKMeD dataset, a sensitivity of 98.52% and an accuracy of 95.43% is achieved [12]

Hua Chen et al. 2021, utilized compact VGG model to analyse both SIPaKMeD and Herlev datasets. The authors reported an accuracy of 97.80% for SIPaKMeD datasets and 94.81% for Herlev dataset [13] In 2021, Anurag Tripathi et al. employed a deep learning classification model, called ResNet-152, on the SIPaKMeD dataset, achieving an accuracy of 94.89% in the classification of cervical cell images [14]

Toto Haryanto et al. 2020, utilized CNN algorithm using Alexnet, modified VGG 19 and Lenet model and achieved an accuracy of 87.32% in alexnet on the datasets of SIPaKMeD [15] In 2019, Rohan Gorantla et al. utilized the Intel and Mobile ODT datasets, achieving results with a kappa score of 0.951 and an accuracy of 96.77%. The authors used the CervixNet methodology and proposed a novel algorithm named Hierarchical Convolutional of Experts algorithm [16]

In 2020, Fatma B A et al. employed a set of algorithms, including Naive Bayes, K-Nearest Neighbors, Support Vector Machine, Random Forest, Multilayer Perceptron, and Logistic Regression. These algorithms were applied to the Herlev dataset, consisting of 917 images, and yielded accuracy values ranging between 83% and 92% [17]

In 2022, Wen Chen et al. applied the ShuffleNetV2, GhostNet, and DenseNet-121 architectures to the SIPaKMed dataset. ShuffleNetV2 exhibited performance with a specificity of 99.08%, precision of 96.03%, recall of 96.23%, and an accuracy of 96.79% when compared to DenseNet-121 with a accuracy of 96.79%. GhostNet attaining a specificity of 99.09%, accuracy of 96.39%, precision of 96.42%, and recall of 96.39%, showcasing its effectiveness relative to DenseNet-121 in the classification task [18]

Ming Fang et al. 2022 conducted a study utilizing both SIPaKMed and Herlev datasets, achieving superior results on SIPaKMed. Employing the DeepCell convolutional neural network, attained an accuracy of 95.628%, an F-score of 95.63%, recall of 95.647%, and precision of 95.685% [19] In 2021, Senturk and Suleyman employed a variety of deep learning models, including SqueezeNet, VGG19, AlexNet, ResNet-50, and Inception V3, on the SIPaKMeD dataset. Among these models, SqueezeNet yielded the highest accuracy, reaching a maximum of 96.90% [20]

In 2023, Ritesh Maurya et al. employed the SIPaKMeD dataset and implemented both Long Short-Term Memory networks and Vision Transfer networks. Their study resulted in an accuracy of 95.80% for the LSTM network and an even higher accuracy of 97.65% for the Vision Transfer network [21] In 2022, Meghana Karrai et al. conducted a study utilizing SIPaKMed, Herlev, and CRIC datasets. Employing transfer learning with the ResNet-50 architecture, they achieved remarkable results on the SIPaKMeD dataset, including a specificity of 99.25% and an accuracy of 99.12% [22]

In 2021, Hritam Basak et al. employed the Mendeley, Herlev, and SIPaKMed datasets, applying the Grey Wolf Optimizer to a Convolutional Neural Network (CNN). Remarkably, the authors achieved high classification accuracies of 99.47%, 98.32%, and 97.87% for the respective datasets [23]

III. Materials and Methods

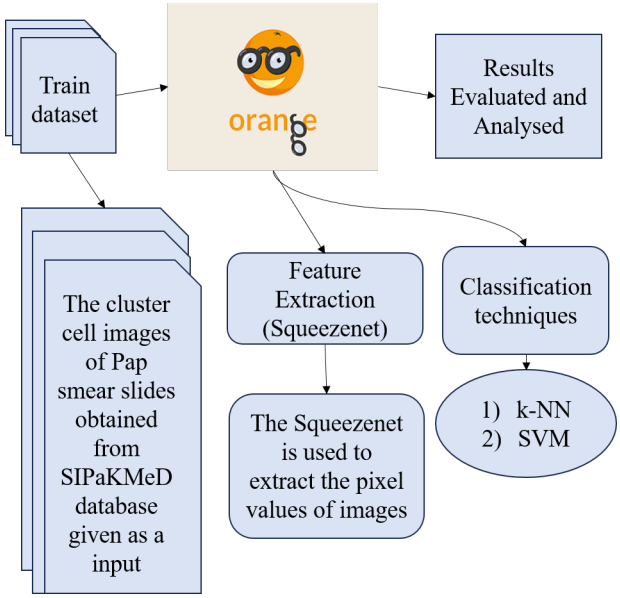


Fig. 1. The proposed workflow

A. Dataset

This study considers 966 cluster cell images of Pap smear slides from the SIPaKMeD database. The cell images are divided into five categories containing Dyskeratotic (DK) (233 images) , Koilocytotic (KC) (238 images), Metaplastic (MP) (271 images), Parabasal (PB) (108 images) and Superficial - intermediate (SI) (126 images).

B. Feature Extraction and Classification techniques

The process of converting raw data into numerical features is referred to as feature extraction. The histopathological

ISBN 979-8-3503-5094-4

images are subjected to feature extraction phase using Squeezenet. The feature length of Squeezenet is 1000 . These feature values are fed to SVM and k-NN for identifying the cancers. The efficiency of SVM is evaluated by changing the kernel functions and the efficiency of k-NN is evaluated by changing the distance metrics.

1) Squeezenet: Squeezenet comprises fifteen layers, with five unique layers:

- Two convolutional layers
- Eightfire layers
- Three maxpool layers
- one global average pool layers
- one output layer(softmax) [24]

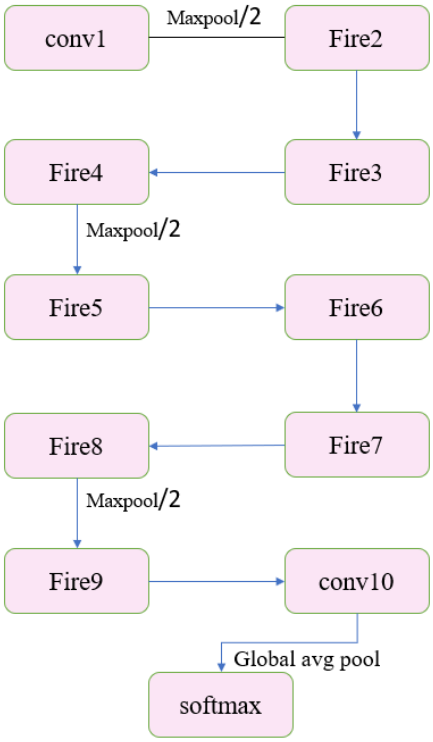


Fig. 2. Squeezenet model

2) Support Vector Machine(SVM): Support Vector Machine, one of the classification methods, seeks to determine the optimal point dividing the two classes. SVM, which was introduced by Vapnik, one of the most prominent machine learning approaches over time. Its primary objective is to identify the optimal line that effectively separates two classes within a dataset. SVM is an extremely strong approach for predicting values that aren't in the training set. It may be used successfully in difficult categorization applications because of this property [25]. SVM is one of the kernel-based algorithms used for pattern recognition and regression. In the context of classification, Support Vector Machine offers a variety of kernel functions, such as sigmoid, linear, polynomial, and radial basis function. We employed sigmoid, linear, radial basis function and polynomial kernels in cervical cancer classifcation

Linear(L) kernel function

$$k = x.y \tag{1}$$

Polynomial(P) kernel function

$$k = (gx.y + c)^d \tag{2}$$

Radial Basis function(RBF) kernel

$$k = exp(-g|x - y|^2) \tag{3}$$

Sigmoid(S) kernel function

$$k = tanh(gx.y + c) \tag{4}$$

3) k-nearest neighbor(k-NN): Based on its simplicity and efficiency, kNN classification has been ranked as one of the top ten data-mining methods. Thus, the kNN approach has been effectively developed in data-mining applications that include regression, missing value imputation, and classification [26]. In this section, we employed distance metrics for classifiaction which includes Manhattan distance metric, Euclidean distance metric and Chebyshev distance metric. These are the most frequently used distance metrics. We intend to evaluate the performance of several distance measures, in order to select the most optimal metric for classification performance.

$$d(x, y) = \left[\sum_{j=1}^{d} |x_j - y_j|^p\right]^{\frac{1}{p}} \tag{1}$$

In Eq.(5) the x and y are two different sample points and d represent the dimension, if p = 2, d(x,y) represent the Euclidean distance metric, if p = 1, d(x,y) represent the Chebyshev distance metric, if p = infinity, d(x,y) represent the Manhattan distance metric [27].

IV. Results and Discussions

Fig. 3. Representative cluster cell images of cervical cancers considered for the analysis

The representative cluster cell images are shown in Fig. 3. It can be observed that the staining pattern is varied between the abnormalities and also there exists a significant difference in number of cells in the images. Squeezenet model extracts 1000 feature values from the cluster cell images of cervical abnormalities and are subjected to classification using SVM by varying kernel functions and k-NN by varying distance and weight metrics. Table.I presents the performance metrics of the machine learning model for distinguishing abnormalities. As can be seen, the SVM with the RBF kernel function achieved 87% higher classification accuracy and recall, which

ISBN 979-8-3503-5094-4

TABLE I
PERFORMANCE OF SVM AND KNN- IN DIFFERENTIATING THE CERVICAL CANCER USING FEATURES OF SQUEEZET MODEL

	CA	F1	Prec	Rec	MCC	Spec	AUC
SVM-L	0.815	0.815	0.817	0.815	0.762	0.942	0.954
SVM-P	0.867	0.867	0.868	0.867	0.830	0.960	0.975
SVM-R	0.870	0.870	0.873	0.870	0.833	0.960	0.978
SVM-S	0.433	0.430	0.432	0.433	0.268	0.828	0.761
kNN-EU	0.816	0.816	0.825	0.816	0.766	0.946	0.948
kNN-EBD	0.822	0.822	0.825	0.822	0.773	0.948	0.951
kNN-M-U	0.818	0.818	0.825	0.818	0.768	0.947	0.946
kNN-MBD	0.827	0.827	0.829	0.827	0.779	0.950	0.949
kNN-CU	0.807	0.808	0.816	0.807	0.756	0.944	0.941
kNN-CBD	0.821	0.821	0.824	0.821	0.772	0.948	0.945

is comparatively higher than the performance of the SVM with linear, polynomial, and sigmoid kernel functions, which obtained an accuracy and recall value of 81.5%, 86.7%, and 43.3%, respectively. The SVM with polynomial kernel function shows similar performance to the SVM with RBF kernel in classifying the abnormalities, with an MCC value of 83%, which corresponds to the MCC (83.3%) value of the SVM with RBF kernel function. The SVM with sigomid kernel exhibits the lowest performance with an accuracy and recall value of 43.3%. The SVM with RBF and polynomial kernel function was the most specificity with 96%. The performance of k-NN model evaluated and analysed by varying distance metrics and weight functions. The k-NN distance metric displays the better performance than the k-NN with uniform function. The k-NN with distance metrics obtained 82.2%, 82.7% and 82.1% classification accuracy and recall values in euclidean, manhattan and chebyshev metrics respectively and k-NN with uniform function obtained 81.6%, 81.8% and 80.7% classification accuracy and recall values in euclidean, manhattan and chebyshev metric respectively. The k-NN with manhattan distance metric shows the better performance than other k-NN used in the study and and its achieved 77.9% in MCC and 95% in specificity for classifying the abnormalities. According to the analysis, the SVM with RBF kernel outperformed the other models used in the study with the highest accuracy, recall, precision, MCC, AUC and specificity of 87%, 87%, 87.3%, 83.3%, 97.8% and 96% in classifying the cervical cancers respectively and SVM with sigmoid kernel function shows the lowest performance in classifying the cervical cancers.

From Fig.4-7 the confusion matrix with RBF kernel function displays the highest classification rate with a percentage of 87% comparatively other confusion matrix in SVM model. The sigmoid kernel displays the lowest performance in identifying the abnormalities with a misclassification rate of 76.7% in overall abnormalities. Also, the SVM with linear and polynomial kernels shows the better performance in identifying the abnormalities with 81.5% and 86.7% classification accuracy respectively.

It is observed from Fig.8-13 the confusion matrix with manhattan distance metric displays the highest classification rate with 82.7% among alll the k-NN models used in the study. The

Confusion matrix for SVM - Linear (showing number of instances)

Actual \ Predicted	Dyskeratotic	Koliocytotic	Metaplastic	parabasal	superficial-intermediate	∑
Dyskeratotic	186	26	7	4	0	223
Koliocytotic	40	171	19	2	6	238
Metaplastic	18	26	220	1	6	271
parabasal	3	0	2	103	0	108
superficial-intermediate	2	9	8	0	107	126
∑	249	232	256	110	119	966

Fig. 4. The confusion matrix for the SVM linear kernel, used with squeezenet features, shows the following classification accuracies using 10 fold cross validation: Parabasal (PB) and Koilocytotic (KC) were correctly classified with 95.3% and 71.8% respectively. DK, KC and MP were correctly classified with over 80%. This model had a classification rate of 81.5% and misclassification rate of 18.5% among all the classes

Confusion matrix for SVM - Polynomial (showing number of instances)

Actual \ Predicted	Dyskeratotic	Koliocytotic	Metaplastic	parabasal	superficial-intermediate	∑
Dyskeratotic	203	10	7	2	1	223
Koliocytotic	20	189	19	2	8	238
Metaplastic	14	15	235	3	4	271
parabasal	7	1	2	98	0	108
superficial-intermediate	2	8	3	0	113	126
∑	246	223	266	105	126	966

Fig. 5. Confusion matrix of SVM with polynomial kernel function achieves a 86.7% of classification rate. It correctly classified 91.03% instances in DK, 79.4% instances in KC, 86.7% in MP, 90.7% in PB and 89.6 In SI instances. The instances displays the presence of 13.3% incorrect predictions out of 988 instances

Confusion matrix for SVM - RBF (showing number of instances)

Actual \ Predicted	Dyskeratotic	Koliocytotic	Metaplastic	parabasal	superficial-intermediate	∑
Dyskeratotic	203	12	8	0	0	223
Koliocytotic	16	192	19	1	10	238
Metaplastic	14	8	243	0	6	271
parabasal	15	2	3	88	0	108
superficial-intermediate	1	9	2	0	114	126
∑	249	223	275	89	130	966

Fig. 6. Confusion matrix of the RBF kernel correctly predicted 91.03%, 79.4%, 86.7%, 90.7% and 89.6% in DK, KC, MP, PB and SI instances respectively. The classifier were misclassified 46 instances as a dyskeratotic (DK), 31 instances as a koilocytotic (KC), 32 instances as a metaplastic (MP), 1 instances as a parabasal (PB) and 16 instances as a superficial-intermediate (SI). A total of 46 instances (DK) were misclassified by the classifier, which is highest misclassification rate among all the classes. This confusion matrix shows 87% classification rate and 13% misclassfication rate out of 966 images

Confusion matrix for SVM - Sigmoid (showing number of instances)

Actual \ Predicted	Dyskeratotic	Koliocytotic	Metaplastic	parabasal	superficial-intermediate	∑
Dyskeratotic	80	67	60	4	12	223
Koliocytotic	49	95	65	4	25	238
Metaplastic	16	40	138	49	28	271
parabasal	13	10	51	34	0	108
superficial-intermediate	24	12	19	0	71	126
∑	182	224	333	91	136	966

Fig. 7. The confusion matrix for the SVM with sigmoid kernel, used with squeezenet features, shows the following classification accuracies: MP and SI were correctly classified with 50.92% and 56.34% respectively. DK, KC and PB were classified with over 30%. This model displys the lowest classification rate which is 43.3% and highest misclassification rate with 76.7% out of 966 images that showing great difficulties in its identification.

ISBN 979-8-3503-5094-4

Confusion matrix for KNN Euclidean - Uniform (showing number of instances)

		Predicted					
		Dyskeratotic	Koliocytotic	Metaplastic	parabasal	superficial-intermediate	∑
Actual	Dyskeratotic	197	16	7	2	1	223
	Koliocytotic	36	180	14	2	6	238
	Metaplastic	26	17	209	8	11	271
	parabasal	14	1	2	91	0	108
	superficial-intermediate	5	6	4	0	111	126
	∑	278	220	236	103	129	966

Fig. 8. Confusion matrix of k-NN with Euclidean metric and uniform weight shows the classification rate of 81.6% in overall abnormalities. This model has misclassified 18.4% images out of 966 images

Confusion matrix for KNN Euclidean - By distances (showing number of instances)

		Predicted					
		Dyskeratotic	Koliocytotic	Metaplastic	parabasal	superficial-intermediate	∑
Actual	Dyskeratotic	194	15	10	3	1	223
	Koliocytotic	29	180	18	2	9	238
	Metaplastic	19	17	215	8	12	271
	parabasal	12	2	2	92	0	108
	superficial-intermediate	2	6	5	0	113	126
	∑	256	220	250	105	135	966

Fig. 9. Confusion matrix of k-NN with Euclidean metric and By distances weight gives a 82.2% accuracy. The plotted confusion matrix displays the presence of 17.8% incorrect predictions out of 966 images

Confusion matrix for KNN Manhattan - Uniform (showing number of instances)

		Predicted					
		Dyskeratotic	Koliocytotic	Metaplastic	parabasal	superficial-intermediate	∑
Actual	Dyskeratotic	198	15	7	2	1	223
	Koliocytotic	32	182	15	4	5	238
	Metaplastic	25	16	211	7	12	271
	parabasal	14	1	2	91	0	108
	superficial-intermediate	5	7	6	0	108	126
	∑	274	221	241	104	126	966

Fig. 10. The confusion matrix for the k-NN with uniform metric, used with squeezenet features, shows the following classification accuracies: DK, PB and SI were correctly identified with 88.78%, 84.2% and 85.72% respectively. The KC and MP were correctly classified with exceeding 70%. This model had a misclassification rate of 18.2% among all the abnormalities

Confusion matrix for KNN Manhattan - By disatances (showing number of instances)

		Predicted					
		Dyskeratotic	Koliocytotic	Metaplastic	parabasal	superficial-intermediate	∑
Actual	Dyskeratotic	194	16	9	3	1	223
	Koliocytotic	25	182	17	4	10	238
	Metaplastic	19	14	218	8	12	271
	parabasal	11	2	2	93	0	108
	superficial-intermediate	2	6	6	0	112	126
	∑	251	220	252	108	135	966

Fig. 11. Confusion matrix of k-NN with Manhattan metric and By distances weight gives 82.7% accuracy using stratified 10 fold cross validation The main diagonal (194, 182, 218, 93, 112) gives the correct predictions. The instances shows the presence of 167 wrong predictions out of 966 instances

Confusion matrix for KNN chebyshev - uniform (showing number of instances)

		Predicted					
		Dyskeratotic	Koliocytotic	Metaplastic	parabasal	superficial-intermediate	∑
Actual	Dyskeratotic	200	11	8	1	3	223
	Koliocytotic	38	178	13	4	5	238
	Metaplastic	23	23	205	6	14	271
	parabasal	15	2	1	89	1	108
	superficial-intermediate	2	10	5	1	108	126
	∑	278	224	232	101	131	966

Fig. 12. Confusion matrix of k-NN with Chebyshev metric and uniform weight achieves a 80.7% of classification rate. It correctly predicted 89.6%, 73.9%, 75.6%, 82.4% and 85.7% in dyskeratotic (DK), koilocytotic (KC), metaplastic (MP), parabasal (PB) and superficial intermediate (SI) respectively. The classifier incorrectly predicted 28.05% as a DK images, 20.53 as a KC, 27% as a MP images, 11.8% as a PB images and 17.55% as a SI images. The percentage of 27% (MP) were misclassified by the classifier, which is the highest misclassification among all the abnormalities

Confusion matrix for KNN chebyshev - By distances (showing number of instances)

		Predicted					
		Dyskeratotic	Koliocytotic	Metaplastic	parabasal	superficial-intermediate	∑
Actual	Dyskeratotic	197	14	7	2	3	223
	Koliocytotic	29	181	18	3	7	238
	Metaplastic	20	18	212	7	14	271
	parabasal	10	1	2	94	1	108
	superficial-intermediate	0	11	5	1	109	126
	∑	256	225	244	107	134	966

Fig. 13. Confusion matrix of k-NN with Chebyshev metric and By distances weight achieves 82.1% accuracy using stratified 10 fold cross validation. It correctly classified 197, 181, 212, 94, and 109 dyskeratotic, koilocytotic, metaplastic, parabasal and and superficial-intermediate images respectively. Also, it displays the presence of 173 incorrect predictions out of 966 instances

chebyshev with uniform metric shows the lowest performance in identifying the abnormalities and it's misclassification rate of 19.3% out of 966 images. The proposed methodology is found to be in par with the cervical cancer detection using textural features [28], [29].

V. Conclusion

The experimental results show that the cluster cell images of Pap smear slides show significant difference between the cervical abnormalities. SVM model with different kernel function outperforms the K-NN with different parameters in differentiating the abnormalities using cluster cell images of Pap smear slide with 87% of classification accuracy and recall rate. High values of AUC (97.8%) and MCC (96%) indicates that the model is highly stable and reliable for identifying the cervical abnormalities in medical image classification. The findings of this study indicate that the employed deep learning models, in combination with the Orange Data Mining Tool, can effectively extract features from cluster cell images and classify accuracy and applicability of these methods in clinical settings. By leveraging advanced technologies, we can improve the diagnosis, treatment, and management of patients leading to better health outcomes and improved quality of life.

References

[1] R. W. Ruddon, Cancer biology. Oxford University Press, 2007.

[2] D. S. Latha, P. V. Lakshmi, and S. Fathima, 'Staging prediction in cervical cancer patients–a machine learning approach', International Journal of Innovative Research and Practices, vol. 2, no. 2, pp. 14–23, 2014.

[3] M. Arbyn et al., 'Estimates of incidence and mortality of cervical cancer in 2018: a worldwide analysis', The Lancet Global Health, vol. 8, no. 2, pp. e191–e203, 2020.

[4] S. Wilailak, M. Kengsakul, and S. Kehoe, 'Worldwide initiatives to eliminate cervical cancer', International Journal of Gynecology & Obstetrics, vol. 155, pp. 102–106, 2021.

[5] G. A. Mishra, S. A. Pimple, and S. S. Shastri, 'An overview of prevention and early detection of cervical cancers', Indian Journal of Medical and Paediatric Oncology, vol. 32, no. 03, pp. 125–132, 2011.

[6] S. Zhang, H. Xu, L. Zhang, and Y. Qiao, 'Cervical cancer: Epidemiology, risk factors and screening', Chinese Journal of Cancer Research, vol. 32, no. 6, p. 720, 2020.

[7] E. Pelkofski, J. Stine, N. A. Wages, P. A. Gehrig, K. H. Kim, and L. A. Cantrell, 'Cervical cancer in women aged 35 years and younger', Clinical therapeutics, vol. 38, no. 3, pp. 459–466, 2016.

[8] P. Tsikouras et al., 'Cervical cancer: screening, diagnosis and staging', J buon, vol. 21, no. 2, pp. 320–325, 2016.

[9] C. L. Siegel et al., 'ACR Appropriateness Criteria® pretreatment planning of invasive cancer of the cervix', Journal of the American College of Radiology, vol. 9, no. 6, pp. 395–402, 2012.

ISBN 979-8-3503-5094-4

[10] . Šarenac and M. Mikov, 'Cervical cancer, different treatments and importance of bile acids as therapeutic agents in this disease', Frontiers in Pharmacology, vol. 10, p. 484, 2019.
[11] Y. Chen et al., 'Deep transfer learning for histopathological diagnosis of cervical cancer using convolutional neural networks with visualization schemes', Journal of Medical Imaging and Health Informatics, vol. 10, no. 2, pp. 391–400, 2020.
[12] A. Manna, R. Kundu, D. Kaplun, A. Sinitca, and R. Sarkar, 'A fuzzy rank-based ensemble of CNN models for classification of cervical cytology', Scientific Reports, vol. 11, no. 1, p. 14538, 2021.
[13] H. Chen et al., 'CytoBrain: cervical cancer screening system based on deep learning technology', Journal of Computer Science and Technology, vol. 36, pp. 347–360, 2021.
[14] A. Tripathi, A. Arora, and A. Bhan, 'Classification of cervical cancer using Deep Learning Algorithm', in 2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS), 2021, pp. 1210–1218.
[15] T. Haryanto, I. S. Sitanggang, M. A. Agmalaro, and R. Rulaningtyas, 'The utilization of padding scheme on convolutional neural network for cervical cell images classification', in 2020 International conference on computer engineering, network, and intelligent multimedia (CENIM), 2020, pp. 34–38.
[16] R. Gorantla, R. K. Singh, R. Pandey, and M. Jain, 'Cervical cancer diagnosis using cervixnet-a deep learning approach', in 2019 IEEE 19th international conference on bioinformatics and bioengineering (BIBE), 2019, pp. 397–404.
[17] F. B. Akyol and O. Altun, 'Detection of cervix cancer from pap-smear images', Sakarya University Journal of Computer and Information Sciences, vol. 3, no. 2, pp. 99–111, 2020.
[18] W. Chen, W. Shen, L. Gao, and X. Li, 'Hybrid Loss-Constrained Lightweight Convolutional Neural Networks for Cervical Cell Classification', Sensors, vol. 22, no. 9, p. 3272, 2022.
[19] M. Fang, X. Lei, B. Liao, and F.-X. Wu, 'A Deep Neural Network for Cervical Cell Classification Based on Cytology Images', IEEE Access, vol. 10, pp. 130968–130980, 2022.
[20] Z. K. Şentürk and U. Süleyman, 'An improved deep learning based cervical cancer detection using a median filter based preprocessing', Avrupa Bilim ve Teknoloji Dergisi, no. 32, pp. 50–58, 2022.
[21] R. Maurya, N. N. Pandey, and M. K. Dutta, 'VisionCervix: Papanicolaou cervical smears classification using novel CNN-Vision ensemble approach', Biomedical Signal Processing and Control, vol. 79, p. 104156, 2023.
[22] M. Karri, C. S. R. Annavarapu, S. Mallik, Z. Zhao, and U. R. Acharya, 'Multi-class nucleus detection and classification using deep convolutional neural network with enhanced high dimensional dissimilarity translation model on cervical cells', Biocybernetics and Biomedical Engineering, vol. 42, no. 3, pp. 797–814, 2022.
[23] H. Basak, R. Kundu, S. Chakraborty, and N. Das, 'Cervical cytology classification using PCA and GWO enhanced deep features selection', SN Computer Science, vol. 2, no. 5, p. 369, 2021.
[24] F. Ucar and D. Korkmaz, 'COVIDiagnosis-Net: Deep Bayes-SqueezeNet based diagnosis of the coronavirus disease 2019 (COVID-19) from X-ray images', Medical hypotheses, vol. 140, p. 109761, 2020.
[25] A. B. Yilmaz, Y. S. Taspinar, and M. Koklu, 'Classification of Malicious Android Applications Using Naive Bayes and Support Vector Machine Algorithms', International Journal of Intelligent Systems and Applications in Engineering, vol. 10, no. 2, pp. 269–274, 2022.
[26] S. Zhang, 'Cost-sensitive KNN classification', Neurocomputing, vol. 391, pp. 234–242, 2020.
[27] S. Zhang, J. Li, and Y. Li, 'Reachable distance function for KNN classification', IEEE Transactions on Knowledge and Data Engineering, 2022.
[28] T. Singh and R. Ashmi, 'Jha Multinomial Logistic Regression and Hybrid GLCM for Cervical Cancer Diagnosis and Prognosis', in International Conference on Innovations in Control, Communication and Information Systems (lCICCI-2017), 2017.
[29] S. Srinivasan, A. B. K. Raju, S. K. Mathivanan, P. Jayagopal, J. C. Babu, and A. K. Sahu, 'Local-Ternary-Pattern-Based Associated Histogram Equalization Technique for Cervical Cancer Detection', Diagnostics, vol. 13, no. 3, p. 548, 2023.

ISBN 979-8-3503-5094-4

Comparison of Deep Learning Models for Efficient Classification of Gastric Abnormalities

Ramesh Munirathinam
Department of Biomedical Engineering
Karpagam Academy of Higher Education
Coimbatore, India
rameshmunirathinam@gmail.com

M Latha
Department of ECE
Rajalakshmi Engineering College
Chennai, India
latha.m@rajalakshmi.edu.in
*corresponding author

M Muthulakshmi
Department of ECE
Amrita School of Engineering
Amrita Vishwa Vidyapeetham
Chennai, India
m_muthulakshmi@ch.amrita.edu
*corresponding author

Maram Prathibha Reddy
Department of ECE
Amrita School of Engineering
Amrita Vishwa Vidyapeetham
Chennai, India

Naveen A
Department of Biomedical Engineering
Karpagam Academy of Higher Education
Coimbatore, India
20bebme029@kahedu.edu.in

M Tamilnidhi
Department of ECE
Karpagam College of Engineering
Coimbatore, India
tamilnidhi.ece@gmail.com

***Abstract*—Gastric cancer stands out as one of the most prevalent malignancies globally, holding the fourth position in terms of cancer-related fatalities. This intricate disease is shaped by a mix of environmental and genetic factors that play a role in its onset and advancement. In this study, 288 histopathological images comprising normal stage, stage_1 and stage_2 obtained from gastric tissue dataset are considered. Squeezenet and Inception V3 deep learning models are used to extract pixel level features. Efficiency of neural network in classifying the tissue changes is accessed by varing the activation and solvers. Results show that the features extracted using squeezenet and inception V3 models has the ability of differentiating tissue changes. to achieve an accuracy of 96in differentiating the tumor cells. High value of 99.1% for MCC clearly indicates that the features extracted using inception v3 features has the ability to identifying the tissue changes. ReLu-L-BFGS-B displays better peformance in identifying the tissue changes with an accuracy of 98.3% and 97.2% using inception v3 and squeezenet deep learning model. From the results, neural network model with logistic-adam and ReLu-L-BFGS-B is found to show comparatively high peformance in detecting the tissue changes in the images with accuracy and AUC values of 98.3% and 100% respectively.. From the study it is concluded that the inception v3 features has better ability than the features obtained from squeezenet model. This may be due to the overfitting of data which results as a results of increased feature length of squeezenet model compared to inception v3 model.**

***Index Terms*—Deep learning, Normal, Stage 1, Stage 2, Orange data mining tool, Neural network, Actvation, Solver**

I. Introduction

Cancer occurs when cells start growing abnormally due to changes in gene, causing them to multiply uncontrollably and potentially harm the body. If not treated, it can become life- threatening [1]. Gastric cancer (GC) stands out as one of the most prevalent malignancies globally, holding the fourth position in terms of cancer-related fatalities. This intricate disease is shaped by a mix of environmental and genetic factors that play a role in its onset and advancement [2]. As per the Global Cancer Observatory project by the International Agency for Research on Cancer, it is estimated that around 1089103 new cases of gastric cancer were diagnosed globally in 2021. Furthermore, there were 768793 deaths associated with GC [3]. Gastric cancer can manifest with various indicators such as difficulty swallowing , reduced appetite , weight loss,, gastrointestinal bleeding, abdominal swelling, and signs of anemia [4]. Males are affected two times more frequently than females, especially in high-risk countries, and they usually occur between the ages of 60 to 80 [5]. The best and standard method for the early detection of gastric cancer and premalignant lesions is endoscopy with appropriate biopsy [6]. After an endoscopic biopsy, the histological confirmation and staging of gastric cancer involve employing CT scans, PET scans, endoscopic ultrasonography, and laparoscopy for a comprehensive assessment [7]. Gastric cancer progresses through stages as it invades various layers of tissue, including the lamina propria or muscularis mucosae, submucosa, muscularis propria, subserosal connective tissue, serosa, and neighboring structures [8]. Systemic chemotherapy, radiotherapy, surgery, immunotherapy, and targeted therapy have all demonstrated effectiveness in treating GC [9]. Orange is an open – source data mining or machine learning program. Orange is useful for visualizing and analyzing exploratory data. It provides a platform selection, predictive modeling, and recommendation systems [10].

II. Literature Review

In literature, Convolutional neural networks were employed by Sin-Ae- lee et al. 2021 to classify gastric cancer from endoscopic images of 470. This produced a roc curve of 0.96

ISBN 979-8-3503-5094-4

[11]. Using morphological features from optical coherence tomography images, Luo et al. conducted experimental research on the imaging of stomach cancerous tissues and classified human stomach cancer. Across all five classifiers—logistic regression, SVM, KNN, Random forest, and the traditional threshold method—the top feature demonstrated above 95% accuracy [12]. Jiang HL et al. employed a deep neural network and transfer-learning techniques for the classification of gastrointestinal endoscopic color images into either benign ulcers or malignancy and their study incorporated Inception, ResNet, and VGGNet models, yielding impressive area under the curve (AUC) values of 0.95, 0.97, and 0.85, respectively [13]. Osamu Lizuka et al. utilized Convolutional Neural Networks for the pathological classification of stomach and colonic epithelial tumors, demonstrating a remarkable accuracy rate of 99% [14]. Using the Inception Resnet V2 model, Bum-Joo Cho et al. were able to classify stomach neoplasms in 5017 endoscopic images, achieving an accuracy of 84.6% [15]. By using hybrid features , Sharmila Joseph et al. were able to classify multiclass intestinal illnesses with a maximum accuracy of 99.2% and a precision of 99.1% [16]. In 2020, Zhigang et al. accomplished near-perfect sensitivity, reaching nearly 100%, and an average specificity of 80.6% when applying deep learning techniques to histopathological diagnosis. Their study involved the analysis of 3,212 images in diverse datasets [17]. In their 2022 study, Xiao Guan et al. employed a dataset comprising 347 patients, with 242 individuals allocated for training purposes and the remaining 105 for testing. Through the utilization of computed tomography, the training phase yielded an impressive area under the curve (AUC) of 0.9978, while the testing phase demonstrated a high AUC of 0.9914 [18]. Using the VGG16, ResNet-18, and GoogLeNet models, Yogapriya et al. evaluated the categorization of GI tract diseases and found that the VGG-16 model had an accuracy of 96.33% [19]. Deep learning models are also found to be highly significant in classifying respiratory and cardiac abnormalities [20], [21].

III. Methodology

The workflow of the proposed methodology is shown in Fig.1. Histopathology images are obtained from Kaggle repository and the feature extraction and classification algorithms are carried out using orange data mining tool.

A. Dataset

The images for the study is obtained from gastric tissue dataset. The data used for the study comprises of 288 images which includes normal stage (98), stage_1 (90) and stage_2 (100).

B. Feature Extraction and Classification techniques

The histopathological images are subjected to feature extraction phase using Squeezenet and Inception V3 deep learning models. The feature length of squeezenet and inception v3 model are 1000 and 2048 respectively. Feature extraction holds significance in image retrieval, image processing, data mining, and computer vision. It involves the extraction of pertinent details from raw data, facilitating the identification of relevant patterns and information[1]. For feature extraction, we employed deep learning models such as Inception v3 and SqueezeNet.

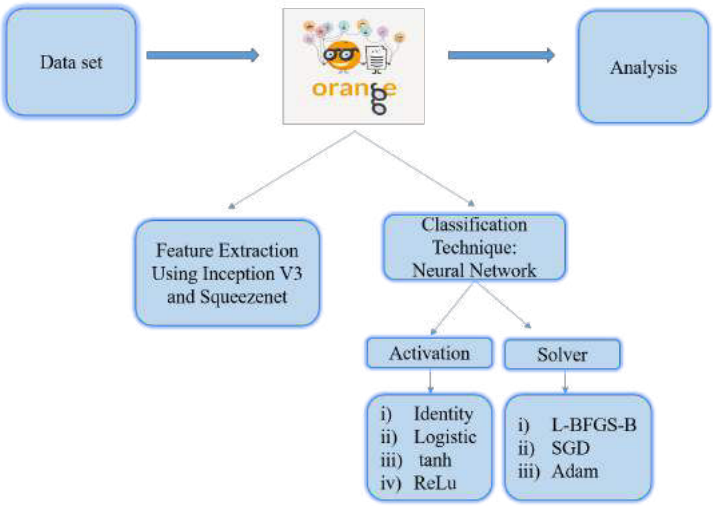


Fig. 1. The prposed work flow

1) Inception V3: A deep learning model for image classification called Inception V3 is based on convolutional neural networks.42 layers make up the inception v3 model.

2) Squeezenet: A kind of deep convolutional neural network (CNN) architecture called SqueezeNet is intended for image categorization applications. It contains 18 layers.The network architecture incorporates ”fire modules” that effectively collect both local and global information in an image by combining 1x1 and 3x3 convolutions.

IV. Results and Discussions

TABLE I
Performance of neural network in differentiating the gastric cancer tissue using features of Inception V3

	CA	F1	Prec	Rec	MCC	Spec
Identity- L-BFGS-B	0.962	0.962	0.962	0.962	0.943	0.981
Identity-SGD	0.948	0.948	0.948	0.948	0.922	0.974
Identity-Adam	0.948	0.948	0.949	0.948	0.922	0.974
Logistic-L-BFGS-B	0.972	0.972	0.972	0.972	0.95	0.986
Logistic-SGD	0.965	0.965	0.96	0.965	0.949	0.982
Logistic-Adam	0.979	0.979	0.980	0.97	0.969	0.989
tanh-L-BFGS-B	0.979	0.979	0.980	0.979	0.969	0.989
tanh-SGD	0.969	0.968	0.970	0.969	0.954	0.984
tanh-Adam	0.965	0.965	0.966	0.965	0.949	0.982
ReLu-L-BFGS-B	0.983	0.983	0.983	0.983	0.974	0.991
ReLu-SGD	0.955	0.954	0.955	0.955	0.933	0.977
ReLu-Adam	0.979	0.979	0.980	0.979	0.969	0.989

The representative histopathology images of normal stage, stage_1 and stage_2 are shown in Fig.2 . It can be observed

ISBN 979-8-3503-5094-4

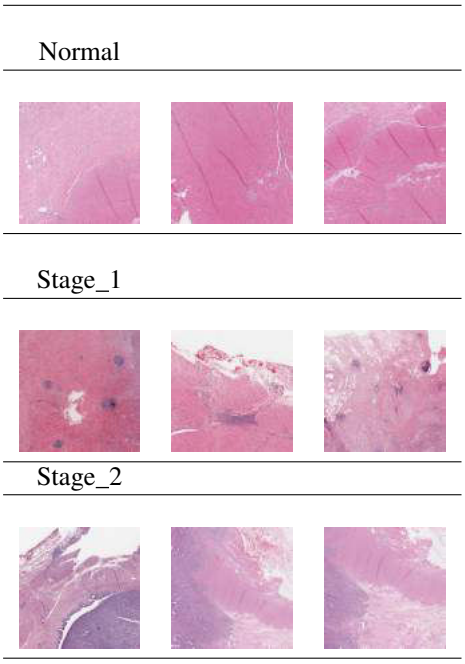


Fig. 2. Representative histopathological images of gastri cancer considered for the analysis

that the tissue changes is varied between the abnormalities. Squeezenet model extracts 1000 feature values from the histopathological images of gastric cancerous tissue and are subjected to classification using Neural network by varying the activation and solvers Table.2 shows the models perfomance inception v3 model features, sequentially. The models performance was achieved using stratified 10-fold cross validation. It can be seen that the activation Identity with L-BFGS-B solver has achieved an accuracy of 96.2% which is comparatively higher than the perfomance of Identity with SGD and adam solver which is 94.8% for both solvers. The activation-Identity with L-BFGS-B solver has the highest TP rate and recall value of 96.2% and it has also achieved 96.2% in precision and 98.1% in specificty, followed by Identity with SGD and Adam found to have the similiar performance in all the perfomance metrics which is produced 94.8% in classification accuray, F1, precison and recall, 92.2% in MCC and 97.4% in specificity. Using logistic activation, the adam solver was most accurate and specificty technique with 97.9% respectively. It can also obtained higher true positive rate and recall value is 97.9% which is comparatively higher than the other solvers in logistic activation. The MCC value 98.9% which is produced by logistic activation with adam solver used to differentiating the abnormalities. Using tanh activation, the solvers are L-BFGS-B, SGD and adam obtained 97.9%, 96.8% and 96.5% classification accuray respectively. The L-BFGS-B solver has higher true positive rate and recall value which is 97.9%, followed by the SGD obtained 96.9%, and Adam solver with 96.5%. The L-BFGS-B was the most specific technique with 98.9%. The ReLu activation model is found to have the highest accuray of 98.3% compared to logistic, identity and tanh models. The ReLu with L-BFGS-B displays the better performance than the ReLu with other SGD and adam models. ReLu with L-BFGS-B has the better performance of classifying the gastric tissues with the MCC values of 99.1%, followed by 97.7% for the SGD model and 98.9% for the adam model. It can also obtained higher true positive rate and recall value which is 98.3% than other activation models. According to the analysis, the activation Relu with L-BFGS-B solver model outperforms the other models in all metrics such as classification accuary, F1, precision, recall, MCC and specificity.

Confusion matrix for NN Identity - L-BFGS-B (showing number of instances)

		Predicted			
		normal	stage_1	stage_2	∑
Actual	normal	96	1	1	98
	stage_1	5	83	2	90
	stage_2	0	2	98	100
	∑	101	86	101	288

Fig. 3. Confusion matrix for neural network with Identity-L-BFGS-B model achieves a 96.2% accuracy using 10-fold cross validation. The instances shows the presence of 11 incorrect predictions out of 288 images

Confusion matrix for NN Identity - SGD (showing number of instances)

		Predicted			
		normal	stage_1	stage_2	∑
Actual	normal	96	1	1	98
	stage_1	5	81	4	90
	stage_2	1	3	96	100
	∑	102	85	101	288

Fig. 4. Confusion matrix for neural network with Identity-SGD model achieves a 94.8% accuracy usin 10-fold cross validation. The main diagnol (96, 81, 96) gives the correct predictions. The instances shows the presence of 15 wrong predictions out of 228 images

Confusion matrix for NN Identity - Adam (showing number of instances)

		Predicted			
		normal	stage_1	stage_2	∑
Actual	normal	95	2	1	98
	stage_1	7	80	3	90
	stage_2	1	1	98	100
	∑	103	83	102	288

Fig. 5. Confusion matrix for neural network with Identity-adam models gives achieves a 94.8% accuracy using 10-fold cross validation. The plotted confusion matrix classified 95, 80 and 98 of normal, stage_1 and stage_2 precisely. It also shows the presence of15 wrong predictions out of 228 images

From Fig.3-14 the identity activation with SGD and adam solver got less accuracy and high number incorrect predictions comapred to other models. The Relu-LFGS-B model got high accuaracy and less number of incorrect predictions comapred to other models and also, its achieved higher number of true positive rate instances comapared to other models in the study

Table.II demonstrates the metrics performance such as classification accuray, F1 score, precision, recall, MCC and

ISBN 979-8-3503-5094-4

Confusion matrix for NN Logistic - L-BFGS-B (showing number of instances)

		Predicted			
		normal	**stage_1**	**stage_2**	**∑**
Actual	**normal**	97	1	0	**98**
	stage_1	4	84	2	**90**
	stage_2	0	1	99	**100**
	∑	**101**	**86**	**101**	**288**

Fig. 6. Confusion matrix for neural network with logistic-L-BFGS-B model achieves a 97.2% accuracy using 10 fold cross validation. The main diagonal (97, 84,99) gives the correction predictions. The plotted confusion matrix shows the presence of 8 incorrect predictions out of 288 images

Confusion matrix for NN Logistic - SGD (showing number of instances)

		Predicted			
		normal	**stage_1**	**stage_2**	**∑**
Actual	**normal**	98	0	0	**98**
	stage_1	5	81	4	**90**
	stage_2	0	1	99	**100**
	∑	**103**	**82**	**103**	**288**

Fig. 7. Confusion matrix for neural network with logistic-SGD model achieves a 96.5% accuracy using 10 fold cross validation. The outcome shows the presence of 10 incorrect predictions out of 288 images

Confusion matrix for NN Logistic - Adam (showing number of instances)

		Predicted			
		normal	**stage_1**	**stage_2**	**∑**
Actual	**normal**	98	0	0	**98**
	stage_1	4	84	2	**90**
	stage_2	0	0	100	**100**
	∑	**102**	**84**	**102**	**288**

Fig. 8. Confusion matrix for neural network with logistic-adam model achieves a 97.9% accuracy using 10 fold cross validation. It correctly classified 98 images in normal, 84 images in stage_1 and 100 images in stage_2. Also, its misclassified 6 instances out of 288 instances

Confusion matrix for NN tanh - L-BFGS-B (showing number of instances)

		Predicted			
		normal	**stage_1**	**stage_2**	**∑**
Actual	**normal**	98	0	0	**98**
	stage_1	4	85	1	**90**
	stage_2	0	1	99	**100**
	∑	**102**	**86**	**100**	**288**

Fig. 9. Confusion matrix for neural network with tanh-L-BFGS-B model achieves a 98.3% accuracy using 10 fold cross validation. The main diagonal (98, 85, 99) gives the correct predictions. The instances displays the presence of 6 incorrect predictions out of 228 instances

Confusion matrix for NN tanh - SGD (showing number of instances)

		Predicted			
		normal	**stage_1**	**stage_2**	**∑**
Actual	**normal**	98	0	0	**98**
	stage_1	5	82	3	**90**
	stage_2	0	1	99	**100**
	∑	**103**	**83**	**102**	**288**

Fig. 10. Confusion matrix for neural network with tanh-sgd model achieves a 96.9% accuracy using 10 fold cross validation. Its correctly predicted 98 instances in normal, 82 instances in stage_1 and 99 instances in stage_2 out of 288 instances

Confusion matrix for NN tanh - Adam (showing number of instances)

		Predicted			
		normal	**stage_1**	**stage_2**	**∑**
Actual	**normal**	97	1	0	**98**
	stage_1	7	81	2	**90**
	stage_2	0	0	100	**100**
	∑	**104**	**82**	**102**	**288**

Fig. 11. Confusion matrix for neural network with tanh-adam model achieves a 96.5% accuracy using 10 fold cross validation. The plotted confusion matrix shows the model performance of outcomes. The outcome displays the 97, 81 and 100 correct predictions out of 288 instances and also shows the 10 misclassified predictions out 288 images

Confusion matrix for NN Relu - L-BFGS-B (showing number of instances)

		Predicted			
		normal	**stage_1**	**stage_2**	**∑**
Actual	**normal**	98	0	0	**98**
	stage_1	3	86	1	**90**
	stage_2	0	1	99	**100**
	∑	**101**	**87**	**100**	**288**

Fig. 12. Confusion matrix for neural network with ReLu-L-BFGS-B model achieves a 98.3% accuracy using 10 fold cross validation. It correctly classified and mis-classified instances 98 instances in normal , 86 instances in stage_1, 99 instances in stage_2 and 5 in all abnormalities respectively.

Confusion matrix for NN Relu - SGD (showing number of instances)

		Predicted			
		normal	**stage_1**	**stage_2**	**∑**
Actual	**normal**	97	1	0	**98**
	stage_1	6	80	4	**90**
	stage_2	0	2	98	**100**
	∑	**103**	**83**	**102**	**288**

Fig. 13. Confusion matrix for neural network with ReLu-SGD model achieves a 95.5% accuracy using 10 fold cross validation. It's correctly predicted 97, 80 and 80 in normal, stage_1 and stage_2 respectively. Also, the confusion matrix displays the presence of 13 incorrect predictions out of 288 instances

ISBN 979-8-3503-5094-4

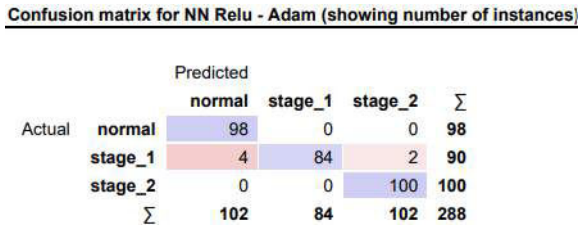


Fig. 14. Confusion matrix for neural network with ReLu-adam model achieves a 97.9% accuracy using 10 fold cross validation. The plotted confusion matrix displays the correctly classified images which is 98, 84 and 100 in normal, stage_1 and stage_2 precisely. This also demonstrates the presence of 6 wrong predictions out of 288 instances

TABLE II
PERFORMANCE OF NEURAL NETWORK IN DIFFERENTIATING THE GASTRIC CANCER TISSUE USING FEATURES OF SQUEEZENET

	CA	F1	Prec	Rec	MCC	Spec
Identity-L-BFGS-B	0.927	0.927	0.928	0.927	0.891	0.964
Identity-SGD	0.934	0.934	0.935	0.934	0.902	0.967
Identity-Adam	0.941	0.941	0.941	0.941	0.912	0.970
Logistic-L-BFGS-B	0.938	0.937	0.937	0.938	0.906	0.969
Logistic-SGD	0.885	0.884	0.893	0.885	0.833	0.942
Logistic-Adam	0.969	0.969	0.969	0.969	0.953	0.984
tanh-L-BFGS-B	0.969	0.969	0.969	0.969	0.953	0.984
tanh-SGD	0.958	0.958	0.959	0.958	0.938	0.979
tanh-Adam	0.965	0.965	0.966	0.965	0.948	0.983
ReLu-L-BFGS-B	0.972	0.972	0.972	0.972	0.958	0.986
ReLu-SGD	0.951	0.951	0.952	0.951	0.927	0.976
ReLu-Adam	0.962	0.962	0.962	0.962	0.943	0.981

specificity obtained using Squeezenet model feature. The model performnace was achieved using stratified 10-fold cross validation. It can be seen that the neural network activation identity with L-BFGS-B, SGD and adam solvers produced an accuracy values are 92.7%, 93.3% and 94.1% respectively. Comparatively, the adam solver has the higher true positive rate or recall value with 94.1%. It ca be also the most specificity with 97.%. Using logistic activation, the adam optimizer got higher accuracy compared to other solvers which is 93.8%. The adam optimizer has obtained higher true positive rate and recall value of 96.9% and also the high sspecificity with 98.4% compared to other optimizers in logistic function. Using tanh activation, the L-BFGS-B solver found to displays better performance in classification accuracy, recall and precison values with 96.9% which is comparetively higher than the other solvers. It obtained the MCC value of 98.4% which is equal to the MCC (98.3%) value of adam solver model. The performance of the ReLu model evaluated and analysed using different solvers. Higher performance of ReLu model achieved using L-BFGS-B with 97.2% classification accuracy and recall. ReLu with L-BFGS-B has the better performance of classifying the gastric tissues with the MCC values of 95.8%, followed by 97.7% for the SGD model and 98.9% for the adam model. According to the analysis, the ReLu-L-BFGS-B model outperforms the other models in all performance metrics such classification accuracy, F1- score, precision and recall with 97.2%. Comparatively, it is the best specific technique with 98.6% than other models.

Finally, from Table.I and Table.II the ReLu-LFGS-B algorithm achieves a better performance in both inception v3 and squeezenet model. Significantly, it achieved the highest classification accuracy, F1 score, recall, MCC and specificty in Inception V3 model feature.

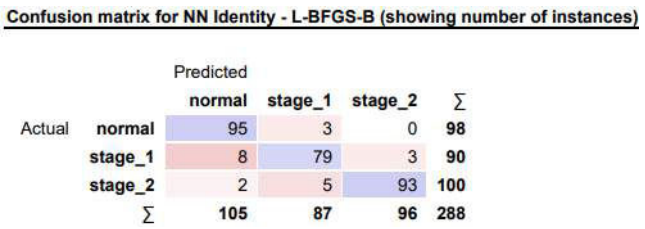


Fig. 15. Confusion matrix for neural network with Identity-L-BFGS-B model achieves a 92.7% accuracy using 10-fold cross validation. The instances shows the presence of 21 incorrect predictions out of 288 images

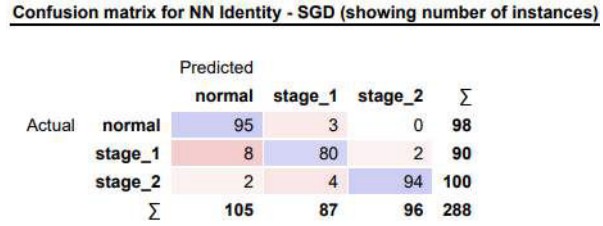


Fig. 16. Confusion matrix for neural network with Identity-SGD model achieves a 93.4% accuracy usin 10-fold cross validation. The main diagnol (95, 80, 94) gives the correct predictions. The instances shows the presence of 19 wrong predictions out of 228 images

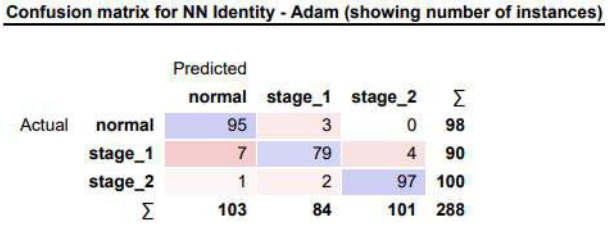


Fig. 17. Confusion matrix for neural network with logistic-adam models gives achieves a 94.1% accuracy using 10-fold cross validation. The plotted confusion matrix classified 95, 79 and 97 of normal, stage_1 and stage_2 precisely. It also shows the presence of 15 wrong predictions out of 228 images

From Fig.15-26 the logistic activation with adam solver got less accuracy and high number incorrect predictions comapred to other models. The Relu-LFGS-B model got high accuaracy and less number of incorrect predictions comapred to other models and also, its achieved higher number of true positive rate instances comapared to other models

ISBN 979-8-3503-5094-4

Confusion matrix for NN Logistic - L-BFGS-B (showing number of instances)

		Predicted			
		normal	stage_1	stage_2	∑
Actual	normal	92	3	3	98
	stage_1	5	82	3	90
	stage_2	0	4	96	100
	∑	97	89	102	288

Fig. 18. Confusion matrix for neural network with Logistic-L-BFGS-B model achieves a 93.8% accuracy using 10 fold cross validation. The main diagonal (92, 82,96) gives the correction predictions. The plotted confusion matrix shows the presence of 18 incorrect predictions out of 288 images

Confusion matrix for NN Logistic - SGD (showing number of instances)

		Predicted			
		normal	stage_1	stage_2	∑
Actual	normal	94	4	0	98
	stage_1	21	66	3	90
	stage_2	2	3	95	100
	∑	117	73	98	288

Fig. 19. Confusion matrix for neural network with logistic-SGD model achieves a 88.5% accuracy using 10 fold cross validation. The outcome shows the presence of 33 incorrect predictions out of 288 images

Confusion matrix for NN Logistic - Adam (showing number of instances)

		Predicted			
		normal	stage_1	stage_2	∑
Actual	normal	97	1	0	98
	stage_1	4	84	2	90
	stage_2	0	2	98	100
	∑	101	87	100	288

Fig. 20. Confusion matrix for neural network with logistic-adam model achieves a 96.9% accuracy using 10 fold cross validation. It correctly classified 97 images in normal, 84 images in stage_1 and 98 images in stage_2. Also, its misclassified 9 instances out of 288 instances

Confusion matrix for NN tanh - L-BFGS-B (showing number of instances)

		Predicted			
		normal	stage_1	stage_2	∑
Actual	normal	97	1	0	98
	stage_1	4	84	2	90
	stage_2	0	2	98	100
	∑	101	87	100	288

Fig. 21. Confusion matrix for neural network with tanh-L-BFGS-B model achieves a 96.9% accuracy using 10 fold cross validation. The main diagonal (97, 84, 98) gives the correct predictions. The instances displays the presence of 9 incorrect predictions out of 228 instances

Confusion matrix for NN tanh - SGD (showing number of instances)

		Predicted			
		normal	stage_1	stage_2	∑
Actual	normal	97	1	0	98
	stage_1	6	82	2	90
	stage_2	0	3	97	100
	∑	103	86	99	288

Fig. 22. Confusion matrix for neural network with tanh-sgd model achieves a 95.8% accuracy using 10 fold cross validation. Its correctly predicted 97 instances in normal, 82 instances in stage_1 and 97 instances in stage_2 out of 288 instances

Confusion matrix for NN tanh - Adam (showing number of instances)

		Predicted			
		normal	stage_1	stage_2	∑
Actual	normal	97	1	0	98
	stage_1	6	83	1	90
	stage_2	0	2	98	100
	∑	103	86	99	288

Fig. 23. Confusion matrix for neural network with tanh-adam model achieves a 96.5% accuracy using 10 fold cross validation. The plotted confusion matrix shows the model performance of outcomes. The outcome displays the 97, 83 and 98 correct predictions out of 288 instances and also shows the 10 misclassified predictions out 288 images

Confusion matrix for NN Relu - L-BFGS-B (showing number of instances)

		Predicted			
		normal	stage_1	stage_2	∑
Actual	normal	96	2	0	98
	stage_1	2	86	2	90
	stage_2	0	2	98	100
	∑	98	90	100	288

Fig. 24. Confusion matrix for neural network with ReLu-L-BFGS-B model achieves a 97.2% accuracy using 10 fold cross validation. It correctly classified and mis-classified instances 96 instances in normal , 86 instances in stage_1, 98 instances in stage_2 and 8 in all abnormalities respectively.

Confusion matrix for NN Relu - SGD (showing number of instances)

		Predicted			
		normal	stage_1	stage_2	∑
Actual	normal	95	3	0	98
	stage_1	8	81	1	90
	stage_2	0	2	98	100
	∑	103	86	99	288

Fig. 25. Confusion matrix for neural network with ReLu-SGD model achieves a 95.1% accuracy using 10 fold cross validation. It's correctly predicted 95, 81 and 98 in normal, stage_1 and stage_2 respectively. Also, the confusion matrix displays the presence of 14 incorrect predictions out of 288 instances

ISBN 979-8-3503-5094-4

Confusion matrix for NN Relu - Adam (showing number of instances)

		Predicted			
		normal	**stage_1**	**stage_2**	**∑**
Actual	**normal**	96	2	0	**98**
	stage_1	5	83	2	**90**
	stage_2	0	2	98	**100**
	∑	**101**	**87**	**100**	**288**

Fig. 26. Confusion matrix for neural network with ReLu-adam model achieves a 96.2% accuracy using 10 fold cross validation. The plotted confusion matrix displays the correctly classified images which is 96, 83 and 98 in normal, stage_1 and stage_2 precisely. This also demonstrates the presence of 6 wrong predictions out of 288 instances

V. Conclusion

The experimental results demonstrates that the histopathological images show significant difference between the gastric tissue changes. In this work the efficiency of squeezenet features and inception v3 features are analyzed and compared using the performance of neural network.The performance of neural network is also evaluated by changing the activation and solver.The neural network with ReLu-L-BFGS-B achives better performance in both squeeznet and inception v3 model feature with an classification accuracy and recall values of 98.3% and 97.2% respectively. Results conclude that the neural network with ReLu-L-BFGS-B model has better performance with inception v3 model and it outperforms the squeeznet model in all metrics such as classification accuracy, recall, F1, precision, MCC and specificity. The high values of AUC (100%) in logistic-adam model and MCC (99.1%) in ReLu-LFGS-B obtained using inception v3 feature model and its clearly indicated tha the model is high stability and realiability for identifying the tissue changes

References

[1] R. W. Ruddon, Cancer biology. Oxford University Press, 2007.

[2] J. Machlowska, J. Baj, M. Sitarz, R. Maciejewski, and R. Sitarz, 'Gastric cancer: epidemiology, risk factors, classification, genomic characteristics and treatment strategies', International journal of molecular sciences, vol. 21, no. 11, p. 4012, 2020.

[3] A. P. Thrift and H. B. El-Serag, 'Burden of gastric cancer', Clinical Gastroenterology and Hepatology, vol. 18, no. 3, pp. 534–542, 2020.

[4] G. Maconi, G. Manes, and G. B. Porro, 'Role of symptoms in diagnosis and outcome of gastric cancer', World journal of gastroenterology: WJG, vol. 14, no. 8, p. 1149, 2008.

[5] R. Sitarz, M. Skierucha, J. Mielko, G. J. A. Offerhaus, R. Maciejewski, and W. P. Polkowski, 'Gastric cancer: epidemiology, prevention, classification, and treatment', Cancer management and research, pp. 239–248, 2018.

[6] V. Pasechnikov, S. Chukov, E. Fedorov, I. Kikuste, and M. Leja, 'Gastric cancer: prevention, screening and early diagnosis', World journal of gastroenterology: WJG, vol. 20, no. 38, p. 13842, 2014.

[7] E. C. Smyth, M. Nilsson, H. I. Grabsch, N. C. T. van Grieken, and F. Lordick, 'Gastric cancer', The Lancet, vol. 396, no. 10251, pp. 635–648, 2020.

[8] J. T. P. D. Hallinan and S. K. Venkatesh, 'Gastric carcinoma: imaging diagnosis, staging and assessment of treatment response', Cancer imaging, vol. 13, no. 2, p. 212, 2013.

[9] S. S. Joshi and B. D. Badgwell, 'Current treatment and recent progress in gastric cancer', CA: a cancer journal for clinicians, vol. 71, no. 3, pp. 264–279, 2021.

[10] S. Kodati and R. Vivekanandam, 'Analysis of heart disease using in data mining tools Orange and Weka', Glob. J. Comput. Sci. Technol. C Softw. Data Eng, vol. 18, no. 1, pp. 16–22, 2018.

[11] S.-A. Lee, H. C. Cho, and H.-C. Cho, 'A novel approach for increased convolutional neural network performance in gastric-cancer classification using endoscopic images', IEEE Access, vol. 9, pp. 51847–51854, 2021.

[12] S. Luo, Y. Fan, W. Chang, H. Liao, H. Kang, and L. Huo, 'Classification of human stomach cancer using morphological feature analysis from optical coherence tomography images', Laser physics letters, vol. 16, no. 9, p. 095602, 2019.

[13] J. H. Lee et al., 'Spotting malignancies from gastric endoscopic images using deep learning', Surgical endoscopy, vol. 33, pp. 3790–3797, 2019.

[14] O. Iizuka, F. Kanavati, K. Kato, M. Rambeau, K. Arihiro, and M. Tsuneki, 'Deep learning models for histopathological classification of gastric and colonic epithelial tumours', Scientific reports, vol. 10, no. 1, p. 1504, 2020.

[15] B.-J. Cho et al., 'Automated classification of gastric neoplasms in endoscopic images using a convolutional neural network', Endoscopy, vol. 51, no. 12, pp. 1121–1129, 2019.

[16] J. Sharmila Joseph and A. Vidyarthi, 'Multiclass Gastrointestinal Diseases Classification Based on Hybrid Features and Duo Feature Selection', Journal of Biomedical Nanotechnology, vol. 19, no. 2, pp. 288–298, 2023.

[17] Z. Song et al., 'Clinically applicable histopathological diagnosis system for gastric cancer detection using deep learning', Nature communications, vol. 11, no. 1, p. 4294, 2020.

[18] X. Guan, N. Lu, and J. Zhang, 'Computed tomography-based deep learning nomogram can accurately predict lymph node metastasis in gastric cancer', Digestive Diseases and Sciences, vol. 68, no. 4, pp. 1473–1481, 2023.

[19] J. Yogapriya, V. Chandran, M. G. Sumithra, P. Anitha, P. Jenopaul, and C. Suresh Gnana Dhas, 'Gastrointestinal tract disease classification from wireless endoscopy images using pretrained deep learning model', Computational and mathematical methods in medicine, vol. 2021, 2021.

[20] S. Palaniappan, S. V. S. Sripriya, A. R. L. Pranathi, and M. Muthulakshmi, 'Diagnosis of Acute Respiratory Syndromes from X-Rays using Customised CNN Architecture', in 2022 IEEE International Conference on Artificial Intelligence in Engineering and Technology (IICAIET), 2022, pp. 1–5.

[21] M. Muthulakshmi, G. Kavitha, and N. Aishwarya, 'Multi-Objective Butterfly Optimization for Feature and classifier parameter's selection in Diagnosis of Heart Failure types Using CMR images', in 2022 IEEE Global Conference on Computing, Power and Communication Technologies (GlobConPT), 2022, pp. 01–06.

ISBN 979-8-3503-5094-4

Investigation of Thyroid Nodule Detection Using Ultrasound Images with Deep Learning

C. Kaushik Viknesh
Department of ECE
CEG Campus, Anna University
Chennai-600025, India
ecekv12@gmail.com

S.Kanimozhi
Department of ECE
CEG Campus, Anna University
Chennai-600025, India
kanimozhiss255@gmail.com

R.Thirumalai selvi
Department of ECE
CEG Campus, Anna University
Chennai-600025, India
srimerraselvi@gmail.com

***Abstract*— The foremost predominant shape of endocrine cancer is thyroid cancer, and its rate has been relentlessly expanding to its worldwide scale. This audit centers on the challenging assignment of distinguishing thyroid knobs utilizing ultrasound images. Ultrasound offers a cost-effective and non-invasive implies of visualizing the thyroid organ and the encompassing tissues. Truly, consideration has transcendently coordinated to experimentally approve imaging highlights utilized in ultrasonography for thyroid knob discovery. The essential goal of computerized knob location is to achieve a level of precision that rivals that of fine needle goal biopsy. In light of this objective, we have presented an inventive approach for knob acknowledgment, tackling the control of convolutional neural systems. Convolutional neural systems (CNNs) have a particular advantage in naturally learning high-level and various leveled reflections from visual information through end-to-end preparing. The adequacy of knob distinguishing proof is too surveyed from three diverse points: multiscale prediction engineering, post-processing strategies, and plans of the loss work. Clinical information assesses the execution of this approach, with comparisons made to the ground truth labels given by restorative experts. Moreover, when a thyroid knob is recognized, it gets to be conceivable to determine whether the condition has spread to adjacent zones, potentially driving to progressed persistent results. In rundown, the executions of our proposed strategy holds significant potential to upgrade the viability and accuracy of thyroid knob discovery and characterization.**

Keywords— Thyroid Gland, Thyroid Cancer, Convolutional neural network, Automated nodule detection, Multiscale prediction, Post-processing method, Design of the loss function, Thyroid nodule detection.

I. INTRODUCTION

Thyroid knobs are lifted regions of tissue or liquid that more often than not frame inside an something else solid thyroid organ. These knobs can be either hyperplastic or tumorous, but as it were a little rate of them turn out to be threatening. Numerous little knobs, particularly in the event that they do not cause any side effects, regularly go unnoticed. In any case, as knobs develop bigger or begin causing side effects, they may in the long run require therapeutic consideration [1]. Figure 1 outlines a goiter with either a single knob (uni-nodular), different knobs (multi- nodular), or diffuse knobs.

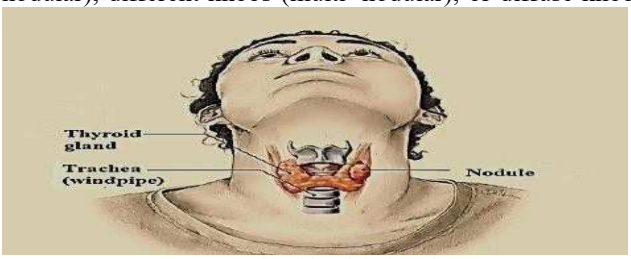


Figure 1. Thyroid nodule detection

These anomalous developments of thyroid tissue often manifest as a lump in the throat and are typically located near the outer edge of the thyroid gland. In cases where they are sizable, a thyroid nodule may become visible as a noticeable lump in the front of the neck [2]. Conversely, a thyroid cyst might present as a depression, often containing solid components. Thyroid cysts usually arise from degenerating thyroid adenomas, which are generally benign but may occasionally contain malignant solid components. Upon the discovery of a nodule during a physical examination, a patient may be referred to an endocrinologist, thyroidologist, or otolaryngologist for further evaluation [3].

Typically, an ultrasound is employed to assess the gland's condition and confirm the presence of a nodule. Functional thyroid disorders, such as Hashimoto's thyroiditis, a common cause of benign nodular goiter, can be diagnosed by measuring thyroid-stimulating hormone and anti-thyroid antibodies. Fine-needle biopsy is also utilized for cytopathology assessment. It is worth noting that thyroid nodules are relatively common in children and young adults, with approximately half of individuals experiencing one at some point. However, they are usually only detected by a healthcare professional during a routine health examination or incidentally while investigating an unrelated condition [4]. The evaluation of these incidental nodules is outlined in the following section, as shown in Table 1.

Table.1 Suggested workups by nodule characteristic

S. No	Workups	Characteristics
1.	Elevated PET activity or regional infiltration or Abnormal lymph nodes	Highly probable ultrasound imaging.
2.	Multiple masses	Probable ultrasound imaging
3.	Single growth in an individual under 35 years of age	Probable ultrasound examination for lesions measuring 1 cm or morein adults, and for all sizes in children. No requirement if Smaller than 1 cm in adults.
4.	Single growth in an individual aged 35 or older	Probable ultrasound imaging for lesions measuring 1.5 cm or more. No requirement for sizes less than 1.5

ISBN 979-8-3503-5094-4

The primary objective of this work is to scrutinize thyroid nodules and ascertain their proximity to adjacent organs or tissues. Additionally, the study aims to explore the key advantages of Convolutional Neural Networks (CNNs), which include the ability to automatically grasp intricate patterns from visual data through end-to-end training. This involves organizing the components associated with the CNN to effectively predict thyroid diseases using deep learning algorithms. Automated thyroid detection using ultrasound images can be achieved using Convolutional Neural Networks (CNNs) [5]. CNNs are commonly used deep learning models for image analysis tasks. Here's a basic overview of how CNNs can be applied to automated thyroid detection:

Data Collection and Annotation: Initiate the process by collecting a dataset of thyroid ultrasound images and labeling them appropriately, distinguishing between "normal" and "abnormal" or specific thyroid conditions. Ensure the dataset contains a sufficient number of examples from different classes for effective training [6].

Preprocessing: Preprocess the ultrasound images to standardize their size, resolution, and format. This ensures consistency during training [7]

Dataset Partitioning: Divide the dataset into two components: a training subset and a testing subset. Train the CNN model on the training subset, followed by the evaluation of performance using the testing subset [8].

CNN Architecture: Design the architecture of the Convolutional Neural Network model, often comprising fully connected layers for classification, pooling layers for down-sampling, and multiple convolutional layers. The specific architecture and hyper-parameters depend on the complexity of the task and available resources [9].

Training: As the model is trained, it becomes proficient at identifying relevant features from ultrasound images and leveraging this information to make predictions based on the provided labels [10].

Validation: Monitor the model's performance using validation data points throughout the training process to avoid overfitting, where the model becomes too specific to the training data and performs ineffectively on new data. Adjustments to training parameters or model architecture can be made based on validation results [11].

Testing and Evaluation: After the training process concludes, the model undergoes evaluation using a test set to gauge its effectiveness with previously unseen data. Various evaluation metrics, including but not limited to accuracy, precision, recall, and the F1-score, serve as valuable tools for assessing the model's capability in detecting thyroid anomalies [12].

Deployment: Following satisfactory performance, the model can be deployed for automated thyroid detection. Input modern ultrasound images into the trained model, which can make predictions regarding the presence of anomalies and potentially identify specific conditions [13].

II. MATERIALS AND METHODS

A. *Strategies of CNN*

A Convolutional neural network is specific neural network architecture primarily designed for deep learning tasks involving visual depiction analysis. CNNs employ a scientific handle utilizing convolution in one of their layers, Replacing traditional network duplication, multilayer perceptrons are adapted versions of CNNs, often being fully connected networks where each node in a layer is linked to every neuron in the subsequent layer. However, this "dense" network structure can be susceptible to overfitting, a problem commonly addressed through regularization techniques such as network reduction (e.g., skipped connections, dropout) or applying penalties to parameters during training (like weight decay) [14].

Neural network layers consist of interconnected nodes or neurons that process input data and generate output. Each layer within a neural network serves distinct roles:

Input Layer: The initial layer where input data is initially processed into the network. It only takes in the input and passes it to the subsequent layers without performing any computations [15].

Hidden Layers: These layers follow the input and precede the output layer. They perform the majority of a network's computations, including tasks like feature extraction and transformation [16].

Output Layer: The final network layer produces the network's output.

While CNNs utilize a hierarchical arrangement of data and filters to construct increasingly complex patterns, they differ from traditional approaches by dividing the input into smaller, more manageable features represented by channels. These channels are then applied over different regions of the input to extract relevant information. As the network progresses through its layers, these features are gradually integrated and organized into more complex patterns. This hierarchical process enables CNNs to effectively capture intricate representations while avoiding the issue of overfitting [17].

The input to a CNN is typically a tensor of the required structure: (number of input channels) x (input height) x (input width). After traversing through a convolutional layer, the image undergoes a transformation, becoming a feature map or activation map defined by the dimensions: (number of feature maps) x (feature map height) x (feature map width). As illustrated in Figure 2, convolutional layers perform convolution on the input and pass the result to the subsequent framework, akin to how neurons in the occipital cortex respond to specific stimuli. Each convolutional unit processes information related to specific features, contributing to the network's ability to identify complex patterns in data [18].

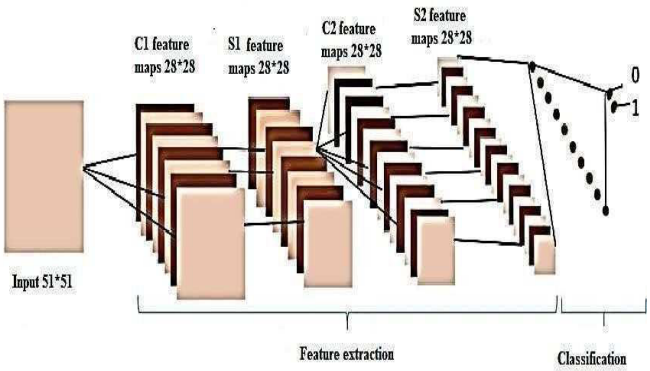


Figure 2. Convolutional Layers

ISBN 979-8-3503-5094-4

In an open field, fully connected feed-forward neural networks are suitable for feature learning and data classification. However, this architecture faces challenges when handling larger inputs, especially high-resolution images. The limitation arises because each pixel in such images constitutes a significant input feature, resulting in a requirement for an enormous number of neurons to effectively process the input. The convolution operation plays a crucial role in reducing the numerical representation sent to the CNN, ensuring that the neural network only receives essential features for image classification [19]. This step not only improves the network's accuracy but also conserves computational resources during training. The result of the convolution operation is commonly referred to as a feature map, convolutional map, or activation map. To generate a feature map, a filter, also known as a kernel, is used, often in the form of a 3 by 3 matrix. The feature map is created by conducting element-wise multiplication between the filter and the input image and summing the resulting values [20]. This process is carried out by sliding the filter over the input image, with the sliding action occurring in predefined steps. When constructing a CNN, the size of the filter and the step size can be manually configured as needed. The network utilizes a combination of nonlinear and pooling layers alongside various convolutional layers. As the image progresses through each layer, the output of the first convolutional layer serves as the input for the subsequent layer. This pattern continues with each subsequent convolutional layer in the network [21]. Nonlinear activation layers are often introduced right after the convolutional layer to introduce non-linearity. This is done because convolution is a linear operation, while images are inherently non-linear [22]. There are different types of activation functions used in CNNs:

Sigmoid: The mathematical expression for the sigmoid nonlinearity is "() = 1/(1+e)." It compresses a real-numbered input into a range between 0 and 1. A significant drawback of the sigmoid function is that when the activation approaches either extreme end (0 or 1), the slope becomes extremely close to zero, posing a problem during back propagation [23].

Tanh: It is an activation function that maps real numbers into a range between -1 and 1. Similar to the sigmoid function, tanh activations can experience saturation, but their outputs are centered on zero [24].

ReLU (Rectified Linear Unit): It is defined by the function "() = max(0, x)," which means it remains inactive at zero and becomes linear for positive values of x. ReLU is known for its reliability and its ability to accelerate training around six times faster than sigmoid and tanh activation functions. One downside of ReLU is that it can be sensitive during training, causing neurons to stop updating altogether when a large gradient flows through it. However, this issue can be managed by selecting an appropriate learning rate [25].

Padding is employed when applying the convolution operation to prevent the reduction in the size of the feature map. This helps maintain the output feature map's dimensions the same as the input image [26]..

By "cushioning" those pictures with zeros around their edges, the compelling measure of the input picture is increased. Subsequently, when the channel is utilized on the picture, it produces a featuremap with indistinguishable measurements to the input picture." (as shownin Figure 3).

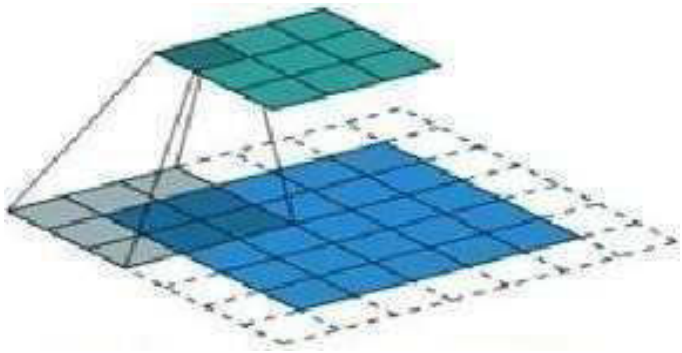


Figure 3. Uncolored area represents the padded area

Padding during the convolution process serves several crucial purposes. Firstly, it helps preserve more of the original data within the image. Additionally, it ensures that the convolution operation effectively considers the edges of the picture [27]. When constructing a CNN, you have the flexibility to specify the padding strategy that best suits your needs, or you can choose to forego padding altogether. The two most frequently used options are "Valid" and "Same." In "Valid" mode, no padding is used, while "Same" mode ensures that padding is added to match the feature map's dimensions with those of the input picture, as illustrated in Figure 4.

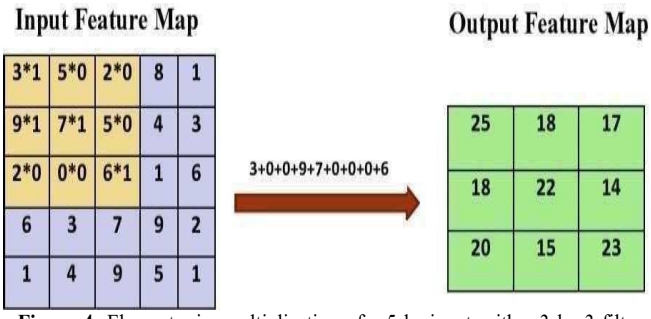


Figure 4. Element-wise multiplication of a 5 by input with a 3 by 3 filter.

The Fully Connected Layer: The fully connected layer aptly describes its function. In contrast to partially connected layers where the input pixel values are not directly linked to their output layer, the fully connected layer establishes direct connections between every node in the output layer and every node in the preceding layer. As depicted in Figure 5, this layer plays a key role in the classification task, utilizing the features extracted from earlier layers and their respective channels. While pooling and convolutional layers often use ReLU functions, fully connected layers typically employ a softmax activation function to effectively classify the inputs and generate probabilities within the range of 0 to 1 [28].

ISBN 979-8-3503-5094-4

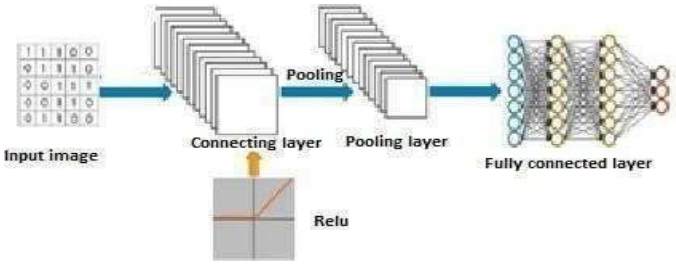


Figure 5. Fully connected layer

At the beginning of this section, I would like to elucidate the method of supervised machine learning, which serves as the foundation for our model. Supervised machine learning represents one of the techniques within the broader domain of machine learning, wherein the model undergoes training using input data and corresponding expected output data [29]. To develop such a model, it is essential to go through the following stages.

Model Construction: The construction of the model revolves around the selection of the machine learning algorithm. In the context of the work, neural networks were employed [30]. The algorithm typically takes the following form: initiate the model object with 'model = Sequential()', followed by the addition of layers with their respective types. After a sufficient number of layers have been added, the model is compiled. At this point, Keras interfaces with Tensor Flow to build the model. During the compilation, it is crucial to define both a loss function and an optimizer algorithm expressed as 'model.compile(loss='loss function name', optimizer='optimizer algorithm name'). The loss function assesses the accuracy of each model prediction [31].

Data Preprocessing: Before commencing model training, it is essential to normalize the data for later use.
Model Training: In this stage, the model is trained using the provided training data and their corresponding expected outputs. The process involves using 'model.fit'. Progress can be monitored in the console as the script runs, and upon completion, the final accuracy of the model is reported.
Model Testing: Following the model's construction and training, it is time to subject it to testing. During this stage, a separate dataset, one that has never been exposed to the model, is introduced. This serves to validate the model's real accuracy.
Model Saving: Once it is determined that the model exhibits the required performance, it can be saved using 'model.save('name_of_file.h5')'.
Model Deployment: Eventually, the saved model is deployed for real-world applications, constituting the stage known as deploying the model for its evaluation. This allows the model to be used for assessing new data and making predictions. Figure 6 shows the flowchart of the overall process [32].

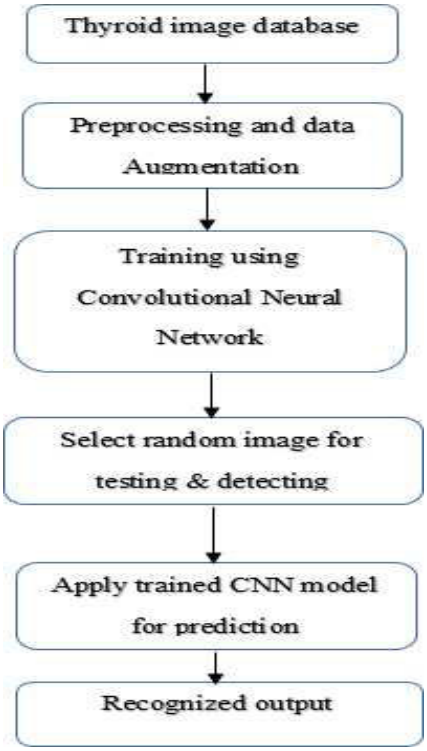


Figure 6. Block Diagram

B. Dataset Preparation

Compile a comprehensive dataset of thyroid images from diverse sources such as hospitals, medical databases, or research institutions. It is crucial to ensure that this dataset includes an adequate number of cases representing both positive instances (indicating thyroid abnormalities) and negative instances (representing non-thyroid conditions). To denote the presence or absence of thyroid abnormalities, each thyroid image should be meticulously annotated with corresponding labels. These annotations serve as the foundation for training the Convolutional Neural Network (CNN) [33]. Typically, the task of examining images and assigning appropriate labels is carried out by annotators, often comprising medical professionals with expertise in thyroid-related diagnoses. As part of the preprocessing stage, enhance the quality and suitability of the images for training purposes using various preprocessing techniques. Some commonly employed preprocessing steps include:

Rescaling: Ensure consistency across the dataset by resizing the images to a uniform size. Square dimensions are often preferred, aligning with the typical requirements of CNN architectures. This step helps standardize the images for effective training and analysis [34].

Normalization: Establish a consistent scale for the pixel values within the images. This can be achieved by dividing the pixel values by 255, resulting in values ranging from 0 to 1. Normalization plays a significant role in stabilizing the training process and facilitating improved convergence during model training [35]. The careful curation of a diverse dataset and the application of preprocessing techniques are crucial steps in ensuring the efficacy of the Convolutional Neural Network for accurate thyroid anomaly detection. These measures contribute to the overall robustness and generalization capability of the model for real-world applications.

ISBN 979-8-3503-5094-4

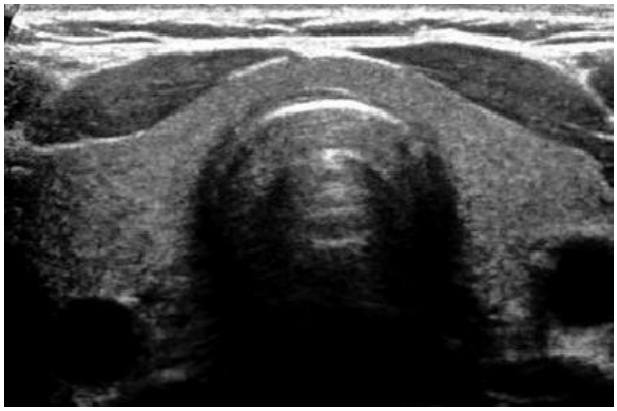


Figure 7. Ultrasound image

Contrast Adjustment: Modify the contrast levels within the image highlights to enhance the model's effectiveness.

Noise Reduction: Employ techniques such as blurring or denoising filters to eliminate any artifacts or noise in the images. The preprocessing strategies outlined above include resizing, normalization, contrast adjustment, and noise reduction applied to the input ultrasound thyroid data image. The resulting processed image is illustrated in Figure 7.

C. Classification

A CNN architecture is composed of several components, including fully connected layers for classification, pooling layers for feature extraction from images, and multiple convolutional layers. Initially, the CNN model is trained using labeled training data. During this training phase, the model acquires the ability to identify patterns and features within ultrasound images indicative of either normal or abnormal thyroid conditions. This training involves the use of optimization techniques like gradient descent, encompassing forward propagation, loss calculation, and backward propagation to iteratively adjust the model's parameters [36]. The testing and training process is outlined in the flowchart, as shown in Figure 8.

Configure the Model: Implement the chosen model architecture with random initial weights.
Forward Propagation: Execute the model on the training images to obtain predicted class probabilities.
Loss Calculation: Employing a loss function, assess the model's performance by evaluating the discrepancy between the forecasted probabilities and the ground truth labels.
Backward Propagation: Employ backpropagation to compute the gradient of the loss concerning the model's parameters.
Validation Set Model Evaluation: Keep track of the model's progress by regularly assessing its performance on the validation set. This systematic process ensures that the CNN model is effectively trained, adjusting its parameters to accurately classify thyroid conditions based on the input ultrasound images. The inclusion of validation set evaluations helps monitor the model's generalization capabilities beyond the training data.

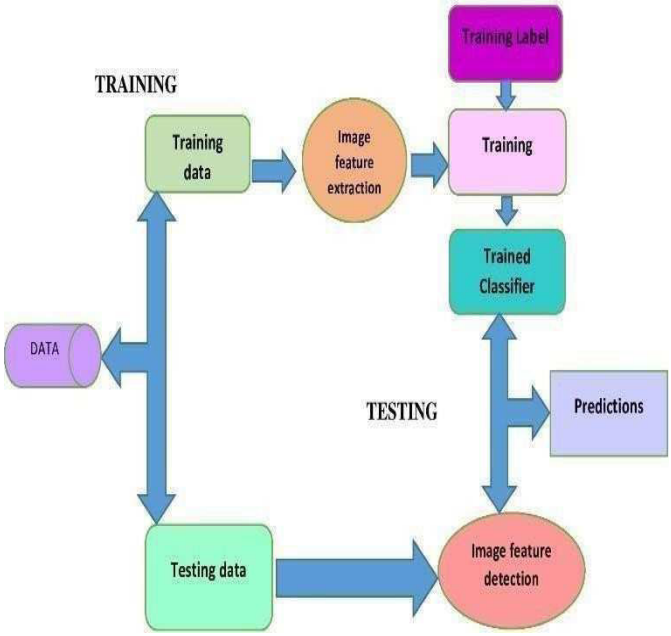


Figure.8. Block diagram for test and training

D. Testing Stage

Test Results: Generate a distinct testing dataset comprising images that were not previously encountered during validation or training phases.
Forward Propagation: Feed the testing images through the trained model to obtain the predicted class probabilities.
Class Prediction: For each image, select a class and label it with the highest probability as the predicted label.
Performance Evaluation: Contrast the forecasted labels with the true labels in the testing dataset to assess the model's precision and other performance measures like precision, recall, and F1 score. This process ensures a thorough evaluation of the model's performance on unseen data, providing insights into its ability to generalize and make accurate predictions beyond the training and validation sets.
, recall, and F1 score.

III. RESULTS & DISCUSSION

A. Analysis Of Thyroid Nodule

The dataset utilized in this study is sourced from the paper titled "Convolutional neural network aimed at thyroid nodule detection in ultrasound images." In addition to the augmentation methods applied, various preprocessing steps are performed on the ultrasound input images, including resizing, normalization, contrast adjustment, noise removal, and more. Following the preprocessing stage, a convolutional neural network is employed for training, enabling the model to learn and identify patterns and features within the ultrasound images indicative of either normal or abnormal thyroid conditions. Consequently, random images will be selected for testing, and the trained CNN model will be used to make predictions, as illustrated in Figure 9. Images representing both benign and malignant cases will also be included in the evaluation, as depicted in Figure 10.

ISBN 979-8-3503-5094-4

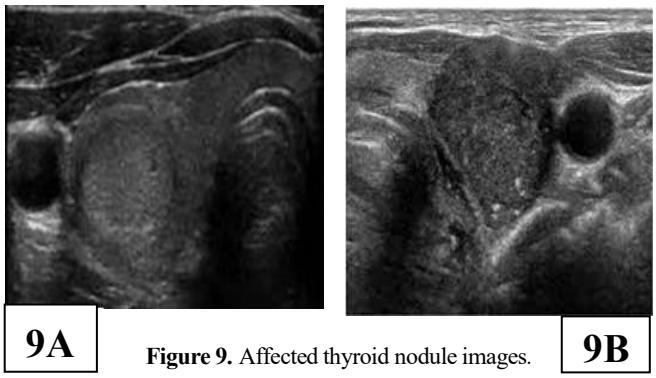


Figure 9. Affected thyroid nodule images.

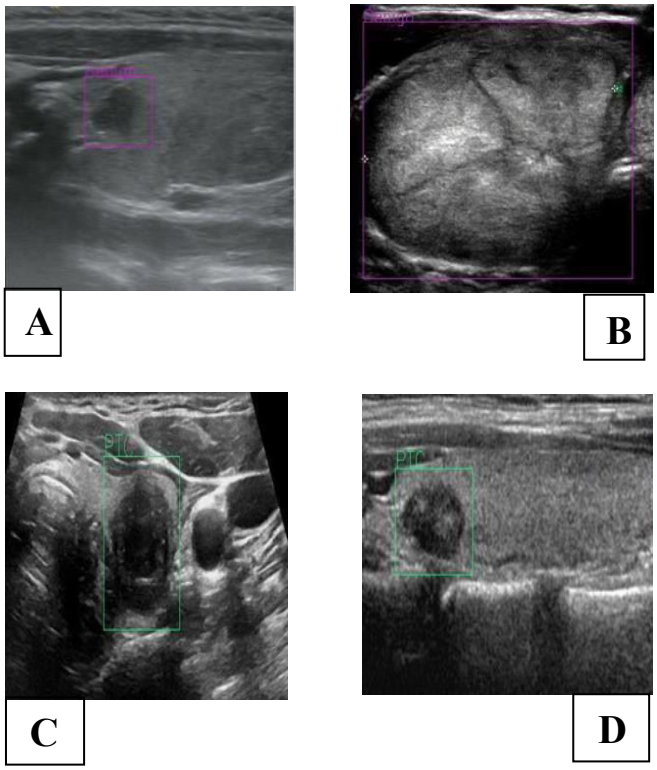


Figure 10. Affected benign and malignant image.

IV. CONCLUSION

In this study, ultrasound images serve as input data and undergo several preprocessing steps, including resizing, standardization, contrast adjustment, noise reduction, and augmentation. Subsequently, a convolutional neural network is employed for training and testing to automatically predict thyroid nodules. This task is particularly challenging due to the dataset's wide variety of nodule sizes and the limited number of thyroid nodules available. The experimental results from the testing and training process demonstrate that our proposed method effectively detects thyroid nodules, emphasizing the importance of accuracy in the detection process. Recently, ultrasound thyroid images have been utilized for classification during the testing and training phases, distinguishing between benign and malignant cases. Looking ahead, we plan to leverage deep learning through convolutional neural networks to ascertain not only whether nodules have extended into adjacent organs but also to assess the extent of organ involvement and the accuracy of identifying affected regions.

REFERENCES

[1] J. Kim, J. E. Gosnell, and S. A. Roman, ‘‘Geographic influences in the global rise of thyroid cancer,’’ Nature Rev. Endocrinol., vol. 16, no. 1, pp. 17–29, Jan 2020.

[2] H. Ye, J. Hang, X. Chen, D. Xu, J. Chen, X. Ye, and D. Zhang, ‘‘An intelligent platform for ultrasound diagnosis of thyroid nodules,’’ Sci. Rep., vol. 10, no.1, Aug. 2020.

[3] O. Moussa, H. Khachnaoui, R. Guetari, and N. Khlifa, ‘‘Thyroid nodules classification and diagnosis in ultrasound images using fine-tuning deep convolutional neural network,’’ Int. J. Imag. Syst. Technol., vol. 30, no.1, pp.185–195, Aug. 2019.

[4] Automatic Thyroid Ultrasound Image Classification Using Feature Fusion Network XIAOHUI ZHAO 1 , XUEQIN SHEN2, WENBO WAN3 , vol 10, no.1, Mar 2022.

[5] H. Ye, J. Hang, X. Chen, D. Xu, J. Chen, X. Ye, and D. Zhang, ‘‘An intelligent platform for ultrasound diagnosis of thyroid nodules,’’ Sci. Rep., vol. 10, no. 1, Aug. 2020, Art. no. 13223.

[6] C. Chen, L. Zhan, X. Pan, Z. Wang, X. Guo, H. Qin, F. Xiong, W. Shi, M. Shi, F. Ji, Q. Wang, N. Yu, and R. Xiao, ‘‘Automatic recognition of auditory brainstem response characteristic waveform based on bidirectional long short-term memory,’’ Frontiers Med., vol. 7, Jan. 2021, Art. no. 613708.

[7] Y. Zheng, S. Xu, Z. Zheng, L. Wu, L. Chen, and W. Zhan, ‘‘Ultrasonic classification of multicategory thyroid nodules based on logistic regression,’’ Ultrasound Quart., vol. 36, no. 2, pp. 149–157, Jun. 2020.

[8] L. Wang, L. Zhang, M. Zhu, X. Qi, and Z. Yi, ‘‘Automatic diagnosis for thyroid nodules in ultrasound images by deep neural networks,’’ Med. Image Anal., vol. 61, Apr. 2020, Art. no. 101665.

[9] C. Chen, K. Zhou, M. Zha, X. Qu, X. Guo, H. Chen, Z. Wang, and R. Xiao, ‘‘An effective deep neural network for lung lesions segmentation from COVID-19 CT images,’’ IEEE Trans. Ind. Informat., vol. 17, no. 9, pp. 6528–6538, Sep. 2021.

[10] C. Chen, R. Xiao, T. Zhang, Y. Lu, X. Guo, J. Wang, H. Chen, and Z. Wang, ‘‘Pathological lung segmentation in chest CT images based on improved random Walker,’’ Comput. Methods Programs Biomed., vol. 200, Mar. 2021, Art. no. 105864.

[11] C. Chen, J. Zhou, K. Zhou, Z. Wang, and R. Xiao, ‘‘DW-UNet: Loss balance under local-patch for 3D infection segmentation from COVID-19 CT images,’’ Diagnostics, vol. 11, no. 11, p. 1942, Oct. 2021.

[12] Z.Chunyu and C. Xianyi, ‘‘Automatic segmentation of thyroid ultrasound images based on adaptive filtering combined with T-snake model,’’ Appl. Res. Comput., vol. 37, no. 3, pp. 944–946, 2020.

[13] R. Song, L. Zhang, C. Zhu, J. Liu, J. Yang, and T. Zhang, “Thyroid nodule ultrasound image classification through hybrid feature cropping network,” IEEE Access, vol. 8, pp. 64064–64074, Mar. 2020.

[14] W. Yang et al., “Integrate domain knowledge in training multi-task cascade deep learning model for benign-malignant thyroid nodule classification on ultrasound images,” Eng. Appl. Artif. Intell., vol. 98, Feb. 2021.

[15] G. Shi et al., “Knowledge-guided synthetic medical image adversarial augmentation for ultrasonography thyroid nodule classification,” Comput. Methods Programs Biomed., vol. 196, Jun. 2020.

[16] V. Kumar et al., “Automated segmentation of thyroid nodule, gland, and cystic components from ultrasound images using deep learning,” IEEE Access, vol. 8, pp. 63 482–63 496, 2020.

[17] D. Koundal, B. Sharma, and Y. Guo, “Intuitionistic based segmentation of thyroid nodules in ultrasound images,” Comput. Biol. Med., vol. 121, 2020.

ISBN 979-8-3503-5094-4

[18] . Z.Jin, Y. Zhu, S. Zhang, F. Xie, M. Zhang, Y. Zhang, X. Tian, J. Zhang, Y. Luo, and J. Cao, ''Ultrasound computer-aided diagnosis (CAD) based on the thyroid imaging reporting and data system (TI-RADS) to distinguish benign from malignant thyroid nodules and the diagnostic performance of radiologists with different diagnostic experience,'' Med. Sci. Monitor, vol. 26, pp. 1643–3750, Jan. 2020

[19] O.Moussa, H. Khachnaoui, R. Guetari, and N. Khlifa, ''Thyroid nodules classification and diagnosis in ultrasound images using fine-tuning deep convolutional neural network,'' Int. J. Imag. Syst. Technol., vol. 30, no. 1, pp. 185–195, Aug. 2019

[20] Y.Xie, Y. Xia, J. Zhang, Y. Song, D. Feng, M. Fulham, and W. Cai, ''Knowledge-based collaborative deep learning for benign-malignant lung nodule classification on chest CT,'' IEEE Trans. Med. Imag., vol. 38, no. 4, pp. 991–1004, Apr. 2019.

[21] D.China, A. Illanes, P. Poudel, M. Friebe, P. Mitra, and D. Sheet, ''Anatomical structure segmentation in ultrasound volumes using cross frame belief propagating iterative random walks,'' IEEE J. Biomed. Health Inform., vol. 23, no. 3, pp. 1110–1118, May 2019.

[22] W.Song et al., "Multitask cascade convolution neural networks for automatic thyroid nodule detection and recognition," IEEE J. Biomed. Health Informat., vol. 23, no. 3, pp. 1215–1224, May 2019.

[23] LWang et al., "Automatic thyroid nodule recognition and diagnosis in ultrasound imaging with the YOLOv2 neural network," World J. Surg. Oncol., vol. 17, no. 1, pp. 1–9, Dec. 2019

[24] A.Illanes, N. Esmaeili, P. Poudel, S. Balakrishnan, and M. Friebe, "Parametrical modelling for texture characterization-a novel approach applied to ultrasound thyroid segmentation," PLoS One, vol. 14, Jan. 2019.

[25] P.Poudel, A. Illanes, E. J. G. Ataide, N. Esmaeili, S. Balakrishnan, and M. Friebe, "Thyroid ultrasound texture classification using autoregressive features in conjunction with machine learning approaches," IEEE Access, vol. 7, pp. 79 354–79 365, 2019.

[26] R. Zhu, Z. Wang, Y. Zhang, B. Yu, M. Qi, X. Feng, C. Wu, Y. Cui, L. Huang, F. Li, and F. Zhou, ''Integrating five feature types extracted from ultrasonograms to improve the prediction of thyroid papillary carcinoma,'' IEEE Access, vol. 7, pp. 101820–101828, Jul. 2019.

[27] P. Poudel, A. Illanes, E. J. G. Ataide, N. Esmaeili, S. Balakrishnan, and M. Friebe, ''Thyroid ultrasound texture classification using autoregressive features in conjunction with machine learning approaches,'' IEEE Access, vol. 7, pp. 79354–79365, Jun. 2019.

[28] A. E. Pablo and T. Michel, ''Normalized mutual information feature selection,'' IEEE Trans. Neural Netw., vol. 20, no. 2, pp. 189–201, Jan. 2019.

[29] Y. Zhang, H. Lai, and W. Yang, "Cascade UNet and CH-UNet for thyroid nodule segmentation and benign and malignant classification," in Proc. Int. Conf. Med. Image Comput. Comput.- Assist. Interv., 2020.

[30] H. Fukui, T. Hirakawa, T. Yamashita, and H. Fujiyoshi, "Attention branch network: Learning of attention mechanism for visual explanation," in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit., 2019.

[31] W. Song et al., "Multitask cascade convolution neural networks for automatic thyroid nodule detection and recognition," IEEE J. Biomed. Health Informat., vol. 23, no. 3, pp. 1215–1224, May 2019.

[32] X. Chen, J. Yu, S. Kong, Z. Wu, and L. Wen, "Joint anchor-feature refinement for real-time accurate object detection in images and videos," IEEE Trans. Circuits Syst. Video Technol., vol. 31, no. 2, pp. 594–607, Feb. 2021.

[33] A. Abbasian Ardakani, A. Bitarafan-Rajabi, A. Mohammadi, S. Hekmat, A. Tahmasebi, M. B. Shiran, and A. Mohammadzadeh, ''CAD system based on B-mode and color Doppler sonographic features may predict if a thyroid nodule is hot or cold,'' Eur. Radiol., vol. 29, no. 8, pp. 4258–4265, Aug. 2019.

[34] A. A. Parsa and H. Gharib, ''Thyroid nodule: Current evaluationx and management,'' in The Thyroid and Its Diseases: A Comprehensive Guide for the Clinician, M. Luster, L. Duntas, L. Wartofsky, Eds. Cham, Switzerland: Springer, Jan. 2019.

[35] R. Zhu, Z. Wang, Y. Zhang, B. Yu, M. Qi, X. Feng, C. Wu, Y. Cui, L. Huang, F. Li, and F. Zhou, ''Integrating five feature types extracted from ultrasonograms to improve the prediction of thyroid papillary carcinoma,'' IEEE Access, vol. 7, pp. 101820–101828, Jul. 2019.

[36] W. Yang, J. Zhao, Y. Qiang, X. Yang, Y. Dong, Q. Du, G. Shi, and M. B. Zia, ''DScGANS: Integrate domain knowledge in training dual-path semi-supervised conditional generative adversarial networks and s3vm for ultrasonography thyroid nodules classification,'' in Proc. MICCAI, Shenzhen, China, Oct. 2019.

ISBN 979-8-3503-5094-4

Tenth International Conference on Biosignals, Images and Instrumentation, SSNCE, Chennai, March 20 – 22, 2024

TRACK 7

Analysis and Optimization of Foot Pressure Patterns in Pes Planus-A Comprehensive Study

Sakthivel Sankaran[1]
Department of BME
Kalasalingam Academy of research and education
Krishnankoil, Tamilnadu, 626126.
sakthivelsankaran1992@gmail.com

T. Thanga Suriya[2]
Department of BME
Kalasalingam Academy of research and education
Krishnankoil, Tamilnadu, 626126.
9921020011@klu.ac.in

M.Peerkkanriyash[3]
Department of BME
Kalasalingam Academy of research and education
Krishnankoil, Tamilnadu, 626126.
9921020007@klu.ac.in

Preethika Immaculate Britto[4]
Department of BME
King Faisal University
Saudi Arabia.
pbritto@kfu.edu.sa

Abstract-- This work is dedicated to designing and developing a specialized Pressure Mapping Analysis and Evaluation measurement module tailored for individuals with Pes Planus, commonly referred to as flat feet. Pes Planus, a medical condition characterized by flattened arches, alters biomechanics and increases pressure on various foot areas. The proposed module aims to offer a comprehensive solution by assessing and analyzing pressure distribution across the foot sole. This analysis will guide the development of customized footwear, alleviating discomfort and preventing related issues. To achieve accurate pressure mapping, we have fixed pressure mapping sensors in the insole of the shoe, ensuring proper output and effective alignment. A graphical monitor is included to visualize foot pressure distribution. In our future endeavors, we plan to create shoes specifically designed for Pes Planus, enhancing proper alignment to further improve the well-being of individuals with flat feet. This advancement holds the potential to reduce pain, enhance stability, and prevent long-term musculoskeletal problems associated with flat feet.

Key words: Pes planus, criteria, plantar pressure, insole, pressure distribution map, graphical representation.

I. INTRODUCTION

This work addresses flat feet, or pes planus, by combining Create advanced technology with medical expertise a holistic solution. The collaborative approach involves close partnerships with medical professionals and researchers to establish stringent criteria for identifying and categorizing flat feet based on severity levels. Using an Arduino Uno and piezoelectric sensors, strategically placed within a shoe, our hardware measurement system accurately gauges plantar pressure while prioritizing wearer comfort. The data collected is processed through specialized software equipped with intricate algorithms, providing a detailed visualization of high-pressure areas in the feet. This collaborative effort ensures a comprehensive method for identifying, assessing, and rectifying flat feet, culminating in the implement of a specialized shoe tailored to individual needs. The hinges on the integration of medical insights, technological innovation, and collaborative efforts, offering a promising solution for individuals with flat feet.

The proposed solution involves designing a comprehensive method to address pes planus (flat feet) and create a specialized shoe. This includes establishing precise criteria for identifying flat feet in collaboration with medical professionals, developing a hardware measurement system with strategically placed pressure sensors, and integrating the collected data into specialized software for thorough analysis. The focus is on creating a personalized shoe solution that addresses specific pressure points and provides tailored support for individuals with pes planus.

The advantage of this approach lies in it's ability to accurately measure plantar pressure, visualize pressure distribution maps, and analyze data to identify regions requiring correction. By utilizing piezoelectric sensor technology and involving medical experts in the criteria-setting process, the solution ensures the development of custom footwear that effectively rectifies pes planus. The comprehensive report generated provides detailed insights, enabling targeted modifications to enhance foot function and alleviate discomfort associated with flat feet.

The outcome of the pressure mapping analysis, particularly using piezoelectric sensors, has provided valuable insights into the pressure dynamics associated with pes planus. The next phase of the work involves the development of a specialized shoe that integrates these findings to provide targeted support. The anticipated result is a custom-designed shoe that not only addresses the unique pressure distribution patterns identified in the analysis but also aims to improve overall well-being and quality of life for individuals with flat feet. The approach ensures a tailored solution, enhancing foot function and comfort for those affected by pes planus.

The current challenge in addressing flat feet stems from a lack of a systematic method for precise identification and severity categorization, compounded by insufficient integration of medical expertise, research, and technology. Existing solutions fall short in providing effective hardware-based measurement systems to accurately capture plantar pressure, resulting in a deficit in designing corrective and comfortable footwear. This initiative seeks to fill these gaps by holistically combining medical insights, cutting-edge sensor technology, and sophisticated algorithms to craft specialized shoes, promising enhanced comfort, pain reduction, and improved mobility for individuals with varying.

II. LITERATURE REVIEW

Ozlem Feyzioglu, et al [1] Researchers compared gait and foot pressure in individuals with Morton's neuroma (MN) and

ISBN 979-8-3503-5094-4

either flat feet (pes planus) or high arches (pes cavus). Surprisingly, there weren't major differences in overall gait or pressure between the groups, suggesting treatments should focus on reducing MN symptoms directly rather than changing foot structure. However, MN seemed to affect gait more in high-arched feet, leading to shorter stance phases, longer swing phases, and slower steps compared to healthy feet. This suggests MN might disrupt gait timing more in people with supinated feet, highlighting the need for further research in this area. Overall, the study emphasizes the of considering foot posture when managing MN but also suggests focusing on symptom reduction rather than solely correcting foot mechanics

Bente E. Bloks et al, [2] have discussed focused on into Hereditary Motor and Sensory Neuropathies (HMSN), commonly known as Charcot-Marie-Tooth disease, involved a comprehensive analysis of foot deformities across four categories. In a study encompassing 52 individuals with HMSN and 586 healthy controls, the research scrutinized plantar pressure patterns, root mean square deviations (RMSD), center of pressure trajectories, and pressure ratios. Results indicated elevated RMSD values in all foot deformity categories compared to controls, alongside distinct spatial and temporal differences in plantar pressure patterns and center of pressure trajectories among individuals with HMSN. Notably, variations in plantar pressure ratios, specifically the fifth metatarsal head pressure ratio, were observed, differentiating both controls and HMSN cases, as well as the four deformity categories. The study advocates for employing RMSD and the fifth metatarsal head pressure ratio as crucial metrics in evaluating overall abnormality and hindfoot varus deformity, respectively, for effective assessment of surgical interventions in individuals with HMSN.

Ben Efrima, MD et al, [3] delved into the etiology of symptomatic pediatric pes planus (PP) deformity, with a specific focus on the reduced os calcis subtalar joint (OCST) anterior facet morphology. Employing weight-bearing computed tomography (WBCT) and distance mapping (DM), the study aimed to categorize PP subtalar subtypes based on Bruckner's A-D classification system. Through the examination of 40 feet in 25 patients, the research revealed excellent intra- and inter-observer agreement in classifying OCST using DM. Notably, the PP cohort exhibited a higher prevalence of types B and D compared to types A and C within Bruckner's classification. These findings hint at a potential correlation between OCST morphology and PP deformity, underscoring the valuable role of DM in such investigations.

Nadir YALÇIN et al [4] have discussed to compare statically obtained radiographic angles with dynamic plantar pressure distribution measurements for evaluating the medial longitudinal arch (MLA) in healthy individuals. The retrospective analysis of 95 subjects revealed no significant gender-based correlation with the arch index. The talo-first metatarsal and talohorizontal angles correlated significantly with the arch index, while the talocalcaneal and calcaneal pitch angles did not. The arch index method was deemed a simple and reproducible pedobaro graphic measurement for MLA evaluation. However, statically obtained radiographic angles showing correlations with the arch index may yield similar results, suggesting the potential utility of both static and dynamic methods in MLA assessment.

Jun Na Zhai [5] have described the assess plantar pressure differences between flexible flatfoot and normal foot during various walking conditions, investigating the need for treatment and changes in plantar pressure while walking on stairs. Fifteen adults with mild flexible flatfoot, fifteen with severe flexible flatfoot, and fifteen with normal foot were examined. Results indicated significantly increased max force and arch index in severe flatfoot compared to normal foot across different walking conditions. The study emphasized the necessity of treatment for severe flexible flatfoot to prevent further deformation, especially as observed during walking downstairs. Plantar pressure differences were measured using the RS scan system, with data analyzed through repeated measures ANOVA.

Andrew K. Buldt et al, [6] have involving to compare plantar pressure among healthy individuals with normal, pes planus (low arch), and pes cavus (high arch) feet. Ninety-two volunteers were classified based on foot posture indices, and Plantar pressure trials were conducted using barefoot walking system. Analyses revealed significant differences, particularly between planus and cavus foot groups, with distinct characteristics in forefoot pressure and force. The largest effect sizes were observed in peak pressure, pressure-time integral, maximum force, and force-time integral for the lateral forefoot in planus feet compared to normal and cavus feet. The findings emphasize the influence of foot posture on plantar pressures during walking.

Caiting Zhang and Yining Xu [7] This systematic review rigorously assessed interventions for pes cavus, utilizing well-defined inclusion and exclusion criteria. The focus was on studies involving individuals with pes cavus and interventions employing assistive devices, with outcomes measured through plantar pressure and medial longitudinal angle (MLA). Only controlled trials were considered, and exclusion criteria targeted patients to the musculoskeletal contraindications to exercise. The search, spanning from 1963 to May 2023, included PubMed, Scopus, and Web of Science, with a thorough review of reference lists. A comprehensive search strategy, guided by Boolean logic, was executed by two independent reviewers, refined by a third librarian. Data collection, analysis, and risk of bias assessment were conducted meticulously, ensuring a robust evaluation. The synthesis process, involving the screening of 499 titles and abstracts, led to the inclusion of three trials [13,21,38], reflecting the systematic and reliable nature of the review.

Sachini n.k and Kodithuwakku arachchige [8] this article aims to offer a succinct overview of the structural anatomy, pathomechanics, assessment, and effective management of flatfeet. Pes planus, a common concern impacting diverse segments of society, leads to issues such as poor postural stability, injuries, and discomfort. Recognizing the importance of early detection, the article emphasizes the significance of employing various feet assessments, each with its unique advantages and disadvantages. Proper management is crucial, and the selection of appropriate methods varies for each individual. Healthcare professionals, occupational therapists, kinesiologists, biomechanists, coaches, and ergonomists will find this condensed summary

ISBN 979-8-3503-5094-4

valuable in understanding the complexities of pes planus and facilitating early detection for optimal care and prevention.

Sakthivel Sankaran et al, [9] in this study expresses that recent technical innovations have increased the comfort of artificial limbs, yet patient satisfaction remains a challenge. This study introduces a novel solution—a bio-composite prosthesis with a sensor-based system replicating anatomical leg movements. Utilizing an accelerometer sensor and microcontroller, the concept of cloning leg actions enhances the prosthetic experience. Results highlight the device's efficiency, dependent on angle suspension, hinge lubrication, and prosthetic material. This sensor-based technology is pivotal in rehabilitation, ensuring quick data transfer with reduced loss. The prototype successfully mimics various leg gestures, improving the amputee's overall experience. Constructed with jute fiber reinforced epoxy bio composite material, it minimizes skin irritation for above-knee amputees. This proposed prosthetic leg enables patients to perform normal leg actions, enhancing daily life for amputees and disabled individuals with improved gait movement.

Sakthivel Sankaran et al, [10] The utilization of smart shirts for monitoring vital physiological parameters is widespread, particularly in the context of miners' safety. With an alarming number of fatalities due to mine collapses, this innovative smart shirt addresses the pressing need to enhance miners' alertness to potential risks. By continuously monitoring barometric pressure, temperature, humidity, and carbon monoxide levels through sensors, the system aims to alert users to hazardous conditions. This proactive approach empowers both miners and supervisors to oversee these critical parameters, offering a more effective safety solution for underground work, as validated by a comprehensive survey of related works.

Sakthivel Sankaran et al, [11] have Cervicalgia, commonly known as neck discomfort, impacts a considerable portion of the global population due to various factors such as poor posture, accidents, bone disorders, and spinal conditions. Although electric stimulation therapy has proven effective, its high cost poses a challenge. In response to this issue, a research initiative has proposed a novel approach by transforming a muscle stimulator into a wearable coat for self-treatment. The prototype is designed to electronically reactivate physiological muscle functions, leading to a clinical reduction in pain. The hardware system comprises a microcontroller, amplifier IC, and batteries. Clinical trials have demonstrated promising results, indicating effective pain relief for individuals suffering from cervicalgia. This wearable device serves as a cost-effective alternative to traditional methods. Future plans involve advancing electric muscle stimulation technology further, with potential integration of additional sensors to enable comprehensive patient monitoring.

Sakthivel Sankaran et al, [12] The foot, a crucial element in daily life, undergoes weight bearing, pressure, and friction. These factors can lead to damage in the outermost layers of the skin due to pressure and friction, resulting in conditions like plantar warts, foot corns, or calluses. Plantar warts, known as Helomadurem and Helomamolle, manifest as hard or soft corns, causing discomfort during walking. Traditional treatments involve applying specific temperature ranges to deliver thermal energy to the affected tissue. Addressing corns and calluses focuses on managing increased mechanical stress rather than treating a skin disease. Pain often arises from pressure on nerve endings, signaling the need for intervention. Treatment involves reducing growth size and tissue replacement using microwave energy (100-200 KHz) applied through a probe to concentrated hardened areas. Contemporary approaches include salicylic acid application and, in severe cases, surgical intervention.

III. METHODOLOGY

To develop a comprehensive method for addressing pes planus (flat feet) and creating a specialized shoe. Establish the criteria for identifying pes planus. This should involve consulting medical professionals and researchers to define the specific characteristics and severity levels of flat feet that need correction. Design a hardware measurement system to assess plantar pressure. This system should consist of pressure sensors strategically placed on the insoles of shoes. The hardware should be capable of accurately measuring pressure points in the feet while ensuring comfort for the wearer. Select the optimal sensor technology for measuring plantar pressure. Capacitive sensors, piezoresistive sensors, or other options should be evaluated based on their accuracy, durability, and suitability for integration into the shoe design. Once the hardware is developed and integrated into the shoe, load the collected data into specialized software for analysis. This software should include algorithms to process the data effectively.

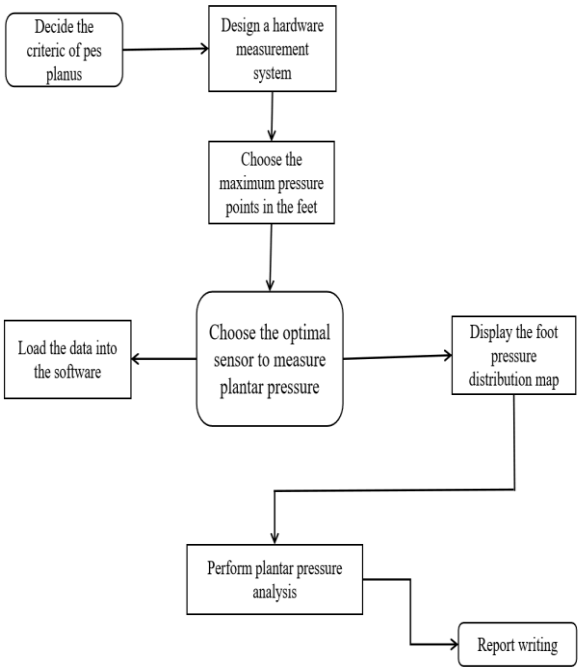


Fig 1. Methodology for analysis pressure on foot

Display the pressure distribution map to visualize high-pressure areas in the feet. Identify regions of concern that require correction. Perform plantar pressure analysis to determine the extent of flat feet and areas that require support or adjustment in the shoe design. Consider factors such as arch support, cushioning, and custom insole design. We, create a comprehensive report that summarizes the findings and recommendations for addressing pes planus through the custom shoe design. This report should include detailed

ISBN 979-8-3503-5094-4

information on pressure points, shoe modifications, and any additional medical advice for individuals with flat feet. To design a specialized shoe that effectively rectifies pes planus by addressing specific pressure points and providing tailored support for individuals with this condition is shown in Fig.1.

IV. IMPLEMENTATION

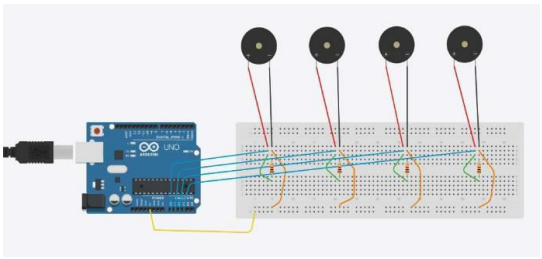


Fig 2. Software implementation

In the software section, the implementation involved crafting code to read and interpret the signals from the piezoelectric sensor data. The Arduino software not only processed the pressure information but also generated a graphical representation, providing a visual insight into the variations in pressure over time as shown in Fig.2.

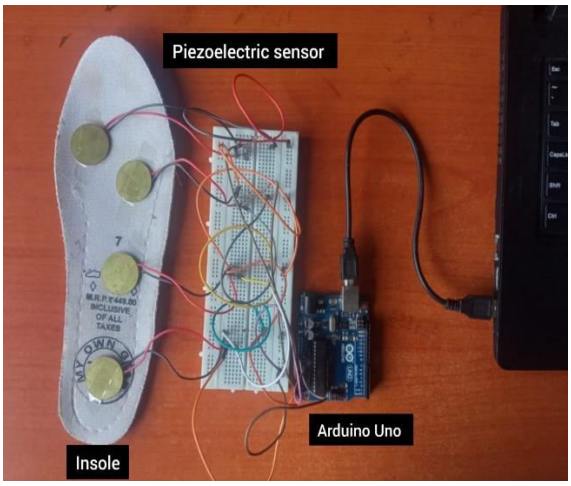


Fig 3. Sensors are attached to the insole

In this work, the goal is to analyze foot pressure utilizing a fixed insole equipped with piezoelectric sensors interfaced with an Arduino Uno microcontroller. The sensors, strategically positioned at A0, A1, A2, and A4, capture pressure data as users walk or apply force. The Arduino Uno processes this information, and the output is translated into a graphical representation using the Arduino software. This hardware implementation enables real-time monitoring and visualization of foot pressure distribution, offering valuable insights for applications ranging from ergonomic design to healthcare diagnostics. The integration of piezoelectric sensors and Arduino Uno enhances the work's versatility, making it a robust solution for pressure analysis within the context of shoe insoles. In this work, the goal is to analyze foot pressure utilizing a fixed insole equipped with piezoelectric sensors interfaced with an Arduino Uno microcontroller as shown in Fig.3 and 4.

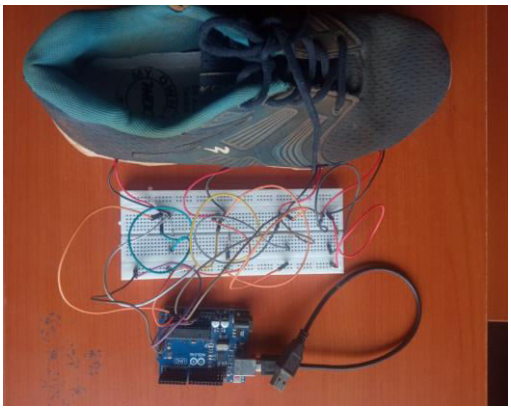


Fig 4. Hardware implementation

The sensors, strategically positioned at A0, A1, A2, and A4, capture pressure data as users walk or apply force. The Arduino Uno processes this information, and the output is translated into a graphical representation using the Arduino software. This hardware implementation enables real-time monitoring and visualization of foot pressure distribution, offering valuable insights for applications ranging from ergonomic design to healthcare diagnostics. The integration of piezoelectric sensors and Arduino Uno enhances the work's versatility, making it a robust solution for pressure analysis within the context of shoe insoles.

Fig 5. Graphical representation software output

Arduino Uno software analyzes and visually represents individuals pressure (value1,2,3,4) from piezoelectric sensor, providing a graphical summary of the applied pressure is shown in Fig.5.

Table 1. Foot sensor values aiding in flat foot analysis

Foot region	Value (X)	Value (Y)	Color
Arch (sensor 1)	203	135	Green
Metatarsals (sensor 2)	200	158	yellow
Toes (sensor 3)	193	130	Orange
Heel (sensor 4)	185	150	Blue

The foot sensor data (Table 1), represented by different regions of the foot with associated colors, provides essential information for analyzing and treating flat feet. The Arch sensor (Green) contributes (203, 135) values, offering insights into pressure in the arch region, while the Metatarsals sensor (Yellow) provides (200, 158) values for assessing toe pressure. The Toes sensor (Orange) delivers (193, 130) values to analyze pressure on the metatarsals, and the Heel sensor (Blue) offers (185, 150) values for evaluating heel

ISBN 979-8-3503-5094-4

pressure. This comprehensive data enables a detailed understanding of pressure distribution, aiding healthcare professionals in diagnosing and customizing treatment plans for individuals with flat feet. The color-coded system enhances visual interpretation, streamlining the analysis of pressure irregularities across different foot regions.

V. RESULT AND DISUCSSION

The utilization of piezoelectric sensors for pressure mapping analysis has provided deep insights into the intricate pressure dynamics associated with pes planus. By strategically placing sensors of varying colors, distinct pressure distribution patterns across the foot during different activities have been discerned. Notably, areas of heightened pressure, particularly at the medial arch and metatarsal head, indicate potential discomfort and injury risk for individuals with pes planus. The integration of these findings into specialized shoe design offers an opportunity to deliver targeted support and alleviate discomfort. Our forthcoming product, the specialized shoe, is already in progress, drawing from the insights gained through pressure mapping analysis with insole technology. Through customization of cushioning and support structures based on pressure mapping data, our aim is to elevate overall comfort and optimize foot function, representing a notable advancement in personalized footwear solutions for pes planus sufferers. Building upon the outcomes of our analysis, the subsequent phase of our work entails developing a specialized shoe tailored to address the identified pressure distribution pattern. This endeavor involves integrating the insights gleaned from the pressure mapping module into the shoe's design to offer targeted support and mitigate potential discomfort associated with pes planus

VI. CONCLUSION

In conclusion, the method outlined here provides a comprehensive approach to address pes planus (flat feet) and design a specialized shoe for individuals suffering from this condition and that is future work of our work. By establishing precise criteria for pes planus identification, designing a hardware measurement system with piezoelectric sensor technology, and analyzing plantar pressure data, we can create custom footwear that effectively rectifies the problem. The combination of accurate pressure measurement, pressure distribution mapping, and thorough analysis enables the development of personalized shoe solutions, tailored to the unique needs of individuals with flat feet. This approach not only aims to alleviate discomfort and improve foot function but also enhances overall well-being and quality of life for those affected by pes planus.

References

[1] Ozlem ¨ Feyzioglu a, Ozgül ¨ Oztürk ¨ a, Selim Mugrabi ˘ b "Is Morton's neuroma in a pes planus or pes cavus foot lead to differences in pressure distribution and gait parameters". Heliyon 9 (2023) e19111, https://doi.org/10.1016/j.heliyon.2023.e19111

[2] Bloks BE, Wilders LM, Louwerens JWK, Geurts AC, Nonnekes J, Keijsers NLW. Quantitative assessment of plantar pressure patterns in relation to foot deformities in people with hereditary motor and sensory neuropathies. J Neuroeng Rehabil. 2023 May 16;20(1):65. doi: 10.1186/s12984-023-01172-1.

[3] Efrima B, Barbero A, Ovadia JE, Indino C, Maccario C, Usuelli FG. Classification of the Os Calcis Subtalar Morphology in Symptomatic Flexible Pediatric Pes Planus Deformity Using Weightbearing CT and Distance Mapping. Foot Ankle Int. 2023 Apr;44(4):322-329. doi: 10.1177/10711007231156605.

[4] Yalçin N, Esen E, Kanatli U, Yetkin H. Evaluation of the medial longitudinal arch: a comparison between the dynamic plantar pressure measurement system and radiographic analysis. Acta Orthop Traumatol Turc. 2010;44(3):241-5. doi: 10.3944/AOTT.2010.2233.

[5] Zhai JN, Wang J, Qiu YS. Plantar pressure differences among adults with mild flexible flatfoot, severe flexible flatfoot and normal foot when walking on level surface, walking upstairs and downstairs. J Phys Ther Sci. 2017 Apr;29(4):641-646. doi: 10.1589/jpts.29.641.

[6] Andrew K. Buldt a b, Saeed Forghany c d, Karl B. Landorf a b, Pazit Levinger e, George S. Murley b, Hylton B. Menz a bFoot posture is associated with plantar pressure during gait: A comparison of normal, planus and cavus feet. https://doi.org/10.1016/j.gaitpost.2018.03.005

[7] Caiting Zhang, Yining Xu, Jiao Li, Fekete Gusztáv and Yaodong GuMixed Comparison of Intervention with Assistive Devices for Plantar Pressure Distribution and Anatomical Characteristics in Adults with Pes Cavus: Systemic Review with Network Meta-Analysis. Vol. 13, no. 17, 2023. https://doi.org/10.3390/app13179699

[8] Kodithuwakku Arachchige SNK, Chander H, Knight A. Flatfeet: Biomechanical implications, assessment and management. Foot (Edinb). 2019 Mar;38:81-85. doi: 10.1016/j.foot.2019.02.004..

[9] Sankaran, sakthivel; murugan, pallikonda rajasekaran design and development of a device for reduction of data loss and time taken for transferring the data into an artificial biocomposite socket prosthesis through arduino Journal of Mechanics in Medicine & Biology, 2023, Vol 23, Issue 10, p1 DOI 10.1142/S021951942240019X

[10] Sankaran, S. et al. (2023). Monitoring of Physiological and Atmospheric Parameters of People Working in Mining Sites Using a Smart Shirt: A Review of Latest Technologies and Limitations. In: Chatterjee, P., Pamucar, D., Yazdani, M., Panchal, D. (eds) Computational Intelligence for Engineering and Management Applications. Lecture Notes in Electrical Engineering, vol 984. Springer, Singapore. https://doi.org/10.1007/978-981-19-8493-8_53

[11] S. Sankaran, J. Deny, P. V, S. K and M. V, "Design and Development of muscle stimulation for cervicalgia patient using wearable coat," *2021 6th International Conference on Communication and Electronics Systems (ICCES)*, Coimbatore, India, 2021, pp. 1279-1286, doi: 10.1109/ICCES51350.2021.9489228.

[12] S. Sankaran, T. Keerthana, S. M, T. A. Prasath and P. I. Britto, "A Comprehensive Study on Design of Frequency Probe for Treatment of Plantar Corns and Warts using Microwave Technology," 2023 2nd International Conference on Automation, Computing and Renewable Systems (ICACRS), Pudukkottai, India, 2023, pp. 76-81, doi: 10.1109/ICACRS58579.2023.10404576.

ISBN 979-8-3503-5094-4

DEVELOPMENT OF IOT BASED ASSISTIVE SYSTEM FOR STROKE SURVIVORS

Dr. Banu Rekha B
Department of Biomedical Engineering
PSG College of Technology
Coimbatore, India
bbr.bme@psgtech.ac.in

Vibin VS
Department of Biomedical Engineering
PSG College of Technology
Coimbatore, India
20d249@psgtech.ac.in

Archidha R
Department of Biomedical Engineering
PSG College of Technology
Coimbatore, India
20d205@psgtech.ac.in

Sree Subiksha M K
Department of Biomedical Engineering
PSG College of Technology
Coimbatore, India
20d241@psgtech.ac.in

Kathiravan S
Department of Biomedical Engineering
PSG College of Technology
Coimbatore, India
21d435@psgtech.ac.in

Abstract— Hand impairment due to paralysis or aging often reduces grasping strength to just 50% of peak force, severely impacting daily tasks. To address this, an advanced glove has been developed, aiding hand-impaired individuals in flexion, extension, and rehabilitation. Pneumatic actuators, composed of latex rubber tubes within a braided sleeve, respond to air pressure, enabling finger bending (flexion) when pressurized and extension when air is released via solenoid valves. This glove allows users finger control, accessible through voice commands, inbuilt switches, or a mobile app, with customizable pressure and duration settings based on individual impairment conditions. IoT integration facilitates remote patient monitoring, employing Google Assistant for voice control and IFTTT for data exchange. Beneficial for stroke survivors, the elderly, and those with neurological disorders, the glove offers versatility in rehabilitation and grip tasks. It features voice control, cost-effectiveness, thumb and wrist support, and wearability. Challenges include managing pneumatic vibrations and ensuring lightweight comfort.

Keywords—pneumatic actuator, voice control, IoT, grasping and physiotherapy, IFTTT

I. INTRODUCTION

Activities of daily living (ADL) involve grasping. Impaired hand function resulting from paralysis and the natural aging process can lead to a significant reduction in grasping strength, often diminishing to just 50% of peak force. Grasping is crucial for daily tasks as it allows us to interact with objects and perform activities efficiently. It facilitates independence and aids in tasks like eating, dressing, and writing. Grasping strength exhibits notable disparities between individuals with typical hand function and those facing hand impairments. Typically, individuals without hand impairments enjoy superior dexterity and grip strength, granting them effortless object manipulation. In contrast, individuals dealing with hand impairments often grapple with diminished grip strength, compromised fine motor control, and restricted hand mobility. These challenges significantly impede their autonomy in carrying out everyday tasks independently. It's important to understand these differences because it helps us create special tools and ways to help people with hand problems. These tools and methods are designed to make it easier for them to do things on their own and have a better life.

Current treatment methods for individuals with hand paralysis or hand impairment are typically administered by trained healthcare professionals, such as physical therapists and occupational therapists.

- **Physical Therapy:**
  - **Range of Motion Exercises:** Therapists work with patients to improve the range of motion in the affected hand and fingers. This involves gently moving the joints through various motions to prevent stiffness and maintain flexibility.
  - **Strengthening Exercises:** Specific exercises are designed to target the muscles in the hand and forearm, gradually increasing strength and grip. Therapists may use resistance bands, putty, or small weights..
- **Hand Splints:**
  - Hand splints are specialized orthotic devices that are custom-made to fit the individual's hand. They provide support and maintain the hand and wrist in a functional position.
  - These splints are designed to prevent joint stiffness, muscle tightness, and contractures that can occur due to lack of use or immobility.
- **Manual Massage:**
  - Massage therapy can help alleviate pain, reduce muscle tension, and improve blood circulation in the hand and forearm.
  - Therapists use various massage techniques, such as effleurage and petrissage, to target specific muscles and promote relaxation.
- **Mirror Therapy:**
  - Mirror therapy is a technique where a mirror is placed between the individual's hands. When the unaffected hand is moved, it creates the illusion that the paralyzed hand is moving.
  - This method helps rewire the brain and improve motor control and function in the paralyzed hand.
- **Sensory Re-education:**
  - Individuals with hand impairment may experience sensory deficits. Sensory re-education involves stimulating the hand's sensory receptors through techniques like tapping, pressure, and texture exploration to improve sensation and tactile awareness.

ISBN 979-8-3503-5094-4

In hand rehabilitation gloves, traditional string-based systems have been commonly used to provide assistance to impaired fingers. However, pneumatic actuators offer distinct advantages due to their lightweight nature. Unlike strings, which can add bulk and restrict movement, pneumatic actuators are compact and lightweight, allowing for greater freedom of motion. Additionally, they offer precise control over force and movement, enabling tailored rehabilitation programs to accommodate varying levels of impairment.

II. Existing Methodologies

Various soft robotic technologies have been developed for hand rehabilitation and assistance:

- Al-Fahaam et al. [1] introduced a wearable power-assisted glove using pneumatic soft actuators, focusing on multi-griping and multi-pinching movements to aid users with impaired hand function.
- Chopade et al. [2] presented a power-assist glove with predictive grasping modes, utilizing pneumatic soft actuators to support various grasping tasks.
- Toya et al. [3] proposed a control algorithm for a Power-Assist Glove, emphasizing assisting grasping force and facilitating at-home rehabilitation.
- Polygerinos et al. [4][5] developed a portable, assistive soft robotic glove for hand rehabilitation, providing sufficient forces to aid passive finger closure.
- Yamamoto et al. [6] researched stand-alone wearable power-assisting suits with sensing and control systems.
- Naruse et al. [7] designed a wearable exoskeleton power assist system for lower back support, applicable to targeted assistance and rehabilitation.
- Roger et al. [8] highlighted the role of wearable robotics, including soft actuators and exoskeletons, in cardiac and stroke rehabilitation.
- Ueki et al. [9] developed a hand-assist robot with multi-degrees-of-freedom for rehabilitation therapy, demonstrating its potential in aiding hand rehabilitation during therapy sessions.

III. Proposed Methodology

A. Pneumatic based assistive grasp system

With current treatments for hand paralysis or impairment relying heavily on manual techniques, a solution has been proposed to aid individuals in both grasping and rehabilitation. This solution revolves around a pneumatic-based assistive grasp system, where pneumatic actuators are integrated into a glove. These pneumatic actuators consist of a latex tube enclosed within a braided sleeve. The purpose of this braided sleeve is to regulate the inflation of the latex rubber tube. The entire system is powered by an air pump that pressurizes the pneumatic actuator [1].

When the actuator, affixed to the glove, receives air pressure, which causes the fingers to bend effectively. Conversely, when the air within the tube is released, the fingers return to their natural position. This approach holds significant promise for assisting individuals with hand impairments by enhancing their grasp and supporting rehabilitation efforts. By incorporating pneumatic technology into wearable devices, significant progress can be made in enhancing the lives of individuals facing hand-related challenges.

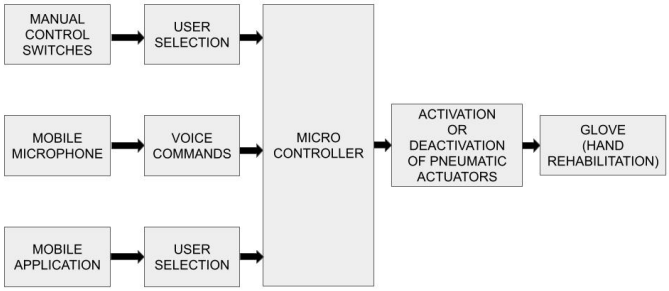


Fig 1. Methodology of Physiotherapy Assist System for Hand Finger Rehabilitation

B. Functions of the pneumatic actuator

Pneumatic actuators within the system have two primary functions: flexion and extension of the fingers.

- Flexion is achieved by pressurizing the pneumatic actuator to a predetermined level using an air pump. This inflation causes the fingers to bend, allowing individuals to grip and hold objects effectively.
- Extension, the opposite motion, is facilitated by releasing air from the tube. This action is controlled by a solenoid valve, resulting in the straightening of the fingers and allowing individuals to release held objects [2].

C. Assistance in Grasping Tasks

- The flexion function of the pneumatic actuators enables individuals to grasp and hold objects securely, aiding them in various daily tasks that involve handling items.
- Extension, on the other hand, allows individuals to release the objects they were holding, offering a comprehensive solution for tasks that require both gripping and letting go.

D. Rehabilitation Benefits:

- Continuous and controlled flexion and extension of the fingers provided by the pneumatic-based system offer valuable rehabilitation exercises.
- These repetitive movements can be tailored to an individual's needs, making them particularly beneficial for individuals seeking to regain or enhance hand function.
- The system not only assists in daily activities but also promotes hand function recovery, serving as a versatile tool to improve both independence and rehabilitation outcomes.

ISBN 979-8-3503-5094-4

Fig 2. Pneumatic Actuator

IV. MATERIALS

A. Air Pump:

The air pump used for the pneumatic system is a compact and efficient device designed to provide pressurized airflow. Operating at a voltage of 12V DC, it delivers a pressure output of 15 psi, making it suitable for a wide range of applications. Its compact design allows for easy integration into various systems, while its efficient operation ensures reliable performance. With versatile usage capabilities, the pump can be employed in soft robotics, air mattresses, pneumatic tools, and more. This 12V DC air pump is a dependable component ideal for applications requiring moderate air pressure.

B. Solenoid Valve:

The solenoid valve is a crucial component in pneumatic systems, serving to control the flow of air by electrically opening or closing the valve. It consists of a coil of wire wound around a movable plunger. When an electric current passes through the coil, it creates a magnetic field that pulls the plunger, allowing air to flow through the valve. Conversely, when the current is turned off, the magnetic field dissipates, and a spring returns the plunger to its original position, shutting off the airflow.

C. ESP32 XiaoC3 Microcontroller Module:

The ESP32 XiaoC3 is a compact microcontroller board based on the ESP32 chip, offering Wi-Fi and Bluetooth connectivity. Despite its small size, it provides GPIO pins, UART, SPI, I2C, ADC, DAC, PWM, and more for versatile connectivity. Programmable with the Arduino IDE, it's ideal for IoT projects, offering low power consumption and open-source firmware for flexibility.

D. OLED Display:

The OLED (Organic Light-Emitting Diode) display is a type of screen technology that offers high contrast, wide viewing angles, and low power consumption. It consists of organic compounds that emit light when an electric current passes through them. The display enables the users to select different modes and controls.

V. CUSTOMIZATION AND CONTROL

The pneumatic glove designed for individuals with impaired hand function offers options for control, allowing users to customize its performance to fit their specific needs. Users are presented with three primary control methods: voice commands, manual mode, and a dedicated mobile application. These methods empower users to tailor the glove's performance, adjusting pressure levels, action duration, and other parameters to meet their specific disabilities and rehabilitation requirements. These control methods help make the glove a versatile and user-friendly assistive system.

A. Voice Commands:

One of the standout features of this pneumatic glove is its voice command functionality. Users can interact with and manipulate the glove through voice commands, making it highly accessible for individuals with limited manual dexterity.

Voice commands empower users to initiate various actions and movements, such as grasping or releasing objects, with just their voice. For example, a simple vocal command like "flexion" can trigger the glove to close and securely hold an object, while saying "extension" can prompt the glove to open its fingers. This intuitive control method allows users to perform tasks and activities with ease, providing a sense of independence and empowerment.

B. Manual Mode:

The pneumatic glove also offers a manual control mode through in-built switches on the assistive system. This mode empowers users to take direct control of the glove's movements, specifically initiating flexion and extension motions as needed. The manual mode provides a high degree of flexibility, allowing users to tailor the glove's actions to meet specific requirements.

For example, if a user wishes to adjust the grip strength for a particular task, they can do so through manual control. By toggling the switches, users can precisely control the glove's movements, ensuring that it behaves in a manner that suits their immediate needs.

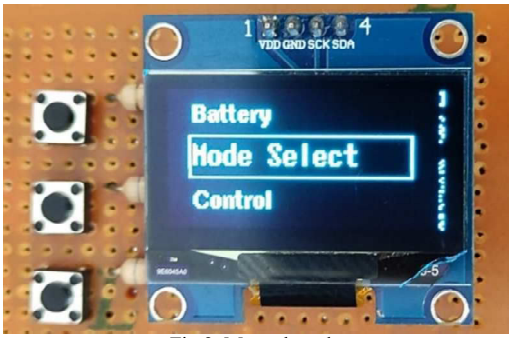


Fig 3. Manual mode

C. Mobile Application:

Complementing the voice command and manual control options, a dedicated mobile application has been developed to enhance the glove's performance. This mobile application serves as a user-friendly interface that allows users to fine-tune the glove's behavior to their preferences and unique needs.

ISBN 979-8-3503-5094-4

All the control methods offers a range of customization options, including the ability to adjust pressure levels, the duration of each action, and the glove's responsiveness to specific commands. Users can configure the glove to align with their individual disabilities and rehabilitation goals.

For instance, users can set the pressure levels to accommodate their comfort and strength, ensuring that the glove's grip is neither too weak nor too forceful. They can also modify the duration of specific actions, optimizing the glove's performance for various tasks. This level of customization is especially valuable for individuals with varying degrees of hand impairment.

VI. IoT Integration for Remote Monitoring

The integration of Internet of Things (IoT) technology into the pneumatic glove for hand rehabilitation enhances the device's functionality. It allows for remote monitoring, which is a valuable feature for both users and healthcare professionals. This integration has the potential to significantly improve the rehabilitation process by offering real-time data and the capability to make necessary adjustments to rehabilitation programs.

A. IoT Integration:

The pneumatic glove is equipped with IoT technology, which enables it to connect to the internet and transmit data wirelessly. This connectivity can be achieved through Wi-Fi or cellular networks, making it accessible in various settings.

B. Remote Monitoring:

The IoT integration allows healthcare professionals to remotely monitor the progress of patients who are using the pneumatic glove for hand rehabilitation. Here's how remote monitoring works:

- Data Collection: The glove continuously collects data related to its usage and the user's hand function. This data can include information about the duration and frequency of glove usage, grip strength, and the types of activities the user is engaging in during rehabilitation.
- Data Transmission: The collected data is securely transmitted to a centralized platform or a cloud-based server using the IoT connectivity. This platform serves as a hub for storing and processing the data.
- Real-time Access: Healthcare professionals, such as physical therapists and physicians, can access this platform through a web-based interface or a dedicated application. This access allows them to view the data in real time.
- Analysis and Assessment: The collected data provides insights into the user's progress and the effectiveness of the rehabilitation program. Healthcare professionals can analyze this data to assess whether the user is making the expected improvements.

C. Tailoring Rehabilitation Programs:

One of the key advantages of IoT-based remote monitoring is the ability to tailor rehabilitation programs to individual needs. Here's how this customization is achieved:

- Adjustment of Parameters: The pneumatic glove may have adjustable parameters, such as pressure levels or the range of motion. Based on the data and user feedback, healthcare professionals can remotely adjust these parameters to optimize the rehabilitation process.
- Goal Setting: Using the collected data, healthcare professionals can set specific goals and milestones for the user. These goals can be tailored to the individual's progress, ensuring that they are neither too easy nor too challenging.
- Progress Tracking: Users can also track their own progress and set personal goals through the mobile application. They can see how they are improving over time, which can be motivating and provide a sense of achievement.

D. Convenience and Efficiency:

IoT-based remote monitoring offers several advantages:

- Convenience: Users can engage in their rehabilitation activities from the comfort of their homes while still receiving professional guidance and support.
- Efficiency: Healthcare professionals can efficiently manage and monitor multiple patients, reducing the need for frequent in-person visits.
- Personalization: The rehabilitation process becomes highly personalized, as adjustments and recommendations are based on the user's real-time data.

E. Privacy and Security:

Given that the glove collects sensitive health data, ensuring data privacy and security is paramount. IoT platforms are designed to implement robust security measures to protect patient information and comply with healthcare data regulations.

VII. Voice Control and Feedback:

The pneumatic glove designed for hand rehabilitation offers voice control capabilities, which are made possible through the integration of Google Assistant. This feature allows users to interact with the glove using voice commands and receive feedback, creating a user-friendly and accessible interface. Voice control is particularly advantageous for individuals with limited hand mobility or those who prefer hands-free operation.

A. Voice Control Capabilities:

- Integration of Google Assistant:

Google Assistant is a virtual assistant developed by Google that provides voice-activated and natural language processing capabilities. It has been integrated into the pneumatic glove to enable voice control. This integration allows the glove to understand and respond to voice commands issued by the user.

- Effortless Interaction:

Users can effortlessly interact with the pneumatic glove by issuing voice commands. These commands can include a variety of instructions related to the glove's functions, such as grasping an object, releasing it, adjusting grip strength, or transitioning between different modes of operation.

- Accessibility for Limited Hand Mobility:

Voice control is a game-changer for individuals with limited hand mobility or those who may find manual button controls challenging. It opens up new possibilities

ISBN 979-8-3503-5094-4

for individuals who may not have the dexterity or strength to use traditional control methods effectively.

B. Advantages of Voice Control and Feedback:

- Accessibility: Voice control makes the glove accessible to a wider range of users, including those with limited hand function, physical disabilities, or conditions that affect motor control.
- Intuitiveness: Voice commands are natural and intuitive, requiring minimal training. Users can communicate with the glove in a way that feels familiar and comfortable.
- Hands-Free Operation: Voice control enables hands-free operation, which is especially beneficial in situations where users need to multitask or operate the glove without manual input.
- Reduced Cognitive Load: Voice control simplifies the user experience by reducing the need for complex button sequences or manual adjustments. Users can focus on their tasks rather than navigating complicated controls.
- Safety and Confidence: Real-time feedback ensures that users are aware of the glove's actions, contributing to safety and user confidence.

VIII. Data Transfer and Automation with IFTTT:

The pneumatic glove designed for hand rehabilitation incorporates two crucial features: seamless data transfer and automation through If This Then That (IFTTT). These elements work together to enhance the functionality and integration of the device into users' daily lives.

A. Data Transfer (Wi-Fi Connectivity):

One key aspect of the pneumatic glove is its ability to exchange data between the glove's microcontroller and the user's smartphone via Wi-Fi. This real-time communication ensures that the glove promptly responds to user commands and facilitates efficient monitoring of progress and adjustments to the rehabilitation regimen.

- Real-Time Communication: Wi-Fi connectivity guarantees real-time communication, allowing immediate response to user commands. This real-time aspect is essential for a natural and responsive user experience.
- Progress Monitoring: The seamless data transfer system efficiently tracks the user's rehabilitation progress. Data related to glove usage, hand function, and exercises is instantly transmitted to the user's smartphone. This data provides insights into the user's rehabilitation journey, enabling them to track their improvements over time.
- Rehabilitation Regimen Adjustments: Access to real-time data empowers both users and healthcare professionals to make informed decisions about adjusting the rehabilitation program. This timely feedback allows for the optimization of rehabilitation, ensuring it is effective and tailored to the user's specific needs.

B. Automation with IFTTT:

In addition to data transfer, the pneumatic glove uses the IFTTT automation platform to improve its functionality and integrate better into daily life.

- Custom Automation Routines: IFTTT allows users to create custom automation routines using a simple "If This Then That" logic. These routines enable the glove to be embedded into the user's daily activities.
- Integration of Online Services and Devices: IFTTT connects various online services, applications, and smart devices, enabling the glove to interact with these services. For example, users can set up applets that remind them to perform hand exercises at specific times during the day or even trigger glove actions based on events.
- Condition-Based Triggers and Event-Driven Actions: Users can define conditions and events that trigger specific glove actions, enhancing the glove's adaptability to different situations and tasks throughout the day.
- Enhancing Accessibility and Personalization: IFTTT's automation capabilities make the glove more accessible to a broader range of users, offering personalized user experiences by allowing tailored automation routines to suit individual requirements.

The combination of seamless data transfer and IFTTT automation empowers users to not only rehabilitate effectively but also integrate the glove into their daily lives seamlessly. Real-time data exchange and customization through IFTTT routines ensure that the glove is responsive, adaptable, and personalized. It transforms the pneumatic glove into a dynamic tool that enhances its utility and accessibility, ultimately improving the user's rehabilitation experience and overall quality of life.

IX. Market Analysis:

A. Market Need [10]:

- India's physiotherapy market size is estimated to expand at a compound annual growth rate (CAGR) of 8.85% from 2022 to 2030 and will reach $1.9 Bn in 2030.
- There is a growing need for effective hand rehabilitation solutions, driven by factors such as aging populations, increased instances of stroke, and the prevalence of hand impairments due to neurological disorders.
- The demand for user-friendly, technologically advanced assistive devices is on the rise, as patients and healthcare professionals seek more personalized and efficient rehabilitation options.

B. Target Audience:

- Stroke survivors looking to regain hand function.
- Elderly individuals with reduced hand mobility and grip strength.
- People with neurological disorders seeking rehabilitation support.
- Physiotherapists and occupational therapists involved in hand rehabilitation.

C. Competitive Landscape:

In the competitive landscape of hand rehabilitation devices, traditional solutions like therapy putty and hand grippers have been prevalent. However, the Syrebo Rehabilitation Glove

ISBN 979-8-3503-5094-4

introduces a modern alternative. It integrates pneumatic soft actuators for targeted assistance and offers adjustable support levels. With its lightweight, ergonomic design, the Syrebo Glove ensures comfort, ease of use, and personalized rehabilitation, setting a new standard in hand rehabilitation technology [11].

D. Market Potential:

- The global assistive technology market is expected to grow steadily due to the aging population and increasing disability rates.
- The hand rehabilitation sector within this market holds substantial potential for growth.
- The proposed glove's unique features and affordable design position it well to capture market share.

E. Challenges:

- Competition from established and emerging players.
- Ensuring user safety and comfort, especially in long-term usage.
- Navigating regulatory approvals for medical devices.

X. CHALLENGES

While the concept of pneumatic glove for rehabilitation and assistive purposes holds great promise, several significant challenges must be addressed during its design and implementation are:

A. Vibration Reduction:

One primary challenge is the generation of vibrations during pneumatic operation, which can potentially cause discomfort or even pain to the user. Mitigating this challenge requires the development of effective vibration-dampening mechanisms to enhance user comfort. This may involve the incorporation of shock-absorbing materials, improved pneumatic system design, or real-time feedback systems to adjust pressure levels and minimize vibrations.

B. Lightweight Construction:

To ensure that users can wear the glove comfortably for extended periods, it is crucial to keep the overall weight of the system to a minimum. Achieving this requires the use of lightweight materials, efficient design, and potentially innovative engineering approaches. Lightweight construction is essential to prevent user fatigue and promote extended use, particularly for individuals with impaired hand function who may rely on the device throughout the day.

XI. CONCLUSION:

In conclusion, the pneumatic-based assistive glove represents a significant advancement in the field of rehabilitation and assistive technology. Its ability to mimic natural hand movements, provide customizable control options, and facilitate remote monitoring through IoT connectivity makes it a versatile tool for stroke survivors, the elderly, and individuals with neurological disorders. Moreover, the integration of voice control and automation through IFTTT ensures that the glove integrates into daily life, enhancing its usability and impact. This technology not only offers hope for improved hand function but also empowers individuals to regain their independence and enhance their quality of life.

However, it is important to recognize and address challenges related to vibration and weight in the wearable design to optimize its usability and effectiveness fully. The ongoing development of such assistive technologies holds great promise for improving the lives of individuals with impaired hand function and advancing the field of rehabilitation and healthcare.

XII. FUTURE WORKS

In future works of the system, a correlation between the pressure given to actuators and the resulting flexion angle can be explored. This correlation would involve conducting validation experiments to establish a relationship between the applied pressure and the resulting movement of the actuators, specifically focusing on the flexion angle achieved in response to varying pressure levels.

The validation process could include experimental trials where different pressure levels are applied to the actuators, and the resulting flexion angles are measured and recorded. These experiments would be conducted across a range of pressure settings to capture the full spectrum of possible movements and to establish a reliable correlation.

REFERENCES

[1] Hassanin Al-Fahaam,Dr. Steve Davis,Prof. Samia Nefti-Meziani— Power Assistive and Rehabilitation Wearable Robot based on Pneumatic Soft Actuators, International Conference on MMAR, Miedzyzdroje, Poland, 2016.

[2] Shripad Shashikant Chopade, Sagar Pradip Kauthalkar, Chaitanya Bhalchandra Bhandari, — (MCKIBBEN'S MUSCLE) Robots Make Our Work Lighter, But We Have Made the Robots Lighter, IJMER Vol. 3, Issue. 5, Sep - Oct. 2013.

[3] Kiminori TOYA, Toyomi MIYAGAWA and Yuji KUBOTA, — Power-assist glove operated by predicting the grasping mode, Journal of System Design and Dynamics, Vol. 5, No. 1, 2011.

[4] P. Polygerinos, Z. Wang, K. C. Galloway, R. J. Wood, an d C. J. Walsh, — Soft robotic glove for combined assistance and at-home rehabilitation, Robotics and Autonomous Systems, Volume 73, 2015 .

[5] P. Polygerinos, S. Lyne, Z. Wang, L. F. Nicolini, B. Mosadegh, G. M. Whitesides, et al., — Towards a Soft Pneumatic Glove for Hand Rehabilitation, 2013 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS).

[6] K. Yamamoto, M. Ishii, H. Noborisaka, and K. Hyodo, "Stand alone wearable power assisting suit-sensing and control systems," in Robot and Human Interactive Communication, 2004. ROMAN 2004. 13th IEEE International Workshop on, 2004, pp. 661-666.

[7] K. Naruse, S. Kawai, H. Yokoi, and Y. Kakazu, "Development of wearable exoskeleton power assist system for lower back support," in Intelligent Robots and Systems, 2003.(IROS 2003). Proceedings. 2003 IEEE/RSJ International Conference on, 2003, pp. 3630-3635

[8] V.L. Roger, A.S. Go, D.M. Lloyd-Jones, E.J. Benjamin, J.D. Berry, W.B. Borden, et al., Heart disease and stroke statistics—2012 update, a report from the American Heart Association, Circulation 125 (2012) e2–e220.

[9] S. Ueki, H. Kawasaki, S. Ito, Y. Nishimoto, M. Abe, T. Aoki, et al., Development of a hand-assist robot with multi-degrees-of-freedom for rehabilitation therapy, IEEE/ASME Trans. Mechatronics 17 (2012) 136–146.

[10] https://www.marketsandmarkets.com/Market-Reports/assistive-robotics-market-37247851.html

[11] https://www.syrebo.com/rehabilitation-equipment-for-home/hand-rehabilitation-robot-sy-hre12.htm

ISBN 979-8-3503-5094-4

MODIFICATION OF ARTIFICIAL WOMB IN COMBINATION WITH MODERN NEONATAL INCUBATOR PRINCIPLES

Dr.M.Neela Harish
Assistant Professor
Department of Biomedical Engineering
Easwari Engineering College
Chennai, India
neelaharish.m@eec.srmrmp.edu.in

S. Prasanthi
Undergraduate Student
Department of Biomedical Engineering
Easwari Engineering College
Chennai, India
sprasanthi1723@gmail.com

Anushiya. K
Undergraduate Student
Department of Biomedical Engineering
Easwari Engineering College
Chennai, India
madhuanu2004@gmail.com

***Abstract*—According to the World Health Organisation(WHO), preterm birth is considered as the main cause for increase in infant mortality rate all around the world. When an infant is born before the completion of 37 weeks of pregnancy, it is termed as a preterm birth. As per reports, in 2022, preterm birth affected about 1 of every 10 infants born in the United States. To provide a suitable environment for the growth and development of such premature infants, even those born at 22 weeks of gestation, an artificial womb along with an artificial placenta and umbilical cord is being developed. At present, such premature infants are supported and monitored using neonatal incubators. In this paper, an artificial womb has been designed incorporating the principles used in modern-day neonatal incubators using biocompatible material**

I. INTRODUCTION

A typical pregnancy usually lasts for about 40 weeks. Preterm babies are those born before the 37th week of gestation. Premature babies often have serious illnesses and they lack nutrients and sufficient oxygen supply. Initially, these babies were treated using an incubator. Preterm births are classified into four categories depending on the gestational period: late, moderately preterm, very preterm and extremely preterm.

The uterus is a female reproductive organ that houses and nourishes the foetus. The placenta is an organ that is formed temporarily inside the uterus during pregnancy. The functions of the placenta include oxygen and nutrient supply and removal of nitrogen through the umbilical cord.

The umbilical cord is a connection between the foetus and the placenta. It is formed during the 3rd week of pregnancy. The umbilical cord has two umbilical arteries and one umbilical vein that is surrounded by an extracellular matrix called Wharton's jelly. The arteries are intended to carry waste from the foetus to the placenta to be eliminated whereas the vein carries oxygen and nutrients from the placenta to the foetus.

The methodology provided, mimics the function of the human uterus, placenta and umbilical cord. In addition to this , a modified neonatal incubator has been used with smart sensors to monitor the baby's response continuously. The neonatal incubator helps to monitor, control and maintain factors like temperature, humidity and heart rate. This combined system can primarily reduce the infant mortality rate.

ISBN 979-8-3503-5094-4

II. LITERATURE SURVEY

Greenberg was the first person to file a patent for his illustration of an artificial uterus. In his illustration, the foetus was placed inside a tank filled with amniotic fluid. The other components used in his research include a machine connected to the umbilical cord, blood pumps, an artificial kidney and a water heater.

The next patent was filed by Cooper William which included a life support system for a premature baby. The system was attached to the baby's placenta through an umbilical cord. It consists of two chambers. The upper chamber contains a placenta connected to an oxygen-containing atmosphere and a nutrient-supplying system where the placenta remains on top of the dome-structured partition. The lower chamber consists of a liquid to keep the baby immersed in the solution.

Another study was made by Kuwabara Y along with his staff who invented a new extra-uterine incubation system using 14 goat foetuses. They developed a technique called ExtraUterine Foetal Incubation (EUFI). In the system developed, the foetus was suspended in an artificial amniotic fluid, connected to an Extracorporeal Membrane Oxygenator.

A study made by a group of students in Mysore in the year 2023 used light bulbs to heat up the womb in which concerns about the safety of the bulbs were not achieved. Their proposed system specifically targeted low risk babies which again was a drawback. Also their estimated sum of implementing the system in hospitals was around Rs.8000 that ultimately could not benefit the government hospitals.

A study carried out by a team in recent times in the UK proposed an innovative smart sensing device that helps the neonatal staff for effective and accurate measurement of new born's heart rate. This was done by placing the newborn on top of a smart mattress integrated with sensors. This system was made possible by using Electric Potential Sensors (EPS) that gives ECG readings within seconds.

Most of the research made with artificial placenta ended up using a plastic bag as a womb to nourish the baby. Here in this paper, instead of a plastic material being used to house the baby, the use of a biobag is recommended. The proposed methodology combines an artificial womb, placenta, umbilical cord, and a neonatal incubator to monitor and control the baby's environment thus making the baby feel the warmth of the mother.

III. OBJECTIVE

As mentioned earlier, most of the ongoing research involves the use of plastic-based materials in the construction of the womb as well as for the placenta and umbilical cord, which can cause allergies to newborn infants in certain cases. This paper primarily focuses on designing an artificial womb and placenta with highly biocompatible materials to provide infants with a safe and comfortable environment. As this artificial womb technology is still under research, at present, premature infants are monitored using a neonatal incubator. Thus, this paper also focuses on incorporating the principles used in the present neonatal incubator in this artificial womb technology.

IV. PROPOSED DESIGN

In recent years, the death of infants due to preterm birth has been increasing constantly. Even if these premature babies survive through the critical initial days, there will be profound consequences in the course of cardiovascular, metabolic, neurological and pulmonary diseases throughout their lives. These consequences are due to the improper development of these premature infants outside the maternal womb condition.

With the vision of providing an environment that mimics the natural human womb condition, to ensure proper growth and development of infants born before 37 weeks of pregnancy, we are designing an artificial womb along with an artificial placenta and umbilical cord. As we intend to design an artificial

ISBN 979-8-3503-5094-4

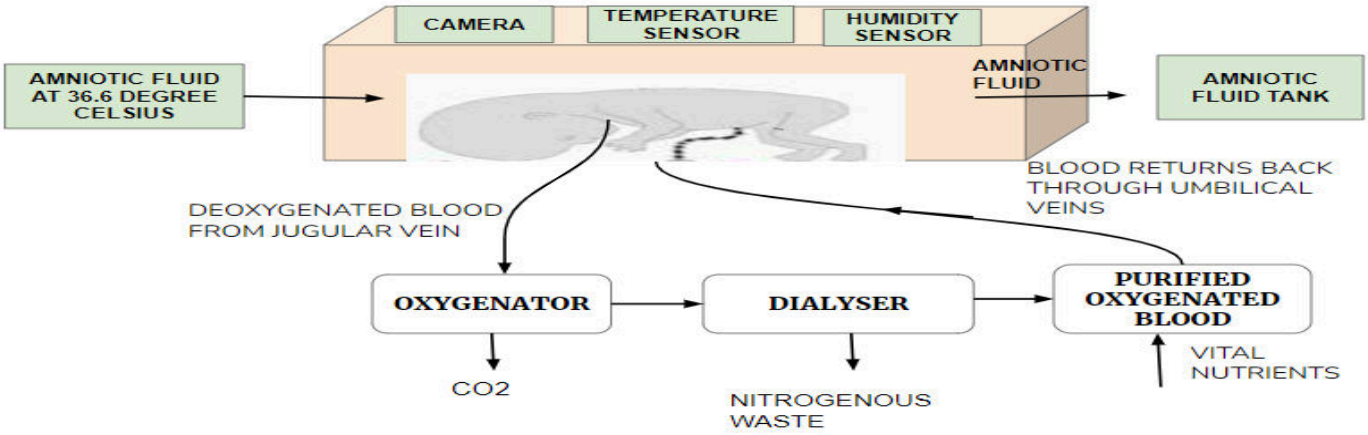


Fig. 1. Proposed artificial womb architecture

womb structure using highly biocompatible material, materials like hydrogel, synthetic scaffolds, nanoscaffolds and decellularized scaffolds are utilised for the womb construction.

In the maternal womb, the foetus floats in the amniotic fluid which is a clear, white fluid. Particles like lanugo, vernix caseosa and exfoliated squamous cells are present in this fluid. However, the major constituents of this amniotic fluid include protein, urea, uric acid, creatinine, bilirubin and estriol. In the case of normal pregnancy, 600-800 mL of amniotic fluid surrounds the foetus. However, the preterm infant has to be provided with a volume of 500-1400 mL to ensure proper growth and development. To maintain the temperature within the womb at 37 degree Celsius, amniotic fluid is pumped in at a temperature of 36.6 degree Celsius.

COMPOSITION OF ARTIFICIAL AMNIOTIC FLUID

CONSTITUENTS	VOLUME
Protein	0.19-0.26 g/dL
Urea	11 mg/dL
Uric acid	9.9 mg/dL
Creatinine	1.8-4.0 mg/dL
Bilirubin	<0.025 mg/dL
Estriol	>60 mg/dL

As mentioned earlier, with the vision of incorporating a part of modern neonatal incubator technology, an array of sensors like temperature sensors and humidity sensors are embedded in the upper portion of the biocompatible womb. The type of temperature sensor used for an artificial womb is LM35. This particular temperature sensor is easy to use and has excellent linearity. With LM35, temperature can be measured more accurately than with a thermistor. The sensor circuitry is sealed, hence it is not subjected to oxidation or other processes. For the measurement of humidity level, the DHT11 sensor is utilised. This is a basic, ultra low-cost digital humidity sensor that measures the relative humidity. It uses a capacitive humidity sensor to measure the surrounding air. These sensors would constantly monitor the temperature and humidity within the womb, as these are the two vital parameters that have to be monitored regularly.

Suppose an abnormality is detected by any one of the sensors in the embedded array. In that case, the alarm buzzer connected to the sensor array gets activated and creates a noise alerting the nurses and the doctors.

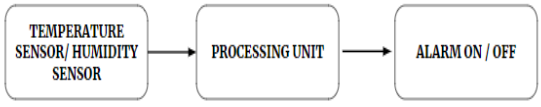


Fig. 2. Real-time sensing of internal environment of the womb

In addition, as these wombs are covered on all sides to maintain the required temperature within them, an

ISBN 979-8-3503-5094-4

ESP32-CAM module has been fixed in the inner layer of the womb for regularly monitoring the infants' conditions through a monitor. This camera module is connected to FTDI, which facilitates the real time monitoring of the infant. With the help of a wifi module, the image data acquired from the camera module is transmitted to the mobile application of both infants' parents as well as medical professionals through IOT.

Fig. 3. Foetus monitoring system

Apart from monitoring the infant's temperature and humidity, the cardiac activity of the infant has to be monitored. To accomplish this, ESP sensor embedded cotton mattress is wrapped around the chest region of the infant. As the mattress is made of cotton, it is biocompatible as well as comfortable for the newborn infants. This ESP sensor is connected to a processing unit which consists of amplifiers and filters for data filtering and acquisition. Then, the analogue output from the processing unit is sent to the microcontroller. An alarm also has been connected to this microcontroller module, which gets activated incase of abnormal heart rate. The analogue output is also digitised and sent to the display.

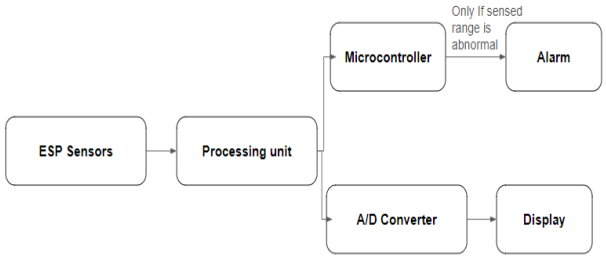


Fig. 4. Foetus cardiac activity monitoring

As the weight of the infant also has to be monitored regularly. For this purpose, a weight monitoring module has been designed. This module consists of a HX711 load cell amplifier , which has the capability to accurately sense the infant's weight, attached to a digital display through an arduino module. However, this module will not continuously monitor the weight, but only when manually activated by the medical staff and doctors.

Fig.4. Weight monitoring system

The artificial placenta and umbilical cord, not only provide nutrition to the foetus but also removes carbon dioxide & nitrogenous waste from the foetus. To design a biocompatible placenta and umbilical cord; materials such as silicone membranes and synthetic polymers are used.

V. CONCLUSION

In recent years, the research on artificial womb technology has increased to a greater extent. However, this technology remains science fiction, as it is not practically feasible. In the future, if this technology comes into practice, the lives of many premature infants can be saved in a more effective manner, which would result in a significant decrease in the infant mortality rate.

VI. REFERENCES

[1] Maria Laura Solerte MD, Ph.D, Medical Surgeon. Ob-Gyn specialist, Italy; Artificial uterus -research background to improve survival and outcome of extremely low birth weight newborns.

[2] Likith Kumar M, Sanjay Kumar M, Manoj M, Akash Kulkarni, Bhargava Dantapu, Gurusatwik Bhatta N, Remya Jayachandran "WARM WOMB: A Wireless Real-time Monitoring System for Neonatal Care in Resource-Limited Settings" 2023 International Conference on Recent Advances in Science and Engineering Technology.

[3] Rodrigo Aviles-Espinosa, Elizabeth Rendon-Morales, Zhenhua Luo, Henry Dore, Oana Anton, Heike Rabe, Robert J.Prance "Neo-SENSE: a non-invasive smart sensing mattress for cardiac

ISBN 979-8-3503-5094-4

monitoring of babies."

[4] Katarina zammer, "Artificial_Wombs_are_Science_Fiction_But artificial_Placentas_are_on_the_Horizon". Spectrum IEEE April 2021

[5] LAMROT HAILEMICHAEL WOLDEAMANUEL AND ARUN RAMAVEERAPATHIRAN; "Design of Multivariable PID Control Scheme for Humidity and Temperature Control of Neonatal Incubator."

[6] Dr. Arun Kumar Marandi1 and Dr. Harshal Shah; "A Convolution Neural Network-based Prediction of Fetal Health" 11th International Conference on System Modelling & Advancement in Research Trends December 2022.

[7] VAISALI CHANDRASEKAR, MOHAMMED YUSUF ANSARI , AJAY VIKRAM SINGH , SHAHAB UDDIN , KIRTHI S. PRABHU4 , SAGNIKA DASH, SOUHAILA AL KHODOR , ANNALISA TERRANEGRA , MATTEO AVELLA AND SARADA PRASAD DAKUA; "Investigating the Use of Machine Learning Models to Understand the Drugs Permeability Across Placenta"

[8] Brian P. Fallon1 , George B. Mychaliska; Development of an artificial placenta for support of premature infants: narrative review of the history, recent milestones, and future innovation.

[9] Shiming Gu, Wei Zhou*, Jun Li, Danyang Yang; "Internet Plus and Assisted Reproductive Information System Promotes the Healthy Development of Reproductive Medicine" 2021 Fifth International Conference on I-SMAC.

[10] Horaci Garcia, Spanish researchers aim to 'trick nature' with artificial womb.

[11] Zhangming Wei , Yi Hu ,Xiang He , Wen Ling , Jinxin Yao, Zhenjuan Li, Qiru Wang, Liping Li ; Biomaterializing the advances in uterine tissue engineering

[12] An Artificial Placenta Experimental System in Sheep: Critical Issues for Successful Transition and Survival up to One Week

[13] Haruo Usuda, Erin L. Fee, Tsukasa Takahashi, Yuki Takahashi, Sean Carter, John P. Newnham, Matthew W. Kemp; THE ARTIFICIAL PLACENTA: SCI-FI OR REALITY?

[14] Elizabeth Chloe Romanis; Artificial womb technology and clinical translation: Innovative treatment or medical research?

ISBN 979-8-3503-5094-4

Emotion–Based Media Playback system for Autistic children

Dr.A.N.Nithyaa
AssociateProfessor
Department of Biomedical engineering
Rajalakshmi Engineering College
Chennai
nithyaa.an@rajalakshmi.edu.in

Dr.R.PremKumar
Professor
Department of Biomedical engineering
Rajalakshmi Engineering College
Chennai
premkumar.r@rajalakshmi.edu.in

J.NizuShahain
PG student
Department of Biomedical engineering
Rajalakshmi Engineering College
Chennai
220311002@rajalakshmi.edu.in

R.Sriraman
PG student
Department of Biomedical engineering
Rajalakshmi Engineering College
Chennai
230311003@rajalakshmi.edu.in

***Abstract*- Children with Autism find it hard to associate with people socially and setting apart themselves. While normal people identify facial expression very quickly, children suffering from Autism Spectral Disorder (ASD) have problems in interacting with their own people. This paper presents an innovative methodology to recognize the facial expressions of children with ASD. An Emotion detection through facial expression recognition and detection of stress level using GSR (Galvanic Skin Response) sensors for ASD children is proposed.**

Keywords- Autism, facial expressions, emotion detection, recognition, Galvanic skin response.

1. INTRODUCTION

Autism Spectral Disorder is a neuro dysfunction having hardship in communication skills with increased repetitive or fixed behaviour. The level of disorder varies from person to person with autism. These people face challenges from mild to severe range. The World Health Organization says, one in 160 children is affected with autism worldwide. They find it difficult to express their emotions. There are Assistive technologies that help people with autism which holds them that reduces the workload of the therapist. Gaming applications also aids in effective ways in training the people with ASD. The influence of music in therapy for autistic children has been investigated and proven to be efficient in aiding in interaction. This project aims at capturing a real time image and recognizing it if it is a human face using Computer Vision technique. After detecting, facial expressions are observed using Neural networks. If a happy face is detected, then AT mega microcontroller and DF mini-player will play suitable audio to keep them relaxed and reduce the level of anxiety. Also a stress detector using Galvanic Skin Response is designed and built.

2.GOAL AND OBJECTIVE

The main -+aim of this technique is capturing a real time image and recognizing it for human face. After facial recognition, facial expressions are observed. If the face is detected with happy, suitable music will be played by AT mega microcontroller and DF mini-player will keep them rest and reduce tension. Also stress detector using GSR is designed and built. For this, different conductance when a person is under stress or when not need to be calculated using a Galvanic Skin Response (GSR) device.

3. LITERATURE SURVEY

Facial Emotion Detection Using Deep Learning Akriti J, et al., Their approach includes the usage of Deep Learning and Artificial Intelligence to achieve emotion detection using an image. The image is first uploaded and the face is detected. And then feature extraction is done from the image. Finally, the classification of the images based on the expression is done. This model has an accuracy of 70.14. This approach only classifies the images into categories, a real-time approach is preferred in the current scenario as it is user-friendly and the result is spontaneous.

Simple Touch sensor-based Game as Ambient A assistive device for mild Autism Spectrum Disorder Children: The possibility of embedding therapeutic modules in assistive technologies is investigated. This step is important to help children with ASD learn effectively and to reduce the therapist's workload. Although a previous systematic review argued that assistive technologies resulted in mixed reviews and outcomes. [2] There were highlighted critical issues in the recent analysis that reported that computer-based assistive devices were useful in terms of improvement of social skills.[3]

Lasri I, et al.,(2019). Facial Emotion Recognition of Students using Convolutional Neural Network: The current model makes use of Haar cascades and CNN-based emotion recognition techniques. Processing and normalization is being done which includes the detection of 7 types of expressions. Their approach, though it is helpful for teachers to assess the emotion, lacks a validation system to verify the emotion experienced by the subject. As autistic children experience a wide range of emotion which is rapid to change, predicting the emotion which is consistent, is important. There are no other means to cross-check the accuracy of the emotion detection system. Models that purely depend on facial recognition and emotion detection software aspects are said to have their own disadvantages. Several studies have shown that there is a significant degree of FER delay in relation to eye avoidance which is common in autistic children as they often hesitate to establish eye contact and have a lack of social interaction. Thus this

ISBN 979-8-3503-5094-4

require the caregiver to recognize the emotion through human-human interaction and imitate or reproduce the same to the required device.

4.EXISTING SYSTEM

Considering students, children with ASD has an advantage of having both parents and teachers taking same kind of efforts at both home and school which is equally supportive. Initially, the teacher should meet the parents before staring the lessons and discuss about the lessons for autism which will effectively help in motivating the respective children with ASD with the help of inputs from the parents side along with the teachers.

4.1 DISADVANTAGE OF EXISTING SYSTEM

The main drawback of this existing system is that the teacher can be able to concentrate on a single child at a time which makes very time consuming and the system becomes eventually less effective.

5. METHODOLOGY

The proposed system is capturing a real time image with facial recognition using Computer Vision technique. Once the human face is recognized and the predicted emotion of the face seems to be happy, then the appropriate music will be played using AT mega microcontroller and DF mini-player in order to make them feel rest and also help them to learn. With this, a stress detector is also designed and built using GSR device.

The suggested system combines GSR sensors for stress level monitoring with computer vision algorithms for real-time face recognition. An AT mega microcontroller and DF mini-player are used to detect expressions and offer appropriate audio feedback. For smooth integration, the system architecture incorporates hardware elements including the DF mini-player, Node MCU, GSR sensors, and Arduino Mega 2560.

5.1 SYSTEM ARCHITECTURE

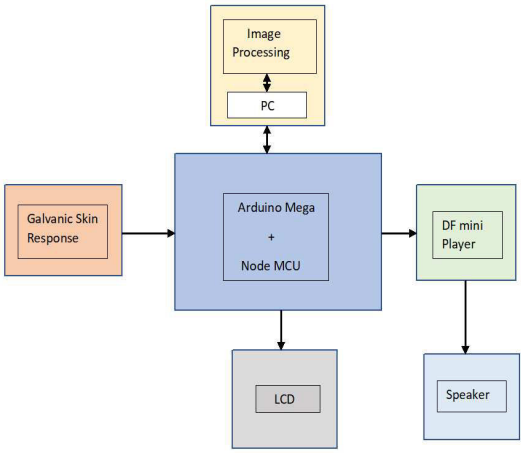


Fig 1. Proposed system architecture

The Fig 1. shows the system architecture of proposed system consisting of a Facial expression recognition and GSR for the inputs and a Liquid Crystal Display (LCD) for the output, Speaker and device with real time data logging in an IoT platform.

The Emotion-Based Media playback system for Autistic children uses a trained dataset. The dataset is taken from Kaggle datasets for facial expression detection. Convolutional Neural Network (CNN) is a part of a Deep neural network. CNN has pre-processed data, useful in image processing, video processing, facial recognition, etc., In this work, CNN is useful in feature extraction, the region of interest is around the mouth. Haar classifier is found to be very effective as it is used to detect faces in images or video sequences due to their high accuracy and low computational cost.

Real-time video or image capturing techniques make use of

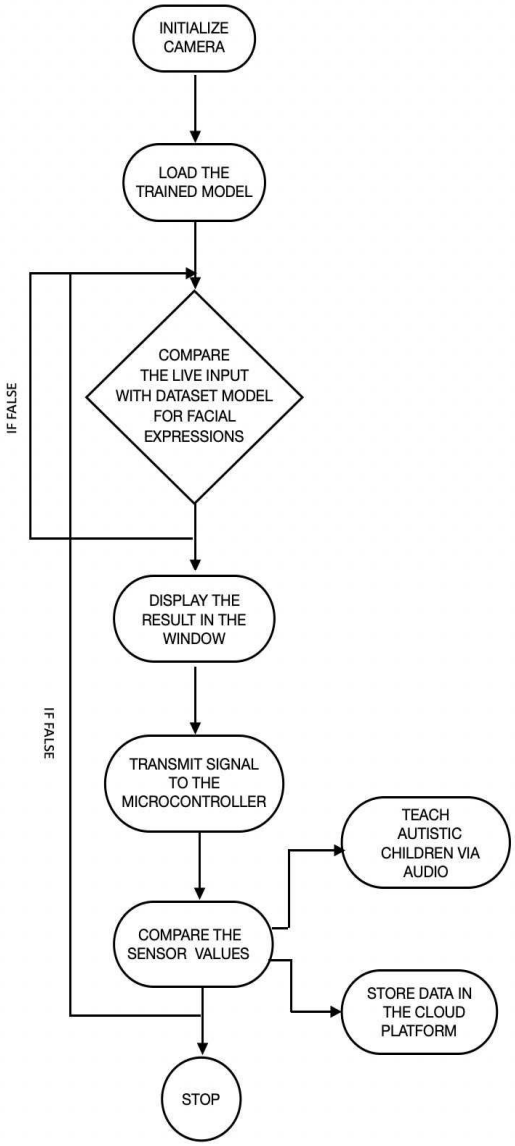


FIG. 2. FLOW CHART

the Haar cascade classifier. In this project, the Haar cascade is used to detect whether the image captured is a human face or not and also whether the face is present or not.

The purpose of this technique is to find images within an object. This method is quick because it finds the integral pictures. It can also distinguish between positive and negative pictures. Negative images relate to irrelevant images, while positive images are nothing but suggest the kind of images we want to identify. It facilitates the classification of faces into human and non-human categories. Its capabilities go beyond just identifying faces to include identifying other aspects like eyes. For this study, we gathered the data sets and compared them with real-time faces using this method. Additionally, it employs adaptive boosting strategies, which are employed to enhance the machine learning algorithm.

5.2 Hardware Details

1) Arduino Mega 2560

   The Arduino Mega 2560 is a microcontroller, it has 54 I/O pins and has many ports so that our device can be connected to many other devices or components.

2) Galvanic Skin Response Sensor

   GSR measures the skin conductance and gives an analog signal as output corresponding to sweat level i.e moisture level with respect to the physiological arousal as per sweating produced by the skin. GSR interface with Arduino will convert the analog signal to digital.

3) Node MCU

   The Node MCU is a microcontroller unit, that has an in-built wifi module. It serves as an open-source IoT prototype. Our proposed system makes use of Node MCU for data logging.

4) Liquid Crystal Display

   LCD is used to display all the system information. In this proposed system a 16 x 2 LCD screen is used. It can show 16 characters in two lines

5) DF mini-player

   DF mini-player is a small mp3 module. It flawlessly integrates with a speaker. The desired audio can be played without any delay. The audio files are stored in a secure digital card.The speaker will be connected to the DF mini-player so that the audio will be played on a loudspeaker. So that the ASD children will respond to the audio signal in a quick manner.

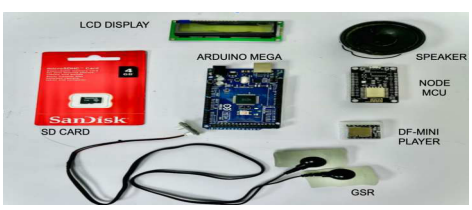


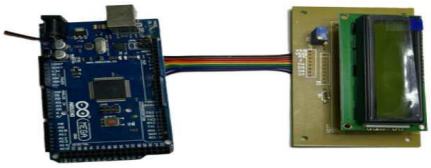
Fig. 3. Hardware Components

5.3 Arduino and Lcd Interface

Lcd display is mounted on a PCB board and it is connected with a potentiometer and a resistor. The use of a potentiometer is to adjust the brightness of the display. Arduino is connected to LCD in serial communication from digital I/O pins.

5.4 Arduino Interface with GSR

The Galvanic Skin Resistance sensor works on the electrical conductance activity on the skin by measuring the voltages



based on the amount of

Fig. 4. Arduino and Lcd Interfacing

sweat secreted by the person. It gives us the analog output. Therefore GSR is connected to the analog pin -A0 of Arduino mega.

5.5 Arduino Interface with DF mini player

Memory card with songs initiated with numbers and saved and inserted into DF mini player. As by output received from the GSR sensor and facial emotion, the song has to be played. Therefore the Rx pin is connected to Tx3 and the Tx pin is connected to the Rx pin. SPK_1 and SPK_2 pins are connected to the speaker.

5.6 Arduino Interface with Node MCU

Node MCU is a low-cost open-source IOT platform. D5 and D6 pins from Node MCU are connected to Tx2 and Rx2 respectively.

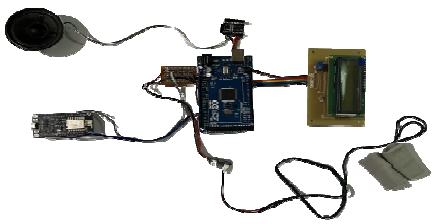


Fig. 5. Soldered Components

5.7 Things Speak- IoT analytics

ThingSpeak is an IoT analytics platform service that allows you to aggregate, visualize, and analyze real-time data streams in the cloud. You can send data from your device to ThingSpeak to instantly visualize live data and send alerts. It is open-source software written in Ruby which allows users to communicate with internet-enabled devices. It facili

ISBN 979-8-3503-5094-4

data access, retrieval, and logging of data by providing an API to both the devices and social network websites. Thing Speak was originally launched by io Bridge in 2010 as a service in support of IoT applications.

5.8 THING SPEAK CHANNEL

A new Email ID is created for the project to log in to the website. The next, step is to create an account on www.thingspeak.com by using the created email. Once an account is created login to it and go to channels and create new channels. Fill up the details which are necessary and a number of plots are needed to be viewed and save the channel.

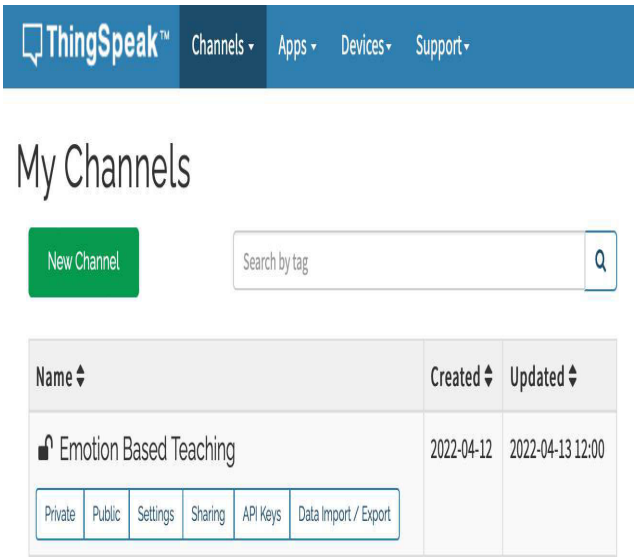


FIG. 6. CHANNEL IN THING SPEAK PLATFORM

API KEYS

To read or write data to your channel using the Thing Speak in code, you need appropriate read and write permissions. The 16-digit API key allows you to read from a private channel and write to a channel. You do not need an API key to read from a public Thing Speak channel. Account-level API keys are described in User Accounts and Channels. Once the channel is created, go to API keys in the channel and from there take the write API key to get the data from Node MCU and store it in the Thing speak platform.

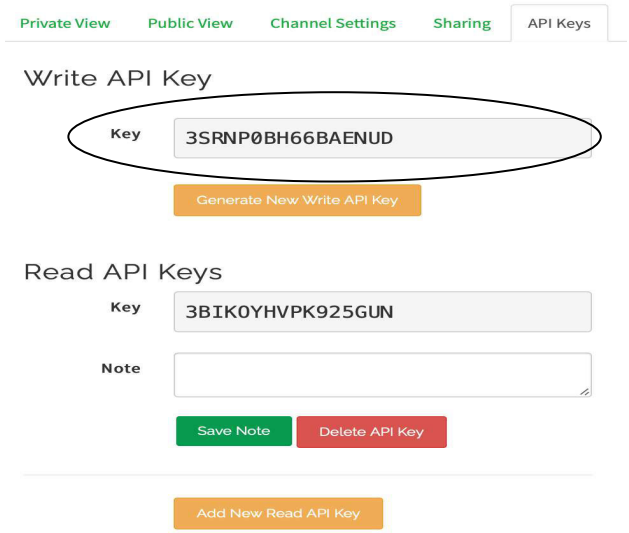


FIG. 7. API KEYS

6. RESULTS AND DISCUSSION

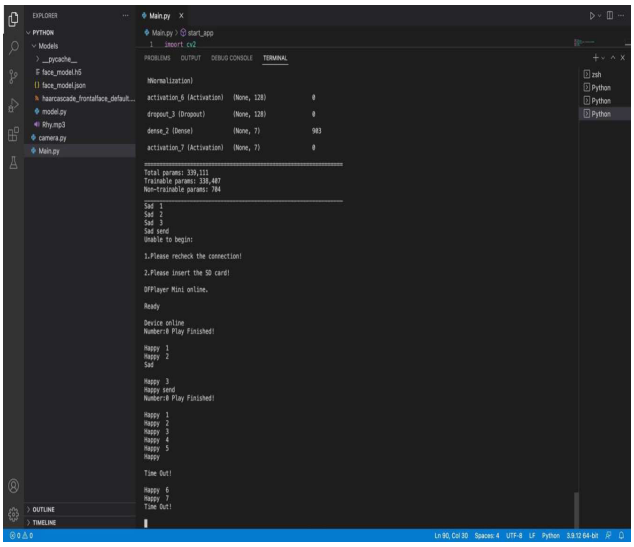
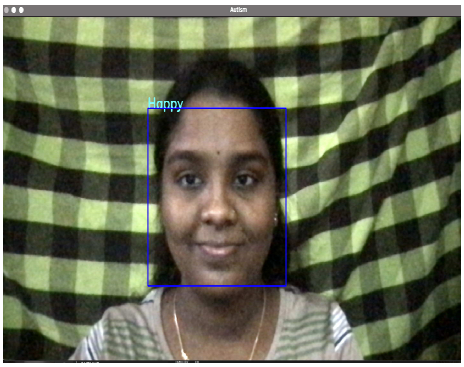


FIG. 8. PYTHON IDE OUTPUT

This FIG.8. describes that When the code is set to run there are a few necessary conditions that need to be checked, so the code checks for the connection sd card set up and, df mini player is online or not and after that the camera is initialized and it starts to recognize the face expression.



ISBN 979-8-3503-5094-4

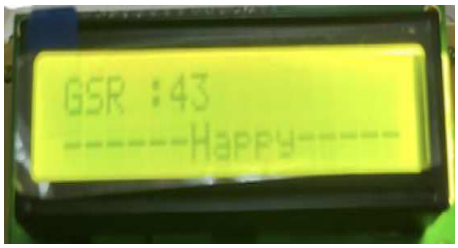
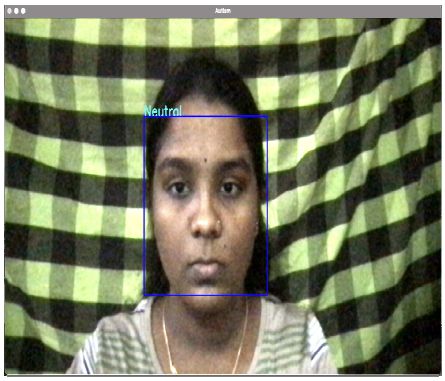


FIG.9. EMOTION IS DETECTED AS HAPPY

Fig.9. Describes that If the face emotion is detected as Happy and the GSR value is below 80 then music will start to play



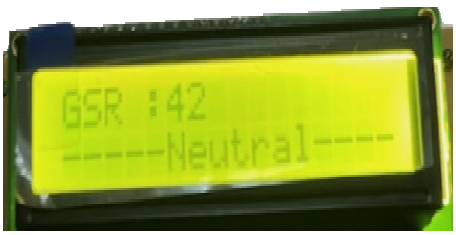


FIG 10. EMOTION IS DETECTED AS NEUTRAL

Fig.10. Describes that If the face emotion is detected as Neutral and the GSR value can be seen LCD and no music played

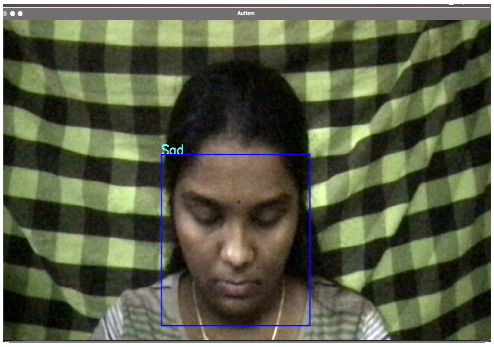


FIG.11.EMOTION IS DETECTED AS SAD

Figure.11. Describes that If face emotion is detected as Sad and the GSR value is above 80 then the music will start to play

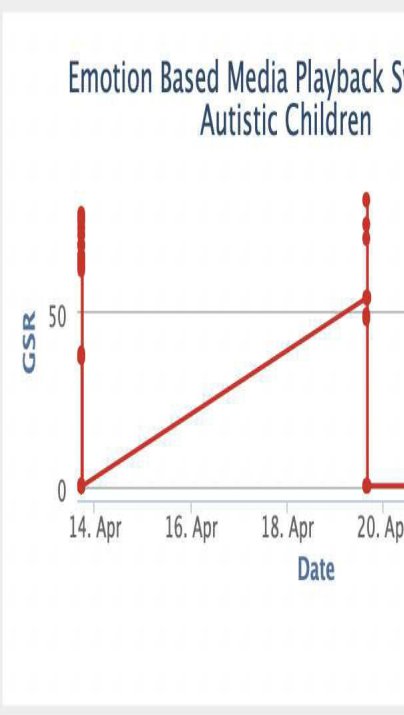


FIG. 12. GRAPHICAL REPRESENTATION OF IOT DATA

ISBN 979-8-3503-5094-4

NAME	GENDER	AGE	FACE EMOTION	GSR VALUE %
SUBJECT 1	Female	9	Happy	39
SUBJECT 2	Female	4	Neutral	41
SUBJECT 3	Male	8	Happy	62
SUBJECT 4	Male	49	Neutral	77
SUBJECT 5	Female	21	Neutral	49
SUBJECT 6	Female	7	Sad	55
SUBJECT 7	Female	21	Sad	39
SUBJECT 8	Female	21	Happy	60
SUBJECT 9	Female	22	Happy	52
SUBJECT 10	Female	41	Happy	80
SUBJECT 11	Female	34	Neutral	70
SUBJECT 12	Male	17	Sad	66
SUBJECT 13	Female	71	Neutral	43
SUBJECT 14	Male	80	Sad	58
SUBJECT 15	Male	53	Sad	67
SUBJECT 16	Male	39	Happy	55
SUBJECT 17	Male	57	Neutral	60
SUBJECT 18	Female	47	Happy	88
SUBJECT 19	Female	44	Sad	50

TABLE. 1. IOT DATA

Created_at	Entry Id	Gsr value %	Face Emotion
2022-04-13T15:15:48+05:30	170	72	Neutral
2022-04-13T15:16:04+05:30	171	71	Happy
2022-04-13T15:16:21+05:30	172	73	Happy
2022-04-13T15:16:39+05:30	173	73	Happy
2022-04-13T15:16:56+05:30	174	73	Happy
2022-04-13T15:17:13+05:30	175	69	Happy
2022-04-13T15:17:30+05:30	176	70	Neutral
2022-04-13T15:17:48+05:30	177	71	Neutral
2022-04-13T15:18:05+05:30	178	82	Sad
2022-04-13T15:18:22+05:30	179	80	Neutral
2022-04-13T15:18:39+05:30	180	78	Neutral
2022-04-13T15:18:56+05:30	181	79	Neutral
2022-04-13T15:19:13+05:30	182	76	Happy
2022-04-13T15:19:30+05:30	183	75	Neutral

TABLE. 2. TEST DATA

7.CONCLUSION

Face recognition is one of the domains of Computer vision to identify and recognizing images. At present, they solve various real-world problems in different fields like healthcare, manufacturing and industries, etc. In this project, face recognition is used to detect emotion and interfaces it with GSR which helps to detect stress levels in children with ASD based on their emotions. From the above output, the system takes both GSR values and Face recognition data and analyses it. Songs are played by DF mini-player based on the obtained output. Real-time testing demonstrates the system's capability to detect facial expressions and stress levels accurately. By analyzing GSR data and facial expressions, the system effectively provides appropriate feedback, as evidenced by the results presented.

ISBN 979-8-3503-5094-4

8. FUTURE SCOPE

This project can be developed for giving therapy for people with nerve damage, autism and other brain related disorders by stimulating the brain nerves by giving them vibrations. A case report by Bressel, Gibbons, and Samaha (2011) reported that whole-body vibration aids in treating some stereotypes in children ageing from four to six.

9. REFERENCES

[1] Mohammed Taj-Eldin, Christian, Ryan, Brendan O'Flynn, and Paul Galvin, "*A Review of Wearable Solutions for Physiological and Emotional Monitoring for Use by People with Autism Spectrum Disorder and Their Caregivers*",pubmed.ncbi [2018]

[2] Muskan Agarwal, Pratyush Saxena, K.B. Dubey, "Facial Emotion detector and music player system", IRJET-V8I7438 [2021].

[3] Sarah Afiqah Mohd Zabidi, Hazlina Md Yusof, Shahrul Naim Sidek, Sukreen Hana Herman, "*Simple Touch sensor-based Game as Ambient A assistive device for mild Autism Spectrum Disorder Children",*ResearchGate [2021]

[4] Saranya Rajan, Poongodi Chenniappan, Somasundaram Devaraj, Nirmala Definition of Autism - Autism Awareness (autismawarenesscentre.com)Madia, *"Facial expression recognition techniques: a comprehensive survey*'',The institution of engineering and technology [2019]

[5] Steve Berggren, Fletcher-Watson Nina Milenkovic, Peter B, "*Emotion recognition training in autism spectrum disorder: A systematic review of challenges related to generalizability*", pubmed.ncbi[2017].

[6] Tarannum Zaki, Muhammad Nazrul Islam, Md. Sami Uddin, Sanjida Nasreen Tumpa, Md. Jubair Hossain,Maksuda Rahman Anti, Md. Mahedi Hasan, "Towards Developing a Learning Tool for Children with Autism",ResearchGate [2017]

[7] Uzma Hasan, Md. Fourkanul, Muhammad Nazrul, Sifat Bin Zama, "*Towards Developing an IoT Based Gaming Application for Improving Cognitive Skills of Autistic Kids*",ResearchGate [2020]

ISBN 979-8-3503-5094-4

CNN - Based Breast Cancer detection using Non-invasive methods

Dr. M. Neela Harish
Assistant Professor, Department of Biomedical Engineering, Easwari Engineering College, Chennai, India
neelaharish.m@eec.srmrmp.edu.in

Dr. S. Sathish
Assistant Professor, Department of Biomedical Engineering, Easwari Engineering College, Chennai, India
sathish.s@eec.srmrmp.edu.in

Dharun K
UG student, Department of Biomedical Engineering, Easwari Engineering College, Chennai, India
dharunkumar07.pro@gmail.com

Keerthivasan R
UG student, Department of Biomedical Engineering, Easwari Engineering College, Chennai, India
vasanrkeerthi@gmail.com

Rudhresh V
UG student, Department of Biomedical Engineering, Easwari Engineering College, Chennai, India
rudhresh1741@gmail.com

***Abstract*— One of the foremost reason for the maximum number of female deaths in world is due to the breast cancer. According to WHO, one among eight women is affected by breast cancer worldwide. This study introduces an advanced non-invasive method for detecting and classifying breast cancer using deep learning. Utilizing CT mammogram images, the research aims to implement a Convolutional neural network (CNN) along with a Region Proposal Network (RPN) for precise localization. The model first pinpoints potential cancerous regions and then classifies them into specific types. Tested on diverse datasets, our approach outperforms traditional methods. By accurately identifying cancer and its subtypes, our model promises to enhance early diagnosis and aid in personalized treatment. This research marks a significant step toward improving breast cancer detection, paving the way for more effective healthcare interventions and better patient outcomes. The proposed framework stands as a promising advancement in the realm of medical diagnostics, contributing to improved patient outcomes and more efficient healthcare interventions.**

Keywords*— *CNN, Image Processing, Breast cancer, Benign, Malignant

I. INTRODUCTION

Although it can strike males as well, breast cancer is a common and possibly fatal illness that mostly affects women. The importance of comprehending breast cancer's frequency, symptoms, and impact on people and society is highlighted by the disease's worldwide burden. Based on current data, breast cancer is the most prevalent cancer among women globally. According to predictions from the World Health Organization (WHO), there will be around 2.5 million new cases of breast cancer worldwide in 2020. Breast cancer continues to be the primary cause of cancer-related fatalities despite improvements in diagnosis and treatment. It caused more than 685,000 fatalities worldwide in 2020[1]. The incidence and death rates of breast cancer vary significantly by geography. While incidence rates are often greater in wealthier nations, death rates are higher in less developed nations.

A painless lump or tumor in the breast or underarm is the most typical sign of breast cancer. While not every bump is malignant, any new or strange changes need to be checked out right once. Breast cancer may be indicated by changes in the breast's size, shape, or appearance, as well as skin abnormalities like redness or dimpling. Inversion, discharge, or unexplained discomfort in the nipples can potentially be warning indicators of breast cancer.

Breast cancer and its treatments, including surgery, chemotherapy, and radiation, can have significant physical effects. These may include fatigue, hair loss, changes in body image, and potential long-term side effects. A breast cancer diagnosis can cause emotional distress, anxiety, and depression. Coping with the uncertainty of the disease, treatment, and potential impact on one's life can be challenging. Breast cancer can also have broader societal and economic consequences, including the strain on healthcare systems, increased healthcare costs, and impacts on workforce productivity as individuals undergo treatment and recovery. The proposed framework stands as a promising advancement in the realm of medical diagnostics, contributing to improved patient outcomes and more efficient healthcare interventions.

The main objective of conducting this study is the rigorous evaluation of diverse datasets will validate the model's generalization and superiority over traditional methods, contributing to the advancement of early diagnosis and

ISBN 979-8-3503-5094-4

personalized treatment strategies in breast cancer. Evaluate the performance of the proposed deep learning model on diverse and representative datasets to ensure generalization across varied patient populations.

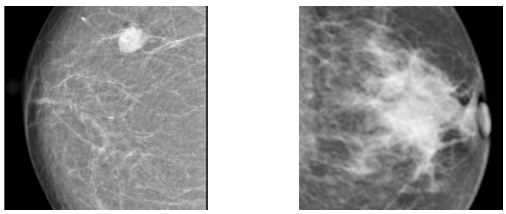


Fig .1a. Benign Breast cancer b. Malignant Breast Cancer

II. LITERATURE REVIEW

Suleiman A. AlShowarah (2023) [2] has implemented a deep learning model for the diagnosis of breast cancer with the help VGG-19 model as a bilayer network. To increase the accuracy of the training model, he used statistical operations such as maximum, minimum and average operations along with VGG-19 to extract the fusion features. He also experimented with various classifiers for the results of the training model to predict the types of breast cancer. The outcomes showed that the random forest model gives high accuracy (88.36%), recall and precision.

A hybrid model that combines CNN architecture with the Ebola optimization search algorithm (EOSA) for breast cancer diagnosis and detection was developed by Tehnan I. A. Mohamed et al. (2023)[3] using gene expression data. A number of pre-processing methods were employed to get the data ready, including the removal of outliers using Array-Array Intensity Correlation (AAIC). To ensure that the expression measures were free of bias, they employed the normalization process. After that, they converted the gene expression data into grayscale images and produced two-dimensional graphics. For the classification procedure, the EOSA-CNN model is used. In terms of accuracy, precision, recall, f1-score, kappa, specificity, and sensitivity with a percentage over ninety, the study's findings demonstrate that the proposed model satisfied all performance requirements for the malignant class.

In this study, an efficient deep learning-assisted AdaBoost approach was presented by Jing Zheng et al. (2020) [4] for the early identification and detection of breast cancer. The AdaBoost technique is used to build the final prediction function of the ensemble classifier. The research has directly increased the accuracy of the neural network's outcome prediction and shown significant promise for speedy generalization. Using the high-deep learning from the DLA-EABA-suggested Convolutional Neural Network deep learning model improved the system's performance. Deep learning approaches may be customized to the specific features of a dataset as they are based on machine learning and each data set necessitates the development of a unique model.

The research work of Xiaofan Cheng et al. (2021)[5]began with the issue of the poor accuracy of conventional medical equipment for diagnosing breast cancer in conjunction with the now widely used deep learning techniques for investigation. The accuracy of diagnosing benign and malignant breast cancer is successfully increased by fusing breast cancer characteristics taken from various networks using a range of convolutional neural network model fusion techniques. To conduct tests, the dataset is separated into binary data that is benign and data that is malignant and magnified 40X, 100X, 200X, and 400X. The 6:2:2 data ratio is used to separate the data into training, validation, and training sets. 20% of the data is utilized for model optimization and verification, 60% of the data is used for training, and the remaining 20% is used to verify the performance of the model. They used Tensor flow as a framework and ResNet-50, VGG-16 and Inceptionv3 as layers for training the model. Ultimately, the categorization outcome achieved a 98.26 percentage accuracy rate. It demonstrates how our strategy may employ computer-assisted diagnostics to help breast cancer patients receive treatment on time and can successfully address issues with physicians and underdeveloped facilities in some distant places.

The paper titled "Breast Cancer Detection Using Extreme Learning Machine Based on Feature Fusion with CNN Deep Features"[6] by Zhiqiong Wang et al, introduces an innovative approach to breast cancer detection by combining the strengths of Extreme Learning Machines (ELM) and Convolutional Neural Networks (CNN). The research focuses on leveraging deep features extracted from CNNs for image representation and subsequently fusing these features with traditional handcrafted features. This fusion aims to enhance the discriminative power of the feature set used in breast cancer detection. The proposed method utilizes ELM as a classifier, taking advantage of its fast training speed and efficient generalization capabilities. The integration of deep features from CNNs and traditional features aims to capture both high-level semantic information and detailed structural characteristics in medical images. The paper likely provides insights into the experimental evaluation of the proposed method, including performance metrics such as sensitivity, specificity, and accuracy. This work is significant for its exploration of a hybrid approach, merging deep learning representations with classical machine learning techniques, to improve the accuracy of breast cancer detection systems

In summarizing the research works, we have found that microscopic images were mostly used for the data set for the training model and it took enormous time to train the model to attain a good accuracy with precision as a drawback of the previous research works. The usage of multilayered networks with different training models yielded better outcomes but the data set of more numbers made it computationally complex.

III. PROPOSED SYSTEM

The methodology adopted for this research constitutes a comprehensive approach, encompassing key stages such as dataset collection, preprocessing, and the development of a sophisticated deep learning model dedicated to the critical task of breast cancer detection and classification. In particular, the model leverages the capabilities of a convolutional neural network (CNN) equipped with a region

ISBN 979-8-3503-5094-4

proposal network (RPN), a combination designed to enhance precision in localization and facilitate detailed Feature extraction [7]. The training phase involves the utilization of a diverse dataset, with a strategic incorporation of transfer learning techniques aimed at optimizing efficiency and

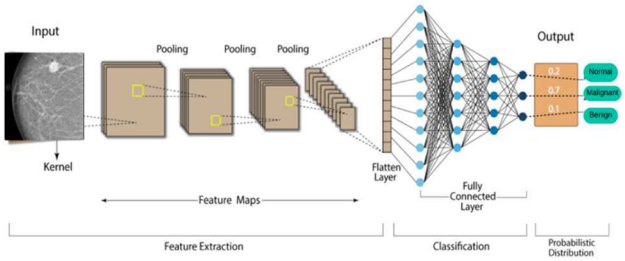


Fig .2.Block Diagram of Proposed system

A notable feature of the proposed methodology is its emphasis on subtype classification within the breast cancer detection framework. This additional layer of classification not only enhances the model's diagnostic capabilities but also contributes to a more nuanced understanding of the diverse manifestations of breast cancer. Rigorous validation procedures are implemented to assess the model's performance, employing metrics such as accuracy and sensitivity [8]. This ensures a thorough evaluation of the model's efficacy and reliability in real-world scenarios.

In adherence to ethical standards, the methodology places a high priority on safeguarding patient privacy and obtaining necessary approvals for the use of medical data. Stringent measures are implemented to uphold confidentiality and comply with regulatory requirements, acknowledging the sensitive nature of healthcare information. By addressing ethical considerations, the research endeavors to establish a framework that not only advances the accuracy of early breast cancer diagnosis but also aligns with ethical norms, reinforcing trust in the application of deep learning technologies in the medical domain. Ultimately, this methodology aspires to contribute to the refinement of personalized treatment strategies by providing a robust and reliable tool for clinicians and healthcare professionals.

IV. SIMULATION EXPERIMENTS AND DISCUSSIONS

A. *Tools used*

Anaconda is a distribution of open-source software for data science, machine learning, and scientific computing [9]. Anaconda includes a variety of tools and libraries such as Python, Jupyter Notebooks, NumPy, SciPy, pandas, and more, which are commonly used in data analysis and scientific research. However, Anaconda comes with its own package and environment management system called conda. Conda helps users install, manage, and update packages and dependencies within their Anaconda environments. Jupyter Notebooks are widely used for interactive data analysis, prototyping, and sharing research findings.

Keras is a Python API framework especially for neural networks and works with as part of the Tensor Flow library. Keras provides a high-level, user-friendly API that allows developers to define and train neural networks without delving into the lower-level details of neural network implementation [10]. Keras offers a modular approach to building neural networks. Users can assemble neural network models by connecting configurable building blocks called layers. Developers can extend and customize Keras by creating their own layers, objectives, and optimizers. This makes it adaptable to various research and application needs.

B. *Technology stack*

Convolutional Neural Networks (CNNs) have emerged as a transformative technology in the field of computer vision, providing a robust framework for image analysis and recognition tasks. In recent research papers, CNNs with sequential algorithm have been extensively employed to address complex challenges, demonstrating their efficacy in tasks such as image classification, object detection, and segmentation [11]. The architecture of CNNs, with their convolutional layers for feature extraction and hierarchical representation learning, has proven highly adaptable to various datasets and domains. Researchers often leverage the power of transfer learning, employing pre-trained CNN models on large datasets like ImageNet and fine-tuning them for specific applications with limited data. The ability of CNNs to automatically learn hierarchical features from raw pixel values has played a pivotal role in advancing the state-of-the-art in image-related research. Their success has not only facilitated breakthroughs in computer vision but has also paved the way for interdisciplinary applications, including medical imaging, remote sensing, and autonomous systems. As the field continues to evolve, CNNs remain a cornerstone in the pursuit of more accurate, efficient, and interpretable visual recognition systems, contributing significantly to the advancement of research across a myriad of domains.

C. *Dataset*

For the experiments, we have taken the CT mammogram image datasets from the open-source web application- Kaggle [12]. The dataset consists of 30 different features of 570 distinct data on both malignant and benign tumours. For training, we have used 50 malignant and 50 benign data from the dataset to proceed the training process for preprocessing, feature extraction and analyses the data to reach optimal accuracy above 85%. The test data set is fed into the model for getting the result.

D. *Process Flow*

The sequence algorithm for Convolutional Neural Networks (CNNs) involves a structured set of steps for developing and deploying these networks for tasks such as image classification [13]. The first step is data collection, where a labeled dataset is amassed to train and evaluate the CNN. Following this, data preprocessing is conducted, involving tasks like resizing images and normalizing pixel values [14].

The next crucial stage involves designing the model architecture, specifying the number and types of layers, including convolutional, pooling, and fully connected layers, to capture hierarchical features in the data. Subsequently, the

ISBN 979-8-3503-5094-4

model is compiled, wherein an optimizer, loss function, and evaluation metrics are chosen. Data splitting comes next, dividing the dataset into training, validation, and test sets. The CNN is then trained on the training set using backpropagation and optimization algorithms. Hyper parameter tuning fine-tunes aspects like learning rate and batch size based on performance on the validation set. Model evaluation on the test set is paramount, employing metrics such as accuracy to gauge the CNN's effectiveness. Once the model proves satisfactory, it can be deployed for predictions on new, unseen data, marking the culmination of the CNN development and deployment sequence.

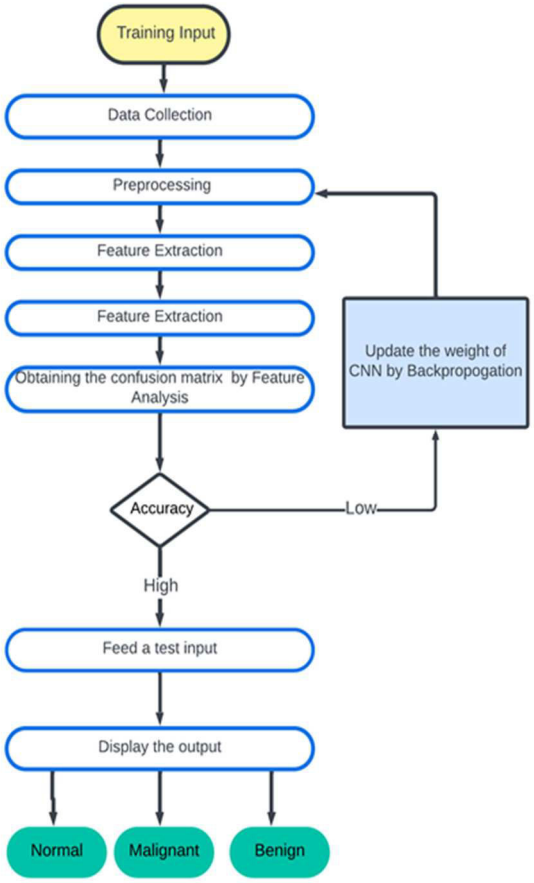


Fig .3. CT Mammogram dataset

id	diagnosis	radius_me	texture_n	perimeter	area_mea	smoothne
842302	M	17.99	10.38	122.8	1001	0.1184
842517	M	20.57	17.77	132.9	1326	0.08474
84300903	M	19.69	21.25	130	1203	0.1096
84348301	M	11.42	20.38	77.58	386.1	0.1425
84358402	M	20.29	14.34	135.1	1297	0.1003
843786	M	12.45	15.7	82.57	477.1	0.1278
844359	M	18.25	19.98	119.6	1040	0.09463
84458202	M	13.71	20.83	90.2	577.9	0.1189
844981	M	13	21.82	87.5	519.8	0.1273
84501001	M	12.46	24.04	83.97	475.9	0.1186
845636	M	16.02	23.24	102.7	797.8	0.08206
84610002	M	15.78	17.89	103.6	781	0.0971
846226	M	19.17	24.8	132.4	1123	0.0974
846381	M	15.85	23.95	103.7	782.7	0.08401
84667401	M	13.73	22.61	93.6	578.3	0.1131
84799002	M	14.54	27.54	96.73	658.8	0.1139
848406	M	14.68	20.13	94.74	684.5	0.09867
84862001	M	16.13	20.68	108.1	798.8	0.117
849014	M	19.81	22.15	130	1260	0.09831
8510426	B	13.54	14.36	87.46	566.3	0.09779
8510653	B	13.08	15.71	85.63	520	0.1075

Fig .4. Flow chart of the Process Flow of Proposed Model

V. CONCLUSION AND FUTURE WORKS

Classification of breast cancer with this sequential based CNN model are quicker than other CNN model such as ResNet model because a sequential algorithm would have a time complexity that scales with the size of the input data in a linear manner. Our proposed model generates a confusion matrix output with the overall accuracy of both type of breast cancer as 89%. Our future work is to increase the accuracy, precision and recall of the system by combining it with another deep learning model without affecting the time complexity of the proposed system.

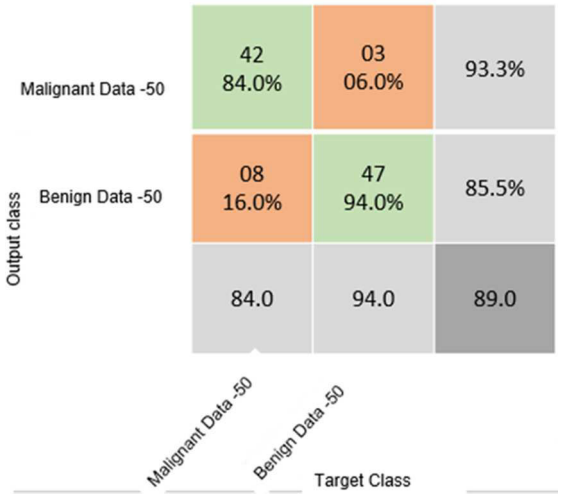


Fig .5. Confusion matrix results of the proposed system

REFERENCES

[1] https://www.who.int/news-room/fact-sheets/detail/breast-cancer

[2] AlShowarah, Suleyman A. "Breast Cancer Detection System using Deep Learning Based on Fusion Features and Statistical Operations International Journal of Advanced Computer Science and Applications, Vol. 14, Iss. 11, 2023

[3] Mohamed TI, Ezugwu AE, Fonou-Dombeu JV, Ikotun AM, Mohammed M. "A bio-inspired convolution neural network architecture for automatic breast cancer detection and classification using rna-seq gene expression dat"a. Scientific Reports, Sep 2023

[4] Zheng J, Lin D, Gao Z, Wang S, He M, Fan J." Deep learning assisted efficient AdaBoost algorithm for breast cancer detection and early diagnosis". IEEE Access, May 2020

[5] Cheng X, Tan L, Ming F." Feature fusion based on convolutional neural network for breast cancer auxiliary diagnosis". Mathematical Problems in Engineering, Sep 2021 .

[6] Wang, Zhiqiong, et al. "Breast cancer detection using extreme learning machine based on feature fusion with CNN deep features." IEEE Access 7, 2019

[7] Keshari, Rohit, et al. "Learning structure and strength of CNN filters for small sample size training." proceedings of the IEEE conference on computer vision and pattern recognition. 2018.

ISBN 979-8-3503-5094-4

[8] Wang, Dong, et al. "Decoupled R-CNN: Sensitivity-specific detector for higher accurate localization." IEEE Transactions on Circuits and Systems for Video Technology 32.9 ,2022.

[9] Yan, Yuxing, and James Yan. Hands-On Data Science with Anaconda: Utilize the right mix of tools to create high-performance data science applications. Packt Publishing Ltd, 2018.

[10] Moolayil, Jojo, Jojo Moolayil, and Suresh John. Learn Keras for deep neural networks. Berkeley: Apress, 2019.

[11] Pitas, Ioannis. Digital image processing algorithms and applications. John Wiley & Sons, 2000.

[12] https://www.kaggle.com/datasets/yasserh/breast-cancer-dataset?resource=download

[13] Dhawan, Atam P. Medical image analysis. John Wiley & Sons, 2011.

[14] Zanddizari, Hadi, et al. "A new preprocessing approach to improve the performance of CNN-based skin lesion classification." Medical & Biological Engineering & Computing 59.5, 2021

ISBN 979-8-3503-5094-4

Sensor-Based Gait Analysis and Machine Learning for Predictive Fall Risk Assessment in Adults

Dr Gayathri T
Department of Biomedical Engineering
KPR Institute of Engineering and Technology
Coimbatore,India
gayathri.t@kpriet.ac.in

Mr. Boopathi Sathivel
Founder and CEO
Prashan Medical Technologies
Coimbatore, India
prashanmedtech@gmail.com

Indiera M
Department of Biomedical Engineering
KPR Institute of Engineering and Technology
Coimbatore, India
indieramurugan@gmail.com

Sanjay Kumar C
Department of Biomedical Engineering
KPR Institute of Engineering and Technology
Coimbatore,India
sanjaychinraj2003@gmail.com

Pravin kumar S
Department of Biomedical Engineering
KPR Institute of Engineering and Technology
Coimbatore,India
kumarpravin16422@gmail.com

SudalaiMani B
Department of Biomedical Engineering
KPR Institute of Engineering and Technology
Coimbatore, India
Sudalai1786@gmail.com

***Abstract*— The need for fall prediction and prevention measures arose from the increasing global ageing population's increased incidence of falls among the elderly. The combination of Internet of Things (IoT) devices and machine learning algorithms is becoming more and more prevalent in these methods. These algorithms are able to identify changes in gait patterns based on data gathered from many sensors, such as accelerometers and gyroscopes, providing early warning signs of approaching falls.**

Keywords*— *Fall prediction, Gait, IOT, Machine learning, Accelerometer, Gyroscope.

I. INTRODUCTION

Falls in the elderly population are a significant concern for health and safety. These technologies enable continuous monitoring of elderly individuals and can assist in identifying fall risk factors in advance, facilitating effective fall prediction and prevention strategies. These strategiesrevolve around the integration of machine learning and IoT devices. The application of machine learning algorithms and IoT devices in fall prediction systems for the elderly hasgarnered considerable attention in recent times [2]. Continuous monitoring via wearable devices like sensors is pivotal in predicting falls. These systems offer real-time data on gait and movement patterns, enabling the early detection of changes indicating an elevated risk of falling[1]. Utilizing ESP 8266 and MPU6050, gait analysis can be conducted to assess an elderly person's walking patterns and identify any irregularities signifying a higher fall risk [4].By integrating machine learning algorithms such as K- nearest neighbors, support vector machines, and random forests with the collected sensor data from elderly individuals, a more precise fall prediction system can be developed [2] [4]. This solution entails placing accelerometers strategically, like at the knee joint, ankle,and hip, to capture the dynamic and intricate nature of daily movements [1].

II. METHODOLOGY

A. *Hardware system*

Hardware system To facilitate gait analysis and fall prediction for elderly individuals, we've integrated ESP8266 sensors and Node MCU modules at critical locations, specifically the hip, knee, and ankle [4]. This setup enhances data accuracy and real-time monitoring compared to the previous design [4].

Precisely positioning MPU6050 sensors at the hip, knee, and ankle joints plays a pivotal role in ensuring accurate data collection for sensor-based gait analysis and fall prediction systems. The sensor at the hip joint is securely attached to the upper thigh or hip area, aligned parallel to the thigh's longitudinal axis. This setup facilitates the measurement of hip movements and gait patterns critical for assessing fall risk. Similarly, at the knee joint, the sensor is affixed to the lower thigh or just above the knee, capturing knee flexion-extension movements during various activities. Meanwhile, the sensor placement around the ankle region enables the monitoring of ankle dorsiflexion-plantarflexion movements, offering insights into dynamic balance and foot clearance during walking. Calibration of these sensors is indispensable to ensure data accuracy by minimizing errors and discrepancies. Calibration techniques involve procedures like zero-offset calibration to remove bias and sensitivity calibration to address sensor response variations. Additionally, temperature compensation techniques may be applied to mitigate environmental influences on sensor readings. These calibration measures are essential for maintaining data accuracy and reliability, thereby enhancing the effectiveness of gait analysis and fall prediction systems. [11].

These sensors communicate wirelessly with a central unit for data acquisition and processing. Node MCU boards effectively handle this task. The collected data fromthese sensors enables comprehensive analysis, including gait patterns and fall prediction. Algorithms utilize motion data to evaluate stability and detect anomalies that may indicate a potential fall. This advanced system provides valuable insights into the movements and balance of elderly individuals, enhancing their safety and well-being [4]. An overview of gait analysis and fall prediction is shown in the figure [Figure-1].

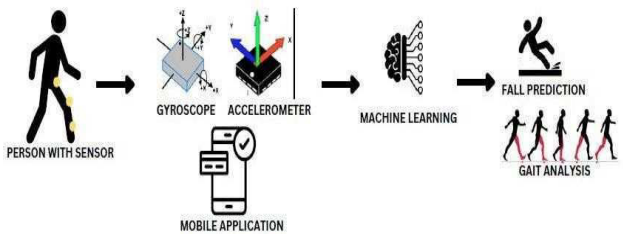


Figure 1. An overview of gait analysis and Fall prediction

ISBN 979-8-3503-5094-4

Calibration stands as a critical process in ensuring the precision and reliability of sensor data, particularly in applications like fall prediction and gait analysis. Its significance lies in its ability to minimize inaccuracies arising from sensor drift, variations in temperature, or inconsistencies in manufacturing. By aligning the sensor's output with known reference values or standardizing its measurements, calibration enhances the accuracy of collected data. Moreover, calibrated sensors provide consistent and repeatable measurements across diverse environments and sessions, facilitating reliable comparisons and longitudinal analysis of gait patterns. Techniques such as zero-offset calibration mitigate biases in sensor readings, while scale factor calibration adjusts sensor sensitivity to maintain linearity with measured physical quantities. Temperature compensation methods address errors induced by temperature fluctuations, ensuring the reliability of measurements under varying environmental conditions. Dynamic calibration approaches characterize sensor responses across dynamic scenarios, aiding in identifying and rectifying nonlinearities or cross-axis sensitivities. Overall, calibration is indispensable for optimizing sensor performance and improving the quality and validity of data gathered for fall prediction and gait analysis.

B. Machine learning

Machine learning stands as a subset of artificial intelligence, dedicated to crafting algorithms and models endowed with the ability to learn from data. Its core strength lies in its capacity to identify inherent patterns and structures within datasets, facilitating the extraction of valuable insights[6]. Through this learning process, machine learning systems gain the ability to generalize their knowledge, enabling them to make predictions and classifications on previously unencountered data. This ability to extrapolate from existing information to make informed decisions forms the crux of machine learning's utility across various domains,including gait analysis and fall prediction [3].

In the training process for machine learning models used in fall prediction and gait analysis, the first step involves preprocessing the labeled dataset. This includes cleaning the data by addressing missing values, outliers, and normalizing features to ensure consistency and enhance model performance. Following preprocessing, the appropriate machine learning algorithms such as k-Nearest Neighbors (k-NN), Support Vector Machines (SVM), or Random Forests are chosen based on factors like dataset size, problem complexity, and computational resources available. These models are then trained using the labeled dataset, where features derived from sensor measurements (e.g., accelerometer and gyroscope data) are used to predict fall or non-fall events. After training, the model's performance is evaluated using various metrics such as accuracy, sensitivity, specificity, and area under the ROC curve (AUC) on a separate testing dataset. This evaluation assesses the model's ability to generalize to unseen data and accurately predict fall events. Challenges in the training process may include optimizing model parameters, selecting the most suitable algorithm, and handling imbalanced datasets. Strategies like hyperparameter tuning and cross-validation techniques are employed to address these challenges and improve model performance.

In fall prediction and gait analysis, a balanced dataset is fundamental to prevent biases and enable the model to learn effectively from both fall and non-fall instances.

Imbalanced datasets, where one class significantly outweighs the other, can lead to skewed predictions, with the model favoring the majority class. This imbalance undermines the model's ability to accurately identify instances of the minority class, such as falls, compromising overall performance. Cross-validation techniques, such as k-fold cross-validation or stratified cross-validation, are essential for assessing model performance while guarding against overfitting. By dividing the dataset into multiple subsets and iteratively training and testing the model on different subsets, cross-validation provides a robust estimate of the model's performance on unseen data. This method helps detect overfitting, where the model performs well on the training data but fails to generalize to new data, by evaluating performance across various iterations and subsets of the dataset. Integrating cross-validation ensures that performance metrics accurately reflect the model's ability to generalize and make reliable predictions in real-world scenarios.

1. Supervised Learning:

In supervised learning, the algorithmlearns to make predictions or classifications by training on labeled data. Each data instance in the training set is associated with a known label or outcome. The algorithm infers a function during training that maps input data to their corresponding labels during training that maps input data to their corresponding labels. [3].

2. Unsupervised Learning:

Unsupervised learning, on the other hand, deals withunlabeled data [3]. The algorithm's task is to uncover patterns, structures, or groupings within the data without anypre-defined labels [5]. It essentially makes self-discoveries and can model probability distributions of the input data [3].

3. Reinforcement Learning:

Reinforcement learning is distinct in that it's centered around an agent learning to make decisions through interaction with an environment. The agent takes actions, and the environment provides feedback in the form of rewards or penalties. The goal is to learn a policy that maximizes the cumulative reward over time.

k-Nearest Neighbor

The k-Nearest Neighbors (k-NN) algorithm is a versatile machine learning approach widely employed for both classification and regression tasks. At its core, k-NN relies on the concept of similarity, seeking the 'k' nearestdata points from the training dataset to a given query point, typically measured using distance metrics like Euclideandistance. For classification, it assigns the majority class among these 'k' neighbors as the prediction for the query point. In regression, it computes the average (or weighted average) of the 'k' nearest neighbors' target values as the prediction. Notably, k-NN is valued for its simplicity and ease of implementation, making it a pragmatic choice when dealing with complex, nonlinear decision boundaries. However, it's sensitive to 'k' and the choice of distance metric, and its computational demands can grow with larger datasets [3]. Figure following illustrates how the K-Mean Nearest Neighbor Algorithm is currently visualized in action. [Figure 2].

ISBN 979-8-3503-5094-4

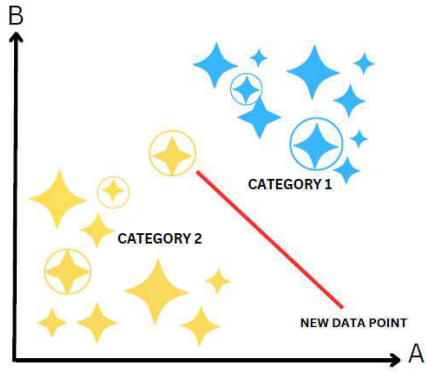


Figure 2: k- Nearest Neighbor

Support Vector Machine

Support Vector Machines (SVM) are a versatile machine learning method capable of handling both classification and regression problems. This technique involves defining a hyperplane as the decision boundary, particularly when dealing with objects from different classes that need to be separated.

It's important to note that these objects may not always be linearly separable, requiring the use of complex mathematical functions known as kernels to effectively segregate them.[5]

SVM's primary goal is to correctly classify objects basedon the examples provided in the training dataset. Someof its advantages include its ability to work with both semi- structured and structured data, its capacity to handle complex functions through appropriate kernel functions, andits inherent resistance to overfitting due to its generalization capabilities. Additionally, SVM can efficiently scale up to handle high-dimensional data and avoids getting trapped in local optima. However, there are certain disadvantages associated with SVM. Its performance tends to degrade with larger datasets due to increased training times. Selecting the right kernel function can be challenging, and SVM is less effective when dealing with noisy datasets. Furthermore, SVM doesn't readily provide probability estimates, and interpreting the final SVM model can be complex. When choosing between machine learning approaches like KNN, Random Forests, and SVM, Logistic Regression is often tried first, followed by Random forest if significant improvement is needed [3]. The support vector machine's operational visualization is displayed in the figure below [Figure-3].

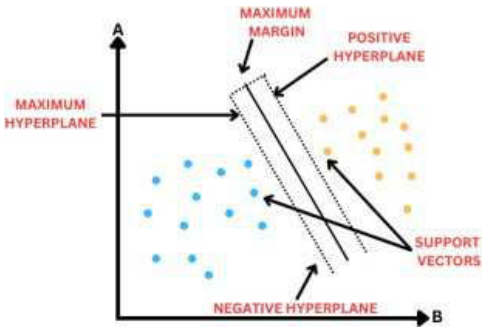


Figure 3: Support vector machine

Random Forest

Random Forest is a powerful supervised classification technique composed of multiple decision trees. In a decision tree, nodes represent tests, and branches display specific test results. It then builds classifications using these pathways. However, decision trees are sensitive to training data; even small changes can greatly affect the tree structure. Random Forest addresses this by using bagging. It randomly selects samples from the dataset to create multiple decision trees. Each tree becomes a class, and voting decides the best prediction within the model. This ensemble approach minimizes overfitting and enhances accuracy while maintaining model diversity [3]. The following figure displays the Random Forest Algorithm's operational visualization [Figure 4].

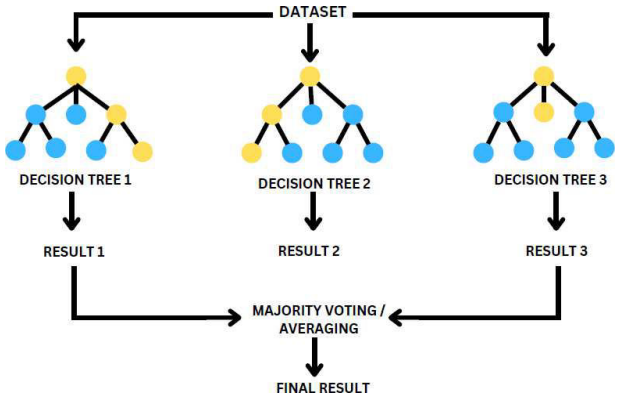


Figure 4: Random Forest

Model performance and challenges

k-NN is appreciated for its simplicity and easy implementation, making it a preferred option across diverse applications. However, its performance heavily relies on the selection of the 'k' parameter, determining the number of nearest neighbors for classification. The choice of 'k' significantly influences the model's accuracy and its ability to generalize to unseen data. Moreover, k-NN can face computational challenges, particularly with large datasets, due to the necessity of computing distances between the query point and all data points in the training set. Despite these challenges, k-NN remains popular for its intuitive approach and straightforward implementation.

SVM boasts several advantages, including resilience to overfitting and effectiveness in high-dimensional spaces. It excels in scenarios with complex decision boundaries where other algorithms struggle. However, SVM's effectiveness may diminish when dealing with noisy or overlapping data, as it relies on identifying a clear margin of separation between classes. The selection of an appropriate kernel function poses another challenge, as different kernels may perform differently based on the dataset's characteristics. Despite these challenges, SVM's capability to handle high-dimensional data and its resistance to overfitting make it invaluable for fall prediction and gait analysis tasks.

ISBN 979-8-3503-5094-4

Random Forest is renowned for its capacity to tackle complex datasets and combat overfitting through ensemble learning. By constructing multiple decision trees and aggregating their predictions, Random Forest enhances model robustness and generalization. Nevertheless, optimal performance often requires meticulous tuning of hyperparameters such as the number of trees in the forest and the maximum depth of each tree. Additionally, interpreting the outcomes of a Random Forest model can be challenging due to its ensemble nature. Despite these hurdles, Random Forest remains a favored choice for its ability to manage nonlinear relationships and capture feature interactions effectively.

C. System Overview

The process begins with data collection from MPU6050 sensors placed at the hip, knee, and ankle, facilitated by the ESP 8266 module. This raw data is then preprocessed to eliminate noise and irrelevant information, ensuring its reliability. Following this, crucial features related to gait patterns and motion angles are extracted from the refined data [4]. MachineLearning algorithms such as K-Nearest Neighbors,Support Vector Machines, and Random Forest are employed to analyze these features, acting as predictive 'detectives' [7]. The models are rigorously evaluated using metrics like accuracy and sensitivity to gauge theireffectiveness [4]. This project presents a comprehensive approach to understanding and predicting fall risk, merging data science and embedded systems seamlessly [8]. Continuous refinement and fine-tuning of models based on evaluation results will lead to more accurate and reliable fall predictions. The following figure illustrates the fall prediction process [Figure 5].

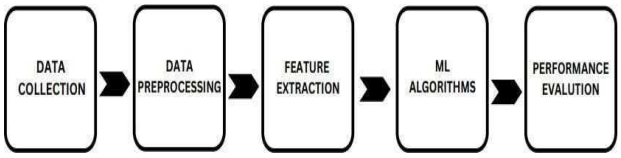


Figure 5: Procedure for Fall prediction

The proposed system for sensor-based gait analysis and fall prediction in adults presents a pioneering approach by seamlessly integrating IoT devices, such as ESP8266 sensors and Node MCU modules, with a diverse range of machine learning algorithms, including K-Nearest Neighbors, Support Vector Machines, and Random Forests. This integration enables continuous monitoring and real-time analysis, significantly enhancing the system's predictive capabilities. Furthermore, the comprehensive sensor placement and calibration techniques ensure precise data collection from critical locations like the hip, knee, and ankle joints, crucial for reliable fall prediction and gait analysis. A key feature of the system is its real-time mobile application, offering a user-friendly interface for both regular users and healthcare professionals, facilitating early fall risk assessment and intervention. The interdisciplinary approach adopted in the system design ensures a holistic understanding of fall risk factors and allows for continuous refinement of predictive models based on evaluation outcomes. Ethical considerations regarding data privacy, security, and fairness are meticulously addressed throughout the development process, emphasizing transparent communication, informed consent, and robust protection mechanisms. With its potential to significantly enhance healthcare outcomes, particularly for the elderly population, by providing early fall risk assessment and prevention measures, the proposed system stands as a promising solution at the forefront of addressing the challenges of fall risk assessment in adults.

D. Mobile application

In conclusion, fusing Internet of Things (IoT) devices Creating a mobile application for sensor-based gait analysis using machine learning involves a methodical step-by-step process [7]. Begin by defining the application's objectives and scope, specifying the useof sensors such as ESP8266 for gait data collection [9]. Proceed to implement mechanisms for efficient data collection and preprocessing to ensure data accuracy [11].

The core of the application lies in integrating a machine learning model, like the K-Nearest Neighbors Algorithm (KNN), for gait analysis [5]. Train the model using labeled data and incorporate real-time analysis features into the mobile app [8]. Design a user-friendly and intuitive interface catering to both regular users and healthcare professionals to ensure a positive user experience [9].

Establish a connection between Arduino and the mobile application to display accelerometer (MPU6050) values [10]. Prioritize security and privacy measures to safeguard user data and uphold ethical standards [4].Conduct thorough testing across various devices to ensure reliability and address any potential issues [6].

The deployment phase involves releasing the application on app stores, with a commitment to ongoingupdates for bug fixes and feature improvements [4]. Provide comprehensive user documentation for effective usage and consider a strategic marketing plan,potentially involving collaboration with healthcare professionals for endorsement [8]. Throughout the entire development process, maintain a focus on ethical considerations to ensure responsible and inclusive use of the application in early fall risk assessment for older adults [5].

III. RESULT

The investigation produced promising outcomes, highlighting the efficacy of the integrated system in forecasting fall risks among the elderly. The hardware configuration, integrating ESP8266 sensors and Node MCU modules strategically positioned, exhibited superior data accuracy and real-time monitoring capabilities compared to previous iterations. Gait analysis, facilitated by MPU6050 sensors, enabled thorough data collection. Following preprocessing, essential features related to gait patterns and motion angles were unveiled. Machine learning algorithms, encompassing K-Nearest Neighbors, Support Vector Machines, and Random Forest, adeptly scrutinized these features, offering valuable insights into fall prediction. The figure-6.1 and figure-6.2 depicts the typical and anomalous outcomes for an app page used normally [Figure-6.1,6.2].

ISBN 979-8-3503-5094-4

Model assessments using metrics like accuracy and sensitivity underscored their proficiency in predicting fall risks. The interdisciplinary fusion of data science and embedded systems demonstrated a holistic comprehension of fall risk, emphasizing continuous refinement and adjustment based on evaluation outcomes for heightened precision. The mobile application, characterized by its user-friendly interface, effectively portrayed accelerometer values, ensuring a positive user experience. Robust security and privacy protocols were instituted to protect user data, and rigorous testing across diverse devices validated the system's reliability. The figure-7.1 and figure-7.2 displays the typical and anomalous outcomes from the admin app page [Figure-7.1,7.2].

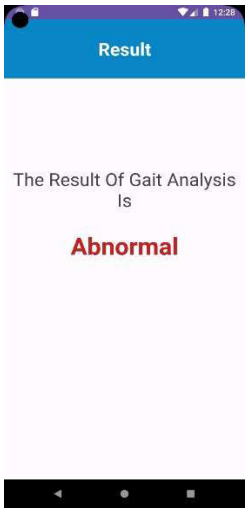


Figure 6.1. Abnormal result displays on normal usage app interface

Figure 6.2. Normal result displays on normal usage app interface

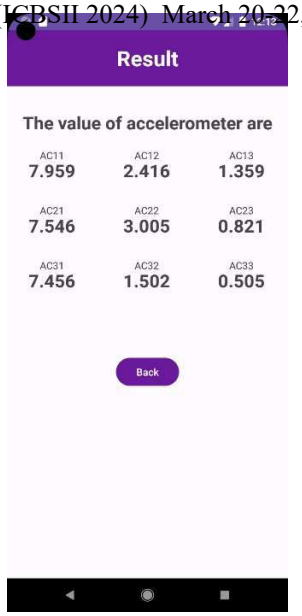


Figure 7.1: Abnormal value displays on the admin usage app interface

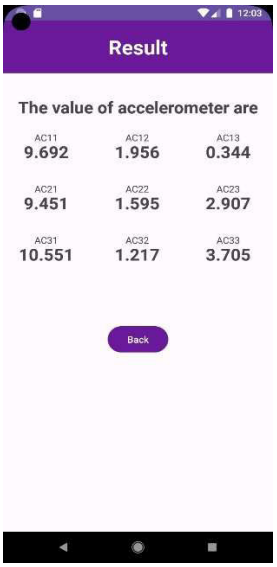


Figure 7.2: Normal value displays on the admin usage app interface

IV.CONCLUSION

In conclusion, fusing Internet of Things (IoT) devices with machine learning for elderly fall prediction holds promise for healthcare enhancement. This study introduces a comprehensive system using ESP8266 sensors, Node MCU modules, and various machine learning techniques. Placed strategically, sensors enable precise gait analysisand real-time monitoring, with predictive power boosted by supervised, hybrid, and ensemble machine learning methods. The system emphasizes continuous monitoringand real-time data analysis, complemented by a user- friendly mobile application providing insights and timely notifications. Overall, the study highlights the importance of quick, data-driven decision-making and continuous observation in fall prevention.The integration of IoT and machine learning offerspotential to elevate elderly living standards, crucial for improving healthcare outcomes with the growing global senior population. Advanced fall prediction systems areessential to address healthcare demands effectively.

ISBN 979-8-3503-5094-4

The proposed fall prediction and prevention system faces several limitations. Firstly, precise sensor placement and calibration are crucial for accurate data collection, but challenges in achieving this may compromise the system's effectiveness. Secondly, ensuring data privacy and security is vital due to continuous monitoring, demanding robust measures to protect sensitive information from unauthorized access. Additionally, generalizing findings beyond the elderly population could be complex due to diverse mobility patterns and medical conditions, necessitating further validation and adaptation for broader applicability. Furthermore, while promising in controlled environments, scalability and real-world implementation pose challenges like large-scale deployment and integration into existing healthcare systems, highlighting the importance of addressing logistical and practical concerns.

Future avenues for enhancing the fall prediction system involve conducting longitudinal studies to assess its long-term effectiveness, tracking individuals over extended periods to capture nuanced changes in gait patterns. Integration of additional sensors, such as pressure sensors or depth cameras, could bolster the system's capabilities, enabling more accurate fall detection and comprehensive gait analysis. Tailoring risk assessment models to individuals' unique characteristics and implementing real-time intervention strategies, like automated alerts, would augment the system's effectiveness in preventing falls. Furthermore, validation across diverse real-world settings, with collaboration from healthcare providers and caregivers, is crucial for evaluating the system's applicability and promoting its adoption in clinical practice.The fusion of sensor-based gait analysis and machine learning for fall risk assessment poses ethical challenges related to data privacy and security, demanding robust protection mechanisms. Transparent communication and informed consent are vital to ensure individuals comprehend the data collection purpose and their rights. To mitigate biases and ensure fairness, algorithms must be deployed equitably across diverse demographic groups for accurate predictions. It's essential to strike a balance between safety and autonomy to avoid excessive surveillance and uphold individual independence. Continuous monitoring, regular audits, and user feedback collection are imperative to maintain ethical standards.

REFERENCE

[1] Daniela Tarnita. (2016) 'Wearable sensors used for human gait analysis', Romanian Journal of Morphology & Embryology, Vol.57,No.2 pp 378-379.

[2] Bet, P., Castro, P. C., & Ponti, M. A. (2019). Fall detection and fall risk assessment in older person using wearable sensors: A systematic review. International journal of medical informatics, 130, 103946.

[3] Steinbach, M., & Tan, P. N. (2009). kNN: k-nearest neighbors. The top ten algorithms in data mining, 151-162.

[4] Arunkumar, P., Karthik, S., Shanthini, J., Karthik, S., and Karthikeyan, N (2021) 'Interpretation of Gait Supervising Mechanism Using Sensor Integrated Makeshift and Analysing Pattern by K-Means Clustering Algorithm', Journal of Medical Imaging and Health Informatics,Vol.11, No.10, pp 2599-2606.

[5] Jørgensen, V., Forslund, E. B., Opheim, A., Franzén, E., Wahman, K., Hultling, C., & Roaldsen, K. S. (2017). Falls and fear of falling predict future falls and related injuries in ambulatory individuals with spinal cord injury: a longitudinal observational study. Journal of physiotherapy, 63(2), 108-113.

[6] Gao, J., Gu, P., Ren, Q., Zhang, J. and Song, X., (2012) 'Abnormal gait recognition algorithm based on LSTM-CNN fusion network', IEEE Journal of Translation Engineering in Health and Medicine,Vol.7, pp.163180–163189.

[7] Yang, M., Zheng, H., Wang, H., McClean, S. and Newell, D., (2012) IGAIT: 'An interactive accelerometer-based gait analysis system. Computer Methods and Programs in Biomedicine' pp.715–723.

[8] Kim, K., Yun, G., Park, S. K., & Kim, D. H. (2019, July). Fall detection for the elderly based on 3-axis accelerometer and depth sensor fusion with random forest classifier. In 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) (pp. 4611-4614). IEEE.

[9] A.H.A., Aruwa., Al Junid., Razak., M.M.M. Arava., S.A.M., Idros, M.F.M., Halim, N.Z. and Khairuddin, N., (2017) 'Wireless integrated gait analysis system forheel-strike and toe-off measurement', Journal of Telecommunication, Electronic and Computer Engineering, vol.9, No.2-8, pp.128–131.

[10] Agur, A. M , Dalley, A. F., & Luca Moore. (1992) 'Lower limb',Clinically oriented anatomy, Vol 3, pp 403-468.

[11] Auvinet, E., Multon, F., Stasse, O., & Lavaste, F. (2014). 'Accelerometer and Gyroscope Based Gait Analysis Using Spectral Analysis of Patients with Osteoarthritis of the Knee', The journal of physical therapy science, Vol.26, No.7, pp 998-1001

ISBN 979-8-3503-5094-4

ISBN 979-8-3503-5094-4

Feeding Aid Stabilizer for People with Essential Tremors Caused by Thalamic Haemorrhage

Allwyn Gnanadas A
Department of Biomedical Engineering
KPR Institute of Engineering and Technology
Coimbatore,India
allwyn3t6@gmail.com

Hamshavarthini S
Department of Biomedical Engineering
KPR Institute of Engineering and Technology
Coimbatore,India
hamshavarthini13112002@gmail.com

Rithanya K V
Department of Biomedical Engineering
KPR Institute of Engineering and Technology
Coimbatore,India
rithanyakv@gmail.com

Kaviya P
Department of Biomedical Engineering
KPR Institute of Engineering and Technology
Coimbatore,India
kaviyakavi55667@gmail.com

Isabella S
Department of Biomedical Engineering
KPR Institute of Engineering and Technology
Coimbatore,India
isabellarosy61@gmail.com

***Abstract*—Essential tremors, affecting approximately 5% of the population, pose significant challenges for the elderly, limiting their daily activities. This neurodegenerative disorder primarily impacts motor movements, with familial factors playing a prominent role. Common symptoms include tremors, memory loss, and rigidity, particularly affecting activities like eating. Affected individuals, especially the elderly people often struggle with food consumption, requiring extended time to complete daily tasks. A few people accept and adapt to these challenges, existing aids, such as heavier gloves and specialized utensils, often compromise user comfort owing to their various backgrounds. To reduce, this project introduces a novel stabilized feeding aid specifically designed to address the difficulties faced by individuals with hand tremors. With a microcontroller that governs a brushless gimbal motor, the system effectively mitigates tremors. A triple accelerometer setup captures motion data, which is fused and utilized as input for the controller, which operates the motor and enhances the smooth operation of the feeding aid. This gyro-stabilized motor's end is connected to a customized 3D-designed adapter for optimal performance. Preliminary results demonstrate the effectiveness of the stabilized feeding aid, exhibiting smoother, tremor-free motion even under varying loads. This innovative approach holds promise for improving the daily lives of those suffering with essential tremors.**

***Index Terms*—Essential tremors, stabilized feeding aid, accelerometer fusing, microcontroller, gyro-stabilized motor**

I. Introduction

Essential Tremor is recognized as one of the most prevalent motor disorders. Task-activated functional imaging, utilizing electromyography (EMG), reveals widespread activations in the motor control network and beyond in Essential Tremor patients, while coherence analysis identifies distinct brain areas contributing to tremor, including the motor cortex, prefrontal cortex, thalamus, cerebellum, and brainstem as shown in the Fig. 1. While the precise mechanism, progression, and associated features of the disease are still yet to be unveiled, an evolving body of evidence challenges the perception of essential tremors as a benign or inconsequential condition [1]. Addressing these issues could only be done by either surgery or through proper assistive aids. Despite recent advancements in treating tremor disorders, a significant gap persists, leaving many patients with functional disabilities crucial for independent living [2] Assistive devices emerge as a promising solution, bridging the gap between symptom burden and current therapies. While not widely considered in therapeutic guidelines, assistive technologies offer a unique opportunity as adjunctive interventions, utilizing new portable technologies at a lower cost. A vibration-damping robotic eating device with two degrees of freedom, employing an active controller structure to absorb vibrations and maintain spoon stability during eating on Simulations performed significantly well and data from tremor patients indicate promising results in stabilizing [3]. To eliminate hand tremors by utilizing inertial measurement units that measure hand tremors and implement targeted tremor control. Given the impact of on daily activities, the noninvasive nature of the device can perform as a practical and innovative solution for enhancing patients' abilities [4]. Hand tremors, often exacerbated during moments of stress or fatigue, can result from the degeneration of dopamine producing neurons in the brain's substantia nigra [9], it is shown in the Fig. 2. This loss of dopamine disrupts the intricate balance between inhibitory and excitatory signals within the basal ganglia, leading to uncoordinated muscle contractions and characteristic tremors. While tremors can affect various body parts, the upper extremities, especially the hands, are particularly susceptible, making everyday tasks that require fine motor control, such as eating, a formidable challenge. Eating, a seemingly effortless action for most becomes an ordeal for individuals with hand tremors. Comparison of cerebellum images obtained using various hardware and imaging conditions. Includes in vivo 1.5T MR image (A) and in vivo 3T MR image (B) of a patient with ET.

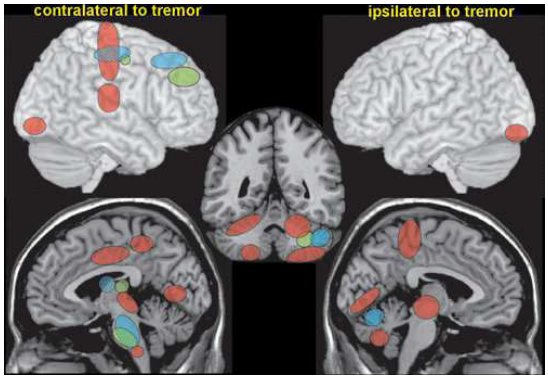


Fig. 1. Activations in brain in patients affected with essential tremors

ISBN 979-8-3503-5094-4

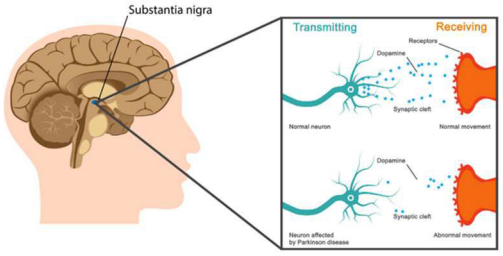


Fig. 2. Dopamine depletion diagram

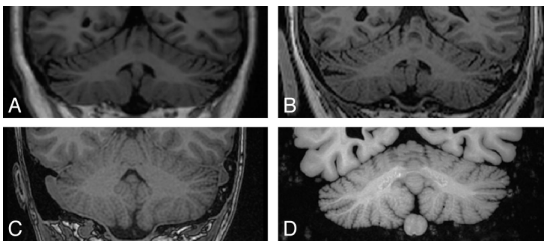


Fig. 3. Comparison of cerebellum images, including in vivo 1.5T and 3T MR images of a patient with ET, postmortem high-resolution 1.5T MR image acquired in situ, and 1.5T MR image

Additionally, postmortem high-resolution 1.5T MR image (C) acquired in situ and 1.5T MR image (D) of a formaldehyde-fixed specimen obtained ex-situ are presented. The direct comparison reveals morphologic details in the cerebellar cortex and demonstrates the potential of different scans (and scanners) to depict geometric complexity and contrast between the cerebellar cortex and white matter [8]. Image details for each modality are specified as shown in the Fig. 3.

The act of conveying food from plate to mouth becomes a painstaking endeavour fraught with difficulties. The tremors, often unpredictable in intensity and frequency, can cause food spillage, misdirected bites, and an overall struggle to bring utensils to the mouth. In the realm of technological solutions, a novel approach emerges with a stabilized feeding aid designed to address the challenges faced by individuals with hand tremors during meals [5], [7], [11]. This innovation aims to alleviate the burden and enhance overall well-being by integrating modern sensor and control technologies. The feeding aid employs components such as the MPU6050 sensor, 2805 Gimbal brushless motor, and Raspberry Pi Pico W microcontroller to counteract the disruptive effects of tremors. The MPU6050 sensor plays a foundational role in this system [6]. As a versatile component capable of measuring both acceleration and gyroscopic movements, it serves as the mechanism for detecting the user's hand tremors. By precisely tracking these tremors, the system can anticipate involuntary movements and engage the 2805 Gimbal brushless motor to counteract the disruptive effects [10]. The 2805 Gimbal brushless motor, in conjunction with the sensor data, provides controlled movements to stabilize the feeding aid. This collaborative effort is orchestrated by the Raspberry

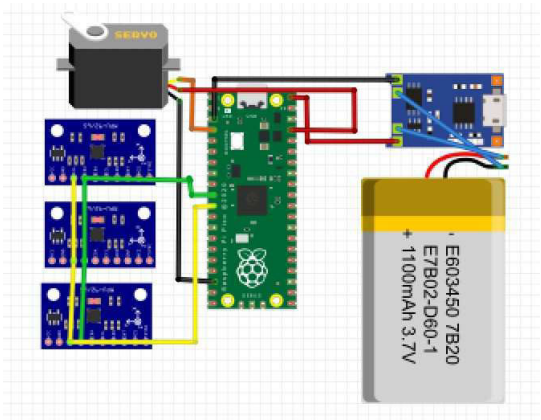


Fig. 4. Stabilized feeding system

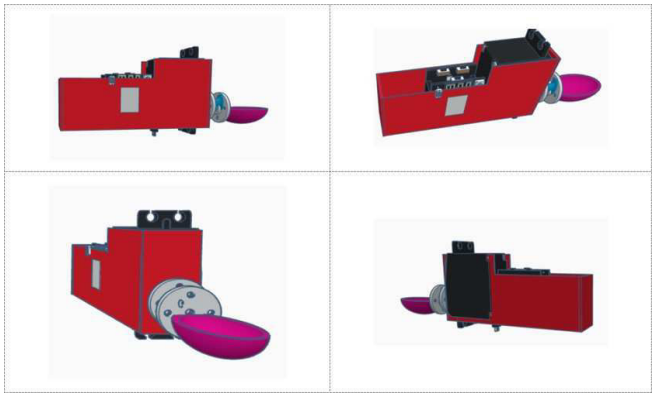


Fig. 5. 3D Designed stabilized feeding aid

Pi Pico W microcontroller, functioning as the brain of the system [12]. In real-time, it adjusts the utensil's position to compensate for the detected tremors, enabling a more seamless and controlled eating experience for individuals grappling with hand tremors. As a comprehensive solution, the stabilized feeding aid seeks to restore a measure of independence to individuals facing the challenges of hand tremors during meals [5], [7], [11]. By addressing the disruptions caused by tremors, this technological innovation aims to enhance the overall well-being of users, offering them a renewed sense of autonomy in performing essential daily activities.

II. Methodology

The methodology of the system comprises the sensing system, the controller and the gyro stabilization setup all encapsulated in a customized 3D setup as shown in the Fig. 4.

A. 3D design

The design is made to make sure that each of the systems is a custom-defined structure that helps the patient to be at complete ease. A picture of the design in a different view is shown in the Fig. 5. This design undergoes a series of iterations, incorporating valuable insights from user feedback and rigorous testing. The primary objective is to

ISBN 979-8-3503-5094-4

create a feeding aid that not only addresses mechanical stability but also prioritizes user comfort and usability.

B. Multisensing

The MPU6050 is a popular Inertial Measurement Unit (IMU) that combines a 3-axis accelerometer and a 3-axis gyroscope in a single chip. To fuse data from these sensors, sensor fusion algorithms are used to obtain a more accurate and stable estimation of the device's orientation and motion. The process of sensor fusion involves combining data from multiple sensors to improve the overall accuracy and reliability of the measurements. Initially, the sensor value is calibrated and then it is subjected to Madgwick filter filtering and the equation from 1 to 6 sequences the steps involved in the algorithm that uses gradient descent to minimize errors in orientation estimates

Initializing Variables:

q: Quaternion representing the orientation

β: Filter gain (adjustable parameter)

Algorithm (for each time step):

gyro: Gyroscope data

accel: Accelerometer data

Calculation of the gyroscope measurement error:

$$\text{gyro\_error} = \text{gyro} - q \otimes \text{conj}(\text{reference\_vector}) \otimes q \quad (1)$$

Update the quaternion using the gyroscope data:

$$q = q + 0.5 \times q \otimes \text{gyro} \times \text{dt} \quad (2)$$

Normalization of the quaternion to prevent drift:

$$q = \text{normalize}(q) \quad (3)$$

Calculation of the accelerometer measurement error:

$$\text{accel\_est} = q \otimes \text{conj}(q) \otimes \text{accel} \otimes \text{conj}(q)$$
$$\text{accel\_error} = \text{accel} - \text{accel\_est} \quad (4)$$

Update the quaternion using the accelerometer data:

$$q = q + \beta \times \text{accel\_error} \times \text{accel} \times \text{dt} \quad (5)$$

Normalization of the quaternion:

$$q = \text{normalize}(q) \quad (6)$$

The filter is very crucial in determining the exact direction very accurately. A single sensor can yield results however, it is insufficient to determine the exact direction which is evident from the figure as shown in the Fig. 6.

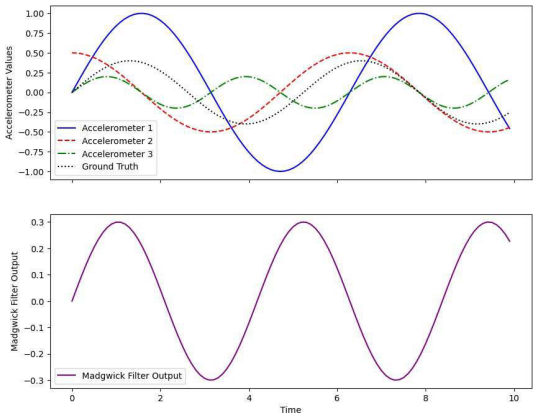


Fig. 6. Sensor fusion using Madgwick filter

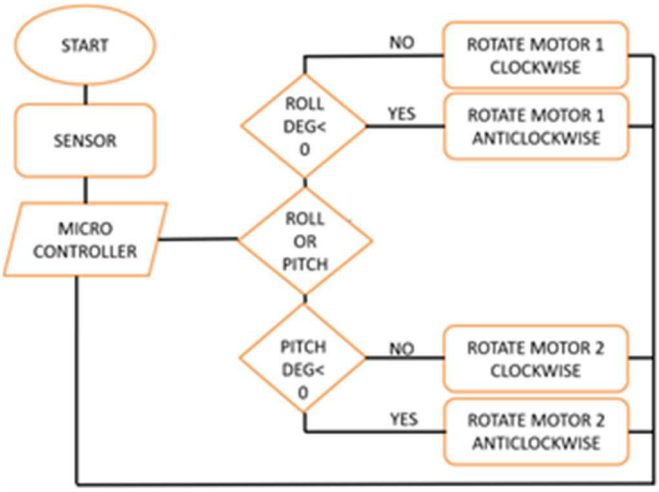


Fig. 7. The logic behind Gyro stabilized system

C. Gyro stabilization system

The Gyro Stabilization System utilizes a Raspberry Pi Pico, MPU6050 sensor, and brushless gimbal servo to create a stable platform. Interfacing the MPU6050, the Madgwick filter for sensor fusion was implemented, providing accurate orientation data. The brushless gimbal servo is controlled by PID algorithms based on the fused sensor information, ensuring precise adjustments for stabilization. The Raspberry Pi Pico runs MicroPython, enabling seamless integration of sensor data and control logic. Initial testing demonstrated successful stabilization, followed by fine-tuning for optimal performance. Safety measures, calibration routines, and optional user interfaces were implemented. The documentation provides a concise overview of the project's components, implementation steps, and additional considerations, laying the foundation for a robust gyro stabilization system. The logic behind the operation is explained in the Fig. 7.

ISBN 979-8-3503-5094-4

III. EXPERIMENTAL RESULTS

The integrated system developed for the stabilized feeding aid demonstrates significant advancements in addressing challenges faced by patients with hand tremors. Leveraging the Raspberry Pi Pico, MPU 6050 sensor, and 2805 Gimbal Brushless motor, the solution offers a reliable mechanism for stabilizing feeding utensils, enhancing the dining experience for individuals with motor control difficulties. Precise detection of Roll and Pitch Angles enables accurate adjustments of the spoon's orientation, effectively compensating for hand tremors and facilitating independent feeding.

The motor angle data in the table reveals consistent and responsive adjustments, emphasizing the system's robust performance across diverse tilt angles.

TABLE I
MOTOR ANGLE DATA

Roll Angle (degrees)	Pitch Angle (degrees)	Motor x Angle (degrees)	Motor y Angle (degrees)
36	54	72	108
45	45	90	90
45	90	90	180
45	81	90	162
54	45	108	90
63	72	126	144
81	90	162	180
90	45	180	90
99	81	198	162
108	72	216	144
108	90	216	180
117	99	234	198
117	90	234	180
126	90	252	180
126	54	252	108
135	90	270	180
135	81	270	162
144	81	288	162
144	90	288	180
144	72	288	144
144	63	288	126
144	99	288	198
153	108	306	216
153	90	306	180
162	135	324	270
162	72	324	144

The 3D scatter plot reveals a positive correlation between Roll Angle and both Motor x Angle and Motor y Angle as shown in the Table I. The consistent variation patterns, spatial alignment, and clustered distribution suggest a synchronized relationship. Peaks and troughs in Roll Angle correspond to analogous changes in the motor angles, reinforcing their close association. The linear trends further emphasize the agreement between Roll Angle and the motor angles, indicating that variations in Roll Angle are consistently reflected in both Motor x Angle and Motor y Angle.

IV. CONCLUSION

The integrated system, encompassing the sensing system, controller, and gyro stabilization setup within a custom 3D design as shown in the Fig. 4 is meticulously

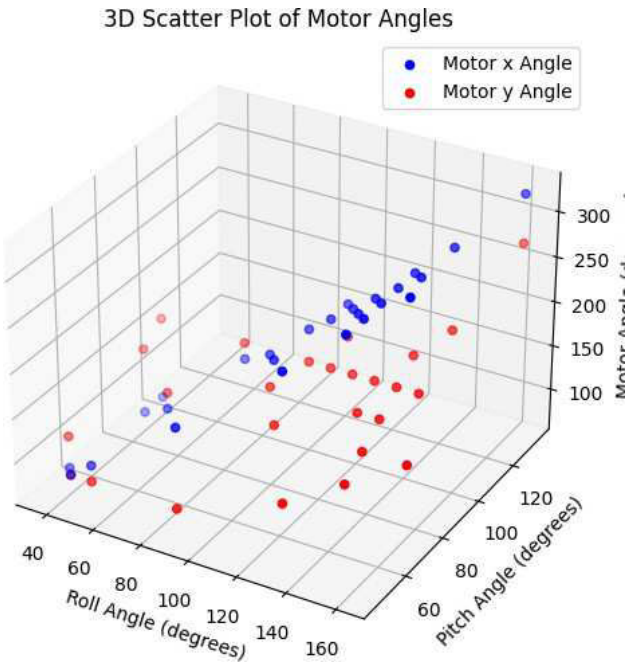


Fig. 8. The scatter plot

engineered for enhanced patient comfort. The iterative 3D design prioritizes mechanical stability, incorporating user feedback and rigorous testing shown in the Fig. 5, Multisensing using the MPU6050 IMU and Madgwick filter ensures precise orientation estimation, as showcased in the sensor fusion plot shown in the Fig. 6. The Gyro Stabilization System, employing Raspberry Pi Pico, MPU6050, and a brushless gimbal servo, exhibits successful stabilization as shown in the Fig. 7. Experimental results demonstrate the system's efficacy in stabilizing feeding utensils for individuals with hand tremors, supported by consistent motor angle adjustments shown in Table I and a positive correlation observed in the 3D scatter plot shown in Fig. 8. The system's holistic approach prioritizes both mechanical stability and user comfort, showcasing its potential impact on improving the quality of life for individuals with motor control difficulties.

REFERENCES

[1] Welton, Thomas, Francisco Cardoso, Jonathan A. Carr, Ling-Ling Chan, Gu¨nther Deuschl, Joseph Jankovic, and Eng-King Tan. "Essential tremor." Nature Reviews Disease Primers 7, no. 1 (2021): 83.

[2] Bhidayasiri, Roongroj, Suppata Maytharakcheep, Saisamorn Phumphid, and Walter Maetzler. "Improving functional disability in patients with tremor: A clinical perspective of the efficacies, considerations, and challenges of assistive technology." Journal of the Neurological Sciences 435 (2022): 120197.

[3] Tas¸ar, B., Tatar, A. B., Tanyıldızı, A. K., & Yakut, O. (2023). FiMec tremor stabilization spoon: design and active stabilization control of two DoF robotic eating devices for hand tremor patients. *Medical & Biological Engineering & Computing, 61*(10), 2757-2768.

[4] Talaei, Fereshteh, and Seyed Mohammad Kargar. "Design and fabrication of a device for reducing hand tremor in parkinson patients during eating." Journal of Medical Signals and Sensors 13, no. 1 (2023): 21.

[5] Altan, A., & Hacioglu, R. (2020). Model predictive control of three-axis gimbal system mounted on UAV for real-time target tracking under external disturbances. *Mechanical Systems and Signal Processing, 138*, 106548.

ISBN 979-8-3503-5094-4

[6] Fang, Y., Wang, H., Wang, X., & Luo, Y. (2020, December). Automated assembly system for miniature bolted joints during the gimbal assembly. In *IEEE 3rd World Conference on Mechanical Engineering and Intelligent Manufacturing (WCMEIM)* (pp. 381-384).

[7] Guo, J., Geng, Y., Wu, B., & Kong, X. (2020). Manifold path selection steering law for pyramid-type single-gimbal control moment gyros. *Aerospace Science and Technology, 106*, 106207.

[8] Klaming, R., & Annese, J. (2014). Functional Anatomy of Essential Tremor: Lessons from Neuroimaging. *American Journal of Neuroradiology, 35*(8), 1450-1457.

[9] Iizuka, Y., Hiraoka, M., Kokubo, M., Sakamoto, T., Karasawa, K., Murofushi, K., ... & Mizowaki, T. (2023). Dynamic tumor-tracking stereotactic body radiotherapy with real-time monitoring of liver tumors using a gimbal-mounted linac: A multi-institutional phase II study. *Clinical and Translational Radiation Oncology, 39*, 100591.

[10] Mughal, H., Javed, A. R., Rizwan, M., Almadhor, A. S., & Kryvinska, N. (2022). Parkinson's disease management via wearable sensors: a systematic review. *IEEE Access, 10*, 35219-35237.

[11] Pan, S., Xu, Z., Lu, M., Chen, L., Liang, Z., & Zhang, J. (2023). The coupling analysis for the gimbal servo system of a control moment gyroscope considering the influence of a flexible vibration isolator. *ISA transactions, 137*, 601-614.

[12] Ren, W., Qiao, Q., Nie, K., & Mao, Y. (2019). Robust DOBC for stabilization loop of a two-axes gimbal system. *IEEE Access, 7*, 110554-110562.

ISBN 979-8-3503-5094-4

Tenth International Conference on Biosignals, Images and Instrumentation, SSNCE, Chennai, March 20 – 22, 2024

TRACK 8

Hybrid Analysis for Detection of Alzheimer Disease Using Computer Aided Diagnostic Tool

N. Kiran Kumar
Department of ECE
CEG Campus, Anna University
Chennai-600025, India
kiran24052001@gmail.com

C. Kaushik Viknesh
Department of ECE
CEG Campus, Anna University
Chennai-600025, India
ecekv12@gmail.com

***Abstract*— Brain image analysis, a significant field in clinical research, is employed to detect brain diseases. Alzheimer's disease is an irreversible, progressive neurological disorder characterized by the loss of neurotransmitters, leading to cognitive and mental decline. Detecting Alzheimer's disease early can facilitate appropriate treatment and prevent brain tissue damage. It is one of the most common causes of dementia, resulting in significant impairments in cognitive functions like memory, reasoning, and social skills, affecting daily life. This neurodegenerative form of dementia advances slowly over time, with decreased brain activity and blood flow being primary contributors to Alzheimer's. In the initial stages of the disease, rapid tissue loss in the hippocampus is observed, which is associated with functional impairment in other brain regions. Consequently, there is extensive research on segmenting clinical images, with various methods, including the Gaussian Mixture Model (GMM) segmentation technique, being employed for this purpose. Gaussian Mixture Models (GMM) can effectively separate the brain's white and gray matter. By extracting features from brain images, these characteristics can then be utilized to classify Alzheimer's disease.**

***Keywords*— Gaussian Mixture Model, Alzheimer's disease, Dementia, Cognitive Normal, Early Mild Cognitive Impairment, Support Vector Machine.**

I. INTRODUCTION

Alzheimer's infection is an ever-evolving neurodegenerative condition that principally influences the cerebrum and causes cognitive decline over the long haul. Language inconveniences, disarray (counting an inclination to get lost successfully), outlook swings, a shortfall of need, self-negligence, and social issues can all be indications of state of the art Alzheimer's sickness [1]. Around 24 million people by and large are impacted by it. Practically 33% of adults more than 85 and one of each and every ten people more than 65 have the issue. Figure 1a and Figure 1b shows the qualification between customary brain and Alzheimer's disease [2].

1(a). Normal Brain **1(b). Alzheimer's Disease**

Fig. 1(a, b). Normal Brain vs Alzheimer's disease

The signs and symptoms of Alzheimer's disease (AD) vary depending on the stage of the disease. In most cases, a steady decline in all, most, or some of the following are symptoms of AD: i). Memory. ii). Understanding and the executives of testing assignments. iii). Language. iv). a comprehension of how space and visual form interact. v). Character and conduct [3]. Alzheimer's illness is welcomed on by a distorted protein gathering in the mind. Both amyloid protein and tau protein collect in the cerebrum, which brings about the passing of synapses. More than 100 billion nerve and other types of cells make up the human brain. Every one of the interchanges fundamental for activities like reasoning, getting the hang of, recollecting, and arranging are done by the nerve cells on the whole. As per researchers, amyloid protein gathers into greater totals known as plaques inside your synapses. The plaques and tangles discourage nerve cell correspondence, which prevents the nerve cells from carrying out their roles. The persistent and progressive loss of nerve cells is what leads to the symptoms of Alzheimer's disease [4]. In view of side effects, Alzheimer's sickness associations and clinical experts utilize an assortment of phrasing to portray the disease's stages. The symptoms of AD worsen over time, despite the fact that the stages have different names. i). CognitiveNormal (CN) ii). Early Mild Cognitive Impairment (EMCI) iii). Mild cognitive impairment (MCI) iv). Late Mild Cognitive Impairment (LMCI) v). Alzheimer's disease (AD) [5]. Cognitive Normal (CN) has mental working is considered to be typical, an individual is supposed to be in this stage. Even though a person with normal cognitive functioning may occasionally forget things or have difficulty with difficult tasks, they are still able to lead normal lives without any issues. Early Mild Cognitive Impairment (EMCI) has individual might have some mental deterioration during the EMCI stage, however influencing their regular activities isn't sufficiently extreme. Despite the fact that they might have issues reviewing names or words, they are in any case fit for completing troublesome exercises and holding their freedom. Neglect, issues reviewing words or finding the appropriate ones, and inconvenience recalling late occasions are possible indications of EMCI. Mild Cognitive Impairment (MCI) has the cognitive decline is more pronounced than it was during the EMCI stage, it is not severe enough to hinder the individual's ability to perform day-to-day activities. Since they may struggle with more challenging tasks like managing their finances or driving, they would need assistance. MCI may be indicated by more obvious memory impairment, difficulty planning and organizing work, and difficulty making decisions. Late Mild Cognitive Impairment (LMCI) has the LMCI stage, an individual's mental hindrance is more articulated than in MCI, yet it isn't yet sufficiently serious to meet the necessities for an Alzheimer's Sickness finding. They could need more noteworthy support with day to day obligations like getting dressed or making dinners since they think that they are more

ISBN 979-8-3503-5094-4

troublesome. More serious memory issues, issues with day to day undertakings like getting dressed and scrubbing down, and changes in state of mind and conduct are possible indications of LMCI. Alzheimer's disease (AD) has neurological condition known as Alzheimer's disease (AD) is characterized by progressive cognitive impairment that impairs daily activities. An individual might battle to review ongoing discussions or occasions in the beginning phases, however in the last option stages, they might lose their capacity to talk and need help with all errands of everyday living. Forgetfulness, difficulties with routine tasks, and difficulty finding words are all possible early symptoms of AD. In the later stages, side effects could incorporate gulping issues, incontinence, and a deficiency of discourse abilities [6]. When diagnosing Alzheimer's disease, healthcare professionals employ a variety of approaches. This is on the grounds that there are many circumstances that can be taken for thought. Alzheimer's disease cannot be diagnosed with a single test; rather, it typically requires a combination of approaches. The following are some of the typical Alzheimer's disease diagnostic techniques: Clinical history and actual assessment are the essential conclusion that should be possible overall assessment, the specialist will get some information about their side effects, clinical history, and family ancestry. Additionally, they will conduct a physical examination to look for any indications of neurological issues [7]. Cognitive tests to assess the patient's memory, thinking, and abilities to think, the specialist might give mental tests. Imaging exams has anomalies in the cerebrum that are related with Alzheimer's sickness, like decay of specific mind districts or the aggregation of amyloid plaques, can be related to the assistance of positron emanation tomography (PET) and attractive reverberation imaging (X-ray) filters. Blood tests should be possible to preclude different sicknesses, similar to thyroid tissues or lacks of nutrient that could cause comparative side effects. Genetic testing has qualities connected to an expanded gamble of fostering Alzheimer's illness can be recognized through hereditary testing. Cerebrospinal fluid (CSF) analysis has sample of the fluid that surrounds the brain and spinal cord is examined. Changes in the measures of explicit proteins connected to Alzheimer's illness can be found through CSF examination [8]. A common imaging test for Alzheimer's disease is an MRI (Magnetic Resonance Imaging). An easy imaging procedure uses solid areas for a field and radio waves to make organized photos of the frontal cortex. X-beam can perceive changes in the psyche that are connected with Alzheimer's sickness, similar to shrinkage of explicit locale of the frontal cortex, improvement of the ventricles (the fluid consumed spaces in the brain), and changes in the plan of the hippocampus (a region of the frontal cortex that is huge for memory). X-ray can be utilized to evaluate blood stream and metabolic action in the mind, as well as giving data about the cerebrum's construction. The utilitarian changes in Alzheimer's sickness patients' cerebrums that happen because of this can be helpful data. Examiners have made explicit X-beam shows, similar to dispersal tensor imaging (DTI) and helpful X-beam (fMRI), to all the more promptly understand the hidden and reasonable changes in the frontal cortex related with Alzheimer's sickness. These techniques can give substantially more quick and dirty information about the movements that occur in the frontal cortex with Alzheimer's ailment. In general, X-rays are important tools for Alzheimer's patients and their families. It can give huge information about the plan and ability of the frontal cortex, helping with chipping away at the precision of finding and screen the development of the ailment after some time [9]. Alzheimer's disease is incurable and incurable. Holding regular working for some time by beginning Alzheimer's treatment quickly might be plausible. Imaging tests are utilized in Alzheimer's sickness since they can give nitty gritty data about the design, capability, and digestion of the mind. These tests can assist with distinguishing changes in the mind that are related with Alzheimer's illness, which can support determination and following sickness movement. It is a useful asset in the examination of clinical imaging information for Alzheimer's sickness and it can possibly further develop finding, therapy, and comprehension of the illness [10].

II. MATERIALS AND METHODS

A. Image Processing:

To remove helpful data from an image, it should initially be changed over into a computerized organization and afterward exposed to various systems. The image processing system typically interprets all images as two-dimensional signals if particular signal processing methods are utilized. The accompanying three stages generally make up picture handling: Bringing in the image involving programming for picture catch. manipulating the image through analysis. Output, which could be a report based on image analysis or a modified image [11]. To deal with the electronic pictures, some item is required. The software used in this work is Python 3.9. Python is an irrefutable level, by and large valuable programming language that is translated. Utilitarian programming, viewpoint arranged programming (counting meta programming and metal items), object-situated programming, and organized writing computer programs are completely upheld by quite a few people of its highlights. This venture utilizes the device PyCharm to execute Python code. Python composing PC programs is done in PyCharm, a planned headway environment. It offers code assessment, a graphical debugger, a basic unit testing, association with variation control structures, and help with Django web improvement. PyCharm is made by Jet Brains, a Czech company. PyCharm is a specific Python Consolidated Improvement Environment (IDE) that offers a far reaching variety of urgent instruments for Python fashioners. These instruments are solidly planned to convey a wonderful environment for valuable Python, web, and data science improvement. The various inbuilt libraries used here are open-source Python module that capabilities with NumPy clusters is called Scikit-picture. Algorithms and utilities are implemented for use in applications in industry, education, and research. In any event, for people who are new to Python's biological system, this module is somewhat easy to utilize. This code is all around elegantly composed. Most of the capabilities are situated in the submodules, which are imported as the skimage bundle. A free bundle called PIL (Python Imaging bundle) for the Python programming language upgrades support for review, changing, and saving an extensive variety of picture document designs. Pillow, a fork of PIL that is supported by Python 3, is thankfully available, is simple to install, and is actively maintained. The library incorporates the crucial devices for handling pictures, like point tasks, separating with an underlying assortment of convolution pieces, and variety space transformations.

ISBN 979-8-3503-5094-4

Matplotlib is a cross-stage information perception and graphical outlining bundle for Python and NumPy, its mathematical extension. Thusly, it offers a feasible open-source option in contrast to MATLAB. The APIs (Application Programming Points of association) for matplotlib license engineers to integrate outlines into GUI applications. NumPy, one of the fundamental Python programming libraries, provides array support. Basically, a picture is a customary NumPy exhibit of pixels and pieces of information. The pixel upsides of an image may thusly be changed by using central NumPy tasks like cutting, veiling, and smart ordering. The picture can be loaded with Skimage and displayed with Matplotlib. OpenCV (Open Source PC Vision Library) is one of the most generally utilized PC vision application libraries. OpenCV-Python is the Python Programming interface for OpenCV. In light of the Python covering toward the front, OpenCV-Python is just difficult to make and send yet furthermore expedient by virtue of the establishment C/C++ code. An open-source, cross-stage structure called Straightforward ITK gives developers admittance to an extensive variety of programming instruments for picture investigation, which permit us to decide an image's qualities. An extensive variety of picture sifting, division, enlistment, and change capacities are accessible with Straightforward ITK. For undertakings including picture smoothing, edge recognition, picture enlistment, picture division, and element extraction, it has different underlying calculations and strategies. Radiomics is a part of clinical picture examination that spotlights on getting quantitative data from clinical pictures to portray growth qualities, decide the guess of an illness, and direct treatment decisions. The geological dispersion, shape, surface, and force of growths and encompassing tissues are completely caught by these qualities, alongside an extensive variety of other data. Because it has a large number of libraries and tools for each stage of the radiomics process, Python is a popular choice among academics and professionals in the field. The creation and utilization of radiomics calculations and applications is made simpler by these libraries' quick information handling, include extraction, measurable examination, and representation capacities.

B. *Image Acquisition:*

The ADNI (Alzheimer's Disease Neuroimaging Initiative) dataset is a massive, open-access database of brain imaging and other clinical data gathered from healthy controls and individuals with Alzheimer's disease or mild cognitive impairment. The data, which were gathered over a period of time, are intended to support research on the causes and progression of Alzheimer's disease [12]. Dark to white are the monochromatic tones that make up the grayscale range. Thus, a grayscale picture has no variety, simply various levels of dark. Grayscale representations are frequently utilized because they simplify the algorithm and require less effort. Grayscale pictures are particular from the slightest bit bi-apparent high contrast pictures, which are pictures with just two tones: high contrast (otherwise called bi-level or paired pictures) with regards to PC imaging. There are a few shades of in the middle of between in grayscale photographs. By measuring the intensity of light at each pixel in accordance with a particular weighted combination of frequencies, monochromatic images can be produced when only one frequency is collected.

It is fundamental for precise investigation of mind construction and capability; because tissue outside of the brain can hinder the analysis. Image quality is improved and noise is reduced with this method. There are various kinds of channels that can be utilized, like Gaussian, middle, and Wiener channels. In medical image analysis, the procedure known as skull stripping involves separating the skull from the rest of the brain in an MRI or CT scan image. The skull is eliminated to disengage the mind and reject the encompassing bone, which can impede the investigation of the cerebrum tissue. Otsu thresholding is a generally involved procedure for picture division, which is the most common way of partitioning a picture into various locales or items. The Otsu thresholding technique works out an ideal limit esteem that isolates the picture into two classes in view of the histogram of pixel forces. The calculation expands the change between the two classes, which relates to the partition of the objects of interest from the foundation.

C. *Block Diagram:*

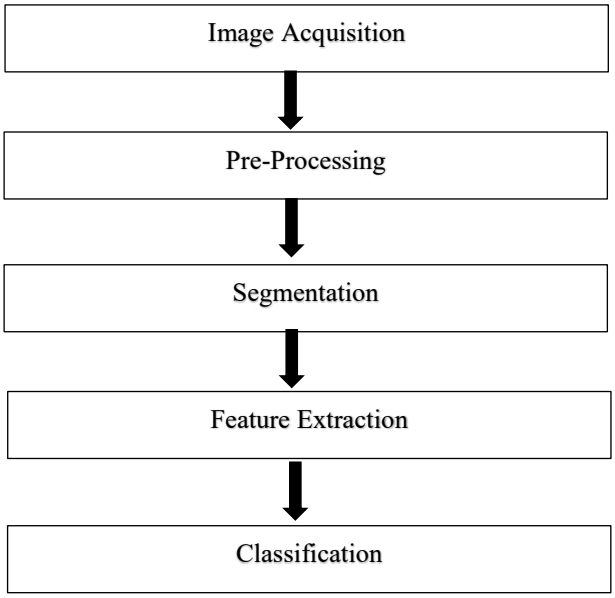


Figure 2. Block Diagram

D. *Image Acquisition:*

The popular method of image denoising known as the Non-local means (NLM) filter can be found in numerous fields, including medical imaging like MRI. The NLM filter is used in MRI to preserve structural details while reducing image noise. This is especially important in MRI, where various noise sources like patient motion and radio frequency interference can affect the images. The NLM filter works by estimating the value of each pixel by locating similar patches of pixels within the image and averaging their values. The similitude between patches is resolved utilizing a weighted normal of the pixel powers, where the loads depend on the distance between the patches and the likeness between their pixel values [13]. The NLM filter has the main advantage of effectively reducing noise without blurring the image's edges or fine details as shown in equation 1.

$$u_i^{NL} = \sum_{j\in\Omega} \omega(i,j) g(j) \ldots\ldots(1)$$

these weights are defined as shown in equation 2,

$$\omega(i,j) = \frac{1}{Z(i)} e^{-\frac{\|g(Nj)-g(Ni)\|_2^2}{h^2}} \ldots\ldots(2)$$

ISBN 979-8-3503-5094-4

Where z(i), normalization constant as shown in equation 3,

$$Z(i) = \sum_j e^{-\frac{||g(Nj)-g(Ni)||_2^2}{h^2}} \quad \ldots\ldots(3)$$

In plain English, division is the most common way of giving pixels marks. A unique label is given to each pixel or piece of a picture that belongs to the same category. In medical images, this method is used to distinguish between brain regions. It is fundamental for exact cerebrum construction and capability examination and can be utilized to distinguish locales of interest for additional investigation [14].

E. GMM Segmentation:

Gaussian Blend Model (GMM) division is a sort of solo machine learning calculation that can be utilized to portion cerebrum pictures into various tissue types, like dim matter, white matter, and cerebrospinal Fluid (CSF). In GMM division, the picture is displayed as a combination of Gaussian likelihood disseminations, where every circulation addresses an alternate tissue type. The calculation then, at that point, allocates every pixel in the picture to one of the tissue types in light of its likelihood of having a place with every circulation. Based on its probability of belonging to each distribution, the algorithm then assigns each image pixel to a tissue type [15]. The probability distribution function of d-dimensions Gaussian Distribution is defined as in equation 4.

$$\aleph(X|\mu,\Sigma\ ) = \frac{1}{(2\pi)^{D/2}|\Sigma|^{\frac{1}{2}}}\exp(-\frac{1}{2}(X-\mu)^T\Sigma^{-1}(X-\mu) \quad \ldots..(4)$$

Where μ = Mean

$\sum$ = Covariance Matrix of the Gaussian

d = Numbers of features in our dataset

x = Number of data points

In Alzheimer's disease, there is a gradual loss of white matter and gray matter in specific regions of the brain, particularly in areas involved in memory and cognitive functions. We propose a method for segmenting white matter and gray matter in brain MRI using the Gaussian Mixture Model (GMM) segmentation technique. TheGMM algorithm utilizes intensity information and spatial relationships in the MRI images to accurately classify the brain tissue into white matter and gray matter regions. Because the loss of white matter in the brain may be used to determine the stageof a person's condition, and the white matter area is being used for the feature extraction process. White matter and gray matter content in the brain [16]. The method of transforming crude information into mathematical elements that can be taken care of while keeping the data in the first dataset is known as component extraction. In computer vision and image processing, a feature is a piece of data that describes an image's content; generally, it demonstrates whether a specific piece of the image has specific qualities. A picture's features could be specific things like points, edges, or objects. The objective of element extraction is to change the sectioned picture information into a bunch of highlights that can be utilized for additional investigation, like item acknowledgment or grouping. GLRLM represents Dark Level Run Length Grid, which is a surface component extraction technique utilized in picture handling. In GLRLM, an image is divided into smaller regions, and within each region, the number of pixels with the same gray-level intensity that follow each other (also called "runs") is counted. These counts are then coordinated into a framework, called the Dark Level Run Length Grid, where every component addresses the quantity of runs of a particular length and dim level force.

$$R(i, j) = (g(i, j) \mid i)),\ 0 \le i \le Ng,\ 0 \le j \le Rmax$$

where Ng is the greatest dim level and Rmax is the most extreme length. The GLRLM might be utilized to remove six surface attributes. These qualities utilize the pixel's dark level to separate between surfaces that have a similar SRE and LRE esteems however contrasting dim level dispersions. As a two layered model, think about a 5X5 picture, with 5 discrete dim levels:

$$I=\begin{bmatrix}5&2&5&4&4\\3&3&3&1&3\\2&1&1&1&3\\4&2&2&2&3\\3&5&3&3&2\end{bmatrix}$$

which P is the GLRLM matrix of I image:

$$P=\begin{bmatrix}1&0&1&0&0\\3&0&1&0&0\\4&1&1&0&0\\1&1&0&0&0\\3&0&0&0&0\end{bmatrix}$$

Let, Ng be the quantity of discrete force values in the picture. Nr is the image's number of discrete run lengths. Np be the quantity of pixels in the picture. (θ) be the number in the picture along point θ,as shown in equation 5 which is equivalent to e

$$\sum_{i=1}^{Ng}\sum_{j=1}^{Nr}P(i,j|\theta)\ \ and\ \ 1 \le Nr(\theta) \le Np \ldots..(5)$$

$(i, j|\theta)$ be the run length matrix for an arbitrary direction θ. p(i, j|θ) be the normalization run length matrix, as shown in equation 6 define as

$$P(i,j|\theta) = \frac{(i,j|\theta)}{Nr(\theta)} \ldots\ldots(6)$$

The features of GLRLM are computed from the GLRLM matrix and can be used to characterizedifferent textures in an image. The GLRLMfeatures are as follows:

Gray-Level Non-Uniformity: measures thevariance of the gray-level intensity values in the image as shown in equation 7.

$$GLN = \frac{\sum_{i=1}^{Ng}(\sum_{j=1}^{Nr}P(i,j|\theta))^2}{Nr(\theta)} \ldots\ldots\ldots\ldots (7)$$

Run length Non-Consistency: gauges the difference of the lengths of runs of pixels with a similar dim level as shown in equation 8.

$$RLN = \frac{\sum_{j=1}^{Nr}(\sum_{i=1}^{Ng}P(i,j|\theta)^2}{Nr(\theta)} \ldots\ldots..(8)$$

Run Rate: gauges the extent of runs of pixels with a similar dim level in the picture as shown in equation 9.

$$RP = \frac{Nr(\theta)}{Np} \ldots\ldots..(9)$$

Low Dim Level Run Accentuation: gauges the circulation of short runs of pixels with low dim level power values as shown in equation 10.

$$LGLRE = \frac{\sum_{i=1}^{Ng}\sum_{j=1}^{Nr}\frac{P(i,j|\theta)}{i^2}}{Nr(\theta)} \ldots..(10)$$

ISBN 979-8-3503-5094-4

High Dark Level Run Accentuation: gauges the appropriation of long runs of pixels with high dim level power values as shown in equation 11.

$$\text{HGLRE}=\frac{\sum_{i=1}^{Ng}\sum_{j=1}^{Nr}P(i,j|\theta)i^2}{Nr(\theta)}\text{.....(11)}$$

Short Run Accentuation: gauges the dispersion of short runs of pixels with any dim level power esteem as shown in equation 12.

$$\text{SRE}=\frac{\sum_{i=1}^{Ng}\sum_{j=1}^{Nr}\frac{P(i,j|\theta)}{j^2}}{Nr(\theta)}\text{.....(12)}$$

Long Run Accentuation: gauges the appropriation of long runs of pixels with any dim level force esteem as shown in equation 13.

$$\text{LRE}=\frac{\sum_{i=1}^{Ng}\sum_{j=1}^{Nr}P(i,j|\theta)j^2}{Nr(\theta)}\text{.....(13)}$$

Gray-Level Difference: gauges the fluctuation of the dim level force values in the picture as shown in equation 14.

$$\text{GLV}=\sum_{i=1}^{Ng}\sum_{j=1}^{Nr}p(i,j|\theta)(i-\mu)^2$$

$$\text{Here, }\mu=\sum_{i=1}^{Ng}\sum_{j=1}^{Nr}p(i,j|\theta)i\text{.....(14)}$$

Run Change: gauges the difference of the lengths of runs of pixels with any dim level as shown in equation 15.

$$\text{RV}=\sum_{i=1}^{Ng}\sum_{j=1}^{Nr}p(i,j|\theta)(j-\mu)^2\text{.....(15)}$$

Run Entropy: gauges the arbitrariness or vulnerability of the dissemination of runs of pixels with any dark level as shown in equation 16.

$$\text{RE}=\sum_{i=1}^{Ng}\sum_{j=1}^{Nr}p(i,j|\theta)\log_2(p(i,j|\theta)+\epsilon)\text{.....(16)}$$

Low Dark Level Run Length Accentuation: gauges the appropriation of short runs of pixels with low dim level power values as shown in equation 17.

$$\text{LGLRE}=\frac{\sum_{i=1}^{Ng}\sum_{j=1}^{Nr}\frac{P(i,j|\theta)}{i^2}}{Nr(\theta)}\text{.....(17)}$$

High Dim Level Run Length Accentuation: gauges the circulation of long runs of pixels with high dark level force values as shown in equation 18.

$$\text{HGLRE}=\frac{\sum_{i=1}^{Ng}\sum_{j=1}^{Nr}P(i,j|\theta)i^2}{Nr(\theta)}\text{.....(18)}$$

Short Run Low Dim Level Accentuation: gauges the dispersion of short runs of pixels with both low dark level force values and short lengths as shown in equation 19.

$$\text{SRLGLE}=\frac{\sum_{i=1}^{Ng}\sum_{j=1}^{Nr}\frac{P(i,j|\theta)}{i^2j^2}}{Nr(\theta)}\text{.....(19)}$$

Short Run High Dim Level Accentuation: gauges the dissemination of short runs of pixels with both high dim level force values and short lengths as shown in equation 20.

$$\text{SRHGLE}=\frac{\sum_{i=1}^{Ng}\sum_{j=1}^{Nr}\frac{P(i,j|\theta)i^2}{j^2}}{Nr(\theta)}\text{.....(20)}$$

Long Run Low Dim Level Accentuation: measures the appropriation of long runs of pixels with both low dim level force values and long lengths as shown in equation 21.

$$\text{LRLGLRE}=\frac{\sum_{i=1}^{Ng}\sum_{j=1}^{Nr}\frac{P(i,j|\theta)j^2}{i^2}}{Nr(\theta)}\text{.....(21)}$$

Long Run High Dim Level Accentuation: gauges the dissemination of long runs of pixels with both high dim level power values and long lengths as shown in equation 22.

$$\text{LRHGLRE}=\frac{\sum_{i=1}^{Ng}\sum_{j=1}^{Nr}P(i,j|\theta)i^2j^2}{Nr(\theta)}\text{.....(22)}$$

Picture grouping is a course of classifying pictures into predefined classes or classifications in view of the visual elements present in the pictures. Alzheimer's sickness arrangement alludes to the method involved with sorting patients into various phases of the illness in light of different clinical and natural markers. The objective of creating grouping models for Alzheimer's sickness is to empower prior and more precise finding of the illness, which can prompt better therapy results and worked on persistent consideration.

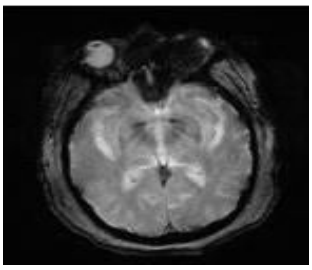
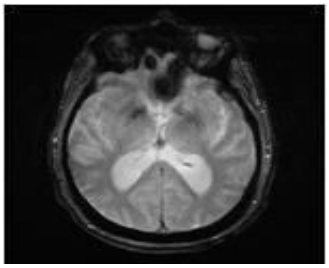
F. Support Vector Machine(SVM):

Order and relapse issues could be addressed using a help vector machine (SVM)-based directed AI strategy. A SVM getting ready system makes a model that sorts new guides to one of two orders when given a movement of planning models, changing over it into a non-probabilistic equal straight classifier. SVM extends the distance between the two classes by arranging getting ready advisers for centers in space. Then, considering which side of the opening they fall, new models are worked into that comparable locale and expected to have a spot with a class. To parcel a couple of classes, SVM makes a hyperplane in complex space. In an iterative manner, SVM constructs an ideal hyperplane to reduce error. Finding a Most limit Insignificant Hyperplane (MMH) that preferably describes the dataset is the central target of SVM. The information focuses that are most close the hyperplane are known as help vectors. By sorting out edges, these focuses will help lay out the partitioning line. The classifier's plan is all the more firmly connected with these thoughts. Isolating the gave dataset as best as plausible is the principal objective. The edge is the partition between the two spots that are the nearest to one another. In the gave dataset, picking a hyperplane with the biggest edge between help vectors is the objective. The accompanying stages are utilized by SVM to search for the most extreme negligible hyperplane. Make hyperplanes that effectively free the classes. Despite the fact that the dark is correctly recognizing the two parties in this situation, the blue and orange have greater order errors. Pick the right hyperplane, as displayed in the picture on the right that has the best seclusion from the two nearest information of interest.

III. RESULTS AND DISCUSSION

Images were acquired from the ADNI database andused for this study.

ISBN 979-8-3503-5094-4

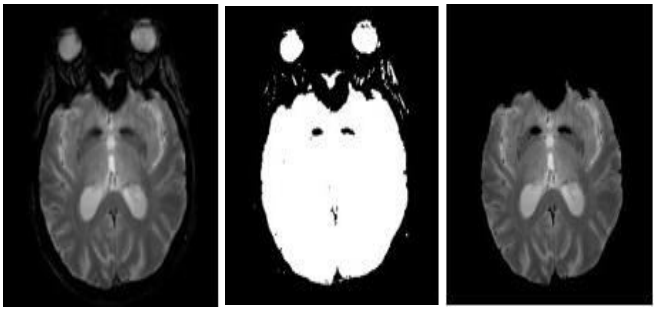



3(a) Cognitive Normal **3(b) Alzheimer's disease**

Fig. 3(a, b). Image Acquisition

Picture obtained from ADNI dataset that is perused by OpenCV. First get the picture in the grayscale to decrease the computational time. Figure 3a and Figure 3b shows the obtaining picture from the dataset. The got pictures are first skull stripped which implies the non-cerebrum tissue has been eliminated from the pictures with OTSU edge. It is fundamental for exact investigation of mind design and capability; as non-cerebrum tissue can disrupt the examination.

The limit picture go about as a cover which will overlay onto the first picture and afterward we track down the biggest region to remove the associated part, which is the mind. Figure 4 shows the skull stripped picture by utilizing otsu edge.



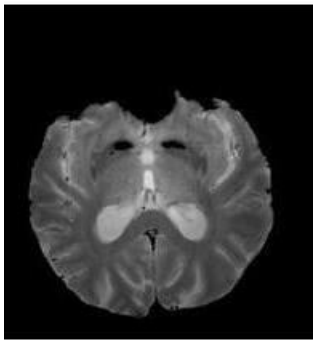
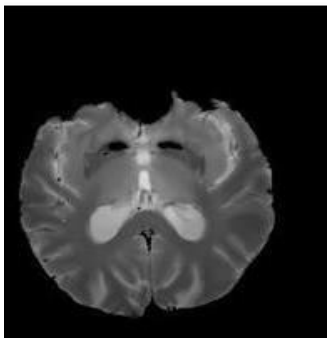
4(a) Original Image 4(b) OTSU Threshold 4(c) Skull Stripped

Fig. 4(a, b, c). Skull stripped image

Figure 4a shows the original Image, Figure shows the OTSU Threshold and Figure 4c shows the Skull Stripped. Picture is then separated to eliminate the commotion for additional handling. Non-neighborhood implies channel is utilized to eliminate commotion. The non-nearby means calculation substitutes a normal of various other pixel values for the worth of a given pixel. Just pixels with patches near the ongoing patch are found the middle value of, and minuscule patches focused on different pixels are contrasted with the fix fixated on the pixel of interest. As an outcome, this approach can effectively recuperate surfaces that ordinary denoising calculations would have obscured. Figure 5a and Figure 5b shows the filtered image from the skull stripped image.

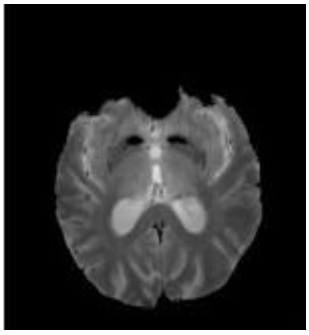
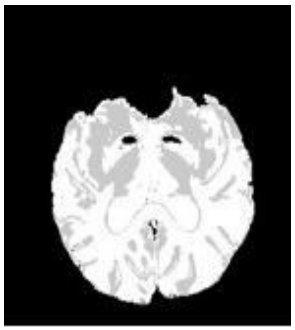
A. GMM Segmentation:

A capacity called a Gaussian Blend is contained a couple of Gaussians, all of which is implied by the string k (1, K), where K is the amount of gatherings in our dataset. The going with limits are the pieces of each Gaussian k in the mix: A mean μ that depicts its centre. A covariance Σ that describes its width. In a multivariate situation, this would be essentially indistinguishable from the ellipsoid's perspectives. A mixing probability π that chooses the size of the Gaussian capacity. Figure 6a and Figure 6b shows the sectioned picture of GMM.

5(a) Original Image **5(b) Filtered Image**

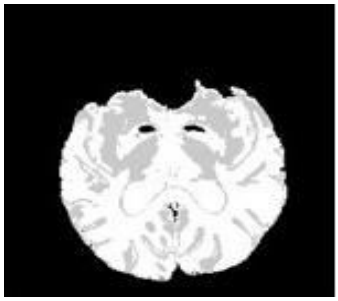
Fig. 5(a, b). Filtered image

6(a) Skull Stripped Image **6(b) GMM Segmentation**

Fig. 6a, b). GMM segmented image

Subsequent to sectioning the cerebrum picture utilizing GMM division, then, at that point, the picture is further procedure to white matter and dim matter division utilizing GMM division. In the wake of fragmenting the white matter and dim matter in the picture, white matter divided picture is further procedure into highlight extraction. Figure 7a and Figure7b shows the white matter fragmented picture.




7(a) GMM Segmentation **7(b) White Matter Segmentation**

Fig. 7(a, b). White matter segmented image

B. Feature Extraction GLRLM:

There are 16 elements extricated from the white matter sectioned picture as shown in Table 1. These elements of the pictures, the order cycle is finished by utilizing Backing Vector Machine (SVM).

Table 1. Feature Extraction GLRLM

ISBN 979-8-3503-5094-4

S. No	Features	Values
1	Gray Level Non-Uniformity	128.24138580891113
2	Gray Level Non-Uniformity Normalized	0.24138666949379567
3	Gray Level Variance	8.598219034896788
4	High Gray Level Run Emphasis	76.3479216664611
5	Long Run Emphasis	1.0729837889821354
6	Long Run High Gray Level Emphasis	82.36428440966732
7	Long Run Low Gray Level Emphasis	0.060533936440617206
8	Low Gray Level Run Emphasis	0.058296038837690176
9	Run Entropy	2.648590422267098
10	Run Length Non-Uniformity	506.85043224884066
11	Run Length Non-Uniformity Normalized	0.9540395640180889
12	Run Percentage	0.9765625
13	Run Variance	0.024360947269456337
14	Short Run Emphasis	0.9822752118122107
15	Short Run High Gray Level Emphasis	74.89594688643398
16	Short Run Low Gray Level Emphasis	0.05774177602753587

C. *Classification for Support Vector Machine (SVM):*

The dataset contains 81 images, each with unique features, and is divided into two categories. Class 0 corresponds to Alzheimer's disease (AD), while class 1 represents mentally healthy individuals (CN). Figure 8 illustrates the distribution of classes and the number of input image datasets utilized in the classification process.

The number of classes and it graphic:

```
cell_df['CLASS'].value_counts()

CLASS
0    41
1    40
Name: count, dtype: int64
```

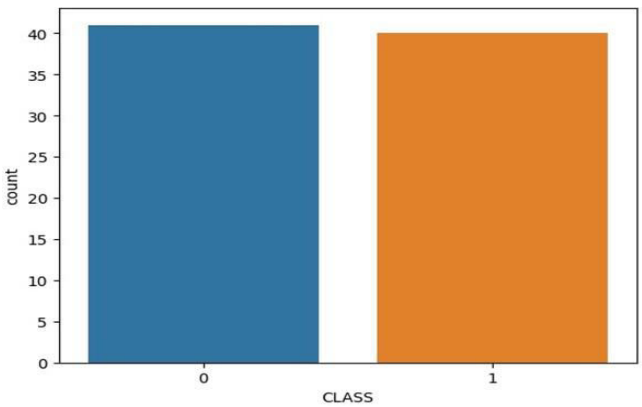


Fig. 8. Graph of classes and count

Table 2. Cell Description

Cell_df.describe()

S. No		GLNU	GLNUN	GLV	HGLRE	LRE
1.	cout	81.0000	81.0000	81.0000	81.0000	81.0000
2.	mean	115.823	0.22537	11.4803	70.4614	1.10920
3.	Std	28.0922	0.04071	3.75124	16.6105	0.04314
4.	Min	34.0891	0.13346	0.19397	20.0411	1.05875
5.	25%	101.084	0.19714	8.76756	59.1372	1.08038
6.	50%	116.1850	0.21898	12.5163	73.4243	1.09753
7.	75%	131.3305	0.25227	14.3216	84.8157	1.12639
8.	Max	181.5465	0.34915	16.7975	95.7975	1.28869

Table 3. Cell Portrayal

Cell_df.describe()

S. No		RE	RLNU	RLN	RLNUN	RP
1.	cout	81.0000	81.0000	81.0000	81.0000	81.0000
2.	mean	0.10336	2.65826	493.094	0.94828	0.97100
3.	Std	0.07115	0.21921	109.283	0.00845	0.00625
4.	Min	0.02259	2.07878	147.882	0.91869	0.94920
5.	25%	0.05639	2.50597	441.415	0.94265	0.96794
6.	50%	0.08533	2.66716	496.238	0.94801	0.97179
7.	75%	0.11580	2.79793	554.644	0.95368	0.97565
8.	Max	0.36138	3.21224	709.844	0.96683	0.98144

Table 2 and Table 3 shows the cell portrayal which is given the count, mean, standard deviation, least, and limit of the given information esteem removed from include extraction. Table 4 and Table 5 shows the upsides of x_ information and y_data which addresses highlights where it contains 81 lines and 16 segments, where 16 sections address the 16 elements that are referenced previously. y_ information addresses a class which has just 0 and 1. To work on the exactness and make the data set simpler to explore.

Table 4. x_data and y_data

#x_data
x_data = cell_df.drop(["CLASS"], axis =1)
#y_data
y_data =cell_df.CLASS.values

S. No		GLNU	GLNUN	GLV	HGLRE	LRE
1.	0	38.1716	0.24225	14.5766	80.2791	1.16018
2.	1	112.731	0.24523	15.1411	50.7639	1.12639
3.	2	76.7450	0.32353	1.83306	85.4041	1.28869
4.	3	101.084	0.20388	14.4028	41.6868	1.1569
5.	4	94.2864	0.18915	13.7349	47.5425	1.150613
6.	…	…	…	…	…	…
7.	76	148.271	0.31513	4.88214	95.4553	1.08176
8.	77	148.194	0.20675	11.4727	66.9789	1.09545

ISBN 979-8-3503-5094-4

9.	78	92.0787	0.18517	13.8527	46.6876	1.07805
10.	79	125.6596	0.25809	6.472899	84.81570	1.15277
11.	80	152.7580	0.21364	15.4552	59.13722	1.11336

Table 5. x_data and y_data of Rows and Columns

#x_data
x_data = cell_df.drop(["CLASS"], axis =1)
#y_data
y_data =cell_df.CLASS.values

S. No		RE	RLNU	RLNUN	RP	RV
1.	0	2.49623	147.885	0.93772	0.96036	0.07262
2.	1	2.560502	430.3473	0.93542	0.96383	0.04843
3.	2	2.07878	223.1470	0.94042	0.95665	0.19426
4.	3	2.680719	467.1479	0.94204	0.96449	0.01810
5.	4	2.764716	467.5993	0.93768	0.96235	0.07005
6.	…	…	…	…	…	…
7.	76	2.182884	444.7400	0.94523	0.97210	0.02893
8.	77	2.850269	682.1811	0.95169	0.97384	0.04085
9.	78	2.824827	472.9220	0.95098	0.97500	0.02598
10.	79	2.588374	470.2546	0.94609	0.96692	0.08283
11.	80	2.768031	674.3139	0.94288	0.96883	0.04749

D. Normalization:

Standardization is a significant stage in SVM arrangement. It includes scaling the info elements to a comparable reach to keep a specific component from ruling the growing experience. Standardization guarantees that all elements contribute similarly to the SVM's dynamic interaction. Table 6 an Table 7 shows the normalization of input value which is extracted from the feature extraction.

Table 6. Normalization

Normalization
x_data = (x_data – np.min(x_data))/(np.max(x_data) – np.min(x_data))

S. No		GLNU	GLNUN	GLV	HGLRE	LRE
1.	0	0.053749	0.000314	0.020508	0.113070	0.001607
2.	1	0.158788	0.000318	0.021303	0.071489	0.001559
3.	2	0.108091	0.000428	0.002555	0.120290	0.001788
4.	3	0.142381	0.000260	0.020263	0.058701	0.001602
5.	4	0.132803	0.000239	0.019322	0.066950	0.001593
6.	…	…	…	…	…	…
7.	76	0.208858	0.00416	0.006850	0.134450	0.001504
8.	77	0.208749	0.000264	0.0016135	0.094332	0.001516
9.	78	0.129693	0.000233	0.019488	0.065746	0.001491
10.	79	0.177002	0.000329	0.009091	0.119461	0.001596
11.	80	0.215178	0.000273	0.021746	0.083285	0.001541

Table 7. Normalization of Rows and Columns

Normalization
x_data = (x_data – np.min(x_data))/(np.max(x_data) – np.min(x_data))

S. No		RE	RLNU	RLNUN	RP	RV
1.	0	0.003489	0.208309	0.001294	0.001325	0.000075
2.	1	0.003580	0.606245	0.001290	0.001330	0.000041
3.	2	0.002901	0.314342	0.001297	0.001320	0.000246
4.	3	0.003749	0.658090	0.001300	0.001331	0.000087
5.	4	0.003867	0.658726	0.001293	0.001328	0.000071
6.	…	…	…	…	…	…
7.	76	0.003048	0.626522	0.001304	0.001342	0.000013
8.	77	0.003988	0.961028	0.001313	0.001344	0.000030
9.	78	0.003952	0.666225	0.001312	0.001346	0.000009
10.	79	0.003619	0.662467	0.001305	0.001335	0.000089
11.	80	0.003872	0.949945	0.001301	0.001337	0.000039

Dividing information into preparing and test information where the preparation set has 70% of picture highlights what's more, test set has 30% of picture highlights. Pictures that are in preparing are 56 endlessly pictures that are in test are 25 pictures which incorporate the two classes (Promotion and CN). The count of test and train image values. The all-out exactness and best hyper parameters and the precision of test and train. With network search CV, you characterize a matrix of values for these hyper parameters, and the procedure plays out a comprehensive inquiry, assessing every blend utilizing cross-approval. By the search grid cv, the best hyper parameters are inferred. In which the exactness of the set is likewise determined which is 72.69%.

IV. CONCLUSION

A significant number of individuals worldwide suffer from the degenerative brain disorder known as Alzheimer's disease. The use of image capturing techniques has become increasingly crucial for both detecting and managing Alzheimer's disease. In this study, Support Vector Machines (SVM) were employed to classify cases of Alzheimer's disease. The SVM model achieved a 72.69% accuracy, indicating its potential to assist in the early identification and diagnosis of this neurodegenerative condition. The SVM model's ability to accurately classify Alzheimer's disease could have significant clinical applications. It can assist healthcare professionals in making critical decisions and enabling personalized treatment plans. Healthcare providers can more precisely allocate resources to ensure that individuals in need of diagnostic testing and specialized care receive it by accurately identifying potential Alzheimer's disease cases.

REFERENCES

[1] J. Feng, S. -W. Zhang and L. Chen, "Extracting ROI-Based Contourlet Subband Energy Feature from the sMRI Image for Alzheimer's Disease Classification," in IEEE/ACM Transactions on Computational Biology and Bioinformatics, vol. 19, no. 3, pp. 1627-1639, 1 May-June 2022.

[2] F. Riaz et al., "Gaussian Mixture Model Based Probabilistic Modeling of Images for Medical Image Segmentation," in IEEE Access, vol. 8, pp. 16846-16856, 2020.

ISBN 979-8-3503-5094-4

[3] M. Bachute, N. Vyas, S. Khanna, K. Modi and C. Katpatal, "Alzheimer's disease detection and classification using machine learning techniques," 4th Smart Cities Symposium (SCS 2021), Online Conference, Bahrain, 2021, pp. 312-318.

[4] Y. Chen, M. Cai, X. Zhou, C. Ning, C. Cao and J. Yang, "A Robust Spatial Information- Theoretic GMM Algorithm for Bias Field Estimation and Brain MRI Segmentation," in IEEE Access, vol. 8, pp. 89617- 89629, 2020.

[5] T. Zhu, C. Cao, Z. Wang, G. Xu and J. Qiao, "Anatomical Landmarks and DAG Network Learning for Alzheimer's Disease Diagnosis," in IEEE Access, vol. 8, pp. 206063-206073, 2020.

[6] M. Dadar et al., "Validation of a Regression Technique for Segmentation of White Matter Hyperintensities in Alzheimer's Disease," in IEEE Transactions on Medical Imaging, vol. 36, no. 8, pp. 1758-1768, Aug. 2017.

[7] V. Patil and S. L. Nisha, "Detection of Alzheimer's Disease Using Machine Learning and Image Processing," 2021 International Conference on Smart Generation Computing, Communication and Networking, Pune, India, 2021, pp. 1-5.

[8] M. Zaabi, N. Smaoui, W. Hariri and H. Derbel, "Medical Imaging-based Techniques for Alzheimer's Disease Detection: A Survey," 2021 18th International Multi- Conference on Systems, Signals & Devices (SSD), Monastir, Tunisia, 2021, pp. 1202-1208.

[9] K. D. T, M. G, J. S and D. K. K, "Deep learning to predict Alzheimer's disease from neuroimaging of MRI analysis: A systematic study," 2020 International Conference on Power, Energy, Control and Transmission Systems (ICPECTS), Chennai, India, 2020, pp. 1-9.

[10] M. Mamun, S. Bin Shawkat, M. S. Ahammed, M. M. Uddin, M. I. Mahmud and A. M. Islam, "Deep Learning Based Model for Alzheimer's Disease Detection Using Brain MRI Images," 2022 IEEE 13th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON), New York, NY, NY, USA, 2022.

[11] R. Anitha, Prakash and S. Jyothi, "A segmentation technique to detect Alzheimer's disease using image processing," 2016 International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT), Chennai, India, 2016, pp. 3800-3801.

[12] H. Fuse, K. Oishi, N. Maikusa, T. Fukami and J. A. D. N. Initiative, "Detection of Alzheimer's Disease with Shape Analysis of MRI Images," 2018 Joint 10th International Conference on Soft Computing and Intelligent Systems (SCIS) and 19th International Symposium on Advanced Intelligent Systems (ISIS), Toyama, Japan, 2018, pp. 1031- 1034.

[13] D. Tosun, N. Schuff and M. Weiner, "An integrated multimodality MR brain imaging study: Gray matter tissue loss mediates the association between cerebral hypoperfusion and alzheimer's disease," 2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Minneapolis, MN, USA, 2009, pp. 6981-6984.

[14] S. -L. Wang, Z. -C. Cai and C. -L. Xu, "Classification for Alzheimer's disease based on SVM using a spatial texture feature of cortical thickness," 2013 10th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP), Chengdu, China, 2013, pp. 158-161.

[15] R. Das and S. Kalita, "Classification of Alzheimer's Disease Stages Through Volumetric Analysis of MRI Data," 2022 IEEE Calcutta Conference (CALCON), Kolkata, India, 2022, pp. 165-169.

[16] A. B and K. Kalirajan, "Alzheimer's Disease Classification using Convolutional Neural Networks," 2023 International Conference on Innovative Data Communication Technologies and Application (ICIDCA), Uttarakhand, India, 2023, pp. 1044-1048.

ISBN 979-8-3503-5094-4

Detection of Fetal and Mother Heart Rate with Uterine Contractions

R.Seetharaman
Department of ECE
CEG Campus, Anna University
Chennai-600025 India
rseetharamanece@annauniv.edu

C.Ambhika
Department of AIML
RMD Engineering College
Chennai-601206 India

S.Gayathri
Department of EEE
CEG Campus, Anna University
Chennai-600025 India

Anika Thirukkonda Karthikeyan
Department of ECE
CEG Campus, Anna University
Chennai-600025 India

L.Harihara Subramanian
Department of Management Studies
CEG Campus, Anna University
Chennai-600025 India

P.S.Harsha Ruban
Department of ECE
Sri Krishna College of Engineering and Technology
Coimbatore-641008 India

S.Sasikumar
Department of ECE
CEG Campus, Anna University
Chennai-600025 India

***Abstract*— The heart rate of a pregnant woman and the fetus can be detected and monitored. The average heart rate of a pregnant woman is around 70 to 90 beats per minute, and for that of a fetus is 110 to 160 beats per minute. Heart rate can be detected using AD8232 heart rate monitoring module. As a pregnant woman approaches labor, their uterine muscles begin contracting. This is known as Uterine Contractions. Since uterine contraction is a form of muscle contraction, we can use AD8226 module for detecting the uterine contractions in a pregnant woman. An Arduino Nano is programmed and used to measure the heart rate as well as muscle contractions. Graphs are obtained using Arduino IDE. The AD8232 Heart rate monitoring module is an Electrocardiogram (ECG) Sensor, and the AD8226 is an Electromyography (EMG) sensor. Electrodes are connected to both the modules. The electrodes need to be fixed at exact positions in the body to be able to detect the heart rate as well as the contractions of the uterus accurately.**

Keywords— Fetal Heart Rate, Mother Heart Rate, Uterine Contractions, AD8232 ECG sensor, AD8226 EMG sensor, Arduino Nano, Electrodes

I. INTRODUCTION

Monitoring the heart rates of a pregnant woman and the fetus is a very crucial aspect of the health and wellness of the mother as well as the child. Constant monitoring of the heart rate ensures that any irregularity or changes in the heart rate of the mother or the fetus can be treated immediately by notifying the health professionals, ensuring a safe pregnancy and well-being of the mother and the child. Fetal heart rate monitoring is highly important for women who have a high-pregnancy risk. A high pregnancy risk for women can be possible if they are diabetic, or may have high blood pressure. It also ensures that the fetus is developing properly at the right stages of pregnancy.
During pregnancy, another important factor to be monitored is the contractions of the uterus. Measuring the uterine contractions of a woman is crucial during labor. The frequency, duration and strength of the contractions can be measured. Excessive uterine activity is a common cause of interrupted fetal oxygenation. Normal contraction frequency is defined as five or fewer contractions in 10 minutes averaged over 30 minutes. Other important features of uterine activity include contraction intensity, duration, resting tone, and time between contractions.

II. COMPONENTS USED FOR THE STUDY

The components used for the monitoring of heart rate of a pregnant woman and the fetus are AD8232 heart rate monitoring module, electrodes, and an Arduino Nano.The components used to measure uterine contractions are AD8226 module along with electrodes as well as an Arduino Nano [1].

The AD8232 heart rate detection module is specifically designed for medical purposes. The method of Electrocardiogram (ECG) is used to detect heart rate of a pregnant woman and a fetus. The sensor amplifies the weak electrical signals generated by the heart and outputs a clean and readable waveform, making it ideal for monitoring heart rates and rhythms [2]. For the measurement of uterine contractions, Electromyography or EMG is used, and that is done with the help of the AD8226 EMG sensor module. The Arduino Nano is programmed and the output graphs of the heart rate and contractions of muscles is viewed in Arduino IDE.

Electrodes are conductive pads or sensors that are placed on the skin to pick up electrical signals from the body. They are connected to one of the modules, AD8232 or AD8226 depending upon the application, which in turn is connected to the Arduino Nano. The electrodes need to be placed in exact positions in the body to be able to detect and measure the heart rate as well as contractions accurately. The use of electrode patches helps in cancelling out the noise, leading to more accurate results. This is because the electrodes would be fixed in the exact position, leading to easier detection of the data. Even if there is a slight movement of

ISBN 979-8-3503-5094-4

the electrodes, the AD8232 and AD8226 modules are designed to produce the required output accurately. When all of these components: AD8232, AD8226, Electrode cables and patches along with an Arduino Nano board are used together, it forms a comprehensive system and helps in effectively detecting and monitoring the heart rate and uterine contractions ensuring a safe delivery [3].

A. AD8232: Heart rate detection module

The AD8232 sensor module is a compact chip designed to measure the electrical movement of the heart rate as shown in Figure 1. It serves as a safe and non-invasive ECG sensor, capable of directly connecting to controllers like Arduino and Raspberry Pi. This sensor efficiently detects and measures heart rate signals even when the electrodes are not in optimal contact with the body. After extraction, the signal is amplified, allowing easy integration with embedded controller devices or analog-to-digital converters (ADCs) with ultralow power consumption.

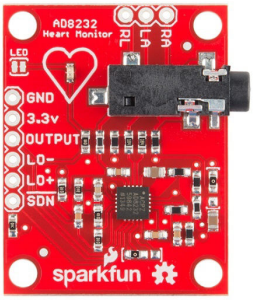


Fig. 1. AD8232 ECG Sensor Module

B. AD8226: EMG Sensor Module

The AD8226 EMG sensor module, as shown in Figure 2 is a small, specialized chip engineered to measure electrical signals generated by muscle activity, known as electromyography (EMG). This sensor offers a safe and non-intrusive way to capture EMG signals and can be seamlessly connected to various controllers, including Arduino and Raspberry Pi. When the electrodes are placed accurately, the AD8226 sensor effectively extracts EMG signals. These signals are then amplified, ensuring accurate detection and measurement. This design allows for easy integration with embedded controller devices or analog-to-digital converters (ADCs), all while maintaining ultra-low power consumption [4].

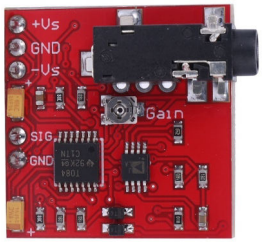


Fig. 2. AD8226 EMG Sensor Module

C. Electrodes

The electrodes act as a sensor module in this study experiment. The electrodes are fixed on the patients' body to detect the necessary information. The ends of the electrodes are connected to either AD8232 or AD8226 to detect heart rate and muscle contraction respectively. The electrodes used must be disposable electrodes for hygiene purposes, to prevent the spread of infections via electrodes. The electrodes are pre-gelled and is suitable for the skin of the patient, without being too harsh or containing harmful chemicals leading to skin damage. Three electrodes are used for this experiment. The three electrodes have a common end which is connected to the end of the AD8232 Module and AD8226 Module. The three electrodes would be labelled and they are placed in those positions in the body of the patient[5].

Electrodes are essential for continuous fetal and maternal monitoring during labor, providing detailed information on well-being. They measure heart and muscle electrical activity, supplying data on fetal and maternal heart rates and uterine contractions, aiding in the assessment of oxygenation, distress, and labor progress. These electrodes are utilized externally on the abdomen for comfort using Doppler ultrasound and internally through electrocardiography (ECG) on the fetal scalp and uterine wall for more precise signals. While external electrodes are less invasive, they may face challenges in obtaining clear signals, while internal electrodes offer reliability.

Connected to modules like AD8232 or AD8226, electrodes amplify, filter, and process signals, interfacing with an Arduino Nano board. This microcontroller controls, communicates, and manages data display, storage, or transmission to a computer or mobile device. This integrated system facilitates effective detection and monitoring of fetal and maternal parameters, ensuring a safe delivery by alerting healthcare providers to complications or distress and providing valuable data for decision-making during labor (Figure 3).

Fig. 3. Electrode connection to AD8226 Module; Electrode cable with patches.

D. Arduino Nano

Arduino Nano microcontroller is a compact microcontroller board that plays a significant role in the field of healthcare by enabling the development of monitoring systems for Fetal Heart Rate (FHR), Maternal Heart Rate (MHR), and Uterine Contractions (Figure 4). It contains the ATMega328 microprocessor. Its versatility, ease of programming, and ability to interface with various sensors make it a valuable

ISBN 979-8-3503-5094-4

tool for creating cost-effective and efficient monitoring solutions during pregnancy and childbirth [6].
The Arduino IDE software is used to program the Arduino Nano board using C/C++ programming. Arduino Nano can be programmed and used for various medical applications, and in this case, it has been used for detection of fetal and mother heart rate along with uterine contractions. It plays a crucial role in this study as it needs to be connected to the AD8232 and AD8226 modules before testing.

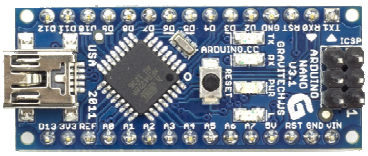


Fig. 4. Arduino Nano

III. TESTING OF AD8232 MODULE WITH ARDUINO NANO AND ELECTRODES

A. Objective

The objective of testing the AD8232 Module with electrodes connected to the body is to check if the results given by the module is accurate. The Arduino Nano microcontroller board is programmed and connected to the AD8232 and graphs have been obtained on Arduino IDE and have also been displayed using an LED display (Figure 5).

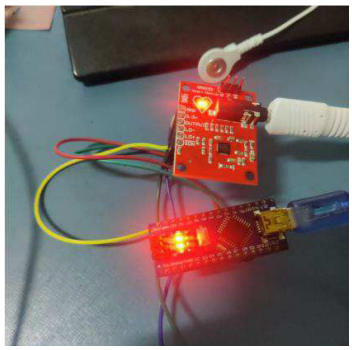


Fig. 5. AD8232 connected to Arduino Nano and Electrode cable

B. Working Process

The AD8232 module operates on the fundamental principles of Electrocardiography (ECG), a technique used to monitor the heart's electrical activity. Electrodes placed on the body detect the heart's electrical signals. The signal keeps varying as the heart muscles contract and relax. The AD8232 processes these weak signals, amplifies them, and applies filtering mechanisms to eliminate noise and interference. These processed signals, now amplified and are without noise or disturbances. The signals are transmitted to the Arduino Nano microcontroller. The Arduino Nano is the microcontroller of the system which is used to further process the signals and analyze the ECG waveform. Through careful programming, the Arduino Nano interprets the data and calculates the heart rate, which is then displayed as output in the form of graph in Arduino IDE as well as text in the LED display. This collaborative process allows for real-time monitoring of heart activity, making it an essential setup in medical applications, and can be used to detect and monitor fetal and mother heart rate efficiently and easily [7].

C. Circuit Connections and Placement of Electrodes on the Body

The AD8232 sensor module requires a stable and regulated power supply to function efficiently. Providing it with a consistent source of power ensures reliable and accurate operation, allowing the sensor to capture and process electrical signals effectively. The AD8232 is given power with the help of a Regulated Power Supply of 3.3 V, by connecting positive supply of Regulated Power Supply with 3.3V pin of AD 8232 and negative supply of Regulated Power Supply with GND pin of AD 8232.

The Arduino Nano is supplied with power by connecting it with a PC /Laptop via a USB cable. The serial connection has been established between the PC/ Laptop and the Arduino Nano board. There are three electrodes used for this study, and they are labelled as L (left), R (right) and COM (common).

The placement of each electrode is for a significant purpose. Place the R electrode on the left arm of the patient, L electrode on the right arm of the patient, and the common (COM) electrode needs to be placed on the right ankle or thigh of the patient. The R and L electrodes are switched up because they are designed in such a manner that is in the doctor's point of view.

The other main connections between the Arduino Nano and AD8232 are: LO- pin of AD8232 needs to be connected to the D9 pin of the Arduino Nano, LO+ of AD8232 is connected to D8 pin of Arduino Nano. The OUTPUT pin of AD8232, which is the Analog output of the sensor is connected to A0 pin of Arduino Nano.

D. Observations for Heart Rate Detection using AD8232

As the AD8232 Module was connected to the Arduino Nano and tested, the output was viewed in the serial monitor and serial plotter of Arduino IDE.
The output from Arduino IDE in the serial plotter, serial monitor and using the LED has been attached below.
The same output graph can also be viewed using the HANTEK PC oscilloscope.

HANTEK PC Oscilloscope is a device which is connected to the Arduino using probes. Using the HANTEK PC

ISBN 979-8-3503-5094-4

Oscilloscope software, we can view the output graphs and even take readings of necessary values.

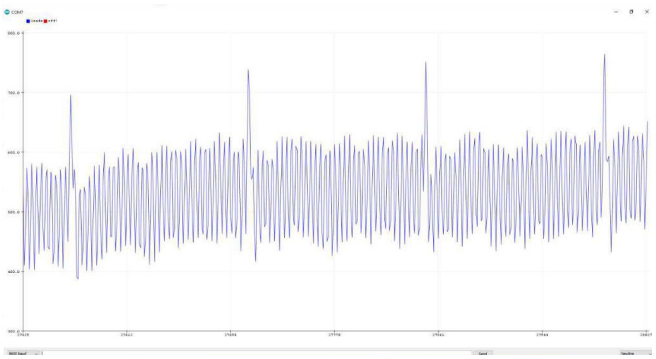


Fig. 6. Serial Plotter in Arduino IDE

The Serial Plotter in Arduino IDE (Figure 6) depicts a periodic signal with occasional noise, showcasing distinct peaks at regular intervals, indicative of a recurring pattern. This real-time visualization tool represents data sent from the Arduino board on the x-axis (time or sequential data points) and the y-axis (Arduino-received values), adjusting dynamically. The inference suggests data from a sensor measuring a natural frequency phenomenon, with peaks potentially denoting amplitude variations. Noise may arise from interference or environmental factors. Further confirmation requires insights into data source, Arduino code, and the connected circuit, achievable through experimentation and modifications.

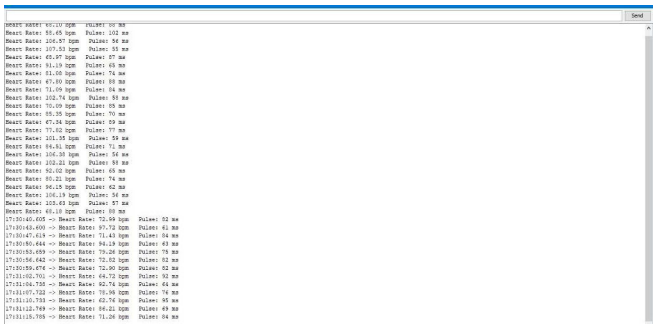


Fig. 7. Serial Monitor in Arduino IDE which displays heartbeat of patient

The real-time data graph, monitored through Arduino IDE Serial Monitor, displays a patient's heartbeat with BPM and pulse volume at different timestamps (Figures 7 and 8). Serving for code debugging and sensor value printing, the Serial Monitor suggests a normal resting heart rate between 60 and 100 BPM, supported by consistent BPM values and stable pulse volume. This indicates a well-functioning cardiovascular system without signs of arrhythmia. To confirm, compare with other vital signs and use the Serial Plotter tool for a graphical representation.

The HANTEK PC Oscilloscope reveals a complex electronic waveform, vital for precise voltage signal analysis. Inferences suggest potential noise and instability, requiring adjustments to controls like time base and vertical scale. Utilizing advanced features ensures a detailed analysis in line with research standards (Figure 9).

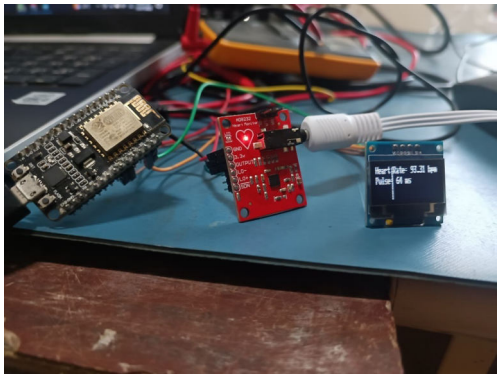


Fig. 8. Display of heartbeat using LED Display

Fig. 9. Output graph using HANTEK PC Oscilloscope

E. Calculations

The heart rate can be calculated manually from the graph output in Arduino IDE (Figure 10).

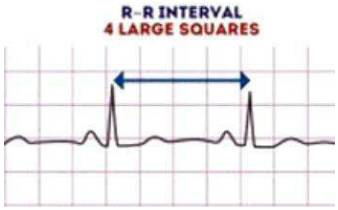


Fig. 10. R-R interval

$$HR = \frac{60}{RR\ interval}$$

Here HR stands for heart rate and RR interval indicates the time period between two peaks.
The output pulse observed RR interval is 0.715s or 715ms (approx.)
Heart rate is calculated from the above formula
= 60/0.715 = 84 beats per minute.

ISBN 979-8-3503-5094-4

IV. TESTING OF AD8226 MODULE WITH ARDUINO NANO AND ELECTRODES

A. Objective

The objective of testing the AD8226 EMG Sensor module along with the Arduino Nano microcontroller board is to detect if the muscle contractions are measured accurately (Figures 11 and 12). Since this is used to detect uterine muscle contractions and relaxations, it has been tested using the arm muscles. The electrode patches need to be properly fixed on the arm. For testing purposes, the arm was contracted and relaxed to check if the output graphs vary [8].

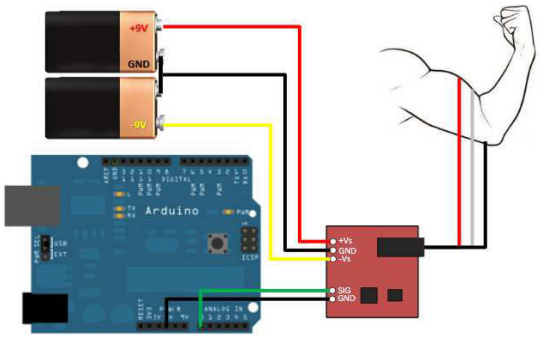


Fig. 11. EMG Electrodes placed on the arm muscles, AD8226 EMG sensor connected to the electrodes as well as the programmed Arduino Nano

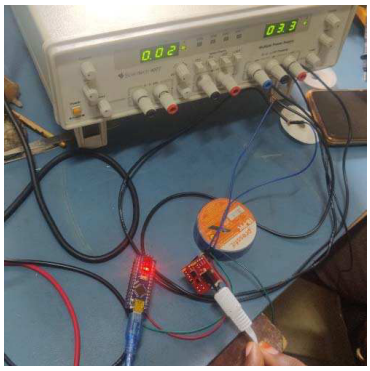


Fig. 12. AD8226 and Arduino Nano

Fig. 13. Placement of electrodes on the arm muscles

B. Working Process

The electrodes attached to the arm of the body collects electrical signals as there is contraction and relaxation of the muscles (Figure 13). The collected signals are passed to the AD8226 EMG Sensor Module and it amplifies the received signal while also removing the noise. The data is sent to the Arduino Nano for processing, and the output graph is viewed on the serial monitor of Arduino IDE. Usually, what we infer from the graph is that it peaks when there is contraction, and there are no peaks when the arm muscles are in a relaxed position [9].

C. Observations from Connecting AD8226 EMG Sensor to Arduino and Electrodes

As the AD8226 module is tested, the output can be viewed using the serial plotter in Arduino IDE.We can infer from the graphs obtained that during muscle contraction, the graph will peak and during the relaxation of arm muscles, the graph will not peak.

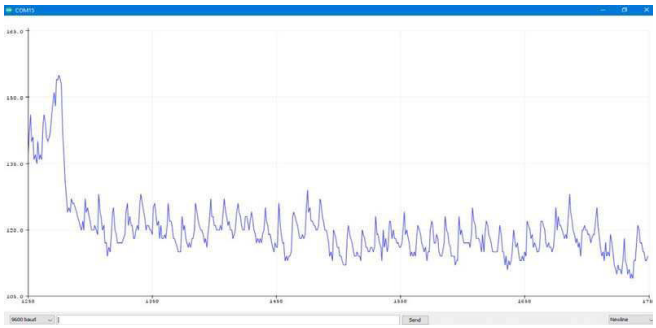


Fig. 14. Output observed in Arduino IDE Serial Plotter when hand is in relaxed position

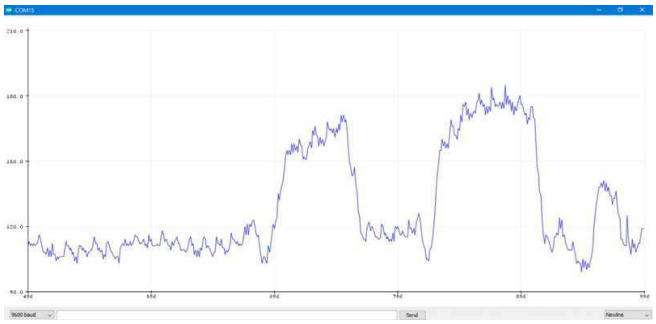


Fig. 15. Output observed in Arduino IDE Serial Plotter when
hand muscles are contracting

In a relaxed hand position, the Arduino IDE Serial Plotter graph illustrates output fluctuations around 0.4 to 0.5 volts (Figure 14). This tool visualizes real-time data from the Arduino board, presenting variables on the x-axis (time or sequential data points) and the y-axis (Arduino-received values), dynamically adjusting with changing values. The graph suggests that the recorded data may originate from an electromyography (EMG) sensor, measuring the electrical activity of hand muscles. Voltage fluctuations correspond

ISBN 979-8-3503-5094-4

with the level of muscle contraction or relaxation; in relaxed states, voltage remains low and stable, while hand clenching or movement leads to increased and variable voltage levels.

The Arduino IDE Serial Plotter graph illustrates the output during hand muscle contraction, displaying a significant signal increase with two prominent peaks before returning to a lower baseline level (Figure 15). This real-time visualization tool represents data from the Arduino board, plotting variables on the x-axis (time or sequential data points) and the y-axis (Arduino-received values), adjusting dynamically. The data likely originates from a sensor, such as an electromyography (EMG) sensor, measuring the electrical activity of hand muscles. Voltage fluctuations indicate the level of muscle contraction or relaxation, with low and stable voltage during relaxation and increased and varied voltage during hand clenching or movement [10].

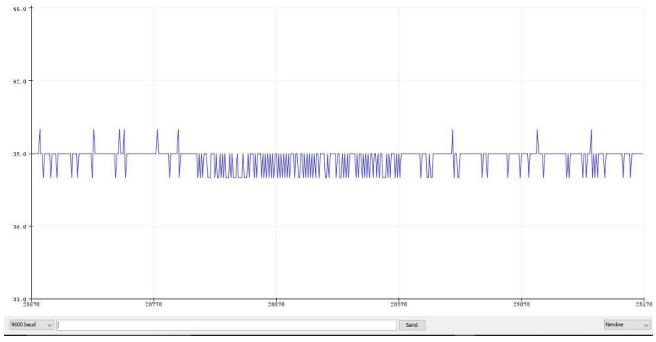


Fig. 16. Contraction and relaxation of muscles

Figure 16 seemingly depicts temporal data, potentially reflecting electrical activity associated with muscle contraction and relaxation. Spikes suggest moments of muscle contraction, while lower, less intense segments may indicate periods of relaxation. An inference is that the data originates from a sensor, possibly an electromyography (EMG) sensor, measuring muscle electrical activity. Voltage fluctuations likely signify the degree of muscle contraction or relaxation, with low and stable voltage during relaxation and increased, variable voltage during contraction

V. CONCLUSION

In conclusion, the fetal and mother heart rate monitoring system can be extremely useful, especially for patients who are unable to go to hospitals for their pregnancy checkups due to health complications. This can be done remotely, at the comfort of the patients' home and the necessary details can be transferred to the health officials. Regular and constant monitoring of heart rates of the fetus as well as mother is very crucial. Continuous monitoring of FHR ensures the baby's oxygenation and overall cardiac health, allowing healthcare professionals to detect any distress promptly. Similarly, monitoring MHR provides crucial information about the mother's cardiovascular condition, aiding in the assessment of her overall well-being. If there is any emergency situation, it can be immediately treated by health professionals without any delay, leading to a safe pregnancy and well-being of the mother and child. Measuring contractions in the uterus is also very important as it could indicate the time of labor, which helps the mother and health professionals to be prepared for any situation. Real-time data acquisition and analysis enable healthcare professionals to make informed decisions, intervene when necessary, and provide appropriate care, ultimately contributing to safer pregnancies and childbirth experiences.

REFERENCES

[1] EKG-Based Heart-Rate Monitor Implementation Using the MSP430G2xx Launchpad Development Kit: Abhishek Joshi, Sourabh Ravindran, Austin Miller. – Texas Instruments.

[2] Anitha S Prasad, Kavanashree N "ECG Monitoring System Using AD8232 Sensor," 2019 International Conference on Communication and Electronics Systems (ICCES), Coimbatore, India, 2019, pp. 976-980

[3] Sletten, J., Kiserud, T. and Kessler, J., 2016. Effect of uterine contractions on fetal heart rate in pregnancy: a prospective observational study. Acta obstetricia et gynecologica Scandinavica, 95(10), pp.1129-1135.

[4] Arduino-Nano Based Low Cost Power Converter Learning Kit: Sivapriyan R, Ajay K.V, Ashwath Koorse N.

[5] Obstetrics: Normal and Problem Pregnancies (Seventh Edition), Chapter 15 - Intrapartum Fetal Evaluation: David Arthur Miller

[6] Rotariu, C., Pasarica, A., Andruseac, G., Costin, H. and Nemescu, D., 2014, October. Automatic analysis of the fetal heart rate variability and uterine contractions. In 2014 International Conference and Exposition on Electrical and Power Engineering (EPE) (pp. 553-556). IEEE.

[7] Rahman, M.A., Li, Y., Nabeed, T. and Rahman, M.T., 2021, March. Remote monitoring of heart rate and ECG signal using ESP32. In 2021 4th International Conference on Advanced Electronic Materials, Computers and Software Engineering (AEMCSE) (pp. 604-610). IEEE.

[8] Kusuma, I.K.B.J., Dharma, M.A.G., Kumara, I.N.S. and Garinto, D., 2023, November. Biceps Muscle Motor Strength Signal Reading Using AD8226 Sensor. In 2023 International Conference on Smart-Green Technology in Electrical and Information Systems (ICSGTEIS) (pp. 145-149). IEEE.

[9] Trivedi, N., Naik, I., Fadia, N., Karia, D. and Ghatte, N., 2021, November. Universal Remote using EMG based Gesture Recognition. In 2021 IEEE Bombay Section Signature Conference (IBSSC) (pp. 1-5). IEEE.

[10] Cohen, W.R., Ommani, S., Hassan, S., Mirza, F.G., Solomon, M., Brown, R., Schifrin, B.S., Himsworth, J.M. and Hayes‐Gill, B.R., 2012. Accuracy and reliability of fetal heart rate monitoring using maternal abdominal surface electrodes. Acta obstetricia et gynecologica Scandinavica, 91(11), pp.1306-1313.

ISBN 979-8-3503-5094-4

Artificial Bee-Optimized CNN for Osteoporosis Detection using Leg X-ray Images

Bakiya A
Department of Electronics and Communicaion Engineering, Vel tech Rangarajan Dr Sagunthala R&D Institute of Science and Technology
Chennai.
bakiya88@gmail.com

Vetrivel V
Department of Mathematics, Indian Institute of Technology Madras, Chennai.
vetri@iitm.ac.in

Kamalanand K
Department of Instrumentation Engineering, MIT Campus Anna University, Chennai.
kamalanand@mitindia.edu

Anitha A
Department of Computer Science, Tamilnadu Dr Ambedkar law University
anitha.dgvc@gmail.com

Aswath S
Department of Electronics and Communicaion Engineering, Vel tech Rangarajan Dr Sagunthala R&D Institute of Science and Technology
Chennai
saswath@veltech.edu.in

Valliammai M
Department of Electronics and Communicaion Engineering, Vel tech Rangarajan Dr Sagunthala R&D Institute of Science and Technology
Chennai.
drvalliammaim@veltech.edu.in

Sridevi S
School of Computer Science and Engineering, Vellore Institute of Technology, Chennai
sridevi.s@vit.ac.in

***Abstract*—Osteoporosis is emerging as a worldwide health concern as people live longer. Although Dual-Energy X-Ray Absorptiometry (DXA) is considered the primary method for diagnosing osteoporosis, its widespread use as a screening tool is limited. This study uses an evolutionary-based deep learning algorithm to forecast osteoporosis using leg X-ray images. A Convolutional Neural Network was enhanced using the artificial bee optimization algorithm to classify normal and osteoporosis cases using leg X-ray images. The effectiveness of the developed Bee CNN was compared to the conventional CNN. The results indicate that the CNN combined with bee optimization achieved a high accuracy rate of 92.32% in accurately classifying normal and osteoporosis X-ray leg images. Integrating artificial bee optimization into the CNN demonstrated superior performance compared to the conventional CNN. Based on these results, the Artificial Bee CNN model proves to be a valuable tool for straightforward osteoporosis prediction in practical clinical environments.**

Keywords— **Osteoporosis, Leg X-Ray images, Artificial Bee optimization algorithm, CNN, classification.**

I. INTRODUCTION

Osteoporosis is characterized by the deterioration of bone tissue microarchitecture and a decrease in bone mass. It is a prevalent metabolic bone disease, rendering individuals more susceptible to fractures [1]. Fractures in areas such as the spine, hips, legs, and wrists are anticipated outcomes of osteoporosis and significantly diminish a patient's quality of life. In severe instances, these fractures can contribute to conditions that elevate the risk of mortality [2]. As life expectancy increases worldwide, the aging population experiences accelerated aging, contributing to osteoporosis as a notable global public health issue affecting millions annually. One of the challenges with osteoporosis is that it often develops without apparent symptoms, making early detection difficult [3].

DXA is one of the well-known diagnostic test techniques for osteoporosis and is used to measure bone mineral density. However, despite DXA scans being the preferred method for diagnosing osteoporosis, their extensive use for general screening is impractical because of their expensive nature and restricted accessibility in developing countries [4]. X-rays are often captured during routine clinical check-ups or before specific medical procedures. Hence, it holds significant medical and economic importance if healthcare professionals employ X-rays as a screening method for osteoporosis in patients [5].

Deep learning has recently found applications in various areas of medical imaging interpretation [6-7]. Additionally, numerous studies have explored the use of deep understanding to create a screening tool for osteoporosis [8-9]. Sukegawa et al. (2022) [10] created ensemble deep-learning models to predict osteoporosis based on panoramic radiographs. The researchers found that the ensemble deep learning models significantly enhance the overall accuracy compared to conventional methods. Jang et al. (2021) [11] have used the simple CNN model to predict osteoporosis based on hip radiography images. The authors have found that the accuracy of the developed model attains 81.2% for the prediction of osteoporosis using hip radiography images.

The main objective of this research was to forecast osteoporosis using leg X-ray images by developing an integrated artificial bee optimization with a CNN model. Through the integration of artificial bee

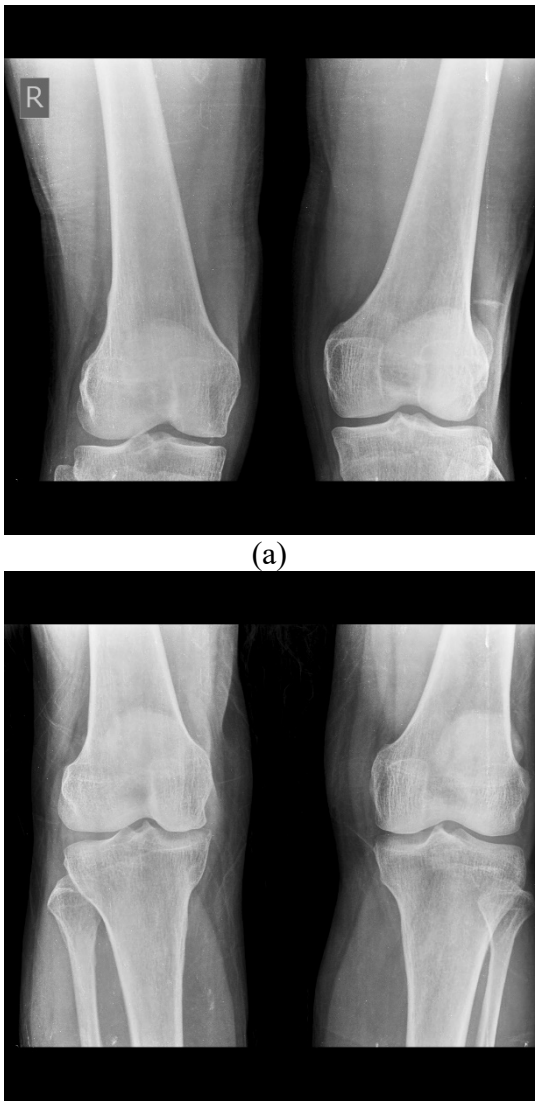
ISBN 979-8-3503-5094-4

optimization methods with CNN, the model demonstrates promise as a screening tool for osteoporosis, addressing previous constraints effectively.

II. METHODOLOGY

A. Dataset

The osteoporosis leg x-images dataset comprises 36 normal and 49 osteoporosis subjects. Participants' ages ranged from 23 to 76 years, with 60% identified as female and 40% identified as male [12]. The weight of participants ranged from 39Kg to 98 Kg. The typical normal and osteoporosis leg X-ray images are depicted in Figure 1(a-b).



(a)

(b)

Figure 1 (a) Healthy knee X-Ray Images and (b) Osteoporosis affected knee X-Ray Images

B. Artificial Bee Optimization Algorithm:

The Artificial Bee Optimization process comprises four primary stages: initialization, employed bee, onlooker bee, and scout bee phases. Different bee types can alternate their functions throughout the terminal condition of maximum iterations. It is worth noting that each food source has an associated counter [13]. When a food source does not show improvement, its corresponding counter is incremented by 1. Nevertheless, should a solution's quality not advance beyond a predetermined limit parameter, the employed bee transitions into a scout bee, resetting the counter to 0 [14].

Phase of Initialization

Assume X is the initial population, represented as $X = \{X_1, X_2, ..X_{SN}\}$, where the individuals are randomly generated throughout the entire space. The determination of the food source X_{ij} within the initial population is determined by the subsequent equation (1) [14].

$$X_{ij} = X_j^{\min} + r \times (X_j^{\max} - X_j^{\min}) \quad \textbf{(1)}$$
$$where \; i = 1, 2, 3.., SN \text{ and } j = 1, 2, 3..., D.$$

Phase of Employed bee

In the phase where bees are actively employed, generating new solutions V_{ij} involves the execution of a random neighbourhood search around X_{ij}. This search is performed using the following equation [14].

$$V_{ij} = X_{ij} + \varphi_{ij} \times (X_{ij} - X_{kj}) \quad (2)$$
$$where \; k = 1, 2.., SN \text{ and } k \neq i$$

On looker bee phase

During this stage, food sources will be selected based on their fitness ratio, which serves as the probability for selection. The calculation for this probability is as follows [14],

$$P_i = \frac{Fitness_i}{\sum_{i=1}^{SN} Fitness_i} \quad (3)$$

$$Fitness_i = \begin{cases} \dfrac{1}{1 + f(X_i)}, if \; f(X_i) \geq 0 \\ 1 + |f(X_i)|, \; else \end{cases} \quad (4)$$

Equations (3) and (4) illustrate that a food source with higher fitness stands a greater chance of being chosen by onlooker bees. Similar to the phase where bees are actively employed, equation (2) is utilized for generating fresh solutions, with the respective counters being adjusted accordingly [14].

Phase of Scout Bee

During this stage, if the counter of a particular solution surpasses a predetermined threshold, the solution is

ISBN 979-8-3503-5094-4

considered unfavourable, prompting the employed bee linked to it to become a scout bee. Further, bee's position will then be randomly determined across the entire space according to equation (1) [15]. In this work, the parameters used for bee optimization algorithms are maximum iteration is 15, the number of scout bees is 10, and neighbourhood radius damp rate is 0.95.

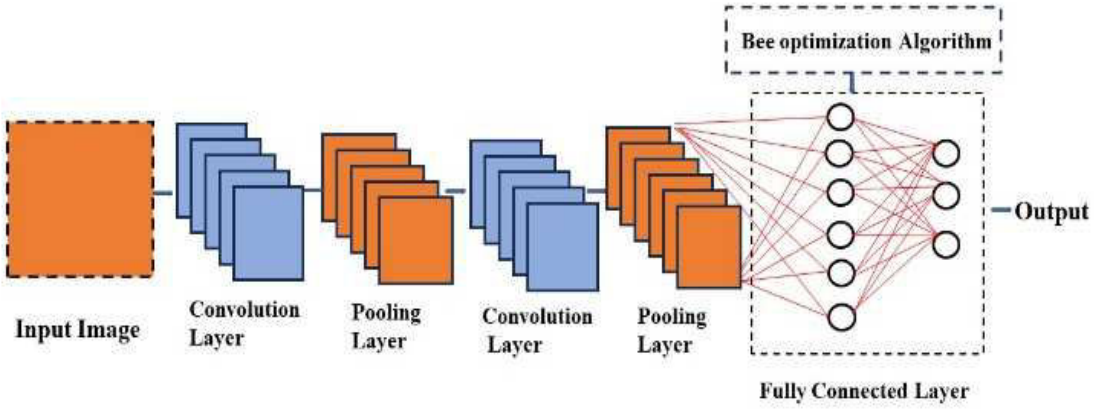


Figure 2 The architecture of Artificial BEE-CNN

C. Artificial Bee Optimization Convolution neural Network (Artificial -BEE-CNN)

CNN is a component of deep learning, resembling the fundamental neural network structure. It is a specialized neural network model designed to process images and videos [16]. By taking the raw pixel data of an image, CNN undergoes training and automatically extracts features, leading to improved classification capabilities. The CNN architecture comprises three convolutions, two pooling, relu activation functions, one dense layer, one softmax layer, and one classification layer. CNN automatically extracts image features, and a fully connected layer is employed to classify them as abnormal or healthy bone. The pooling and convolution layers utilize a 3x3 filter to remove noise and extract image features. Subsequently, the classification process is carried out by the dense layer [17].

During the training phase of a neural network using the gradient descent algorithm, encountering local optima can affect performance. Furthermore, the effectiveness of a trained neural network classifier relies heavily on the initial assignment of weights. To mitigate this issue, the study employs the bee algorithm to select the optimal set of weights from various initial configurations [18]. Specifically, the overall error is utilized as a fitness function for the bee algorithm to determine the weight set with the lowest error. The architecture of the Artificial BEE-CNN is depicted in Figure 2. The Artificial BEE-CNN generates initial solutions corresponding to the weights of filters in convolution and fully connected layers, serving as solutions for the employed bees. The classification error of the CNN classifier assesses the quality of classification and serves as the fitness measure for these solutions. Employed bees disseminate and update recently discovered solutions with onlooker bees. Onlooker bees select solutions with higher fitness, update them, and reassess their fitness. This iterative process continues until the maximum number of iterations is reached [18-19]. The training parameters applied to Artificial BEE-CNN are outlined in Table 1.

Table 1 Artificial BEE-CNN training parameters

Training Parameters	Specification
Number of Epochs	20
Minibatch Size	128
Learning Rate	1.00e-03
Initial Learning rate	0.01
L2-regularization	0.0001
Gradient Threshold method	I2norm

III. Results and Discussion

This section discusses the practical outcomes of utilizing Artificial BEE-CNN to differentiate between normal and Osteoporosis leg X-ray images. The execution and implementation of Artificial BEE-CNN were carried out using the licensed version of MATLAB 2023a.

ISBN 979-8-3503-5094-4

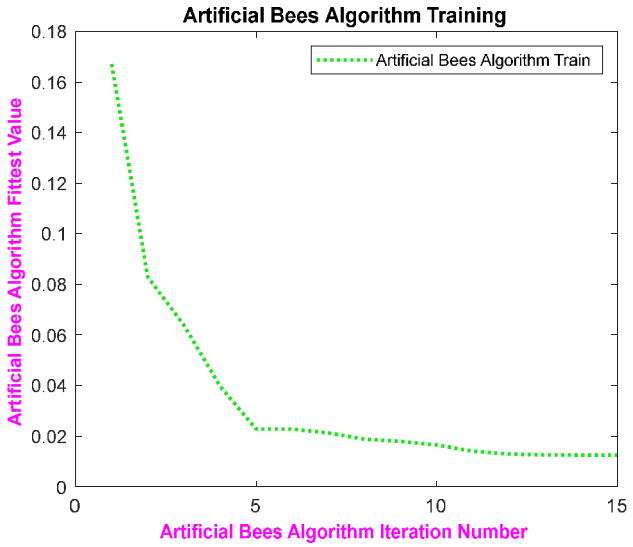
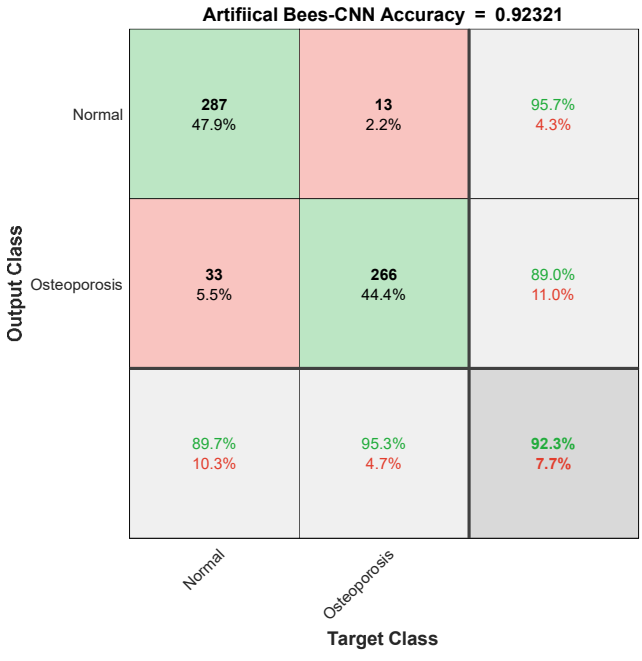
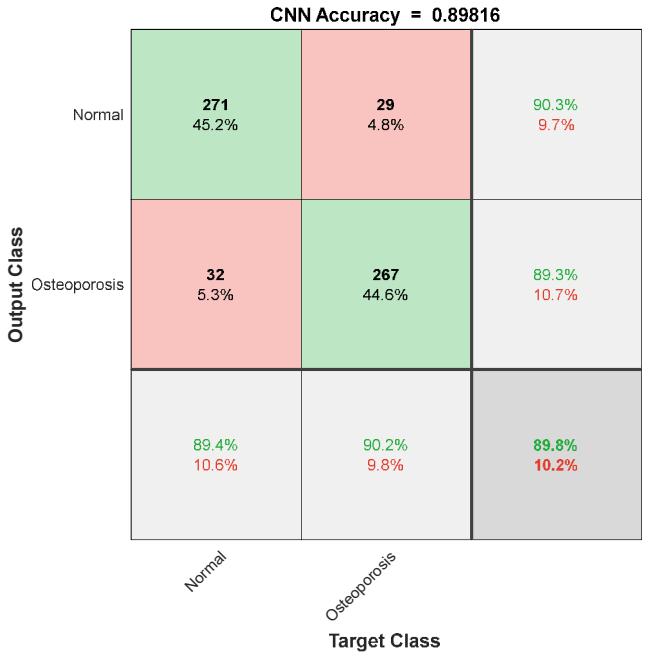


Figure 3 Convergence Curve of Artificial Bee Optimization Algorithm

Normal and osteoporosis leg X-ray images are classified using the Artificial BEE-CNN and the conventional CNN. The model is trained with a specific image size, and subsequent layers relate to a kernel filter size to enhance the learning process. Relu activation functions, population-based batch normalization, and max pooling with a designated pool size are applied. Additionally, artificial bee optimization is integrated into the fully connected layer to update the weights, followed by the softmax layer and classification layer for the osteoporosis leg X-ray image classification. The confusion matrices depicting the maximum accuracy achieved by both the Artificial Bee-CNN and Conventional CNN are shown in Figure 4 (a-b)



(a)



(b)

Figure 4 (a) Artificial Bee-CNN Confusion Matrix and (b) Conventional CNN Confusion Matrix

Table 2 illustrates the comprehensive performance metrics for the Artificial BEE-CNN and the conventional CNN in classifying normal and osteoporosis leg X-ray images. The results indicate that the artificial bee-CNN achieves a significantly higher accuracy (92.32%) than traditional CNN's (89.82%). Further, the sensitivity of Artificial Bee-CNN attains 95.67% when compared to the sensitivity of the traditional CNN (89.82%).

Table 2 The performance matrices of Artificial BEE-CNN and Conventional CNN

Performance Metrices (%)	BEE-CNN	Conventional CNN
Accuracy	92.32	89.82
Sensitivity	95.67	90.33
Specificiaty	88.96	89.30
PPV	89.69	89.44
NPV	95.34	90.20
F1 Score	92.58	89.88

IV. CONCLUSION

With increasing lifespans, osteoporosis is emerging as a significant global health issue. Although, (DXA) is the prevailing diagnostic technique for osteoporosis, its applicability for screening is constrained. This research endeavors to tackle this challenge by utilizing an evolutionary-based deep learning algorithm to forecast osteoporosis through leg X-ray images. A Convolutional Neural Network was enhanced using the

ISBN 979-8-3503-5094-4

artificial bee optimization algorithm to classify typical osteoporosis cases based on the leg X-ray images. The developed Artificial Bee CNN's performance was compared to the conventional CNN's. The results indicate that the CNN and bee optimization achieved a high accuracy rate of 92.32% in effectively classifying typical and osteoporosis leg X-ray images. The integration of artificial bee optimization into the CNN showed superior performance compared to the conventional CNN. According to these discoveries, the Artificial Bee CNN model demonstrates potential as a beneficial screening instrument for straightforward osteoporosis prediction in practical clinical environments.

ACKNOWLEDGEMENT

This study received funding from Science and Engineering Research Board Teacher Associateship for Research Excellence (SERB-TARE) (Grant Number :TARE/2022/000500).

REFERENCES

[1]. Scott, J. C. (1997). Epidemiology of osteoporosis. Journal of clinical rheumatology: practical reports on rheumatic & musculoskeletal diseases, 3(2 Suppl), 9-13.

[2]. Christiansen, C. (1991). Consensus development conference: prophylaxis and treatment of osteoporosis. Am J Med, 90, 107-110.

[3]. Smith, A. D. (2019). Screening of bone density at CT: an overlooked opportunity. Radiology, 291(2), 368-369.

[4]. Yamamoto, N., Sukegawa, S., Kitamura, A., Goto, R., Noda, T., Nakano, K., ... & Ozaki, T. (2020). Deep learning for osteoporosis classification using hip radiographs and patient clinical covariates. Biomolecules, 10(11), 1534.

[5]. Pisani, P., Renna, M. D., Conversano, F., Casciaro, E., Muratore, M., Quarta, E., ... & Casciaro, S. (2013). Screening and early diagnosis of osteoporosis through X-ray and ultrasound based techniques. World journal of radiology, 5(11), 398.

[6]. Smets, J., Shevroja, E., Hügle, T., Leslie, W. D., & Hans, D. (2021). Machine learning solutions for osteoporosis—a review. Journal of bone and mineral research, 36(5), 833-851.

[7]. Damilakis, J., Adams, J. E., Guglielmi, G., & Link, T. M. (2010). Radiation exposure in X-ray-based imaging techniques used in osteoporosis. European radiology, 20, 2707-2714.

[8]. Zhang, B., Yu, K., Ning, Z., Wang, K., Dong, Y., Liu, X., ... & Zhang, S. (2020). Deep learning of lumbar spine X-ray for osteopenia and osteoporosis screening: A multicenter retrospective cohort study. Bone, 140, 115561.

[9]. Nakamoto, T., Taguchi, A., & Kakimoto, N. (2022). Osteoporosis screening support system from panoramic radiographs using deep learning by convolutional neural network. Dentomaxillofacial Radiology, 51(6), 20220135.

[10] Sukegawa, S., Fujimura, A., Taguchi, A., Yamamoto, N., Kitamura, A., Goto, R., ... & Furuki, Y. (2022). Identification of osteoporosis using ensemble deep learning model with panoramic radiographs and clinical covariates. Scientific reports, 12(1), 6088.

[11] Jang, R., Choi, J. H., Kim, N., Chang, J. S., Yoon, P. W., & Kim, C. H. (2021). Prediction of osteoporosis from simple hip radiography using deep learning algorithm. Scientific reports, 11(1), 19997.

[12] Majeed Wani, Insha ; Arora, Sakshi (2021), "Knee X-ray Osteoporosis Database", Mendeley Data, V2, doi: 10.17632/fxjm8fb6mw.2.

[13]. Gao, W., Liu, S., & Huang, L. (2012). A global best artificial bee colony algorithm for global optimization. Journal of Computational and Applied Mathematics, 236(11), 2741-2753.

[14]. Karaboga, D., Gorkemli, B., Ozturk, C., & Karaboga, N. (2014). A comprehensive survey: artificial bee colony (ABC) algorithm and applications. Artificial intelligence review, 42, 21-57.

[15]. Öztürk, Ş., Ahmad, R., & Akhtar, N. (2020). Variants of Artificial Bee Colony algorithm and its applications in medical image processing. Applied soft computing, 97, 106799.

[16]. Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. Advances in neural information processing systems, 25.

[17]. Hu, B., Lu, Z., Li, H., & Chen, Q. (2014). Convolutional neural network architectures for matching natural language sentences. Advances in neural information processing systems, 27.

[18]. Banharnsakun, A. (2019). Towards improving the convolutional neural networks for deep learning using the distributed artificial bee colony method. International journal of machine learning and cybernetics, 10(6), 1301-1311.

[19]. Banharnsakun, A., Achalakul, T., & Sirinaovakul, B. (2010, December). Artificial bee colony algorithm on distributed environments. In 2010 second world congress on nature and biologically inspired computing (NaBIC) (pp. 13-18). IEEE.

AN INTELLIGENT BRAILLE COMMUNICATION USING TRANSLATION GLOVE

Subathra Y.[1]
Assistant Professor,
Rajalakshmi Engineering College,
subathra.y@rajalakshmi.edu.in

Silviya Jebastina A.[2]
Undergraduate Student,
Rajalakshmi Engineering College,
silviyajebastina.a.2019.bme@rajalakshmi.edu.in

Swetha R.[3]
Undergraduate Student,
Rajalakshmi Engineering College,
swetha.r.2019.bme@rajalakshmi.edu.in

Dharanika G R.[4]
Undergraduate Student,
Rajalakshmi Engineering College,
200301015@rajalakshmi.edu.in

Abstract – **This paper introduces an innovative method to enhance communication for deaf-blind individuals by employing an intelligent Braille communication system, which is made feasible with the assistance of a translation glove. The dead-blind population poses significant challenges to communication, and this solution aims to bridge that gap. The wearable glove functions as a communication interface by translating text messages into the Braille alphabet, a widely used system by literate blind people. The design incorporates mini vibrational motors on the palmar side of the glove to enable the perception and interpretation of incoming messages through tactile feedback patterns. Additionally, pushbuttons on the dorsal side of the glove allow users to convey pre-stored messages. Since, the glove is equipped with a GPS module during emergencies, it enables the user to transmit their current location. Compared to existing technologies for visually and hearing-impaired individuals, the proposed method offers superior features such as convenience, affordability, portability, and haptic feedback for effective interpretation of incoming messages. Overall, the proposed system offers a promising new approach to enhancing communication for deaf-blind individuals**.

Keywords: Braille, assistive technology, wearable device, deaf-blind, sensory impairment, haptic feedback.

I. INTRODUCTION

Deaf-blindness is a unique sensory disability that affects individuals who experience both profound hearing loss and complete vision loss or severe visual impairment. This dual sensory loss significantly impacts their ability to communicate, access information, and interact with the world around them. It is estimated by an NGO that there may be more than 5 lakh deafblind adults and children living across 23 states and 110 districts in India. Unlike individuals who are only deaf or blind, deaf-blind individuals face a complex set of challenges that require specialized support and innovative solutions. Typically, a deafblind individual won't be both deaf and blind, but both senses will be compromised to the point where daily life becomes significantly more difficult. As the two senses function together and one would typically assist make up for the loss of the other, these issues can arise even if hearing loss and vision loss are minor. Deaf-blindness can manifest in different ways, encompassing both congenital and acquired cases. Some individuals are born with combined hearing and vision loss, while others may acquire deaf-blindness later in life due to illnesses, injuries, or degenerative conditions. There are instances where individuals gradually lose both their hearing and vision over time, often associated with conditions such as Usher syndrome or other progressive disorders. Additionally, some individuals may experience a dual sensory loss, initially being blind and later acquiring deafness, or vice versa [1]. This population finds it more difficult to engage in society, adapt, access education, etc. due to their combined hearing and vision impairments, which in certain circumstances may also be accompanied by additional physical or intellectual disabilities.

A. Tactile communication

Deafblind individuals heavily depend on tactile communication as they face limitations in visual and auditory interactions [2]. Consequently, a tactile communication interface that offers tactile feedback becomes the preferred choice from this standpoint [3]. It falls under the category of sensory substitution [4] where the devices utilize sensory input in the form of stimulation (vibration, heat, electro cutaneous) [5]. Several technologies have been created in the last ten years for a variety of helpful uses that use sensory substitution [6][7]. They utilize a range of technological techniques [8] and assistive technologies [9][10] incorporating actuators and tactile sensors. Among the various tactile communication methods available, Braille is a widely recognized option. It serves as a tactile approach specifically designed for reading and writing, catering to the needs of individuals who are blind or deafblind [11][12].

B. Braille Cell Unit

Braille code is a universal language that is used by people with visual impairment including people who are blind, deafblind, or who have low vision. It is a tactile alphabet system that involves a combination of dots raised above the surface, arranged in a matrix of 6

ISBN 979-8-3503-5094-4

cells in a 3x2 configuration [13] as shown in Fig. 1. Each cell corresponds to a braille letter, number, or punctuation mark [14][15]. It also has its own single-cell patterns for several commonly used words and letter combinations. There are three distinct types of braille code. Namely Grades 1, 2, and 3.

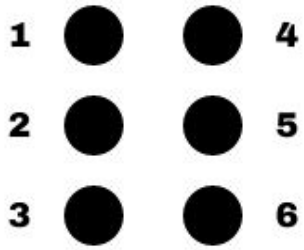


Fig. 1. Braille cell

C. Technical Background

There have been numerous research initiatives aimed at bridging the assistive technology barrier for deafblind communication. Over the last few decades, mechanical hands for automatic fingerspelling and other glove systems with a variety of alphabets have been developed [16]. Some of them have been withdrawn.

One such design is a Wearable Assistive Tactile Communication Interface [17] that incorporates integrated touch sensors and actuators. This innovative design utilizes flexible piezoresistive sensors and vibrotactile actuators, seamlessly integrated to form a tactile communication glove. The glove is equipped with Bluetooth functionality, enabling deafblind individuals to effectively communicate using Braille codes.

Similar to this, a dual-function wearable gadget (Tacsac) [18] utilizes capacitive tactile sensing and integrated tactile feedback to enhance user interaction. The system incorporates a vibrotactile module to interpret messages for the user, while a capacitive touch-detecting layer serves as a touch interface for transmitting messages to the user's mobile phone through a dedicated mobile app installation.

Much similarly, a wearable Malossi alphabet interface for deafblind people has been designed. This interface comprises a wearable peripheral, namely a glove embedded with sensors and actuators, serving as an input/output device. Its primary function is to convert text into sequences of tactile stimuli and vice versa, allowing users to send messages, execute commands, and communicate with other individuals effectively [19].

D. Overview of the proposed solution

As shown in Fig.2 The proposed solution entails a glove equipped with a specialized module, designed to be worn on the left hand of the individual. The glove features vibrational motors positioned on the palm side and pushbuttons on the back side of the hand. This setup enables the system to receive incoming messages through GSM technology and interpret them accordingly.

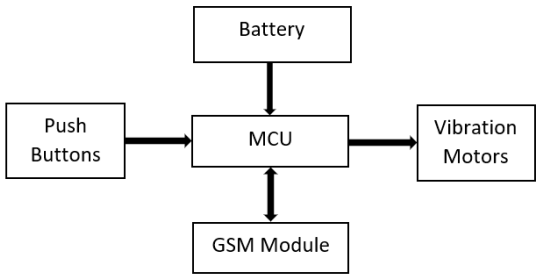


Fig.2. Overview of the system

The system can read incoming messages through GSM. The Arduino Nano MCU then decodes the alphabet and converts it to braille code. The interpreted alphabet is presented to the user as a vibrotactile feedback pattern, which they may read and comprehend. In this device, a Grade 1 Braille coding pattern is featured. It comprises 26 basic alphabet letters and punctuation. It is primarily used by people who have barely started reading Braille. Pre-defined messages are stored in the microcontroller. They are activated when the relevant push buttons are pressed.

II. METHODOLOGY

Hardware implementation consists of a rechargeable 4V Li-Ion battery which has a nominal voltage of 3.7V. This battery is boosted by the HW-357 battery charging module, charged via USB type-C port. The main control unit of the system is ATmega328P microcontroller which interacts with both input and output units. The SIM800L GSM module is the important component that is used to transmit and receive SMS (Short Message Service). The vibration motors and push buttons are connected to MCU. The Arduino is programmed using C embedded programming language.

A. Vibration Motors

This prototype involves six mini-vibration motors that rotate at a speed of 1100 RPM. These motors are connected to the digital output pins of the MCU through MOSFET drivers. The MOSFET drivers are used for high-speed switching applications. The messages are received by the GSM module, which is converted to the according braille alphabets in the microcontroller.

The haptic feedback provided by the vibration motors will allow the individual to perceive the message without the need for any additional visual aids. The user will then be able to read the message by experiencing the vibrating motors and interpreting the Braille characters. Here the haptic feedback is sensed as vibrotactile feedback.

ISBN 979-8-3503-5094-4

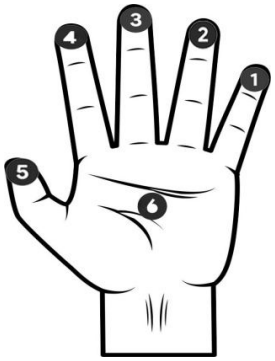


Fig.3. Placement of vibration motors in the glove (Palmar side)

B. Vibrotactile Feedback

In vibrotactile feedback, sounds are perceived. A substance moving back and forth produces sound, which is a type of mechanical energy. As opposed to what we typically imagine, sound waves can move through any substance, including human skin. Humans have a particular kind of skin receptor called a Pacinian Corpuscle that is primarily aimed to pick up these vibrations. Devices that provide vibrotactile feedback have one or more actuators that make skin contact. The receptive fields of skin receptors that detect vibrations are substantially larger than those that detect tactile stimuli. In comparison to tactile feedback devices, vibrotactile devices dramatically reduce the need for actuators.

C. Push buttons

The user can send pre-stored messages when they want to communicate with others. They are also called tactile switches. A tactile switch click activates these pre-stored messages. This prototype comes with four push buttons which are connected to the analog input pins of the MCU. They are pre-programmed with frequent messages. One of the push buttons is implemented to communicate the live location and the aid message in the event of an emergency.

Fig.4. Placement of pushbuttons in the glove

(Dorsal side)

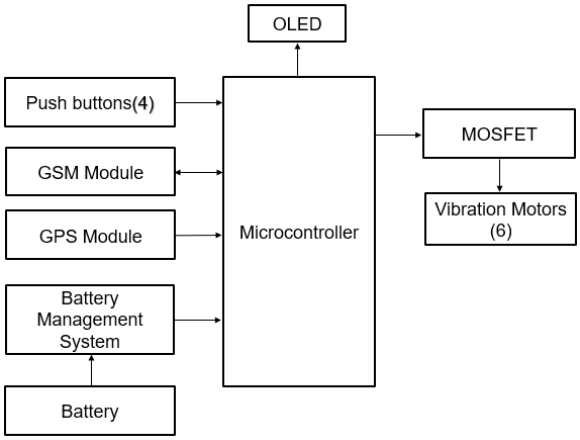


Fig.5. Block diagram of the system

III. RESULTS AND DISCUSSION

As shown in Fig.6 and Fig.7, the hardware is placed in the gloves. The prototype performs accordingly as mentioned in the flowcharts (Fig.10). When the module is activated, the GSM is initialized. The incoming messages from other people have been interpreted accordingly via the vibrotactile feedback pattern by the user. Since vibration motors are localized at each finger, the vibrations from them are easily recognized. Similarly, the pre-stored messages sent by the user using push buttons are successfully received by the pre-programmed number/caretaker. The screenshot of the received sample messages is shown below in Fig.8 and Fig.9

Fig.6. Final Prototype

(Dorsal side of the glove)

ISBN 979-8-3503-5094-4

Fig.7. Final Prototype (Palmer side of the glove)

Fig.8. Message received by the user

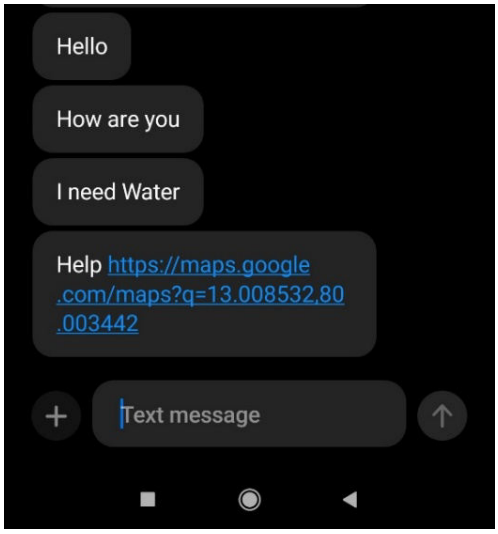


Fig.9. Message sent by the user

TABLE I. The Active and Inactive Vibration Motors During the Incoming Sample Message

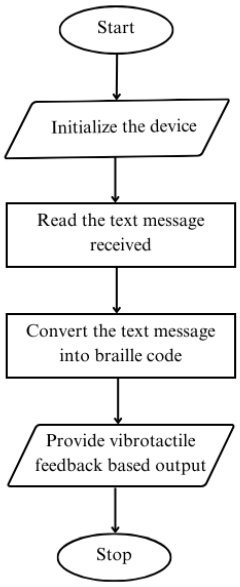
ALPHABET	ACTIVE MOTORS	INACTIVE MOTORS
S	M2, M3&M4	M1, M5&M6
A	M1	M2, M3, M4, M5&M6
T	M2, M3, M4&M5	M1&M6
U	M1, M3&M6	M2, M4&M5
R	M1, M2, M3&M5	M4&M6
D	M1, M4&M5	M2, M3&M6
A	M1	M2, M3, M4, M5&M6
Y	M1, M3, M4, M5&M6	M2

IV. CONCLUSION AND FUTURE WORK

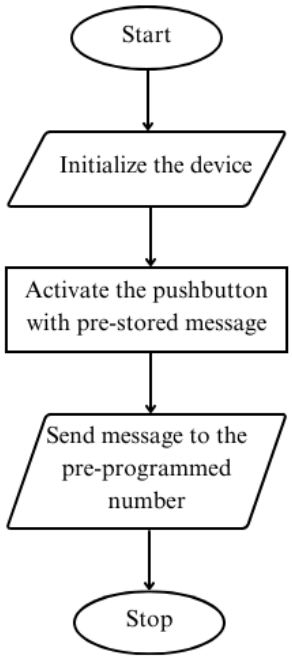
The braille interface enables a haptic feedback system for visually and hearing-impaired individuals to read incoming text messages via vibration motors. The use of open-source hardware components such as the Arduino Nano and GSM modules makes the design more affordable and available to a larger audience. Since it is simple to use and explore, the braille pad provides a user-friendly interface for hearing and visually impaired people. This initiative promotes social inclusion by allowing visually impaired people to communicate with their family and friends via text messages. The proposed design is more beneficial in comparison to the existing methods and can be developed into a marketable product readily. Further, this design can be fabricated by integrating the hardware module into the glove which makes it more aesthetic and wearable. The system can be improved to include a Braille translator, converting printed text to Braille and enabling users to read books, newspapers, and other printed items.

APPENDIX

Fig.10. Flowcharts of two functions



ISBN 979-8-3503-5094-4



REFERENCES

[1] Silvia Gabarro-Lopez and Johanna Mesch (2020)*"Conveying Environmental Information to Deafblind People: A Study of Tactile Sign Language Interpreting"*, Front. Educ., Sec. Educational Psychology, Volume 5|

[2] F. Sorgini, R. Caliò, M. C. Carrozza and C. M. Oddo, *"Haptic-assistive technologies for audition and vision sensory disabilities", Disab. Rehabil. Assistive Technol.,* vol. 13, no. 4, pp. 394-421, May 2018

[3] S. Luo, J. Bimbo, R. Dahiya and H. Liu, *"Robotic tactile perception of object properties: A review", Mechatronics*, vol. 48, pp. 54-67, Dec. 2017.

[4] P. Bach-Y-Rita, *"Tactile sensory substitution studies", Ann. New York Acad. Sci.,* vol. 1013, no. 1, pp. 83-91, 2004.

[5] L. Cancar, A. Díaz, A. Barrientos, D. Travieso, and D. M. Jacobs, *"Tactile-sight: A sensory substitution device based on distance-related vibrotactile flow", Int. J. Adv. Robotic Syst.,* vol. 10, no. 6, pp. 272, Jun. 2013.

[6] D. Dakopoulos and N. G. Bourbakis, *"Wearable obstacle avoidance electronic travel aids for blind: A survey", IEEE Trans. Syst. Man Cybern. C Appl. Rev.,* vol. 40, no. 1, pp. 25-35, Jan. 2010.

[7] S. K. Nagel, C. Carl, T. Kringe, R. Märtin, and P. König, *"Beyond sensory substitution-learning the sixth sense", J. Neural Eng.,* vol. 2, no. 4, pp. R13, 2005.

[8] N. Yogeswaran et al., *"Piezoelectric graphene field effect transistor pressure sensors for tactile sensing",* Appl. Phys. Lett., vol. 113, no. 1, Jul. 2018.

[9] W. T. Navaraj, H. Nassar and R. Dahiya, *"Prosthetic hand with biomimetic tactile sensing and force feedback", Proc. IEEE Int. Symp. Circuits Syst. (ISCAS)*, pp. 1-4, May 2019.

[10] O. Ozioko, M. Hersh and R. Dahiya, *"Inductance-based flexible pressure sensor for assistive gloves", Proc. IEEE SENSORS,* pp. 1-4, Oct. 2018.

[11] A. Alfadhel, M. A. Khan, S. Cardoso de Freitas and J. Kosel, *"Magnetic tactile sensor for braille reading", IEEE Sensors* J., vol. 16, no. 24, pp. 8700-8705, Dec. 2016.

[12] A. Russomanno, S. O'Modhrain, R. B. Gillespie and M. W. M. Rodger, *"Refreshing refreshable braille displays", IEEE Trans. Haptics,* vol. 8, no. 3, pp. 287-297, Jul. 2015.

[13] T. Choudhary, S. Kulkarni, and P. Reddy, *"A Braille-based mobile communication and translation glove for deaf-blind people,"* 2015 International Conference on Pervasive Computing (ICPC), Pune, India, 2015, pp. 1-4,

[14] K. Fukuda et al., *"A 4 V operation flexible braille display using organic transistors carbon nanotube actuators and organic static random-access memory", Adv. Funct. Mater.,* vol. 21, no. 21, pp. 4019-4027, 2011.

[15] H.-J. Kwon, S. W. Lee, and S. S. Lee, *"Braille dot display module with a PDMS membrane driven by a thermopneumatic actuator", Sens. Actuators A Phys.,* vol. 154, no. 2, pp. 238-246, Sep. 2009.

[16] Hersh, M.A., Johnson, M.A. (eds) (2003). *Hearing-aid Principles and Technology. In: Assistive Technology for the Hearing-impaired, Deaf, and Deafblind. Springer,* London.

[17] O. Ozioko, P. Karipoth, M. Hersh and R. Dahiya, June 2020, *IEEE Transactions on Neural Systems and Rehabilitation Engineering,* vol. 28, no.6,pp.1344-1352, doi:10.1109/TNSRE.2020.2986222

[18] O. Ozioko, W. Navaraj, M. Hersh, and R. Dahiya, *"Tacsac: A Wearable Haptic Device with Capacitive Touch-Sensing Capability for Tactile Display," Sensors*, vol. 20, no. 17, p. 4780, Aug. 2020, doi: 10.3390/s20174780.

[19] Nicholas Caporusso – *AVI'08: Proceedings of the working conference on Advanced visual interfaces* May 2008 Pages 445-448

ISBN 979-8-3503-5094-4

Swift Fingerprint Enhancement Technique with Parabolic Mask Filtering Algorithm

R.Seetharaman
Department of ECE
CEG Campus, Anna University
Chennai-600025 India
rseetharamanece@annauniv.edu

K.Soumya
Department of ECE
SRMIST-*Ramapuram Campus*
Chennai-600089 India

D.Nedumaran
CISL
University of Madras
Chennai-600025 India

M.Selvakumar
CISL
University of Madras
Chennai-600025 India

C.Ambhika
Department of AIML
RMD Engineering College
Chennai-601206 India

R.Jeffery
Department of ECE
SRMIST-Ramapuram Campus
Chennai-600089 India

L.Harihara Subramanian
Department of Management Studies
CEG Campus, Anna University
Chennai-600025 India

Abstract: **The aim of fingerprint image enhancement is to enhance image quality by eliminating noise while preserving essential minutiae points. This study introduces a rapid fingerprint enhancement algorithm that combines local ridge orientation estimation, local frequency estimation, and parabolic mask filtering techniques. The algorithm underwent testing on diverse fingerprint images, with findings demonstrating its effectiveness in significantly improving raw fingerprint images, particularly those of low quality, within a shorter execution time.**

Keywords—Fingerprint Biometrics, Enhancement, Orientation, frequency, parabolic mask, Binarization, performance evaluation

I. INTRODUCTION

Biometrics stands as a well-established cornerstone in the realm of authentication and recognition. Among the various biometric techniques available, fingerprint biometrics reigns supreme as the most straightforward and dependable means of authentication. Its widespread acceptance is attributed to its remarkable attributes, including high recognition rates, cost-effectiveness, compactness, minimal variance, portability, and seamless integration into mobile devices. Consequently, fingerprint biometrics finds extensive application across a multitude of high and medium-security authentication systems.

In a fingerprint identification system, the matching process hinges heavily on the reliability of minutiae points, which encompass ridge bifurcations and ridge terminations. The accuracy of the fingerprint identification system is predominantly contingent on the quality of the images captured by these devices, as high-quality images significantly reduce the necessity for additional image processing techniques in biometric systems. However, consistently obtaining high-quality images in real-time scenarios presents a formidable challenge. For instance, as depicted in Figure 1, a low-quality fingerprint image shows indistinct valleys and ridges with discontinuities in the ridge flow patterns. To tackle these challenges, various image enhancement techniques have been explored and documented in the literature [1-7].

Low-quality fingerprint images pose significant challenges, primarily characterized by blurriness and unclear spatial details resulting from suboptimal capturing or enrollment procedures [8]. Despite the presence of built-in enhancement algorithms in some capturing devices designed to prevent the need for repeated enrollment [9], challenges persist. Additionally, FPGA/DSP-based systems have incorporated intricate enhancement algorithms into their hardware platforms [10]. Nevertheless, the issues associated with low-quality fingerprints remain unresolved, leaving ample room for advancements in both hardware and software domains.

Numerous techniques have been explored to enhance the quality of fingerprint images. These include methods like Histogram Equalization [11], Fourier Transform [12], Gabor Filters [13-15], and Short-Time Fourier Transform (STFT) [16, 17]. Additionally, adaptive fingerprint image enhancement approaches, such as those based on image clustering, feature extraction, and characteristics [18, 19], have been investigated. Hsieh et al. [20] introduced a local orientation and global texture-based wavelet algorithm to improve the clarity and continuity of ridge patterns. Gottschlich et al. [21] presented curved Gabor filters for enhancing low-quality fingerprints, utilizing estimated orientation and ridge frequencies. Khan and Khan [22] employed a Directional Filter Bank (DDFB) without decimation to decompose input fingerprint images into directional components. They subsequently applied Principal Component Analysis (PCA) to normalize these directional fingerprint images, using them to reconstruct enhanced images. Another approach involved Morlet wavelet-based image enhancement to enhance ridge continuity and eliminate spurious minutiae and gaps, resulting in improved authentication system accuracy [23]. Furthermore, fuzzy logic-based enhancement techniques were explored to refine the clarity of low-quality fingerprint images.

ISBN 979-8-3503-5094-4

Fig. 1 Sample low quality fingerprint images from FVC databases

II. Related Work

Enhancing low-quality images poses a formidable challenge, mainly due to the complexities involved in identifying ridge-occupied regions for enhancement while excluding irrelevant areas. The use of orientation estimation techniques becomes particularly valuable in this context, as they effectively eliminate undesired image regions from the enhancement process, thereby enhancing processing speed. Low-quality images, characterized by attributes such as oiliness, wetness, moisture, scars, low contrast, and high contrast, present unique difficulties in the enhancement process. Hence, this work undertakes the development of an effective enhancement algorithm, leveraging orientation mapping and local frequency estimation, to achieve optimal results in enhancing low-quality fingerprint images.

This paper is structured into the following sections:

Section III provides an in-depth explanation of the proposed fingerprint enhancement method. Section IV demonstrates the application of the proposed algorithm to fingerprint images. Section V presents the experimental results, including a performance comparison across FVC databases. Finally, Section VI offers concluding remarks summarizing the findings of this study.

III. Enhancement Techniques

The proposed fingerprint enhancement methodology is based on three key steps: local ridge orientation estimation, local ridge frequency estimation, and the application of a parabolic mask filtering enhancement process. In this study, we employ advanced orientation and frequency estimation techniques to formulate the parabolic mask filtering enhancement algorithm.

A. Local Ridge Orientation Estimation

Fingerprint orientation mapping entails creating a 2D orientation field that visualizes the directional flow patterns of the fingerprint ridges. In this context, a modified polar complex moment (MPCMs) technique is employed to derive the directional mapping of ridge orientation.

In MPCMs, significant local ridge orientations are depicted using the phase angle (θ) in conjunction with grayscale information (magnitude) within regularly spaced image blocks. To address periodicity and discontinuity issues, the phase angle is doubled to construct the vector field for each orientation value. Equation 1 is utilized to generate the orientation field O, which is derived from a complex-valued orientation field.

$$e^{i(2\theta)} = \cos(2\theta) + i\sin(2\theta)$$

$$O(k,l) = e^{i(2\theta)} = \cos(2\theta) + i\sin(2\theta) \quad (1)$$

Now, a discrete complex function is employed in a new domain defined by xk and yl, where xk = (k - K/2) / (K/2) and yl = (l - L/2) / (L/2). This function is utilized to calculate the translation-corrected orientation field O(k, l), where k ranges from 0 to K-1, and l ranges from 0 to L-1. Consequently, the computation of discrete polar harmonics transform (PHT) coefficients is carried out as follows:

$$\mathrm{M}_{p,q} = \frac{n}{KL}\sum_{k=0}^{K}\sum_{l=0}^{L}\psi_{p,q}(\mathrm{x}_k, \mathrm{y}_k)\, O(k,l)w(k,l) \quad (2)$$

where $\psi_{p,q}(\mathrm{x}_k, \mathrm{y}_k)$ is the basis function, *w(k, l)* is the segmented matrix element, *n* is the integer chosen based on the ridge thickness (frequencies) and image intensity $f(x, y)$. Combining MPCMs with the basis function yields a vector field composed of complex values, as represented in Equation (3).

$$f(r,\varphi) = \sum_{p=-\infty}^{\infty}\sum_{q=-\infty}^{\infty}\mathrm{M}_{p,q}[R'_p(r)\Phi^*_q(\varphi)] \quad (3)$$

In Equation (3), the values of *p* and *q* typically vary to formulate the PCMs of different orders. However, in the current study, we maintain constant values for *p* and *q*, while varying the value of *n* in relation to higher-order moments at higher frequencies. This approach is employed to achieve orientation estimation for low-quality fingerprint images, as depicted in Figure 2.

(a) (b) (c)

Fig. 2 (a) Original image, (b) orientation mapped over the image and (c) orientation mapped image.

B. Local Ridge frequency Estimation

To estimate the local-ridge distance, we first calculate the local orientation of the fingerprint ridges. The local frequency estimation is then determined using the MPCMs method, following these algorithmic steps:

Step 1:

To compute the individual pixel values of a fingerprint image $I(i, j)$, consider that the local orientation $f(r,\varphi)$ within blocks $w \times w$ is centered at pixel $I(i, j)$. In this context, $f(r,\varphi)$ denotes the orthogonal direction relative to the dominant direction $L(i, j)$ of the Fourier spectrum, with a window size of $w \times w$.

Step 2:

In this step, we determine the influential direction (line), denoted as $L(i, j)$, which is orthogonal to the local orientation $f(r,\varphi)$ at each individual pixel point in the

ISBN 979-8-3503-5094-4

fingerprint image. Each $I(i,j)$ value is composed of a 40-pixel neighborhood centered on the given point. Within each *L(i, j)*, specific pixels are responsible for identifying the local minimum points B_r (where $r = 1,2,3,.....k-1$ corresponding to fingerprint ridges) and local maximum points (where $v = 1,2,3,.....\gamma-1$ corresponding to valleys). Based on this assumption, we utilize k and γ to sequentially determine the number of ridges and valleys in the fingerprint images, as expressed by Equations (4) and (5)

$$V_r = B_r - B_{r+1} \; r = 1,2,3,.....k-1 \tag{4}$$

$$R_v = T_v - T_{v+1} \; v = 1,2,3,.....\gamma-1 \tag{5}$$

Step 3:

The estimation of ridge frequencies within fingerprint images involves measuring the distances between appropriately selected points. In this stage, we compute two sequences of distances for each line, denoted as *L*. Based on these criteria, we explore the pixels along line *L*, leading to the reordering of certain elements within the sequence. Let $\varepsilon(j) = \sum_{i-1}^{p-1} x_j - x_i)$,, where *x* represents the pixels along the chosen line *L*. Subsequently, we arrange the elements (B_r T_v - pixels) of line *L* in a suitable order, following the mathematical procedure outlined below.

$$\forall_{j=1}^{p-1} \left|\varepsilon(j)\right| > \left|\varepsilon(j) + x_{j+1}\right| = \begin{cases} \text{then } cond_1 \\ \text{otherwise } j = j+1 \end{cases} \tag{6}$$

Here, we have $cond_1$, implying that $x_j = x_j + x_{j+1}$ and $x_j + 1$ are excluded from the equation, $p = k \; or \; p = \gamma$ is utilized to solve Equation (6) with respect to the selected sequence. The frequency of the local fingerprint for line *L*, centered at point $I(i,j)$ and with the least-squares estimation of the local ridge orientation $\theta(i,j)$, can be directly computed using Equation (7).

$$ds(i,j) = \frac{\sum_{v=1}^{\gamma} R_v + \sum_{r=1}^{k} V_r}{\gamma + k + 2} \tag{7}$$

Step 4:

To mitigate noise in low-quality images, the frequency smoothing step ($w \times w$ is the size of the window) is employed to compute reliable frequency values. Ultimately, the total ridge frequency is determined using Equation (8).

$$f(i,j) = \frac{\sum_{u=i-w/2}^{i+w/2} \sum_{v=j-w/2}^{j+w/2} ds(u,v)}{\omega^2} \tag{8}$$

C. Parabolic Mask Filtering

he two aforementioned processes, namely local orientation estimation and local frequency estimation methods, are conducted to facilitate the parabolic mask filtering enhancement process while preserving the inherent characteristics of the fingerprint ridges. Leveraging the values obtained from local ridge orientation mapping $f(r,\varphi)$ and local frequency estimation $f(i,j)$, we enhance the overall enhancement procedure $I(i,j)$.

Following the estimation of the local orientation and frequency of the ridges, a mask with a size of 3×9 is applied to the ridge direction. This mask introduces a directional low-pass filter effect, with its coefficients derived from a parabolic curve. The mask's formation is determined through the following steps: The *0th* row of the mask represents a concave structure ($y = x^2$), while the other two rows (-1 and 1) feature convex structures ($y = -x^2$) in the form of parabolic curves.

As a result, there is a strong correlation among the pixels in the middle row, while the correlation with the neighboring pixels is lower. Likewise, the pixel of interest exhibits a higher correlation with the adjacent sections of the first and last rows of the mask, in comparison to the other two rows. This means that if there are minor discontinuities along the central ridge direction, they are mitigated, resulting in a smoothing effect on the ridge, as illustrated below.

$$j' = j + (u\cos\theta + v\sin\theta)$$
$$u = -4,-3,....,3,4 \; , \quad v = -1,0,1 \tag{9}$$

$$i' = i + (-u\cos\theta + v\sin\theta)$$
$$u = -4,-3,....,3,4 \; , \quad v = -1,0,1 \tag{10}$$

Equations (9) and (10) encompass the functionality of applying the mask to low-quality fingerprint images, with the pixel situated at the ridge direction of the mask calculated as per Equation (12).

$$K = \sum_{u=-1}^{1} \sum_{v=-4}^{4} W(u,v) \; f(r,\varphi) \; f(i,j) \tag{11}$$

$$I(i,j) = \frac{1}{K} \sum_{u=-1}^{1} \sum_{v=-4}^{4} W(u,v) I(i',j') \tag{12}$$

D. Binarization

Here, a binarization operation is employed for a visual evaluation of the enhancement process. In this stage, the threshold is determined by taking the mean value of the image blocks. The fingerprint image is divided into 9x9 blocks, and the mean grayscale value ($I(i,j)$ is computed using the following equation.

$$I_{LM}(i,j) = \frac{1}{w * w} \sum_{u=-w/2}^{w/2} \sum_{v=-w/2}^{w/2} I(u,v) \tag{13}$$

If $I(i,j) > I_{LM}(i,j)$, then the pixel is assigned the color white; otherwise, it is labeled as black.

ISBN 979-8-3503-5094-4

$$C = \begin{cases} 255 & \text{(White)} \quad if \ I(i,j) > I_{LM}(i,j) \\ 0 & \text{(Black)} \quad \text{otherwise} \end{cases} \tag{14}$$

It automatically transforms the fingerprint images into a bi-level format, with the foreground region represented by black pixels and the background by white pixels, as illustrated in Figures 3(c), 4(c), 5(c), and 6(c).

IV. Implementation of Proposed Method in FVC Database Images

In the FVC Database, low-quality fingerprint images were selected to assess the performance of the rapid enhancement algorithm. The results were evaluated through visual inspection of the binarized output of the enhanced images. Figures 3(a) and 4(a) depict the original low-contrast images obtained from FVC 2004 DB2_A 33_5.tif and FVC 2004 DB2_A 34_1.tif, respectively. Likewise, Figures 5(a) and 6(a) showcase other low-quality images with oiliness and shading, sourced from FVC 2004 DB2_A 34_7.tif and FVC 2002 DB2_A 107_5.tif, respectively.

Figures 3(b), 4(b), 5(b), and 6(b) showcase the fingerprint images after applying the parabolic mask filtering enhancement. This process effectively bridges the discontinuities in the fingerprint ridges throughout the image, especially in ridge-occupied regions, while excluding the background regions. This enhancement is achieved through the utilization of local orientation and local frequency estimation techniques. Furthermore, Figures 3(c), 4(c), 5(c), and 6(c) display the resulting binarized images, which demonstrate improved visual clarity due to the performance of the parabolic mask filter.

The primary objective of all pre-processing techniques is to facilitate accurate matching, particularly in terms of high-quality minutiae points. While many algorithms excel at matching high-quality images, the task of matching low-quality images remains a persistent challenge. It's worth noting that in the FVC2000 database, approximately 20% of the issues related to low-quality images are responsible for approximately 80% of the false non-match results.

Similarly, in the case of FVC2002 and FVC2004 databases, approximately 10% of the issues associated with low-quality images account for approximately 60% of the false non-match results. The application of the proposed enhancement algorithm has notably improved the performance ratio, as evidenced by the computed performance metrics across the FVC databases.

V. Experimental Results of Perfomance Metrics

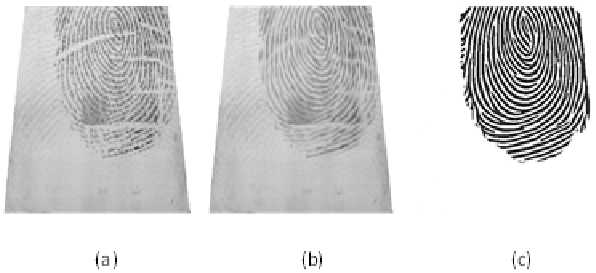


Fig. 3 (a) Original low contrast FVC 2004 DB2_A 33_5.tif, (b) Enhanced by parabolic mask filtering and (C) the binarized fingerprint image for visual examination

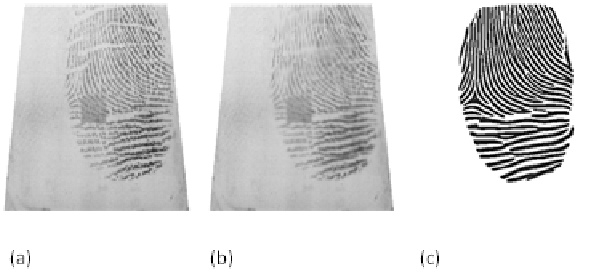


Fig. 4 (a) Original low contrast FVC 2004 DB2_A 34_1.tif, (b) Enhanced by parabolic mask filtering technique and (C) the binarized fingerprint image for visual examination

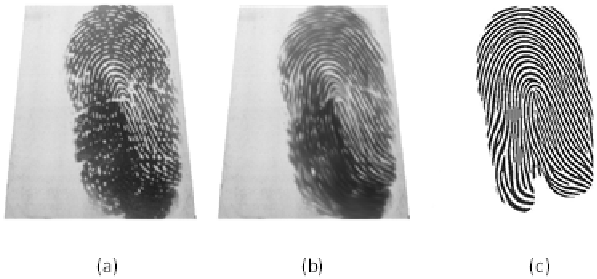
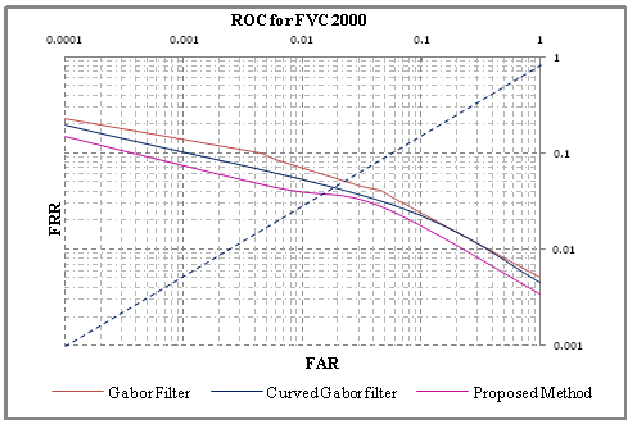
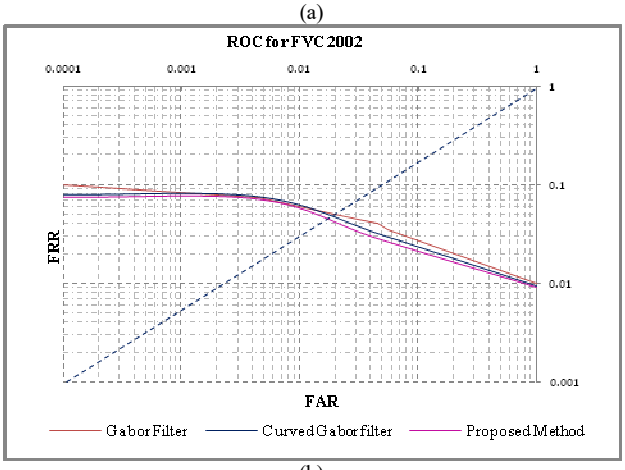


Fig. 5 (a) Original low contrast FVC 2004 DB2_A 34_7.tif, (b) Enhanced by parabolic mask filtering technique and (C) the binarized fingerprint image for visual examination

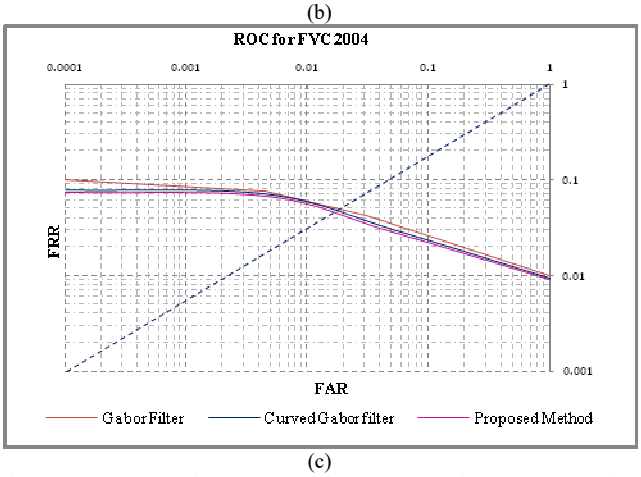
ISBN 979-8-3503-5094-4



(a)



(b)



(c)

Fig. 6 (a) ROC Curve for FVC 2000, (b) ROC Curve for FVC 2002 and (c) ROC Curve for FVC 2004

The evaluation of the biometric-based identification system's performance involves the graphical presentation of Receiver Operating Characteristic (ROC) curves. These curves are derived from the false rejection rate (FRR) and false acceptance rate (FAR) of the respective FVC 2000, FVC 2002, and FVC 2004 databases, as illustrated in Figure 6. In this context, the results obtained through the proposed method are compared with those obtained using traditional Gabor filter and Curved Gabor filter methods.

Another important parameter used to assess the performance of verification or identification system is the Equal Error Rate (EER). The EER represents the point where the false acceptance rate (FAR) equals the false rejection rate (FRR). In the context of fingerprint matching analysis, an average EER ranging from 2% to 3.5% is considered indicative of good matching performance. For the FVC 2000 database, the proposed algorithm achieves an EER of 2.82%, while the Gabor filter and Curved Gabor filter methods yield EER values of 3.92% and 3.43%, respectively. Similarly, for the FVC 2002 database, the EER of the proposed algorithm is 2.65%, whereas EER values of 3.00% and 2.85% are observed for the Gabor filter and Curved Gabor filter methods, respectively. This trend persists in the case of the FVC 2004 database, with EER values of 2.62% for the proposed algorithm, 2.85% for the Gabor filter method, and 2.70% for the Curved Gabor filter.

In summary, the average EER values for the Gabor filter (3.25%) and Curved Gabor filter (2.99%) methods are higher than that of the proposed Modified Polar Complex Moments (MPCM) approach (2.70%), as detailed in Table 1

TABLE 1 Equal error rate (EER) for all existed method with FVC database

Methods	EER for FVC 2000 Database	EER for FVC 2002 Database	EER for FVC 2004 Database	Average EER
Gabor Filter	3.92%	3.00%	2.85%	3.25%
Curved Gabor filter	3.43%	2.85%	2.70%	2.99%
Proposed method	2.82%	2.65%	2.62%	2.70%

In summary, Gabor filter-based techniques excel in enhancing normal-quality fingerprint images by effectively correcting errors present in the images. However, they have limitations, particularly with respect to the values of the vector K, making them less suitable for low-quality fingerprint images. On the other hand, our proposed parabolic method, coupled with advanced MPCMs-based orientation techniques, significantly improves the quality of low-quality fingerprint images and efficiently identifies a maximum number of minutiae points in a short duration, typically less than 300 ms. The algorithm, which encompasses local orientation estimation, local frequency estimation, parabolic mask filtering, binarization, and performance metrics estimation, has been meticulously developed and rigorously tested using the C++ builder compiler on an Intel Xeon Processor (3.2 GHz) based Workstation Computer with 2GB of RAM.

VI. CONCLUSION

We have introduced a rapid fingerprint image enhancement algorithm that combines efficient orientation estimation, local frequency estimation, and parabolic mask filtering techniques. We have also compared the outcomes with a well-established technique, namely the Gabor-based enhancement method. The experimental results obtained from this study unequivocally demonstrate the superior enhancement capabilities of the proposed method, particularly for low-quality fingerprint images. This is substantiated by the performance metrics and visual inspection results.

REFERENCES

[1] Shams H, Jan T, Khalil AA, Ahmad N, Munir A, Khalil RA. 2023. Fingerprint image enhancement using multiple filters. PeerJ Comput. Sci. 9:e1183http://doi.org/10.7717/peerj-cs.1183

[2] Ram Prakash Sharma, and Somnath Dey, Two-stage quality adaptive fingerprint image enhancement using Fuzzy C-means clustering based fingerprint quality analysis, Image and Vision Computing, Vol. 83

ISBN 979-8-3503-5094-4

No. C, Mar 2019, pp 1–16https://doi.org/10.1016/j.imavis.2019.02.006

[3] Erwin et al, The Enhancement of Fingerprint Images using Gabor Filter 2019,J. Phys.: Conf. Ser. 1196 012045doi:10.1088/1742-6596/1196/1/012045

[4] T. M. Khan, M. A.U. Khan and Y. Kong, "Fingerprint image enhancement using multi-scale DDFB based diffusion filters and modified Hong filters," Optik, Vol. 125, 2014, pp. 4206–4214.

[5] D. Ezhilmaran and M. Adhiyaman, A Review Study on Fingerprint Image Enhancement Techniques, International Journal of Computer Science & Engineering Technology (IJCSET), Vol. 5 No. 06 Jun 2014, pp. 625-631, ISSN : 2229-3345.

[6] S. K. Oh, J. J. Lee, C. H. Park, B. S. Kim and K. H. Park, "New Fingerprint Image Enhancement Using Directional Filter Bank," Journal of WSCG, Vol.11, No.1, 2003.

[7] J. Yang, N. Xiong, and A. V. Vasilakos, "Two-Stage Enhancement Scheme for Low-Quality Fingerprint Images by Learning From the Images," IEEE Transactions on Human-Machine Systems, Vol. 43, No. 2, March 2013, pp. 235-248.

[8] E. Blotta and E. Moler, "Fingerprint image enhancement by differential hysteresis processing," Forensic Science International, Vol. 141, 2004, pp. 109-113

[9] S. Shigematsu, H. Morimura, Y. Tanabe, T. Adachi, and K. Machida, "A Single-Chip Fingerprint Sensor and Identifier," IEEE Journal of Solid-State Circuits, Vol. 34, No. 12, December 1999, pp. 1852-1859.

[10] M. Lopez, E. Canto and M. Fons, "Hardware-Software Co-design of a Fingerprint Image Enhancement Algorithm," IEEE 32nd Annual Conference on Industrial Electronics, IECON, 2006, pp. 3496-3501.

[11] S. Greenberg, M. Aladjem and D. Kogan, "Fingerprint Image Enhancement using Filtering Techniques," Real-Time Imaging, Vol. 8, pp. 227–236, 2002.

[12] S. Tarar and E. Kumar, "Fingerprint Image Enhancement: Iterative Fast Fourier Transform Algorithm and Performance Evaluation," International Journal of Hybrid Information Technology, Vol. 6, No. 4, July, 2013, pp. 11-20.

[13] J. Wang and X. Sun, "Fingerprint Image Enhancement using a fast Gabor filter," IEEE Proceedings of the 8th World Congress on Intelligent Control and Automation, 2010, pp.6347-6350.

[14] L. Hong, . Y. Wan and A. Jain, **"**Fingerprint Image Enhancement: Algorithm and Performance Evaluation," IEEE Transactions On Pattern Analysis and Machine Intelligence, Vol. 20, No. 8, August 1998, pp. 777-789.

[15] J. Yang, L. Liu, Ti. Jiang and Y. Fan, "A modified Gabor filter design method for fingerprint image enhancement," Pattern Recognition Letters, Vol. 24, 2003, pp. 1805–1817.

[16] S. Chikkerur, A. N. Cartwright and V. Govindaraju, "Fingerprint enhancement using STFT analysis," Pattern Recognition, Vol. 40, 2007 pp. 198–211.

[17] P. K. Shimna and B. Neethu, "Fingerprint Image Enhancement Using STFT Analysis," IOSR Journal of Electronics and Communication Engineering, Vol. 10, No. 2, Apr. 2015, pp. 61-68.

[18] E.-K. Yun and S.-B. Cho, "Adaptive fingerprint image enhancement with fingerprint image quality analysis," Image and Vision Computing, Vol. 24, 2006, pp. 101-110.

[19] J. S. Bartunek, M.Nilsson, B. Sallberg and I. Claesson, "Adaptive Fingerprint Image Enhancement With Emphasis on Preprocessing of Data," IEEE Transactions on Image Processing, Vol. 22, No. 2, February 2013, pp. 644-656.

[20] Hsieh, C., Lai, E., Wang, Y.: An effective algorithm for fingerprint image enhancement based on wavelet transform. Pattern Recognition 36, (2003), pp.303–312.

[21] C. Gottschlich, "Curved-Region-Based Ridge Frequency Estimation and Curved Gabor Filters for Fingerprint Image Enhancement," IEEE Transactions on Image Processing, Vol. 21, No. 4, April 2012, pp. 2220-2227.

[22] M. A.U. Khan and T. M. Khan, "Fingerprint image enhancement using data driven Directional Filter Bank," Optik, Vol. 124, 2013, pp. 6063– 6068.

[23] P.Viswanathan and P.V. Krishna, "Morlet Wavelet Fingerprint Invariant Automated Authentication System," Int. J. of Recent Trends in Engineering and Technology, Vol. 4, No. 1, Nov 2010, pp. 1-5.

ISBN 979-8-3503-5094-4

Comparative Analysis of Brain Tumor Detection in MRI Images: A Study of Neural Networks for Enhanced Classification Performance

Sowjanya B
Department of Biomedical Engineering
Sri Sivasubramaniya Nadar College of Engineering,
Kalavakkam, Chennai, Tamil Nadu, India
sowjanya2010391@ssn.edu.in

Suganthi Lakshmanan
Department of Biomedical Engineering
Sri Sivasubramaniya Nadar College of Engineering,
Kalavakkam, Chennai, Tamil Nadu, India
suganthil@ssn.edu.in

***Abstract*— Medical imaging plays an important role in the early detection and treatment of various diseases, including tumor detection via Magnetic Resonance Imaging (MRI). This research paper presents an extensive investigation into the application of different neural network models for the classification of MRI images to enable tumor detection. We evaluate the performance of Backpropagation Neural Networks (BPN), Multiple layer perceptron (MLP), and Convolutional Neural Networks (CNN) in this context.**

Results give away that MLP achieves the highest accuracy at 96.69%, with good precision of 98.82% and specificity of 98.41%. CNN follows closely with 94.70% accuracy and notable sensitivity of 96.94%. These findings reveal the potential of neural networks, specifically CNN and MLP, to enhance the diagnostic capabilities of medical imaging technologies for faster and more accurate tumor detection, therefore leading to better patient outcomes.

***Keywords*—** *Back Propagation Network, Multilayer Perceptron Edge Detection, PCA, CNN, Tumor Classification*

I. INTRODUCTION

Healthcare has undergone a significant transformation because to medical imaging, especially magnetic resonance imaging (MRI), which offers precise and non-invasive views of internal anatomical structures. The application of magnetic resonance imaging (MRI) has grown dramatically in recent years, particularly in the prompt identification and diagnosis of a wide range of illnesses, including malignancies. Modern healthcare now depends on the ability to identify MRI images for tumor diagnosis rapidly and reliably.

The frequency of brain tumors in India is on a steady upward direction. India witnesses a growing range of reported cases of brain tumors across different age groups every single year. In 2018, brain tumors were listed as the 10th most prevalent type of tumor in the Indian population. According to data from the International Association of Cancer Registries (IARC), India registers more than 28,000 cases of brain tumors annually, leading to over 24,000 reported fatalities from this condition each year. The presence of a brain tumor is a grave medical concern and can result in fatality if not identified early and promptly treated.

Survey study has been carried out in India [3] which resulted that Tumors of the central nervous system (CNS) constitute a relatively small portion, approximately 2%, of all malignancies. However, due to their impact on morbidity and mortality, particularly among young and middle-aged individuals, they have a substantial influence on death-adjusted life years compared to other malignancies. CNS tumors exhibit considerable diversity in terms of age of onset, clinical outcomes, histology and location. Recent advancements in diagnostic imaging, surgical methods, radiotherapy technology, and the development of novel chemotherapeutic and targeted treatments have significantly improved the prognosis for these tumors.

This research paper is dedicated to investigating the application of machine learning techniques, specifically Backpropagation Neural Networks (BPN), Multilayer Perceptron (MLP), and Convolutional Neural Networks (CNN), for the classification of MRI images in the context of tumor detection. Tumor detection through MRI not only offers the advantage of early diagnosis but also plays a vital role in treatment planning and monitoring the development of diseases. As a result, developing effective and efficient methods for classifying MRI images is of high importance in the medical field.

This work aims to evaluate the effectiveness of BPN, MLP, and CNN for tumor detection in MRI image categorization. We use a well-structured dataset containing images of both tumor and non-tumor patients to achieve this goal. We hope to provide a comprehensive knowledge of the strengths and weaknesses of these three different neural networks in this important medical application by comparing the results obtained from them.

Ibrahim et al.'s [2] explanation of the BPN classification process breaks down into three steps: preprocessing, Principal Component Analysis (PCA)-based dimensionality reduction, and back-propagation neural network classification. The paper compares the suggested method to baseline techniques using a dataset of 3x58 MRI brain sagittal images for training and testing.

The results of this study could help researchers and medical practitioners choose the best neural network architecture for classifying MRI images, which would advance the field of tumor identification in medical imaging. In the end, our work supports the continuous goal of boosting early tumor detection and improving patient outcomes, both of which can significantly affect healthcare and patient well-being.

ISBN 979-8-3503-5094-4

II. LITERATURE REVIEW

Ahmed et al. presented a method that investigated the application of deep features from pre-trained Convolutional Neural Networks (CNNs) to survival time prediction, providing further evidence in favor of improving its performance via domain-specific fine-tuning. The study made use of an online standard dataset that is widely available. In the context of leave-one-out based cross-validation, the results showed an accuracy of almost 81%. [4]

T.R Thamizhvani et al demonstrates research on creating a standalone application for skin tumor diagnosis. It emphasizes a thorough feature extraction process from segmented images. A meticulous evaluation involving three classifiers and multiple trials with the ANOVA test, the study highlights the Multiclass Support Vector Machine as the optimal classifier. [1]

While Yue Wang and co-authors (1997) introduced a probabilistic neural network-based technique for the unaided quantification and categorization of brain tumors within magnetic resonance images, N. Benamrane and colleagues (2006) introduced a method to identify and characterize anomalies found in medical images.[5]

Mohan and Subashini in 2018 conducted a research that identified good strategies for categorizing MRI brain tumors through a hybrid approach. This method gives medical practitioners an extra diagnostic tool and a different viewpoint to improve treatment protocols. It is specific to a given type of tumor and depends on a limited number of imaging datasets. [6]

Chun-Ming Gu and Hui Ji presented a detection technique for identifying abnormal and normal MR brain pictures in March 2012. This method uses an Artificial Neural Network (ANN) for classifier training, Principal Component Analysis (PCA) for feature dimensionality reduction, and Discrete Wavelet Transform (DWT) for feature extraction. To train the ANN, the Adaptive Back Propagation (ABP) technique was chosen. [7]

Gu et al. introduced a method for classifying MRI brain tumor images into benign and malignant categories, leveraging deep features and machine learning. They compared CNN and traditional classifiers, reporting superior performance akin to Kang et al.'s 2021 findings. Additionally, they proposed a novel approach based on local constraint-based convolutional dictionary learning, outperforming established methods for distinguishing normal and abnormal brain tumors using REMBRANDT datasets.[9]

III. CONCEPT

The method used in this research includes the principal component analysis (PCA) which is then used in BPN and MLP to increase the accuracy, and a CNN model is constructed which is then used as a comparative study to choose the best model suitable for this application. this section highlights the concepts which forms the base of this study

A. Principal component analysis (PCA):

Using this dimensionality reduction technique, high-dimensional data can be converted into a lower-dimensional format while maintaining the most crucial information. It achieves this by identifying and highlighting the principal components or directions in the data with the most variance, allowing for a denser representation of the data.[8]

B. Backpropagation Network (BPN):

The architecture consists of an input layer, a hidden layer with 128 neurons, and an output layer with a single neuron for binary classification. The activation function used throughout the network is the sigmoid function, which compresses the network's output between 0 and 1.

Both the input and hidden layers have randomly initialized weights and biases. The training process is regulated by the learning rate and the quantity of training epochs. Forward propagation uses the current weights and biases to compute each layer's output during training. By comparing the expected output with the true labels, the error is computed.

The goal of backpropagation is to minimize the error by adjusting the network's biases and weights in the opposite direction as the error gradient. Using the gradients of the error with respect to each parameter, scaled by the learning rate, the weights and biases are updated during this process. To increase the network's performance on the training data, the training loop iterates for a predetermined number of epochs, gradually adjusting the network's parameters. The basic architecture of BPN is shown in Figure 1

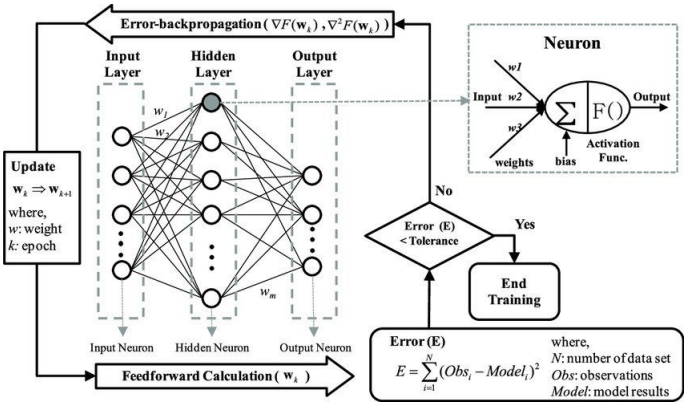


Figure 1: BPN Model

C. Multilayer Perceptron (MLP):

The architecture of the Multilayer perceptron starts by defining the MLP model with an input layer consisting of 30 neurons with Rectified Linear Unit (ReLU) activation, followed by an output layer with 2 neurons and an activation function(softmax), apt for binary classification.

Adam optimizer is employed to compile the model along with a sparse categorical cross-entropy loss function. Training is performed on the training data for 1000 epochs, with a portion of the data reserved for validation. After training, the model predicts labels for the test data, and various performance metrics such as precision, accuracy, specificity and sensitivity which are calculated using scikit-learn functions.

The confusion matrix is calculated and visualized to assess the model's performance. In addition, true negatives, true positives, false negatives, and false positives are extracted

ISBN 979-8-3503-5094-4

from the confusion matrix, enabling the calculation of more detailed metrics such as precision, sensitivity, and specificity, providing a thorough evaluation of the model's effectiveness in binary classification tasks. The preliminary architecture of MLP is shown in Figure 2

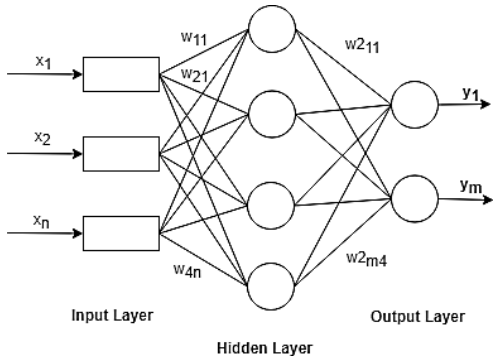


Figure 2: MLP

C. Convolutional Neural Network (CNN)

Figure 3 demonstrates the architecture of CNN, The model onsets with an input layer, which takes input data in the structure of images with the specified shape. Zero padding is applied to the input images to maintain spatial dimensions during convolution operations. The convolutional layer follows, with 32 filters of size 7x7 pixels, using a stride of 1x1 and ReLU activation.
Batch normalization is then applied to normalize the activations of the previous layer. The feature maps are down sampled using two max-pooling layers, which reduces their spatial dimensions without sacrificing important information.

Following the pooling layers, the feature maps are flattened into a one-dimensional array, and a fully connected layer with a single neuron and sigmoid activation is added for binary classification. The output of this layer represents the probability of tumor presence in the input image. Overall, the architecture consists of CNN is used in a sequential pipeline for tumor detection, using convolutional and pooling layers for feature extraction and a fully connected layer for classification.

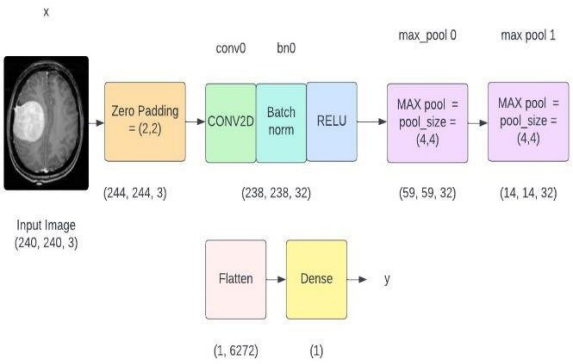


Figure 3: Architecture of CNN

IV. METHODOLOGY

In this study, we conduct a comparative analysis of various neural networks, including BPN, MLP, and CNN, to determine the most effective network for the application of brain tumor detection in MRI images. The dataset used in this research is obtained from Kaggle [10], which consists of 253 brain images, comprising both tumor and non-tumor images, forming the foundation of our research .

Data augmentation techniques are utilized in recognition of the importance of dataset diversity and size. This process includes creating variations of the original dataset by the application of operations such as rotation, flipping, or scaling. Data augmentation helps mitigate overfitting and enhances the model's ability to handle a wider range of inputs. The shape of the train, validation, and test samples for BPN and MLP is given in table 1 and CNN is given in table 2

TABLE 1: Shape of train, test and Validation sets for BPN and MLP

X train shape	[760, 100]
Y train shape	[706, 1]
X Validation shape	[152, 100]
Y Validation Shape	[152, 1]
X test shape	[152,100]
Y Test shape	[152,1]

TABLE 2: Shape of train, test and Validation sets for CNN

X train shape	[760, 240, 240, 3]
Y train shape	[706, 1]
X Validation shape	[152, 240, 240, 3]
Y Validation Shape	[152, 1]
X test shape	[152,240, 240, 3]
Y Test shape	[152,1]

Pre-processing Techniques such as Cropping is used where the input image is to focus on the brain and tumor region, The image is first converted into grayscale to simplify further processing, A slight Gaussian blurring is applied the grayscale image, Thresholding is done to create a binary image highlighting the area of interest. A series of erosions and dilations are applied to remove small regions of noise and refine the binary image. Grayscale dilation is performed with the dilation operation in Equation. 1 with f(x) as image function and b(x) as the structuring function in Euclidian space

$$(f \oplus b)(x) = \sup[f(y) + b(x - y)] \quad for\ all\ y\ belongs\ to\ E$$

Equation: 1

ISBN 979-8-3503-5094-4

Contours are identified in the thresholded image, and the largest contour, assumed to correspond to the brain and tumor region, is selected. From the contour, the extreme points are identified A new image is created by cropping the original image using the extreme points, resulting in an image that primarily displays the brain and tumor region.

The image is then resized since the original images are varied in size, the resized image has a consisted shape of (240,240,3) where 240 X 240 represents the width and height and 3 denotes the number of color channels. Ensuring consistent and uniform data, the pixels are normalized, this step scaled pixel values to the range of 0-1, facilitating model training and convergence

This input image is given to the CNN model, for the other models such as the BPN and MLP two more steps are followed which are the Edge Detection using Sobel and Principal Component Analysis, Sobel edge detection is a widely used technique in image processing for detecting edges and boundaries within an image

The image preprocessing begins with Gaussian smoothing, which reduces noise through a convolution operation with a 2D Gaussian kernel, as defined by Equation 2:

$$G(x,y) = \frac{1}{2\pi\sigma^2} e^{\frac{-(x^2+y^2)}{2\sigma^2}}$$

Equation: 2

Next, the Sobel operators are employed to calculate the first-order image derivatives in both the horizontal and vertical directions, as outlined in Equation 3:

$$G(x,y) = \sqrt{Gx(x,y)^2 + Gy(x,y)^2}$$

$$\phi(x,y) = \arctan\left(\frac{Gx(x,y)}{Gy(x,y)}\right)$$

Equation: 3

Following gradient calculation, the algorithm proceeds with non-maximum suppression, preserving local maximum values of the gradient magnitude. This step refines the edges and retains distinct features. To conclude, the edges are further enhanced by analyzing gradient magnitudes in various directions and applying a hysteresis threshold, leading to the identification of continuous edge contours while filtering out weaker and noisy edges.

The Sobel operator-based edge detection technique excels in highlighting edges and features within images

The Edges are the formed as an array and given to the PCA algorithm for dimensionality reduction, this analyses the variability in the data and extracts essential features, reducing data complexity. By applying PCA, we aim to improve the efficiency and accuracy of the BPN and MLP models.

Therefore, the suggested methodology is shown in Figure 4

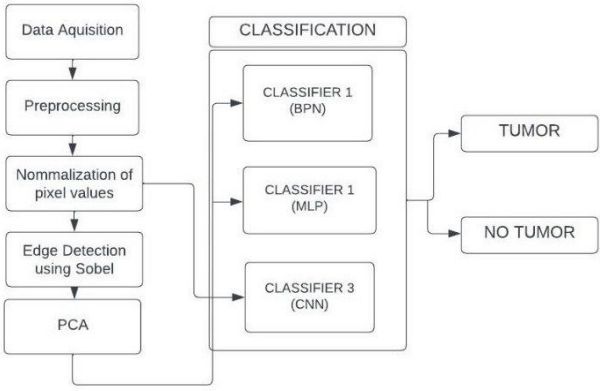


Figure 4 : Suggested methodology for MRI image classification

.

V. RESULTS AND DISCUSSIONS

In the proposed method, The MRI image dataset from Kaggle is used for the Classification into tumor and non- tumor images for further diagnosis

Pre-processing Techniques such as Cropping is used where the input image is to focus on the brain and tumor region, The image is first converted into grayscale to reduce the complication further processing, A slight Gaussian blurring is applied the grayscale image, Thresholding is done to create a binary image highlighting the area of interest. To eliminate isolated noise patches, a sequence of erosions and dilations is used. and refine the binary image.
The cropped images are shown in Figure 5

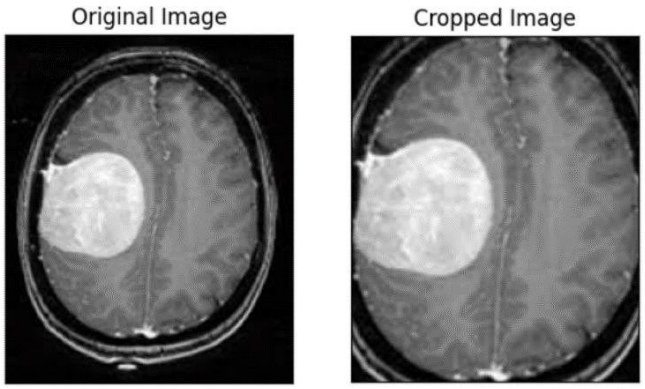


Figure 5: Original image and Cropped image

Feature Extraction is brought out by edge detection by the Sobel operator The image preprocessing involves Gaussian smoothing to reduce noise using a 2D Gaussian kernel (Equation 3) Subsequently, Sobel operators calculate first-order image derivatives in both horizontal and vertical directions (Equation 4). Non-maximum suppression refines edges by retaining local gradient magnitude maxima. Finally, gradient analysis and a hysteresis threshold connect continuous edge contours

ISBN 979-8-3503-5094-4

The edge detection features are given below in Figure 6.

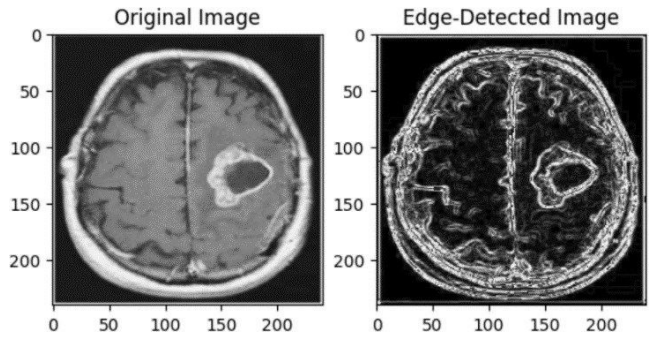


Figure 6: Original image and Edge detected image

Principal Component Analysis (PCA) is utilized for dimensionality reduction after applying Sobel edge detection to the images. PCA is a statistical technique that helps reduce the dimensionality of data while retaining the most important information. It works by converting the high-dimensional data into a lower-dimensional space by identifying and selecting the most significant features (components) that capture the maximum variance in the data. In this context, PCA is applied to the preprocessed images to reduce the dimensionality of the image data while preserving the essential information.

Classifier 1 is the back propagation network in which both training and testing occurs, The model architecture includes an input layer, a hidden layer with 128 neurons, and an output layer with a single neuron for binary classification. Throughout the network, the sigmoid activation function is used to constrain outputs between 0 and 1, with its derivative facilitating backpropagation. Weights and biases are initialized randomly for the input and hidden layers, and training parameters such as learning rate and epochs are set for control.
During training, forward propagation computes layer outputs based on current parameters, while error is determined by comparing predicted outputs with actual labels. Backpropagation adjusts weights and biases in the alternate direction of the error gradient, aiming to reduce discrepancies. This iterative process updates parameters using error gradients scaled by the learning rate, gradually improving model performance over the specified epochs.
The confusion matrix of the classifier 1 given in Figure 6.

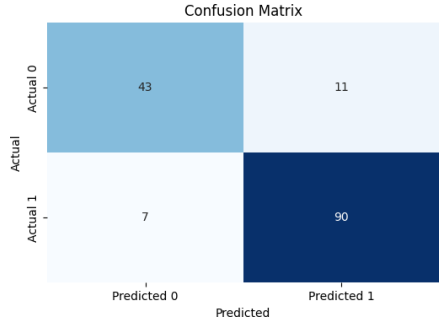


Figure 7: Confusion matrix for BPN

Classifier 2, is the MLP which is initialized with an input layer of 30 neurons utilizing the Rectified Linear Unit (ReLU) activation function, and an output layer with 2 neurons and a activation function (Softmax), for binary classification tasks. Compilation is done using the Adam optimizer and sparse categorical cross-entropy loss.
Training occurs over 1000 epochs on the training dataset, with a fraction allocated for validation. After the training, the model predicts labels for the test dataset, and various performance metrics such precision, accuracy, specificity and sensitivity are computed using scikit-learn functions. To analyze the model's performance, a confusion matrix is computed and visually represented.

Additionally, key metrics including true negatives, true positives, false negatives, and false positives are extracted from the confusion matrix, enabling precise calculation of precision, sensitivity, and specificity. This evaluation provides a detailed assessment of the model's effectiveness. the confusion matrix of the classifier 2 given in Figure 8.

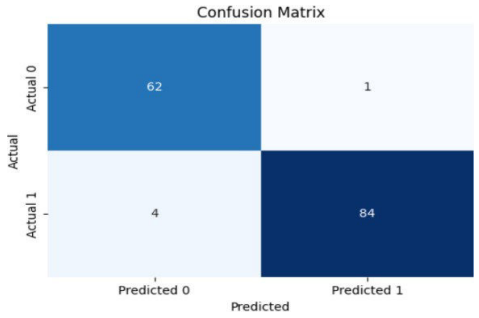


Figure 8: Confusion matrix for Multilayer Perceptron

Classifier 3 is the Convolutional neural. It takes image data with a specified input shape, applies zero-padding to maintain spatial dimensions, and then builds a series of convolutional layers (Conv2D) accompanied by batch normalization and activation function (RELU). Max-pooling layers are used to reduce spatial dimensions. The flattened output is given to a fully connected layer with a sigmoid activation for binary classification. This model is created using the Keras library in Python, the confusion matrix of the classifier 3 given in Figure 9.

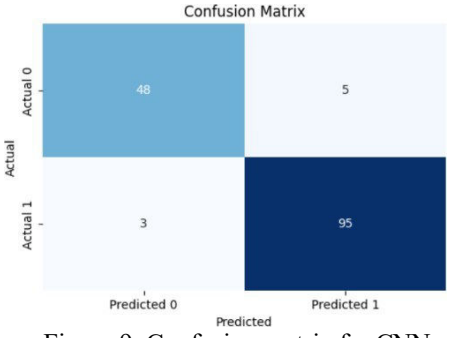


Figure 9: Confusion matrix for CNN

The comparative analysis of the 3 classifiers is given by a graphical representation of all the 4 important parameters

ISBN 979-8-3503-5094-4

such as the precision, accuracy, specificity, and sensitivity in figure 10, the performance parameters of the various models in given in table 3

TABLE 3: Performance parameters of Various models

Model	Accuracy	Precision	Sensitivity	Specificity
CNN	94.70%	95%	96.94%	90.57%
MLP	96.69%	98.82%	95.45%	98.41%
BPN	88.08%	89.11%	92.78%	79.63%

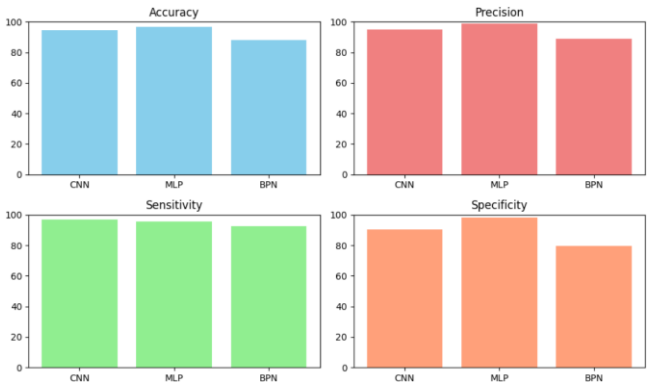


Figure 10: Graphical analysis of Performance parameters of CNN, MLP and BPN

VI. CONCLUSION

The Multilayer Perceptron (MLP) model exhibits superior performance in MRI image classification for tumor detection compared to Convolutional Neural Networks (CNN) and Backpropagation Neural Networks (BPN). With an accuracy of 96.69%, MLP outperforms both CNN (94.70%) and BPN (88.08%).

Moreover, MLP achieves the highest precision at 98.82%, indicating its capability to accurately identify true positive predictions among all positive predictions, surpassing CNN (95%) and BPN (89.11%). While MLP's sensitivity of 95.45% slightly trails behind BPN (92.78%), it significantly outperforms CNN (96.94%).

Additionally, MLP demonstrates the highest specificity at 98.41%, indicating its ability to accurately identify true negative predictions among all actual negative instances, surpassing CNN (90.57%) and BPN (79.63%).

In contrast to previous research, our study evaluates the performance of Multilayer Perceptron (MLP), Convolutional Neural Networks (CNN), and Backpropagation Neural Networks (BPN) in MRI image classification for tumor identification. We discovered that MLP performed better than CNN and BPN, with an accuracy of 96.69%, exceeding the accuracy of 94.70% for CNN and 88.08% for BPN. Additionally, MLP being almost at par with CNN suggests its greater emphasis on suitable feature extraction. The present-day literature presents novel techniques, such as 2D CNN and auto-encoder networks, demonstrating their promise for high accuracy and resilience in MRI image classification, whereas our study highlights the effectiveness of MLP. When taken as a whole, this research advances the area of medical image analysis by providing effective methods for tumor detection from MRI scans

REFERENCES

[1] T. R. Thamizhvani, Suganthi Lakshmanan & R. Sivaramakrishnan (2018) Mobile application-based computer-aided diagnosis of skin tumours from dermal images, The Imaging Science Journal, 66:6, 382-391, DOI: 10.1080/13682199.2018.1492682

[2] W. H. Ibrahim, A. A. A. Osman and Y. I. Mohamed, "MRI brain image classification using neural networks," 2013 INTERNATIONAL CONFERENCE ON COMPUTING, ELECTRICAL AND ELECTRONIC ENGINEERING (ICCEEE), Khartoum, Sudan, 2013, pp. 253-258, doi: 10.1109/ICCEEE.2013.6633943.

[3] Archya Dasgupta et al."Indian data on central nervous tumors: A summary of published work"

[4] Ahmed, K. B., Hall, L. O., Goldgof, D. B., Liu, R., and Gatenby, R. A. (2017). Fine-tuning convolutional deep features for MRI based brain tumor classification, International Society for Optics and Photonics. Med. Imaging Comput. Aided Diagnosis 10134:101342E. doi: 10.1117/12.2253982

[5] Chaplot, L.M. Patnaik, N.R. Jagannathan "Classification of magnetic resonance brain images, January 2006, Pages 86-92

[6] Mohan, G., and Subashini, M. M. (2018). MRI based medical image analysis: survey on brain tumor grade classification. Biomed. Signal Process. Control 39, 139–161. doi: 10.1016/j.bspc.2017.07.007

[7] Young, Jonathan & Mendelson, Alex & Cardoso, Manuel Jorge & Modat, Marc & Ashburner, John & Ourselin, Sebastien. (2015). Improving MRI Brain Image Classification with Anatomical Regional Kernels. 45-53. 10.1007/978-3-319-27929-9_5.

[8] A tutorial on Principal Components Analysis Lindsay I Smith February 26, 2002.

[9] Gu, X., Shen, Z., Xue, J., Fan, Y., and Ni, T. (2021). Brain Tumor MR Image Classification using Convolutional Dictionary Learning with Local Constraint, front. Neurosci. 15:679847. doi: 10.3389/fnins.2021.679847

[10] https://www.kaggle.com/datasets/navoneel/brain-mri-images-for-brain-tumor-detection

ISBN 979-8-3503-5094-4

Tenth International Conference on Biosignals, Images and Instrumentation, SSNCE, Chennai, March 20 – 22, 2024

TRACK 9

Emotional Resonance in Brainwaves: EEG based Classification of Emotional Dynamics

Manishaa K
Department of Electronics Engineering,
Madras Institute of Technology, Anna University,
Chennai, India.
manishaa.kandasamy@gmail.com

C Sridevi
Department of Electronics Engineering,
Madras Institute of Technology, Anna University,
Chennai, India.
sridevijegan@gmail.com

Kiran B
Department of Electronics Engineering,
Madras Institute of Technology, Anna University,
Chennai, India.
kiranbalamurugan@gmail.com

M Tharun Ganesh Roy
Department of Electronics Engineering,
Madras Institute of Technology, Anna University,
Chennai, India.
tharunaniruth46@gmail.com

***Abstract*— This project focuses on combining various machine learning algorithms to classify emotions based on electroencephalogram (EEG) data. In the fields of affective computing, human-computer interface, and healthcare, emotion recognition is significant. The DREAMER (Database for Emotional Analysis in Music Videos) and GAMEEMO datasets, both of which include EEG signals captured during particular stimuli are used in the study. The two datasets are compared at the initial phase of the project in order to figure out which is the most appropriate for additional investigation. The study involves feature extraction, preprocessing, artifact identification, and dataset comparison analysis after dataset selection. Using the selected dataset, several machine learning techniques are used for emotion classification, which include Decision Tree, Random Forest, AdaBoost, Naïve Bayes, and Linear SVM. The results indicate AdaBoost is effective in classifying emotions with the maximum accuracy of 91.7%. Additionally, Adaboost has a F1 score of 94.1% and precision of 88% which tends to be the highest among other algorithms used. Various performance metrics such as F1 Score, Sensitivity, Specificity, Recall and ROC curve are determined for these algorithms, which classify emotions into stress and non-stress classes. Further studies include exploring multimodal approaches and transfer learning to enhance model performance and accurately predict emotions.**

Keywords*— *Emotion, DREAMER, GAMEEMO, machine learning, stress, non-stress

I. INTRODUCTION

Deeply rooted deeply rooted in human nature, emotions are the vital thread that connects our perceptions, interactions, and reactions in the dynamic real world. Emotions of all kinds, from love to fear, from joy to grief, enhance our experiences and have an impact on how we interact with and view the world. The function of technology in the dynamic interplay of human emotions has become more and more important. The need to give technological systems the capacity to perceive and react to human emotions becomes increasingly pressing as our interactions with them become more complex. Our ability to comprehend and incorporate the intricate language of human emotions is critical to the development of intelligent and sympathetic systems.

In the quest to incorporate the language of emotions and visions, researchers have turned to physiological signals and non-physiological signals as a powerful medium for recognizing emotions.

When it comes to non-physiological indicators, body language nuances and widely recognized facial expressions offer complex windows into emotional recognition. In parallel, physiological signals directly reveal the brain processes linked to emotions, with the Electroencephalogram (EEG) being the most prominent example. Physiological elements are added by the Electrocardiogram (ECG), galvanic skin response, blood pressure, and breathing rate, which illustrates the complex interaction between emotions and the autonomic nervous system.

Two-dimensional and three-dimensional models offer frameworks for the classification of human emotions. In two-dimensional models, emotions are commonly described using two domains: arousal and valence. The term “valence” describes the positive or negative aspects of an emotion, such as joy, love, rage, and grief. On the other side, arousal corresponds to the degree of physiological activation or intensity linked to an emotion, ranging from states of calm and relaxation to states of excitement and anxiety. These models are also referred to as the “valence-arousal model” for emotion classification.

The three-dimensional model is often referred to as the PAD model (pleasure- arousal- dominance. The subjective degree of enjoyment or displeasure associated with the specific emotion or overall emotional state is represented by the pleasure parameter. Arousal measures the degree of physiological activation or intensity linked to an emotion, just as in the two-dimensional model. Dominance refers to a person’s degree of influence over his or her emotions. High level of dominance indicates that the person has control over his or her emotions in the circumstance where as low level of dominance indicates that the person is being controlled by the external factors.

ISBN 979-8-3503-5094-4

EEG signal offers high temporal resolution, is non-invasive and is widely used in neuroscience applications making it an efficient dataset among all the physiological signals. Various EEG signal datasets DEAP, SEED, AMIGOS, DREAMER, SEMAINE, MELD, are publicly available. These are collected using low-cost portable wearable EEG headset. A standardized technique for placing the electrodes on the scalp of a subject during electroencephalography (EEG) recordings is the 10-20 system. This approach guarantees that electrode locations are constant and repeatable for many people and research projects. The term "10-20" refers to the approximate distances, 10% or 20% of the total front-back or right-left distance of the skull, between neighboring electrodes. Electrodes are positioned along the scalp in both the lateral (right-left) and anterior-posterior (front-back) directions at percentages of 10% and 20% of the total distance.

Sequentially, the emotions are classified using these datasets by employing various deep and machine learning algorithms. In most of the studies, the emotions are classified using traditional deep learning methods such as CNN, DNN, and traditional classifiers namely SVM, Decision tree etc., This paper focuses on both the modern and traditional classifiers and performs comparative analysis with various performance metrics.

the paper is structured as follows. Section II discusses the linked works. Section III depicts the proposed paper's block diagram. The proposed technique is presented in Section IV. Section V discusses the methodology's findings and inferences. Section VI discusses improvisation in the future. Section VII contains the paper's conclusion.

II. Literature Survey

The course of the work described here has been done in the following research areas of brain-computer interface and affective computing. Tiantian Lv et.al [1] introduced a novel EEG emotion classification approach involving the use of the AdaBoost classifier, achieving a high average recognition rate of 90.8%. AdaBoost improves accuracy over more conventional techniques like Bagging-SVM and SVM by merging several weak classifiers. The work places a strong emphasis on feature selection using genetic algorithms to improve emotion classification. In general, the suggested method outperforms other methods in terms of stability and generalization when it comes to EEG-based emotion recognition.

Katsigiannis and Ramzan [2] created the DREAMER dataset by recording the EEG signals of the 23 participants when they were presented with 18 movie clips that elicited specific emotions. The participants also gave self-assessments of their emotional states for each stimulus. Through supervised machine learning techniques that involve utilizing support vector machines (SVMs), a baseline for participant-wise emotion recognition using EEG and ECG-based data, as well as their fusion, was developed.

Four distinct computer games—boring, calm, horror, and funny—were displayed to the participants in a new data set named GAMEMO by Alakus, Gonen, and Turkoglu [3]. After playing each game for five minutes, the participants performed a Self-Manikin Form (SAM) and the datasets were collected, resulting in a total of twenty minutes of EEG data for each of the twenty-eight study participants. The datasets were then analyzed via pattern recognition and signal processing through classification methods such as k-Nearest Neighbors, Support Vector Machine (SVM), and Multi-Layer Perceptron Neural Network. The ANOVA, Tukey, and Friedman tests were used to assess how the games affected the players. The statistical study of the respondents' ratings included the computation of the ratings' standard deviation, variance, mean, and distribution.

Qing et.al [3] proposed a novel approach to emotional activation mechanism based on machine learning and EEG signals in an attempt to address the inadequate awareness of the emotion stimulation process. To verify the efficacy of the proposed approach, experiments were carried out on the DEAP and SEED datasets, examining aspects such as emotion activation curves, interpretability, and accuracy enhancement. Decision tree, KNN, and random forest are some of the machine learning algorithms that were employed.

Koelstra et.al [4] produced a multimodal dataset called the DEAP (Database for Emotion Analysis using Physiological Signals) dataset that comprises of evaluations of participants, frontal face videos for a subset of the subjects, EEG, and peripheral physiological markers with the purpose of analyzing human affective states. A unique strategy for stimuli collection is proposed by combining video highlight recognition, affective tag retrieval, and online rating tool. The algorithms used include Gaussian Naïve Bayes as well as a decision fusion algorithm that integrates the classification results of different modalities. The DEAP dataset is a publicly available data set.

Islam et.al [5] conducted a rigorous review of the contemporary systems for emotional recognition using EEG signals. The study was conducted on a variety of existing datasets that consist of a combination of physiological and non-physiological signals. The methods and techniques used in the investigation include but are not limited to Independent Component Analysis (ICA), Second-Order Blind Identification (SOBI), Algorithms for Multiple Unknown Signals Extraction (AMUSE) and Joint Approximate Diagonalization of Eigenmatrices (JADE). Blind Source Separation algorithm has been used for artifact removal and extraction of required neural signals. Techniques like Convolutional Neural Network (CNN), Deep Neural Network (DNN), etc., and classifiers such as Decision Tree, Random Forest, etc., have been utilized for emotion classification.

In a study by Chen et al. [6], the influence of the various features that were extracted and the various frequency bands on the smoothness of emotion categorization was examined. With the help of Support Vector Machines (SVM), about six different kinds of time domain and frequency domain features were extracted and examined. It was also investigated how the signals were acquired from the active zones. Reconstructed cortex sources are determined to be preferred options for emotion recognition, according to the analyses. The TYUT 2.0 dataset, which contains EEG data from 16 participants for 250 trials, and the DEAP dataset, which contains 32-channel EEG signals of 32 subjects for 40 trials, have been used in the study. While the TYUT 2.0 dataset provides five emotion classes—sadness, anger, surprise, happiness, and neutral—each with 50 trials, the DEAP dataset divides emotions into four classes based on their valence-arousal ratings.

To improve emotion recognition using EEG, Yan, Chen, and Deng [7] created the Rhythmic Time Model for Emotion Recognition Model (RT-ERM). They employed 10-fold cross-validation for validation and LSTM networks to integrate valence and arousal metrics. By comparing the categorization findings across various rhythms and time scales, the RT-ERM model was optimized through the analysis of EEG data for frequency bands and rhythm. In comparison to traditional techniques, the study showed that the RT-ERM LSTM-based

ISBN 979-8-3503-5094-4

deep learning network had better recognition accuracy.

III. WORK FLOW

The work flow of the research work is shown in Figure 1

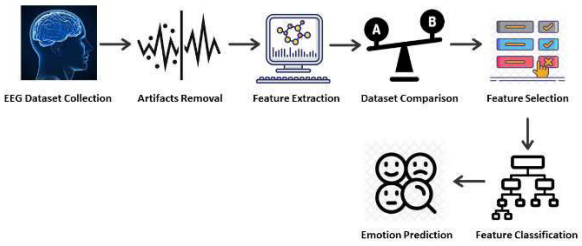


Figure 1: Block Diagram

The proposed work involves the classification of emotions using DREAMER and GAMEEMO EEG datasets. The datasets are preprocessed to ensure the removal of artifacts, such as EOG signals and power line noise etc. Subsequently, techniques such as filtering and Independent Component Analysis (ICA) have been employed to remove the artifacts and enhance the quality of the signal.

Following artifact removal, distinct features have been extracted using the preprocessed data. The features include Power Spectral Density (PSD), normalized PSD and normalized mean. The choice of these features suggests an interest in capturing the frequency distribution and mean amplitude of the EEG signals, which are indicative of underlying brain activity associated with the emotional states.

The next step involves a comparative analysis of the features extracted from both DREAMER and GAMEEMO datasets. This comparison aims to identify the most efficient and suitable dataset for the project. The criteria for efficiency could include the dataset's ability to provide discriminative features, minimize noise, and enhance the overall accuracy of emotion classification.

Once the efficient dataset is selected, the study proceeds to the emotional classification phase. The identified features from the chosen dataset are fed into various classifiers to distinguish between stress and non-stress emotions

IV. PROPOSED METHODOLOGY

The proposed study comprises of the following functions: GAMEEMO Dataset Collection, DREAMER Dataset Collection, Preprocessing GAMEEMO, Preprocessing DREAMER, Feature Extraction, Comparison of Datasets and Classification of Emotions.

A. GAMEEMO Dataset Collection

GAMEEMO is a database of EEG signals collected from 28 subjects playing four different computer games that evoke emotions such as boredom, calmness, horror, and humor. The EEG signals are collected with wearable and portable EEG device called 14 channel Emotiv Epoc+. The electrode locations used in obtaining the EEG signals are AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, and AF4. Two of these locations, P3 and P4, were used as reference electrodes to arrange the device to correct signals properly.

Subjects played emotionally four different computer games (boring, calmness, horror, and funny) for five minutes and totally 20 minutes long EEG data is available for each subject. The dataset includes raw EEG data as well as each subject's rating score and Self-Assessment Manikin (SAM) form. The SAM form is a widely used tool for measuring subjective emotional experience, and it was used in the study to rate each computer game based on the scale of arousal and valence. Emotions were described using the two-dimensional model which is valance and arousal model. Valance refers to the degree of positivity or negativity of an emotion, while arousal refers to the degree of physiological activation or intensity linked to an emotion. The games played by the subjects are:

1. G1: A train simulator game called "Train Sim World" that was found to be boring and difficult to understand by most of the subjects.
2. G2: A game called "Unravel" where the subjects controlled a yarn that tries to fix the broken bonds of people. This game was found to be calming and relaxing by the subjects.
3. G3: A horror game called "Slender – The Arrival" that was selected to induce negative emotions such as fear and anxiety.
4. G4: A funny game called "Goat Simulator" that was selected to induce positive emotions such as joy and happiness.

Figure 2: Train Sim World

Figure 3: Unravel

Figure 4: Slender – The Arrival

ISBN 979-8-3503-5094-4

Figure 5: Goat Simulator

B. *DREAMER Dataset Collection*

Dreamer is a multimodal database which records the electroencephalogram (EEG) signal and electrocardiogram (ECG) signal by means of audio-visual stimuli. 23 individuals' signals were recorded and after each stimulus, they self-assessed their affective state in terms of dominance, arousal, and valance. All the signals were recorded with readily available, affordable, wearable, wireless, portable, and off-the-shelf equipment that may make it possible to apply affective computing techniques in regular applications. For EEG and ECG, the Emotiv EPOC wireless EEG headset and the Shimmer2 ECG sensor were utilized.

This dataset uses a three-dimensional model of valence arousal and dominance to classify emotions. The valence and arousal parameters are the same as the two-dimensional model and the addition of dominance provides insight on how people perceive their ability to influence or control their emotional states. Participants in the DREAMER dataset are exposed to carefully chosen movie clips intended to elicit certain emotions, such as calmness, surprise, amusement, fear, excitement, disgust, happiness, anger, and sadness. The display of audio-visual stimuli that elicit emotional reactions is a part of this process. Subjective annotations are provided by the participants' self-reported emotional experiences, which capture the qualitative elements of their sentiments during the viewing of each video clip. There are physiological recordings in the dataset to supplement these subjective measurements. Quantitative information about the participants' physiological reaction to emotional stimuli is provided by these objective assessments. The movie clips and their associated emotions are listed below in the table.

TABLE 1: MOVIE CLIPS AND THEIR TARGET EMOTIONS

EMOTION	MOVIE
Calmness	Searching for Bobby Fischer
	Pride and Prejudice
Surprise	D.O.A.
	The Departed
Amusement	The Hangover
	Modern Times
Fear	The Ring
	Psycho
Excitement	300
	The Bourne Identity
Disgust	National Lampoon's VanWilder
	The Fly
Happiness	Wall-E
	Remember the Titans
Anger	Crash
	Gentlemans Agreement
Sadness	My Girl
	The Shawshank Redemption

C. *Preprocessing GAMEEMO*

Preprocessing GAMEEMO involves the following steps:

1. Data Cleaning: The first step is to collect and prepare the data for training the machine learning model. This involves loading the raw data and sampling the data. The sampling step is a crucial aspect of EEG data preprocessing. The sampling rate determines how many data points are recorded per unit time. Here, EEG signal is sampled around 250 Hz, meaning that 250 data points are collected per second.
2. Artifact Detection: The next step is to detect the artifacts. GAMEEMO involves artifacts such as power line noise and Electrooculography (EOG) signals which arises due to eye movement of the subjects during the EEG signal collection.
3. Power Line Noise Removal: Once the artifacts are detected, they are removed using multiple methods. Power line noise is removed using filtering. Normally, brainwaves lie in the range of 3.5-30 hertz. Hence, all the frequencies above 30 Hz are filtered out using a bandpass filter.
4. EOG Signal Removal: After the noise is removed, EOG artifacts are removed using Independent Component Analysis (ICA).
    - EEG signals can often be affected by a variety of deviations, including contractions of the muscles, eye movements, blinks, and electrical interference from nearby devices.
    - By assuming that the EEG signal is a linear combination of independent source signals, ICA may successfully isolate these artifacts from the brain-related EEG signals.
    - The quality of the EEG data can be improved by ICA by eliminating undesired artifacts and noise by identifying the independent components that correspond to brain activity
    - . As a result, further stages of the study, such feature extraction and emotion categorization, are more reliable.
    - ICA is a flexible method that does not require any prior understanding of the artifacts' origins.
    - It is appropriate for processing a variety of datasets, such as DREAMER, since it can adjust to various artifact kinds and is resilient against fluctuations in EEG data.

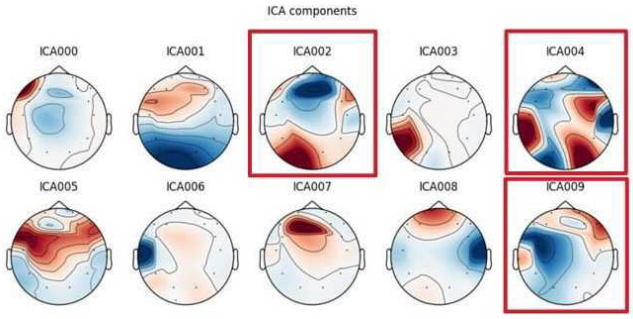


Figure 6: Independent Component Analysis

Figure 6 shows the ICA components that were obtained from the raw EEG signals after applying Independent

ISBN 979-8-3503-5094-4

Component Analysis (ICA) to the GAMEEMO dataset. By dividing the combined EEG signals into statistically independent parts, ICA efficiently isolates brain signals from noise and artifacts. In figure 6, the noisy components and other artifacts removed are ICA007, ICA009 and ICA009.

5. Plotting the Power Spectral Density: Once the data is sampled and artifacts are removed, the power spectral density (PSD) of the pre-processed data is plotted.

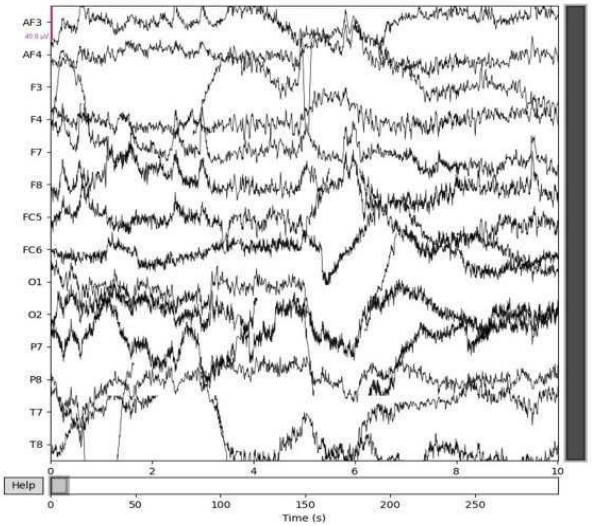


Figure 7: GAMEEMO – Raw EEG signal

The GAMEEMO dataset's raw signal offers an in-depth view of brain activity by collecting relevant neural data and artifacts like noise and other physiological data. These distortions may obscure the underlying brain signals, making it difficult to directly derive significant conclusions. Preprocessing methods are therefore crucial for improving the signal quality by getting rid of these artifacts.

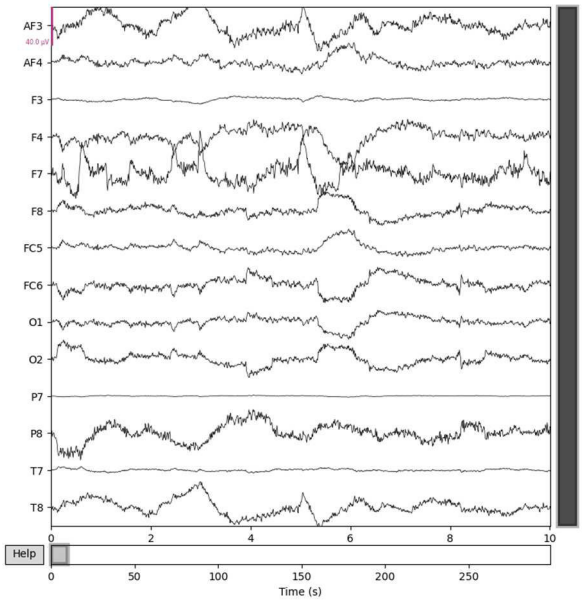


Figure 8: GAMEEMO – Filtered EEG signal

After using a variety of preprocessing methods to eliminate artifacts and enhance the overall signal quality, the preprocessed signal is shown in Figure 8. These pre-processing methods involve artifact removal techniques, and filtering.

D. *Preprocessing DREAMER*

Preprocessing DREAMER involves the following steps:

1. Data Cleaning: The EEG data from the DREAMER dataset is loaded. Missing data are identified and handled appropriately. The step also involves sampling the EEG signal at the rate of 256 Hz. The EEG data is organized into frames and each frame contains 640 data points and there are a total of 12 frames in the dataset.
2. Filtering: Filters are designed for the specific bands of frequencies and are applied to the EEG signal dataset. The EEG frequency bands used here include Delta (0-4.0 Hz), Theta (4-7.5 Hz), Alpha (7.5-13 Hz), Beta (13-20 Hz), and Gamma (20-40 Hz).
3. Power Spectral Density Calculation: After filtering, the power spectral density for the respective bands (delta, theta, alpha, beta, gamma) are calculated. Welch's method, a technique for estimating the PSD of a signal by dividing the signal into overlapping segments, is employed here.
4. Normalization: The dataset consists of two types of sub-datasets, one which is collected in the neutral condition and the other one recorded during stimulated condition. Normalized power spectral density is obtained by dividing stimulated power spectral density by baseline power spectral density.

E. *Feature Extraction*

The feature extraction stage of EEG signal processing creates a concise and useful set of features from preprocessed raw data. This phase is essential to improving the efficiency of machine learning models. The dimensionality of the data is decreased and meaningful representations of the underlying brain activity are preserved by identifying relevant features.

In this study, a total of 14 EEG signals was chosen for feature extraction. These signals were recorded from specific electrode locations, including AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, and AF4. These electrode locations were selected based on their relevance to capturing brain activity associated with emotional states.

In the case of the GAMEEMO and DREAMER datasets, both time and frequency domain features have been extracted to examine various aspects of the EEG signal.

For the GAMEEMO and DREAMER datasets, the Power Spectral Density (PSD) has been computed, in the frequency domain. The study offers a thorough understanding of the spectral features of the EEG signals by depicting the power distribution across several frequency bands.

Additionally, the time domain features that were extracted for the DREAMER dataset include key parameters consisting of the mean across multiple frequency bands, such as delta, theta, alpha, beta, and gamma. These time domain features provide an overview of the properties of the EEG signal within particular frequency ranges, which is useful information for further analysis.

Relative Power Spectral Density has been computed for the DREAMER dataset in order to further enhance the feature set. This provides a comparative analysis for the baseline and stimulated emotions.

ISBN 979-8-3503-5094-4

F. Comparison of Datasets

GAMEEMO and DREAMER datasets are compared using the extracted features. The comparative analysis is tabulated below.

TABLE 2: COMPARISON OF DATASETS

DREAMER	GAMEEMO
DREAMER dataset consists of EEG signals recorded in baseline as well as stimuli condition	GAMEEMO dataset consists of EEG signals recorded only in the stimuli condition
PSD of DREAMER consists of power ranging from -60 dBm to 60dBm	PSD of GAMEEMO consists of power ranging from -40 dBm to 40dBm
All 23 participants were exposed to different video clips and their EEG data was collected	All 28 participants were exposed to the same four games
Participants provided self-assessment of their affective states after each stimulus, which was used to validate the EEG signals recorded	No such self-assessment or validation was included
Artifacts in the signal include only EOG and ECG signals	Artifacts in the signal include EOG, ECG, and distortions due to movement of head, arms, and hands

From the above comparison, we can infer that the DREAMER dataset has a wide range of frequencies which helps in the efficient analysis of the EEG signals and in all other aspects, the DREAMER dataset is more suitable for further classification of emotions. Hence, we proceed with further classification with the DREAMER dataset.

G. Classification of Emotions

This emotional analysis uses a two-dimensional model involving valence and arousal to classify a wide range of emotions from the DREAMER dataset. STRESS emotions and NON-STRESS emotions are the two primary groups into which the emotions are categorized. This paradigm classifies emotions into NON-STRESS category such as calm, happiness, surprise, and excitement, and STRESS emotions such as anger, fear, sadness, and disgust.

On further examination of these emotions, STRESS emotions encompassing anger, fear, and disgust are associated with High Arousal Low Valence (HALV) and sadness is associated with Low Arousal Low Valence (LALV).

Conversely, NON-STRESS emotions such as excitement, surprise, and happiness correspond to High Arousal and High Valence (HAHV), indicating positive and intense feelings and calmness comes under Low Arousal and High Valence (LAHV), representing positive states.

The emotional states and their plot in two-dimensional valence arousal plane is shown in the figure below.

Figure 9: Two-Dimensional Valence-Arousal Plane

1. K-Fold Splitting

In traditional train and test split scenarios, the dataset is divided into predefined training and test data, through which they affect the reliability of model evaluation. Since the dataset is split randomly, there is a chance that the model may not be trained for a specific pattern. This randomness in the splitting of data will lead to inaccurate predictions.

K-fold cross-validation is the strategy we utilize to get over the limitations of the classic train and test split technique. A k-fold cross validation involves splitting the dataset into k-folds, or subgroups. Then, each of the k-folds is used exactly once as the validation data in the further k iterations of the cross-validation procedure, with the remaining (k-1) folds serving as the training data.

K-fold cross-validation addresses the shortcomings of traditional train-test splits by systematically iterating through different combinations of training and test datasets. Every fold uses a distinct subset of data for training and evaluation of the model, resulting in a more comprehensive and trustworthy assessment.

The dataset has been split into ten segments and the cross-validation procedure is carried out ten times in this study, where the value of k is 10. The remaining components are utilized as the training set, while a different portion is used as the test set each time.

2. Algorithms Used

After splitting test and train data various algorithms are employed to classify the emotions.

The algorithms include Naïve Bayes, Linear SVM, Decision Tree, Random Forest and AdaBoost. Subsequently these algorithms are comparatively evaluated using various performance metrics. For comparison purpose basic algorithms (Naïve Bayes, Linear SVM), weak classifiers (Random forest, Decision Tree) and strong classifier (AdaBoost) were employed to classify the emotions.

a) Naïve Bayes

- Based on the Bayes theorem, the supervised learning technique known as the Naïve Bayes algorithm is used to solve classification issues. The primary application of this approach is with high-dimensional training datasets.
- This classifier is probabilistic, implying that the model's prediction is dependent on the likelihood of the object; it is also one of the quickest machine learning algorithms that generates predictions quickly.

ISBN 979-8-3503-5094-4

- *Bayes' Theorem*:
  - The Bayes Theorem estimates the likelihood of an event based on the likelihood of an earlier event. The following equation represents the mathematical formulation of Bayes' theorem:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

  - P(A|B) is Posterior probability: Probability of hypothesis A on the observed occurrence B.
  - P(B|A) is Likelihood probability: Probability of the evidence given that the probability of a hypothesis is true.

b) Linear Support Vector Machine (SVM)

- The main objective of the SVM algorithm is to find the optimal hyperplane in an N-dimensional space so that data points can be divided into different feature space classes
- The hyperplane aims to preserve the greatest feasible buffer between the closest points belonging to different classes.
- The hyperplane's dimension is determined by the number of features.
- This is the formula for the linear hyperplane:

$$w^T + b = 0$$

- Given a data point xi, the distance to the decision border is given by:

$$d_i = \frac{w^T x_i + b}{||w||}$$

- Linear SVM is given by:

$$\hat{y} = \begin{cases} 1 : (w^T x + b) \geq 0 \\ 0 : (w^T x + b) < 0 \end{cases}$$

c) Decision Tree

- For classification and regression tasks, decision trees are the most used supervised learning algorithm.
- By laying out the potential courses of action, it helps to generate a visual depiction of the decision-making process.
- The algorithm starts at the root node, or S, in a decision tree, which contains the entire dataset, in order to predict the class of the provided dataset.
- The optimal attribute within the dataset is chosen through the application of Attribute Selection Measure (ASM). The best selected characteristic may have many values, which are represented by subsets of the dataset S.
- The optimal attribute is contained in a decision tree node that is created.
- New nodes are generated recursively using the subsets that are formed, until the nodes are no longer capable of being further classification.

d) Random Forest

- A supervised learning approach called Random Forest fits several decision tree classifiers on different subsets of the entire dataset.
- To increase the accuracy of the prediction and classification, it averages the outcomes of these independent decision trees. For regression, the final output is determined by majority vote, and for classification, by averaging.
- Since random forests construct an uncorrelated forest of decision trees using bagging and feature randomness, they are seen as an extension of the bagging technique.

e) AdaBoost

- AdaBoost, also known as adaptive boosting, is a boosting algorithm that prioritizes the under-fitted training cases.
- With the help of this supervised learning technique, classification tasks can be completed more quickly and effectively by combining weak learners into strong learners.
- It works by the stagewise addition of the weaker models.
- The model is trained by first initialsing the dataset and equal weight is assigned to all of the data points in the dataset.
- These data points are provided as input to the model and the data points that are classified incorrectly are identified.
- The weight of these data points are increased and the process is repeated until the required results are obtained.

V. RESULT AND DISCUSSION

This project demonstrates a systematic approach to EEG-based emotion classification using GAMEEMO and DREAMER datasets. The process involves removing the artifacts and extracting the time domain and frequency domain features.

Building on previous approaches, we examined the effectiveness of several machine learning algorithms for EEG-based emotion classification in our experiment. For every algorithm—AdaBoost, Random Forest, Decision Tree, Linear SVM, and Naïve Bayes—we evaluated measures including accuracy, precision, recall, and F1 score. AdaBoost outperformed existing methods with an accuracy of 91.6%, which was the greatest accuracy. This increase in accuracy demonstrates how well our method works to enhance emotion categorization from EEG data.

In order to determine which of the two well-known EEG datasets—DREAMER and GAMEEMO—was best suited for emotion classification, our study also included a detailed comparison of the two. We were able to increase the overall accuracy of emotion classification by selecting the dataset with better noise reduction and discriminative feature capabilities through this dataset comparison. Furthermore, AdaBoost showed high recall, F1 score, and accuracy, indicating its superiority over the current approach. As a result, our methodology delivers

ISBN 979-8-3503-5094-4

improved performance in emotion classification based on EEG signals. AdaBoost, while enhancing upon existing developed methodologies.

A. Naïve Bayes

Among all the classifiers, Naïve Bayes shows the lowest accuracy of 58.3%. Naive Bayes is a simplified algorithm and classifies based on the probabilistic approach. An average F1 Score of 70.6% indicates a balanced performance, but there is still a possibility to improve by reducing false positives and raising sensitivity. Approximately 66.7% of positive predictions are right, indicating a moderate level of accuracy in positive prediction making. 75% of positive occurrences are captured by the model, suggesting an acceptable performance for positive case identification. As discussed earlier, its assumption of independence among the extracted features might limit its performance, especially in our dataset with complex dependencies. The AUC-ROC of Naive Bayes is 0.36, which implies that the model has limited ability to discriminate between positive and negative instances.

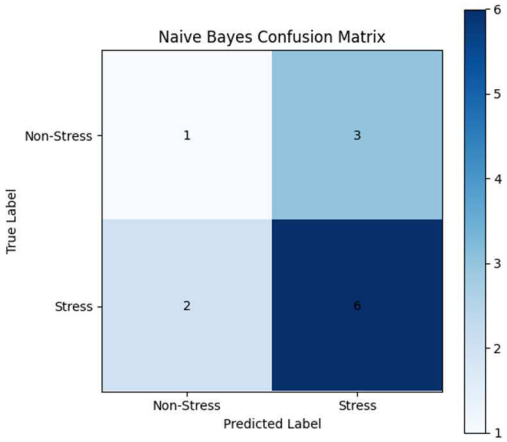


Figure 10: Confusion matrix for Naïve Bayes

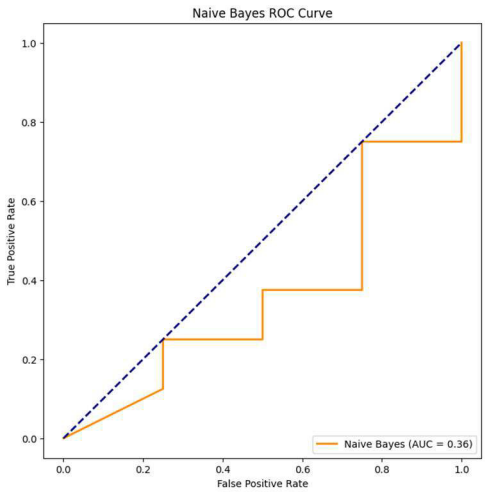


Figure 11: ROC Curve for Naïve Bayes

B. Linear SVM

This classifier shows a moderate level of accuracy meaning that approximately two-thirds of the instances are predicted accurately.

However, the reduced accuracy may indicate difficulties in handling the complex datasets or potential misclassifications. The accuracy of positive predictions is moderate, with approximately 66.7% of the positive predictions being correct. With a 100% recall score, the model does exceptionally well at capturing all positive situations.

An F1 score of 80% indicates a fair trade-off between accuracy and recall, suggesting that the model strikes a harmonic balance between reducing false positives and accurately recognizing positive cases.

The model's weak ability to distinguish between positive and negative instances is indicated by its AUC-ROC of 0.53.

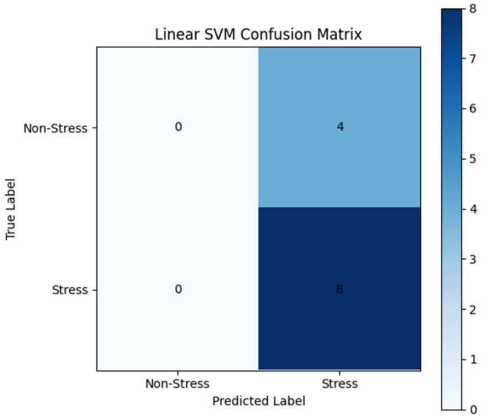


Figure 12: Confusion Matrix for Linear SVM

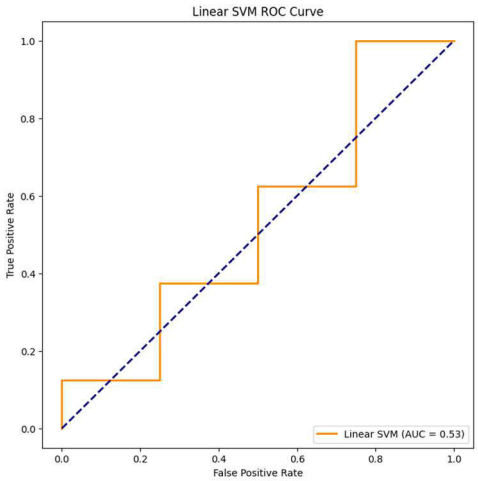


Figure 13: ROC Curve for Linear SVM

C. Decision Tree

The Decision Tree algorithm has a slightly higher accuracy of 75% compared to Linear SVM. The F1 Score of 82.4% implies a significant ability in identifying positive instances and reducing false positives, and also demonstrates an appropriate balance between precision and recall.

Approximately 77.8% of positive predictions are correct, indicating a high level of accuracy in positive prediction-making.

The model shows great capability in detecting positive cases, effectively identifying 87.5% of positive instances. The model

ISBN 979-8-3503-5094-4

gives improved accuracy when it comes to complex data.

However, there may be some incorrectly predicted instances which may be corrected through some other algorithms. With the area under the ROC curve being 0.84 for Decision Tree, it showcases a good discriminative behaviour between positive and negative instances.

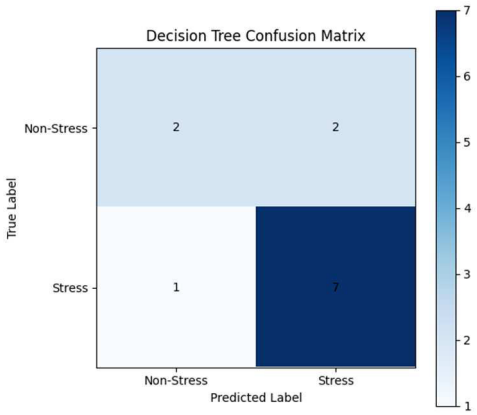


Figure 14: Confusion Matrix for Decision Tree

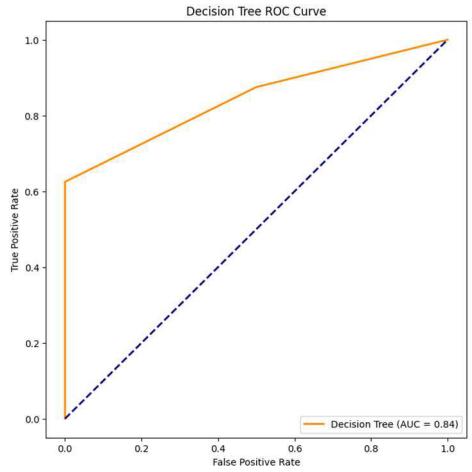


Figure 15: ROC Curve for Decision Tree

D. *Random Forest*

Random Forest classifier outperforms both the linear SVM and decision tree with an accuracy of 83%. About 80.7% of positive predictions turn out to be correct, indicating a high degree of accuracy for positive predictions. With a perfect recall score, the model demonstrates exceptional performance in identifying all positive experiences.

The improved accuracy is due to the ensemble nature of Random Forest, which aggregates the predictions from multiple decision tress. This aggregated prediction enhances the predictive accuracy than the former algorithms.

Hence it is evident that Random Forest performs well in capturing underlying patterns of the data. The Random Forest model's AUC-ROC of 0.78 suggests that discriminatory performance is mediocre.

It indicates a reasonable ability to distinguish between positive and negative instances. It also showcases the ensemble's effectiveness in mitigating overfitting.

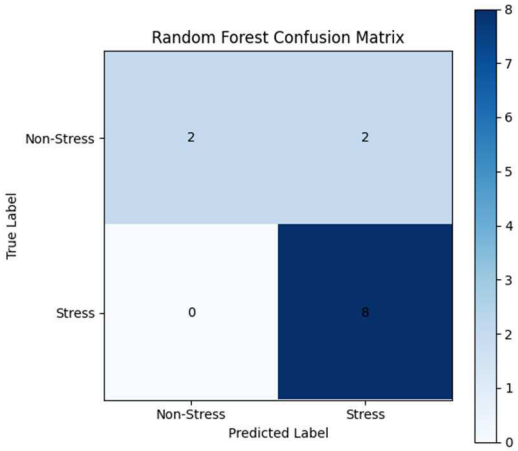


Figure 16: Confusion Matrix for Random Forest

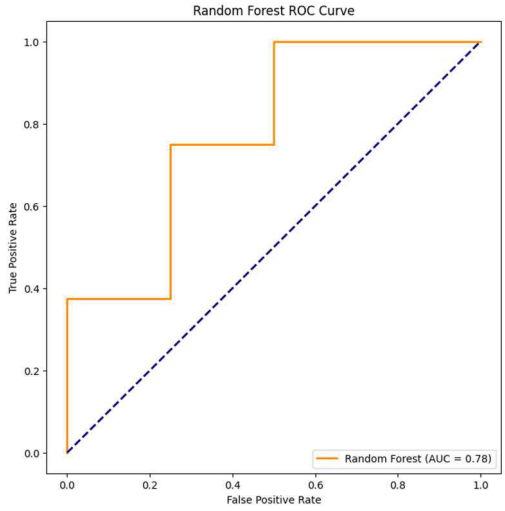


Figure 17: ROC Curve for Random Forest

E. *AdaBoost*

In classifying emotions through complex EEG datasets, AdaBoost stands out with the highest accuracy of 91.7% among the presented algorithms.

A remarkable ratio of precision and recall is indicated by the very high F1 Score of 94.1%, indicating that this model has few false positives and excellent sensitivity.

Positive predictions have a high accuracy rate of 88.9% of positive predictions turn out to be appropriate.

With a 100% recall score, the model does exceptionally well at identifying all positive instances. This classifier improves the accuracy by performing boosting which focuses on improving the classification of mis-predicted instances in the previous iterations.

This adaptable nature contributes to its high accuracy and strong ability to classify emotions in the complex EEG dataset.

With an AUC-ROC of 0.88, the AdaBoost classifier indicates strong discriminatory behavior.

The model excels in distinguishing between positive and negative instances and indicates a high ability to rank instances accurately.

ISBN 979-8-3503-5094-4

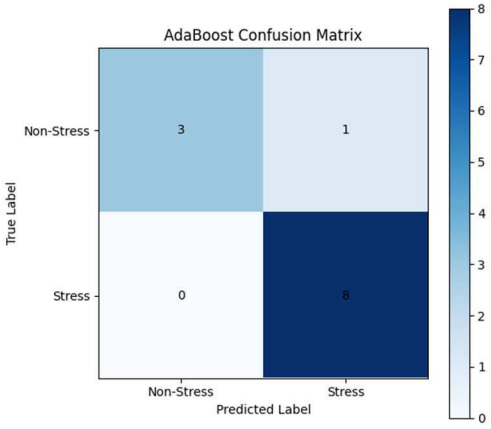


Figure 18: Confusion Matrix for AdaBoost

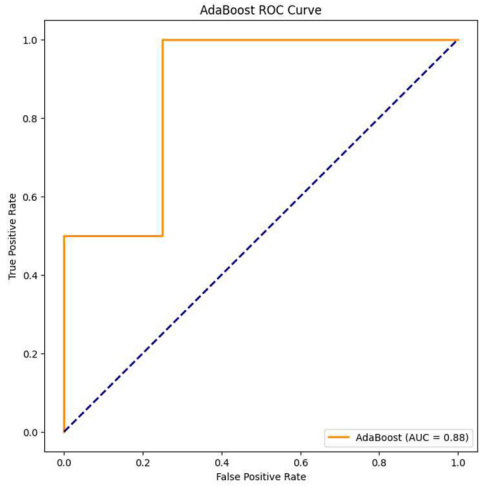


Figure 19: ROC Curve for AdaBoost

TABLE 3: COMPARISON OF ALGORITHMS

ALGORITHMS	F1 SCORE	ACCURACY	PRECISION	RECALL	SPECIFICITY
ADABOOST	0.941176	0.916667	0.888889	1.000	0.25
RANDOM FOREST	0.888889	0.833333	0.800000	1.000	0.75
DECISION TREE	0.823529	0.750000	0.777778	0.875	0.50
LINEAR SVM	0.800000	0.666667	0.666667	1.000	0.50
NAIVE BAYES	0.705882	0.583333	0.666667	0.750	0.00

The table above summarizes the performance parameters of the various classifiers that were obtained during the study.

VI. CONCLUSION

Aiming at studying the performance of machine learning algorithms rather than the traditional deep learning algorithms for EEG-based emotion classification, in this paper, we found that AdaBoost classifier was more accurate when compared to all the other algorithms. Since it involves aggregation of the predictions of multiple classifiers, it exhibited an excellent performance.

The existing approaches for EEG-based emotion classification typically fail to compare multiple datasets for effectiveness; instead, they rely on singular datasets. However, in order to determine which of the two well-known EEG datasets—DREAMER and GAMEEMO—is best suited for emotion classification tasks, our research presents an innovative approach by methodically comparing both.

Furthermore, our research shows better performance with an accuracy of 91.7%, even if existing approaches might use specific machine learning techniques like AdaBoost. This is superior to the existing technique, which used AdaBoost to achieve an accuracy of 90.8%.

VII. FUTURE SCOPE

Deep learning and multimodal data integration will be major developments in EEG-based emotion classification in the future. More complete models may be produced by integrating EEG signals with additional physiological indications such as heart rate variability and facial expressions.

Our comparative study of the DREAMER and GAMEEMO datasets offers information for future research into multimodal data fusion methods. Furthermore, using deep learning methods like CNNs and RNNs has promise for identifying complex patterns in EEG data. Our research with several classifiers highlights the value of different modeling strategies. Accepting these developments can help applications in virtual reality, mental health monitoring, and human-computer interaction by improving the stability and accuracy of emotion classification.

This integration matches the intricate complexities of the human experience while also improving the complexity of emotion perception.

REFERENCES

[1] T. Lv, J. Yan, H. Xu, “An EEG Emotion Recognition Method Based on AdaBoost Classifier,” IEEE Chinese Automation Congress (CAC), 2017, doi: 10.1109/CAC.2017.8243867

[2] S. Katsigiannis and N. Ramzan, ‘‘DREAMER: A database for emotion recognition through EEG and ECG signals from wireless low-cost offthe-shelf devices,’’ IEEE J. Biomed. Health Informat., vol. 22, no. 1, pp. 98–107, Jan. 2018, doi: 10.1109/JBHI.2017.2688239.

[3] T. B. Alakus, M. Gonenb, I. Turkogluc, “Database for an emotion recognition system based on EEG signals and various computer games -GAMEEMO,” Biomedical Signal Processing and Control vol.60, Jul 2020, Art. No. 101951, doi: 10.1016/j.bspc.2020.101951

[4] C. Qing, R. Qiao, X. Xu, and Y. Cheng, ‘‘Interpretable emotion recognition using EEG signals,’’ IEEE Access, vol. 7, pp. 94160–94170, 2019, doi: 10.1109/ACCESS.2019.2928691.

[5] S. Koelstra, C. Muhl, M. Soleymani, J.-S. Lee, A. Yazdani, T. Ebrahimi, T. Pun, A. Nijholt, and I. Patras, ‘‘DEAP: A database for emotion analysis; using physiological signals,’’ IEEE Trans. Affect. Comput., vol. 3, no. 1, pp. 18–31, Jan. 2012, doi: 10.1109/T-AFFC.2011.15

[6] M. R. Islam, M. A. Moni, M. M. Islam, M. Rashed-Al-Mahfuz, M. S. Islam, M. K. Hasan, M. S. Hossain, M. Ahmad, S. Uddin, A. Azad, S.A. Alyami, M. A. R. Ahad, P. Lio, “Emotion Recognition From EEG Signal Focusing on Deep Learning and Shallow Learning Techniques,” vol 9, IEEE Access 2021, doi: 10.1109/ACCESS.2021.3091487

[7] G. Chen, X. Zhang, Y. Sun, J. Zhang, “Emotion feature analysis and recognition based on reconstructed EEG sources,” IEEE Access 8 (2020) 11907–11916, doi: 10.1109/ACCESS.2020.2966144

[8] J. Yan, S. Chen, S. Deng, “A EEG-based emotion recognition model with rhythm and time characteristics,” Brain Informatics Springer Open, 2019(6:7), doi: 10.1186/s40708-019-0100-y

[9] Z. Lan, O. Sourina, L. Wang, R. Scherer, and G. R. Müller-Putz, “Domain adaptation techniques for EEG-based emotion recognition: A comparative study on two public datasets,” IEEE Transactions on Cognitive and Developmental Systems, 2018, doi:10.1109/tcds.2018.2826840

[10] R. Nawaz, K. H. Cheah, H. Nisar, V. V. Yap, “Comparison of different feature extraction methods for EEG-based emotion recognition,” Biocybernetics and Biomedical Engineering 40 (2020) 910-926, doi: 10.1016/j.bbe.2020.04.005

ISBN 979-8-3503-5094-4

[11] W. L. Zheng, J. Y. Zhu, B. L. Lu, "Identifying Stable Patterns over Time for Emotion Recognition from EEG," IEEE 2016

[12] M. M. Rahman, A. K. Sarkar, M. A. Hossain, M.S. Hossain, M. R. Islam, M. B. Hossain, J. M. W. Quinn, M. A. Moni, "Recognition of human emotions using EEG signals: A review," Computers in Biology and Medicine 136 (2021) 104696, doi: 10.1109/MMSP.2007.4412807

[13] N. Zhuang, Y. Zeng, K. Yang, C. Zhang, L. Tong, and B. Yan, ''Investigating patterns for self-induced emotion recognition from EEG signals,'' Sensors, vol. 18, no. 3, p. 841, Mar. 2018, doi: 10.3390/s18030841.

[14] L. De Lathauwer, B. De Moor, and J. Vandewalle, ''An introduction to independent component analysis,'' J. Chemometrics, vol. 14, no. 3, pp. 123–149, 2000, doi: 10.1002/1099-128X(200005/06)14:33.0.CO;2-1.

ISBN 979-8-3503-5094-4

Cardiac Arrhythmia Diagnosis Using Deep Learning: A 1D CNN-GRU Approach with Multiclass SVM and DWT Analysis

Hepsiba. D
Division of Biomedical Engineering
Karunya Institute of Technology and Sciences
Coimbatore, Tamil Nadu, India
hepsiba@karunya.edu

L.D. Vijay Anand
Division of Robotics Engineering
Karunya Institute of Technology and Sciences
Coimbatore, Tamil Nadu, India
vijayanand@karunya.edu

Jencia. J
Division of Biomedical Engineering
Karunya Institute of Technology and Sciences
Coimbatore, Tamil Nadu, India
jenciaj@karunya.edu.in

Vinotha. R
Division of Robotics Engineering
Karunya Institute of Technology and Sciences
Coimbatore, Tamil Nadu, India
vinothar@karunya.edu.in

Amy Fedora F
Division of Biomedical Engineering
Karunya Institute of Technology and Sciences
Coimbatore, Tamil Nadu, India
famy@karunya.edu.in

Reshma Immaculate S
Division of Biomedical Engineering
Karunya Institute of Technology and Sciences
Coimbatore, Tamil Nadu, India
reshmaimmaculate@karunya.edu.in

***Abstract*—Arrhythmia is a cardiovascular disease characterized by an abnormal heartbeat in rate or rhythm. An electrocardiogram (ECG) is an important part of cardiac arrhythmia diagnosis in clinical practice. Recording ECG for the detection of arrhythmias can pose challenges under certain circumstances. This paper's objective is to employ deep learning methods in cardiac arrhythmia diagnosis based on ECG signals, with the least amount of data pre-processing. The proposed algorithm architecture is characterized by pre-processing, feature extraction and 1-Dimension Convolutional Neural Network (1D CNN) Gate Recurrent Unit (GRU) classification for arrhythmia detection. In this paper, the following arrhythmia categories have been considered: Normal beats, premature contractions originating in the ventricles (PVCs), disruptions in the electrical conduction through the left bundle branch (LBBB), interruptions in the conduction through the right bundle branch (RBBB), and artificially initiated depolarizations (paced beats). Classifiers produce vectors, and the highest element in each vector indicates the degree of membership in the corresponding class. The proposed approach involves employing a 1D CNN coupled with GRU for classification, alongside a multi-Class SVM, resulting in an impressive accuracy of 99.97%. Though the characteristics of Discrete Wavelet Transform (DWT) are similar, their output from Multiclass Support Vector Machine (Multi SVM) is hardly different.**

Keywords—ECG, 1D CNN, DWT, GRU, Multi SVM classifier

I. INTRODUCTION (*HEADING 1*)

ECG is a non-invasive diagnostic tool with well-established utility for various purposes [1-3]. It depicts variations in the heart's electrical activity over time and contains vital physiological data that is commonly used to assess heart function. Electrodes on the body surface are used to record the cardiac electric field. Measurement involves utilizing the voltage (or potential difference) between two electrodes. A "lead" is a pair of electrodes that form an imaginary line within the body, facilitating the measurement of electrical signals along that line. ECG signals exhibit periodicity as they consist of a repeating sequence of waves over time: beginning with a P wave, followed by Q, R, and S (QRS complex), and concluding with a T wave, as illustrated in Figure 1. A 'U' wave can be noticed hardly ever. The QRS complex assumes significance within the context of an ECG. Its analysis is related to the beat-to-beat classification process. Arrhythmia is a cardiovascular disease characterized by an abnormal heartbeat in rate or rhythm. During an arrhythmia, the heart may beat excessively fast, too slowly, or in an irregular rhythm. While most arrhythmias are harmless, some can pose a danger and may even be fatal.

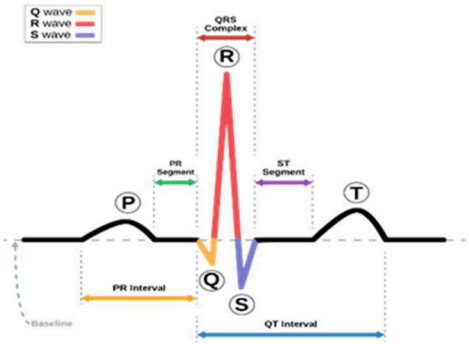


Fig 1. ECG Waveform

Recent research in the field of cardiac arrhythmia diagnosis has focused on leveraging deep learning methods for analysing ECG signals with minimal data pre-processing. Arrhythmia, a cardiovascular disorder marked by irregular heartbeats, presents challenges in accurate detection using conventional methods. In a recent study, researchers aimed to develop an efficient algorithm for arrhythmia diagnosis by integrating deep learning techniques. During an arrhythmia, the heart might struggle to effectively pump an adequate amount of blood to the body. adequate blood supply has the potential to result in harm to various organs, including the brain and heart. Cardiac arrhythmias are divided into two significant categories based on their severity: those that pose a risk to life and those that do not. Arrhythmias falling into the life-threatening category, such as ventricular fibrillation and tachycardia, carry the potential to lead to cardiac arrest and, consequently, mortality. Patient experiencing these situations need prompt medical attention. Even if non-life-threatening

ISBN 979-8-3503-5094-4

arrhythmias are not likely to cause heart failure, it's important to get treatment quickly to prevent any harm to the heart. In specific instances, irregular heart rhythms may manifest sporadically in a patient's daily routine. To capture these infrequent occurrences, the Holter device is frequently employed to collect extended ECG data. Consequently, automatically identifying abnormal heartbeats from extensive ECG datasets emerges as a critical and significant undertaking. Extensive research has been conducted on the initial form of arrhythmia, leading to the development and implementation of detection algorithms utilized in automatic external defibrillators. Recognizing ectopic beats represents a crucial phase in identifying the second type of arrhythmia.

Furthermore, the manual examination of beats on a beat-by-beat basis is tedious and time-consuming, particularly when studying long-term ECG recordings. This challenge becomes more noticeable in scenarios like wearable healthcare monitoring or bedside monitoring, where real-time diagnosis poses a difficulty for junior doctors. As a result, physicians typically examine and interpret ECG signals using computer-assisted approaches.

As mentioned earlier, the advancement of new sensing technologies has led to not only a substantial volume of ECG data but also a complex data structure, rendering conventional software and hardware ineffective. In fact, according to some reports, the total amount of digital healthcare data collected worldwide in 2012 is 500 petabytes (1015). Furthermore, it is predicted to grow to 25 Exabyte's by 2020 [9]. Hence, there is a need for solutions capable of efficiently handling and analysing diverse, complex, and large datasets within a reasonable timeframe and storage. Big data analytics, a general term for vast and complex datasets, is critical for handling the massive volume of healthcare data and enhancing the excellence of services provided to patients.

One of the issues in this context is data classification, which requires effective distributed processing platforms as well as machine learning approaches and modern data mining. Therefore, in this study, a deep learning algorithm is employed to tackle the challenges associated with classifying ECG beats.

Researchers utilize the well-established Massachusetts Institute of Technology - Boston's Beth Israel Hospital (MIT–BIH) Arrhythmia Database for experimentation, comparing their results with existing academic literature. The conclusive results suggest that the proposed model not only exceeds the current standards in accuracy but also shows competitiveness regarding sensitivity and specificity.

II. Related Works

Ali Isina et al. introduced a framework highlighting the crucial role of ECG in clinical practice for diagnosing heart arrhythmias. The research utilizes a deep learning framework that has been pre-trained on a varied image dataset to automate the diagnosis of ECG arrhythmias. The primary objective is to create a straightforward, dependable, and user-friendly deep learning technique [1].

Paweł Pławiak et al. addressed the global impact of cardiovascular disease, affecting over 50 million people. Using 744 ECG signal segments from the MIT-BIH Arrhythmia database, the DGEC system is demonstrated with sensitivity of 94.62 percent (40 errors out of 744 classifications) and a precision of 99.66 percent (across 17 arrhythmia ECG classes) [2].

In a different investigation, the Deep 1D model attained a total recognition accuracy of 91.33 percent for 17 cardiac arrhythmia disorders (classes), and the classification time for each individual sample is 0.015 seconds. The study employed 1000 fragments of ECG signals from 45 individuals, using data from the MIT-BIH Arrhythmia database for a single lead (MLII). The results are among the most promising to date compared to existing research, suggesting the potential use of the technique in cloud computing and mobile devices [3].

Nicholar Clark et al. conducted research to develop a wearable medical device designed for real-time arrhythmia detection. The device utilizes a three-lead ECG sensor for data collection, processes ECG signals, and promptly sends wireless notifications to the patient's healthcare provider. The hardware setup comprises an economical Texas Instruments TMS320C5515 and a Raspberry Pi 3 Model B. In case three or more consecutive Premature Ventricular Contractions (PVCs) are identified, the device sends an urgent report email to the patient's doctor [5].

The automatic diagnosis of cardiac illnesses heavily relies on the classification of ECG signals [22-27]. Recent breakthroughs in artificial intelligence have shown that deep neural networks can perform feature extraction directly from data. This research introduces an ECG arrhythmia classification approach using a two-dimensional (2D) deep CNN [6]. Despite the preventability of most cardiovascular problems, fatality rates continue to rise due to incorrect treatment resulting from misdiagnoses.

Deep Learning, a robust approach in Machine Learning, is most widely used in applications such as Speech Recognition, Bioinformatics, and Computer Vision. In this paper, the results are compared to various machine learning algorithms, including Artificial Neural Network (ANN), SVM, Naive Bayes, and K-Nearest Neighbor [7]

The work proposed by Swapna G et al. aims to diagnose cardiac arrhythmia using ECG signals through the application of deep learning techniques [12-18]. Unlike traditional analysis methods that rely on feature extraction, deep learning algorithms utilize approaches such as CNN, recurrent structures like Gated Recurrent Unit (GRU), Long Short-Term Memory (LSTM), as well as a combination of CNN and recurrent structures to detect irregularities. All experimental trials are conducted for 1000 epochs with a learning rate ranging from 0.01 to 0.5 [8].

III. DataSet

The dataset employed in this study is sourced from the MIT-BIH Arrhythmia Database. Two channels are present in the MIT-BIH Arrhythmia Database. The datasets employed for training and testing the automatic classification of ECG signals encompass a variety of distinct features. Some features are derived from laboratory tests, while others are based on clinical symptoms. The MIT-BIH, on the other hand, is one of the most popular and helpful databases. Researchers have utilized this database to assess numerous algorithms for the detection and categorization of arrhythmias. For the categorization of the ECG signal, several approaches have been presented. It's crucial to appropriately recognise aberrant ECG beats. As a conducting system exists in the heart, arrhythmia is caused by a disturbance in the conducting system. The task of detecting arrhythmia early is critical.

ISBN 979-8-3503-5094-4

IV. METHODOLOGY

This section deals with the methodology adopted in detecting arrhythmia in the ECG signal.

A. Wavelet Transform

Wavelet transform is a widely used tool for analysing non-stationary signals as it enables the characterization of time-frequency data. The Fourier transform is incapable of providing time information. The use of a fixed window size in the Short Time Fourier Transform (STFT) results in a constant time-frequency resolution [11]. Because the window size of the Wavelet Transform varies, the time-frequency resolution varies as well. Wavelets demonstrate exceptional temporal resolution for high frequencies and effective frequency resolution for low frequencies.

B. Discrete Wavelet Transform

In the process of DWT, a scaling function is employed to produce the wavelet transform, necessitating that the scaling function be orthogonal to its discrete translations during construction. Implementation of the DWT involves the utilization of low-pass and high-pass filters. Both filters are applied to pass the signal x[n]. The low-pass filter, h(n), produces an approximate coefficient (A1), while the high-pass filter, g(n), produces a detailed coefficient (D1). Signal decomposition occurs in stages. There are two types of low-pass and high-pass filters: half-band low-pass and high-pass filters. The former allows only lower frequencies to pass through, rejecting higher frequencies, while the latter permits only high frequencies, rejecting low frequencies. Following the filtration process, the signal experiences down-sampling by a factor of two as shown in Figure 2.

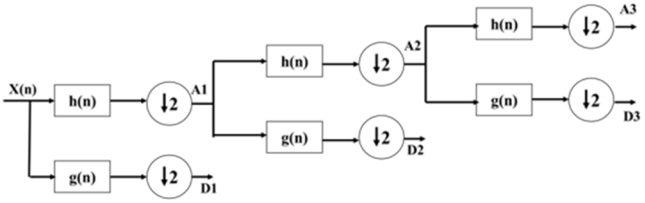


Fig.1 Decomposition at three levels

C. Proposed Methodology

The proposed work as shown in Figure 3 includes three phases: Pre-processing, feature extraction, and classification.

a) Pre-Processing:

The ECG records obtained from the MIT-BIH Arrhythmia Database undergo pre-processing. Because the database contains both low and high-frequency noise in the database, it is essential to eliminate it, this task achieved during the pre-processing stage. The pre-processing stage consists of two key steps: (i) Normalization of the ECG signal involves adjusting the amplitude of ECG signals to achieve a mean of zero and a standard deviation of one. This aids in minimizing DC offset and standardizing amplitude differences across files. (ii) The Adaptive Wiener Filter, a non-linear digital filtering method, is commonly employed to remove noise from signals. Noise reduction is a standard pre-processing technique aimed at improving the results of subsequent processing steps, such as signal edge detection. Wiener filtering is favored in digital signal processing for its ability to preserve edges while effectively reducing noise under specific conditions.

b) Feature Extraction:

The most critical stage in pattern recognition is to collect features. In categorization, the proper selection of characteristics is critical. For feature extraction, the presented work employs the DWT.

- To extract the QRS complex signal, choose 276 samples around the R peak.
- Conducting signal decomposition up to five levels, comprising detail coefficients (D1 through D5) and approximation coefficients that constitute the decomposed signal (A5).
- Two statistical features are calculated from each sub-band as follows : In each sub band, the skewness and absolute value of the wavelet coefficient. The features are retrieved from each beat and fed into the ANN, GRU, and Multiclass SVM as inputs. The Fully Connected layer within a 1D-CNN adheres to the structure of a conventional Multi-Layer Perceptron, featuring a softmax activation function in the output layer. In this arrangement, the term "fully connected" signifies that each neuron in the previous layer establishes a connection with every neuron in the subsequent layer.

c) Classification:

- 1D Convolution Neural Network (1D-CNN): The Fully Connected layer within a 1D-CNN adheres to the structure of a conventional Multi-Layer Perceptron, featuring a softmax activation function in the output layer. In this arrangement, the term "fully connected" signifies that each neuron in the previous layer establishes a connection with every neuron in the subsequent layer. The outcome of the feature extraction stage captures the high-level features of the input image.
- Gate Recurrent Unit: In terms of parameterization, the gating mechanism is a replica of a simple RNN. Back propagation through time (BTT) stochastic gradient descent, which aims to minimize a loss function, is also used to update the weights corresponding to these gates. Consequently, each alteration in parameters incorporates information about the entire network's condition. Consequently, the final state variable encapsulates all details regarding the present input and preceding hidden states. The signals propelling the gating signals exhibit redundancy. The predominant driving signal should be the internal status of the network. Moreover, every adaptive modification in parameters encompasses elements of the system's internal state.
- Multi Class SVM : In a standard classification task, there exists a collection of training and testing data containing diverse data instances. Each instance within the training set possesses a singular target value alongside multiple attributes. The goal of Multi SVM is to construct a model capable of predicting the target value of data instances within the testing set solely relying on their characteristics. An instance of supervised learning, multi SVM involves known labels to ascertain the system's correctness. This labeled data serves the purpose of indicating a desired response, validating the system's accuracy, or aiding the system in learning to behave appropriately.

Combining a 1D Convolutional Neural Network (CNN) with a Gate Recurrent Unit (GRU) offers a robust method for classifying arrhythmia diseases. The CNN extracts important features from electrocardiogram (ECG) signals, capturing spatial patterns crucial for identifying various arrhythmias.

ISBN 979-8-3503-5094-4

These features are then processed by the GRU, which excels at capturing temporal dependencies within sequential data like ECG signals. By integrating both spatial and temporal information, the combined CNN-GRU model can effectively discern patterns indicative of different arrhythmias with high accuracy. This approach holds promise for improving the reliability of arrhythmia diagnosis in clinical settings, offering a powerful tool for healthcare professionals to better understand and manage cardiovascular disorders.

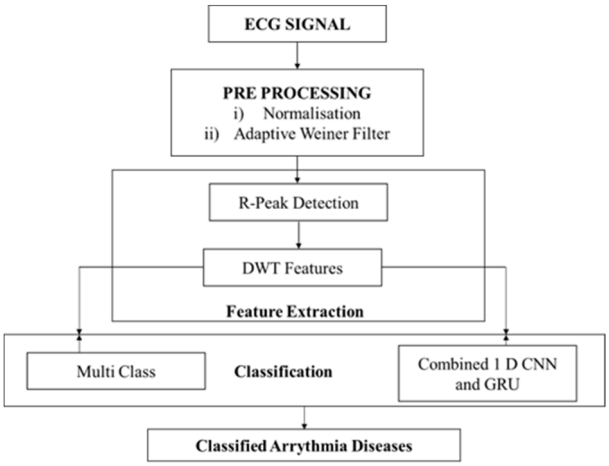


Fig 3. Proposed Method for Arrhythmia Detection and Classification

V. Results and Discussion

The performance of the classifier is evaluated using 12 files from the MIT-BIH arrhythmia database in this research. 100, 102, 103, 109, 111, 113, 118, 208, 217, 221, 231 and 233 are the records utilised. Normalization helps in reducing the impact of noise on the ECG signal. Figures 4 illustrate the ECG signal prior to normalization and Figure 5 indicates the signal after normalization. R-peaks in an ECG signal correspond to the highest amplitude point in the cardiac cycle, representing the depolarization of the ventricles as shown in Figure 6. The high-frequency components obtained from DWT as shown in Figure 7, represent the fine details of the ECG signal, capturing rapid changes associated with specific cardiac events like R-peaks. Meanwhile, the low-frequency components provide a smoother approximation of the overall signal.

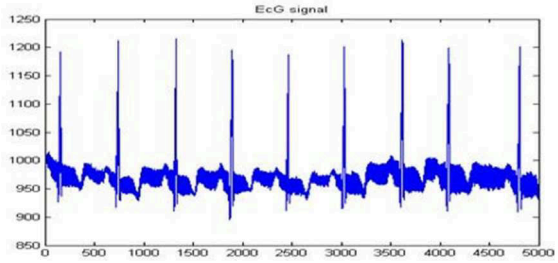


Fig 4. Input Signal

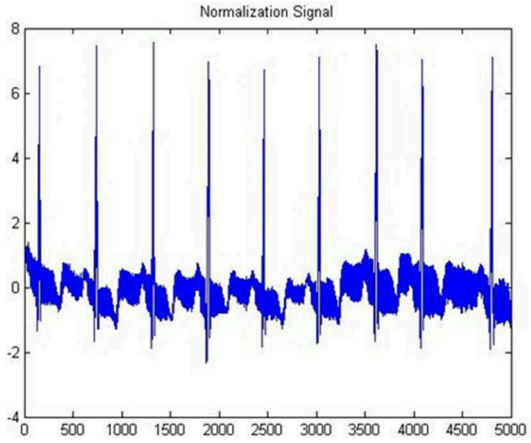


Fig 5. Normalized ECG signal

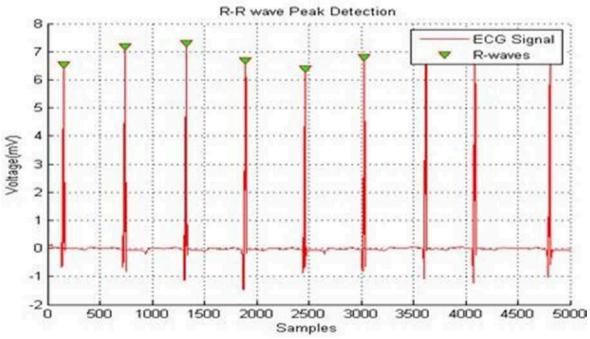


Fig 6. R Peak Detection

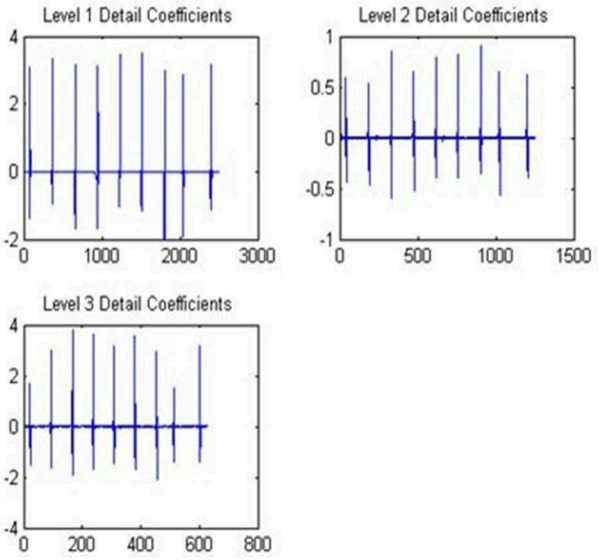


Fig 7. Decomposed Signal

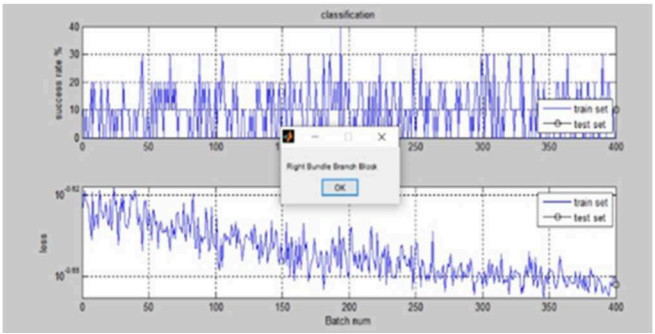


Fig 8. Detection of Right Bundle Branch Block

ISBN 979-8-3503-5094-4

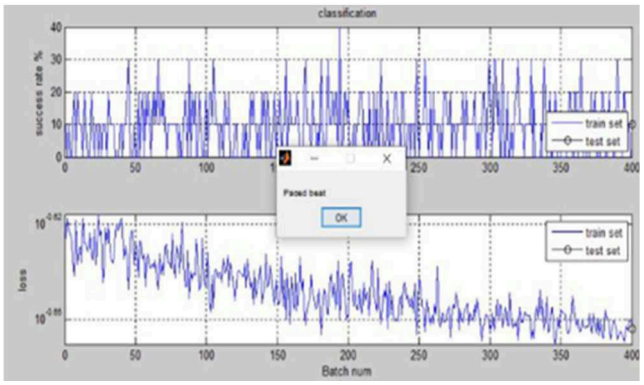


Fig 9. Detection of Paced Beat

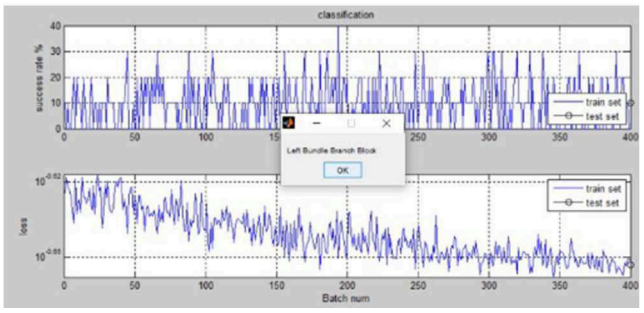


Fig 10. Detection of Left Bundle Branch Block

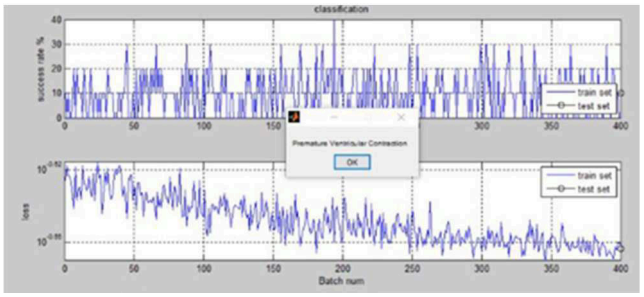


Fig 11. Detection of Premature Ventricular Contraction

The DWT is employed to extract features, and these features are subsequently categorized using ANN and SVM. The network is trained with a total of 3271 beats, and testing is conducted with 28059 beats. The different types of arrhythmias are detected and classified such as right bundle branch block, paced beats, left bundle branch block and premature ventricular contraction as shown in Figure 8, Figure 9, Figure 10 and Figure 11. Table I presents the overall data quantity for each arrhythmia type, along with the training and testing matrices. Table II and III display the confusion matrices for ANN and SVM. Fig 8 -11 displays the results of detecting various cardiac arrhythmias, including RBBB, Paced Beat, LBBB, and PVC, using an automated algorithm applied to ECG signals. Each arrhythmia type is represented by a distinct waveform pattern on the graph, allowing for visual identification and classification.

In the plot, different sections of the ECG signal correspond to specific arrhythmia categories, with characteristic morphologies indicative of each condition. The algorithm accurately detects and labels these segments, enabling clinicians to quickly assess the presence of specific arrhythmias in the ECG recording. The algorithm extracts features like waveform morphology and temporal relationships from the ECG signal corresponding to different arrhythmia types such as RBBB, Paced Beat, LBBB, and PVC. These features are used as inputs for ANN and Support Vector Machine SVM classifiers. The ANN learns to map feature vectors to specific arrhythmia classes through interconnected neurons, while the SVM constructs a hyperplane in feature space to separate arrhythmia classes. By training on labelled ECG data, both classifiers can accurately classify unseen signals, aiding in the diagnosis of cardiac arrhythmias.

In the ANN classification, there are misclassifications of 122 P beats ,345 N beats, 206 LBBB beats, 117 RBBB beats, and 626 PVC beats from each class. On the other hand, the SVM classifier identifies 122 P beats, 490 N beats, 97 LBBB beats, 39 RBBB beats, and 60 PVC beats. The performance of both ANN and SVM is presented in Table IV

The evaluation of classifier performance is computed using the following metrics:

$$\text{Accuracy} = (TP+TN)/(TP+TN+FP+FN) \quad (1)$$

$$\text{Sensitivity} = TP/(TP+FN) \quad (2)$$

$$\text{Specificity} = TN/(TN+FP) \quad (3)$$

$$\text{Positive Predictivity} = TP/(TP+FP) \quad (4)$$

In the context of these metrics: True Positive (TP) signifies all actual events correctly classified as such. True Negative (TN) encompasses all non-actual events correctly identified as not real events. False Positive (FP) includes all non-actual occurrences incorrectly classified as real events. FN (False Negative) accounts for all actual events incorrectly unrecognized as real events. Sensitivity measures the successful detection of all genuine events, Specificity gauges the accurate rejection of non-real events, and Positive Predictive Value represents the fraction of true events among all detected events. A method for classifying heart rhythm beats in an automated manner is implemented. Because the method relies solely on ECG intervals, it is not only faster and less influenced by noise than other proposed methods, but it also classifies heart illnesses more accurately.

TABLE I. Training and Testing of Arrythmia Beats

Arrythmia Type	Total number of Sample	Training Beats	Testing Beats
Normal	14678	1347	13331
P	4256	549	3707
RBBB	3783	413	3370
LBBB	5156	673	4483
PVC	3457	289	3168

TABLE II. Confusion Matrix [DWT with ANN]

Class	Confusion Matrix [DWT with ANN]				
	Normal	**P**	**RBBB**	**LBBB**	**PVC**
Normal	13721	9	31	91	43
P	197	3720	27	31	19
RBBB	27	30	2148	47	547
LBBB	121	57	0	3800	17
PVC	0	71	59	37	1875

TABLE III. Confusion Matrix [DWT -SVM]

Class	Confusion Matrix [DWT -SVM]				
	Normal	**P**	**RBBB**	**LBBB**	**PVC**
Normal	12176	13	27	9	7

ISBN 979-8-3503-5094-4

P	27	3515	0	35	19
RBBB	175	45	2171	47	442
LBBB	131	57	5	3517	2
PVC	157	7	7	6	1972

TABLE IV. Assessment of ANN and SVM Performance Utilizing DWT

Method and Classifier	Average Accuracy	Average Sensitivity	Average Specificity	Average Positive Productivity
DWT – ANN	98.30%	96.26%	98.97%	95.17%
DWT – SVM	99.97%	98.40%	99.25%	97.85%

The deep learning approach, utilizing a combination of 1D CNN and GRU, offers the advantage of automatically learning discriminative features from raw ECG signals without the need for handcrafted feature extraction. This contrasts with conventional schemes, which often rely on manually designed features that may not fully capture the complexity of ECG patterns. Furthermore, the use of multiclass SVM in conjunction with the deep learning architecture provides a robust classification framework that can effectively distinguish between different arrhythmia categories. In contrast, conventional schemes typically employ traditional machine learning algorithms such as decision trees or support vector machines, which may struggle to handle the high-dimensional and nonlinear nature of ECG data. Overall, the deep learning approach presents several advantages over conventional schemes when applied to cardiac arrhythmia diagnosis. By leveraging the power of neural networks to automatically learn features from raw ECG data, combined with robust classification techniques and complementary feature extraction methods, the proposed approach offers a promising avenue for improving the accuracy and reliability of arrhythmia detection on standard databases.

VI. Conclusion

The performance of classifiers is assessed in this work utilizing five different types of ECG beats. To begin, the ECG signal is pre-processed to reduce unnecessary noise. The skewness and absolute values of DWT coefficients are recovered as well. From each beat, a total of 24 characteristics are retrieved. The feature sets are provided to the 1D CNN and the Multi class SVM separately. The suggested method uses a 1D CNN with GRU as a classifier and a Multi Class SVM to achieve a 99.97 percent accuracy. This demonstrates that combining the DWT with the Multi Class SVM produces improved results and can be employed in a cardiac detection system. In future work, there is potential for improving upon the achieved accuracy of the proposed architecture. Exploring alternative model architectures, including variations in CNN and GRU layer configurations, as well as investigating different types of recurrent units, could enhance the model's capability to capture intricate patterns present in ECG signals. Furthermore, incorporating data augmentation methods such as signal noise addition, signal scaling, and time warping could broaden the model's exposure to diverse input data, thereby improving its generalization performance. Additionally, exploring regularization techniques like dropout or L2 regularization holds promise for preventing overfitting and bolstering the robustness of the model. These avenues for future research offer opportunities to advance the effectiveness and reliability of deep learning-based cardiac arrhythmia diagnosis.

References

[1] Islam, Md Shofiqul, Khondokar Fida Hasan, Sunjida Sultana, Shahadat Uddin, Julian MW Quinn, and Mohammad Ali Moni. "HARDC: A novel ECG-based heartbeat classification method to detect arrhythmia using hierarchical attention based dual structured RNN with dilated CNN." Neural Networks 162 (2023): 271-287.

[2] Alipo-on, Jacqueline RT, Francesca IF Escobar, Myles JT Tan, Hezerul Abdul Karim, and Nouar AlDahoul. "ECG-Based Heartbeat Classification Using Convolutional Neural Networks." International Journal of Biomedical and Biological Engineering 17, no. 12 (2023): 344-351.

[3] Choudhury, Avishek, Shankar Vuppu, Suryabhan Pratap Singh, Manoj Kumar, and Sanjay Nakharu Prasad Kumar. "ECG-based heartbeat classification using exponential-political optimizer trained deep learning for arrhythmia detection." Biomedical Signal Processing and Control 84 (2023): 104816.

[4] Cao, Minh, Tianqi Zhao, Yanxun Li, Wenhao Zhang, Peyman Benharash, and Ramin Ramezani. "ECG Heartbeat classification using deep transfer learning with Convolutional Neural Network and STFT technique." In Journal of Physics: Conference Series, vol. 2547, no. 1, p. 012031. IOP Publishing, 2023.

[5] Malakouti, Seyed Matin. "Heart disease classification based on ECG using machine learning models." Biomedical Signal Processing and Control 84 (2023): 104796.

[6] Hirnschrodt, Michael, Christoph Hintermüller, Hermann Blessberger, and Clemens Steinwender. "Delineating QRS detector parameter based ECG-Beat classification." In Current Directions in Biomedical Engineering, vol. 9, no. 1, pp. 603-606. De Gruyter, 2023.

[7] Boda, Somaraju, Manjunatha Mahadevappa, and Pranab Kumar Dutta. "An automated patient-specific ECG beat classification using LSTM-based recurrent neural networks." Biomedical Signal Processing and Control 84 (2023): 104756.

[8] Hintermüller, Christoph, Michael Hirnschrodt, Hermann Blessberger, and Clemens Steinwender. "Delineating QRS detector parameter based ECG-Beat classification." Current Directions in Biomedical Engineering 9, no. 1 (2023).

[9] Allam, Jaya Prakash, Saunak Samantray, and Samit Ari. "Empirical wavelet transform and deep learning-based technique for ECG beat classification." In Advanced Methods in Biomedical Signal Processing and Analysis, pp. 109-128. Academic Press, 2023.

[10] Dwivedi, Shashank, and Abuzar Mohammad. "Heartbeat Pattern and Arrhythmia Classification: A Review." Frontiers in Biomedical Technologies 11, no. 1 (2024): 130-148.

[11] Qi, Meng, Hongxiang Shao, Nianfeng Shi, Guoqiang Wang, and Yifei Lv. "Arrhythmia classification detection based on multiple electrocardiograms databases." Plos one 18, no. 9 (2023): e0290995.

[12] Xiaolin, Li, Fang Xiang, Rajesh C. Panicker, Barry Cardiff, and Deepu John. "Classification of ecg based on hybrid features using cnns for wearable applications." In 2023 IEEE 5th International Conference on Artificial Intelligence Circuits and Systems (AICAS), pp. 1-4. IEEE, 2023.

[13] Dhyani, Shikha, Adesh Kumar, and Sushabhan Choudhury. "Analysis of ECG-based arrhythmia detection system using machine learning." MethodsX 10 (2023): 102195.

[14] Zhou, Fei-yan, Yu-hao Sun, and Ya-wen Wang. "Inter-patient ECG arrhythmia heartbeat classification network based on multiscale convolution and FCBA." Biomedical Signal Processing and Control 90 (2024): 105789.

[15] Dwivedi, Shashank, and Abuzar Mohammad. "Heartbeat Pattern and Arrhythmia Classification: A Review." Frontiers in Biomedical Technologies 11, no. 1 (2024): 130-148.

[16] Parveen, Nikhat, Manisha Gupta, Shirisha Kasireddy, Md Shamsul Haque Ansari, and Mohammad Nadeem Ahmed. "ECG based one-dimensional residual deep convolutional auto-encoder model for heart disease classification." Multimedia Tools and Applications (2024): 1-27.

[17] Alamatsaz, Negin, Leyla Tabatabaei, Mohammadreza Yazdchi, Hamidreza Payan, Nima Alamatsaz, and Fahimeh Nasimi. "A lightweight hybrid CNN-LSTM explainable model for ECG-based

ISBN 979-8-3503-5094-4

arrhythmia detection." Biomedical Signal Processing and Control 90 (2024): 105884.

[18] Fang, Ancheng, Fan Pan, Weichuang Yu, Linkun Yang, and Peiyu He. "ECG-based emotion recognition using random convolutional kernel method." Biomedical Signal Processing and Control 91 (2024): 105907.

[19] Wang, Yanan, Shuaicong Hu, Jian Liu, Gaoyan Zhong, and Cuiwei Yang. "A multi-module algorithm for heartbeat classification based on unsupervised learning and adaptive feature transfer." Computers in Biology and Medicine (2024): 108072.

[20] Zhang, Liping, Shukai Chen, Wei Ren, Geyong Min, and Kim-Kwang Raymond Choo. "Accurate authentication based on ECG using deep learning." Journal of Computer Security Preprint (2024): 1-22.

[21] Chen, Wei-Wen, Chien-Chao Tseng, Ching-Chun Huang, and Henry Horng-Shing Lu. "Improving deep-learning electrocardiogram classification with an effective coloring method." Artificial Intelligence in Medicine (2024): 102809.

[22] Chen, Zhikang, Danni Yang, Tianrui Cui, Ding Li, Houfang Liu, Yi Yang, Sheng Zhang, Sifan Yang, and Tian-Ling Ren. "A novel imbalanced dataset mitigation method and ECG classification model based on combined 1D_CBAM-autoencoder and lightweight CNN model." Biomedical Signal Processing and Control 87 (2024): 105437.

[23] Waris, Saiyed Faiayaz. "Analysis of ECG Signals to Prediction of Ischemic Heart Disease Using Hybrid Neuro-fractal Analysis." International Journal of Intelligent Systems and Applications in Engineering 12, no. 5s (2024): 78-90.

[24] Iqbal, JL Mazher. "A novel deep learning approach for early detection of cardiovascular diseases from ECG signals." Medical Engineering & Physics 125 (2024): 104111.

[25] Aziz, Saira, Sajid Ahmed, and Mohamed-Slim Alouini. "ECG-based machine-learning algorithms for heartbeat classification." Scientific reports 11, no. 1 (2021): 18738.

[26] Xu, Xuexiang, and Hongxing Liu. "ECG heartbeat classification using convolutional neural networks." IEEE access 8 (2020): 8614-8619.

[27] Shi, Haotian, Chengjin Qin, Dengyu Xiao, Liqun Zhao, and Chengliang Liu. "Automated heartbeat classification based on deep neural network with multiple input layers." Knowledge-Based Systems 188 (2020): 105036.

[28] Ahamed, Md Atik, Kazi Amit Hasan, Khan Fashee Monowar, Nowfel Mashnoor, and Md Ali Hossain. "ECG heartbeat classification using ensemble of efficient machine learning approaches on imbalanced datasets." In 2020 2nd International Conference on Advanced Information and Communication Technology (ICAICT), pp. 140-145. IEEE, 2020.

ISBN 979-8-3503-5094-4

A compact spatial attention model for automated epileptic seizure detection using multichannel EEG

Swathy Ravi[1], Ashalatha Radhakrishnan[2]
[1,2] Dept of Neurology,
Sree Chitra Tirunal Institute for Medical sciences and Technology
swathyravi23@gmail.com,
ashalatharadhakrishnan@gmail.com

***Abstract*— Epilepsy is a non-communicable brain disorder that affects a wide population across the globe. It is diagnosed by the visual inspection of EEG which is tedious and time consuming. Automated seizure detection based on multichannel EEG is crucial for the timely identification and treatment initialisation of the disease. A seizure detection model which incorporates compact spatial attention convolutional neural network is proposed in this paper. Model is composed of 3 blocks. A CNN block that automatically captures EEG spatial features, an attention block which focus on discriminative features and a classifier which successfully classifies the seizure and non-seizure classes. Experimenting on CHBMIT dataset, our model achieved test accuracy of 91.35 %, F1score of 97.09%, sensitivity of 96.15, selectivity of 98.08%, MCC of 94% and precision of 98.04%.**

Keywords*— *Electroencephalography, Epilepsy, Attention mechanism, convolutional neural network

I. INTRODUCTION

Epilepsy, a non-communicable neurological disorder is characterized by sudden, recurrent, unprovoked seizures. Seizures result from sudden abnormal electrical impulses of brain leads to sensations and sometimes impaired consciousness [1]. According to World Health Organization (WHO) epilepsy has affected nearly 50 million population across the globe and if given proper diagnosis and care, up to 70% of them are thought to be able to enjoy seizure-free lives [2], [3]. In current clinical practices, electroencephalography (EEG) is used as golden standard for epilepsy diagnosis. The occurrence of abnormal electrical activity (seizures) is identified by the visual analysis of long-term EEG recordings. EEG records electrical activities generated by the brain, acquired using either non-intrusive or implanted devices. However, even for an expert, many hours of visual examination and analysis of lengthy EEG recordings are required to identify seizure activity in a single patient, making the process labor intensive and sometimes error prone [4]. Due to these limitations, development of automated seizure detection applications and system gained attention with objective to reduce workload of neurologists as well as to accurately detect epilepsy, thus improving the quality of life of patients.

Extensive researches have been carried out in recent years on epilepsy, more importantly on seizure detection and prediction tasks. In [5] authors classified normal, preictal, ictal states using continues wavelet transform higher order spectra and texture parameters. In Temko et al.'s work, SVM classifier was used to distinguish seizure and non-seizure classes [6]. Li et al. in [7] presented a nonlinear decomposition-based patient specific model for seizure detection using classifiers such as LDA, KNN and SVM. By using DWT algorithm in [8] they have extracted entropy features, to use SVM as classifier in their seizure prediction model.

Conventional machine learning techniques used handcrafted features for classification, which requires high level of expertise. Deep learning techniques have advanced rapidly and current studies focus on DL methods for seizure detection tasks. After frequency domain transformation of raw EEG signals, Hossain et al. in [9] explored the spatial and temporal signal characteristics using deep convolutional neural network. A binary classifier with single channel EEG was presented using 1D CNN in [10]. A one-dimensional feature fusion CNN model was developed in [11] for automatic detection of epilepsy. In [12] the authors explored spectral graph theory for detecting seizure onsets in real time. Though satisfying results have been obtained in these studies, EEG is highly time variant and it exhibits broad variability for inter and intra subject. EEG signals from different channels exhibit different properties which contribute to spatial variability of signals.

To address these challenges, we introduce a compact spatial attention mechanism focusing on relevant contextual features of multichannel scalp EEG. Attention mechanism mimics the attention process observed in our brain by allocating computational resources to focus on most relevant information while ignoring the irrelevant. This mechanism can be used in combination with neural network architectures such as CNN or LSTM for obtaining relevant information. The workflow diagram of our model is shown in figure 1. Initially from the long-term scalp EEG signals, ictal and interictal segments are extracted. After processing, these segments are given to spatial attention model. Finally, the prediction was done by using a binary classifier

II. MATERIALS AND METHODS

A. EEG Dataset

We have used CHB MIT database in our study. It is a freely accessible benchmark dataset [13] for seizure detection and

ISBN 979-8-3503-5094-4

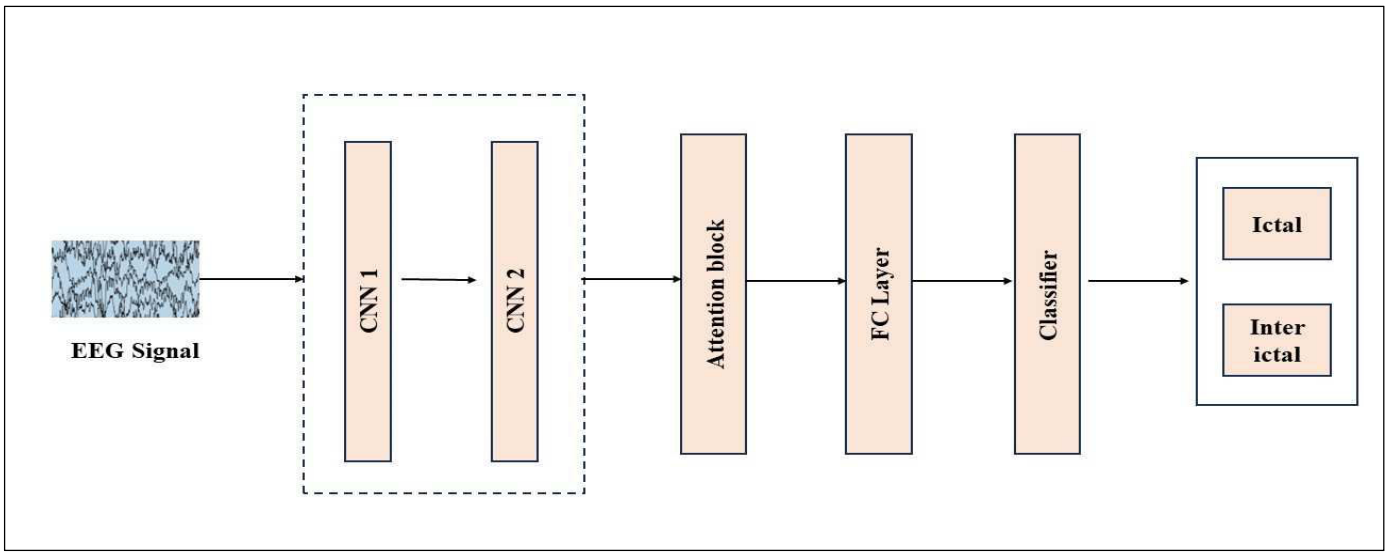


Figure 1: Work flow diagram of proposed model

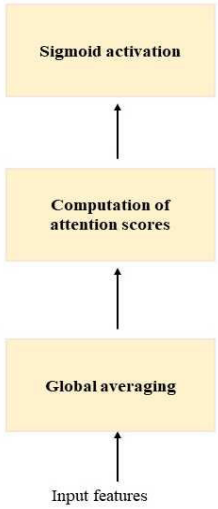


Figure 2: Work flow diagram of attention block

prediction tasks. The dataset includes large number of EEG recordings of refractory epilepsy patients that are annotated. There are 686 scalp EEG recordings of 23 epileptic patients [18 females and 5 males] in the age range of 1.5-22 years, collected jointly by the researchers at Boston children's hospital and Massachusetts Institute of Technology [14], [15]. The signals were recorded using 18 or more lead multichannel electrodes with sampling rate 256 HZ and 16-bit resolution. The dataset contains 976.5 hours of EEG signal of which 198 seizures were recorded. The start and end time of seizure events are annotated by their experts. For our study we have extracted ictal and interictal segments based on their marking. Interictal segments are extracted at least 1 hr apart from the ictal signals.

Additionally, for addressing data imbalance problem we have used sliding window technique as data augmentation step. The conclusive dataset contains ictal and interictal segments of equal 30 seconds length.

B. Preprocessing

In deep learning algorithms, data is first normalised before entering the network so as to bound data in the same order of magnitude. Raw EEG data is standardised by z score normalisation and are binary encoded for the machine to understand. This is the only preprocessing step done in our study.

C. Model description

EEG signals have high intra and inter subject variability. As different neurons interact with each other in different way, signals in every electrode channel in multichannel EEG representing distribution of electrical activity across the entire brain region, vary depending on placement of electrodes and functional specialisations. These signals also exhibit significant difference in strength. The surrounding electrodes contributes for this variability which can be interpreted as spatial interaction of EEG signals. With attention block, we generated attention map by making use of this EEG inter channel spatial relationship.

The proposed shallow spatial CNN attention model mainly comprised of convolutional block (CNN) and an attention block [figure 1]. 2 convolutional and pooling layers constitute CNN block. Number of filters and kernel size of CNN is set to 32 and 3 in both the layers respectively. EEG segments after normalisation are fed into CNN 1, to extract local features by convolving the input signals. ReLu is used as activation function to introduce non linearities to the extracted features. Subsequently max pooling layer with stride 2 is included for feature dimensionality reduction. Further the output is given to CNN2 followed by global average pooling

ISBN 979-8-3503-5094-4

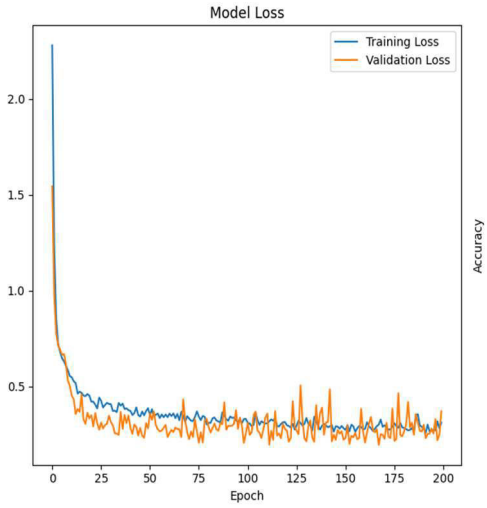


Figure 3: Model loss function

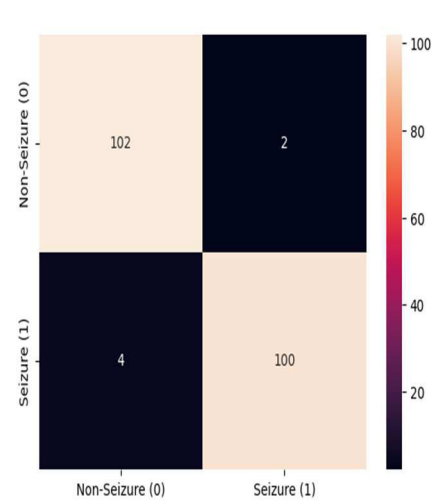


Figure 4 : Confusion matrix

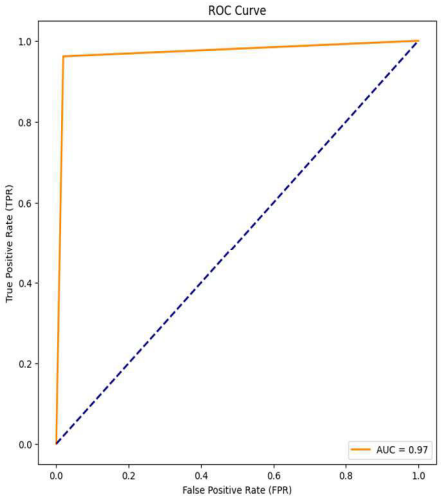


Figure 5: ROC Curve

which results average feature maps across the entire sequence, considered as global summary of spatial features. After each CNN layer a batch normalisation layer and dropout layer are provided to overcome the issue of overfitting. Further the convolutional block is followed by attention block as in figure 2.

In attention block, spatial feature information obtained from CNN are merged initially through average pooling, which computes average of each channel separately. Average activation for each channel is calculated along the temporal dimension/channel axis resulting in an efficient and enhanced spatial feature descriptor. Dot product is performed between the feature descriptor and input, generating attention scores. The operation basically performs weighted summation of hidden layer vector expression along the temporal dimension resulting a single weighted sum for each channel. The weight represents importance of each point in the EEG temporal sequence. Finally softmax function is applied for normalising the score and to obtain final weight according to the relevance. The resultant output tensor is mathematically represented as

$$C_i = \sum_{j=1}^{k} a_{i,j} \cdot h_{i,j} \quad (1)$$

where C represents output feature, interpreted as spatial attention features generated based on their relevance in global temporal context. a represents attention weights after sigmoid activation and h is hidden layer vector expression from CNN.

The architecture is finalised by one or more fully connected layer with softmax activation function for classifying into input datatype.

III. RESULTS AND DISCUSSION

Keras and tensor flow library in python 3 of google colaboratory notebook is used for conducting our experiments. We evaluated the approach of spatial CNN-Attention by performing stratified hold out validation strategy on noisy scalp EEG of CHBMIT dataset. For each patient, 2 or more files are randomly selected and are split into validation and test data, ensuring the inclusion of both seizure and non-seizure representations. Remaining files of each patient are utilised for training the model. We implemented our model using Adam optimiser with binary cross entropy as loss function and trained using batch size of 32 for 200 epochs. Accuracy, f1 score, sensitivity, selectivity, precision, mathew's correlation coefficient (mcc) calculated from the following equations (2-7) are used as evaluation indices.

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (2)$$

$$\text{Precision} = \frac{TP}{TP+FP} \quad (3)$$

$$\text{Sensitivity} = \frac{TP}{TP+FN} \quad (4)$$

$$\text{Specificity} = \frac{TN}{TN+FP} \quad (5)$$

$$\text{F1 - score} = \frac{2}{\frac{1}{precision}+\frac{1}{sensitivity}} \quad (6)$$

$$\text{MCC} = \frac{TP\ X\ TN}{\sqrt{(TP\ +FP\ )(TP\ +FN)\ (TN\ +FP\ (TP\ +FN}} \quad (7)$$

ISBN 979-8-3503-5094-4

where TP, TN, FP, FN represents true positive rate, true negative rate, false positive rate and false negative rate respectively. Figure 3 illustrates model loss function, that indicates how well the model fits to training as well as to new data. Notably, overall accuracy of 91.35% with test loss of 0.28, emphasise model's ability in making accurate predictions. Additionally, based on evaluation of confusion matrix depicted in figure 4, sensitivity of 96.15%, specificity (selectivity) of 98.08%, F1score of 97.09%, precision of 98% is obtained. Moreover, a high-quality value of 0.94 for MCC indicates strong concordance between the model's predictions and actual binary classes. It implies that our model classifies the seizure and non-seizure classes with minimal misclassifications. Receiver operating characteristics curve (ROC) plotted between TPR and FPR in figure 5 yields an area under the curve of 97.12, which further evaluates the diagnostic potential of the classifier,

On comparing our results with similar studies in table 1, our model performs better than existing models, achieving best results in classification reports in less computational time and expense. All the papers used for comparison employed similar deep learning architectures. The superior performance of our model can be attributed to several key factors in the model architecture. Specifically, inclusion of spatial attention mechanism allows the model to focuses on most significant information within the EEG segments across spatial dimensions. This enhances robustness of seizure detection by de-emphasizing noise and irrelevant information during prediction and classification. Generalizability of our model can further be improved by incorporating multiple datasets.

IV. CONCLUSION

We implemented a CNN based spatial attention model, an end-to-end deep learning network for seizure detection in our study. Our experimental findings outperform existing standard methods with the inclusion of spatial attention layer. The model exhibited a high true positivity rate (sensitivity) and low error rate and proved to be having high potential in addressing the challenges faced in existing models.

Table 1: Comparison of results with other models

Work	Dataset	Performance		
		Acc (%)	Sen (%)	Spec (%,)
[16]	CHBMIT	-	88	88
[17]	CHBMIT	92.9	94.3	91.5
[18]	CHBMIT	-	87.00	88.60
[19]	TUH	80.5	97.4	*88.1*
[20]	Clinical	82.86	80	-
Proposed model	**CHBMIT**	**91.35**	**96.15**	**98.08**

V. REFERENCES

[1] C. E. Stafstrom and L. Carmant, "Seizures and epilepsy: an overview for neuroscientists," *Cold Spring Harbor perspectives in medicine* , vol. 5, no. 6, 2015.

[2] C. E. Elger and C. Hoppe, "Diagnostic challenges in epilepsy: seizure under-reporting and seizure detection," *The Lancet Neurology,* vol. 17, no. 3, pp. 279-288, 2018.

[3] W. H. Organization, "Epilepsy," 2017. [Online]. Available: http://www.who.int/mediacentre/factsheets/fs999/en/.

[4] A. S. Zandi, M. Javidan, G. A. Dumont and R. Tafreshi, "Automated real-time epileptic seizure detection in scalp EEG recordings using an algorithm based on wavelet packet transform," *IEEE Transactions on Biomedical Engineering,* vol. 57, no. 7, pp. 1639-1651, 2010.

[5] U. R. Acharya , R. Yanti, J. W. Zheng, M. M. Rama Krishnan, J. H. Tan and R. J. M, "Automated diagnosis of epilepsy using CWT, HOS and texture parameters," *International journal of neural systems,* vol. 23, no. 3, 2013.

[6] A. Temko, E. Thomas, W. Marnane, G. Lightbody and G. Boylan, "EEG-based neonatal seizure detection with support vector machines," *Clinical Neurophysiology,* vol. 122, no. 3, pp. 464-473, 2011.

[7] M. Li, X. Sun and W. Chen, "Patient-specific seizure detection method using nonlinear mode decomposition for long-term EEG signals," *Medical & Biological Engineering & Computing,* vol. 58, pp. 3075-3088, 2020.

[8] T. Gandhi, B. K. Panigrahi and S. Ana, "A comparative study of wavelet families for EEG signal classification," *Neurocomputing,* vol. 74, no. 17, pp. 3051-3057, 2011.

[9] S. M. Hossain, S. U. Amin, M. Alsulaiman and G. Muhammad, "Applying deep learning for epilepsy seizure detection and brain mapping visualization," *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM),* vol. 15, no. 1s, pp. 1-17, 2019.

[10] Ö. Yıldırım, U. B. Baloglu and R. Acharya, "A deep convolutional neural network model for automated identification of abnormal EEG signals," *Neural Computing and Applications,* vol. 32, pp. 15857-15868, 2020.

[11] Q. Hongshuai, B. Deng, J. Wang, G. Yi, R. Wang and Z. Zhang, "Deep multi-scale feature fusion convolutional neural network for automatic epilepsy detection using EEG signals," in *39th Chinese control conference (CCC)*, 2020.

[12] M. Fan and C. A. Chou, "Detecting abnormal pattern of epileptic seizures via temporal synchronization of EEG signals," *IEEE Transactions on Biomedical Engineering,* vol. 66, no. 3, pp. 601-608, 2018.

[13] J. Guttag, "CHB-MIT Scalp EEG Database (version 1.0.0). PhysioNet.," 2010. [Online]. Available: https://doi.org/10.13026/C2K01R..

[14] A. H. Shoeb and J. . V. Guttag, "Application of machine learning to epileptic seizure detection," in *27th international conference on machine learning (ICML-10)*, 2010.

[15] A. Shoeb , H. Edwards, J. Connolly, B. Bourgeois, S. T. Treves and J. Guttag, "Patient-specific seizure onset detection," *Epilepsy & Behavior* , vol. 5, no. 4, pp. 483-498, 2004.

[16] P. Fergus, A. J. Hussain, D. Hignett, D. A. Jumeily, K. A. Aziz and H. Hamdan, "A machine learning system for automated whole-brain seizure detection," *Applied Computing and Informatics* , vol. 12, no. 1, pp. 70-89, 2016.

[17] M. Kaleem, D. Gurve, A. Guergachi and S. Krishnan, "Patient-specific seizure detection in long-term EEG using signal-derived empirical mode decomposition (EMD)-based dictionary approach," *Journal of neural engineering,* vol. 15, no. 5, 2018.

[18] X. Yao, X. Li, Q. Ye, Y. Huang and Q. Cheng, "A robust deep learning approach for automatic classification of seizures against non-seizures," *Biomedical Signal Processing and Control,* vol. 64, p. 102215, 2021.

ISBN 979-8-3503-5094-4

[19] X. Zhang, L. Yao, M. Dong, Z. Liu, Y. Zhang and Y. Li, "Adversarial Representation Learning for Robust Patient-Independent Epileptic Seizure Detection," *IEEE journal of biomedical and health informatics* , vol. 24, no. 10, pp. 2852-2859, 2020.

[20] W. Choi, M. J. Kim, M. S. Yum and D. H. Jeong, "Deep Convolutional Gated Recurrent Unit Combined with Attention Mechanism to Classify Pre-Ictal from Interictal EEG with Minimized Number of Channels," *Journal of Personalized Medicine* , vol. 12, no. 5, p. 763, 2022.

ISBN 979-8-3503-5094-4

EEG-Based Cognitive Impairment Classification using Support Vector Machine Algorithm

Karuppathal E[1]
Biomedical Engineering,
Rajalakshmi Engineering College,
Chennai- 602 105, India
karuppathal.e@rajalakshmi.edu.in
*Corresponding author

Kalpana R[2]
Biomedical Engineering,
Rajalakshmi Engineering College,
Chennai- 602 105, India
kalpana.r@rajalakshmi.edu.in

Deepsika D[3]
Biomedical Engineering,
Rajalakshmi Engineering College,
Chennai- 602 105, India
deepsika63@gmail.com

Senthil nadhan J[4]
Consultant Neurologist, Fortis Malar hospital, Chennai, India
sendil_dr@yahoo.co.in

***Abstract*— Electroencephalography (EEG) offers a non-invasive means to evaluate brain activity by recording brain signals, commonly employed in assessing cognitive functions and related dysfunctions triggered by external stimuli such as light, sound, and cognitive tasks. This study aims to detect cognitive impairment (CI) using resting-state EEG data obtained from 70 participants via 16-channel EEG recordings. Given that CI correlates with both physiological and psychological alterations in the brain, relevant features are extracted using algorithms like phase locking value (PLV) and band power. For binary classification, a support vector machine (SVM) algorithm is created that uses a radial basis function (RBF) kernel. The study's outcomes reveal that the classifier can discern EEG signals with an accuracy of 85.71%. In optimizing the feature set for the tasks at hand, 12 EEG channels, which exhibit relatively higher activity, are selected, leading to a 25% reduction in the feature set. Despite this reduction, SVM maintains the same classification accuracy of 85.71%. Notably, the proposed method demonstrates superior accuracy compared to contemporary research efforts in the field.**

***Keywords*— Phase locking value; Resting state EEG; Support vector machine; Cognitive impairment; Band power.**

I. INTRODUCTION

The brain being the vital organ in the central nervous system (CNS) [1], controls motor activities based on the input received from the sense organs. It has four lobes frontal, parietal, occipital and temporal [2] which are responsible for reasoning, problem-solving and processing the temperature, pressure, pain, memory [3], audio and visual information respectively.

Any imbalance in these activities leads to cognitive impairment (CI) [4]. It can be mild or severe; in first case there is a mild difference in the cognitive function while in second there is a loss in the ability to understand, recollect or to take decision. It occurs due to factors like ageing, cerebral stroke, accidental brain injury and chronic illness [5]. To identify these abnormalities, a variety of methods are utilized, including magnetic resonance imaging (MRI), computed tomography (CT), electroencephalography (EEG) [6], and screening tests including the Mini-Cog, MMSE, and Montreal Cognitive Assessment (MoCA) [7].

Among these, imaging techniques though accurate, are expensive and have restrictions in frequency of usage. MoCA test, MMSE are assessment tools based on questionnaires and evaluate the response of the patients and care-takers [8 & 9]. Depending on the score value, the stage of CI is assessed. Both tests show a good accuracy but specificity remains about 60% to 90% [10]. Measuring electrical activity on different regions of lobes, gives the information about brain function. EEG is one modality that records the brain activities in terms of signal. Through the use of an analog to digital convertor, this signal is transformed into digital format and shown on the monitor. There are five distinct frequency bands that make up brain waves: theta, alpha, beta, and gamma band. Theta band lies between 4-8 Hz has amplitude greater than 20μV. It arises during emotional, unconscious and deep meditation state [11]. Alpha band frequency range lies between 8-13 Hz occurs during relaxed state predominantly in occipital lobe. Beta band frequency range lies between 13-30Hz are associated with active things. Whereas cognitive and motor functions are indicated by brain signals that are above 30 Hz. This is termed as gamma band. [12].

II. COGNITIVE IMPAIRMENT DETECTED FROM EEG SIGNALS

EEG is typically used to investigate cognitive imbalance in cases of dementia, Alzheimer's disease (AD), and mild cognitive impairment (MCI) [13]. In a study, Czigler, B. et al. used conventional analysis and quantified linear and non-linear EEG complexity to identify early-stage AD using EEG. The brain signals of 12 AD patients and 24 healthy people were recorded by the authors under two different conditions: with eyes open and with their eyes closed. Omega-complexity, synchronization likelihood, and frequency spectra are obtained using the following parameters: 10% cosine window FIR, band pass filter in the range of 0.5 - 45 Hz, and sampling frequency of 1000 Hz. High and low degrees of synchronization are correlated with low and high omega values, respectively. In AD, synchronization was noted between 0.5 and 25 Hz [14]. When 32 channels EEG is used to record brain function of 75 subjects including normal, MCI and AD by McBride et al., an accuracy of above 80% was achieved by using quadratic kernel function. For this, twenty-four features from regional spectral information and complexity analysis were used along with support vector machine (SVM) algorithm to make binary discrimination between the groups like i. MCI Vs. Healthy control (HC), ii. AD Vs. HC, and iii. MCI Vs. AD [15].

Analyzing EEG recordings in various states, such as during rest, with open eyes, and with closed eyes, helps reveal the connectivity patterns among distinct brain regions. In

ISBN 979-8-3503-5094-4

different frequency bands, measurement of phase synchronization exhibits the level of functional connectivity. Reduction in alpha activity was observed by Toth.et.al., while examining EEG of 25 subjects between eye closed and eyes open condition [16]. Investigating the power and coherence of 35 individuals with MCI and 34 HC during both rest and a working memory task, the functional dysconnectivity in cerebral cortex are evident by the higher EEG power of person with MCI compared with the HC in rest condition and in working memory task, the person with MCI has higher coherence [17]. Estimating coherence through spectral power obtained at different lobes also differentiates normal, MCI and AD. Hogen.et.al recorded EEG of 20 volunteers. The upper alpha coherence value is found to be diminished between central and right temporal cortex in the person with AD. Even mild AD will inhibit upper alpha in regions of temporal lobes [18].

To scrutinize the functional connectivity of the brain besides the phase locking index (PLI), graph derived metrics, the power spectrum density (PSD) is also considered to be a vital parameter. Multilayer perceptron (MLP) classifier shows an accuracy of 91.02% [19].

From this discussion, neuronal communication, memory, and synchronization between lobes are the indication for assessing cognitive skill. Since this is a functional aspect of analysis, EEG proves to be one efficient modality.

III. Workflow

The aim of this work is to classify cognitive impaired subjects from the normal category using EEG signals. After acquiring EEG signals, band pass filter is used to separate five different frequency bands. To understand phase synchrony between different regions of the brain, phase locking value (PLV) is used to compute the relevant features. In addition to this power spectrum analysis also gives rise to 16 features for each frequency band. Finally, the signals are classified for normal or cognitive imbalance. The overall work flow is illustrated in Figure 1.

IV. DATA COLLECTION

In total 70 volunteers in the age group between 25 and 70 years participated in this study. Out of this 35 are normal (51.971 ± 11.377 years) and the remaining 35 are cognitive impairment (59.714 ± 6.644 years) as diagnosed and suggested by neurologist. The EEG was acquired using a commercial FDA approved 16-channel EEG system in standard 10-20 electrode montage placement (Longitudinal-Bipolar montage), signals are acquired at a sampling rate of 250 Hz and then stored for offline analysis. Interfacing impedance was adjusted and maintained based on participant condition. EEG signals are recorded in air-conditioned and sound-proof room. Subjects with cardiovascular illness having any kind of implants, have undergone surgery in the past five years and pregnant women are excluded from this study.

On arrival, all the subjects were asked to relax for 15 minutes. Then they are permitted into the recording room. After carefully placing all the electrodes on their scalp, brain signals are recorded during three states: awake, eyes close and eyes open. On each state, brain signals are recorded for 60 seconds for further processing. All participants were adequately briefed on the nature and objectives of the tests. Prior to initiating the procedure, written consent was secured from each participant. This study is approved by the institutional ethical committee with registration number REC/IEC/002/2021.

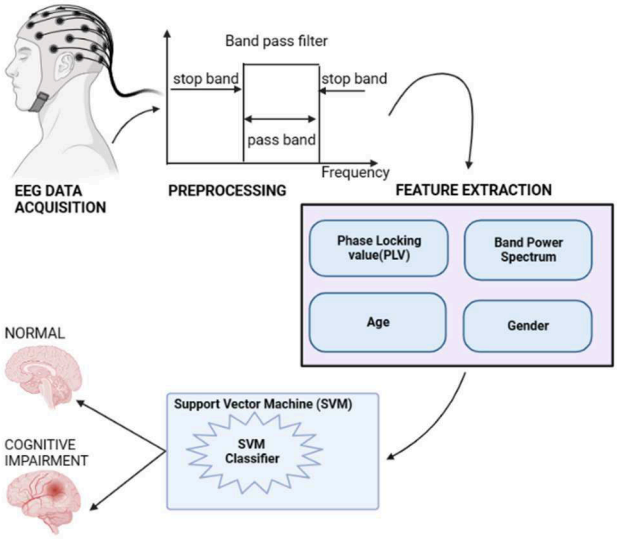


Fig. 1. Block diagram of proposed work

V. Preprocessing

EEG signals obtained exhibit a wide frequency range. In order to comprehend the nature of responses and synchrony during each examination state, these signals undergo preprocessing via a band-pass filter. Different frequency bands, including theta (4.1-8 Hz), alpha (8.1-13 Hz), beta (13.1-30 Hz), low gamma (30.1-50 Hz), and high gamma (50.1-80 Hz), are extracted. The sample steps followed for recorded EEG signal preprocessing is shown in Figure 2.

VI. Synchronization of Brain Activity

The brain, functions by integrating and differentiating the information it perceives. This happens due to interconnection of sets of neurons, leading to simultaneous response of neurons in different regions to any external stimuli. The electrical voltage thus arising when measured though EEG show synchronization and de-synchronization between two channels [20]. In some cases, summation of activities from synchronized input enhances the transmission of information effectively [21].

The closing and opening of different ionic channels lead to extracellular ion concentration. Therefore, quantification of coupling between neurons or nerve cluster becomes one important parameter to assess cognitive skills [22]. To quantify this activity, two methods are adopted in this paper. One is through assessing coupling between EEG signals in terms of amplitude and phase, that could be measured though a technique called PLV. This gives better insight in assessing brain connectivity in specific region (electrode position) on different states [23].

ISBN 979-8-3503-5094-4

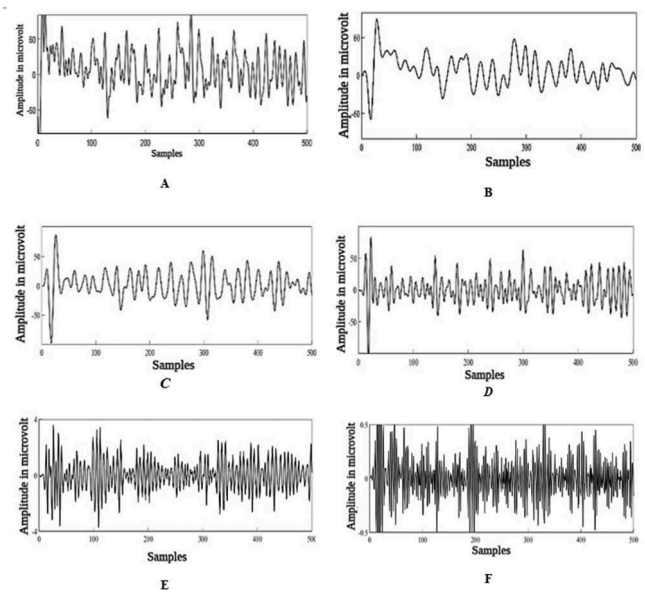


Fig. 2. Sample EEG signal of an individual (A) Raw EEG signal (B) Theta band (C) Alpha band (D) Beta band (E) Low gamma band (F) High gamma band

Second is by computing power of different frequency band of brain signal on different location helps in assessing the dominant frequency and its interconnection between channels. This way PLV and power estimation together brings information regarding brain activity to assess cognitive skill.

1. Phase locking value metric

The PLV (Phase Locking Value) serves as a statistical tool for exploring alterations in long-range synchronization of neural activity in EEG data induced by tasks [24]. This metric offers a means to assess temporal relationships among neural signals, independent of their amplitude [25]. In simpler terms, when the EEG signals from two channels (electrodes) exhibit more synchronous rises and falls during an experimental condition compared to a baseline value, it indicates increased synchronization or enhanced connectivity between these electrodes. Conversely, if it is lower than the baseline value, it signifies de-synchronization or reduced connectivity between the two electrodes [26].
The PLV computation steps as follows:

a) Extraction of phase component using Hilbert transform

The process of convolving the signal with the impulse response ($1/\pi t$) of the linear time-invariant filter is referred to as the Hilbert transform. It is normally applied for narrow band signals [27] and splits that into magnitude and phase component. Considering EEG signals as series of crests and troughs, Hilbert transform works on it in the range of $-\pi$ to π, assuring their limits for peak of crest and bottom of trough. When Hilbert transform is applied for all the 5 bands of 16 channel EEG signal, it separates into corresponding magnitude and phase components. This results in 16 ×16 matrix for an individual in specific band and under specific condition.

b) PLV metric estimation

The 16-channel longitudinal bipolar montage is strategically chosen for spatial specificity, capturing distinct neural activities in functional brain regions and facilitating differentiated analysis of frequency bands. Each channel's impact is critical for understanding temporal dynamics, statistical considerations in classification, and exploring electrode pair interactions for functional connectivity insights. PLV is a measure used to quantify the phase synchronization between any two EEG channels [28]. After computing phase difference, it is summed up as in equation (1). This is called phase locking value, that indicates the synchronization between two regions of brain as a specific activity. PLV computation is done using equation 1

$$PLV = \frac{1}{n}\left(\sum_{m=1}^{n} e^{j(\theta 1-\theta 2)}\right) \quad \text{-----(1)}$$

where n is the total number samples at time point m, θ1 is the phase of one EEG channel and θ2 is the phase of another EEG channel.

If a stimulus during an experiment causes the signals in two electrodes to rise and fall simultaneously or with a specific time delay, the phase difference ($\Delta\theta$) will exhibit consistency across different trials. On the other hand, if there is no discernible pattern regarding when the signals in these two electrodes rise and fall in relation to each other, the phase difference ($\Delta\theta$) will be random [28]. Figure 3 illustrate the process of PLV computation.

If the PLV value between the two channel seems to more, then it can be understood that there is a high functional connectivity among that region of brain for a particular stimulus [29].

1. Band power estimation

Power spectrum represents distribution of signal strength at different frequency. As the brain electric signals based on the frequency content shows various functions [30]. After segmenting the acquired EEG signal into 5 bands (theta, alpha, beta, low gamma, high gamma), Fourier transform as in equation (2) is used to compute the average power in all the five bands. This gives 5 power values for one channel and 80 values for 16 channels in one state. Since EEG recordings is performed on three states for all the individual, it gives rise to 240 metrics for every subject.

Formula for the computation of Discrete Fourier Transform (DFT) is equation 2

$$X(k) = \sum_{n=0}^{N-1} x(n)\, e^{-j\,2*pi*n*k/N} \quad \text{-----(2)}$$

Where x(k) represents complex result of the DFT at the frequency index k. frequency domain x(n) represents input sequence of length N at the discrete time index n.

ISBN 979-8-3503-5094-4

VII. DATASET CREATION

After completing PLV and band power computation, 240 data for everyone. The data set is formed taking 240+240+2=482 data (PLV and power data from 5 bands in all three cognitive tasks) while the last two is for age and gender which is shown in Table I . Thus, we have created a database containing 482 columns (last column for target, 1 for normal and 0 for test) and 70 rows (1 row for one participant). This is used for classification procedure.

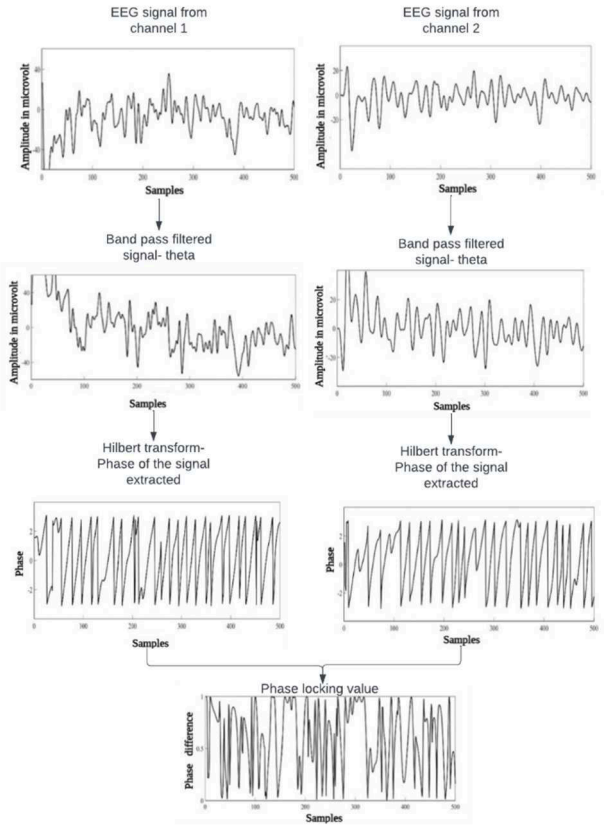


Fig. 3. The illustration showing PLV estimation method

VIII. SUPPORT VECTOR MACHINE CLASSIFIER

The SVM algorithm is one of the supervised machine learning tool for two-class classification [31 & 32], in which the models are trained using labelled dataset. Once training is done, the model is tested via the testing data and predicts the output [33]. Generally, SVM classifies data within a hyper plane in the feature space, employing various kernel functions like Gaussian, Radial Basis Function (RBF), linear, and polynomial [34]. Support vectors in SVM are determined. based on the proximity of test data to the hyper plane [35].

In the present work, SVM classification is performed by using LIBSVM (library for support vector machine) [36] in MATLAB. Using the trial-and-error method, the optimal kernel function is obtained. In the current study, RBF kernel is used (as given in equation 3) in the samples which essentially cannot be a linear sample.

TABLE I. Data set created for SVM classifier from EEG signal for an individual

Cognitive task		Feature					Count
		Theta	Alpha	Beta	Low gamma	High gamma	
Band power	Awake	16	16	16	16	16	80
	Eyes closed	16	16	16	16	16	80
	Eyes open	16	16	16	16	16	80
PLV	Awake	16	16	16	16	16	80
	Eyes closed	16	16	16	16	16	80
	Eyes open	16	16	16	16	16	80
Age		-	-	-	-	-	1
Gender		-	-	-	-	-	1
						Total	482

These sample is mapped into a higher dimensional space. Two parameters, C and GAMMA (γ) are used by RBF kernels. The error parameter, denoted by C and γ represent the curvature width of the radial basis function and γ was set to 0.01, while C was set to 1 [37]. This value is determined by conducting many experiments for different values. The following equation (3) showing the RBF kernel is the form P(r, s).

$$P(r,s) = \exp(-\gamma^{*}|r-s|^{2}) \quad \text{-----(3)}$$

Overall architecture of the constructed SVM classifier is illustrated in Fig. 4. A total of 480 features are obtained from the EEG signal (PLV & band power) under the cognitive tasks like awake, eyes open and eyes closed. Out of this, dataset (as illustrated in Table I), 80% of the data is used for the training and the remaining used for the testing. The network now able to predict the result with an accuracy of 85.71%, which is computed as per equation 4. The RBF kernel's efficiency is particularly noteworthy, with a processing time of 0.5 seconds per sample, compared to the Linear kernel's 1.5 seconds per sample. This demonstrates a remarkable threefold improvement in processing speed, highlighting the advantage of utilizing the RBF kernel in terms of both accuracy and computational efficiency.

The implementation was performed on a desktop equipped with an Intel(R) Core(TM) i7-10700K CPU at 3.80GHz and 3.70 GHz processor and a Windows 64-bit operating system. The capabilities of this hardware configuration likely contributed significantly to the overall efficiency of the SVM classifier.

$$\text{Accuracy} = \frac{(T_{+}+T_{-})}{(T_{+}+T_{-}+F_{+}+F_{-})} \quad \text{-----(4)}$$

where $T+$ is true positive, $T-$is true negative, $F+$is false positive and $F-$is false negative [38]. Since temporal, occipital, and frontal lobes gets more activated during eye open and eye closed [39], twelve numbers of EEG channel arising from these regions are selected. Features arising from

ISBN 979-8-3503-5094-4

these channels amount to 60 per frequency band and per task. This totals to 362 features per individual. When SVM is now tested and trained in the ratio 80:20 data, accuracy of the network remains the same. The result of these two networks, when it takes 20% of data (14 numbers) is presented in confusion matrix as shown in Fig. 5.

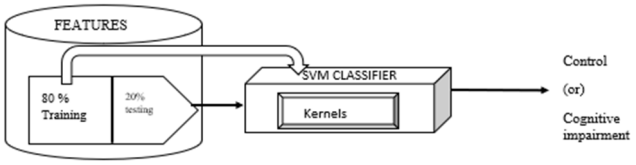


Fig. 4. Architecture of SVM binary classifier

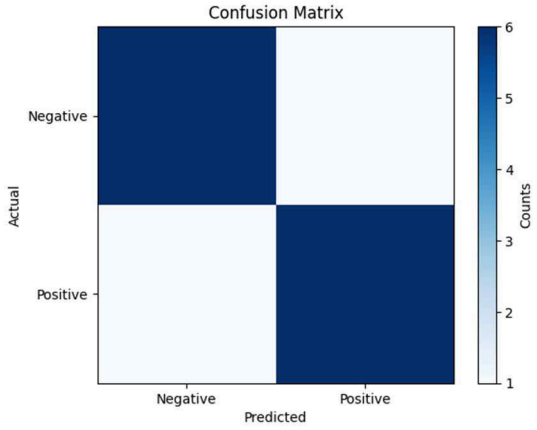


Fig. 5. Obtained confusion matrix from SVM binary classifier

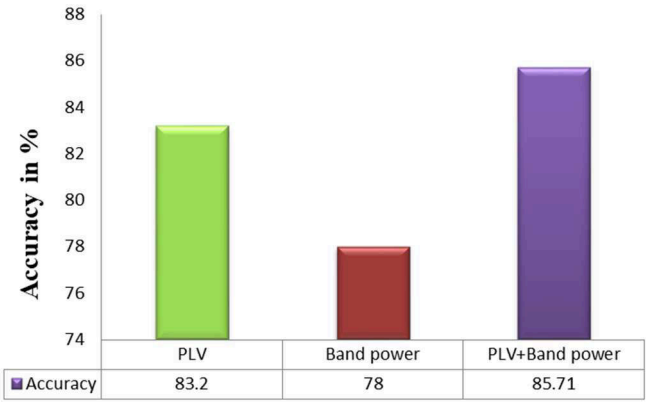


Fig. 6. Classification accuracy through PLV and band power features

The classification Accuracy is plotted against the features like band power, PLV and combination of both which is shown in Fig. 6.

IX. VALIDATION

All volunteered participant's cognitive status was assessed using the MoCA score. The MoCA test is a cognitive power screening tool, and and the maximum score is typically 30 points. A normal MoCA score is generally considered to be around 26 or higher. Scores below 18 may indicate presence of cognitive impairment. The mean and standard deviation of MoCA score of control group and test group is 27.971±1.464 and 18.228±5.851 respectively. This is now compared with the result of SVM classifier to validate the proposed method that takes 3 different inputs as PLV, band power and combination of these two. The result shows that a minimum error of 14.29±2.341 % is obtained when training SVM classifier with combination of PLV and band power.

X. DISCUSSION

Attempted in this work is to classify CI from resting state EEG. After acquiring EEG signal, two different approaches; PLV and power spectrum are used to extract relevant features. For this all the five band frequencies are used. The first one, lobes into localized coupling between phase and amplitude. While the second lobes into dominant frequencies. When these features were used for classifying CI using SVM algorithm, the results show (as given in Fig. 6) higher accuracy when PLV and band power features are used together. Since CI is a neurodegenerative disease, it shows changes both structural and neuronal connectivity when viewed through imaging modality like MRI and PET [40]. This is reflected as functional changes when viewed through EEG. Though PLV takes neuronal connectivity into account, the state of the subject is quantified by band power computation. Thus, when physiological and psychological aspects are combined, the classification accuracy increase.

For the tasks used in this study, frontal, occipital and temporal is higher. Therefore 12 channels arising from these their lobes are selected. The feature set now becomes 60 in each cognitive task under each frequency band. Now this total to 362 features. When the same SVM is used with similar testing and training procedure (as said in section VIII), classification accuracy remains unaltered. This way even by reducing the data by 25%, there is no compromise in classification accuracy. To our knowledge, the proposed method yields better accuracy when compared with contemporary research work, as presented in Table II. The classification accuracy achieved in this study is highly competitive when compared to results reported in other references that utilized the same dataset.

XI. CONCLUSIONS

The conclusion of the work is the combined information related neuronal synchrony through PLV and mental states through band power estimates, able to discriminate CI from controls. Also, effective features are identified (with 75% of original data size) that can produce the same result without compromising in the accuracy. From Figure 6.2 it could be observed that, PLV metric yields maximum accuracy. This is however further increased by band power features. The PLV is an indicative measure of neuronal synchrony. Since the study aims in differentiating cognitive ability using simple tasks, the neuronal synchrony proves to be an effective feature. However this model could be extended for classifying different stages of CI and even for predicting/early diagnosis. In future, resting state analysis can be extended with task-related activation pattern analysis. In spite of this, implemented SVM is not tested on some complex cognitive task oriented EEG datasets. In addition, SVM-RBF is mainly used for binary classification of EEG signals. We hope to use this model for multi class

ISBN 979-8-3503-5094-4

classification of multiple stages of CI and this could even help for predicting/early diagnosis.

TABLE II. Comparison of the performance of the proposed work with earlier published results of SVM classifiers

Authors	Number of subjects and Dataset	Features	SVM kernel	Obtained Accuracy
Li et al.,2021[41]	N=49; Train 60% and test 40%	Clustering coefficient, global efficiency and average node degree features used for two different SVM models: DTF(directed transfer function) directed network and PSI (phase synchronization index) undirected network	RBF	HC Vs amnestic MCI (aMCI) accuracy 80.0% for DTF model and 66.6% for PSI model
Trinh T-T et al., 2021 [42]	N=74; Train 60% and test 40%	Spectral power extracted from Working memory task	RBF	MCI Vs HC : 80.39%; AD Vs HC: 78%
Sedghizadeh M.J et al., 2022 [43]	N=35; 5-fold cross validation	Percent phase locking, amplitude coherence, PAC features extracted from oddball olfactory task	RBF	Over all accuracy between HC Vs aMCI and AD: 73 %
Miao Shi et al., 2020 [44]	N=40; Train 70% and test 30%	Time domain feature such as eigenvalues and root mean square values extracted from only 3 channels (central lobe)	RBF	85.9%
Proposed work	N=70; Train 60% and test 40%	Power & PLV	RBF	85.71%

Funding
This study has not received any financial support.

Conflicts of Interest
All authors disclose that they have no conflicts of interest.

REFERENCES

[1] Korn, T. and Kallies, A, "T cell responses in the central nervous system", Nature Reviews Immunology, Vol.17 No.3, pp.179-194, 2017.

[2] Bos, M.G., Peters, S., van de Kamp, F.C., Crone, E.A. and Tamnes, C.K, "Emerging depression in adolescence coincides with accelerated frontal cortical thinning", Journal of Child Psychology and Psychiatry, Vol.59, No.9, pp.994-1002, 2018

[3] Alnajjar, F., Khalid, S., Vogan, A.A., Shimoda, S., Nouchi, R. and Kawashima, R, "Emerging cognitive intervention technologies to meet the needs of an aging population: a systematic review", Frontiers in Aging Neuroscience, Vol.11, pp.291, 2019.

[4] Tolea, M.I. and Galvin, J.E, "Sarcopenia and impairment in cognitive and physical performance," Clinical interventions in aging, pp.663-671, 2015.

[5] Kapasi, A. and Schneider, J.A, "Vascular contributions to cognitive impairment, clinical Alzheimer's disease, and dementia in older persons," Biochemical et Biophysica Acta (BBA)-Molecular Basis of Disease, Vol.1862 No.5, pp.878-886, 2016.

[6] Lai, C.Q., Ibrahim, H., Abdullah, J.M., Azman, A. and Abdullah, M.Z, "Convolutional neural network utilizing error-correcting output codes support vector machine for classification of non-severe traumatic brain injury from electroencephalogram signal", IEEE Access, Vol.9, pp.24946-24964, 2021.

[7] Saito, H., Yamashita, M., Endo, Y., Mizukami, A., Yoshioka, K., Hashimoto, T., Koseki, S., Shimode, Y., Kitai, T., Maekawa, E. and Kasai, T, "Cognitive impairment measured by Mini-Cog provides additive prognostic information in elderly patients with heart failure," Journal of Cardiology, Vol.76 No.4, pp.350-356, 2020.

[8] Ciesielska, N., Sokołowski, R., Mazur, E., Podhorecka, M., Polak-Szabela, A. and Kędziora-Kornatowska, K, "Is the Montreal Cognitive Assessment (MoCA) test better suited than the Mini-Mental State Examination (MMSE) in mild cognitive impairment (MCI) detection among people aged over 60? Meta-analysis," Psychiatr Pol, Vol.50 No.5, pp.1039-1052, 2026.

[9] Arevalo-Rodriguez, I., Smailagic, N., i Figuls, M.R., Ciapponi, A., Sanchez-Perez, E., Giannakou, A., Pedraza, O.L., Cosp, X.B. and Cullum, S, "Mini-Mental State Examination (MMSE) for the detection of Alzheimer's disease and other dementias in people with mild cognitive impairment (MCI)," Cochrane Database of Systematic Reviews, (3), 2015.

[10] Langa, K.M. and Levine, D.A. "The diagnosis and management of mild cognitive impairment: a clinical review," Jama, Vol.312 No.23, pp.2551-2561, 2014.

[11] Siuly, S., Li, Y. and Zhang, Y, "EEG signal analysis and classification," IEEE Trans Neural Syst Rehabilit Eng, 11, pp.141-14, 24, 2016.

[12] Dubbelink, K.T.O., Stoffers, D., Deijen, J.B., Twisk, J.W., Stam, C.J. and Berendse, H.W, "Cognitive decline in Parkinson's disease is associated with slowing of resting-state brain activity: a longitudinal study," Neurobiology of aging, Vol.34 No.2, pp.408-418, 2013.

[13] Aarsland, D., Creese, B., Politis, M., Chaudhuri, K.R., Ffytche, D.H., Weintraub, D. and Ballard, C, "Cognitive decline in Parkinson disease," Nature Reviews Neurology, Vol.13 No.4, pp.217-231, 2017.

[14] Czigler, B., Csikós, D., Hidasi, Z., Gaál, Z.A., Csibri, É., Kiss, É., Salacz, P. and Molnár, M, "Quantitative EEG in early Alzheimer's disease patients—power spectrum and complexity features", International Journal of Psychophysiology, Vol.68 No.1, pp.75-80, 2008.

[15] McBride, J.C., Zhao, X., Munro, N.B., Smith, C.D., Jicha, G.A., Hively, L., Broster, L.S., Schmitt, F.A., Kryscio, R.J. and Jiang, Y, "Spectral and complexity analysis of scalp EEG

characteristics for mild cognitive impairment and early Alzheimer's disease", Computer methods and programs in biomedicine, Vol.114 No.2, pp.153-163, 2014.

[16] Tóth, B., File, B., Boha, R., Kardos, Z., Hidasi, Z., Gaál, Z.A., Csibri, É., Salacz, P., Stam, C.J. and Molnár, M, "EEG network connectivity changes in mild cognitive impairment—preliminary results", International Journal of Psychophysiology, Vol.92 No.1, pp.1-7, 2014.

[17] Jiang, Z.Y, "Study on EEG power and coherence in patients with mild cognitive impairment during working memory task", Journal of Zhejiang University Science B, Vol.6 No.12, pp.1213-1219, 2005.

[18] Hogan, M.J., Swanwick, G.R., Kaiser, J., Rowan, M. and Lawlor, B, "Memory-related EEG power and coherence reductions in mild Alzheimer's disease," International journal of psychophysiology, Vol.49 No.2, pp.147-163, 2003.

[19] Varone, G., Boulila, W., Lo Giudice, M., Benjdira, B., Mammone, N., Ieracitano, C., Dashtipour, K., Neri, S., Gasparini, S., Morabito, F.C. and Hussain, A, "A Machine Learning Approach Involving Functional Connectivity Features to Classify Rest-EEG Psychogenic Non-Epileptic Seizures from Healthy Controls," Sensors, Vol.22 No.1, pp.129, 2021.

[20] Antonenko, P.D. and Keil, A, "Assessing working memory dynamics with electroencephalography," Cognitive load measurement and application: A theoretical framework for meaningful research and practice, pp.93-111, 2017.

[21] Timofeev, I., Bazhenov, M., Seigneur, J. and Sejnowski, T, "Neuronal synchronization and thalamocortical rhythms in sleep, wake and epilepsy," Jasper's Basic Mechanisms of the Epilepsies [Internet]. 4th edition, 2012.

[22] Wen, D., Zhou, Y. and Li, X, "A critical review: coupling and synchronization analysis methods of EEG signal with mild cognitive impairment," Frontiers in aging neuroscience, Vol.7, pp.54, 2015.

[23] Dvorak, D., Shang, A., Abdel-Baki, S., Suzuki, W. and Fenton, A.A, "Cognitive behavior classification from scalp EEG signals", IEEE transactions on neural systems and rehabilitation engineering, Vol.26 No.4, pp.729-739, 2018.

[24] Soni, B.K., Mishra, D., Richaria, G. and Beg, M.S, "Brain Computer interface: A survey on Feature Extraction Techniques for Brain Signal," Brain, Vol.5 No.11, 2016.

[25] Wang, Z., Tong, Y. and Heng, X, "Phase-locking value-based graph convolutional neural networks for emotion recognition", IEEE Access, Vol.7, pp.93711-93722, 2019.

[26] Solomon, E.A., Kragel, J.E., Sperling, M.R., Sharan, A., Worrell, G., Kucewicz, M., Inman, C.S., Lega, B., Davis, K.A., Stein, J.M. and Jobst, B.C, "Widespread theta synchrony and high-frequency desynchronization underlies enhanced cognition", Nature communications, Vol.8 No.1, pp.1704, 2017.

[27] Freeman, W.J, "Hilbert transform for brain waves", Scholarpedia, Vol.2 No.1, pp.1338, 2007.

[28] Brunner, C., Scherer, R., Graimann, B., Supp, G. and Pfurtscheller, G. "Online control of a brain-computer interface using phase synchronization," IEEE Transactions on Biomedical Engineering, Vol.53 No.12, pp.2501-2506, 2006.

[29] di Biase, L., Ricci, L., Caminiti, M.L., Pecoraro, P.M., Carbone, S.P. and Di Lazzaro, V, "Quantitative High Density EEG Brain Connectivity Evaluation in Parkinson's Disease: The Phase Locking Value (PLV)," Journal of Clinical Medicine, Vol.12 No.4, pp.1450, 2023.

[30] Karuppathal, E., Kalpana, R. and Srinivasan, A.V. "Brainwave entrainment through external sensory stimulus: a therapy for insomnia", Int. J. Medical Engineering and Informatics, Vol. 13, No. 4, pp.323–333, 2021.

[31] Cortes C, Vapnik V, "Support-Vector Networks", Mach Learn 20:273–297, 1995

[32] Mahesh, B, "Machine learning algorithms-a review," International Journal of Science and Research (IJSR), Vol.9 No.1, pp.381-386, 2020.

[33] Sen, P.C., Hajra, M. and Ghosh, M, "Supervised classification algorithms in machine learning: A survey and review", In Emerging Technology in Modelling and Graphics: Proceedings of IEM Graph, pp. 99-111, 2020.

[34] Al-Mejibli, I.S., Alwan, J.K. and Abd Dhafar, H, "The effect of gamma value on support vector machine performance with different kernels", International Journal of Electrical and Computer Engineering, Vol.10 No.5, p.5497, 2020.

[35] Toledo-Pérez, D.C., Rodríguez-Reséndiz, J., Gómez-Loenzo, R.A. and Jauregui-Correa, J.C, "Support vector machine-based EMG signal classification techniques: A review," Applied Sciences, Vol.9 No.20, p.4402, 2019.

[36] Shen, X.J., Mu, L., Li, Z., Wu, H.X., Gou, J.P. and Chen, X, "Large-scale support vector machine classification with redundant data reduction", Neurocomputing, Vol.172, pp.189-197, 2016.

[37] Jahed Armaghani, D., Ming, Y.Y., Salih Mohammed, A., Momeni, E. and Maizir, H. "Effect of SVM Kernel Functions on Bearing Capacity Assessment of Deep Foundations," Journal of Soft Computing in Civil Engineering, Vol.7 No.3, pp.111-128, 2023.

[38] Witten I, Frank E, Hall M, Data Mining, 3rd edition, Boston: Morgan Kaufmann, 2011.

[39] Marx, E., Deutschländer, A., Stephan, T., Dieterich, M., Wiesmann, M. and Brandt, T, "Eyes open and eyes closed as rest conditions: impact on brain activation patterns," Neuroimage, Vol.21 No.4, pp.1818-1824, 2004.

[40] Johnson, K.A., Fox, N.C., Sperling, R.A. and Klunk, W.E, "Brain imaging in Alzheimer disease," Cold Spring Harbor perspectives in medicine, Vol.2 No.4, pp.006213, 2012.

[41] Li, X., Yang, C., Xie, P., Han, Y., Su, R., Li, Z. and Liu, Y, "The diagnosis of amnestic mild cognitive impairment by combining the characteristics of brain functional network and support vector machine classifier," Journal of Neuroscience Methods, Vol.363, pp.109334, 2021.

[42] Trinh T-T, Tsai C-F, Hsiao Y-T, Lee C-Y, Wu C-T and Liu Y-H, "Identifying Individuals With Mild Cognitive Impairment Using Working Memory-Induced Intra-Subject Variability of Resting-State EEGs", Front. Comput. Neurosci, 15:700467, 2021.

[43] Sedghizadeh, M.J., Aghajan, H., Vahabi, Z. et al, "Network synchronization deficits caused by dementia and Alzheimer's disease serve as topographical biomarkers: a pilot study", Brain Struct. Funct, vol. 227, pp. 2957–2969 , 2022.

[44] Shi, Miao, Wang, Chao, Li, Xian-Zhe, Li, Ming-Qiang, Wang, Lu and Xie, Neng-Gang. "EEG signal classification based on SVM with improved squirrel search algorithm", Biomedical Engineering / Biomedizinische Technik, vol. 66, no. 2, pp. 137-152, 2021.

ISBN 979-8-3503-5094-4

Dynamic Connectivity Patterns in Resting State and Task-Based MEG: An Instantaneous Amplitude Correlation Approach

Suhas M.V
Department. of ECE
Manipal Institute of Technology
Manipal Academy of Higher Education
Manipal, India
suhas.mv@manipal.edu

N. Mariyappa
MEG Research Center
Department of Neurology.
NIMHANS
Bengaluru, India
mari.thilak@gmail.com

Anitha H
Department. of ECE
Manipal Institute of Technology
Manipal Academy of Higher Education
Manipal, India
anitha.h@manipal.edu

Sanjib Sinha
MEG Research Center
Department of Neurology.
NIMHANS
Bengaluru, India
sanjib_sinha2004@yahoo.co.in

Dr. Ravindranadh Chowdary M
Department of Neurology.
NIMHANS, Bengaluru, India
mundlamuri.ravi@yahoo.com

Dr. Raghavendra K
Department of Neurology.
NIMHANS, Bengaluru,India
raghoo2k@gmail.com

Dr. Ajay Asranna
Department of Neurology.
NIMHANS, Bengaluru, India
ajayasranna@gmail.com

Dr Viswanathan L.G
Department of Neurology.
NIMHANS, Bengaluru, India
vishwanath.iyer89@gmail.com

***Abstract*— Understanding dynamic functional connectivity is pivotal in unraveling the intricate dynamics of neural activity. This research leverages Magnetoencephalography (MEG) and Instantaneous Amplitude Correlation (IAC) to explore the evolving patterns of dynamic functional connectivity in the human brain. The study meticulously compares IAC outcomes during resting state and task-based MEG, offering insights into the adaptability of brain connectivity. A comprehensive literature review contextualizes the study within existing research, highlighting the relevance of dynamic functional connectivity analysis. The MEG data preprocessing employs advanced techniques, including artifact reduction and source estimation. The IAC analysis, featuring tensor factorization and k-means clustering, reveals distinctive connectivity patterns in various frequency bands. Results demonstrate pronounced transitions between connectivity states, particularly in the beta frequency bands during resting state MEG. This comparative analysis enriches our understanding of neural dynamics and connectivity fluctuations, paving the way for potential clinical applications. The study underscores the need for broader validation through expanded datasets, emphasizing the implications for cognitive neuroscience and clinical practices.**

Keywords*— *magnetoencephalography, dynamic functional connectivity, instantaneous amplitude correlation, resting state, task-based, brain frequency bands.

I. INTRODUCTION

Magnetoencephalography (MEG) stands at the forefront of neuroimaging techniques, offering a unique window into the dynamic interplay of neural activity [1],[2]. Central to our understanding of brain function is the concept of dynamic functional connectivity, a dynamic framework that elucidates the temporal evolution of interactions between different brain regions [3]. Dynamic functional connectivity shows us that the brain's network architecture is not static but undergoes continuous reorganization in response to internal and external stimuli [4-9].

In the field of MEG data analysis, an important area of study is comparing how dynamic functional connectivity looks in different brain states. The two states under consideration here are the resting state, characterized by the absence of explicit cognitive tasks, and the task-based state, marked by engagement in specific cognitive activities [10],[11]. By undertaking such a comparative analysis, we aim to delineate the unique signatures of dynamic functional connectivity during periods of intrinsic neural processing as opposed to cognitive task engagement.

To understand this complex web of neural connections, Instantaneous Amplitude Correlation (IAC) emerges as a valuable tool [12]. IAC enables the investigation of instantaneous fluctuations in amplitude correlation between different brain regions, offering a high-temporal-resolution lens into the ebb and flow of neural communication. This study hinges on the premise that scrutinizing dynamic functional connectivity through IAC can provide unprecedented insights into the brain's adaptability and response to varying states of activity.

This research goes beyond the technical details of brain imaging methods. It aims to shed light on the basic principles that govern how the brain's network works and changes. Understanding how the brain dynamically configures its connectivity patterns in response to different states is crucial not only for advancing our theoretical understanding of neuroscience but also for envisioning potential applications in clinical settings.

In the following sections, we delve into a comprehensive literature survey to give context to our study in the wider world of dynamic functional connectivity research. By building on existing knowledge, we position our investigation as a valuable contribution to the understanding of neural dynamics and connectivity fluctuations in the human brain.

II. RELATED WORKS

The literature survey aims to comprehensively review pertinent studies in the field, outlining their methodologies, key findings, and contributions to dynamic functional connectivity research in MEG studies.

Baker [13] pioneered dynamic functional connectivity exploration in resting-state MEG using hidden Markov models (HMM) and amplitude envelopes. Identifying eight distinct brain states, HMM revealed dynamic states consistent with known Resting State Networks (RSNs) and unveiled cross-network interactions. Further validation across diverse populations and cognitive contexts is necessary.

Antonakakis [14] focused on time-varying Cross-Frequency Coupling (CFC) patterns in resting state MEG, utilizing phase-amplitude coupling (PAC) to identify dynamic network transitions in mild traumatic Brain Injury (mTBI) patients. While promising for mTBI diagnosis, generalizability to other clinical populations and specificity in different frequency bands require further exploration.

ISBN 979-8-3503-5094-4

Sanfratello [15] conducted dynamic functional network connectivity analysis in resting-state MEG data, revealing three cluster states and significant differences in meta-state metrics. Additional research is needed to understand the complex interactions between dynamic connectivity alterations and symptom severity in schizophrenia patients.

Tewarie et al. [16] evaluated high temporal resolution metrics, including Phase Difference Derivative (PDD), Wavelet Coherence (WC), and Instantaneous Amplitude Correlation (IAC), showcasing their efficacy in detecting rapid connectivity changes. However, further investigation is warranted to compare their efficacy across different experimental paradigms and clinical populations.

Antonakakis [17] investigated abnormal alterations in beamforming source-reconstructed Dynamic Functional Connectivity (DFC) from mTBI patients, highlighting reduced mean node degree patterns within specific brain regions. Future research should aim to elucidate the neurobiological underpinnings of these alterations and their implications for functional recovery.

Jiang et al. [18] introduced a Time-Varying Dynamic Network (TVDN) framework applicable to fMRI and MEG/EEG data, demonstrating its ability to capture brain activity dynamics and detect state switching. Further validation across different imaging modalities and experimental paradigms is necessary to establish its reliability and generalizability.

In summary, diverse methodologies, from hidden Markov models to high temporal resolution metrics, have been employed in existing research to investigate dynamic functional connectivity in MEG studies. While these studies collectively contribute to understanding dynamic neural processes and connectivity fluctuations in various clinical and cognitive contexts, further research is needed to validate findings, explore underlying mechanisms, and translate findings into clinical applications.

III. METHODOLOGY

In this section, we delineate the comprehensive methodology employed to analyse dynamic functional connectivity in MEG, focusing on the preprocessing of resting state and task-based datasets, source estimation, and the application of IAC for detailed connectivity analysis.

A system with a 12th Gen Intel Core i9 processor, 128 GB RAM, and an NVIDIA RTX A5000 GPU was used to implement the algorithm. MATLAB R2023a was used as the main tool for analysis supported by Brainstorm [19], which is documented and freely available for download online under the GNU general public license.

A. Data

The dataset consisted of two resting state MEG data procured from the Brainstorm repository [20] and one task-based data procured from FieldTrip [21]

B. MEG Preprocessing

Prior to specific analyses for resting state and task-based MEG, both datasets underwent a rigorous preprocessing pipeline using the Brainstorm toolbox compatible with MATLAB®. This included:

- Maxfilter (Temporal Signal Space Separation): Applied for the removal of external noise interference and movement correction. This step ensured the data's integrity by minimizing artifacts caused by external sources.
- Notch Filtering: Power line noise frequencies at 60Hz 120Hz and 180Hz were filtered out to eliminate potential sources of interference (Fig. 1).
- Biological Artefacts Detection and Removal: Utilizing Independent Component Analysis (ICA), ocular and cardiac artifacts were identified and removed, further enhancing data quality (Fig 2).
- During the manual inspection, a thorough examination was carried out to identify and exclude epochs with artifacts, thereby ensuring the integrity of the dataset was maintained (Fig 3).

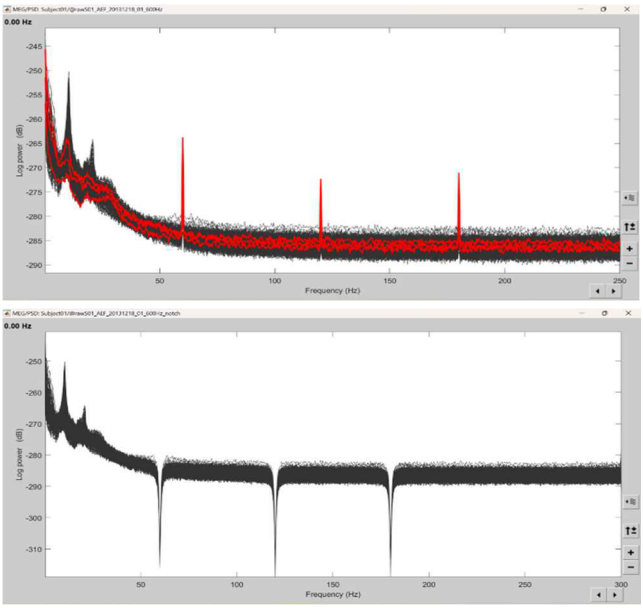


Fig. 1: (Top): MEG Power line interference at 60Hz, 120Hz, 180Hz seen on power spectral density plot using Welch method. (Bottom): PSD of the signal after notch filtering shows the spikes removed at the interfering frequencies.

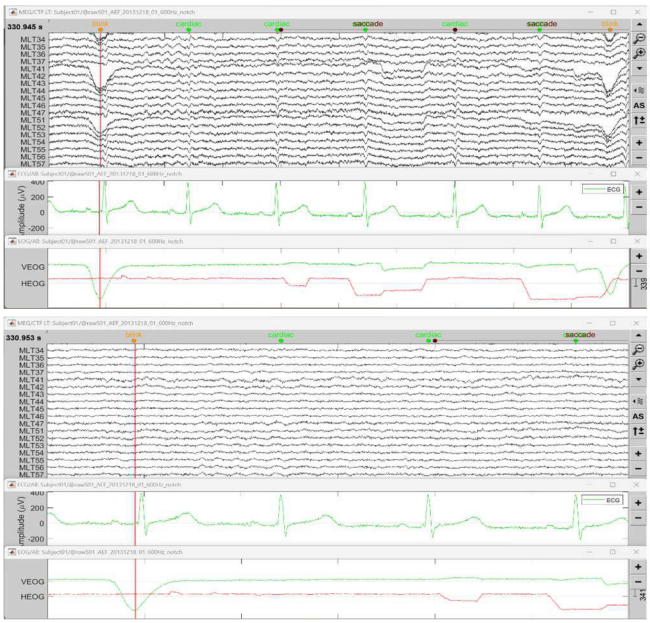


Fig. 2: (Top): Eye blink saccade and heartbeat interference observed corresponding to the EOG and ECG respectively (Bottom): Signal after application of signal space projectors.

ISBN 979-8-3503-5094-4

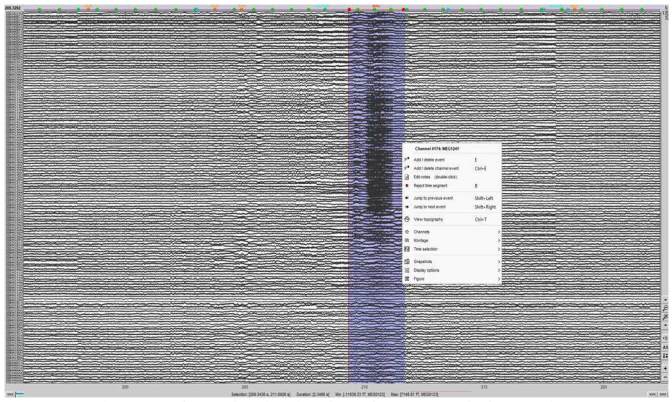


Fig. 3: Bad epoch rejection based on manual observation.

Dynamic connectivity analysis was performed at the source space level using individual MRI-based predefined atlases. The following steps were executed:

- Spatial Normalization: Cortical surfaces were extracted from individual MRIs using FreeSurfer [22] to obtain spatially normalized cortical surfaces (Fig 4).
- ROI Creation: The cortex of each participant was segmented into a 3D grid of voxels with specific resolution. Regions of Interest (ROIs) based on predefined atlases were created using MRI voxels.
- Lead Field Matrix Calculation: The lead field matrix for a unit dipole was calculated at each grid place using an overlapping sphere modelling method.
- LCMV Beamformer Estimation: Employing the linearly constrained minimum-variance (LCMV) [23] beamformer, 68 source time series were estimated, excluding cerebellum and sub-cortical structures from subsequent connectivity analysis (Fig. 5).

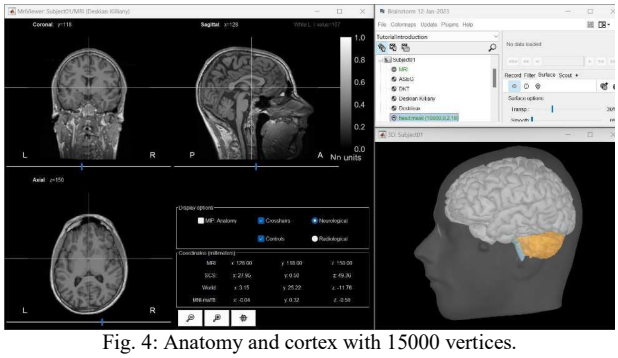


Fig. 4: Anatomy and cortex with 15000 vertices.

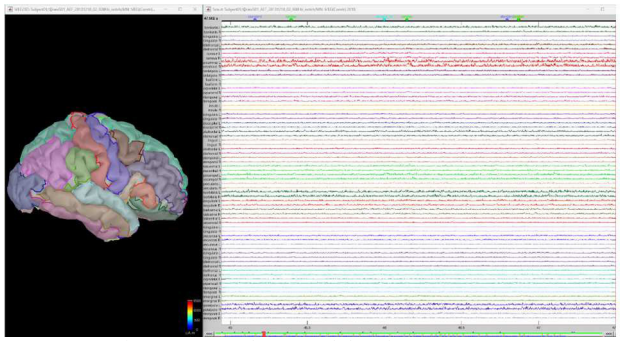


Fig. 5: Extracted temporal data from 68 ROIs.

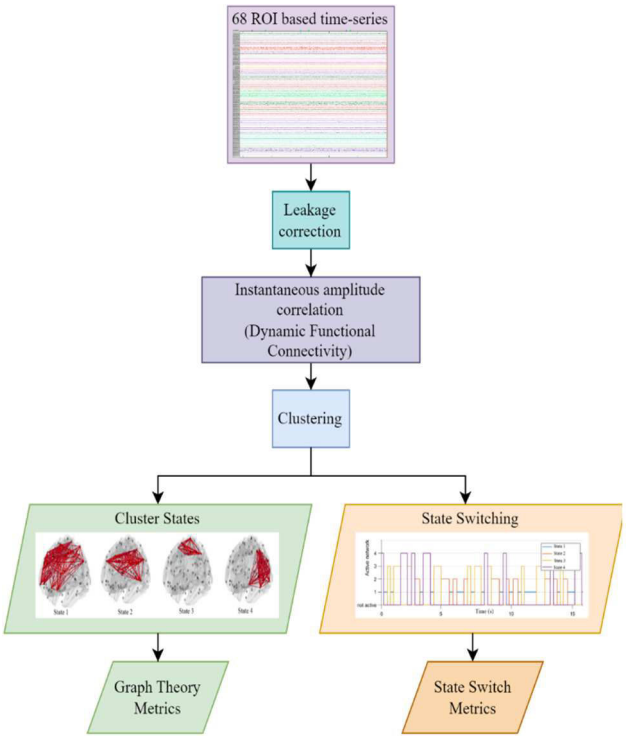


Fig. 6: Flowchart of dynamic functional connectivity computation

C. *Instantaneous Amplitude Correlation (IAC) Analysis*

Pre-processed MEG data files for individual subjects, filtered to specific brain frequency bands, were loaded for IAC computation. Z-score normalization was applied to the loaded data, cantering it around its mean and standardizing it for further analysis.

IAC was computed between different pairs of MEG scouts using the formula:

$$\mathrm{IAC}_{ij}(t) = \hat{\mathrm{E}}_i(t) \circ \hat{\mathrm{E}}_j(t) \quad (1)$$

where $\circ$ denotes the Hadamard product, and $\hat{\mathrm{E}}_i(t)$ and $\hat{\mathrm{E}}_j(t)$ are the Z-scores normalized amplitude envelopes of the respective time-series. This provided a measure of similarity in amplitude variations between two signals at each time point.

To account for potential interference or "leakage" between signals, a correction was applied through pairwise orthogonalization. This correction ensured that the correlation analysis accurately reflected the underlying neural activity.

Tensor Factorization and Clustering:

To extract underlying patterns within the MEG data, we employed tensor factorization using CP (CANDECOMP/PARAFAC) decomposition [24]. The method captures the complex multi-dimensional relationships inherent in the data. Unlike other decomposition methods, such as Singular Value Decomposition or Principal Component Analysis, CP decomposition is specifically tailored to handle multi-dimensional data structures, making it well-suited for analysing dynamic functional connectivity in MEG data. By decomposing the IAC data into a tensor representation, CP decomposition allows us to identify and characterize the intricate temporal dynamics of connectivity patters among the 68 MEG scouts over time.

ISBN 979-8-3503-5094-4

Fig. 7: A sample time zoomed temporal representation of state switching. A non-overlapping state switching is observed in the zoomed 50 second section between 410 to 450 seconds.

The factor matrix obtained from tensor factorization was clustered using k-means clustering. The number of clusters was determined based on the desired granularity of analysis; in this case, seven clusters corresponding to the seven networks of the brain were chosen. The seven clusters obtained from k-means clustering represent distinct functional networks within the brain, each playing a unique role in supporting various cognitive functions. These Commonly recognized networks include the default mode network, sensorimotor network, visual network, auditory network, executive control network, salience network, and dorsal attention network, interact dynamically to facilitate cognitive processes and respond to environmental stimuli. The steps of IAC computation are shown in Fig. 6.

D. Performance Parameters Evaluation

For both resting state and task-based MEG data, temporal and graphical metrics were evaluated on the output obtained from clustering of tensors:

We examined the duration of each state with State Span Analysis, investigated state transitions between clusters using State Transition Analysis to understand timing and frequency, and calculated the cumulative distance covered during state transitions for a quantitative measure of connectivity changes in MEG recordings.

While it is also important to measure the graph theory metrics such as degree of nodes, clustering coefficient, global efficiency, modularity, betweenness centrality from the clustered data at selected locations from the instantaneous amplitude correlation data, we limited our analysis to visually observe these graphs.

IV. Results and Discussion:

While the analysis steps were largely identical for both resting state and task-based MEG data, the differentiation lay in the interpretation and contextual understanding of the results. Resting state MEG analysis focused on characterizing the intrinsic connectivity patterns of the brain during periods of rest, providing insights into the brain's baseline functional organization and dynamics. Task-based MEG analysis emphasized how functional connectivity patterns changed in response to a specific cognitive task, highlighting the adaptability of the brain's functional connectivity during the task.

A. Resting State MEG Analysis:

The analysis of resting state MEG revealed intricate and distinct functional connectivity patterns during periods of intrinsic neural processing. The application of IAC in different frequency bands exposed unique signatures of connectivity dynamics. The clustering of the 68 MEG scouts based on their patterns of activity resulted in the identification of discrete connectivity states. State span analysis provided insights into the duration of each connectivity state, offering a temporal perspective on the stability of these patterns. Fig. 7 illustrates a time-scaled temporal depiction of state transitions. Within the magnified 50-second interval from 410 to 450 seconds, distinct state switches without overlap are evident.

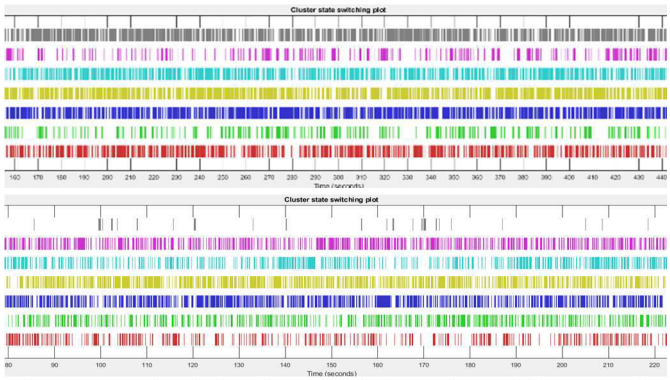


Fig. 8: Temporal variation of networks during rest into 7 different states in subject 1 (top) and subject 2(bottom) in β frequency band

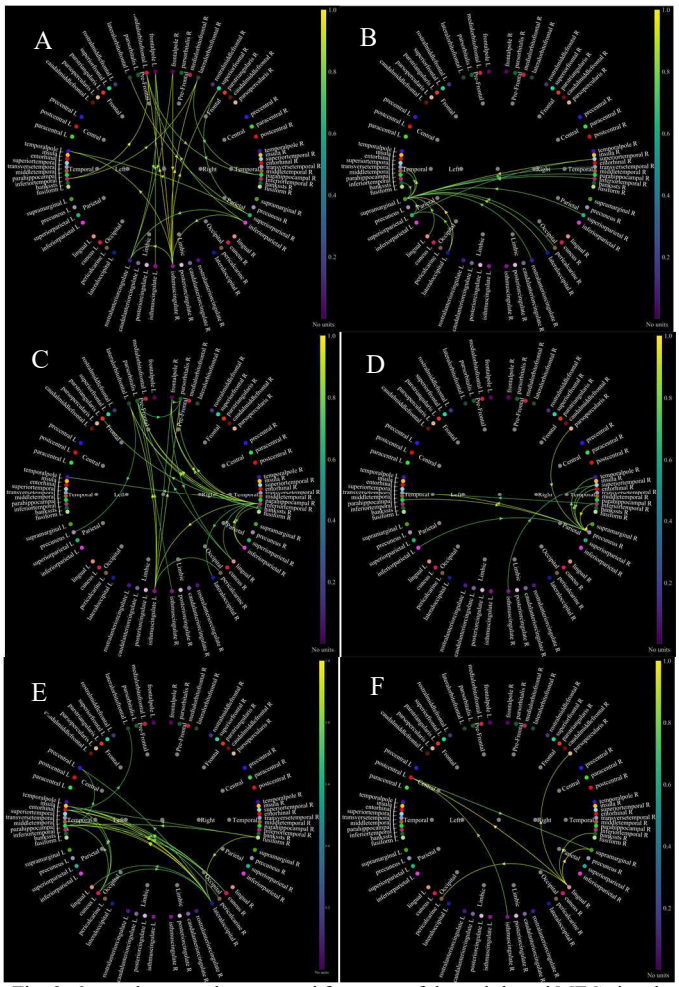


Fig. 9: 6 sample networks extracted from one of the task-based MEG signals at β frequency. The figures a-f represent the state network 1 to 6 respectively.

ISBN 979-8-3503-5094-4

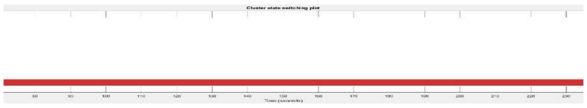


Figure 10: Temporal variation of network states during task in α (8-12 Hz) frequency band.

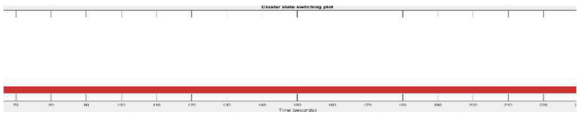


Fig. 11: Temporal variation of network states during task in γ (35-48 Hz) frequency band.

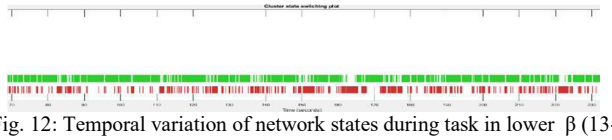


Fig. 12: Temporal variation of network states during task in lower β (13-20Hz) frequency band.

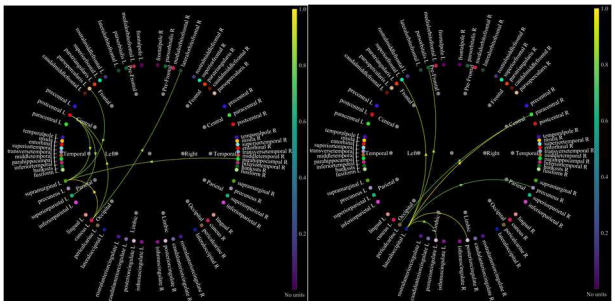


Fig. 13: Networks observed during task in β band. Left: state 1, right: state 2

Beta Frequency Bands:

In the beta frequency bands, the network exhibited dynamic transitions between multiple states during resting state MEG (Fig. 8). The state transition analysis illuminated the frequency and nature of transitions between these connectivity states. Interestingly, the total distance travelled during state transitions was notably longer, indicating substantial changes in connectivity during periods of rest. The observed transitions in the beta frequency bands may suggest a complex interplay of functional connectivity, possibly related to spontaneous neural activity and network reorganization. Out of 7 subnetworks, 6 networks during the rest are shown in Fig. 9.

B. Task-Based MEG Analysis:

Turning our attention to task-based MEG analysis, the application of IAC during cognitive tasks provided insights into how functional connectivity patterns adapted in response to specific cognitive demands. The clustering of MEG scouts in different frequency bands, particularly in the alpha and beta ranges, highlighted connectivity states during task engagement.

Alpha and Gamma Frequency Band:

In the alpha and gamma frequency band, the network predominantly remained in a single state during task performance (Fig. 10, 11). This sustained state in the alpha and gamma range suggests a stable and task-specific connectivity pattern, potentially reflecting focused cognitive engagement during the task.

Beta Frequency Bands:

Contrasting with resting state dynamics, the beta frequency bands exhibited transitions between two distinct states during task-based MEG (Fig. 12,12). Notably, these transitions appeared to be characterized by a longer span in the new state. This observation hints at the adaptive nature of the brain's connectivity in response to cognitive tasks, with the beta frequency range playing a significant role in facilitating these dynamic transitions.

Table I shows the state span analysis for resting state and task-based MEG dynamic connectivity. Resting state spans all seven states, indicating diverse interactions within the brain. During tasks, the β frequency network switches between two states, with longer transitions observed in new state 2. This suggests distinct mechanisms governing connectivity. Longer transitions in beta bands during rest may reflect ongoing cognitive processes, while stable alpha connectivity during tasks may indicate focused neural processing. These findings highlight the dynamic nature of brain connectivity and the differential mechanisms underlying resting state and task-based neural dynamics.

Table II presents a detailed analysis of state transitions within the beta frequency band during both task and rest conditions. The observations are organized into matrices, each corresponding to specific states (1 to 7). During the rest period, distinctive patterns of state transitions are observed, with multiple transitions occurring without overlap between states. Interestingly, reduced or no transitions are observed between states 2, 3, and 7 during certain segments of the rest period. This observation suggests a potential stabilization or synchronization of neural activity within these states, indicating a specific functional configuration or network dynamics. Conversely, during the task period, a clear transition pattern between two states within the beta frequency band is evident, reflecting task-driven modulation of neural connectivity. Overall, these findings underscore the dynamic nature of neural dynamics during rest and task conditions, with implications for understanding brain function and connectivity fluctuations.

Table III outlines the total distance travelled during task and rest conditions within the beta frequency band. In one of the rest periods for a total duration of 340 seconds, the distance travelled is observed to be 28,980 units, indicating a longer distance covered during this rest segment. In another rest period lasting 300 seconds, the total distance travelled is 17,832 units, again indicating a longer distance covered compared to other segments. During the task period lasting 283 seconds, the total distance travelled is 961 units, suggesting a shorter distance covered in comparison to the rest periods. In summary, the total distance travelled during beta frequency band activities, with rest periods consistently shows longer distances covered than the task period.

TABLE I. STATE SPAN ANALYSIS DURING TASK AND REST IN BETA FREQUENCY BAND

Task/rest	Rest(samples)	Rest(samples)	Task(samples)
Cluster 1	57492	36030	95107
Cluster 2	20760	43720	244769
Cluster 3	25140	60778	17
Cluster 4	47192	60981	35
Cluster 5	56455	46570	26
Cluster 6	54104	47503	28
Cluster 7	78857	3418	18
Comment	Goes through all the states with largest span in state 7	Goes through all the states with reduced span in state 7	Goes through 2 states with larger span spent in new state 2 which was not observed in α band.

ISBN 979-8-3503-5094-4

TABLE II. STATE TRANSITION ANALYSIS DURING TASK AND REST IN BETA FREQUENCY BAND

Task/ rest	Observation (number of transitions)								Comments
Rest	States	1	2	3	4	5	6	7	Transition between multiple states
	1	0	78	136	271	313	327	581	
	2	67	0	97	56	79	88	152	
	3	111	105	0	99	120	82	157	
	4	267	76	86	0	290	305	420	
	5	315	77	93	333	0	360	596	
	6	357	76	99	268	390	0	471	
	7	589	127	163	417	582	499	0	
Rest	States	1	2	3	4	5	6	7	Multiple state transition with no transitions between state 2,3 and 7
	1	0	189	147	245	215	148	1	
	2	168	0	302	333	247	215	0	
	3	163	306	0	558	297	312	0	
	4	244	332	561	0	330	348	35	
	5	204	249	293	334	0	281	1	
	6	165	188	333	341	271	0	21	
	7	1	1	0	39	2	15	0	
Task	States	1	2	3	4	5	6	7	Transition between 2 states
	1	0	468	0	1	0	0	0	
	2	468	0	0	0	0	0	0	
	3	0	0	0	0	0	1	1	
	4	1	0	0	0	0	0	0	
	5	0	0	0	1	0	0	1	
	6	0	0	1	0	0	0	0	
	7	0	0	1	0	1	0	0	

TABLE III. TOTAL DISTANCE TRAVELLED DURING TASK AND REST IN BETA FREQUENCY BAND

Task/rest	Total Duration	Observation	Comments
Rest	340	28980	longer distance travelled
Rest	300	17832	longer distance travelled
Task	283	961	shorter distance travelled

V. CONCLUSION

The insights gained from this study significantly contribute to cognitive neuroscience, offering a nuanced understanding of the brain's functional organization during rest and cognitive engagement. Dynamic patterns across frequency bands underscore the brain's adaptability, providing valuable insights into neural network dynamics.

While this study offers valuable insights, its limitations must be acknowledged. Statistical validation with a larger dataset is crucial for broader applicability, especially in clinical practice. Further research could explore implications for specific cognitive functions and clinical biomarkers.

Future studies could incorporate comparative analysis between IAC and other methods for dynamic functional connectivity assessment. This would offer a broader perspective on brain activity dynamics. By comparing methods, researchers can elucidate strengths, limitations, and underlying mechanisms.

In summary, while the comparative analysis of resting state and task-based MEG dynamics utilizing IAC provides valuable insights, incorporating comparative analysis with other methods would enrich our understanding of the dynamic nature of functional connectivity. This comprehensive understanding deepens our knowledge of how the brain dynamically adapts its connectivity patterns during both rest and cognitive tasks, opening avenues for further investigations into the cognitive and clinical significance of these dynamic connectivity dynamics.

REFERENCES

[1] E. Carrette and H. Stefan, “Evidence for the role of magnetic source imaging in the presurgical evaluation of refractory epilepsy patients,” Frontiers in Neurology, p. 933, 2019.

[2] S. Baillet, “Magnetoencephalography for brain electrophysiology and imaging,” Nature Neuroscience, vol. 20, no. 3, pp. 327–339, 2017.

[3] D. Vidaurre, A. J. Quinn, A. P. Baker, D. Dupret, A. Tejero-Cantero, and M. W. Woolrich, “Spectrally resolved fast transient brain states in electrophysiological data,” Neuroimage, vol. 126, pp. 81–95, 2016.

[4] R. M. Hutchison et al., “Dynamic functional connectivity: promise, issues, and interpretations,” Neuroimage, vol. 80, pp. 360–378, 2013.

[5] V. L. Morgan, C. Chang, D. J. Englot, and B. P. Rogers, “Temporal lobe epilepsy alters spatio-temporal dynamics of the hippocampal functional network,” NeuroImage: Clinical, vol. 26, p. 102254, 2020.

[6] A. Fallahi et al., “Dynamic functional connectivity in temporal lobe epilepsy: a graph theoretical and machine learning approach,” Neurological Sciences, vol. 42, no. 6, pp. 2379–2390, 2021.

[7] H. Li, Q. Zhang, Z. Lin, and F. Gao, “Prediction of epilepsy based on tensor decomposition and functional brain network,” Brain Sciences, vol. 11, no. 8, p. 1066, 2021.

[8] G. M. Duma et al., “Investigation of dynamic functional connectivity of the source reconstructed epileptiform discharges in focal epilepsy: A graph theory approach,” Epilepsy Research, vol. 176, p. 106745, 2021.

[9] X. Pang et al., “Abnormal static and dynamic functional connectivity in left and right temporal lobe epilepsy,” Frontiers in Neuroscience, p. 1906, 2021.

[10] A. M. Hermundstad et al., “Structural foundations of resting-state and task-based functional connectivity in the human brain,” Proceedings of the National Academy of Sciences, vol. 110, no. 15, pp. 6169–6174, 2013.

[11] M. C. Stevens, “The contributions of resting state and task-based functional connectivity studies to our understanding of adolescent brain network maturation,” Neuroscience & Biobehavioral Reviews, vol. 70, pp. 13-32, 2016.

[12] P. Tewarie et al., “Tracking dynamic brain networks using high temporal resolution MEG measures of functional connectivity,” Neuroimage, vol. 200, pp. 38–50, 2019.

[13] A. P. Baker et al., “Fast transient networks in spontaneous human brain activity,” eLife, vol. 3, p. e01867, 2014.

[14] M. Antonakakis et al., “Mining cross-frequency coupling microstates from resting state MEG: an application to mild traumatic brain injury,” in 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), IEEE, 2016, pp. 5513–5516.

[15] L. Sanfratello, J. M. Houck, and V. D. Calhoun, “Relationship between MEG global dynamic functional network connectivity measures and symptoms in schizophrenia,” Schizophrenia Research, vol. 209, pp. 129–134, 2019.

[16] P. Tewarie et al., “Tracking dynamic brain networks using high temporal resolution MEG measures of functional connectivity,” Neuroimage, vol. 200, pp. 38–50, 2019.

[17] M. Antonakakis et al., “Aberrant whole-brain transitions and dynamics of spontaneous network microstates in mild traumatic brain injury,” Frontiers in Computational Neuroscience, vol. 13, p. 90, 2020.

[18] F. Jiang et al., “Time-varying dynamic network model for dynamic resting state functional connectivity in fMRI and MEG imaging,” NeuroImage, vol. 254, p. 119131, 2022.

[19] F. Tadel, E. Bock, G. Niso, J. C. Mosher, M. Cousineau, D. Pantazis, R. M. Leahy, and S. Baillet, "MEG/EEG Group Analysis With Brainstorm," Frontiers in Neuroscience, Feb. 2019.

[20] R. Oostenveld, P. Fries, E. Maris, and J.-M. Schoffelen, "FieldTrip: Open Source Software for Advanced Analysis of MEG, EEG, and Invasive Electrophysiological Data," Computational Intelligence and Neuroscience, 2011; 2011:156869.

[21] F. Tadel et al., “Brainstorm: A User-Friendly Application for MEG/EEG Analysis,” Computational Intelligence and Neuroscience, vol. 2011, ID 879716, 2011.

[22] “Coordinatesystems—free surfer wiki [internet].” [Online]. Available: https://surfer.nmr.mgh.harvard.edu/fswiki/CoordinateSystems

[23] B. D. Van Veen et al., “Localization of brain electrical activity via linearly constrained minimum variance spatial filtering,” IEEE Transactions on Biomedical Engineering, vol. 44, no. 9, pp. 867–880, 1997.

[24] L. Gauvin, A. Panisson, and C. Cattuto, “Detecting the community structure and activity patterns of temporal networks: a non-negative tensor factorization approach,” PloS One, vol. 9, no. 1, p. e86028, 2014.

ISBN 979-8-3503-5094-4

A Robust Algorithm for respiration rate monitoring during cognitive load using PPG Signals

M.Muthu Pavithran
B.Tech student Department of Biomedical Engineering
Sri Ramakrishna Engineering College, Coimbatore., India
muthupavi2002@gmail.com

M.S.Ganeshmurthy
PhD scholar Department of Instrumentation and Control Engineering
National Institue of Technology, Tiruchirappalli
Tiruchirappalli, India
ganzmurthy@gmail.com

Dr. R. Periyasamy*
Department of Instrumentation and Control Engineering
National Institue of Technology, Tiruchirappalli
Tiruchirappalli, India
periyasamyr@nitt.edu

***Abstract*— This paper introduces a new method for respiration rate (RR) monitoring, utilizing the OB1203 PPG biosensor module and a custom signal processing algorithm. Traditional methods often involve manual intervention or wearable straps, posing challenges, particularly in emergencies. In contrast, the PPG-based system offers a contactless and user-friendly approach, particularly in pandemics. The RR monitoring method includes data acquisition, noise filtering, and feature extraction for systolic peak detection, with real-time RR display for seamless monitoring. The algorithm's validation used the BIDMC dataset & CAPNOBASE Dataset for calibration and validation, comparing with impedance respiratory and capnography based respiratory signals. Results showcase a superior performance of the proposed algorithm, as evidenced by low RMSE values of 0.856 and 1.92 for BIDMC, and CAPNOBASE datasets, respectively. Additionally, the project explores the RR-cognitive load correlation. Through a study involving tasks with varying mental effort, we contribute to cognitive load assessment. This innovative approach holds promise for vital sign monitoring efficiency, especially in emergencies, and cognitive load assessment in clinical and healthcare settings.**

Keywords—Respiration rate (RR), OB1203 photo plethysmography (PPG) biosensor module, BIDMC dataset, CAPNOBASE dataset, Cognitive load

I. INTRODUCTION

Respiration rate (RR) is a crucial vital sign that can serve as a key indicator of various medical conditions, encompassing adverse cardiac events, pneumonia, clinical deterioration, and response to stressors such as emotional stress, cognitive load, heat, cold, physical effort, and exercise-induced fatigue. Traditional methods for measuring RR include patient monitors, manual counts, capnography, Inductance plethysmography, bio-impedance, or piezoelectric sensors. However, these techniques often necessitate manual intervention for counting or require patients to wear sensor straps, presenting challenges, particularly in emergency situations.

In light of these considerations, a novel and non-contact method for continuous RR monitoring is imperative, especially in scenarios like pandemics. Among various approaches, photo plethysmography (PPG) emerges as a promising solution, surpassing other methods in terms of ease of use and adaptability. PPG-based RR monitoring eliminates the need for manual intervention or cumbersome sensor straps, offering a more seamless and patient-friendly experience. The following work is centred on leveraging the advantages of PPG-based respiration rate monitoring to enhance the efficiency and accessibility of vital sign monitoring, particularly in critical and emergency situations.

The methodology focuses on using a PPG module (OB1203) for data acquisition and a novel signal processing algorithm for advanced analysis of the acquired signals, which is tested and validated by using the BIDMC dataset of 53 subjects (PPG, impedance respiratory signal, and electrocardiogram (ECG)) from Physionet database and CAPNOBASE dataset of 42 subjects (PPG, electrocardiogram (ECG), and capnography signals). The algorithm is modified to acquire the respiration rate in real-time which is then used for the future scope on correlating respiration rate data with cognitive load in patients. The resulting methodology offers a novel, non-invasive approach to vital sign monitoring with significant potential for cognitive load assessment in clinical and healthcare applications.

A. *Objective*

- To develop a robust algorithm to detect the respiration rate from PPG signal.
- To validate the developed algorithm with the publicly available datasets.
- To study the relationship between the respiration rate and cognitive load in human subjects.

B. *Related Works*

One interesting, but unexplored, method for monitoring cognitive strain is the use of respiratory measurements. Basic and variable parameters of respiration rate (RR), partial pressure of end-tidal carbon dioxide (petCO2), and performance measures were measured in 63 male pilot candidates during a challenging cognitive task and recovery period in a study by Mariel Grassmann et al. 2016[1]. The findings demonstrated that while petCO2 dramatically dropped, mental strain was linked to a rise in respiration rate and a decrease in respiratory variability. While petCO2 did not reach the baseline, RR and respiratory variability indicated partial and total effects of recovery, respectively. A good performance was associated with a higher level of reactivity in RR. According to these results, respiratory measurements might be a helpful addition to standard metrics for determining mental stress during the pilot selection process.

ISBN 979-8-3503-5094-4

Numerous studies have shown that respiratory frequency is affected by various cognitive tasks and tends to increase in proportion to their difficulty level. This indicates that respiratory frequency is the ventilatory variable that predominantly reflects the cognitive load. During a state of rest, performing mental arithmetic, inhibition tasks, and working memory exercises can lead to an increase in respiration, while tidal volume may remain unchanged or decrease. Therefore, monitoring respiratory frequency may prove useful in detecting cognitive load in different situations Andrea Nicolò et al.(2020) [2].These studies utilized either contact-based respiration rate detectors or capnography to measure respiration rate. Previous studies did not use contactless PPG-based RR method to investigate the connection between respiration rate and cognitive load.

And the various methods have been proposed by the researchers to detect respiration rate from PPG signals. Karlen et al.(2013) [3] employed a smart fusion approach that involved extracting respiratory-induced variations (frequency, intensity, and amplitude) on PPG signal using Incremental-Merge Segmentation (IMS) and analyzing the frequency of each variation using fast Fourier transform (FFT). They demonstrated that the breathing rate was the arithmetic mean of frequencies from all three induced variations.

Mohammod Abdul Motin (2018) [4] and his team have developed a new algorithm to estimate heart rate (HR) and respiratory rate (RR) simultaneously from a photoplethysmography (PPG) signal. The algorithm is based on ensemble empirical mode decomposition with principal component analysis (EEMDPCA). To evaluate the performance of the algorithm, they tested it on datasets comprising 310 epochs from 35 subjects (recorded using electrocardiogram, PPG, and respiratory signals from MIMIC database) and 632 epochs from 42 subjects (recorded using the same signals from Capnobase database).

Talha Iqbal et al (2022) [5]. proposed a new algorithm that is resistant to signal loss and can handle low-quality signals from the sensor to estimate the respiratory rate using a photoplethysmography sensor signal for health monitoring. It employs selective windowing, preprocessing and signal conditioning, modified Welch filtering and postprocessing to improve accuracy and robustness to noise.

A model to predict respiration rates using PPG signals and a machine learning approach was created by Xuhao Dong et al. (2023) [6]. Even with low-signal-quality PPG signals, the model makes advantage of signal quality metrics to increase estimation accuracy. To build a very reliable model that takes signal quality considerations into account, the researchers employed a technique based on the whale optimization algorithm (WOA) and a hybrid relation vector machine (HRVM). Real-time RR estimation is done by the model using PPG signals. The researchers used simultaneously recorded PPG signals and impedance respiration rates from the BIDMC dataset to evaluate the effectiveness of the suggested model.

All the algorithms mentioned earlier are capable of predicting respiration rates based on the PPG signal, and they have been validated using publicly available datasets. However, their accuracy still poses a challenge to their implementation in clinical settings. On the other hand, the proposed method addresses the existing limitations and offers a better solution.

II. METHODOLOGY

The experimental setup utilizes an OB1203 photo plethysmography (PPG) sensor module for collecting physiological data from subjects. The PPG data is transmitted to a computer for signal processing via a USB cable. The Python programming language is employed for the analysis of acquired PPG signals, specifically for calculating the respiration rate.

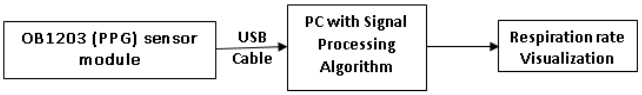


Fig. 1. Experimental setup for respiration rate detection using PPGsignal

A. Abbreviations and Acronyms

Photo-plethysmography (PPG) is a non-invasive optical technique that detects changes in the volume of blood in the microvascular tissue bed. It is obtained from the skin by illuminating it with red (700nm) and infrared (940nm) light using light-emitting diodes (LEDs). The intensity of light transmitted or reflected by the underlying blood vessels is measured. They exhibit characteristic changes in relation to the cardiac and respiratory cycle. The respiratory-induced modulation is reflected in both amplitude and frequency components of the PPG signal. During the inhale cycle, the intra-thoracic pressure changes cause a decrease in the stroke volume of the left ventricle, which leads to a smaller PPG amplitude. Similarly, during expiration, the left ventricle stroke volume increases, which results in an increased pulse amplitude. Thus, PPG signals can be used to estimate respiration rate. The time intervals between consecutive respiratory-related peaks or troughs in the PPG waveform correspond to the respiratory rate. Respiration-induced modulations can be observed by analysing changes in the baseline of the PPG waveform.

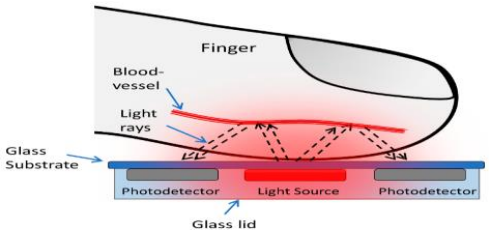


Fig. 2. Acquisition of photoplethysmogram (PPG) signal from the reflection method (*Source : https://www.mdpi.com/2079-6374/9/3/87*)

B. Hardware Setup

OB1203 Photoplethysmography (PPG) Biosensor: The OB1203 photo plethysmography (PPG) biosensor seamlessly integrates light sources, drivers, analog-to-digital conversion, and I2C communication within a compact optical package. This multifunctional biosensor is capable of providing essential physiological data, including heart rate (HR), oxygen saturation (SpO2), respiration rate (RR), and heart rate variability (HRV) – a key indicator of cognitive load. Distinguished by its advanced design, the OB1203 biosensor incorporates a longer wavelength far-red LED (690nm). This choice of wavelength enhances performance by reducing sensitivity to variations in light absorption caused by differences in skin tones.

Fig. 3. OB1203 Photo Plethysmography (PPG) biosensor *(Source : https://www.renesas.com/us/en/document/dst/ob1203-datasheet)*

ISBN 979-8-3503-5094-4

C. *Proposed Algorithm To Detect The Respiration Rate From Ppg Signal*

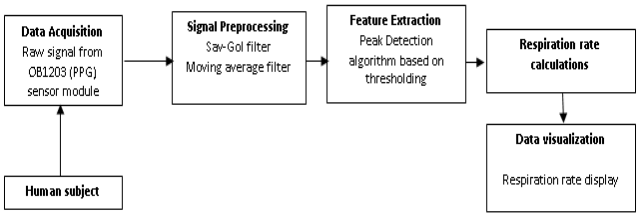


Fig. 4. Signal Processing block diagram for respiration rate detection using PPG sensor module

1) Data acquisition: The data acquisition process is continuous, providing an extensive dataset for subsequent analysis. The PPG signal is acquired using the OB1203 photo plethysmography (PPG) sensor module from subjects by placing their index finger on the module. The data is sent into the algorithm for signal processing.

2) Pre-processing: Data pre-processing is a crucial phase to guarantee the quality and reliability of the acquired signals. Raw PPG signals obtained from the sensor could include a variety of noise, such as interference from the environment, and motion artifacts. These undesirable components are reduced or removed using noise reduction techniques like Savitzky-Golay filter (Sav-Gol filter) and Moving average filter.

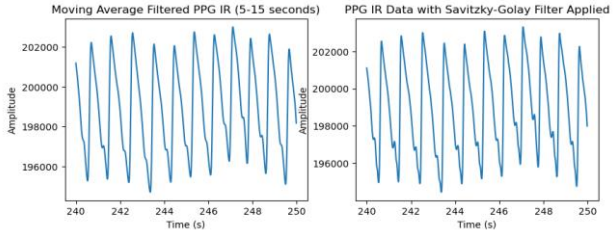


Fig. 5. Pre-processed PPG signal

a) Savitzky-Golay Filter: The Savitzky-Golay filter is a smoothing technique used for noise reduction in signal processing. It is particularly effective in preserving the shape of the original signal while minimizing noise. The filter works by fitting a polynomial within a sliding window and estimating the smoothed value at the center of that window. This method is advantageous for its ability to effectively reduce noise without causing significant distortion to the signal. The formula for the Savitzky-Golay filter involves convolving the signal with a set of coefficients derived from a polynomial fit within the specified window:

$$y[n] = 1/Q \sum_{i=-Q}^{Q} c_i * x[n+i] \quad (1)$$

Here, y[n] is the filtered output, x[n] is the input signal, and ci are the Savitzky-Golay coefficients.

b) Moving Average: The moving average is a simple and commonly used method for smoothing time-series data. It involves calculating the average of a set of consecutive data points within a specified window and assigning this average value to the central data point. This process effectively reduces high-frequency noise and highlights trends in the data. The formula for the moving average is straightforward, consisting of the sum of data points within the window divided by the window size.

$$y[n] = 1/M \sum_{i=0}^{M-1} c_i * x[n-i] \quad (2)$$

Here, y[n] is the moving average output, x[n] is the input signal, and M is the size of the moving average window.

3) Feature Extraction: A specified segment of pre-processed signals are used to detect systolic peaks by adjusting threshold values using a peak detection algorithm. These detected peaks are then joined together to form a wave-like pattern consisting of a number of peaks that represent the respiration rate. The peaks in the systolic waveform are calculated using the peaks of peaks algorithm, which detects the systolic waveform tips based on height and width thresholding values. This threshold represents the minimum data value that a point must exceed to be considered a peak. These peaks can then be used to identify cycles in the respiratory rates.

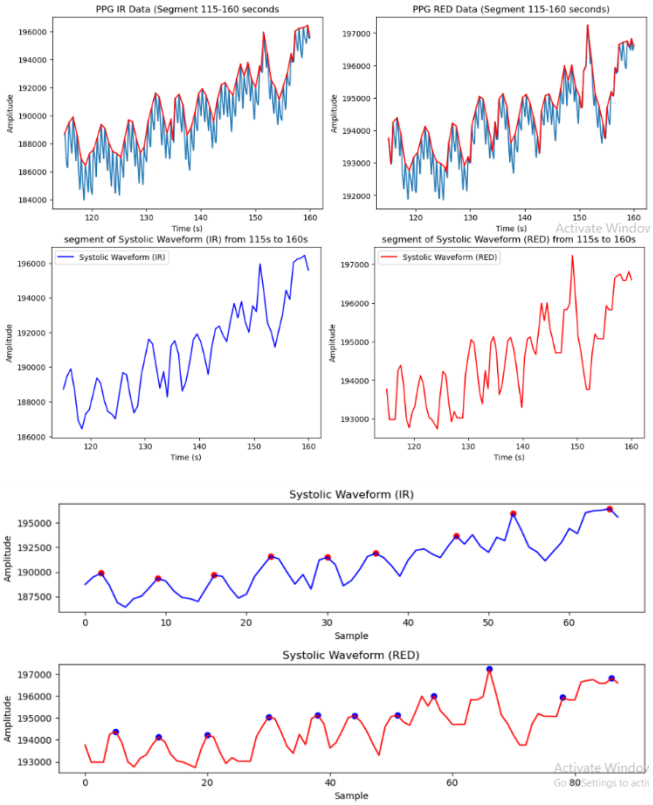


Fig. 6. Feature extraction from PPG signal (IR and RED) of specified segment of time by peak detection on systolic waveform based on thresholding

4) Respiration Rate (RR) Calculations: To calculate the respiration rate per minute for a particular period based on the number of detected systolic waveform peaks from the segmented PPG signal is given by

$$\textbf{Respiration Rate (RR)} = \frac{\text{Number of systolic waveform peaks}}{\text{Total duration of the PPG signal (s)}} \times 60$$

D. *Calibration and validation*

The proposed algorithm for real-time respiration rate monitoring is calibrated and validated using two open-source dataset called BIDMC dataset and the CAPNOBASE dataset.

ISBN 979-8-3503-5094-4

This entails comparing the measurements acquired using the suggested algorithm with the reference standard, which consists of the impedance-based respiratory signals in the BIDMC dataset and the capnography-based respiration rate annotation in the CAPNOBASE dataset. By continuously measuring and analysing the PPG, the system can provide continuous updates on the subject's respiration rate, ensuring accuracy and dependability.

E. To Study The Relationship Between The Respiration Rate And Cognitive Load In Human Subjects

Cognitive load, stemming from mental processes such as problem-solving or information retention, significantly influences the cardiopulmonary cycle. When cognitive load is high, it can cause changes in breathing rate and depth, resulting in alterations in respiratory parameters. This can include increased respiratory rates, shallower breaths, or irregular patterns. By understanding the connection between cognitive load and respiration, the cognitive load of an individual can be assessed.

Fig. 7. Subject 2 performing task 1

1) Protocol Followed To Study The Relationship Between The Cognitive Load And The Respiration Rate: The study involved three male subjects aged between 20 and 23 years. The aim of the study was to determine how a minor psychological stressor affects respiration rate. The subjects were given three tasks, which lasted a total of 9 minutes. The first minute was for the sensor to respond, followed by seven minutes of task performance, and then one minute of rest for the subjects. After completing the tasks, the subjects were instructed to sit quietly and relax.

Two variants of the mental task and one control condition were used. QUIZZ website is used for performing the task 1 and 2. QUIZZ is a website that offers interactive quizzes on various mathematical topics with timers. In the first condition (Speaking), the subjects were asked to read the questions and answers aloud. They were instructed to complete the task as quickly and accurately as possible. In the second condition (Keyboard), the subjects were asked to perform the same task but to record their answers on their mobile phones alone instead of speaking them out loud. Finally, in the third condition (Count), the subjects were asked to count aloud in reverse by ones, starting from 300 and ending with 1, and to say the English alphabet in reverse. They were instructed to complete the task at their own pace. This task was included to assess the effects of speaking and a cognitively demanding task on respiration.

Fig. 8. Subject 3 performing task 2

Fig. 9. Subject 2 performing task 3

III. RESULTS AND DISCUSSIONS

From this experimental the signal and data obtained from the PPG is analysed and the respiration rate is measured. The results are displayed in the below figure

```
Estimated Respiration Rate (IR_): 12.00 bpm from 115s to 160s
Estimated Respiration Rate (RED_): 13.33 bpm from 115s to 160s
Estimated Average RR(IR and RED): 12.67 bpm from 115s to 160s
```

Fig. 10. Respiration rate displayed in the console

The algorithm has been validated using a publicly available datasets, which resulted in a Root Mean Square Error of 0.856 and 1.92. For the majority of the subjects in the dataset, there is only a difference of less than or equal to +or-1 bpm between the actual respiration rate given in the dataset and the respiration rate measured using the proposed algorithm.

TABLE I. COMPARISION OF THE RMSE OF THE PROPOSED ALGORITHM FOR THE CAPNOBASE DATASET WITH THE EXISTING LITERATURE

S. No	TITLE	METHOD	TECHNIQ-UE USED	RMSE (bpm)
1.	Multiparameter respiratory rate estimation from the Photoplethysmogram, 2013 [3] Walter Karlen et al	Photoplethysmogram (PPG) based Respiration detection	Incremental-Merge Segmentation Algorithm	3.0
2	Ensemble Empirical Mode Decomposition With Principal Component Analysis: A Novel Approach for Extracting Respiratory Rate and Heart Rate From Photoplethysmographic Signal,2018 [4] Mohammod Abdul Motin et al	Photoplethysmogram (PPG) based Respiration detection	Ensemble Empirical Mode Decomposition with Principal Component Analysis (EEMD-PCA)	2.77

ISBN 979-8-3503-5094-4

S. No	TITLE	METHOD	TECHNIQ-UE USED	RMSE (bpm)
3	**Proposed Algorithm**	**Photoplethysmogram (PPG) based Respiration detection**	**Systolic waveform peak detection based on thresholding**	**1.92**

TABLE II. COMPARISION OF THE RMSE OF THE PROPOSED ALGORITHM FOR THE BIDMC DATASET WITH THE EXISTING LITERATURE

S. No	TITLE	METHOD	TECHNIQ-UE USED	RMSE (bpm)
1.	Photoplethysmography -Based Respiratory Rate Estimation Algorithm for Health Monitoring Applications.2022 [5] Talha Iqbal et al	Photoplethysmogram (PPG) based Respiration detection	Selective windowing and modified Welch filtering	2.47
2	Whale Optimization Algorithm with a Hybrid Relation Vector Machine: A Highly Robust Respiratory Rate Prediction Model Using Photoplethysmography Signals,2023 [6] Xuhao Dong et al	Photoplethysmogram (PPG) based Respiration detection	Whale Optimization Algorithm (WOA) and a Hybrid Relation Vector Machine (HRVM)	1.79
3	**Proposed Algorithm**	**Photoplethysmogram (PPG) based Respiration detection**	**Systolic waveform peak detection based on thresholding**	**0.856**

.The proposed algorithm for Photoplethysmogram (PPG)-based respiration detection in this study demonstrates superior performance compared to existing literature. In contrast to the methods employed in the referenced papers, the proposed algorithm utilizes a novel approach based on systolic waveform peak detection through thresholding. The key innovation lies in the efficient identification of respiratory cycles, leading to a remarkably low Root Mean Square Error (RMSE) value of 1.92 for CAPNOBASE dataset and 0.856 for BIDMC dataset. This achievement outperforms the RMSE values reported in the existing literature, including the Incremental-Merge Segmentation algorithm (RMSE: 3.0)[3], Ensemble Empirical Mode Decomposition with Principal Component Analysis (EEMD-PCA) (RMSE: 2.77)[4], Selective windowing and modified Welch filtering (RMSE: 2.47)[5] and Whale Optimization Algorithm (WOA) and a Hybrid Relation Vector Machine (HRVM) (RMSE: 1.79)[6].

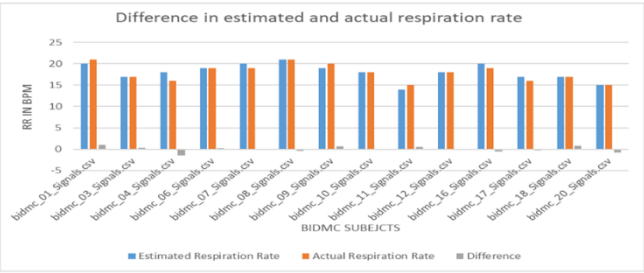


Fig. 11. Estimated respiration rate of the subjects in the BIDMC dataset by the algorithm and actual respiration rate given in the dataset and the difference between them

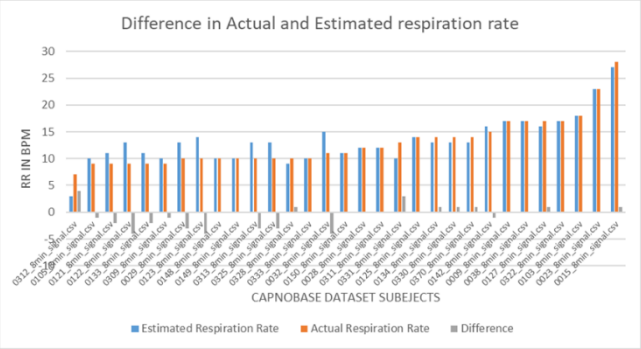


Fig. 12. Estimated respiration rate of the subjects in the CAPNOBASE database by the algorithm and actual respiration rate given in the dataset and the difference

Further the algorithm is modified such that it displays the real time respiration rate of the subject which can used for further assessment of cognitive load of the subject

```
Estimated Respiration Rate (IR_): 10 bpm
Estimated Respiration Rate (RED_): 15 bpm
Estimated Average RR(IR and RED): 13 bpm
Estimated Respiration Rate (IR_): 12 bpm
Estimated Respiration Rate (RED_): 16 bpm
Estimated Average RR(IR and RED): 14 bpm
Estimated Respiration Rate (IR_): 11 bpm
Estimated Respiration Rate (RED_): 16 bpm
Estimated Average RR(IR and RED): 13 bpm
Estimated Respiration Rate (IR_): 11 bpm
Estimated Respiration Rate (RED_): 15 bpm
Estimated Average RR(IR and RED): 13 bpm
Estimated Respiration Rate (IR_): 11 bpm
Estimated Respiration Rate (RED_): 15 bpm
Estimated Average RR(IR and RED): 13 bpm
Estimated Respiration Rate (IR_): 10 bpm
Estimated Respiration Rate (RED_): 15 bpm
Estimated Average RR(IR and RED): 12 bpm
```

Fig. 13. Real time respiration rate displayed in the console

In order to verify if the proposed algorithm can calculate the real-time changes in respiration rate, a subject is asked to participate in a six-minute cycle of breathing. The first two minutes involve normal breathing, followed by two minutes of heavy breathing, then a resting period where the subject is asked to close their eyes and relax, and finally, one minute of sudden heavy breathing.

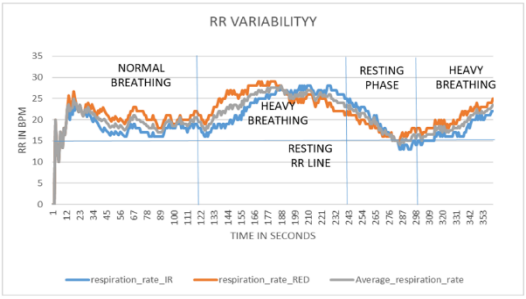


Fig. 14. Respiration rate variability graph

The graph illustrates the efficacy of the proposed algorithm in identifying abrupt changes in respiration rate. Notably, during relaxation, it is observed that the resting respiration rate reaches a minimum, surpassing the normal respiration rate [2]. This phenomenon can be attributed to the relaxation-induced reduction in cortisol, a stress hormone known to influence the cardiopulmonary cycle. Thus, this method is applied for the assessment of cognitive load in an individual in further studies.

ISBN 979-8-3503-5094-4

By following the protocol, three male subjects of the age group between 20-23 years were involved in the study. They were instructed to complete three tasks within a set time frame. The final results include the average respiration rate for each minute of the three tasks is given below.

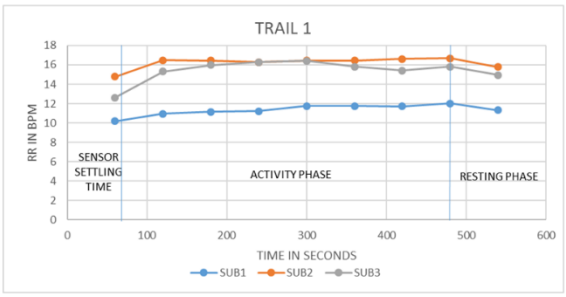


Fig. 15. Mean RR variability graph of all three subjects in trail 1

The findings of this study suggest that respiration rate variability during cognitive load tests can provide valuable information under certain circumstances. However, the degree to which this information can be interpreted may vary depending on the particular condition and the cognitive ability of the subject being studied. The results show that cognitive tasks produced significant increases in the respiration rate of subject 3 during both trial 1 and trial 2. A slight increase in the respiration rate of subject 2 in trial 1 and a decrease in it in trial 2 suggests that vagal interference may have contributed to the change in respiratory rate, not the cognitive load level. Notably, subject 1 respiration remained the same for both trial 1 and trial 2, which suggests that the questionnaires did not elevate subject 1's and subject 2's cognitive load levels but subject 3's respiration is increased in trial 2 in the absence of vagal interference. This may be due to the individual's cognitive ability to respond to the questionnaires.

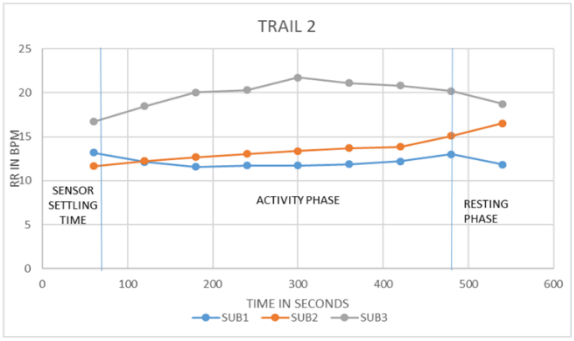


Fig. 16. Mean RR rate variability graph of all three subjects in trail 2

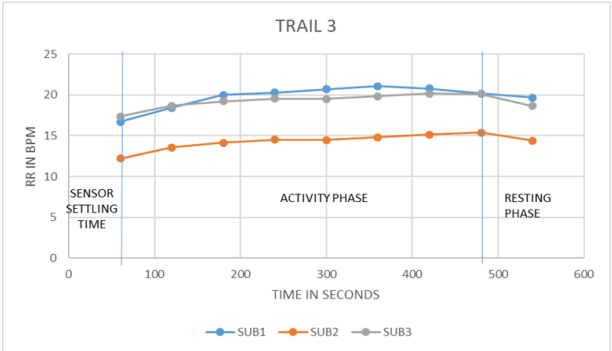


Fig. 17.Mean RR rate variability graph of all three subjects in trail 3

In trial 3, all the subjects showed a clear increase in respiratory rate, indicating that vagal tone solely contributed to the elevated respiration. Thus when the task relies on verbal reports of answers, it may not be possible to assess respiratory rate variability. Finally, the data from this study support the assertion that respiratory rate variation may represent cognitive load during cognitive load tasks in the absence of vocal interference[7]. However, further studies are necessary to confirm these findings in a larger population.

IV. Conclusion

In summary, a novel algorithm was developed using Systolic waveform peak detection based on thresholding for real-time respiratory rate detection. It was validated using an open-source datasets, with a root mean square error (RMSE) of 0.856 and 1.92. Real-time data was collected using a PPG sensor module and analysed to determine the relationship between respiration rate variability and cognitive load levels. The study's outcomes suggest that changes in respiratory rate could indicate cognitive load when people are subjected to psychologically demanding tasks without voice interference. The exploration of diverse participant profiles will fortify the reliability of these findings, ultimately paving the way for broader applications and a deeper comprehension of the intricate interplay between physiological signals and cognitive processes. However, it is essential to acknowledge certain limitations inherent in the proposed algorithm. High noise levels in PPG signals can pose challenges and reduce accuracy, particularly in real-world scenarios. Further refinements and adaptations are warranted to enhance the algorithm's resilience to noisy environments. Despite these limitations, the algorithm presented in this study represents a significant advancement in non-contact continuous respiratory monitoring.

References

[1] Mariel Grassmann, et.al, “The role of respiratory measures to assess mental load in pilot selection”, Ergonomics, 59:6, 745-753, 2016

[2] Nicolò A, Massaroni C, Schena E, Sacchetti M., “The Importance of Respiratory Rate Monitoring: From Healthcare to Sport and Exercise. Sensors”.;20(21):6396. Nov 2020

[3] Karlen W, Raman S, Ansermino JM, Dumont GA, “Multiparameter respiratory rate estimation from the photoplethysmogram”, IEEE Trans Biomed Eng. 60(7):1946-53. July 2013

[4] M.A. Motin, C. K. Karmakar and M. Palaniswami, "Ensemble Empirical Mode Decomposition With Principal Component Analysis: A Novel Approach for Extracting Respiratory Rate and Heart Rate From Photoplethysmographic Signal," in IEEE Journal of Biomedical and Health Informatics, vol. 22, no. 3, pp. 766-774, May 2018

[5] Iqbal, T., Elahi, A., Ganly, S. et al. “Photoplethysmography-Based Respiratory Rate Estimation Algorithm for Health Monitoring Applications.” J. Med. Biol. Eng. 42, 242–252. 2022

[6] Dong, X.; Wang, Z.; Cao, L.; Chen, Z.; Liang, Y, “Whale Optimization Algorithm with a Hybrid Relation Vector Machine: A Highly Robust Respiratory Rate Prediction Model Using Photoplethysmography Signals”, Diagnostics, 13, 913. 2023

[7] R.P. Sloan, J.B. Korten, M.M. Myers,”Components of heart rate reactivity during mental arithmetic with and without speaking,Physiology & Behavior”,Volume 50,Issue 5,Pages 1039-1045,ISSN 0031-9384, 1991

ISBN 979-8-3503-5094-4

IMPLEMENTATION OF A NOVEL PERSONALIZED NON-INVASIVE GLUCOMETER USING MACHINE LEARNING

Rajasingam N
Biomedical Engineering
KPR Institute of Engineering and Technology
Coimbatore, Tamil Nadu, India
rajasingamnk@gmail.com

Padmanaban S
Biomedical Engineering
Sri Sivasubramiya College of Engineering
Chennai, Tamil Nadu, India
padhu.jaya01@gmail.com

Bhuvaneswari B
Biomedical Engineering
KPR Institute of Engineering and Technology
Coimbatore, Tamil Nadu, India
bhuvaneswaribhvui2002@gmail.com

Madhan Kumar U
Biomedical Engineering
KPR Institute of Engineering and Technology
Coimbatore, Tamil Nadu, India
madhangh17@gmail.com

Imrankhan R
Biomedical Engineering
KPR Institute of Engineering and Technology
Coimbatore, Tamil Nadu, India
rimrankhan18@gmail.com

Kesavarani G
Biomedical Engineering
KPR Institute of Engineering and Technology
Coimbatore, Tamil Nadu, India
kesavigiri3092002@gmail.com

***Abstract*— In recent years, diabetes has become more prevalent, and patients suffer from pain and infection due to invasive procedures. Non-invasive glucose monitoring devices improve patient comfort by reducing pain and infection compared to invasive methods. This paper presents a novel non- invasive glucose monitoring system by employing NIR technology and a machine learning algorithm to provide an accurate and continuous monitoring system that simultaneously displays a result on a mobile app in a user-friendly way. This can lead to an optimized system than conventional methods that require blood samples, which is significant for diabetic therapy. By incorporating the proposed technology, a non-invasive as well as user-friendly system is implemented that improves the quality of life of diabetic patients with the potential to completely transform glucose monitoring, reduce the burden on diabetics, and provide healthcare professionals with a valuable tool for improved patient care.**

Keywords*— *Diabetes, Glucose monitoring, Non-Invasive, Machine Learning, Blood Glucose, Glucometer

I. INTRODUCTION

Diabetes is considered to be one of the most prevalent chronic metabolic conditions in the world that affects human. It can be due to a genetic factor, immune disorders and other problems which reduce the production of insulin in human body. Diabetes can be differentiated into two types: Type-1, Type-2 diabetes. Type-1 diabetes is considered as minimal or no secretion of insulin, this is diagnosed in childhood [13] and

Type-2 diabetes is a condition where body cells do not respond to the insulin and decline in insulin production by the pancreas, this is diagnosed in adult [14]. Additionally, gestational diabetes is another type of diabetes that arises during pregnancy and continues after the delivery, Women with this diabetes have high risk to develop Type-2 diabetes. Ifthis type of diabetes is not properly managed it can lead to serious health complications. It is a life-threatening disease on a global scale. In 2021, there were approximately 537 million adults in the World affected by Diabetes as reported by the International Diabetes Federation (IDF).

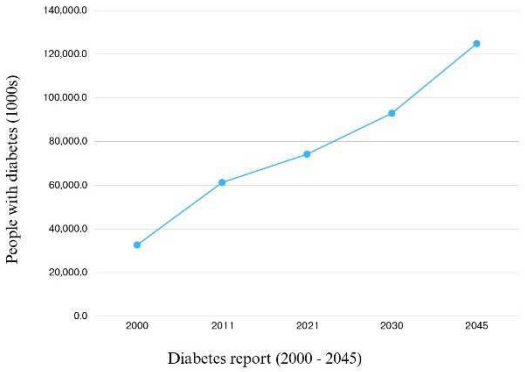


Figure 1: India diabetes report [8]

As a result, 783 million number of people will be increased by 2045. According to World Health Organization (WHO),

ISBN 979-8-3503-5094-4

India is estimated to have the second-largest number of diabetes cases that depicted in Figure 1. Over 77 millionpeople above 18 years are suffering from diabetes. Diabetescan be characterized by a variety of symptoms such asextremely thirsty, frequent urination, fatigue, and blurredvision. It is important to note that smoking and excessive alcohol consumption can further increase the complications in individuals with diabetics. A proper management of diabetes involves a combination of lifestyle changes, such as a healthy diet and regular exercise, along with prescription medications and insulin therapy. Diabetes can affect various organs, to reduce the complication and risk, frequent monitoring is highly important. This requires a collaborative approach between individuals and healthcare professionals. Innovative changes in healthcare technology, especially with the introduction of non-invasive glucometers enhanced by machine learning (ML) algorithms in mobile applications can improve the life style by diagnosing the diabetes in the early stages and reduce the complications. Blood glucose measurement methods can be categorized [6] into invasive,non-invasive [1] and minimally invasive approaches [3]. Toobtain accurate results, invasive techniques usually involveextracting a blood sample from the fingertip using a lancet and then analysing it with a glucometer. Moreover, blood is also collected from upper arm, forearm, fingertip, and thigh, however the accuracy of lab blood glucose test which is taken in venous blood is more accurate than other invasive methods. Overall, the paper focuses on the continuous monitoring of blood glucose level and analysed using machine learning for accurate prediction.

II. METHODS

A. *Invasive Method*

Invasive techniques can be used to measure glucose level in blood through pricking the finger tips. In laboratories invasive method in preformed in venous blood, it gives more accurate than portable invasive method [11]. While laboratory measurements for glucose can take several hours with high precision, portable devices can provide results within a few minutes. This method causes pain and discomfort.

B. *Minimally Invasive Method*

This approach is used to measure the blood glucose level without standard finger pricks. This reduces the discomfort and improve convenience. It tracks the glucose level 24/7 for every minute through tiny sensor inserted on skin either on abdomen or arm [12].

C. *Non-Invasive Method*

Non-invasive techniques tend to measure the blood glucose level without pricking the skin or extracting blood samples, it gives more comfort and convenient than traditionalmethods for diabetic patients. Several techniques have been researched for non-invasive glucose monitoring [2],[4],[7] although achieving accuracy [10] for non-invasive values remains challenging [9]. In certain diabetic cases, it requires continuous monitoring and some group of people may reject the invasive treatment. Non-invasive portable device can change the life of diabetic patients by monitoring their glucoselevel without a drop of blood. Electromagnetic spectrum that includes visible region, Infrared (IR), Ultra-Violet (UV). In this NIR has ability to penetrate dermis of the skin and interactwith blood component. NIR lies between 700-1300nm in Electromagnetic spectrum. When NIR is passed through blood it is absorbed by glucose molecules in blood, through this concentration of glucose is estimated [5]. Advantage of non-invasive method is no need of finger pricking or blood sample, and it helps for continuous, quick and real time monitoring.

III. METHODOLOGY

The development of non-invasive glucometer system combines hardware, machine learning algorithm and mobile app for real time monitoring. The application of Non-invasive blood glucose monitor is designed with NIR technology and machine learning approaches to monitor accurate bloodglucose level. Initially, the NIR emits near-infrared light and penetrates through skin, where it interacts with glucose molecules in blood. The interaction of NIR and glucose can help to analyse the concentration of glucose in blood. Continuous improvement and refinement of algorithms areused to reduce the error and improve the accuracy rate. To predict the accurate glucose level, the value is calculated by transmitter and receiver, then it is analysed and tested in machine learning model. The result is displayed in mobile app for monitoring real time glucose level and helps in self- monitoring, this is considered to be a user-friendly and cost- effective model for many diabetic people.

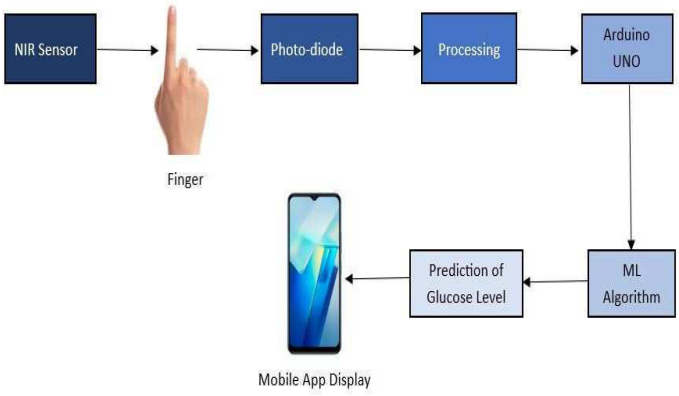


Figure 2: Block Diagram for proposed Non-Invasive Glucometer

A. *Hardware Implementation*

The system's primary hardware elements are microcontroller, an LED (transmitter), a photodiode, an operational amplifier. The NIR detection circuit for blood glucose level monitor is made up of a transmitter circuit and a receiver circuit. The Arduino microcontroller powers the transmitter and receiver at 5 V.

ISBN 979-8-3503-5094-4

1) Receiver Circuit:

An operational amplifier, a photodiode, and a noise filter make up the receiver circuit figure 3. For receiver circuit voltage is supplied by the Arduino UNO. Photodiode is used to detect the light coming from transmitter circuit and convert it into electrical current. Filter is used to reduce noise or unwanted frequencies. To enhance the output, operational amplifier is used.

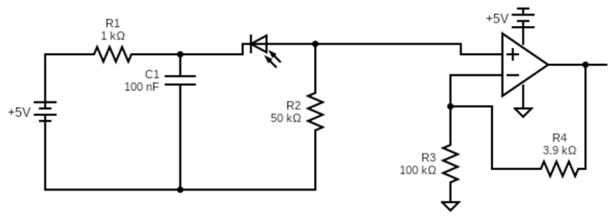


Figure 3: Receiver Circuit

2) Transmitter Circuit:

Transmitter circuit in figure 4 consists of LED, resister, power supply is given to the circuit from microcontroller.Near-infrared is passed through the skin, and transmitted light is received by the receiver. While the intensity and calibration of transmitted light result is calculated. NIR (800nm-1800nm) has ability to interact and change the concentration of glucose level in blood.

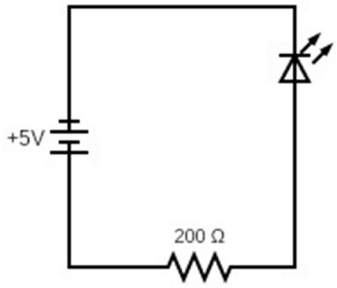


Figure 4: Transmitter Circuit

B. *Software Implementation*

To analyse the program's key features, which revolve around the continuous, real-time monitoring and visualization of users' blood sugar levels. The software is designed to function together with a non-invasive device, making it easier to collect and analyse crucial glucose data. The app's user-friendly interface makes it simple for users to check theircurrent blood sugar levels. Furthermore, our program assists users in understanding their health trends by providing graphical representations of glucose fluctuations over time.Bluetooth or similar compatible wireless technology is utilizedto establish a connection between the app and the non-invasiveglucometer device, ensuring data accuracy and reliability. The non-invasive glucose meter and the app connect via Bluetooth or another suitable wireless technology, ensuring data accuracy and dependability. The non-invasive glucose meter and the app connect via Bluetooth or another suitable wireless technology, ensuring data accuracy and dependability. The program begins with a user registration procedure, in which users create accounts and give personal information. This registration process plays a critical role in ensuring data security and personalizing user experiences. Software is carefully implemented to ensure a reliable and effective platform for glucose monitoring. To protect sensitive personal data, data privacy and security are our top priorities. By streamlining the procedure and making it easier for users to check their blood sugar levels, wireless technology integration eventually improves healthcare management and overall well- being.

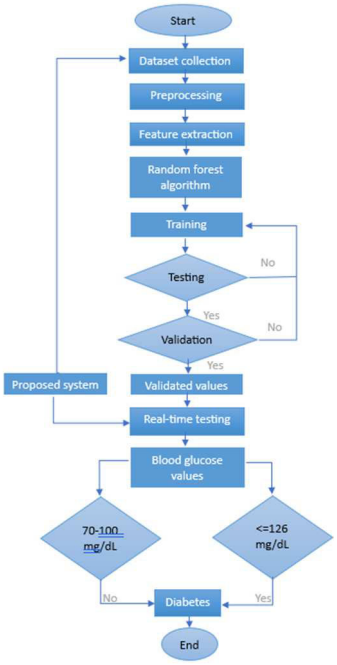


Figure 5: Flowchart of non-invasive model

C. *Selection of Machine learning Algorithm and Dataset*

A non-invasive glucometer employing machine learning encompasses various process, such as data collection and preprocessing, feature extraction, algorithm selection, and performance evaluation. Confusion matrix and metrics are performed to improve the exactness of the value. Non-invasive blood glucose level is predicted by Random-Forest algorithm, that combines numerous decision tree for improving the closeness of the actual value.

Data Collection: Initially, a dataset is compiled, consisting of non-invasive glucose measurements and their corresponding blood glucose levels, gathered through sensors or devices, obviating the necessity for invasive methods.

Data Preprocessing: Clean the data to handle missing values and outliers. Normalize or standardize the features to ensure consistent scaling. Dataset is separated into train and test set. Important process to check the quality of the data is Data cleaning, transformation, Integration and dimensionality reduction.

ISBN 979-8-3503-5094-4

Feature Extraction: Diabetic patients attributes are extracted from the collected data, including variables like age, pregnancy status, blood glucose value, body mass index (BMI), and insulin level. These features are vital as they can significantly affect blood glucose levels. Even skin thickness can affect the value of blood glucose level, when NIR is passed through the skin.

Algorithm Selection: Machine learning model Random Forest applied to the dataset. In Machine learning, Random Forest is a method used for classification and regression. It captures intricate relationships. This algorithm improves the performance and accuracy of the value through training and testing.

Model Training: Once the algorithms are selected, the next step is to train each of them on the training data. Dataset is split as training dataset and testing dataset. To achieve the accurate value for prediction, fine-tune the hyperparameters of algorithm. Training involves to find the optimal parameters that enable the model to make accurate predictions.

Model Evaluation: Testing data is used to evaluate the perform of trained models. Regression metrics like Mean Absolute Error and Mean Squared Error is calculated to measure how accurately the models predict blood glucose levels. This step helps to compare the models and to decide which one is accurate to the predicted value.

Confusion Matrix Analysis: Confusion Matrix is used to evaluate the performance of the classification model in N x N matrix, where N represents the number of targets. This matrix offers a comprehensive breakdown of predictions, encompassing True and Positives (TP), False and Positives (FP), True and Negatives (TN), False and Negatives (FN). Metrics such as support, $F_{1\text{-score}}$, recall, and precision are computed to attain the model's effectiveness in predicting blood glucose levels accurately. Representing TP as A, FP as B, FN as C, TN as D.

Table 1: Train-Test Ratio (90-10)

Actual / Prediction	Diabetic	Non-diabetic
Diabetic (mg/dL)	368	40
Non-diabetic(mg/dL)	62	144

Table 1, shows train-test ratio as 90-10 analysis of machine learning algorithm to classify the data as diabetics or not diabetics. Data-split is represented in percentage. Accuracy percentage of train dataset is 73 percentage, and for test dataset is 83 percentage. Overall accuracy of this model is 75%.

Table 2: Train-Test Ratio (80-20)

Actual / Prediction	Diabetic	Non-diabetic
Diabetic(mg/dL)	315	33
Non-diabetic(mg/dL)	52	137

Table 2, shows train-test ratio as 80-20 analysis of machine learning algorithm to classify the data as diabetics or not diabetics.

$$\text{Precision} = A/(A+B) \quad (1)$$

$$\text{Recall} = A/(A+C) \quad (2)$$

$$F_{1\text{-score}} = (2*(\text{Precision}*\text{Recall})) / (\text{Precision} + \text{Recall}) \quad (3)$$

Table 3: Classification Report

Metrices	**Precision** (percentage)		**Recall** (percentage)		**$F_{1\text{-score}}$** (percentage)	
	(90-10)	**(80-20)**	**(90-10)**	**(80-20)**	**(90-10)**	**(80-20)**
Diabetic (mg/dL)	86	88	90	94	88	91
Non-diabetic (mg/dL)	78	86	70	74	74	80

Table 3, shows the performance of important metrics like Recall, $F_{1\text{-score}}$, Precision for train-test set (90-10) ratio and train-test set (80-20) ratio.

Model Deployment: The selected model is deployed into the non-invasive glucometer for real-time and continuous glucose level predictions. The device can take non-invasive measurements and provide glucose level estimates based onthe trained model. This deployment is a critical step in makingthe technology practical and useful.

Continuous Monitoring and Refinement: Creating a reliable non-invasive glucometer is an ongoing process. We continuously collect new data to keep the model up to date and improve its accuracy over time. This step involves retraining the model with new data and potentially refining the feature selection or engineering process to adapt to changing conditions.

IV. RESULT AND DISCUSSION

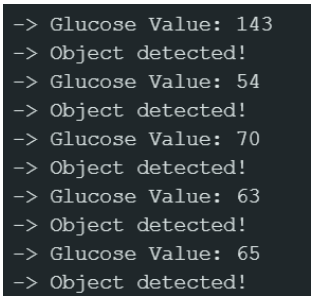


Figure 6: Overview of non-invasive model gives glucose value in mg/dL

The above figure 6 is the result from Arduino program, value collected from Diabetic and Non-diabetic people through this non-invasive model, which is relatively near to their invasive test. and it is tabulated in table 4.

ISBN 979-8-3503-5094-4

Table 4: Predicted Result

Subjects	Non-invasiveGlucose value (mg/dL)
1	143
2	54
3	70
4	63
5	65
6	121
7	96
8	113
9	75
10	68

The findings and examination underscore the preciseness of measurements of blood glucose levels with this level of precision, it is feasible to extract knowledge from the patient database and offer personalized treatments based on the patient's blood sugar levels.

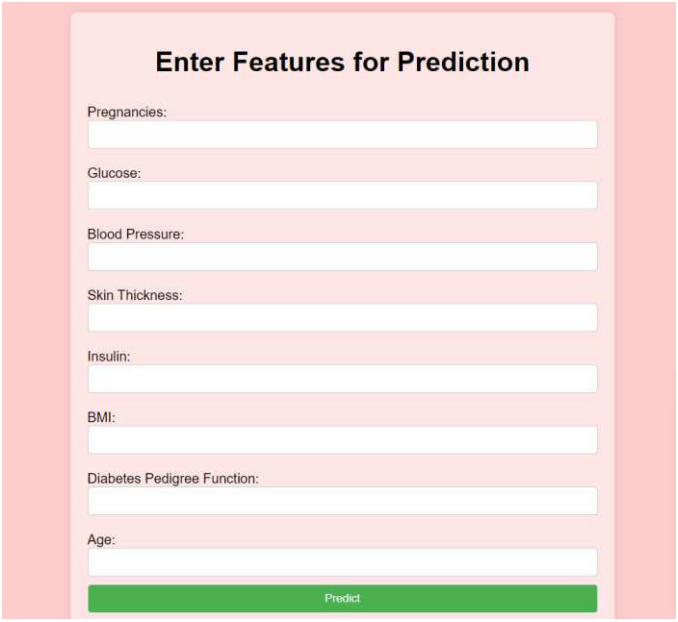


Figure 7. Prediction page of mobile application

The above figure 7 shows the patient initially signs in using their login ID, along with their username and password. The user model can be adapted to monitor and recognize the blood glucose levels and other features of the user. By leveraging the accuracy of current technology, the patient's diabetes statuscan be accurately diagnosed.

V. CONCLUSION

The predicted output shows the nearby glucose value. Accuracy of this value can be achieved through machine learning algorithm. Our further our is to obtain accurate glucose value. Predictive analysis, wearables integration, enhanced accuracy, and wider ecosystem participation in health are among the promises of the technology's future. Healthcare is become increasingly digital, but data security is still important. A new age of preventative and individualized healthcare is being assist in by glucose monitoring in thisfuture environment. In this paper, novel non-invasive glucose monitoring system is designed to realize the optimised continuous monitoring of diabetes by incorporating Near-Infrared (NIR) and machine learning model. Results shown that the proposed approach not only improves the glucoseregulation but also leads the enhanced personal care on public health.

VI. REFERENCES

[1] So, Chi-Fuk, Kup-Sze Choi, Thomas KS Wong, and Joanne WY Chung. "Recent advances in noninvasive glucose monitoring." Medical Devices: Evidence and Research (2012): 45-52.

[2] Dziergowska, Katarzyna, Magdalena Beata Łabowska, Marlena Gąsior-Głogowska, Barbara Kmiecik, and Jerzy Detyna. "Modern noninvasive methods for monitoring glucose levels in patients: A review." Bio-Algorithms and Med-Systems 15, no. 4 (2019): 20190052.

[3] Ginsberg, Barry H. "An overview of minimally invasive technologies." Clinical Chemistry 38, no. 9 (1992): 1596-1600.

[4] Bezuglyi, M., N. Bezuglaya, O. Kuprii, and I. Yakovenko. "The non-invasive optical glucometer prototype with ellipsoidal reflectors." In 2018 IEEE 59th International Scientific Conference on Power and Electrical Engineering of Riga Technical University (RTUCON), pp. 1- 4. IEEE, 2018.

[5] Sunny, Saina, and S. Swapna Kumar. "Optical based non invasive glucometer with IoT." In 2018 International Conference on Power, Signals, Control and Computation (EPSCICON), pp. 1-3. IEEE, 2018.

[6] Shang, Trisha, Jennifer Y. Zhang, Andreas Thomas, Mark A. Arnold, Beatrice N. Vetter, Lutz Heinemann, and David C. Klonoff. "Products for monitoring glucose levels in the human body with noninvasive optical, noninvasive fluid sampling, or minimally invasive technologies." Journal of diabetes science and technology 16, no. 1 (2022): 168-214.

[7] Deshmukh, Vidya, and Suvarna Chorage. "Investigational Outcomes of Normal and Diabetic Human Volunteers using Microwave based Non-invasive Blood Glucometer." Int. J. Emerg. Technol. Adv. Eng 12 (2022): 177-184.

[8] https://diabetesatlas.org/data/en/country/93/in.html

[9] Van Enter, Benjamin Jasha, and Elizabeth Von Hauff. "Challenges and perspectives in continuous glucose monitoring." Chemical Communications 54, no. 40 (2018): 5032-5045.

[10] Tonyushkina, Ksenia, and James H. Nichols. "Glucose meters: a review of technical challenges to obtaining accurate results." Journal of diabetes science and technology 3, no. 4 (2009): 971-980.

[11] Clynes, Mary, Colleen O'Neill, and Sara Raftery. "Blood sugar measurement (using a glucometer)." Foundation Skills for Caring: Using Student-Centred Learning (2018): 180.

[12] Huang, Xinshuo, Jingbo Yang, Shuang Huang, Hui-jiuan Chen, and Xi Xie. "Minimally invasive technology for continuous glucose monitoring." Bio-Design and Manufacturing 5, no. 1 (2022): 9-13.

[13] Katsarou, Anastasia, Soffia Gudbjörnsdottir, Araz Rawshani, Dana Dabelea, Ezio Bonifacio, Barbara J. Anderson, Laura M. Jacobsen, Desmond A. Schatz, and Åke Lernmark. "Type 1 diabetes mellitus." Nature reviews Disease primers 3, no. 1 (2017): 1-17.

[14] DeFronzo, Ralph A., Ele Ferrannini, Leif Groop, Robert R. Henry, William H. Herman, Jens Juul Holst, Frank B. Hu et al. "Type 2 diabetes mellitus." Nature reviews Disease primers 1, no. 1 (2015): 1-22.

ISBN 979-8-3503-5094-4

ISBN 979-8-3503-5094-4

Tenth International Conference on Biosignals, Images and Instrumentation, SSNCE, Chennai, March 20 – 22, 2024

TRACK 10

SIMULATIONAL STUDY FOR DESIGNING LUNG ON-CHIP

Perumal D
Department of *Biomedical Engineering*
Veltech Rangarajan Dr.Sagunthala R&D Institute of Science and Technology
Chennai, India
vtu11791@veltech.edu.in

Sweta Jha
Department of *Biomedical Engineering*
Veltech Rangarajan Dr.Sagunthala R&D Institute of Science and Technology
Chennai, India
vtu13630@veltech.edu.in

Thiyam Deepa Beeta
Department of *Biomedical Engineering*
Veltech Rangarajan Dr.Sagunthala R&D Institute of Science and Technology
Chennai, India
thiyamdeepa@gmail.com

Smriti Ghimire
Department of Biomedical Engineering
Veltech Rangarajan Dr.Sagunthala R&D Institute of Science and Technology
Chennai, India
vtu21477@ veltech.edu.in

Shelishiyah R
Department of *Biomedical Engineering*
Veltech Rangarajan Dr.Sagunthala R&D Institute of Science and Technology
Chennai, India
shelishiya@veltech.edu.in

***Abstract*—**The organ-on-chip (OOC) platform enables faster, better, and less expensive drug development, disease modeling, customized treatment, and insights into human health by providing flexibility and robustness in drug testing. Animal models, on the other hand, have failed to provide effective and efficient drug testing results since different species have distinct characteristics. Here, this study reports a lung-on-chip (LOC) aiming to mimic the basic physiological response invitro human lungs while breathing. This microfluidic device performed the mechanical movement of 3D cyclic stretching inspired by breathing movements. This device consisted of 2 parts, the fluidic, and pneumatic parts. This proposed microfluidic chip had two methods of operating, one was breathing mode and another one was medium exchange mode. The simulation of the lung-on-chip was done in Ansys workbench by using static structural and Computational Fluid Dynamics (CFD). The proposed microfluidic chip has great potential for drug testing and new drug development and has a wide range of applications.

***Keywords*—** LOC, Microfluidic chip, CFD, Static structure, Drug testing, Ansys workbench.

I. INTRODUCTION

Organ-on-chip(OOC), also termed the micro-physiological system, is defined as a microfluidic device lined with living human cells that recapitulate the complex structures of organs and their physiological responses in vitro. OOC platform provides flexibility and robustness in drug testing which helps enable faster, better, and cheaper drug development, disease modeling, personalized medicine, and insights into human health [1]. The main elements of OOC are the microfluidics system, living cell tissue, stimulation or drug delivery system, and sensing system. Lung-on-chip was the first OOC invented by Donald Ingber in 2007 [2]. Simultaneously, single organs have been integrated into a single chip consequently multi-organ chips have been developed. Multi-organ chips combine numerous organ units into a single chip for various activities, such as drug absorption in the stomach, drug metabolism in the liver, and drug excretion in the kidney [3,4]. The subject of LoC technology is growing quickly and has the potential to completely change how we study and treat lung illnesses. This technology has the potential to create novel medicines and enhance respiratory health by tackling present issues and venturing into uncharted territory.

II. LUNG ON CHIP

Lungs are respiratory organs found in a thoracic cavity [5] and the primary objective of the lung is to supply oxygen to the blood. The fundamental purpose of the respiratory system is to deliver O_2 eliminate CO_2 via gas exchange and divide it into airways and lung parenchyma. The airways of the bronchus divide into bronchioles and then further into alveoli. The parenchyma comprises alveoli, alveolar ducts, and bronchioles and assists in gas exchange. The alveolus membrane is superimposed with primary alveolar epithelial cells which comprise alveolar type I (ATI) and alveolar type II (ATII) cells [6][7]. ATI cells are large, flat, and squamous cells composed of 90% of the alveolar surface and are responsible for forming a thin alveolar barrier for gas exchange and immune response. Similarly, ATII cells are cuboidal and help synthesize storage and secrete surfactant into the alveolar hypo phase to lower alveolar surface tension, improving lung compliance and preventing alveolar collapse.

Earlier, animals such as rats were used as a basis for drug testing. However, animal models failed to give effective and efficient drug testing results because different species have distinctive features [4], [6], [8]. Some of the needs for the project are to study the physiological responses of the human lung in vitro [9], to minimize the ethical issues and save the patient's life caused by the wrong dose or side effects of drugs, to accelerate the development of new drugs and personalizes medicine, to examine the response of new

ISBN 979-8-3503-5094-4

medicine to the human lung in vitro, and to determine the precise dose of drugs for the patient.

Herein, in vitro, a layer of 40 µm thick PDMS membrane is attached to the pneumatic part that acts as a micro diaphragm [7], [10]. When negative pressure is applied via an electro-pneumatic setup to the breathing access port, the PDMS membrane cyclically deflects into the respective breathing cavity. The micro diaphragm's movement is transposed to the porous alveolar membrane and the alveolar membrane will cyclically deflect in response to the movements of the micro diaphragm.

The current study aims to design a human multifunctional lung-on-chip in vitro by employing 3D printing technology [11], [12] for drug screening and optimization of different lung diseases.

III. MATERIALS

A. Materials

PDMS is a silicone that is also known as Polydimethylsiloxane. PDMS is inert to proteins and cells while also being gas-permeable. Hence, it is ideal for cellular research [13]. To fabricate microfluidic chips, PDMS base and curing agents (Sylgard 184 and Dow Corning) were mixed in a 10:1 w/w ratio and then degassed in a vacuum desiccator to remove air bubbles and moisture. Then it was cast on molds and cured using an oven at 60°C for at least 24 hours. PDMS has the chemical formula $(H_3C)_3[Si(CH_3)_2O]_n Si(CH_3)_3$, where n denotes the number of repeating monomers $[SiO(CH_3)_2]$.

IV. METHODOLOGY

A. Ansys workbench

Ansys Workbench is an integrated simulation platform that aids in the integration of data across engineering simulations to produce more accurate and efficient models. The following flowchart in figure 1 illustrates the methodology followed:

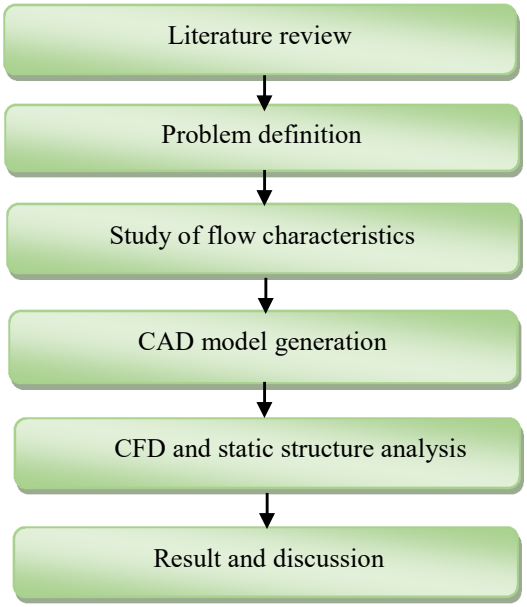


Fig. 1 *Flowchart of a sequence of simulation*

B. CFD Analysis

CFD is a very powerful method that helps to get approximate accurate solutions to fluidic dynamic problems [13] and has a wide range of applications. Finite difference, finite element, and spectral are three main methods to solve the problem in CFD.

C. Static Structure

The static structure is used to analyze the effects of loads on a model while ignoring significant inertia and damping effects in the Ansys platform [14]. It determines the directional stresses, strains, forces, and directional and total deformation caused by the applied loads. In a static structural analysis, externally applied forces and pressures, steady-state inertial forces-imposed displacements, and temperature are some of the types of loads that can be applied to the design.

Here, we have determined the directional deformation, stress, strains, and total deformation by varying pressure values. The formula to calculate equivalent stress and strain is represented by equations (1) and (2) respectively:

$$\text{Equivalent stress} = \frac{[(\sigma_1 - \sigma_2)^2 + (\sigma_2 - \sigma_3)^2 + (\sigma_3 - \sigma_1)^2]^{1/2}}{2} \quad \ldots\ldots (1)$$

Where, $\sigma_1, \sigma_2, \sigma_3$ are principal stresses associated in principal directions along the x, y, and z axes respectively.

The equivalent stress, based on the von Mises criterion, is a widely used parameter to assess the overall stress state in a material subjected to complex loading conditions.

$$\text{Equivalent strain}(\varepsilon_{eq}) = 1/3[2((\varepsilon_x - \varepsilon_y)^2 + (\varepsilon_x - \varepsilon_z)^2 + (\varepsilon_y - \varepsilon_z)^2)]^{1/2} \quad \ldots\ldots (2)$$

Where, $\varepsilon_x, \varepsilon_y, \varepsilon_z$ are principal strains in x,y, and z directions respectively.

In the context of material deformation, the equivalent strain (ε_{eq}) serves as a comprehensive indicator, especially when dealing with multiple strain components.

Geometrical Modelling

The mathematical depiction of an object's geometry is known as geometric modeling. It employs the usage of curves in the creation of models. It can be viewed in both 2D and 3D perspectives. 2D and 3D. Figures 2,3,4 and 5 show the top view, side view, top plate, middle plate, micro diaphragm layer, and bottom plate of the geometry respectively. This microfluidic chip comprises fluidic and pneumatic parts. The fluidic part has a top and middle plate containing three inlets, an outlet, and cell culture wells. The diameter of the cell culture well where the porous membrane is placed is 3 mm.

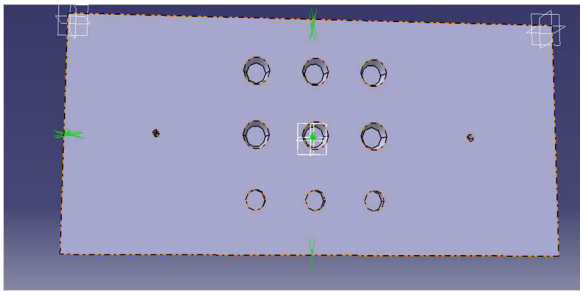


Fig.2 *Top plate*

ISBN 979-8-3503-5094-4

The top plate has a thickness of 2 mm, a length of 75 mm, and a width of 33 mm.

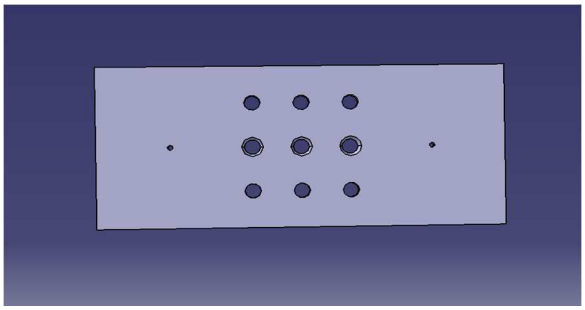


Fig. 3 Middle-plate

The middle plate has a thickness of 1.5 mm, a length of 75 mm, and a width of 33 mm.

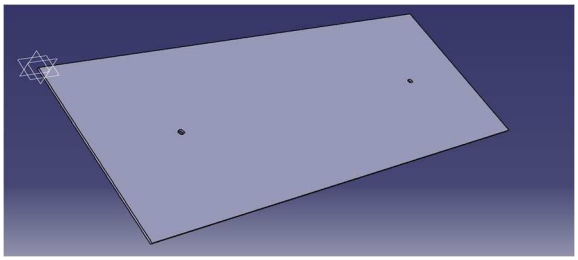


Fig. 4 Microdiaphragm layer

The microdiaphragm layer has a thickness of 0.04 mm, and two accessing holes of diameter of 1 mm each.

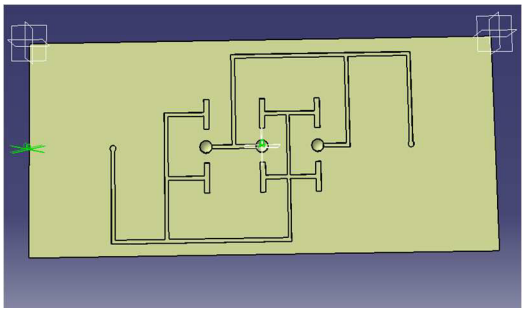


Fig. 5 *Bottom Plate*

The bottom plate has a thickness of 2 mm, a length of 75 mm, and a width of 33 mm.

Grid Generation for CFD Analysis and Static Structure

Meshing for CFD

Mesh or grid generation was performed once the CAD model had been imported and triangular mesh was used. Following face meshing and body scaling, mesh component coordinates of boundary named inlet, outlet, and walls were chosen. The number of elements and nodes formed for human lung-on-chip are 11630 elements and 6221 nodes respectively.

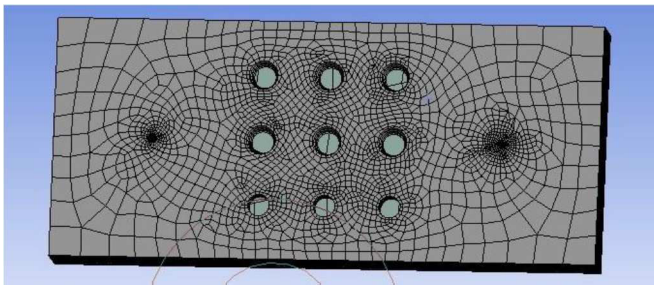


Fig. 6 *Meshing for CFD*

Meshing for static structure

Mesh generation or grid generation was done once the CAD model had been imported. Face meshing and body sizing were chosen along the specified boundary at the mesh part coordinate named walls, inlets, and outlets. The number of elements and nodes formed for human lung-on-chip is 5367 elements and 16725 nodes respectively.

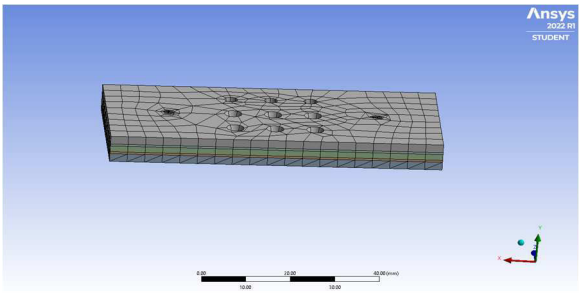


Fig. 7 *Meshing for static structure*

D. Preprocessing for CFD Analysis

The pre-processing procedure of the human lung-on-chip is depicted in Figure 8.

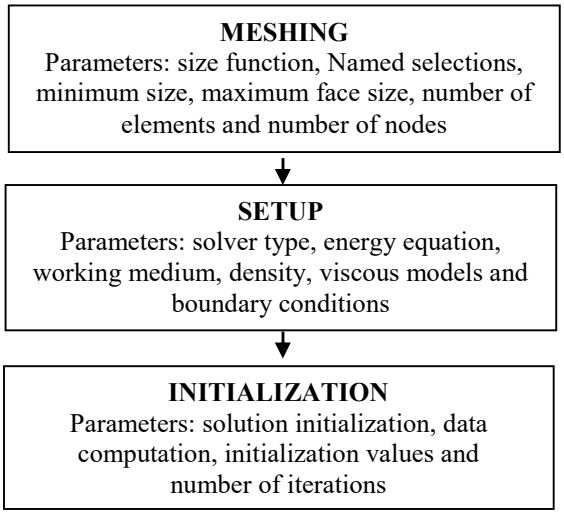


Fig. 8 *Flowchart of preprocessing for CFD analysis*

Static Structural Tabulation

Table 1. shows the static structural analysis setup.

ISBN 979-8-3503-5094-4

TABLE 1. STATIC STRUCTURAL ANALYSIS SETUP

S. No.	MESHING	
1	Face sizing	33
2	Number of nodes	16754
3	Number of elements	5397
4	Active bodies	5
	SETUP	
1	Pressure	50 Pa, 100 Pa, 100 KPa
2	Fixed support	At the bottom part

V. RESULTS

Validation of parameters

The human breathing lung-on-chip model was designed in CATIA v5 software. The deformation of the micro diaphragm and fluid pressure in the microcavity was analyzed using Ansys workbench. The static structure was used to analyze the 3D cyclic deflection of the micro diaphragm during breathing mode. Similarly, modeling fluidic flow during medium exchange operation mode as well as analyzing pressure at different points of the microcavities was done using computational fluid dynamics. The directional deformation, equivalent elastic strain, equivalent stress, and total deformation of the micro diaphragm were analyzed for four different pressure values. Similarly, the fluidic pressure in the microcavities was analyzed at a different velocity.

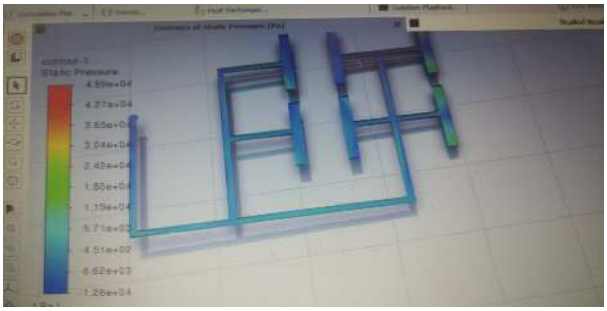


Fig. 9 *Computational fluidic analysis*

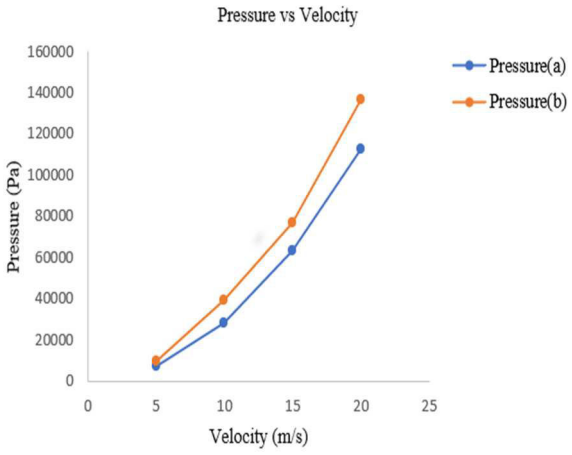


Fig. 10 *Pressure vs velocity*

Static Structural outputs

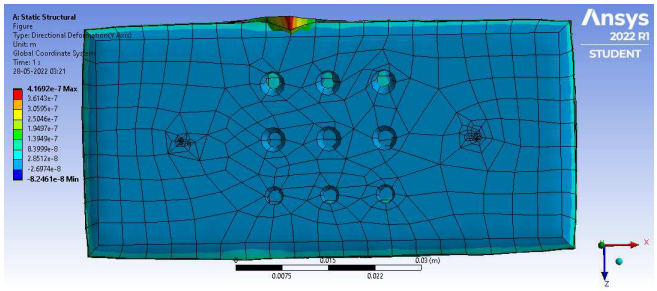


Fig. 11 Directional deformation

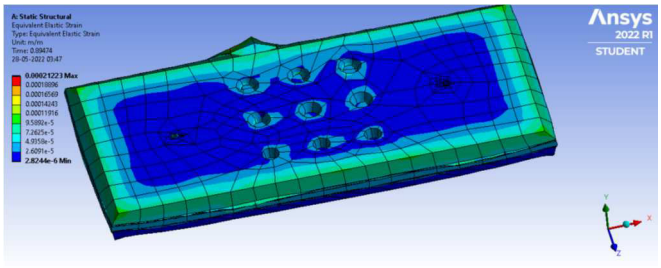


Fig. 12 Equivalent elastic strain

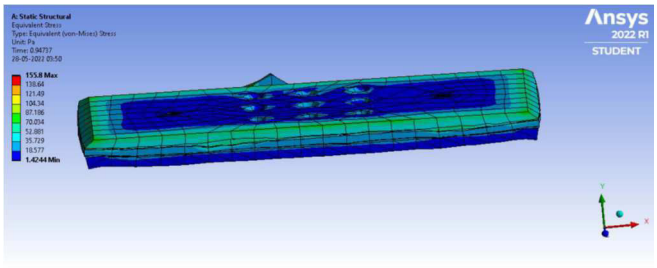


Fig. 13 Equivalent stress

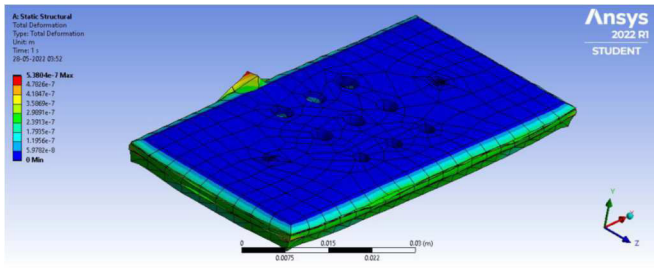


Fig. 14 *Total deformation*

The above figures 11,12,13 and 14 represent the different parameters like directional deformation, equivalent elastic strain, equivalent stress, and total deformation respectively. It was observed that the deformation of the model was increased with the increase in pressure value. As the deformation of the model was increasing, consequently strain and stress were also increasing. The static structure of the microfluidic chip was analyzed for different pressure values and the output achieved mentioned in below table 2:

ISBN 979-8-3503-5094-4

TABLE 2. STATIC STRUCTURAL OUTPUT

S. No.	PARAMETERS	PRESSURE		
		50 Pa	100 Pa	100 KPa
1	Directional deformation	4.16e-7	8.34e-7	0.0007
2	Equivalent elastic strain	0.0002	0.0004	0.42
3	Equivalent stress	155.8	311.59	3.11e5
4	Total deformation	5.38e-7	1.07e-6	0.0009

When the pressure was applied through the access hole, the deflection of the micro diaphragm in the respective breathing cavity was different. The strain and stress of the micro diaphragm at different pressure values were plotted as the pressure vs strain and pressure vs stress graphs have been shown below in Figures 15 and 16.

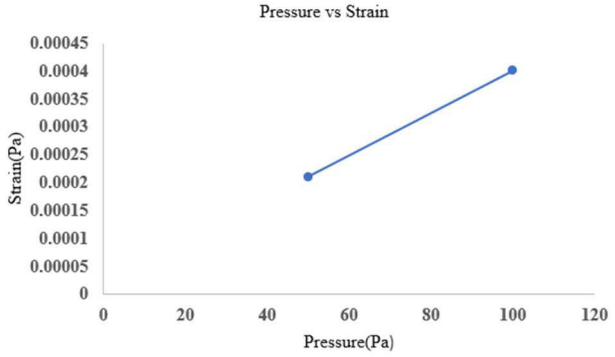


Fig. 15 *Pressure vs strain*

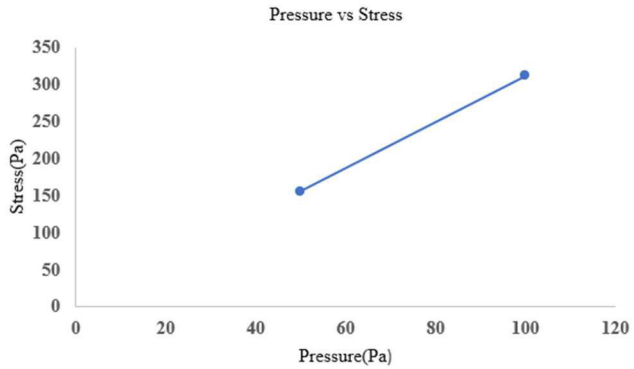


Fig. 16 *Pressure vs stress*

VI. DISCUSSION

During static structural analysis, the fixed support was given at the back of the body surface so that applied pressure won't change the position of the body and the deformation of the micro diaphragm can be analyzed accurately. The pressure was applied from the top plate to the pneumatic part through an access hole which causes 3D cyclic movement of the micro diaphragm in the respective breathing cavity. As a result, cell culture well cyclically deflected in response to the movement of the micro diaphragm. In the present model, there can be a possibility of inhomogeneous cellular spatial distribution and cell loss. To readdress this, we suggested a semi-open approach that allowed exact and precise management of the cells seeded on the porous membrane. The simulation was done using Ansys workbench software to validate the micro diaphragm deflection. To analyze the 3D cyclic deflection of micro diaphragm during breathing mode static structure was used. Similarly, the computational fluidic dynamic was used to model fluidic flow during medium exchange operation mode as well as to analyze pressure at different points of the microcavities. Our future work will implement the cell culture results to determine the exact response of the particular results on the participants.

VII. CONCLUSIONS

The bioinspired human breathing lung-on-chip structure was inspired by the human diaphragm. The present study demonstrated that this microfluidic chip is simple and easy to handle as well as its compatibility makes it unique from the existing models. The practical demonstration of micro diaphragm deflection was done in Ansys workbench software and had been validated. While analyzing the chip, the pressure was applied through an access hole from the top plate to the pneumatic plate and found that the 3D cyclic deflection of the micro diaphragm was directly proportional to the applied pressure. In the future, a microfabricated chip will be made as a compatible and important tool for effective and efficient in vitro drug testing and screening. The potential of in vitro is to minimize ethical issues and save the lives from wrong dose or side effects of drugs. On the other hand, in-vivo drug testing may cause severe effects on the participants due to the wrong dose or drugs endangering the subject's lives. This can be solved in an effective way using in-vitro methods.

VIII. REFERENCES

[1] R. A. K. Possomato-Vieira, José S. and Khalil and O. 2. 0. E. S. E. and S. Modeling, "Human-Derived Organ-on-a-Chip for Personalized Drug Development," *Physiol. Behav.*, vol. 176, no. 12, pp. 139–148, 2017, doi: 10.2174/1381612825666190308150055.Human-Derived.

[2] W. Harvard, "Human Organs-on-chips." https://wyss.harvard.edu/technology/human-organs-on-chips/.

[3] C. Ma, Y. Peng, H. Li, and W. Chen, "Organ-on-a-Chip: A New Paradigm for Drug Development," *Trends Pharmacol. Sci.*, vol. 42, no. 2, pp. 119–133, 2021, doi: 10.1016/j.tips.2020.11.009.

[4] H. G. Yi, H. Lee, and D. W. Cho, "3D printing of organs-on-chips," *Bioengineering*, vol. 4, no. 1, 2017, doi: 10.3390/bioengineering4010010.

[5] B. Raheel, Chaudhry; Bruno, *Anatomy, Thorax, Lungs*. 2023.

[6] J. D. Stucki *et al.*, "Medium throughput breathing human primary cell alveolus-on-chip model," *Sci. Rep.*, vol. 8, no. 1, pp. 1–13, 2018, doi: 10.1038/s41598-018-32523-x.

[7] A. O. Stucki *et al.*, "A lung-on-a-chip array with an

ISBN 979-8-3503-5094-4

integrated bio-inspired respiration mechanism," *Lab Chip*, vol. 15, no. 5, pp. 1302–1310, 2015, doi: 10.1039/c4lc01252f.

[8] D. Huh, G. A. Hamilton, and D. E. Ingber, "From Three-Dimensional Cell Culture to Organs-on-Chips," *Trends Cell Biol*, vol. 21, no. 12, pp. 745–754, 2011, doi: 10.1016/j.tcb.2011.09.005.From.

[9] C. Xu, M. Zhang, W. Chen, L. Jiang, C. Chen, and J. Qin, "Assessment of Air Pollutant PM2.5 Pulmonary Exposure Using a 3D Lung-on-Chip Model," *ACS Biomater. Sci. Eng.*, vol. 6, no. 5, pp. 3081–3090, 2020, doi: 10.1021/acsbiomaterials.0c00221.

[10] V. Kumar *et al.*, "An In Vitro Microfluidic Alveolus Model to Study Lung Biomechanics," *Front. Bioeng. Biotechnol.*, vol. 10, no. February, pp. 1–13, 2022, doi: 10.3389/fbioe.2022.848699.

[11] Y. Zhu *et al.*, "A Biomimetic Human Lung-on-a-Chip with Colorful Display of Microphysiological Breath," *Adv. Mater.*, vol. 34, no. 13, 2022, doi: 10.1002/adma.202108972.

[12] K. C. Lin, C. Z. Yen, J. W. Yang, J. H. Y. Chung, and G. Y. Chen, "Airborne toxicological assessment: The potential of lung-on-a-chip as an alternative to animal testing," *Mater. Today Adv.*, vol. 14, p. 100216, 2022, doi: 10.1016/j.mtadv.2022.100216.

[13] S. Andreia, Sofia Barros; Ana, Costa ; Bruno, "Building three-dimensional lung models for studying pharmacokinetics of inhaled drugs," *Adv. Drug Deliv. Rev.*, vol. 170, pp. 386–395, 2021.

[14] "Ansys Workbench." https://www.ansys.com/products/ansys-workbench.

ISBN 979-8-3503-5094-4

Hydrogel-Based Bilayered Scaffold Incorporating Hydroxyapatite and Glycosaminoglycans for Advanced Osteoarthritis Therapy

Abiramy Soundararajan
Department of Biomedical Engineering
Rajalakshmi Engineering College
Chennai,India
abiramy.s@rajalakshmi.edu.in

Varsha Santhosh
Department of Biomedical Engineering
Rajalakshmi Engineering College
Chennai,India
varsha26san@gmail.com

Nandhitha T
Department of Biomedical Engineering
Rajalakshmi Engineering College
Chennai,India
nandutamil02@gmail.com

Abstract*— Osteoarthritis, the most prevalent musculoskeletal disorder, ranks among the ten most disabling diseases in developed nations. Due to its lack of nerves, lymphatic vessels, and blood vessels, the articular cartilage has no innate capacity for self-healing. The severity of the disease progression and the underlying subchondral bone influence the treatment modalities, which range from management therapies to pharmaceutical and surgical interventions. However, regenerating hyaline cartilage in a damaged articulating joint continues to pose a difficulty, necessitating the development of tissue engineering strategies. This study aims to create bilayered scaffolds with spatial variation that mimic native osteochondral tissue. Hydroxyapatite and glycosaminoglycans have been incorporated into the underlying subchondral bone and superficial cartilage of the freeze-thaw hydrogel. The hydrogel's porous structure was revealed by scanning electron micrographs (SEM). Additionally, the scaffold's interconnected porosity was determined using the liquid displacement method. Swelling and degradation studies were conducted to examine the biphasic scaffold's water uptake potential and weight loss profile. The release profile demonstrates the scaffold's sustained release of chondroitin sulfate. The biphasic scaffold's confirmation of the presence of hydroxyapatite and chondroitin sulfate was established through staining techniques, and their concentrations were also quantified using biochemical methods. In vitro studies using mammalian chondrocytes will be conducted to evaluate the scaffold's chondrogenic potential.

Keywords*— *Osteoarthritis; biphasic; osteochondral; hydroxyapatite; chondroitin sulphate; chondrocytes.

I. Introduction

Osteoarthritis (OA) stands as a persistent degenerative condition affecting the joints, ranking as the predominant type of arthritis and a leading cause of disability among the elderly population.OA is distinguished by the gradual erosion of articular cartilage, the resilient tissue that serves as a protective layer on the extremities of bones within joints [1,3]. This depletion of cartilage results in pain, stiffness, and a decline in joint functionality. The osteochondral complex is a two-layer structure that consists of articular cartilage and subchondral bone. Articular cartilage is a specialized tissue that is composed of water, chondrocytes (cartilage cells), type II collagen, and proteoglycans. Subchondral bone is a type of bone that is located directly beneath the articular cartilage [2].It is composed of water, collagen type I, and hydroxyapatite (HA), a mineral that gives bone its strength and rigidity [1]. The bone-cartilage interface is a complex tissue due to variations in its morphology, composition, and mechanical characteristics between bone and cartilage.

OA can manifest as a result of either traumatic events or degenerative factors, and it is characterized by the loss of cartilage, the thickening of subchondral bone, and the development of osteophytes, which are bony growths found at the joint's margins. In the initial stages of osteoarthritis (OA), conservative approaches involving medications and physical therapy are typically employed to alleviate pain and slow down the progression of joint damage [5]. However, these conservative treatments do not have the capacity to fully reinstate the original function of damaged cartilage. In more severe instances of OA, joint replacement, and osteotomy may be contemplated to reinstate joint function. Nevertheless, it's essential to acknowledge that these surgical procedures come with their own set of limitations, such as the risk of complications and the inability to completely restore the native function of articular cartilage [11, 13,22].

Tissue engineering strategies offer a promising approach to restoring and regenerating damaged articular cartilage [21]. Tissue engineering involves using biomaterials, cells, and bioactive molecules to create scaffolds that can mimic the properties of native cartilage[6,7,15]. Hydrogels are one type of biomaterial that is commonly used as a scaffold for articular. Hydrogels represent biocompatible, three-dimensional porous network constructs that can be tailored to have specific physicochemical and biological properties. Freezing and thawing techniques can be used to create hydrogels with increased mechanical strength [9,10].A scaffold ideal for mending osteochondral injuries should possess the capability to simultaneously rejuvenate and foster the recovery of both cartilage and subchondral bone [4]. In this work, a matrix was designed using a freeze-thaw technique. The chondral layer (articular cartilage) of the scaffold was composed of glycosaminoglycans, while the underlying calcified tissue (subchondral bone) was composed of nanohydroxyapatite. Through a comprehensive examination of its physicochemical properties and biological performance, we have demonstrated its potential to mimic the native extracellular matrix environment and promote chondrogenic differentiation. Furthermore, the scaffold exhibited sustained release kinetics, suggesting its suitability for extended use in osteoarthritis treatment .This research provides a foundation for further innovation in regenerative medicine and tissue engineering approaches for osteoarthritis therapy.

II. Materials and Methods

A. Materials

PVA/Poly (vinyl alcohol) with a molecular weight of 146,000-186,000Da, Gelatin with a molecular weight of 50,000Da and Chondroitin sulphate (CAS No.

ISBN 979-8-3503-5094-4

9082-07-9, Product No. Y0000280) were obtained from Sigma Aldrich.

B. Synthesis of nanohydroxyapatite

The hydroxyapatite (nHAp) utilized in this study was synthesized in-house via a wet chemical precipitation method employing calcium nitrate tetrahydrate ($Ca(NO3)2·4H2O$) and diammonium hydrogen phosphate ($(NH4)2HPO4$) as precursors. Initially, the precursors were dissolved separately in deionized water to form solutions with desired concentrations, which were then mixed under controlled stirring conditions. Subsequently, the mixed solution underwent a precipitation reaction under continuous stirring and the pH was adjusted using dilute ammonium hydroxide solution results in the formation of hydroxyapatite nanoparticles.

C. Fabrication of Hydrogel Scaffold

a) Polymer mixture

An 8:2 mixture of PVA and gelatin was dissolved in water and blended at 60 degrees Celsius for five hours to create an aqueous solution. The polymeric blend was then cooled down to ambient temperature prior to the introduction of particular biochemical signaling molecules [19,20]. For the subchondral portion, 2% nanohydroxyapatite (nHA) was added and mixed, while for the chondral portion, 2% glycosaminoglycan (GAG) was added and mixed [16,24].

b) nBiphasic hydrogel preparation

The precursor polymer solution containing nHA and GAG was carefully poured into a 6-well plate, taking care to prevent the introduction of any air bubbles. Subsequently, the solution was frozen at -4°C for duration of 24 hours and allowed to thaw at room temperature for 3 hours. This freeze-thaw sequence was duplicated two times, followed by a 48-hour process of lyophilization. The resultant hydrogels were preserved in a vacuum at room temperature. Following a 20-minute freezing interval, the GAG-rich layer was overlaid onto the nHA layer. This process was repeated to produce the biphasic hydrogel.

III. Characterization

A. Analysis of Morphology

The analysis of the biphasic hydrogel scaffold's morphology and porosity was performed through the utilization of a scanning electron microscope. In preparation for SEM examination, the scaffolds were gold sputter-coated and then imaged at an acceleration voltage of 3 kV.

B. Swelling Assessment

To evaluate the hydrogel scaffold's water absorption capacity, the scaffold was placed in simulated body fluid (SBF) at a temperature of 37°C. At designated intervals of 15, 30, 45, 60, 120, 240, and 480 minutes, the scaffolds were removed from the solution, and their weights were recorded to calculate the swelling ratio.

C. Study of Degradation

The degradation characteristics of the hydrogel scaffolds (n=3) were assessed by placed them in Phosphate-buffered saline (PBS) with a pH of 7.4 for a duration of six weeks. The experimentation was conducted at a temperature of 37°C with continuous agitation and the PBS was refreshed every 48hrs. At one-week intervals, the hydrogels were extracted from the solution and subjected to lyophilization. The weighed dried scaffolds underwent examination of their morphological changes using SEM analysis. Subsequently, the percentage of weight loss was calculated.

D. Release Profile Assessment

The study assessed the release kinetics of chondroitin sulfate-loaded hydrogels by placing the hydrogels in PBS (pH 7.4) and subjecting them to constant agitation at 37°C within a water bath shaker over a five-day period. At regular time intervals, PBS was collected and the amount of chondroitin sulfate was assessed at 525 nm using the Dimethylmethylene Blue Assay.

E. Determination of Porosity

The porosity of a scaffold can be determined using the liquid displacement method. This method involves immersing the scaffold in a known volume of liquid, such as hexane, and measuring the change in volume [14]. The porosity is then calculated using the following formula (1)

$$P\% = (Vol_1 - Vol_3)/(Vol_2 - Vol_3)*100 \quad (1)$$

- P% is the porosity of the scaffold
- Vol1 is the original liquid volume.
- Vol2 is the liquid volume after immersing the scaffold.
- Vol3 is the residual liquid volume after the scaffold has been removed.

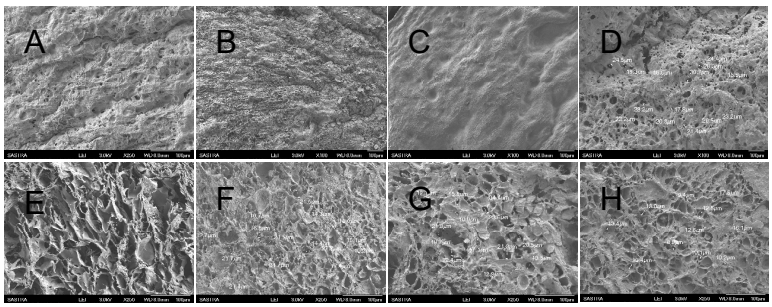


Fig.1: SEM image : Top surface of A) plain hydrogel B) hydrogel loaded chs C) hydrogel incorporated nHA D) Biphasic scaffold . Cross sectional image of E)) plain hydrogel F) hydrogel loaded chs G)) hydrogel incorporated nHA H) Biphasic scaffold

ISBN 979-8-3503-5094-4

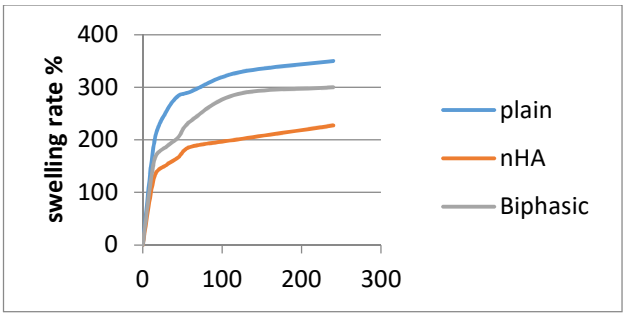


Fig.2.Hydrogel Scaffold Swelling Behaviour

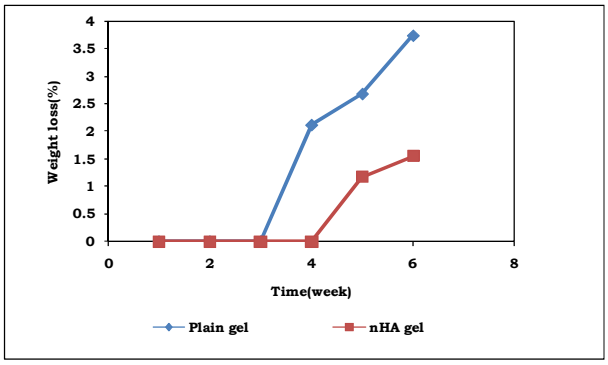


Fig.3.Degradation Profiles of PVA /gelatin blend and nHA/PVA/Gelatin Hydrogel scaffolds

F. Qualitative Analysis

Biochemical staining techniques were employed to confirm the presence of hydroxyapatite (nHA) and chondroitin sulfate (ChS) within the biphasic scaffolds. Hydrogel were placed in 1 ml of staining solution containing alizarin red S and alcian blue for one hour. Subsequently, they underwent a series of four to five rinses with double-distilled water to eliminate any surplus stain. Finally, the scaffold's nHA and ChS content was determined by assessing the intensity of the staining color.

G. Assessment of Biomolecule content

The determination of nanohydroxyapatite (nHA) content was carried out through thermogravimetric analysis (TGA), during which the mass variation in relation to temperature was monitored using a TGA instrument. An alumina pan was loaded with five milligrams of the sample, and under a nitrogen atmosphere, the sample underwent controlled heating from room temperature (23°C) to 1000°C at a ramp rate of 10°C per minute. The determination of nanohydroxyapatite (nHA) content within the scaffold was subsequently derived from the data obtained from the thermogravimetric curves.

IV. Results and Discussion

A. Analysis of Morphology:

SEM was employed to analyze the surface characteristics and pore structures of hydrogel scaffolds made from PVA/gelatin, PVA/gelatin with nHA, PVA/gelatin with chitosan (chs), and PVA/gelatin with both nHA and chitosan (chs). Fig.1A-D shows the surface morphology of the scaffolds, while Fig.1E-H shows their cross-sectional images. The SEM images confirmed that the hydrogel have a porous structure.

B. Swelling Assessment

Fig. 2 displays the swelling percentages of different hydrogel scaffolds in a pH 7.4 simulated body fluid (SBF) buffer. Notably, a distinction was observed in the swelling percentages between hydrogel scaffolds loaded with nanohydroxyapatite (nHA) and biphasic hydrogel scaffolds. This discrepancy can be attributed to the presence of negatively charged chondroitin sulfate, which attracts a significant amount of water into the scaffold In contrast, the inclusion of hydroxyapatite tends to decrease the scaffold's ability to swell.

C. Sudy of Degradation

The deterioration of hydrogel scaffolds was evaluated in vitro for 6 weeks. A detailed record of the weekly weight loss percentage was maintained, and the corresponding visual representation can be found in Fig.3. Notably, the hydrogel scaffolds loaded with nanohydroxyapatite (nHA) exhibited a slower degradation rate in comparison to the plain PVA/gelatin hydrogel. It was evident that these scaffolds maintained their initial weight for the first 2 weeks and then displayed a gradual weight loss over the subsequent weeks, ultimately reaching a noticeable decrease by the 6th week.

D. Release Profile Assessment

The release pattern of chondroitin sulfate from the scaffold was investigated over a 7-day duration, as illustrated in Fig. 4.Initially, within the first day, the biphasic scaffold (nHA/Chs/PVA/gelatin) and the Chs/PVA/gelatin scaffold exhibited chondroitin sulfate release percentages of 71% and 84%, respectively. Over the course of 7 days, both types of scaffolds consistently released chondroitin sulfate, with cumulative releases of 62% for the biphasic scaffold and 73% for the Chs/PVA/gelatin scaffold.

E. Liquid Displacement Method

A liquid displacement technique, employing hexane as the displacing fluid, was utilized to evaluate the interconnected porosity of the scaffolds. Table I summarizes the porosity measurements for different category of scaffolds. Among the three types of hydrogel scaffolds, the biphasic scaffold exhibited the highest porosity of 83%, while the nHA incorporated scaffold and the plain hydrogel scaffold showed porosities of 71% and 82%, respectively.

Table I. Porosity Evaluation in PVA/Gelatin, nHA/PVA/Gelatin, and nHA/Chs/PVA/Gelatin Scaffolds.

Sl.No	Hydrogel	Porosity %
1	plain hydrogel	82
2	Hydrogel incorporated Nha	71
3	Biphasic hydrogel	83

ISBN 979-8-3503-5094-4

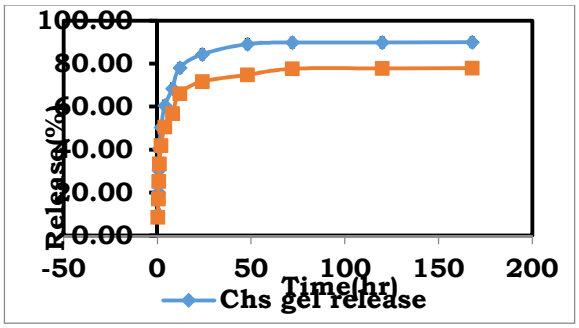


Fig.4. Chondroitin Sulfate Cumulative Release Profiles from Hydrogel Scaffolds

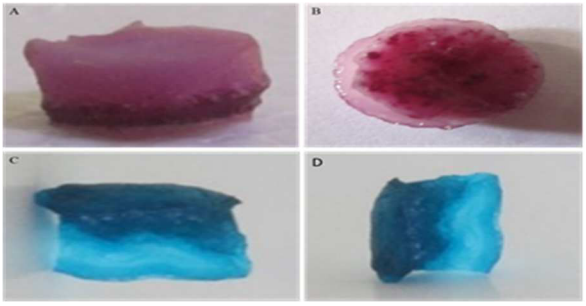


Fig.5 . Qualitative analysis A and B) Utilizing alizarin red S dye. C and D) Employing alcian blue stain for the hydrogel scaffold.

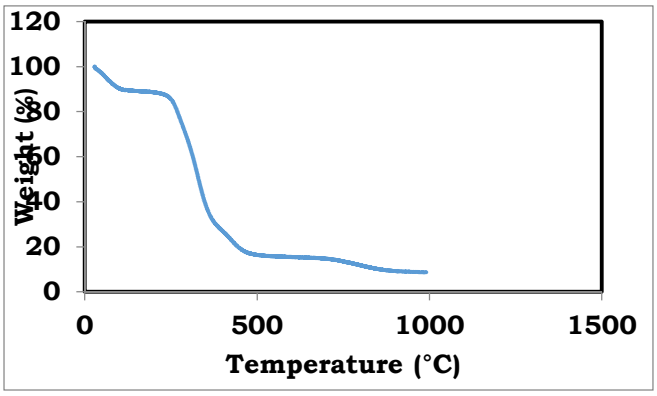


Fig.6 : Thermogravimetric analysis (TGA) of biphasic scaffold

F. Qualitative Analysis

Confirmation of the presence of biomolecules within the hydrogel was achieved through specific staining techniques. In Fig. 5 A-B, represents staining with alizarin red S while Fig. 5 C-D illustrates the staining using alcian blue of the hydrogel.

Alcian blue selectively binds to the sulfate and carboxylate groups of chondroitin sulfate, producing a distinct, intense blue coloration. This staining technique facilitates the identification and differentiation of the chondral layer. Similarly, the hydroxyapatite layer of the bilayered scaffold was stained with a red hue due to the formation of the alizarin red S-calcium complex upon interaction with nanohydroxyapatite (nHA).

G. Assessment of Biomolecule content:

For the quantification of nanohydroxyapatite (nHA) content in the hydrogel scaffold, a thermogravimetric analysis (TGA) was performed (Fig. 6). The results indicated that approximately 8% of the initially loaded nHA content was encapsulated in the scaffold after heating the sample to 1000ºC.

V. CONCLUSION

A novel biphasic hydrogel scaffold has been fabricated by employing the freeze-thawing method, featuring a subchondral layer of nanohydroxyapatite (nHA) and a chondral layer of chondroitin sulfate (CS).Scanning electron microscopy (SEM) revealed the scaffold's intricate porous structure, with an estimated porosity of 83%. This highly porous network facilitates nutrient and oxygen transport, enabling cell migration and proliferation essential for tissue regeneration. The incorporation of chondroitin sulfate (CS) and hydroxyapatite (nHA) within the biphasic scaffold was not only visually confirmed through staining procedures but also quantitatively assessed using both biochemical and thermogravimetric analyses. Swelling and degradation studies further characterized the scaffold's properties, revealing its ability to absorb and retain water, mimicking the natural hydration properties of cartilage tissue. Additionally, the scaffold exhibited a gradual degradation profile, ensuring a controlled release of degradation products and a sustained support for tissue regeneration. The release kinetics of chondroitin sulfate (CS) from the biphasic scaffold was evaluated, demonstrating the scaffold's ability to control the release of this critical component. This controlled release profile is essential for promoting cartilage tissue regeneration by providing a sustained supply of CS, which plays a crucial role in chondrocyte proliferation and extracellular matrix (ECM) synthesis.The biphasic scaffold exhibits promising physicochemical properties, including a highly porous structure, biocompatibility, and controlled degradation profile. These characteristics make it a potential candidate for serving as an articular cartilage substitute, offering a promising therapeutic approach for restoring damaged cartilage tissue.

ACKNOWLEDGMENT

The authors wish to acknowledge Rajalakshmi Engineering College and Saveetha dental College for infrastructural support.

REFERENCES

[1] Y.Li, M.Zhou, W.Zheng, J.Yang, and N.Jiang,"Scaffold-based tissue engineering strategies for soft–hard interface regeneration," *Regenerative Biomaterials*, vol. 10, Nov. 2022.

[2] M.Bhattacharjee, J.Coburn, M.Centola, S.Murab, A.Barbero, D. Kaplan, I.Martin, and S.Ghosh, "Tissue engineering strategies to study cartilage development, degeneration and regeneration," *Advanced Drug Delivery Reviews*, vol. 84, pp. 107–122, 2015.

[3] S.Frenkel, B.Toolan, D.Menche, M.Pitman, and J.Pachence, "Chondrocyte Transplantation Using a Collagen Bilayer Matrix for Cartilage Repair," *The Journal of Bone and Joint Surgery*, vol. 79, no. 5, pp. 831–836, 1997.

[4] X.Guo, J.Liao, H.Park, A.Saraf, R.Raphael, Y.Tabata, F. Kasper, and A.Mikos,"Effects of TGF-β3 and preculture period of osteogenic cells on the chondrogenic differentiation of rabbit marrow mesenchymal stem cells encapsulated in a bilayered hydrogel composite," *Acta Biomaterialia*, vol. 6, no. 8, pp. 2920–2931, 2010.

ISBN 979-8-3503-5094-4

[5] R.Reyes, A.Delgado, E.Sanchez, A.Fernandez, A.Hernandez, and C. Evora,"Repair of an osteochondral defect by sustained delivery of BMP-2 or TGFβ1 from a bilayered alginate-PLGA scaffold," *Journal of Tissue Engineering Regenerative Medicine*, 2012.

[6] N.Castro, C.O'Brien, and L.Zhang,"Biomimetic biphasic 3-D nanocomposite scaffold for osteochondral regeneration,"*AIChE Journal*, vol. 60, no. 2, pp. 432–442, 2013.

[7] R.Schek, J.Taboas, S.Segvich, S.Hollister, and P.Krebsbach, "Engineered Osteochondral Grafts Using Biphasic Composite Solid Free-Form Fabricated Scaffolds,"*Tissue Engineering*, vol. 10, no. 9-10, pp. 1376–1385, 2004.

[8] T.Cao, K.Ho, and S.Teoh, "Scaffold Design and in Vitro Study of Osteochondral Coculture in a Three-Dimensional Porous Polycaprolactone Scaffold Fabricated by Fused Deposition Modeling," *Tissue Engineering*, vol. 9, supplement 1, pp. 103–112, 2003.

[9] C.Hassan and N.Peppas, "Cellular PVA hydrogels produced by freeze/thawing," *Journal of Applied Polymer Science*, vol. 76, no. 14, pp. 2075–2079, 2000.

[10] C.Jiang, H.Chiang, C.Liao, Y.Lin, T.Kuo, C.Shieh, Y.Huang, and R. Tuan, "Repair of porcine articular cartilage defect with a biphasic osteochondral composite," *Journal of Orthopaedic Research*, vol. 25, no. 10, pp. 1277–1290, 2007.

[11] G.Chen, T.Sato, J.Tanaka, and T.Tateishi, "Preparation of a biphasic scaffold for osteochondral tissue engineering," *Materials Science and Engineering: C*, vol. 26, no. 1, pp. 118–123, 2006.

[12] H.Da, S.Jia, G.Meng, J.Cheng, W.Zhou, Z.Xiong, Y.Mu, and J.Liu, "The Impact of Compact Layer in Biphasic Scaffold on Osteochondral Tissue Engineering," *PLoS ONE*, vol. 8, no. 1, 2013.

[13] K.Shimomura, Y.Moriguchi, C.Murawski, H.Yoshikawa, and N. Nakamura,"Osteochondral Tissue Engineering with Biphasic Scaffold: Current Strategies and Techniques," *Tissue Engineering Part B: Reviews*, vol. 20, no. 5, pp. 468–476, 2014.

[14] Z.Pan, P.Duan, X.Liu, H.Wang, L.Cao, Y.He, J.Dong, and J.Ding, "Effect of porosities of bilayered porous scaffolds on spontaneous osteochondral repair in cartilage tissue engineering," *Regenerative Biomaterials*, vol. 2, no. 1, pp. 9–19, 2015.

[15] S.Badylak, D.Freytes, and T.Gilbert, "Extracellular matrix as a biological scaffold material:Structure and function,"*Acta Biomaterialia*, vol. 23, pp. S17–S26, 2015.

[16] M.Azami, M.Moosavifar, N.Baheiraei, F.Moztarzadeh, and J.Ai, "Preparation of a biomimetic nanocomposite scaffold for bone tissue engineering via mineralization of gelatin hydrogel and study of mineral transformation in simulated body fluid,"*Journal of Biomedical Materials Research Part A*, vol. 100A, no. 5, pp. 1347–1355, 2012.

[17] S.Suri and C.Schmidt,"Photopatterned collagen–hyaluronic acid interpenetrating polymer network hydrogels," *Acta Biomaterialia*, vol. 5, no. 7, pp. 2385–2397, 2009.

[18] D.Eglin, S.Maalheem, J.Livage, and T.Coradin,"In vitro apatite forming ability of type I collagen hydrogels containing bioactive glass and silica sol-gel particles,"*J Mater Sci: Mater Med*, vol. 17, no. 2, pp. 161–167, 2006.

[19] E.Crispim, "Hydrogels based on chemically modified poly(vinyl alcohol) (PVA-GMA) and PVA-GMA/chondroitin sulfate: Preparation and characterization," *expresspolymlett*, vol. 6, no. 5, pp. 383–395, 2012.

[20] S.Nanda, N.Sood, B.Reddy, and T.Markandeywar, "Preparation and Characterization of Poly(vinyl alcohol)-chondroitin Sulphate Hydrogel as Scaffolds for Articular Cartilage Regeneration," *Indian Journal of Materials Science*, 2013, pp. 1–8.

[21] Q.Loh and C.Choong,"Three-Dimensional Scaffolds for Tissue Engineering Applications: Role of Porosity and Pore Size," *Tissue Engineering Part B: Reviews*, vol. 19, no. 6, pp. 485–502, 2013.

[22] J.Amirian, N.Linh, Y.Min, and B.Lee, "Bone formation of a porous Gelatin-Pectin-biphasic calcium phosphate composite in presence of BMP-2 and VEGF," *International Journal of Biological Macromolecules*, vol. 76, pp. 10–24, 2015.

[23] K.Chang, L.Cheng, G.Ho, Y.Huang, and Y.Lee, "Fabrication and characterization of poly(γ-glutamic acid)-graft-chondroitin sulfate/polycaprolactone porous scaffolds for cartilage tissue engineering," *Acta Biomaterialia*, vol. 5, no. 6, pp. 1937–1947, 2009.

[24] E.Hago and X.Li, "Interpenetrating Polymer Network Hydrogels Based on Gelatin and PVA by Biocompatible Approaches: Synthesis and Characterization," Advances in Materials Science and Engineering, 2013, pp. 1–8.

[25] C.Yang, L.Xu, Y.Zhou, X.Zhang, X.Huang, M.Wang, Y.Han, M. Zhai, S.Wei, and J.Li, "A green fabrication approach of gelatin/CM-chitosan hybrid hydrogel for wound healing,"*Carbohydrate Polymers*, vol. 82, no. 4, pp. 1297–1305, 2010.

ISBN 979-8-3503-5094-4

Metallonanoparticles with antibacterial and anticarcinogenic activity extracted from *Cinnamomum zeylanicum* bark

Dr. D. Kalyani[1*],
corresponding author Faculty, Division of Biomedical Engineering, Department of ECE, Anna University, CEG, Chennai, Tamil Nadu, India
kalyanichandru2007@gmail.com

Geetha Gayathri[2]
Division of Biomedical Engineering, Department of ECE, Anna University, CEG, Chennai, Tamil Nadu
gayu.gg0054u @gmail.com

Kailaash RM[3]
Division of Biomedical Engineering, Department of ECE, Anna University, CEG, Chennai, Tamil Nadu, India
kailaash06@gmail.com

Christopher Dennis J [4]
Division of Biomedical Engineering, Department of ECE, Anna University, CEG, Chennai, Tamil Nadu, India
christopherdennis0603@gmail.com

Abstract: ***Nanoparticles are one of a kind for their size, shape and diverse applications. A blend of metallic nanoparticles with herbs improves its properties and application than bulk materials. Synthesis of zinc oxide from Cinnamomum zeylanicum bark is done by sol-gel treatment. UV-visible spectroscopy results clearly indicate the absorption peak of ZnO nanoparticles. Fourier transform infrared spectroscopy confirmed the presence of phytochemicals and zinc oxide vibration in crystalline structure. X-ray diffraction exhibited a hexagonal wurtzite structure and confirmed the presence of crystalline structure while scanning electron microscopy report revealed the average particle size of about 70 nm and of spherical shape. Moreover, the antibacterial activity of Cinnamon derived nanoparticles against two main bacterial strains such as Staphylococcus aureus, Bacillus subtilis were assessed by agar well diffusion method. By cytotoxicity assay, lethal effect was determined. Also, we determined the anticancer activity of Cinnamon in HepG2 cell line. Thus the study reveals the synthesized Cinnamon derived zinc oxide nanoparticles serve as an excellent alternative source of antibacterial agent and an anticancer property, can be used for cancer treatment.***

Keywords: ***Cinnamomumzeylanicum, ZnO nanoparticles, SEM, MTT, agar diffusion assay, antibacterial activity, anti-cancer property.***

I. INTRODUCTION

An emerging area of interest for researchers across the globe includes nanotechnology and the synthesis of nanoparticles. Metallic Nanoparticles possess unique properties and diverse functions compared to bulk materials especially in biomedical engineering applications. Hazardous effect of nanoparticles produced by chemical methods was not safe for biomedical applications which were well documented, leading to finding an alternate approach. Due to increase in the demand for the production of nanoparticles in diverse field paved a way for the blossom of green biosynthesis where plant extricates has thrown into greater consideration and even than the microbial biosynthesis because of prerequisite and to maintain an aseptic environment [1]. Plant mediated nanoparticles were always beneficial over synthetic approaches due to cost-effective, eco-accommodating and safe for human applications [2].

Cinnamomum zeylanicum, a tiny tropical and evergreen tree mostly noted for its bark, furnishes the world with the regularly known zest, cinnamon. The bark of the tree is commonly used in cooking as a spice to enhance the flavor of food [3]. In ancient Ayurveda medicine, it acted as a medicinal oil to cure cold, cough and due to its antimicrobial property widely used for stomach related discomforts [4].

Many scientific reports on chemical synthesis of silver and gold nanoparticles were well documented [5] [6] [7].However,the biosynthesis of zinc nanoparticles utilizing plant extricates has not yet been investigated much. Arinjoy Datta et al [8] reported on the potential of Parthenium hysterophorus leaf extracts in the production of zinc oxide nanoparticles and evaluated their antimicrobial potential. Santhosh Kumar and co-workers [9] explored antibacterial activity of nanoparticles from Passiflora caerulea against urinary tract infection. Matinise and co-workers [10] reported the synthesized nanoparticles using the Moringa oleifera's natural extract and investigated their morphological, structural, optical, and electrochemical activity.

The current research includes biosynthesis of zinc nanoparticles from the bark of natural spice cinnamon zeylanicum. To characterize nanoparticles we performed XRD, FTIR, UV-visible spectroscopy, SEM and EDX and also tested on two bacterial strains gram positive and negative strains to screen antibacterial activity of cinnamon bark mediated ZnO nanoparticles. Also, we found the inhibitory effect of nanoparticles on cancer cell growth.

ISBN 979-8-3503-5094-4

II. MATERIALS AND METHOD

A. Cinnamomum zeylanicum bark extraction

From the local market C.zeylanicum bark was acquired,dried under daylight and powdered. Then, 2.5 g of Cinnamon powder weighed and dissolved in 100 ml double distilled water and boiled at 60ºC for 5 minutes.

B. Biosynthesis of Zinc Oxide nanoparticles

By sol-gel treatment, Zinc Oxide nanoparticles are prepared by weighing 2 g of Zinc acetate dihydrate and 8 g of sodium hydroxide, dissolved in double distilled water and for five minutes stirred constantly. Later, the content was titrated against ethanol untill white precipitate was formed. Then 1ml of Cinnamon zeylanicum bark extract was added to 50 ml of Zinc Oxide solution, kept at room temperature for 8 hrs and later centrifuged for 15 minutes at 5,000 rpm. While drying in a hot air oven for 5 hours at 60˚C, Zn(OH)2 were converted to ZnO. Further, the powder is weighed and used for characterization studies.

C. Characterization studies

X-ray diffraction studies were done in Bruker AXSD8, where crystalline phase and particle size were determined with scanning range of 10˚-70˚ and bond angle of 3˚. Absorption spectra of ZnO nanoparticles were obtained within a range of 200-700 nm by using a double beam UV-visible spectrometer (Rayleigh, UV-2100. With aid of Buker 66V/S FTIR spectrophotometer, probed for existence of functional groups between range of 4000 cm^{-1} to 400 $cm.^{-1}$ Surface Topography were observed by scanning electron microscope LEO 1430VP and EDX by HITACHI S-4160, with a voltage of 20 KV. For each experiment, samples ranging from 0.5 mg to 1.5 mg were used.

D. Antibacterial method

Agar plate with a volume of microbial species Bacillus subtilis and Staphylococcus aureus were swabbed on it. Holes of diameter 6 to 8mm were punched antiseptically and volume of 25–100 µl of cinnamon derived zinc oxide nanoparticles were loaded into the well. Further, addition were assigned to aromatic rings, C≡C and C-N stretches were depicted in Table1[16] [17]. Since these active functional groups were commonly observed in phytochemicals like flavonoids, alkaloids, tannins of plant extract might be responsible for synthesizing nanoparticles [8].of 10µg of tetracycline at 37ºC as a control was done and incubated for 24 hours. Then, agar plates were incubated under suitable sterile conditions for the growth of microorganisms. Later, diameter of the zone of inhibition was measured which implies the diffusion of antimicrobial agents and its inhibition effect on microbial strains.

E. MTT assay

HepG2, human liver cancer cell line (NCCS, Pune) were maintained at humidified atmosphere of 5% CO_2 at 37°C in DME medium supplemented with 10% fetal bovine serum, penicillin (100µl) and streptomycin (100µg/ml). Assay reflects metabolic events of live cells by measuring NADH dependent oxidoreductase activity and inhibitory effect of nanoparticles on HepG2 (Mosmann1983). Cell concentrations of about 1×105/well were plated in 1 ml of medium/well in 24-well plates. Cinnamon zeylanicum derived nanoparticles at a series of concentrations (µg/ml) were loaded to wells. After 24 hrs of incubation, 20µl of MTT solution of 5 mg/ml was added to individual wells for further 4 hrs incubated 37^0C.Thus, viable cells were counted by measuring absorbance at 570 nm spectrophotometrically. The inhibitory effect of zinc oxide nanoparticles on hepatic tumor cells were measured by the formula (1)

$$\text{Cell Viability \%} = \frac{Abs\ of\ treated\ cells}{Abs\ of\ control\ cells} \text{x}100 \quad (1)$$

III. RESULTS AND DISCUSSION

A. DIFFRACTION STUDIES

To know the size, peak intensity, positions of peak and width of zinc oxide nanoparticles extracted from cinnamon bark XRD studies were done. Figure 1 reveals the diffraction crests situated at 28.02˚, 33.49˚, 37.716˚, 41.8˚, 47˚and 56.3˚ were distinguished to start from (002), (101), (102), (100), (110) and (113) planes. The presence of planes in XRD spectrum strongly shows that ZnO has wurtzite structure when compared with JCPDS card No.89-7102 and also confirmed that nanopowders doesn't contain any attributes other than ZnO peaks [11] and wurtzite phase was probably due to lattice stability and iconicity [12]. The size of the nanoparticle was achieved by Debye-Scherrer equation (2).

$$d = (0.9\lambda)/\beta cos\theta \quad (2)$$

where 0.9-Scherrer's steady, λ-wavelength of X beams,θ-Bragg diffraction edge and β, full-width half-maximum extreme of diffraction peak, average size of particles found to be 5.26 nm

B. FTIR SPECTROSCOPY

The spectrum in the range of 400-4000 cm^{-1} was shown in Figure 2 clearly states absorption peaks were closer to 490 cm^{-1} confirms presence of ZnO. The peak between 450-500 cm^{-1} represented a metal oxide bond [13] [14]. The observation of peak between 1380-1390 cm-1 implies aldehyde group and occurrence at 3241 cm-1 due to alcoholic and phenolic OH bonds [15]. Existence of peaks at 1453cm^{-1},2412cm^{-1} and 1260 cm^{-1}

C. UV-VISIBLE SPECTROSCOPY

UV-visible spectroscopy was done to reveal occurrence of ZnO NPs. Absorption spectra in the UV-Visible region was due to oscillation of conduction band electrons around the nanoparticle surface. Thus surface plasmon resonance effect was due to the specific vibration mode which in turn depends on size and shape of nanoparticles [14]. Figure 3 depicts the spectrum recorded between 200 to 700 nm. Henceforth, the peak noticed at 375 nm might be due to nanoparticle resonance [15].

D. SEM AND EDX ANALYSIS

Figures 4a and 4b revealed the existence of spherical shape particles with smooth edges of diameter 70 nm at two different magnifications of 10µm and 50 µm with the aid of JSM-6360 scanning electron microscope [8]. In Figure 5,

ISBN 979-8-3503-5094-4

the intense peaks were found with weight percentages of 39.67 for Zinc and 60.33 for oxygen and their atomic percentage contributions were 13.86 of Zinc and 86.14 of oxygen were shown in Table 2

E. AGAR WELL DIFFUSION ASSAY

To screen antimicrobial activity of cinnamon extracts by measuring zone of inhibition on Staphylococcus aureus and Bacillus subtilis bacterial strains [16] with varied volumes such as 25, 50, 75 and 100µl by agar disc diffusion method were shown in Figures 6 and 7. ZnO nanoparticles extracted from Cinnamon revealed diameter of zone of inhibition ring more than 11mm in all the series of dilutions which clearly states that there was a significant reduction in growth rate of both the bacterial strains when compared with untreated ZnO nanoparticles. High incidences of growth inhibition observed in Staphylococcus compared to Bacillus were due to the simplicity of its outer membrane.Unlike Staphylococcus, Bacillus inherits thick lipopolysaccharide encompassing peptidoglycan cell wall [18]. Thus, the study explored cinnamon derived ZnO nanoparticle has good potential due to presence of phytochemicals in its bark and inhibitory action was due to oxidative stress that disrupts membrane of pathogens[17].Thus, ZnO nanoparticles possess antibactericidal properties against both gram positive and gram negative bacteria.

F. CELL LINE AND CULTURE

Figures 8a to 8e depicts the effect of different concentrations of zinc oxide nanoparticles from cinnamon bark on HepG2 cell lines. The percentage of cell viability showed 15.95% of viable cancer cells and 84.05% of lethal cancer cells. To achieve IC50, only 12.5µg/ml of zinc nanoparticles was required for cancerous cell to exhibit lethality [19]. Decline in survival rate of HepG2 of about 84% was noticed at concentration of 250µg/ml Cinnamon nanoparticles. Hence, the extract significantly reduced cancerous cell growth as the dosage increased in a gradual manner. Polyphenols and flavonoids in Cinnamon extract were responsible for its cytotoxic activity which prevented cancerous cell growth, henceforth act as cancer preventing agents [20].Thus, polyphenols may act synergistically with the cinnamaldehyde present in the extract and thereby induce the increased cytotoxic activity compared to the commercial cinnamaldehyde [14] [15].

IV. CONCLUSION

Sol-gel method was used to prepare zinc oxide nanoparticles from Cinnamomum zeylanicum bark extract. The XRD spectrum depicted the size of ZnO nanoparticles as 5.26 nm and also wurtzite phase.FTIR spectrum showed that the characteristic absorption peaks of ZnO were closer to 490 cm^{-1} and presence of functionally active reactive groups. This clearly confirmed that zinc nanoparticles can be a suitable candidate for tailoring, thereby enhancing the properties of nanoparticles.In SEM analysis, nanoparticles were spherical in shape with smooth edges and its average diameter was around 70 nm. The steady increase in antimicrobial activity of extract was observed as an increase in extract concentration.The zone of inhibition ranges from 12-15 mm for sensitive Staphylococcus aureus bacteria. By IC50, 12.5µg/ml of sample to kill 50% of cancerous cell observed from MTT assay Thus, cinnamon bark derived metallic zinc nanoparticles were effective in exhibiting their antimicrobial effect against staphylococcus and bacillus species responsible for causing dental caries.

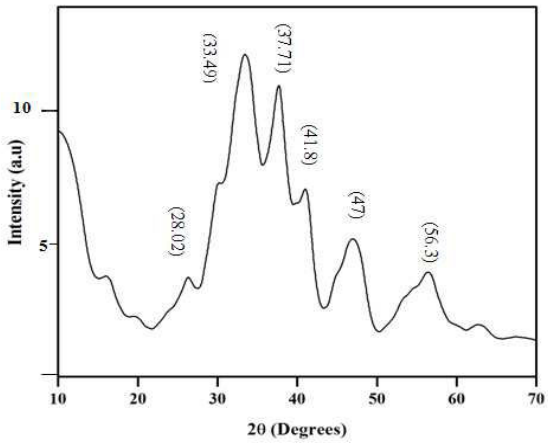


Fig 1 XRD pattern of ZnO nanoparticles from Cinnamon bark

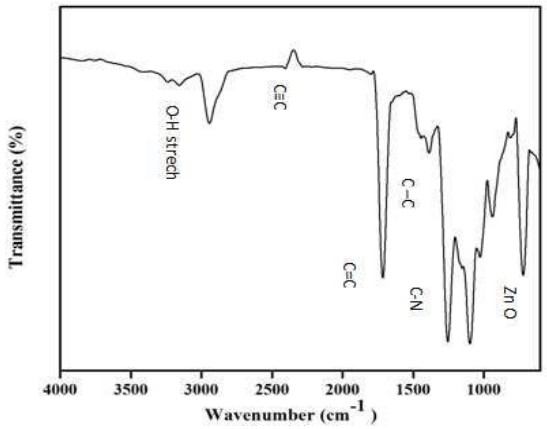
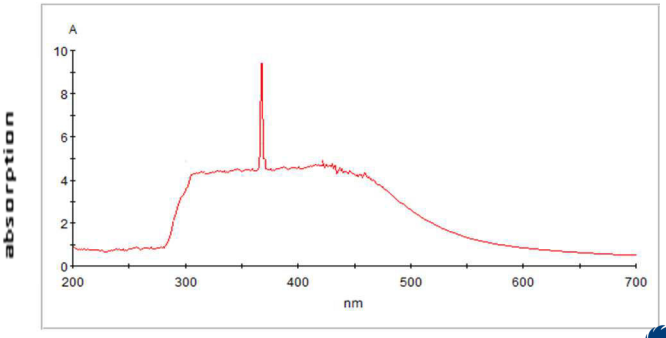


Fig 2 FTIR spectra of ZnO nanoparticles extracted from Cinnamon bark.

TABLE I: represents FTIR spectra shows functional groups of ZnO nanoparticles extracted from Cinnamon bark.

S. No	Peak (cm^{-1})	Vibration	Functional groups
1	3241	O-H stretch	Alcohol,phenol
2	2980	C-H stretch	Alkanes
3	2412	C≡C stretch	Alkynes
4	1453	C=C stretch	Aromatic ring
5	1260	C-N stretch	Aromatic amine
6	1128	C-O stretch	Secondary alcohol
7	1050	C-N stretch	Aliphatic amines



ISBN 979-8-3503-5094-4

Fig 3 absorption spectra of ZnO nanoparticles from cinnamon bark by UVspectroscopy.

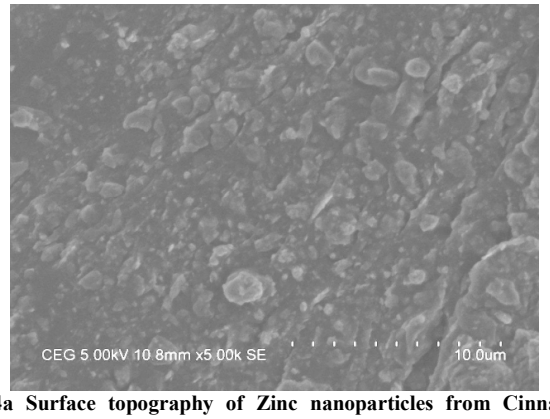


Fig 4a Surface topography of Zinc nanoparticles from Cinnamon bark by SEM analysis at 10 µm magnification.

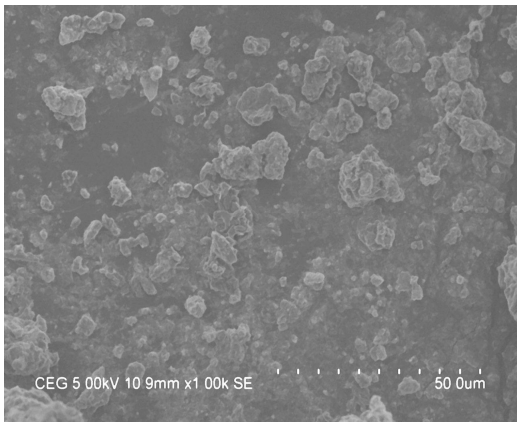


Fig 4bSurface topography of Zinc nanoparticles from Cinnamon bark by SEM analysis at 50 µm magnification.

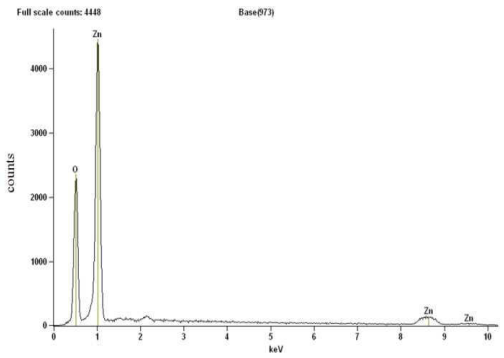


Fig 5 representation of EDX of zinc nanoparticles from Cinnamon extracted from bark.

TABLE II Elemental composition of ZnO nanoparticles extracted from Cinnamon by EDX analysis.

S. NO	Element	Net Counts	Weight (%)	Atom (%)
1	O	23530	60.33	86.14
2	Zn	1610	39.67	13.86
3	Zn	54053	---	---
			Total	100.00

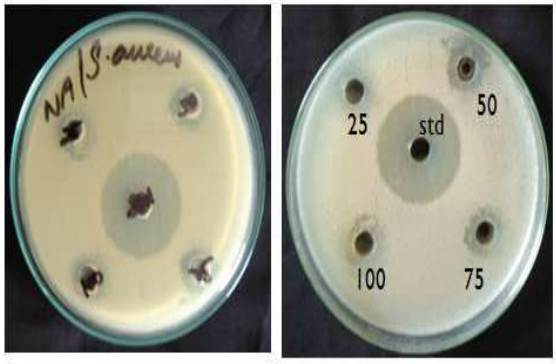


Fig 6 Zone of inhibition of zinc nanoparticles of Cinnamon bark against Staphylococcus aureus by agar well diffusion method

Fig 7 Zone of inhibition of zinc nanoparticles of Cinnamon bark against Bacillus subtilis by agar well diffusion method.

TABLE III The diameter of the zone of inhibition ring of ZnO nanoparticles extracted from cinnamon bark against gram positive (S.aureus) and gram negative (B. subtilis) bacterial species.

S.no	Microbes	Zone of inhibition (mm)				
		25 µl	50 µl	75 µl	100 µl	Control
1	S.aureus	12	13	13	14	32
2	B.subtilis	12	14	14	15	25

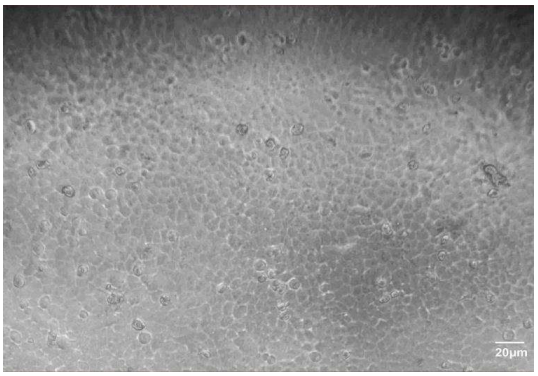


Fig 8a Control of zinc nanoparticles to HepG2 cancer cell medium.

ISBN 979-8-3503-5094-4

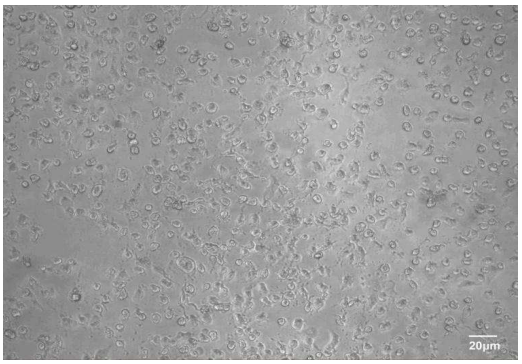


Fig 8b 25 µg/ml of zinc nanoparticles to HepG2 cancer cell medium.

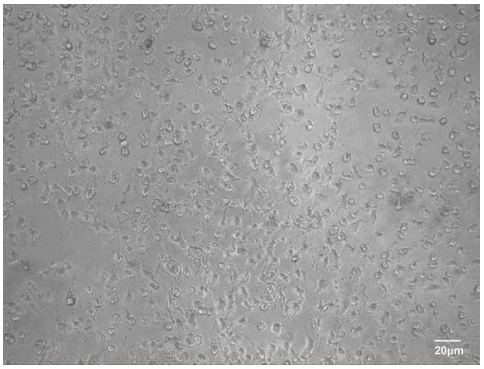


Figure 8c 50 µg/ml of zinc nanoparticles to HepG2 cancer cell medium.

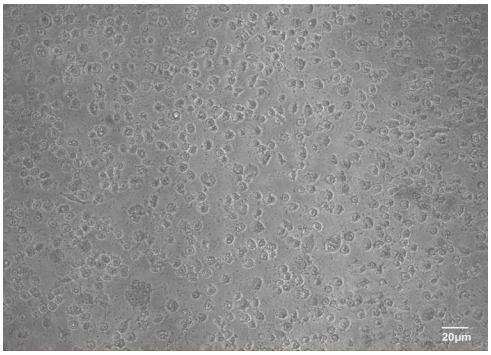


Fig 8d 100 µg/ml of zinc nanoparticles to HepG2 cancer cell medium.

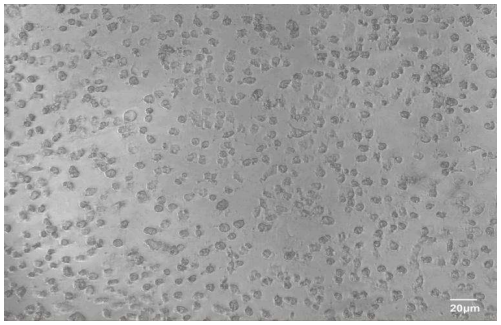


Fig 8e 250 µg/ml of zinc nanoparticles to HepG2 cancer cell medium

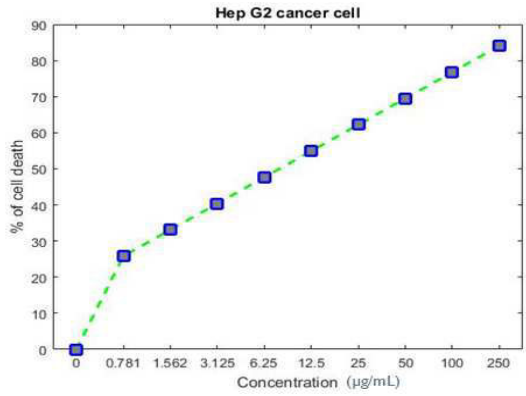


Fig 9 Cytotoxicity of zinc nanoparticles to HepG2 cancer cell medium

REFERENCES

[1] Sorna PremaRajendran and KandasamySengodan,Synthesis and Characterization of Zinc Oxide andIron Oxide Nanoparticles Using *Sesbania grandiflora* Leaf Extract as Reducing Agent'. HindawiJournal of Nanoscience 2017, Article ID 8348507.

[2] Monika Sienkiewicz, Anna Głowacka , Edward Kowalczyk, Anna Wiktorowska-Owczarek ,Marta Jóźwiak-Bębenista and Monika Łysakowska,`The Biological Activities of Cinnamon,Geranium and Lavender Essential Oils' *Molecules* 2014, *19*, 20929-20940; doi:10.3390/molecules191220929.

[3] Rodr´iguez, C. Ner´in, R. Batlle,` New Cinnamon-Based Active Paper Packaging against *Rhizopusstolonifer*Food Spoilage'*journalof Agriculture and Food Chemistry* .2008, *56,* 6364–6369

[4] Laura E. Hill, Carmen Gomes, T. Matthew Taylor,` Characterization of beta-cyclodextrin inclusion complexes containing essential oils (trans-cinnamaldehyde, eugenol, cinnamon bark, and clove bud extracts) for antimicrobial delivery applications' Elsevier journal 2012, Food Science and Technology 51 (2013) 86e93

[5] M. Sathishkumar, K. Sneha, S.W.Won, C.-W. Cho, S. Kim, Y.-S. Yun,` *Cinnamon zeylanicum*bark extract and powder mediated green synthesis of nano-crystalline silver particles and its bactericidal activity' Elsevier journal 2009, Biointerfaces 73 (2009) 332–338

[6] Munnaluri Gauthami1, Nivedita Srinivasan1, Neelam M. Goud, KasilingamBoopalan and Kavitha Thirumurugan,` Synthesis of Silver Nanoparticles using *Cinnamomumzeylanicum*Bark Extract and its Antioxidant Activity' Springer journal 2015, *Nanoscience & Nanotechnology-Asia*, *5,* 2-7

[7] Mehran Torki& Mohsen Akbari &KeyomarsKaviani,` Single and combined effects of zinc and cinnamon essential oil in diet on productive performance, egg quality traits, and blood parameters of laying hens reared under cold stress condition' International Journal of Biometeorology 2016, 59:1169–1177

[8] .ArinjoyDatta, Chandi Patra, HimaniBharadwaj, Sukhjeet Kaur, Neha Dimri and RobinkaKhajuria,`Green Synthesis of Zinc Oxide Nanoparticles Using Partheniumhysterophorus Leaf Extract and Evaluation of their Antibacterial Properties' Journal of Biotechnology & Biomaterials 2017,7:3

[9] J.Santhoshkumar, S.VenkatKumar,S. Rajeshkuma, `Synthesis of zinc oxide nanoparticles using plant leaf extract against urinary tract infection pathogen' Elsevier journal 2017, 59–465

[10] N. Matinise, X.G. Fuku, K. Kaviyarasu, N. Mayedwa, M. Maaza,` ZnO nanoparticles via Moringaoleifera green synthesis: Physical properties & mechanism of formation', Elsevier journal 2017, 339–347

[11] Surabhi Siva Kumar, PutchaVenkateswarlu, VankaRanga Rao and GollapalliNageswara Rao,` Synthesis, characterization and optical properties of zinc oxide nanoparticles' Springer open journal 2013, 3:30

[12] S.Vijayakumar, G.Vinoj, B. Malaikozhundan, S. Shanthi, B. Vaseeharan, *Plectranthusamboinicus*leaf extract mediated synthesis of zinc oxide nanoparticles and its control of methicillin resistant *Staphylococcus aureus* biofilm and blood sucking mosquito larvae,Spectrochim. Acta Part A: Mol. Biomol. Spectroscopy. 137 (2015) 886–891.

[13] J.N. Hasnidawani, H.N. Azlina, H. Norita, N.N. Bonnia, S. Ratim ,E.S. Ali,`Synthesis of ZnO Nanostructures Using Sol-Gel Method' Elsevier journal 2016, 211 – 216

[14] P.C. Nagjyothi, T.V.M. Sreekanth, C.O. Tettey, Y.I. Jun, S.H. Mook, Characterization, antibacterial, antioxidant, and cytotoxic activities of ZnO nanoparticles using *CoptidisRhizoma*, Bioorg, Med. Chem. Let. 24 (17) (2014) 4298-430.

[15] R.Dobrucka, J. Długaszewska , Biosynthesis and antibacterial activity of ZnO nanoparticles using *Trifoliumpratense*flower extract, Saudi J. Biol. Sci. (2015).

[16] https://orgchemboulder.com/Spectroscopy/irtutor/aromaticsir.shtml

[17] A.M. Awwad, B. Albiss, A.L Ahmad, Green synthesis characterization and optical properties of zinc oxide nanosheets using *Olea europea*leaf extract, Adv. Mat. Let. 5(9) (2014)520-524.

ISBN 979-8-3503-5094-4

[18] Songee Beak, Hyeri Kim, and Kyung Bin Song,Sea Squirt Shell Protein and Polylactic Acid Laminated Films Containing Cinnamon Bark Essential Oil, Journal of food science,(2008)

[19] Germana Esposito, Roberta Teta, Roberta Miceli, Luca S. Ceccarelli , Gerardo Della Sala, Rosa Camerlingo, Elena Irollo , Alfonso Mangoni, Giuseppe Pirozzi ,Valeria Costantino,` Isolation and Assessment of the in Vitro Anti-Tumor Activity of Smenothiazole A and B, Chlorinated Thiazole-Containing Peptide/Polyketides from the Caribbean Sponge, Smenospongiaaurea`,marine drugs 2015, 13, 444-459

[20] Soumya J Koppikar, Amit S Choudhari, Snehal A Suryavanshi, Shweta Kumari, Samit Chattopadhyay, RuchikaKaul-Ghanekar,`Aqueous Cinnamon Extract (ACE-c) from the bark of Cinnamomum cassia causes apoptosis in human cervical cancer cell line (SiHa) through loss of mitochondrial membrane potential' BMC Cancer journal 2010, 10:210

ISBN 979-8-3503-5094-4

Applications of Polymers in Tissue Engineering – A Review

Trupti Jain
Department of Biotechnology
Sathyabama Institute of Science and Technology
Chennai, India
truptijain371@gmail.com

Arun Karthick Selvam
Department of Biomedical Engineering
Sri Sivasubramaniya Nadar College of Engineering, Kalavakkam
Chennai, India
arunkarthicks@ssn.edu.in

Vijayavarsine R
Department of Biotechnology
Sathyabama Institute of Science and Technology
Chennai, India
varsinerajesh@gmail.com

R Tejashree
Department of Biotechnology
Sathyabama Institute of Science and Technology
Chennai, India
tejashreerani8622@gmail.com

Mithila R
Department of Biotechnology
Sathyabama Institute of Science and Technology
Chennai, India
Mailmithila23@gmail.com

Haritha S
Department of Biotechnology
Sathyabama Institute of Science and Technology
Chennai, India
haritha2000rtn@gmail.com

Monica V
Department of Biomedical Engineering
Sri Sivasubramaniya Nadar College of Engineering, Kalavakkam
Chennai, India
monica1931@bme.ssn.edu.in

*Devi Baskar
Department of Biotechnology
Sathyabama Institute of Science and Technology
Chennai, India
famidevi@gmail.com

***Abstract*—Tissue engineering has developed as a groundbreaking field pointing to reestablish, repair, or recover damaged tissues and organs. Among the different materials utilized in tissue building, polymers have picked up critical consideration due to their flexible properties, tunability, and biocompatibility. Polymer scaffolds act as a 3D framework that forms the extracellular matrix, providing all kinds of support and guiding tissue growth. The choice of polymer materials, including natural polymers such as chitosan, collagen and synthetic polymers like polyethylene glycol (PEG), polylactic-co-glycolic acid, affects the biodegradability, mechanical strength and cellular interactions of the scaffold. Integration of manufacturing technologies such as electrospinning and 3D printing has increased the accuracy and complexity of scaffolds. Polymer-based drug delivery systems help to release medicine slowly, which can help tissues heal and reduce swelling. Special polymers that can respond to the body's signals can deliver medicine directly to the places it's needed. This makes the treatment work better and reduces bad effects.**

Additionally, polymers are important for protecting transplanted cells and helping them become part of the body's tissues. Hydrogels are a type of material that can hold a lot of water. They are good for holding cells because they are like natural tissue. Although we have made progress, there are still problems to solve. We need to make sure that polymer materials are strong and work well with our bodies. We also need to make sure they last a long time and don't cause a bad reaction from our immune system. Current studies are looking at making plastic mixtures better, trying out new ways to make them, and making new materials that are a mix of different things to solve these problems. This review mainly focuses on the application of polymers in the tissue engineering field and the current trends has been applied.

Keywords— polymers, tissues, scaffold, lyophilization, electrospinning.

I. INTRODUCTION

Human body is made up of complicated systems and millions of cells which interact with each other and perform co-ordinately [1]. Tissue plays an important role in our human body it covers and provide shape and helpful for conserving heat and energy storage. Tissues are broadly classified into four types, they are connective tissue, epithelial tissue, nervous tissue and muscle tissue. Human body in nature possess various properties and functions such as the restoration of cells and also tissues have the capacity to heal by itself. This self-healing capacity refers to regeneration of tissue or development of tissue on its own after damage but some tissues does not possess self-healing capacity, so tissue engineering is used for repairing those tissues. Tissue engineering is the method that is used for repairing damaged tissue when it cannot heal on its own and it plays a very important role. Tissue engineering is a growing field in which, a wide range of materials are used. Among the materials, polymers are predominantly used because of their advantages over the other existing materials.

II. TYPES OF POLYMERS USED IN TISSUE ENGINEERING

A. Natural polymers

Natural polymers have an excellent interaction with cells due to their bioactive properties, which allows to improve the functionality of cells in biological systems [2]. Besides their advantages, the disadvantages are low mechanical and thermal stability which increases the risk of spreading diseases and the difficulties in cleaning. In order to overcome the disadvantages of natural polymers, synthetic polymers have been rapidly developed as biomaterials.

B. Synthetic polymers

Synthetic polymers have many characteristics that are required in tissue engineering applications due to their well-controlled macromolecular architectures and uniformity. Biomaterial scaffolds can be classified into three main

ISBN 979-8-3503-5094-4

material classes, namely architectural materials with fibrous and porous structures, hybrid or composite materials and hydrogels.

Advantages

• Polymers, particularly within the frame of hydrogels and nanofibrous platforms, can make three-dimensional structures that closely mirror the local tissue environment. This 3D engineering is significant for directing cell behaviour and advancing tissue recovery. Polymers are broadly accessible, making them available for analysts and clinicians. This accessibility encourages investigate, improvement, and execution in tissue building applications.

• Numerous polymers utilized in tissue designing are cost-effective and promptly accessible. This common sense makes them appropriate for large-scale generation and execution in different restorative application

• The tunable nature of polymers empowers the plan of platforms with particular geometries, surface properties, and corruption profiles. This customization is fundamental for tending to the assorted necessities of diverse tissues and organs.

• The points of interest of utilizing polymers in tissue designing contribute to their noteworthiness as flexible and successful materials for making frameworks and builds that back tissue recovery and repair. The capacity to tailor properties, control corruption, and coordinated bioactive particles makes polymers a profitable device within the headway of regenerative medication.

Disadvantages

• The long-term solidness and solidness of polymer-based inserts may be a concern. Over time, changes within the fabric properties, such as debasement or misfortune of mechanical quality, may affect the execution of the designed tissue.

• Polymers can imitate certain viewpoints of the extracellular framework (ECM), they frequently need the complexity and energetic intelligent found in local tissues. This confinement may influence cell signalling, separation, and large tissue usefulness.

• Tending to these drawbacks frequently includes a multidisciplinary approach, combining progresses in fabric science, science, and designing. Analysts proceed to investigate modern materials, creation methods, and procedures to overcome these challenges and optimize the utilize of polymers in tissue building for moved forward clinical results

III. TISSUE ENGINEERING TECHNIQUES

Various techniques utilized in application of natural polymers for tissue engineering are:

A. Freeze-drying or lyophilization

Freeze-drying which is also known as lyophilization, could be a method commonly utilized in tissue designing, especially when utilizing polymers. This process includes solidifying a fabric and after that evacuating the solidified dissolvable (as a rule water) through sublimation, taking off behind a permeable structure. In tissue building, freeze-drying with polymers is utilized to form platforms and lattices that can back cell development, give basic judgment, and encourage tissue recovery.

B. Solvent-Casting

Solvent casting could be a common procedure utilized in tissue designing for preparing polymers into lean movies or platforms. This strategy includes the disintegration of a polymer in a dissolvable, casting the polymer arrangement into a shape, and hence expelling the dissolvable to take off behind a strong polymer structure. Dissolvable casting is especially beneficial in tissue building due to its effortlessness, flexibility, and capacity to deliver structures with controlled porosity and geometry.[3]

Natural Polymers are applied in various oral applications, namely

1. Caries management
2. Management of periodontal diseases
3. Regeneration of dental structures by fabrications of membranes and scaffolds
4. Manufacturing and improvement of oral appliances and dentures
5. Delivery of drugs to oral systems

The advantages of using natural polymers in the field of medicine are their compatibility with the biological systems, production of these natural polymers have less or o effect on the environment as well on humans and a low immunogenic rejection of the natural polymer implant[4]. Natural polymers have the advantage of being produced and used at low cost. However, Natural polymers do have certain disadvantages, such as poor mechanical properties, high degradation rate, insufficient resources which inevitably lead to low reproducibility of natural polymers.

However, various steps have been made to tackle the issues. In order to overcome the disadvantages of natural polymers, they are combined with semi-synthetic or synthetic polymers.

IV. CURRENT SCENARIO AND APPLICATIONS OF POLYMER IN TISSUE ENGINEERING

Polymers plays an essential role in tissue engineering by offering properties for developing advanced biomedical devices some applications include:

A. Scaffolds

Scaffolds are the platforms where cell development and regeneration of tissue takes place. Scaffolds are the three-dimensional matrix which mimics the original extracellular matrix of the tissue [5]. Bone, tendon, skin, neural tissue, skeleton muscle, and vascular tissue are the scaffolds used in healing off tissues. Biopolymers are considered as an outstanding candidate material for various in vivo and invitro applications such as wound healing, medical devices for implant, tissue engineering and other health care products because of their biophysical and biochemical properties. Construction of scaffold is generally favoured with polymers that are naturally biodegradable since they are highly useful in developing an ideal scaffold that is necessary for cell development and tissue regeneration [6]. Ideal scaffolds are referred to scaffold that is identical to the natural extracellular matrix of the tissue. An ideal scaffold must own various biological properties like it must promote cell attachment, proliferation and migration, biocompatibility, biodegradability, non-toxicity, water and oxygen

ISBN 979-8-3503-5094-4

permeability and mechanical properties like tensile strength, stiffness and elastic modulus and the type of tissue that a scaffold is intended to be put into determines its features. For example, hard tissues like bone require a rigid polymeric scaffold, soft tissues like nerves demand a malleable polymeric scaffold etc. Different types of polymers are used in development of scaffold like natural polymers and synthetic polymers. In natural polymers due to their bioactive qualities they increase the interactions and functionality of the cells and few examples of natural polymers include protein, polysaccharide, polynucleotides etc. though natural polymers have magnificent biological properties for creating an ideal scaffold they lack mechanical properties which in turn developed the need for synthetic polymers which can be altered according to the need of application due to which it possess controlled macromolecular architectures and uniformities and examples of synthetic polymers include polyester, epoxy, polyethylene etc. and natural polymers classification include polysaccharide and protein-based polymers and classification of synthetic polymers include non-degradable and degradable polymers.

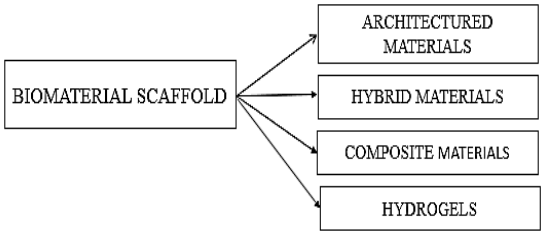


Fig 1: Types of biomaterials used in making of scaffold.

Fig 1 represents the various biomaterials used in developing a scaffold. Architectured materials, hybrid and composite materials and hydrogels are few of the main material categories of the biomaterial scaffolds. Synthetic biodegradable polymers are predominant source of architectured materials such as polylactic acid, polyglycosilic acid etc. and through various fabrication techniques, self-assembly, freeze drying etc., these materials can be established. The blend of both synthetic and biodegradable polymers is categorised as hybrid polymers which constitute of few solvents and composite materials constitute of either layered structures or multi layered structures and through various process and fabrication method both hybrid materials and composite materials can be made. The crosslinked synthetic polymer or natural polymer networks through physical and/or covalent bonds is classified as hydrogels, they exhibit high potential for tissue repair and to the injury site and also it is used for delivering macromolecules of smaller invasions.

B. Electrospinning

Electrospinning [7] is a technique used to create artificial nanostructured scaffolds that mimic the cell's natural extracellular matrix. Electrospun nanofibers have gained popularity in tissue engineering scaffolds because of their high surface area, high porosity and interconnected pores which enables better activity of cells and growth of tissues. The process involves using an electric field to draw charged fibers of polymer solution or melt into extremely fine fibers that form a nonwoven mat that mimics the extracellular matrix (ECM) structure in tissues.

A typical setup includes [8] a syringe pump, high-voltage power supply and a grounded collector. A continuous jet flow of the polymer solution is formed when a high voltage is applied. The solvent evaporated as it reached the collector, this results in the formation of nanofibers.

Electrospinning is a technique that is extremely flexible as many different methods are used based on the type of fibers required. Some of them are explained as follows.

Needle electrospinning:

In this method, the metal needle is connected to the positive electrode and the collector to the negative electrode. Taylor cone is formed when a high voltage is applied between the metal tip and the collector which in turn eject the polymer solution from the tip of the Taylor cone. The solvent from the ejected polymer solution evaporated as it travels and the nanofibers are deposited over the collector.

Needleless electrospinning:

In this method, there are no needles involved. Instead, the electrospinning process is initiated on the surface of an open vessel filled with the polymer solution. Here fibers originate from a stationary or rotating spinneret. To ensure stable Taylor cone formation and eventually fiber generation, the key properly is to concentrate the electric charges on the liquid surface.

In this method, the metal needle is connected to the positive electrode and the collector to the negative electrode. Taylor cone is formed when a high voltage is applied between the metal tip and the collector which in turn eject the polymer solution from the tip of the Taylor cone. The solvent from the ejected polymer solution evaporated as it travels and the nanofibers are deposited over the collector.

Needleless electrospinning:

In this method, there are no needles involved. Instead, the electrospinning process is initiated on the surface of an open vessel filled with the polymer solution. Here fibers originate from a stationary or rotating spinneret. To ensure stable Taylor cone formation and eventually fiber generation, the key properly is to concentrate the electric charges on the liquid surface.

C. Bioprinting

Bioprinting is a developing technique where bio-inks are used to 3D-print or construct models that mimic living tissues. These bio-inks are made up of living cells and bio-materials that can imitate extracellular matrix. The process of 3D bio printing follows a digital file that acts as a blueprint to print a product layer-by-layer. Bioprinting is a recent technology, that is already showing vast help in the applications of

ISBN 979-8-3503-5094-4

regenerative and personalized medicine, tissue engineering, drug discovery, cosmetics, etc. Steps involved in bioprinting: Before bioprinting anything, we need a model that is digital and is scannable by the 3D printer. this model can either be designed from scratch or be scans of a certain existing biological object. we will also need to prepare the suitable bio-inks that would be required for the process. Then the bio-inks are loaded in one or more printheads, after which the parameters are set appropriately so that the product is developed desirably. Lastly the most important is crosslinking the produced product. This is typically done with the help of ionic solutions or UV light.

Uses of bioprinting include; drug development, printing of artificial organs, and constructing prints to heal wounds, etc. Though bioprinting is versatile and extremely useful it is still an evolving field. The first biocompatible 3D printer was developed in the early 1980s, and a cell-embedded 3D printer in 2003 by Thomas Boland, we notice the research in the field has travelled a long way in a short period and will continue to accelerate and discover the latest bioprinting technology.

V. CONCLUSION

In conclusion, the application of polymers in tissue engineering speaks to a significant progression within the field, advertising a flexible and promising stage for the plan and improvement of biomimetic platforms. The one-of-a-kind properties of polymers, such as tunable mechanical characteristics, biocompatibility, and controlled corruption, empower the creation of platforms that closely mirror the characteristic extracellular network. This biomimicry cultivates a conducive microenvironment for cell development, multiplication, and tissue recovery. Furthermore, the capacity to join bioactive particles and tailor the surface properties further improve the usefulness of these polymer-based develops. As investigate proceeds to disclose unused bits of knowledge and developments, the integration of polymers in tissue designing holds incredible potential for tending to complex challenges in regenerative pharmaceutical, clearing the way for the creation of more compelling and personalized helpful arrangements. Long term of tissue designing without a doubt rest on the proceeded investigation and refinement of polymer-based approaches, introducing in a modern period of progressed and custom-made regenerative procedures.

Basically, natural and synthetic polymers are used for regeneration of tissues. Though the synthetic polymers possess many advantages over the natural polymers, their degradation profile restricts their applications in tissue engineering. In order to overcome this disadvantage, the synthetic polymers are mixed with natural polymer to enhance the degradation profile of the polymer. Hence researchers have to either improve the degradation profile of the synthetic polymers or improve the mechanical properties of natural polymers. This will lead to the new polymer materials with enhanced properties that can be used as an effective material for tissue regeneration.

REFERENCES

[1] S. Kliem, M. Kreutzbruck and C. Bonten, “Review on the biological degradation of polymers in various environments”,. Materials, vol. 13,pp. 1-18, October 2020.

[2] A. M. Coenen, K. Bernaerts, J. A. Harings, S. Jockenhoevel and S. Ghazanfari, S, “Elastic materials for tissue engineering applications: Natural, synthetic, and hybrid polymers’, Acta Biomater, vol. *79*, pp. 60–82, October 2018.

[3] A.S.A. Mohammed, M.Naveed and N. Jost, “Polysaccharides; Classification, Chemical Properties, and Future Perspective Applications in Fields of Pharmacology and Biological Medicine (A Review of Current Applications and Upcoming Potentialities)”, *J.* Polym. Environ., vol. *29*, pp. 2359–2371, January 2021.

[4] A. Iovene, Y. Zhao, S. Wang and K. Amoako, “Bioactive Polymeric Materials for the Advancement of Regenerative Medicine”, J. Funct. Biomater, vol. 12, pp. 14, February **2021**.

[5] Ghanbarzadeh, B.; Almasi, H. Biodegradable Polymers; IntechOpen: London, UK, 2013.

[6] M. I. Echeverria Molina, K. G. Malollari and K. Komvopoulos, “Design Challenges in Polymeric Scaffolds for Tissue Engineering”, Front. Bioeng. Biotechnol., vol. 9, pp. , June 2021.

[7] H. Mina, S. Hajir Bahrami, M. Ranjbar Mohammadi and P. B. Milan “Smart electrospun nanofibers containing PCL/gelatin/graphene oxide for application in nerve tissue engineering”, Mater. Sci. Eng. C, vol. 103, pp. 1-8, October 2019

[8] Y. K. Doğan, A. Demirural and T. Baykara, “Single-needle electrospinning of PVA hollow nanofibers for core–shell structures”, SN Applied Sciences, vol. 1(5), pp. 1 April 2019.

ISBN 979-8-3503-5094-4

Review on Biosensor: A Diagnostic tool for cardiovascular diseases

Meghna Rathish
Department of Biotechnology
Sathyabama Institute of Science and Technology
Chennai, India
meghnarathish3@gmail.com

Arun Karthick Selvam
Department of Biomedical Engineering
Sri Sivasubramaniya Nadar College of Engineering, Kalavakkam
Chennai, India
arunkarthicks@ssn.edu.in

Neeraja T
Department of Biotechnology
Sathyabama Institute of Science and Technology
Chennai, India
neeruthiyagarajan@gmail.com

Madhumitha B.S.
Department of Biotechnology
Sathyabama Institute of Science and Technology
Chennai, India
madhu2004bs@gmail.com

Haritha S
Department of Biotechnology
Sathyabama Institute of Science and Technology
Chennai, India
haritha2000rtn@gmail.com

**Devi Baskar*
Department of Biotechnology
Sathyabama Institute of Science and Technology
Chennai, India
famidevi@gmail.com

Shivangi Pandey
Department of Biomedical Engineering
Sri Sivasubramaniya Nadar College of Engineering, Kalavakkam
Chennai, India
Shivangi1950@bme.ssn.edu.in

Abstract— **For the last few years there has been a sudden increase in the occurrence of cardiovascular disease and it has been the constant reason for the increase in mortality rate all over the world. For the diagnosis of CVD, the earlier diagnostic tools used were ECG and immunoassay analysis which in turn delayed the results and caused more inconveniences. Hence to reduce the mortality rate it is better to create biomarkers using biosensors that can easily detect the heart rates and we can also quickly diagnose the presence of CVD. The Biomarkers are made of Biosensors which is an analytical device comprised of nanoparticles with novel bioelectronics and nanotechnology platforms. A variety of biosensor techniques are also available to diagnose the presence of CVD. This review mainly focuses on the basics of Biosensors, Traditional diagnostic tools for CVD, and the current trends followed using biosensors for the identification of CVD mainly Myocardial Infarction (heart attacks).**

Keywords—biosensor, biomarker, myocardial infarction, nanomaterials, diagnosis

I. INTRODUCTION

A. Biosensor-Principle and Design

Biosensor is an analytical device that can measure biological, chemical reactions by creating signals equal to the quantity of an analyte present in the reaction. The common applications of biosensors are drug discovery, detecting the presence of microorganisms that causes diseases and markers that are used as indicators in body fluids like blood, sweat, saliva, and urine [1]. The biosensor mainly comprises of 3 basic components:

1) Bioreceptors - It is the biological recognition system which provides high degree of selectivity for the analyte that has to be measured.

2) Analyte - A particular molecule that has to be measured. E.g. DNA, Enzymes, or Antibodies. This component present in the biosensors plays the role of converting biological signals to measurable electrical or optical signals.

3) Microelectronics interaction is found between the analyte and the bioreceptor which produces a signal that is further converted by the transducer and is finally quantified on the display where all these 3 components are connected in a series arrangement [2]. The purpose of the biosensor is to obtain reliable, quick, and accurate information about the analyte. Biosensors play a significant role in the medicinal field, food safety, bioprocessing, and industrial monitoring as well as in agriculture [3].

B. Classification of Bio-receptors

Bio-receptors are broadly classified into three types. They are physical (piezoelectric and thermoelectric), Optical (direct optical detection, labelled optical detection) and Electrochemical (amperometric, potentiometric, impedimetric and voltametric) bio-receptors [4] (Fig.1).

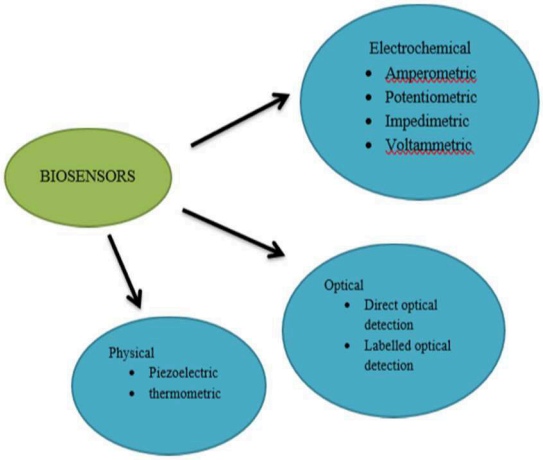


Fig. 1 Bio-receptors Classification

ISBN 979-8-3503-5094-4

C. Detection of Cardiovascular Disease with Biosensors

The strong influence of cardiovascular diseases mainly depends on the rapid turnaround time that is the duration of the patient's sample till the results are obtained. The main role of the biosensors is that they are used for the diagnosis of cardiovascular diseases as biomarkers. The biosensors are made up of nanoparticles which are important for the precise detection and diagnosis and act as markers of heart diseases.

II. BIOSENSOR AND BIOMARKERS IN MYOCARDIAL INFARCTION

Worldwide the primary cause of death is non-communicable diseases like cardiovascular diseases. Early diagnosis is very important in preventing CVD and also helps in reducing time, cost, and inconvenience [5]. Myocardial infarction, MI (heart attack) is a cardiovascular disease where the heart muscle is damaged due to the deficiency of oxygen as the muscle does not get enough blood [6]. Myocardial infarction can be divided into three categories, STEMI (ST-elevated myocardial infarction), unstable angina and NSTEMI (non-ST-elevated myocardial infarction) [7].

In recent years, electrocardiogram (ECG) has been the major technique for diagnosis of MI but it is not enough as only 57% patients exhibit ECG changes and it is limited [6]. One such example is the case of NSTEMI which involves partial blockage of arteries but in the ECG the ST segment doesn't seem elevated [7]. Also, Traditional methods include heavy medical instruments for electrocardiography and immunoassay analysis which have disadvantages such as time consumption and inconvenience [8].

A. Cardiac Biomarkers

An efficient method for myocardial infarction detection is the change in levels of cardiac biomarkers. These markers include enzymes, proteins, and hormones and they will emerge in blood as the heart is undergoing severe stress for a long period. For the detection of AMI, biomarker levels are usually detected in saliva or plasma [9]. Some of the Cardiac biomarkers involved are:

- Cardiol troponin T (cardiac troponin C (cTn C), cardiac troponin I (cTn I, for myocardial tissue) and cardiac troponin T (cTn T) which are the three regulatory proteins) [10]
- Brain natriuretic peptide (BNP)
- C- reactive protein (CRP) (long plasma half-life protein) [9]
- Creatine kinase (CK-MB)
- Myoglobin (Mb) (intracellular O2 binding hemoprotein) [9]
- Creatinine phosphokinase (CPK) etc.

B. Biosensors

Biosensors are involved in the detection of blood markers and contribute to quick diagnosis. In 1950, L.C. is the first biosensor used by Clark which was used to measure the level of glucose in the blood. A combination of biosensors with immunoassays and nanomaterials is employed to form highly sensitive sensing technologies [10]. Involvement of nanotechnology has improved biosensors where nanotubes, nanowires, metal nanoparticles, nano-patterned surfaces are used as electrodes are beneficial because of their large active surface area, increased efficiency of the transducer, improved amplified enzymatic signals, selectivity and specificity [11]. Cardiovascular disease biosensing can be categorized into wearable devices, point-of-care testing, and implantable devices [8]. Electrochemical biosensors are the first scientifically and commercially successful biosensors for multiple analytes. Some electrochemical biosensors are:

- A electrochemical biosensor which is label-free,using porous Graphene Oxide (GO) was constructed by Kazemi and colleagues for cTnI (troponin) detection where specifically it detects targets keeping the detection limit within 0.07 ng/mL-1 and a range of 0.1 to 10 ng mL-1.In addition, graphene oxide and gold nanoparticles are immobilized on a GC-Glassy Carbon electrode are used for large surface area to load cTnI antibody where that label-free electrochemical immunosensor was designed by Liu and others. At last, GO combined with ferrocene molecules, covered the antibody detection. It will detect the concentration 0.05 ng mL-1 cTnI.
- An aptasensor utilize ferrocene-modified silica nanoparticles (Fc-SiNPs) was formed by Jo and others. [9]
- CRP concentration in plasma increases and it was first reported in 1982 by an electrochemical CRP biosensor by Gupta and colleagues. It has a 11ng mL-1 detection limit with a suitable specificity. For CRP detection, an electrochemical aptasensor was formed by wang and others where silica microspheres are combined with AuNPs, which keeps signal molecules such as Zn2+ and produce a wider surface area for inactive anti-CRP. [9]
    - At the initial stage of AMI, the first biomarker detected is cardiac myoglobin (the first non-enzymatic protein). A two-dimensional nanostructure Black phosphorus (phosphorene), similar to graphene as it is black and scaly, has linked atoms in the form of sheets and electrical conductivity. To detect myoglobin, an electrochemical aptamer-functionalized black phosphorus is designed [9].
    - Changes in the level of Non-coding RNA are also used for diagnosis of MI. Kumarswamy and others showed evidence that lncRNA named LIPCAR present in late circulatory levels indicates the initial diagnosis of MI and chronic heart failure. Wu and others created a miRNA biosensor based on hybridization chain reaction. [9]

ISBN 979-8-3503-5094-4

Some optical biosensors are Surface plasmon resonance (SPR): the process using SPR was discovered in 1902. SPR sensor was created for detecting cardiac troponin T in serum. The usage of gold nanoparticles combined with an anti-cTnT detection antibody enhanced the result. An SPR immunosensor was utilized for point-of-care testing enabling advanced diagnosis of AMI. [12]

III. Advantages and Disadvantages of Biosensors

A. Advantages

- Biosensors offer high sensitivity and high selectivity.
- They are smaller in size and compact enabling on-site testing.
- Allows detection of multiple analytes.
- Ability to recognize different analytes such as protein, nucleic acids, etc.

B. Disadvantages

- Biosensors are affected by environmental conditions such as pH, temperature, etc
- They are capable of false results or false positives that lead to error
- It requires monitoring to maintain the stability and accuracy of the result
- Requires costly materials and instruments for development
- Some biosensors have a short and limited lifespan

IV. Applications of Biosensors

Recently, the development of bio-sensing technologies for rapidly diagnosing CVD markers has gained much attention. With the development of nanotechnologies and bioelectronics, new biosensor platforms are being developed to provide on-spot detection, precise measurement, and periodic monitoring of disease progression.

Traditionally, several sensing methods have been used based on the chemical, electrochemical, optical, and electromechanical properties. Recently, heartbeat signals and biomarkers based on cardiac blood are commonly used in clinics, as well as these biomarkers can be used for disease prediction. The emerging CVD biosensors can be worn and implanted, and also used in bioelectronics monitoring that allows for continuous measurement of the cardiac markers.

A. Classical Diagnostic Tools

For CVDs, standardized techniques of diagnosis depend on classical methods such as surface plasmon resonance (SPR), Liquid Chromatography (LC), immunoassays and enzyme-linked immunosorbent assay (ELISA) are difficult to perform because of their long procedure which in turn will consume time, less specificity and sensitivity.

B. Nanotechnology in CVD Diagnosis

The rapid growth of the nanotechnology industry is driven by the increased findings of affinity-based systems, which reduce analysis time and integrate multiple assays in a single model or unit of device, known as a lab on a chip (LOC). Microfluidics and the emergence of nanotechnology are leading to advances in diagnostics. Nano-based systems are a key driver of the rapidly expanding field of nanotechnology, which involves a multi-disciplinary research effort that includes nanofluids, micro electrons, and analytical chemicals.

The analysis of cardiac biomarkers in blood is an essential part of the current standard of care, especially when dealing with ACS/MI and HF. Following guidelines for clinical practice and clinical research, cTn, NPs, and CRP are all suggested for use as cardiac biomarkers which are quite specific to both HF and ACS/MI. Specifically, cTnI and cTnT are the preferred biomarkers for ACS/MI and the assessment of myocardial damage; they are also primarily utilized as prognostic variables of acute heart failure.

C. Biosensors a part of the Diagnostic tool

Researchers detected certain biomarkers using different types of biosensors, such as piezoelectric or calorimetric (physical), field-effect transistor, optical (fluorescence, luminescence, refractive index), electrochemical (amperometry, potentiometry, conductometry), capacitor-based biosensors, and optical (luminescence, luminescence, and conductometry). Among these biosensors, electrochemical biosensors are used as an alternative to traditional analysis methods because of their numerous advantages such as quick and simple to use, accurate, reasonably priced, and able to test multiple chemical compounds.

D. Nanomaterials Used in Biosensors

For the detection of clinically important biomarkers, the following nanomaterials have been used in biosensors: nanomaterials made up of carbon (graphene, carbon nanotubes (CNT), metal nanoparticles (Ag, Au, Pt, Cu), semiconductors quantum dots (Cd, Se, Ge, Te), metal oxides (ZnO, CeO2 and Al2O3) and silicon/ indium/gallium, etc.

By having both nanomaterials and analytical technology with different sensing mechanisms, various nanosensors are present with many advantages such as high sensitivity, good photostability, and favorable biocompatibility. Among these, researchers focused on electrochemical, optical, paper, and pressure-based nanosensors for diagnosing CVDs by detecting the required biomarkers or parameters of cardiovascular function.

V. Current Trends

The fact is that one biomarker is inefficient for the early detection of CVD. Rather, a single platform with several bio-markers detection techniques will improve the quality of life for those with CVD and lessen the financial strain on medical professionals. Recent research works were reviewed to get a greater comprehension of implementing optical-based biosensors like photonic crystals (pc), surface plasmon resonance (SPR), fluorescence-based techniques, fiber optics, and Raman spectroscopy (RS) biosensors to enable ultrasensitive identification of biomarkers of cardiac function.

The use of novel biosensors for identification, including thermoelectric, piezoelectric, optical, and

ISBN 979-8-3503-5094-4

electrochemical biosensors, is presently in its earliest stages however holds the potential to expand dramatically. The prevalence of CVDs has risen, and more novel research focusing on CVDs has emerged from the discovery of possible biomarkers besides to the already-known biomarkers.

VI. Conclusion

Biosensors are practical substitutes for conventional techniques that have been developed. The convergence of 3D bioprinting, microarray technology, and microfluidics has resulted in the development of novel cell-based biosensing technologies that are more accurate, sensitive, and productive. In vitro or small animal model studies conducted in vivo continue to serve the basis of many of these investigations. Its therapeutic value in identifying cardiovascular events will become clearer with more research. It is still difficult to design new biosensors that can identify diverse cardiovascular indicators.

The applications of novel detecting biosensors have to be expanded widely due to the integration of nanotechnology. In future, wearable cardiac monitors are predicted to play a significant role in medical care as they gain prominence as diagnostic tools.

References

[1] P. Mehrotra, “Biosensors and their applications - A review,” J Oral Biol Craniofac Res, vol. 6, pp. 153-159, August 2016.

[2] C.Karunakaran, R.Rajkumar and K. Bhargava.Introduction to biosensors. Biosensors and bioelectronics, 2015,pp.1-68.

[3] P.R.Coulet,.What is a Biosensor?. Biosensor principles and applications,”2019, pp. 1-6.

[4] V. Naresh, N.Lee, “A Review on Biosensors and Recent Development of Nanostructured Materials-Enabled Biosensors,” Sensors (Basel), vol 2, pp. 1-16, February 2021.

[5] M. Ouyang, L. Tong, M.Sarwar, A.Bhimaraj and Di Carlo, “A review of biosensor technologies for blood biomarkers toward monitoring cardiovascular diseases at the point-of-care,” Biosensors and Bioelectronics, vol. 171, pp. 1-31, January 2021.

[6] F. Yousefi, A. Movahedpour, Z. Shabaninejad, Y. Ghasemi, S. Rabbani, A. Sobnani-Nasab and H Mirzaei, “Electrochemical-based biosensors: new diagnosis platforms for cardiovascular disease,” Current medicinal chemistry,vol.27,pp. 2550-2575, 2023.

[7] M. Pawula, Z.Altintas and I. E. Tothill, “SPR detection of cardiac troponin T for acute myocardial infarction,” Talanta, vol. 146, pp. 823-830, January 2016.

[8] A. Pourali., M. R., Rashidi,Barar, G. Pavon-Djavid and Y. Omidi, “ Voltammetric biosensors for analytical detection of cardiac troponin biomarkers in acute myocardial infarction. TrAC,” Trends in Analytical Chemistry, vol. 134,2021

[9] S.Nazir and R.A.Iqbal, “Biosensor for rapid and accurate detection of cardiovascular biomarkers: Progress and prospects in biosensors,” Biosensors and Bioelectronics: X, vol.14,2023 pp 1-14

[10] J. A. Reyes-Retana and L. C Duque-Ossa, “Acute myocardial infarction biosensor: a review from bottom up,” Current problems in cardiology,vol. 46,2021.

[11] X. Tang, Y. Zhu, W. Guan, W.Zhou and P. Wei, “Advances in nanosensors for cardiovascular disease detection,” Life Sciences, vol. 305, 2022

[12] A. Qureshi, Y.Gurbuz and J. H. Niazi, “Biosensors for cardiac biomarkers detection: A review,” Sensors and Actuators B: Chemical, vol. 17 ,2012 ,pp. 62-76.

ISBN 979-8-3503-5094-4

3D printing in orthopedics and healthcare Applications – A Review

L. K. Tejhashwini
Department of Biotechnology
Sathyabama Institute of Science and Technology
Chennai, India
tejhashwinilk@gmail.com

Arun Karthick Selvam
Department of Biomedical Engineering
Sri Sivasubramaniya Nadar College of Engineering, Kalavakkam
Chennai, India
arunkarthicks@ssn.edu.in

K. Preetham Sai
Department of Biotechnology
Sathyabama Institute of Science and Technology
Chennai, India
Pree2605@gmail.com

S. Srimathi
Department of Biotechnology
Sathyabama Institute of Science and Technology
Chennai, India
srimathisrini.05@gmail.com

Shivangi Pandey
Department of Biomedical Engineering
Sri Sivasubramaniya Nadar College of Engineering, Kalavakkam
Shivangi1950@bme.ssn.edu.in

Devi Baskar
Department of Biotechnology
Sathyabama Institute of Science and Technology
Chennai, India
famidevi@gmail.com

*Haritha S
Department of Biotechnology
Sathyabama Institute of Science and Technology
Chennai, India
haritha2000rtn@gmail.com

***Abstract*—This article reviews the immense potential of 3D printing using biomaterials in orthopedics and its innovative applications in other medical specialties. By analyzing current research and case studies, the review highlights the benefits of this innovation that has led to the development of solutions for complex anatomical challenges and improved patient outcomes by providing personalized implant design, and enhanced biocompatibility. Using advanced additive manufacturing techniques, surgeons can optimize surgical interventions, tailor treatment approaches, and set a new standard of care for patients with musculoskeletal and other special conditions. The combination of biomaterials and 3D printing technology has brought about some significant advancements in the fields of orthopedics and other medical specialties. It highlights key advancements, challenges, and future directions in these rapidly evolving fields.**

Keywords—3D printing, Orthopedics, Ocular prostheses, Biocompatibility, Surgical innovation, Personalized implants

I. INTRODUCTION

Biomaterials are bio-compatible materials that enhance the healing process of body parts by supporting the repair and regeneration of the body tissues of those parts damaged by both external and *in vivo* factors (cancer, etc). Sometimes they even replace the affected body parts as a whole by physically and functionally replacing them in the long term. In the last 50 years, three generations of materials have evolved and have been put to use individually or in combinations (using cross-links) through various techniques according to their subjective need in a field. They are classified into the first-generation, second-generation, and third-generation biomaterials. Currently, we are in the fourth generation of biomaterials.

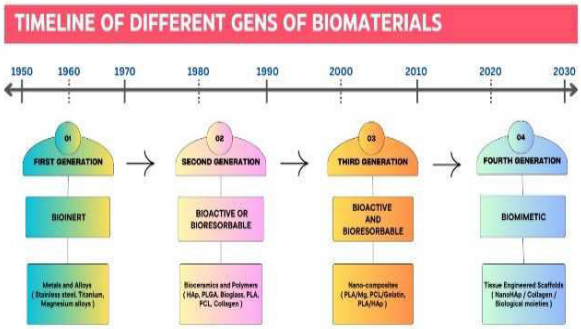


Fig 1. GENERATIONS OF BIOMATERIAL

The successive generation of biomaterials doesn't mean they are better than those belonging to the previous generation. For example, metallic stents (bio-inert material) are still used to replace knee fractures, cardiovascular defects, etc at present due to their easy availability. However, cons arose when they weren't able to interact with surrounding tissues, and adverse immune reactions were caused by them especially pure metals despite being strong and providing unmatched mechanical strength. Similarly, the second-generation biomaterials were bioactive, and porous and flexible structures (usually used in joints) but posed a challenge when it came to being applied during the case of critical repairs (ocular area) due to the lack of presence of biomolecules in them. Then came the third- generation of biomaterials with excellent bioactivity, which improved the functionality of parts promoted self-healing properties, and contained biomolecules. Complete replacements were healthily possible by these. One main ban was that it lacked anti-neoplastic drugs(taxanes). Now at present 4th fourth-generation biomaterials(bio-mimetic) have been developed and are put to use. These possess a high level of biocompatibility. They are developed to be multifunctional to tackle the combined anti-tumor and anti-infective effects. Smart biodevices are extensively preferred. They also have a targeted action for specific pathways and satisfy electrophysiology. Yet we aren't at the peak of biomaterial evolution. This present generation also possesses a few

ISBN 979-8-3503-5094-4

challenges. These still do not give us the controlled release of drug kinetics and also call for better fabrication techniques.

Musculoskeletal disorders can range from fractures and sports injuries to degenerative conditions like arthritis and spinal disorders caused due to trauma, disease, congenital abnormalities, or degenerative conditions. In addition, bone diseases alone account for 50% of chronic diseases in people over the age of 50. Together they bring great harm to the health and impact one's quality of life. It is imperative to address the significant demands that exist today in the field of orthopaedics which calls for constant advancements. In orthopedics, 3D printing enables the creation of implants customized to match a patient's specific needs with precise geometries and mechanical properties. This promotes osseointegration, thereby reducing the risk of complications related to the implant.

Despite the remarkable progress achieved in the field of biomaterials and 3D printing, significant challenges persist. Optimization of material properties, scalability of manufacturing processes, and translation of laboratory findings into clinical practice are among the key hurdles that must be overcome to fully realize the potential of these technologies. Additionally, regulatory considerations, ethical concerns, and cost-effectiveness remain important factors to consider in the widespread adoption of biomaterials and 3D printing in healthcare.

In this comprehensive review, we explore the evolution of biomaterials and 3D printing in medicine, highlighting their applications across various medical specialities and discussing current challenges and future directions. By examining recent advancements in material science, manufacturing techniques, and biomedical engineering, we aim to provide insights into the transformative potential of biomaterials and 3D printing in shaping the future of healthcare. Through interdisciplinary collaboration, innovative research, and strategic investment, we envision a future where biomaterials and 3D printing technologies play pivotal roles in delivering personalized, precise, and effective medical interventions, ultimately improving patient outcomes and enhancing quality of life.

II. ROLE of 3D PRINTED MATERIALS IN ORTHOPEDICS

The development of biomaterials has advanced as the number of surgical treatments in the orthopedic field has expanded. Biomaterials have osteobiologic properties, including osteogenic properties, osteo-conduction and osteo-induction. In Orthopedics, biomaterials are utilized in the creation of implants. These implant devices correct posture and movement(orthosis) or substitute a part of the body (prosthesis) while maintaining functionality [1]. Due to their ability to promote effective bone healing and improve patient outcomes osteobiologic and metallic biomaterials are being preferred more. Moreover with immediate consideration of the mechanical properties the suitable biomaterial is chosen according to the need for either a temporary or permanent replacement of the body part [4].
In the early stage, the most important characteristic for an orthopedic device to have was biological inertness, which resonates with being not affected by the biological environment. Some of our body fluids are highly corrosive and some are sensitive and can create harmful reactions and cause adverse effects like inflammation, etc. yet, bio-inertness does not help in bone healing. But for decades now biomaterials have continuously evolved. The recent advances including rapid prototyping using newly incorporated computer-aided design (CAD), computer-aided engineering (CAE), and computer-aided manufacturing (CAM) tools along with certain/suitable biomaterials in the orthopedic industry, particularly the manufacturing process of ortho-prosthetic aids are mind-blowing. On top of that these orthotic and prosthetic devices are manufactured efficiently in terms of cost, quality time, and quality, highly customizable and accurate in terms of matching the patient's anatomy, which elevates their self and social integrity making them feel more confident in themselves.

The manufacturing process is a carefully drawn out process where the subject's morphology is taken followed by individual manufacture through adjustments done to the prototype. However, the key revolution took place during the application of rapid prototyping technologies (RPT). Techniques such as Fused Deposition Modeling (FDM), Selective Laser Sintering (SLS), Laminated Object Manufacturing (LOM), and 3D printing (3DP) are some major examples of the available methods in the manufacturing industry [3].
In the 1970s the demand for attractive orthotic devices boomed. Due to this, new techniques like plastic coating (applying a tinted rubber-based plastic film) were developed. In the early 1980s, the rise of additive manufacturing technologies (AMT), popularly known as 3D printing technologies came into play. In the following years, other AMTs were introduced, such as Fused Deposition Modeling (FDM), laminated object manufacturing (LOM), Selective Laser Sintering (SLS), 3D printing, and variable rapid prototyping (Polyjet Technology) [2,4].

A. Fused Deposition Modelling (FDM)

This process involves extruding semi-molten material (Polycarbonate (PC) or Acrylonitrile butadiene styrene (ABS), known for their thermoplastic characteristics are commonly used materials in FDM.) through a moving extrusion head, creating layers in the X and Y axes to build a 3D model. Two extrusion nozzles compose the movable extrusion head- one to deposit the build material and the other that contains the support material. The process involves constructing layers by extruding the perimeter first and then filling the delimited zone. The main advantage of FDM technology is that it utilizes low-cost materials but the manufacture time is relatively high. Since then, several applications have shown up for FDM in the biomedical field for upper and lower limb orthoses, hand prostheses, facial prostheses, and drug delivery systems. The recent advancements in the FDM technology involve continuous development in the FDM materials, improvement in the speed and precision of the FDM printers, advancements in FDM that enable the use of simultaneous use of multiple materials during the printing process, innovations in support structures that reduces material waste and enhance the overall quality and the integration with other technologies like artificial intelligence and machine learning [4].

ISBN 979-8-3503-5094-4

B. Selective Laser Sintering (SLS):

The first SLS system was introduced by the DTM Corporation (Desktop Manufacturing) in the 90s. This SLS method uses CO_2 lasers to fuse powdered polymers and builds an object layer by layer. PA (polyamide)12, ABS (acrylonitrile butadiene styrene) and PC (polycarbonate) are the most commonly used thermoplastics. With high accuracies and low discrepancies, SLS has been successfully applied in manufacturing Ankle Foot Orthoses (AFO). It is also proven effective in creating splints for the upper extremities that show good results in manufacturing quality and improved design for comfort. SLS is promising in producing comfortable and functional rehabilitation devices despite limited clinical validations. The recent advancements in the SLS technology focus on refining the surface finish of SLS-printed parts and the integration with Industry 4.0 that incorporates smart technologies such as IoT sensors [4].

C. Powder Bed and Inkjet Head 3D Printing:

3DP refers to three-dimensional printing in which the product is made of powder layers that stick together using an adhesive. In this process, firstly a powder layer is spread onto the build platform followed by the deposition of liquid binder through an inkjet printhead. This forms a patterned layer by lowering the platform and repeating the process. The 3DP process is somewhat similar to SLS. In 3DP, a printing head places liquid adhesive in the material whereas, in SLS a CO2 laser is used to fuse the layers. The accuracy of this process is lower than SLS but is comparatively cost-effective and quick which makes its role predominant in the prototyping industry. Similar to SLS, it uses thermoplastics like ABS as it has the required properties to be used in orthotic and prosthetic applications. It was investigated whether this technology was suitable to produce functional prostheses, it was proven that, despite the manufacturing levels being limited, patients felt relatively more comfortable with prostheses made with 3DP machines than the traditional handmade ones but the resistance was not studied in that work and therefore the durability of the product was unknown. The 3DP technology was used for patient-specific maxillofacial implants that reported a reduction in operation times. The versatility and precision of 3DP make it particularly interesting for its application in regenerative medicine and tissue engineering [4].

III. APPLICATIONS IN OTHER MEDICAL SPECIALITIES

One of the notable uses of fused deposition modelling (FDM) is to create synthetic trabecular bone models using PLA/HA composite materials. PLA, a degradable polymer, is combined with hydroxyapatite (HA) to mimic the mechanical properties and biodegradability of real bone. While FDM allows for customized structures, incorporating HA reduces printing accuracy but enhances mechanical properties [5].

A study delves into the viability of utilizing 3D printing technology to fabricate patient-specific incisor teeth using biopolymers, specifically polylactic acid (PLA) and polyamide 11 (PA11). By conducting comprehensive mechanical tests, including tensile, compression, wear, and creep evaluations under conditions simulating real incisor teeth, the study compares the performance of these two biopolymers. Remarkably, the results demonstrate that PLA showcases superior mechanical properties compared to PA11. This investigation underscores the potential of 3D printing in dentistry, particularly in crafting implants using biocompatible materials, therefore paving the way for more personalized and effective dental treatments [6].

Another case study exemplifies the utilization of Fused Deposition Modelling (FDM) technology in the creation of a breast implant pattern. With dimensions measuring 170 mm × 120 mm × 80 mm, the implant's design and fabrication process were guided by principles of Design for Additive Manufacturing (DfAM). This involved meticulous selection of biomaterials and attention to FDM-specific design parameters, such as geometry optimization and optimal build orientation. The implementation of these strategies ensured the successful production of the breast implant pattern.

This technique of fused deposition modelling (FDM) has found itself useful in developing drug-loaded cardiovascular prostheses, targeting the prevention of infections associated with vascular grafts according to a study. By utilizing FDM technology, medicated vascular grafts are fabricated using thermoplastic polyurethane (TPU) combined with rifampicin (RIF) as a model drug. The incorporation of RIF into TPU filaments is achieved through hot melt extrusion (HME), enabling controlled drug release for up to 80 days and exhibiting antimicrobial properties against Staphylococcus aureus. Additionally, the methodology is extended to include dipyridamole (DIP) as another model drug. This innovative approach demonstrates the potential of FDM in producing personalized vascular grafts with tailored drug delivery capabilities, offering a promising solution to enhance the efficacy and safety of cardiovascular implants [7].

3D printing is applied in bone repair and regeneration in the craniofacial region of humans and animals. Human studies involved 81 patients with craniofacial bone defects, primarily treated with titanium or hydroxyl-apatite scaffolds. Animal studies utilized various biomaterials and cells, offering insights through histological and biochemical findings. It underscores the promising results of 3D printing for craniofacial bone repair, emphasizing the need for further well-designed clinical trials to validate its efficacy. Emerging as a transformative technology, 3D printing holds significant potential for tissue engineering in treating craniofacial bone defects, offering personalized solutions with customized scaffolds and biomaterials [8].

3D printing, combined with the sublimation technique, has been applied to manufacture eyeball prostheses. Data acquisition about the patient's anatomy, medical imaging (both 2D and 3D), and segmentation using CT scanning of craniomaxillofacial structures were done. The virtual model alone was not enough, so with the help of advanced tools (Autodesk Meshmixer) were utilized to produce 3D physical models. This helped medical specialists and scientists together to develop the prototype of the customised oculopalpebral prosthesis. Eye colour, eyebrows, and other anatomical and physical details were matched. Surgical glue was used to fix the prosthesis prototype without any inconvenience to the patient. FDM 3D

ISBN 979-8-3503-5094-4

(PLA) was utilized to match and create the facial features of the patient at relatively low manufacturing costs [9].

Research also explored the use of Selective Laser Sintering (SLS) technology in fabricating oral tablets containing isoniazid, an antitubercular drug, with the incorporation of carbonyl iron as a multifunctional ingredient. Carbonyl iron serves as both a magnetic and heat-conductive component, offering advantages in the manufacturing process. Through SLS 3D printing, tablets with optimized quality attributes are produced by adjusting printing parameters such as laser scanning speed, hatching space, and surface/chamber temperature. The study demonstrated that tablets containing carbonyl iron require lower laser energy input for sintering while exhibiting enhanced drug release under a magnetic field. These findings suggested that magnetic nanoparticles hold promise as conductive materials for facilitating the SLS 3D printing of pharmaceutical dosage forms [10].

A synthetic osteo-regenerative biomaterial called hyper-elastic "bone" (HB) to mend current deficiencies in osteoregenerative products was introduced. HB comprises 90% hydroxyapatite and 10% polycaprolactone or poly (lactic-co-glycolic acid) and can be rapidly printed using 3D printing at high speed (up to 275 cm3/hour) at room temperature. The resulting 3D-printed HB exhibits elastic mechanical properties, and high absorbency, and supports cell viability, proliferation, and osteogenic differentiation of bone marrow–derived human mesenchymal stem cells in vitro. *In vivo*, evaluations in mouse, rat, and non-human primate models demonstrate HB's biocompatibility, vascularization, integration with surrounding tissues, and rapid ossification without the need for additional biological factors. This study highlights the potential of HB as a versatile and effective solution for osteo-regenerative applications [11].

3D printing of controlled release of high-dose pharmaceutical dosage form was done using Selective Laser Sintering. The dosage form contained a test model substance, crystalline paracetamol, and a small amount of dye was also added. After analyzing the pore space, drug release, and dissolution modelling, the printlet was characterized. It revealed that many degrees of freedom were available for tuning into its functional properties, particularly the dissolution performance. Two degrees of freedom were found to be substantial: macrostructure shaping in terms of active surface: volume ratio and microstructure shaping in terms of pore space structure, which aids modification. This approach is advantageous in terms of controlling the dissolution performance of the pharmaceutical dosage [12].

IV. DISCUSSION

In recent years, there has been a significant surge in research and development focused on leveraging biomaterials to revolutionize the treatment of various chronic diseases. From immunomodulation to injectable biomaterials and supramolecular biomaterials, scientists are exploring innovative approaches to address complex health conditions such as type 1 diabetes, bone defects, cancer, and heart attacks. Let's delve into these advancements and their potential implications: (13)

A. Immunomodulation:

One of the most promising areas of research involves the use of immunomodulating biomaterials to fine-tune the immune response, particularly in autoimmune diseases like type 1 diabetes. Recent studies have showcased remarkable progress, with researchers successfully reversing type 1 diabetes in mice using injectable synthetic biomaterials. This breakthrough offers hope for developing effective treatments that could potentially halt or even reverse the progression of autoimmune disorders in humans. However, translating these findings from preclinical models to clinical applications remains a crucial challenge that requires further investigation [13]. Interestingly, Orthopedic biomaterial-associated infections, driven by Staphylococcus aureus biofilms, posed clinical challenges. Using a defective rat bone model, one study observed the immune response to infections like those continuous decrease in T cells and a simultaneous increase in immunosuppressive myeloid-derived suppressor cells (MDSCs). The results show us the complex host-pathogen interaction in orthopedic infections and potential diagnostic and therapeutic areas.

B. Injectable Biomaterials:

The emergence of injectable biomaterials represents a paradigm shift in drug delivery, enabling precise targeting of therapeutic agents while bypassing immune system barriers. Both synthetic and naturally derived injectable biomaterials are being explored across diverse medical domains, including orthopedics, oncology, and cardiology. By delivering therapeutics directly to the site of action, these biomaterials hold immense potential for enhancing treatment outcomes and minimizing adverse effects. However, optimizing the biocompatibility, stability, and controlled release kinetics of injectable biomaterials poses significant technical hurdles that necessitate ongoing research and development efforts. (3) Among injectable biomaterials, injectable hydrogels have made it to the top of the list due to their design flexibility that promotes bone repair. However, hydrogels alone cannot induce molecular signals to induce self-healing. Therefore, a group of researchers recently incorporated the addition of inorganic nanoparticles with the injectable hydrogel. That was proven to increase the osteoconductive properties of bone. Out of the inorganic materials that can be used these researchers used MBGNs (mesoporous bioactive glass nanoparticles) to modify the hydrogel. Further to reckon with they have come up with an injectable, dual-crosslinked hydrogel by self-crosslinking periodate-oxidized dextran (oxDex) and phenylboronic acid-grafted gelatin (PBA-Gel) [14].

C. Supramolecular Biomaterials:

These smart biomaterials offer unprecedented versatility, allowing for dynamic adjustments in response to specific biological signals. By mimicking natural biological processes, such as cell signaling pathways, supramolecular biomaterials hold promise for tailored interventions in injury repair and disease management. Nonetheless, harnessing the full potential of these biomaterials requires interdisciplinary collaboration between materials scientists, bioengineers, and medical researchers to overcome challenges related to stability, reproducibility, and scalability [13].

Supramolecular hydrogels have recently proved to be an amazing biomimetic material. The features of hydrogels are well suited for cell culture, release of drugs with control, tissue adhesion, and molecular sensing. They are used as artificial gel substitutes. (e.g., vitreous humour and synovial fluids), and these properties cannot be achieved by permanently cross-linked covalent hydrogels. As a

ISBN 979-8-3503-5094-4

consequence, supramolecular hydrogels have widely gained popularity in these years, especially in biomedical applications [15].

TABLE 1: RECENT MODIFICATIONS IN APPLIED BIOMATERIALS AT PRESENT

S.NO:	*FEW EXAMPLES*		
	BIOMATERIALS	***RECENT MODIFICATION***	***APPLICATION***
1.	Hyrogels	Albumins's negatively charged surface can bind to a lot of bodily substances such as vitamins, drugs, etc. Albumin-hydrogel formed by gelations induced by heat, Ph, cross-linkings.[16]	These albumin hydrogels are used in specific drug delivery, 3D cell culture, tumour treatment , and tissue engineering. [16]
2.	Dental implants	Dental implants are coated with many nanocoats. E.g.: Zirconia / PEEK coated dental implants. [17]	PEEK- Fabrication of implant, temporary and healing abutments; and healing caps. ZIRCONIA - Less damage, especially in patients susceptible to periodontal disease. [17]
3.	Collagen	Modifying acid - soluble collagen using excessive succinic anhydride in alkaline condition to protect all its active amino groups. Telopeptides of this modified collagen are removed using pepsin to give SPSC (Succinylated Pepsin-Soluble Collagen). It forms a biconjugate (SPSC - PNIPAAm) [18]	Biotherapeutics, cosmetics, applied in In-situ gelling cell delivery scaffold and Kaolin Flocculant. [18]
4.	Implants made of silicon and rubber	Agarose is grafted upon Silicon - Rubber surface to prevent / inhibit the growth of bacterial biofilms on them. [19]	Finds various application in specific tissue engineering, drug delivery, wound dressing, etc, due to its hydrophilicity and anti-biofilm capacity. (Prevents infection issues and its non - toxic) [19]

In conclusion, the ongoing advancements in immunomodulation, injectable biomaterials, and supramolecular biomaterials along with developments to tackle orthopedic biomaterial-associated infections represent groundbreaking strides toward personalized and precision medicine. By harnessing the unique properties of biomaterials, researchers aim to develop innovative therapies capable of addressing the complex challenges associated with chronic diseases.

V. CONCLUSION

In conclusion, the field of biomaterials and 3D printing in orthopedics and beyond has undergone a remarkable evolution, from the early stages of bio-inert materials to the current era of bio-mimetic and smart biomaterials. The advent of 3D printing technologies has revolutionized the landscape of orthopedic treatments, offering customized implants with precise geometries and mechanical properties that promote osseointegration and reduce the risk of complications. Moreover, advancements in biomaterials have led to the development of osteobiologic materials with osteogenic properties, fostering bone healing and improving patient outcomes. The application of biomaterials and 3D printing extends beyond orthopedics, encompassing various medical specialties such as dentistry, cardiovascular medicine, and tissue engineering. These technologies enable the fabrication of patient-specific implants, drug-loaded prostheses, and synthetic biomaterials tailored to address specific clinical needs. From synthetic trabecular bone models to personalized incisor teeth and breast implant patterns, 3D printing offers unparalleled flexibility and precision in manufacturing medical devices and pharmaceutical dosage forms. Furthermore, recent innovations in immunomodulating biomaterials, injectable hydrogels, and supramolecular biomaterials hold promise for revolutionizing the treatment of chronic diseases, including autoimmune disorders and orthopedic biomaterial-associated infections. By fine-tuning the immune response, delivering therapeutics directly to target sites, and sensing and responding to physiological cues, these smart biomaterials offer tailored interventions with enhanced efficacy and minimized adverse effects. However, despite the significant progress achieved thus far, challenges remain in optimizing the biocompatibility, stability, and controlled release kinetics of biomaterials and 3D-printed constructs. Moreover, translating these advancements from preclinical models to clinical applications requires interdisciplinary collaboration, regulatory oversight, and continued investment in research and clinical translation. In summary, the ongoing advancements in biomaterials and 3D printing technologies hold immense potential for revolutionizing patient care across various medical specialties. By harnessing the unique properties of biomaterials and leveraging cutting-edge manufacturing techniques, researchers aim to develop innovative therapies capable of addressing the complex challenges associated with chronic diseases and improving the quality of life for patients worldwide. As we look toward the future, further exploration and refinement of these technologies are essential to unlock their full potential. Continued research efforts, interdisciplinary collaboration, and investment in infrastructure and regulatory frameworks will be crucial in overcoming existing challenges and translating scientific discoveries into tangible clinical benefits. With concerted efforts from researchers, clinicians, industry partners, and regulatory agencies, the field of biomaterials and 3D printing holds the promise of transforming healthcare and

ISBN 979-8-3503-5094-4

advancing personalized medicine for the betterment of society as a whole.

REFERENCES

[1] Yubo Tao, Ling Pan, Dexi Liu, and Peng Li, "A case study: Mechanical modeling optimization of a cellular structure fabricated using wood flour-filled polylactic acid composites with fused deposition modeling", Composite Structures, Volume 216, 15 May 2019, pp . 360-365, May 2019

[2] Hana Beyene Mamo, Marcin Adamiak, and Anil Kunwar "3D printed biomedical devices and their applications: A review on state-of-the-art technologies, existing challenges, and future perspectives" Materials Science and Engineering, Volume 106, January 2020

[3] Sandeep Kumar, Monika Nehra, Deepak Kedia, Neeraj Dilbaghi, K.Tankeshwar, and Ki-Hyun Kim, "Nanotechnology-based biomaterials for orthopedic applications: Recent advances and prospects", September 2019

[4] Saeideh Kholgh Eshkalak, Erfan Rezvani Ghomi, Yunqian Dai, Deepak Choudhury, and Seeram Ramakrishna, "The role of three-dimensional printing in healthcare and medicine", 2020

[5] Dan Wu, Andrea Spanou, Anna Diez-Escudero, and Cecilia Persson, "3D-printed PLA/HA composite structures as synthetic trabecular bone: A feasibility study using fused deposition modeling", Journal of the Mechanical Behavior of Biomedical Materials, Volume 103, March 2020

[6] M. Arun, N. Sathishkumar, K. Nithesh Kumar, S.S. Ajai, and S. Aswin, "Development of patient-specific bio-polymer incisor teeth by 3D printing process: A case study", Materials today: proceedings, Volume 39, Part 4, pp. 1303-1308, 2021

[7] Hugo I. Medellin-Castillo and Jorge Zaragoza-Siqueiros, "Design and Manufacturing Strategies for Fused Deposition Modelling in Additive Manufacturing: A Review", Chinese journal of mechanical engineering, Article no.53, 2019

[8] Niamh K. Martin, Juan Domínguez-Robles, Sarah A. Stewart, Victoria A. Cornelius, Qonita Kurnia Anjani, Emilia Utomo, Inmaculada García-Romero, Ryan F. Donnelly, Andriana Margariti, Dimitrios A. Lamprou a and Eneko Larrañeta, "Fused deposition modeling for the development of drug-loaded cardiovascular prosthesis", International Journal of Pharmaceutics, Vol 595, February 2021

[9] E. Vázquez-Silva, D.D. Bohorquez-Vivas, P.G. Peña-Tapia, F.P. Moncayo-Matute, P.B. Torres-Jara and D.P.Moya-Loaiza, "Oculopalpebral prosthesis prototype design using the additive manufacturing technique: A case study", JPRAS Open Volume 39, Pages 228-236, March 2024

[10] Yu Zhang, Rishi Thakkar, JiaXiang Zhang, AnQi Lu, Ishaan Duggal, Amit Pillai, Jia-Wei Wang, Niloofar Heshmati Aghda, Mohammed Maniruzzaman " Investigating the Use of Magnetic Nanoparticles Alternative Sintering Agents in Selective Laser Sintering (SLS) 3D Printing of Oral Tablets", ACS publications, 2023

[11] Adam E. Jakus, Alexandra L. Rutz, Sumanas W. Jordan, Abhishek Kannan, Sean M. Mitchell, Chawon Yun, Katie D. Koube, Sung C. Yoo, Herbert E. Whiteley and Ramille N. Shah, "Hyperelastic "bone: A highly versatile, growth factor–free, osteoregenerative, scalable, and surgically friendly biomaterial", Science Translational Medicine, Volume 8 pp.358, 2016

[12] MichaelMaroulakos, George Kamperos, Lobat Tayebi, Demetrios H alazonetis and Yijin Ren, "Applications of 3D printing on craniofacial bone repair: A systematic review", Journal of Density, Volume 80, pp.1-14, January 2019

[13] Jorge Barrios-Muriel, Francisco Romero-Sánchez, Francisco Javier Alonso-Sánchez, David Rodríguez Salgado, "Advances in Orthotic and Prosthetic Manufacturing: A Technology Review", Materials, Volume 3, Issue 2, January 2020

[14] lXin Zhang, Kai Nan, Yuankai Zhang, Keke Song, Zilong Geng, Donglong Shang, Xin Guan and Lihong Fan, "A novel injectable hydrogel prepared from phenylboronic acid-modified gelatin and oxidized-dextran for bone tissue engineering", International Journal of Biological Macromolecules, Volume 261, Part 1

[15] J.Y.C. Lim, Q. Lin, K. Xue and X.J. Loh "Recent advances in supramolecular hydrogels for biomedical applications", Materials Today: Proceedings, Volume 3, September 2019

[16] Fanhui Kong, Nabila Mehwish, Bae Hoon Lee "Emerging albumin hydrogels as personalized biomaterials" Acta Biomaterialia Volume 157, pp 67-90, February 2023

[17] A.Heboyan, F. Bennardo "New biomaterials for modern dentistry." BMC Oral Health 23, 817, 2023

[18] Ke Li, Delai Kang, Yu Li, Weizhe Zhu, Lian Zhu, Juntao Zhang, Che ngzhi Xu, Benmei Wei, Haibo Wang "A fluorescent sensing platform based on collagen peptides-protected Au/Ag nanoclusters and WS2 for determining collagen triple helix integrity" Analytica Chimica Acta Volume 1247, 22 March 2023

[19] Wenting Chu, Yuhong Ma, Yuning Zhang, Xinjie Cao, Zhongyu Shi, Ying Liu, Xuejia Ding Significantly improved antifouling capability of silicone rubber surfaces by covalently bonded acrylated agarose towards biomedical applications Colloids and Surfaces B: Biointerfaces Volume 222, February 2023.

ISBN 979-8-3503-5094-4

Chitosan Based Biopatches Impregnated with Amaranthus Spinosus Herbal Drug for Wound Healing Application

Swathy M
Department of Biomedical Engineering
KPR Institute of Engineering and Technology
Coimbatore,India
swathy25591@gmail.com

Ramya E*
Department of Biomedical Engineering
KPR Institute of Engineering and Technology
Coimbatore,India
easura1643@gmail.com

Shruthi M G
Department of Biomedical Engineering
KPR Institute of Engineering and Technology
Coimbatore,India
mgshruthi.sm@gmail.com

Lokeshan R
Department of Biomedical Engineering
KPR Institute of Engineering and Technology
Coimbatore,India
ramlokesh256@gmail.com

Abinesh S
Department of Biomedical Engineering
KPR Institute of Engineering and Technology
Coimbatore,India
sanjuabi7038@gmail.com

***Abstract*— Wound healing is a process in which the body repairs and restores damaged or injured tissue to its normal function and structure. This complex and dynamic process involves several stages and cellular events. Supporting the natural wound healing process and promoting optimal healing requires providing proper wound care, including cleaning the wound, keeping it protected and moist, managing underlying health conditions, and addressing any complications that may arise. Therefore it is essential to develop bio patches with necessary wound healing properties [1], [2]. This article highlights chitosan's properties, making it an ideal choice for wound healing. It presents a comprehensive review of Chitosan-based bio patches impregnated with herbal drugs as a novel approach in wound care. Chitosan, a biopolymer derived from chitin, possesses biocompatibility, biodegradability, and antimicrobial properties [15]. Herbal drugs, derived from various species, are known for their wound healing properties, including anti-inflammatory, antimicrobial, and antioxidant effects [3], [11]. Silver nitrate nanoparticles, with their antimicrobial properties and ability to stimulate cell proliferation, have shown promise in wound healing [9]. Incorporating these nanoparticles into wound dressings and coatings can promote faster wound closure and reduce the risk of infections. A well-designed bio patch, by promoting tissue regeneration and reducing inflammation, can help minimize the formation of excessive scar tissue and reduce pain associated with the wound site. The process protein extraction essential for fabrication of such biopatch is shown in figure 1. It creates a barrier that protects the wound from external contaminants and promotes tissue regeneration, providing a conducive environment for wound healing.**

Keywords: Wound healing, Chitosan, bio patch, tissue regeneration.

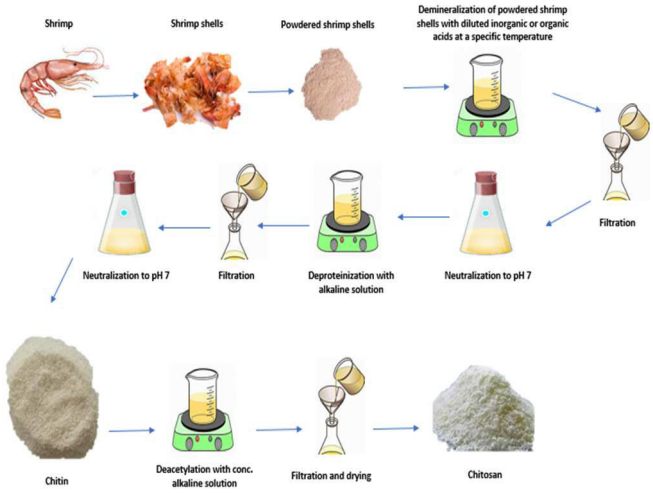


Figure. 1. Extraction of biomaterial for preparation of wound healing biopatch

I. INTRODUCTION

Wound healing is a complex and dynamic biological process vital for the restoration of damaged tissues and the maintenance of overall health. Over the years, significant advancements in medical science and materials engineering have led to the development of innovative wound healing materials aimed at enhancing and expediting this natural healing process. This review aims to provide a comprehensive overview of key materials, their mechanisms of action, and their real-world effectiveness in promoting wound closure, minimizing scarring, and reducing the risk of infections [5]. In recent years, there has been growing interest in harnessing the potential of natural biomaterials in the quest for effective wound healing solutions. Among these biomaterials, chitosan, a biopolymer derived from chitin, and Amaranthus, a group of versatile and nutrient-rich plants, have emerged as promising candidates for wound healing applications [7]. This comprehensive review delves into the fascinating world of wound healing materials prepared by combining chitosan and Amaranthus-derived components. Through a detailed exploration of their properties,

ISBN 979-8-3503-5094-4

mechanisms of action, and recent advancements in their development, this review aims to provide a comprehensive understanding of how these synergistic natural materials can accelerate wound closure, reduce inflammation, and enhance tissue regeneration [10].

Wound healing is a complex and highly regulated biological process by which the body repairs a damaged tissue. It involves a series of coordinated events that aim to restore the structural and functional integrity of the injured area. The process of wound healing can be broadly categorized into four overlapping phases such as haemostasis, inflammation, proliferation, remodelling. Where haemostasis is the initial response to injury, where blood vessels constrict to minimize blood loss. Platelets aggregate at the site to form a clot, sealing the wound temporarily. Clotting factors in the blood promote clot formation. In the inflammation phase inflammatory cells, such as neutrophils and macrophages, are recruited to the wound site. These cells help clear debris, foreign materials, and potential pathogens from the wound. Inflammation also helps create an environment conducive to tissue repair. In the proliferation phase, fibroblasts migrate to the wound site and begin to produce collagen, a protein essential for tissue strength. New blood vessels (angiogenesis) form to provide oxygen and nutrients to the healing tissue. Epithelial cells at the wound edges proliferate and migrate to cover the wound surface. The final phase of wound healing is remodelling which is also known as maturation phase. In this phase the collagen fibers in the wound undergo reorganization.

Wound healing biomaterials are substances or materials designed and used to support the natural wound healing process, accelerate tissue repair, and improve the overall outcome of wound treatment. These materials are often employed in various medical applications, including wound dressings, tissue engineering, and regenerative medicine. Wound dressings are perhaps the most widely used wound healing biomaterials. They can be made from a variety of materials, including gauze, foam, hydrocolloids, hydrogels, and films. These dressings provide a protective barrier, maintain a moist environment, and facilitate the removal of exudate (fluid) from the wound, which is important for healing. Some of the natural biomaterials used in wound healing are Chitosan, collagen, alginate, chitin, starch, silk, agar etc. which mimic the extracellular matrix; synthetic polymers such as polyurethane and Polyglycolic acid for dressings and scaffolds; hydrogels that maintain a moist wound environment; nanomaterials like nanoparticles and nanofibers extracted through green synthesis for controlled drug delivery [14]; biological scaffolds derived from decellularized tissues; antimicrobial biomaterials like silver-infused dressings; hydrogels [13], growth factor delivery systems; and artificial skin substitutes. The selection of a particular biomaterial depends on factors like wound type, size, and desired therapeutic outcome, with ongoing research aiming to improve wound care and tissue regeneration through innovative biomaterial applications. The contribution of a biopatch in effective wound healing is shown in the Figure 2.

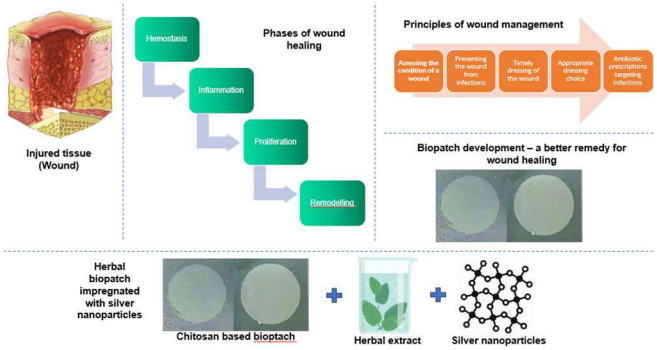


Figure. 2. Wound healing Phases and contribution of a biopatch

II. METHODOLOGY

A. Extraction of chitosan from shrimpshells

Shrimp shells (20 g) undergo several washes with distilled water and drying at 40 degrees Celsius for 48 hours. 5 g of the dried shrimp shells is powdered, immersed in 50 mL of 2 M Hcl solution, and heated at 70 degrees Celsius for 24 hours to eliminate the impurities and other minerals present. Proteins that were present was then removed by heating in 10 mL of 2 M NaOH solution. The resulting filtrate is washed with acetone and distilled water before drying at 60 degrees Celsius for 48 hours, producing 5 g of dark brown chitin. Subsequently, 5 g of the extracted chitin is deacetylated by heating for 4 hours at 150 degrees Celsius with 10 mL of 50% NaOH aqueous solution. Chitosan is obtained after drying at 40 degrees Celsius for 48 hours (1.18g). The chitosan preparation procedure is shown in Figure 3.

Figure. 3. Chitosan extraction from shrimp shells

B. Extraction of herbal essence from amaranthus spinosus

Fresh leaves of Amaranthus spinosus are collected, cleaned, dried, ground into a fine powder, and boiled with distilled water[3]. The extract is obtained through a standard filtration method.

The extraction of amaranthus was done through processes including maceration, Soxhlet extraction, filtration, and purification extraction. Maceration is the process of soaking the plant material in a solvent for an extended period to dissolve desired compounds, while Soxhlet extraction continuously circulates solvent through the material. Filtration separates solid particles from the liquid extract. Following these methods, purification extraction techniques like chromatography or crystallisation can further refine the extracted compounds, isolating specific elements based on

ISBN 979-8-3503-5094-4

their properties [11]. These varied methods offer opportunities to obtain extracts with different purities and concentrations, enabling the extraction of various beneficial compounds present in Amaranthus which is shown in Figure 4 for wound healing applications.

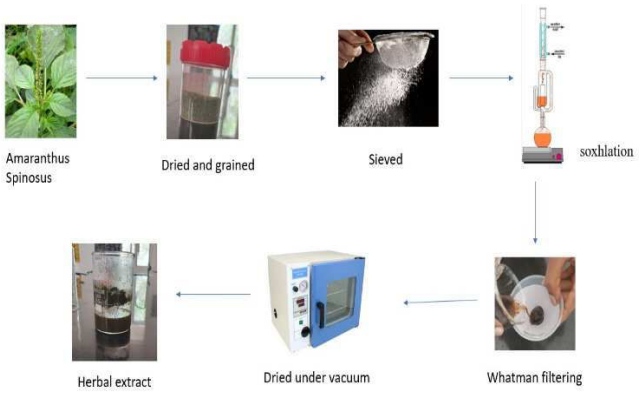


Figure. 4. Herbal extract for bio patch preparation

C. Fabrication of patch using simple solvent casting technique

The biopolymers based bio patch with chitosan and gelatin were weighted and dissolved separately in distilled water to obtain a fine aqueous concentration. Gelatin is a protein obtained by the partial hydrolysis of collagen, which is derived from animal connective tissues such as bones and skin. It is a commonly used ingredient in the preparation of bio patches due to its biocompatibility, biodegradability, and ability to form a stable gel. Glycerol is compatible with a wide range of other wound healing agents, which can improve their efficacy and delivery to the wound site. It is used for cross linking the materials that are used in preparation of bio patch. The bio patch were fabricated by a solvent casting technique [12]. Solvent casting is a common method used to create wound healing patches or films which is shown in figure 5. In this method, a polymer or a combination of polymers is dissolved in a suitable solvent, such as ethanol or chloroform, to form a solution. The solution is then cast onto a substrate, such as a glass plate or a silicon wafer, and allowed to dry. Once the solvent evaporates, a thin film or patch of the polymer is left behind. This patch can be cut into the desired shape and size and then applied to the wound. The patch slowly dissolves in the presence of body fluids, releasing any drugs or growth factors incorporated into the patch as shown in the Figure 5.

Figure. 5. Biopatch Preparation

III. RESULT

X-ray diffraction

XRD can be used to identify the presence of different phases or crystalline structures in a bio patch. Bio patches may contain various materials, such as polymers, ceramics, or minerals that can exhibit different phases or crystal structures depending on their composition and processing. It provides information on the molecular orientation or alignment of materials in a bio patch. Many bio patches are composed of anisotropic materials, such as aligned fibres or crystals, which can exhibit different properties depending on their orientation. It can be used to analyse the molecular orientation of materials in a bio patch, which can be relevant for understanding its mechanical, electrical, or biological properties. It helps to identify the specific phases present in the bio patch, which can provide insights into its composition, stability, and potential bioactive properties. XRD analysis helps to optimize the material properties for skin regenerative applications.

The XRD analysis of chitosan was performed using a D 8 Advance X-Ray Diffractometer with 2 Theta angular range. The resulting diffractogram is shown in Figure 6.

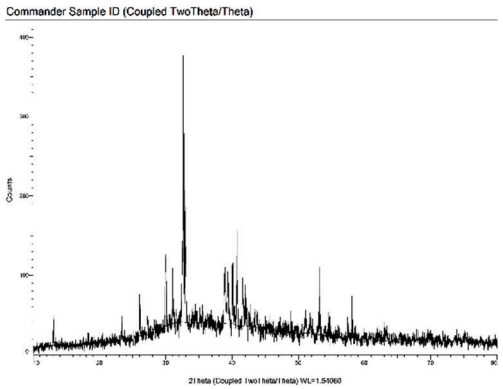


Figure.6. XRD of Chitosan

The diffractogram of chitosan exhibited distinct diffraction peaks at 2 Theta values if 19 degree and 30 degree indicative of its crystalline nature. The positions and intensities of these peaks are nearly consistent with previous reports in the literature for chitosan Synthesis and characterization of chitosan–silver nanocomposite. The crystalline peaks were observed at 2θ values of 28 degree and 32 degree and can be attributed to the characteristic crystal lattice planes of chitosan, which are consistent with its known crystal structure.

IV CONCLUSION

The exploration of chitosan-based bio patches infused with Amaranthus spinosus for wound healing applications represents a significant stride in material characterization. This innovative approach not only showcases the intrinsic properties of the bio patch but also highlights the motivation behind its development. The amalgamation of Chitosan, a biocompatible and biodegradable polymer, with the therapeutic properties of Amaranthus spinosus presents a

ISBN 979-8-3503-5094-4

multifaceted solution for addressing crucial aspects of wound healing.

The incorporation of Amaranthus spinosus enriches the bio patch with enhanced antimicrobial features, anti-inflammatory effects and the capacity to stimulate tissue regeneration. Leveraging the inherent benefits of chitosan, such as biocompatibility, biodegradability and facile fabrication, further amplifies the potential of this bio patch as an advanced wound care solution.

To validate the efficacy and safety of this groundbreaking approach, additional research and clinical studies are imperative. These endeavors will not only substantiate the claims of improved wound healing outcomes but also contribute to the translation of this chitosan-based bio patch with Amaranthus spinosus into practical clinical applications. The completion of protein and herbal essence extractions, coupled with meticulous variations in the ratios of chitosan, gelatin, and Amaranthus spinosus extract during bio patch preparation, underscores the comprehensive nature of this study.

As the bio patch has been meticulously crafted through solvent casting and particulate leaching methods, it sets the stage for potential future applications in advanced wound care. The culmination of this research signifies not just the completion of material characterization but also paves the way for a promising future in the realm of wound healing. The groundwork laid by this study serves as a springboard for continued exploration, offering a glimpse into the exciting prospects and possibilities that lie ahead in the field of biomaterials for wound care.

REFERENCES

[1] Anandhavelu, S., Doble, M., Jeon, J.H., Kim, H.S., Perumal, G., Sellappan, L.K. and Vikraman, D. (2022) 'Biopolymer film fabrication for skin mimetic tissue regenerative wound dressing applications', International Journal of Polymeric Materials and Polymeric Biomaterials, Vol.71, No.3, pp.196-207.

[2] Banerjee, R., Jadhav, K., Jadhav, S., Osmani, R.A. and Singh, E. (2021) 'Biopolymers and biocomposites: Nature's tools for wound healing and tissue engineering', Woodhead Publishing in Materials, pp. 573-630.

[3] Baral, M.A., Chakraborty, S. and Chakraborty, P. (2010) 'Evaluation of anthelmintic and anti-inflammatory activity of Amaranthus spinosus Linn', International Journal of current pharmaceutical research, Vol.2, No.4, pp.44-7.

[4] Biswas, S., Islam, M.M., Rashid, T.U., Sakib, M.N. and Shahruzzaman, M. (2020) 'Chitosan based bioactive materials in tissue engineering applications-A review', Bioactive materials, Vol.5, No.1, pp.164-183.

[5] Chiesa, E., Conti, B., Dorati, R., Genta, I., Pisani, S. and Tottoli, E.M. (2020) 'Skin Wound Healing Process and New Emerging Technologies for Skin Wound Care and Regeneration', Pharmaceutics, Vol.12, No.8, pp.735.

[6] Costa, E.F., Di Stasi, L.C. and Magalhães, W.V. (2022) 'Recent Advances in Herbal-Derived Products with Skin Anti-Aging Properties and Cosmetic Applications', Molecules, Vol.27, No.21, pp.7518.

[7] Sanmugam, A., Abbishek, S., Kumar, S. L., Sairam, A. B., Palem, V. V., Kumar, R. S., and Vikraman, D. (2023). Synthesis of chitosan based reduced graphene oxide-CeO2 nanocomposites for drug delivery and antibacterial applications. Journal of the Mechanical Behavior of Biomedical Materials, 145, 106033.

[8] Paswan, S.K., Rao, C.V. and Srivastava, S. (2020) 'Wound healing, antimicrobial and antioxidant efficacy of Amaranthus spinosus ethanolic extract on rats', Biocatalysis and Agricultural Biotechnology. Vol.26, pg.101624.

[9] Chellamani, K.P., Sundaramoorthy, P. and Sureshram, T.(2012) 'Characterisation of Poly vinyl alcohol (PVA)/Silver nitrate nanomembranes for their suitability in wound dressing applications', International Journal of Emerging Technology and Advanced Engineering.Vol.2,No.11, pp.176-184.

[10] Sanmugam, A., Sellappan, L. K., Manoharan, S., Rameshkumar, A., Kumar, R. S., Almansour, A. I., and Vikraman, D. (2024). Development of chitosan-based cerium and titanium oxide loaded polycaprolactone for cutaneous wound healing and antibacterial applications. International Journal of Biological Macromolecules, 256, 128458.

[11] Ravindran, J., Arumugasamy, V. and Baskaran, A.(2019) 'Wound healing effect of silver nanoparticles from Tridax procumbens leaf extracts on Pangasius hypophthalmus', Wound Medicine.Vol.27,No.1, pg.100170.

[12] Sellappan, L. K., and Manoharan, S. (2024). Fabrication of bioinspired keratin/sodium alginate based biopolymeric mat loaded with herbal drug and green synthesized zinc oxide nanoparticles as a dual drug antimicrobial wound dressing. International Journal of Biological Macromolecules, 259, 129162.

[13] Güiza-Argüello, V.R., Solarte-David, V.A., Pinzón-Mora, A.V., Ávila-Quiroga, J.E. and Becerra-Bayona, S.M.(2022) 'Current advances in the development of hydrogel-based wound dressings for diabetic foot ulcer treatment',Polymers.Vol.14,No.14, pg.2764.

[14] Gayathri, T., Kumar, S. L., Sangavi, S., Yudhika, M., and Swathy, M. (2023). Green synthesis of copper oxide nanoparticles using Carica papaya and their antimicrobial activity. Materials Today: Proceedings.

[15] Que, A.M., Tang, T.N., Thanh, T.P.T., Duong, T.T., Vu, T.B., Doan, H.N. and Nguyen, H.T.(2023), 'Chitosan oligomer/silver nanoparticles/PCL composite electrospun membrane for severe burn wound healing', Vietnam Journal of Science, Technology and Engineering.Vol.65,No.4, pp.62-70.

ISBN 979-8-3503-5094-4

Tenth International Conference on Biosignals, Images and Instrumentation, SSNCE, Chennai, March 20 – 22, 2024

TRACK 11

Beta waves monitoring for the prognosis of seizures in epileptic pediatric patients

Janya Jagan
Cypress High School
Cypress, USA
0000-0002-3968-8944

Abstract—**Epilepsy is a neurological disorder that causes abnormal recurrent seizures with sudden bursts of the brain's electrical activity. Loss of consciousness, injury, or memory impairment are possibly such seizure's serious consequences. Therefore, predicting seizures before they occur can help patients improve their quality of life. One possible way to predict seizures is to monitor the patient's brain waves using electroencephalography (EEG), recording the brain's electrical signals. Studies have suggested that before or during a seizure, brain waves, such as beta waves, may change in frequency or amplitude. Here I show that in pediatric patients with epilepsy, beta waves increase significantly before and during seizures compared to their baseline levels. I analyzed the EEG recordings of 35 seizure sets. I measured the beta wave power during the 30-second interval before and during the seizure. I found that the beta wave power increased by an average of more than 400% in the 30-second interval before the seizure onset, compared to the baseline. My results indicate that beta waves can be a reliable source of truth for pediatric patients' epilepsy seizure prediction and suggest that beta waves may play a pivotal role in the seizures' neural processes. This study contributes to the field of seizure prediction, a challenging yet important task for epilepsy research. The simple, non-invasive method of EEG analysis can be an extremely useful tool for such prediction for pediatric patients with epilepsy. My findings also warrant further investigation to understand the seizure generation and propagation mechanisms in the brain.**

Keywords—epilepsy, beta waves, seizure prediction, prognosis, electroencephalogram.

I. INTRODUCTION

Epilepsy is a neurological disorder characterized by recurrent episodes of dysfunctional high neural activity, known as seizures [1]. A seizure is defined as a sudden uncontrolled electrical activity between the nerves and neurons that results in physical restraints and disruptions of stiffness, numbness, twitching, and weakness [1-3]. The muscle movement becomes limited in the body and creates a struggle in approaching essential tasks. The jerky movement and convulsions are known to be brief and extended and occur as often as multiple times a day, to as rare as minimal times a year [2-3]. Epilepsy is one of the most common disorders, known to affect 1 in 26 people, and in today's world, around 50 million people are struggling with epilepsy [1, 4]. Specifically, within the pediatric population, over 470,000 adolescents are struggling with epilepsy [3]. Approximately 70% of the world's population with epilepsy, has been confirmed to live a healthy and epilepsy-free life when given the opportunity to be diagnosed and treated to completion [4, 5]. Epilepsy cannot be reliably detected using clinical seizure monitoring. Additional work is required in understanding how wearables can be used for detection of epilepsy. Such wearable devices hold success rates in coded algorithms which were able to detect over nine seizure types in 94 patients, proving the potential of such seizure detection devices. However, primarily in experimental trials, the devices are not heavily utilized publicly [18].

An electroencephalogram (EEG) is a technique using the electrode array that tests whether there are brain abnormalities through surface/skin electrical activity or brain waves by measuring brain waves (plot of the waves, a recording, and more) [6]. The EEG array contains small metal disks of electrodes that are attached to wires and firmly attached and implanted into the head [6]. The electrodes read for small charges of electrical activity which are then recorded as waves on a larger device or produced on paper for interpreting by the physician [6]. During the EEG, the patient's brain activity will be recorded, and is considered a painless process [7]. The EEG tests take place under the supervision of a clinical neurophysiologist, and such [7]. The EEG can be used to identify seizures, memory issues, or any other brain/head injuries/conditions [7]. Specifically to epilepsy, the physician can assist in finding out the type of epilepsy through the frequency and evidence from the EEG and accordingly proceed with a method of treatment [7]. Within an EEG, there are various waves such as alpha, theta, delta, and beta [8]. For this study, the beta waves were the focus as the frequency of the beta wave has been shown to increase in an EEG when the patient is experiencing symptoms of arousal, stimuli, anxiety, stress, and many more [9, 10, 11]. Beta waves have an average ranged frequency of 13-30 Hz [8, 11].

The current standard of care for epilepsy patients has been seizure medicines, epilepsy surgeries, epilepsy centers, dietary therapies, and various dietary therapies [12]. From anti-seizure medications to dietary therapies, it can be understood that these treatments are considered to be more reactive than proactive. The medication and treatment do not create an alert or warning system to ensure the patient is aware of when they will be having a seizure. Most commonly, the medication and treatment do not aid in ensuring the patient in a public setting and such is made aware of the oncoming seizure and may take proactive measures to ensure the seizure does not react. It has been brought to awareness that solid odors such as garlic or roses can be used as a preventative measure to halt an oncoming seizure [13]. Additionally, extra doses of medication prescribed by the physician may be of assistance in ensuring that the seizure can be subdued [13]. With this being a present-day issue, having a proactive measure of an alert or warning system may be used to ensure that the patient is aware of an oncoming seizure and will be able to immediately use preventative measures to prevent the oncoming seizure through a mechanism called seizure prediction.

This study shows a method of seizure prediction with the usage of EEG recordings to see if beta waves significantly increase when an oncoming seizure is about to occur for pediatric patients. A prominent spike in the beta waves can provide evidence as to the beta waves being an indicator of a possible seizure. The objective of this prognosis is to innovate a warning or detection system for such spikes in amplitude.

ISBN 979-8-3503-5094-4

II. METHODS

A. Database Background

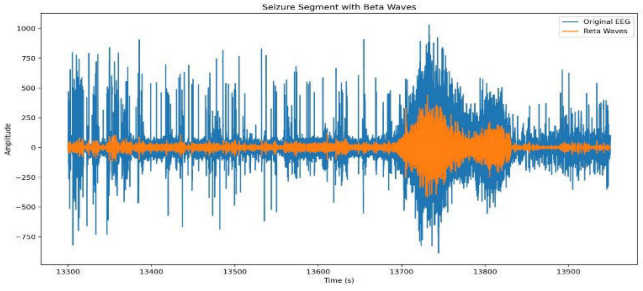
From the Children's Hospital Boston, EEG recordings for pediatric patients with intractable seizures were recorded to monitor their seizures and wave frequencies. The data is available from the online database called CHB-MIT Scalp EEG Database. There were 22 patients in total and various recordings of these patients were monitored and filtered/extracted. There were 5 male and 17 female patients in this study and under each patient 9 to 42 recordings were stored due to hardware limitations resulting in gaps between the data. These subjects were recorded and observed for multiple days and were in the withdrawal stage of their anti-seizure medication. Throughout the timeframe of this research study, 182 seizures were monitored and annotated. All signals were sampled at 256 samples per second with 16-bit resolution. The study used the International 10-20 system of EEG electrode positions, and most files contain 23 EEG signals (24 or 26 in a few cases). For that specific study, the patients were monitored to understand the seizure symptoms as well as categorize them in order to determine whether surgical intervention was required [14].

However, in the case of this study, bandpass filter was applied to isolate the beta frequency range, from all applicable channels available. 33 seizure data from the EEG recordings is utilized, to understand whether the beta waves demonstrate high anxiety, stress, arousal, and stimuli, and if such data can in turn be understood as a proactive signal about an upcoming seizure. The data is analyzed within a 30-second interval before and throughout the full duration of the seizure, to hypothesize and understand if the patient will gradually have an increase in the beta waves from the symptoms that were mentioned. "After 30 seconds" results were not taken once the seizure ended, as the beta wave results would continue to be high as a result of the increase of emotions clouding the patient providing inaccurate data.

III. RESULTS

The majority of results demonstrate the hypothesized result of beta waves increasing over the 30-second interval and are shown in the graphs and tables below for pediatric patients. "Before the Seizure" and "During the Seizure" data is collected and displayed in a way of comparing the increase of beta waves between these areas. "After the Seizure" data was collected, but not used because it would not have any use in seizure prediction.

A. Data Analysis and Values Calculated



In **Fig.1.** (above), file, chb07_19.edf, based on the beta wave frequency range (typically ranging from 13-30 Hz), the beta frequency range was isolated.

Then to do a visual analysis, the chart was plotted with the Amplitude in the Y axis and time in the X axis that had time 200 seconds post-seizure offset. The blue waves represent the ranges from at least 200 seconds before the seizure onset and original data of the EEG showing all the waves. The orange represents the Beta Waves that are demonstrated throughout that portion of the recording. Between the second's time frame of 13688 to 13831, the seizure has occurred. The significant spike of the waves in that region shows how the beta waves significantly increase within the range of the seizure. The .edf file data was extracted and categorized as segments based on the seizure start and end time provided in the database contents.

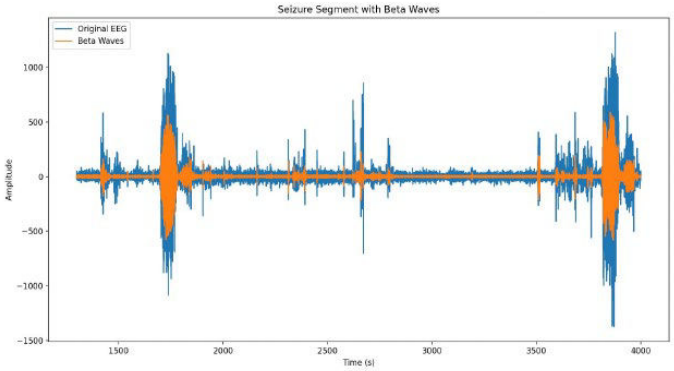


Fig. 2. represents a Seizure Segment with Beta Waves for chb04_28.edf.

Like Figure 1, this is another data set with an X-axis as time in seconds and the measured amplitude for the y-axis. It can be seen how the timeframes of Seizure 1(1679 seconds to 1781 seconds) contain a seizure with high spikes. Similarly, Seizure 2 had timeframes of 3782 to 3898 seconds, which is shown in Figure 2 above with a high amplitude wave.

Before the seizure occurs, beta waves are hypothesized and demonstrated in the results to be closer to 0 in amplitude. As the wavelength continues to waver above and below 0, this proves that the patient has a low beta wave frequency. Originally the data is analyzed to find the average of the amplitude before the seizure occurs. This value would have demonstrated the comparison of when the beta waves were low and high. However, because the waves continue to fluctuate before and during the seizure the average of the amplitudes ended up being 0. Such a calculated value would not give any meaning or comparison of the beta waves.

Following this, data was collected to find the average positive amplitudes of the beta waves before and during the seizure. Then, the difference between both the values is calculated to give evidence for how there is a significant increase in beta wave amplitudes from a normal timeframe to the timeframe with a seizure. This process can be done with only negative amplitudes too.

TABLE I. – Fig.2. analysis of chb04_28.edf, but for all 35 seizure sets. Percentage of difference of average positive amplitude of beta waves during and before seizures, over beta waves' avg positive amplitude before seizure.

File Name	Seizure Start Time	Seizure End Time	During Seizure	Before Seizure	Difference	Percentage
chb07_19	13688	13831	5311.890591	975.3814957	4336.509095	444.60%
chb07_12	4920	5006	2695.723824	1287.171357	1408.552467	109.43%
chb07_13	3285	3381	5463.240931	1298.17288	4165.068051	320.84%
chb01_03	2996	3036	1581.284772	658.863818	922.4209539	140.00%
chb01_04	1467	1494	1100.375423	622.6811707	477.6942527	76.72%
chb01_15	1732	1772	539.8632533	366.0445132	173.8187401	47.49%
chb01_16	1015	1066	1704.07651	702.9746963	1001.101813	142.41%
chb01_18	1720	1810	1106.618507	354.0792786	752.5392279	212.53%
chb01_21	327	420	1497.75982	505.8002592	991.9595609	196.12%
chb01_26	1862	1963	1687.324911	653.1016178	1034.223293	158.36%

ISBN 979-8-3503-5094-4

chb02_16	130	212	1322.728562	597.3781683	725.3503941	121.42%
chb03_01	362	414	2650.736285	467.6625779	2183.073707	466.81%
chb03_02	731	796	2511.813976	516.4303048	1995.383671	386.38%
chb03_03	432	501	2251.941108	919.7733335	1332.167775	144.84%
chb03_04	2162	2214	1774.842159	1185.560802	589.2813566	49.70%
chb03_34	1982	2029	2755.174674	1172.218523	1582.956151	135.04%
chb03_35	2592	2656	2364.760605	568.1501522	1796.610453	316.22%
chb03_36	1725	1778	2448.01144	228.1132706	2219.89817	973.16%
chb04_08	6446	6557	1153.390988	383.5137177	769.8772708	200.74%
chb04_28	1679	1781	7687.818855	237.9718602	7449.846995	3130.56%
chb04_28	3782	3898	7366.462918	792.5042744	6573.958644	829.52%
chb05_06	417	532	4854.861423	1129.386471	3725.474952	329.87%
chb05_13	1086	1196	5539.324938	650.0131701	4889.311768	752.19%
chb05_16	2317	2413	6855.924153	904.5361309	5951.388022	657.95%
chb05_17	2451	2571	6223.568295	559.9202746	5663.648021	1011.51%
chb05_22	2348	2465	6458.191918	1147.251338	5310.940581	462.93%
chb08_02	2670	2841	1320.655857	1208.274206	112.3816515	9.30%
chb08_05	2856	3046	1338.991084	948.3077043	390.6833794	41.20%
chb08_11	2988	3122	1833.672116	1467.098173	366.5739439	24.99%
chb08_13	2417	2577	2012.123981	1026.885411	985.2385705	95.94%
chb08_21	2083	2347	1547.245719	1399.443729	147.8019901	10.56%
chb09_06	12231	12295	7004.617133	1949.625787	5054.991346	259.28%
chb09_08	2951	3030	6967.155714	995.519993	5971.635721	599.85%
chb09_08	9196	9267	6964.355393	1255.462555	5708.892838	454.72%
chb09_19	5299	5361	8694.053035	946.696218	7747.356817	818.36%

The difference between the average of positive amplitudes before and during seizures is measured and provided. This demonstrates the average of positive amplitudes “before seizure” and the average of positive amplitudes “during seizure”. Then the difference was calculated between these values to show how big of a variation there is in beta waves between these portions in the section titled “Difference”. Finally, the percentage is shown as a comparison of the difference between the “During” period and the “Before” period, to demonstrate the significance.

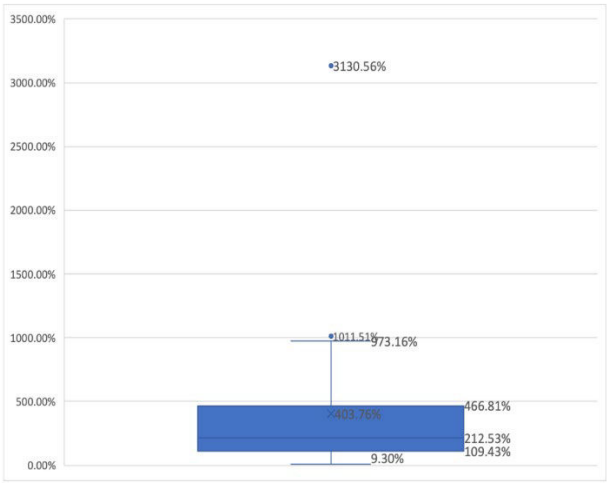


Fig. 3. As can be seen from the box plot, the minimum value is 9.3% and the maximum value is 973.16% with a couple of outliers. The mean percentage increase is 403.76% and the median is 212.53%.

As seen from **Table 1**, the ‘Difference’ represents the increase in the amplitude of the Beta waves during the seizure time compared to that of the 30 seconds of normal time prior to the seizure. Primarily, the value represents that the increase in the amplitude of the Beta waves can be considered as a signal to notify about the upcoming seizure. This is a key finding that can be considered a stepping stone for further research in this area.

IV. DISCUSSIONS

The results of this analysis and study support the hypothesis that the power of the beta waves increases before and during seizures in pediatric patients with epilepsy. This is consistent with other earlier studies about increased beta activity in the pre-ictal and ictal periods of epileptic seizures [15, 16]. Beta waves are considered to reflect cortical and subcortical network activation that involves seizure generation and propagation [15]. The Beta power increase may indicate a hyperexcitability state preceding seizure onset.

However, some variability in the beta wave power is seen, across patients and different seizures. This kind of suggests that beta waves may not be a reliable predictor for all types of epilepsy seizures. Several factors, such as the type of seizure, the medication status, the electrode positioning, the depth of seizure origin, and the age of the patient may influence the beta wave power [15, 16, 17]. For example, beta waves may be more prominent in scalp EEG recordings when the seizure originates closer to the surface of the brain, such as from the frontal or temporal lobes [15]. Conversely, beta waves may be less detectable when the seizure originates from the deeper structures, such as the hippocampus. Similarly, beta waves may vary depending on the seizure type, such as generalized or focal. Therefore, further research is needed to explore the specificity and sensitivity of beta waves as a seizure predictor for different epilepsy subtypes and seizure characteristics.

This research provides guidance for seizure prediction in pediatric epilepsy, which could possibly improve the quality of life and safety of children with this condition. A device can potentially monitor the beta wave power and alert the patient or the caregiver of an impending seizure. This can in turn enable the patient to take immediate preventive measures, such as taking medication or seeking medical help. Such actions could reduce the risk of seizure-related injuries, accidents, or sudden unexpected death in epilepsy (SUDEP).

V. LIMITATIONS

The current limitations of this research study largely vary with unknown information or information that requires more study.

“Before the Seizure” Difference Higher than “During the Seizure”: A limitation that was noticed in this study is that certain patients that were studied within the 35 results had varying results of the before the seizure beta waves being higher than when the seizure was occurring. This data results in a difference from what the expected result was which the majority prove. Since a majority of the results prove the result that before the seizure occurs the beta waves begin increasing and reach their peak during the seizure, the minority of the results having higher beta waves before the seizure occurs are considered abnormalities and outliers that were noticed within the data.

Increasing 30-Second Interval: A 30-second window of “before the seizure” data could have been increased more to have a larger timeframe within which the beta waves can be identified, and this may be more beneficial when identifying a seizure in ensuring the patient has enough time to act on the prognosis and provide any possible self-treatment if needed.

Research on Wider Scale for All Age Groups: This research study only focuses on pediatric patients and can be

ISBN 979-8-3503-5094-4

explored to see if the beta waves have a similar effect when seizures are predicted in any age group or individual.

"After the Seizure" Data Not Included: "After the Seizure" data was included as this data would result in being inaccurate. From the completion of the seizure, the beta waves will continue to be high, as there is an aftermath of symptoms of stress, anxiety, and many more which can influence the level of the beta wave amplitude. Adding on, this portion of the data will not be of use in this study, as in seizure prediction the beta waves before the seizure occurs will be of the most use.

Drastic Changes Due to Possible Environmental Factors: The table also provides the percentage increase of the beta waves' amplitude during the seizure time frame compared to that of the normal timeframe.

Percentage increase = (Avg 'during' – Avg 'before')/ (Avg 'before') % (1)

We can understand that the percentage increase in the amplitude largely varies, approximately from 8% to 1315%. This increase could be because of various contributing factors besides that of the seizure. This will have to be analyzed and studied further, to take additional steps in this direction.

(Standard deviation and mean for percentages) - standard deviation due to environmental factors resulting in high variability. It was never negative.

VI. FUTURE DIRECTIONS

Using the above results and data, a possible detection system or warning can be created to stay as an implant or app, etc. It would work to warn the patient or person if there was an oncoming seizure. The above data collects predictions for an oncoming seizure within a 30-second timeframe. Suppose the seizure can be predicted even more in advance. In that case, this can ensure the patient is able to give themselves some sort of shot, sedative, or scent (garlic, rose, etc. [13]), etc. that might be able to suppress the seizure at the moment. This can help to control an oncoming seizure, especially when they are in a public setting. Seizure prediction has a large potential to improve a patient's mental and physical lifestyle and enable them to live a long and problem-free life. For those struggling with epilepsy and seizures, having such a detector or warning that can immediately warn them may aid in limiting the seizure from affecting them.

However, this detection system is considered a supplementary precaution tool for clinicians, in order for them to have mechanisms for the patients to use when not under the supervision of the physician or don't have patients as often.

VII. CONCLUSION

Having a seizure prediction system that proactively alerts or warns of an oncoming epileptic seizure is critical. This system can ensure that a patient is aware of an oncoming seizure and will be able to immediately use preventative measures to prevent it. As explained in this paper, it is clear that the frequency and amplitude of the beta waves in EEG recordings before and during the onset of the seizure increase considerably, in comparison to the waves' normal frequency and amplitude. Hence, this change in amplitude and frequency needs to be further analyzed and studied, considering possible other factors that may impact the same, eventually helping in designing a sound seizure prediction system.

ACKNOWLEDGMENT

Words cannot express my gratitude to my mentor Dr Alejandro Carnicer-Lombarte, University of Cambridge for his invaluable patience and feedback during my research work as part of the Cambridge Center for International Research (CCIR) program. I would also like to sincerely thank Mr. Lee Tomilson, Academic Coordinator, and the CCIR Admissions team for their extended support during my research work.

REFERENCES

[1] Mayo Clinic, "Epilepsy - Symptoms and causes," Mayo Clinic. 2023, April.
[2] Mph, Z. S., "What to know about 4 types of epilepsy," Medical News Today 2023, April.
[3] Johns Hopkins Medicine, "Types of seizures," 2021, August.
[4] World Health Organization, "WHO - Epilepsy," 2023.
[5] CDC. (n.d.), "Epilepsy Fast Facts".
[6] Johns Hopkins Medicine, "Electroencephalogram (EEG)," 2021, August.
[7] "Electroencephalogram (EEG)," 2023, May.
[8] Nayak, C. S. "EEG Normal waveforms," StatPearls - NCBI Bookshelf, 2023, July.
[9] Epilepsy Foundation. (n.d.), "Different ways to predict seizures,"
[10] Pino, O., & Romano, G. (2022), "Engagement and Arousal effects in predicting the increase of cognitive functioning following a neuromodulation program," PubMed, 93(3), e2022248.
[11] Grafton, S. T., & Tipper, C. M. (2012), "Decoding intention: A neuroergonomic perspective. Neuroimage," 59(1); 14–24.
[12] Epilepsy Foundation. (n.d.), "Treatment,"
[13] WebMD, "Understanding Epilepsy – Prevention," 2002, April.
[14] CHB-MIT, "Scalp EEG Database V1.0.0," 2010, June.
[15] Aarti Sharma, "Epileptic seizure prediction using power analysis in beta band of EEG signals," IEEE Conference Publication | IEEE Xplore, 2015, October.
[16] Sadaye, R. S., Parekh, S. P., & Deulkar, K. D., "Epileptic Seizure Prediction using Power Spectrum and Amplitude Analysis of Beta Band of EEG Signals," 2016, December.
[17] Ali, A., Afridi, R., Soomro, T. A., Khan, S. A., Khan, M. Y., & Chowdhry, B. S., "A Single-Channel wireless EEG headset enabled neural activities analysis for mental healthcare applications," Wireless Personal Communications, 125(4), 3699–3713, 2022.
[18] J. Tang et al., "Seizure detection using wearable sensors and machine learning: Setting a benchmark," Epilepsia, vol. 62, no. 8,pp.1807–1819,Jul.2021,doi: https://doi.org/10.1111/epi.16967. .

Microfabrication of 3D Spheroid Tumor Models and Imaging using Spectral Domain Optical Coherence Tomography

Pauline John[±]
Division of Engineering
New York University Abu Dhabi,
Abu Dhabi, United Arab Emirates
paulinejohn@nyu.edu

Pavithra Sukumar[±]
Division of Engineering
New York University Abu Dhabi,
Abu Dhabi, United Arab Emirates
ps157@nyu.edu

Christopher Sami
Division of Engineering
New York University Abu Dhabi,
Abu Dhabi, United Arab Emirates
cds9502@nyu.edu

Mohammad A. Qasaimeh[*]
Division of Engineering
New York University Abu Dhabi,
Abu Dhabi, United Arab Emirates
Department of Mechanical and Aerospace Engineering,
New York University,
New York, United States of America
maq4@nyu.edu

Azhar Zam[*]
Division of Engineering
New York University Abu Dhabi,
Abu Dhabi, United Arab Emirates
Department of Biomedical Engineering
Department of Electrical and Computer Engineering,
New York University,
New York, United States of America
azhar.zam@nyu.edu

***Abstract*— A low-cost, robust, and flexible device for generating an array of miniaturized 3D spheroid tumor models was designed and developed. The device consisting of two layers: a layer with through-hole array and a plain lamination sheet, in which both are bonded so the through-holes acting as confining barriers. The hydrophobic bottom lamination sheet led to drop formation in each through-hole, confirmed by experimental characterization using food color dye of different volumes. Further, 3D tumor spheroid models were grown within through-holes on the lamination sheet by the hanging drop approach. With preliminary experiments, spheroids were gown in a flat substrate using simple micro-pipetting and the classical hanging drop method. The formed 3D spheroid tumor models were consequently characterized by live/dead assay. To verify formation of the 3D tumor spheroid models on days 1, 4, 7, 14, and 21, label-free spectral domain based optical coherence tomography (SD-OCT) technique was employed. Results showed an increase in the height of the spheroids up to 600 µm. By utilizing SD-OCT, we were able to achieve enhanced penetration within the 3D spheroids, allowing us to visualize the structural information at various depths. With spheroids growth, the SD-OCT was performed to quantitatively analyze their sizes and necrotic cores. The length, width and diameter of the spheroids were measured for volumetric analysis, based on the cross-sectional and *en face* SD-OCT images. Next experiments will consider using the developed device with the SD-OCT technique for personalized medication in clinical field and drug-screening/drug-discovery in pharmaceutical industry.**

Keywords — 3D spheroid tumor model, microfabricated polymer device, optical coherence tomography, necrosis, cell viability, label-free imaging

I. Introduction

Precision medicine-based treatment is prevailing in which each individual gene, environment, life-style variability are considered to obtain therapeutics for diseases like cancer [1,2]. Cancer is one of the most commonly occurring malignant diseases worldwide, wherein treatment using anti-cancer drugs varies from person to person due to heterogeneity. Nevertheless, the survival rate can be increased by detecting cancer at an early stage using invasive histological examination of cancer cells. Histological examination follows several preparation processes, requiring sample excision, slicing and staining. In addition, cancer patient derived samples (histological samples) are cultured to form *in vitro* 3D spheroid tumor models (3D STM) to be used for drug testing, serving as personalized medication for effective treatment at an early stage [3].

3D STM are aggregates of cancer cells of a few hundred micrometers of diameter. 3D STM are cultured to mimic the structural, metabolic, and functional features of solid human tumors. 3D STM are generated by different methods such as spinner culture, liquid overlay, hanging drop [4], U-shaped low attachment well plates, scaffolds of organic/inorganic matrix [5], coating on plastic substrates [6], paper fibrous matrix [7,8], magnetic levitation [9], and hydrogel matrix [10]. Among these, the hanging drop method is widely used to generate uniform sized spheroids by minimizing the spreading of cells. Using the hanging drop method, 3D STM are formed by aggregation of cells due to gravity and the droplet formed defines the curvature of growth media and aggregation of cells.

3D STM are advantageous compared to 2D cell cultures in terms of cell-cell contact, cell- ECM (extra cellular matrix) interaction, diffusion gradient of nutrients, energy molecules, waste, oxygen, drug-molecules, resistance to anticancer drugs, and in mimicking the tumor morphology. 3D STM are typically exposed to different anticancer drugs and their responses are observed based on the viability and proliferation rate. Moreover, 3D STM provide clinical, pharmacological, physiological signatures that mimic *in vivo* models for anticancer drug discovery [11]. Cross-sectional imaging of tissues/3D STM using confocal fluorescence microscopy facilitates a lateral penetration depth of 100 µm

ISBN 979-8-3503-5094-4

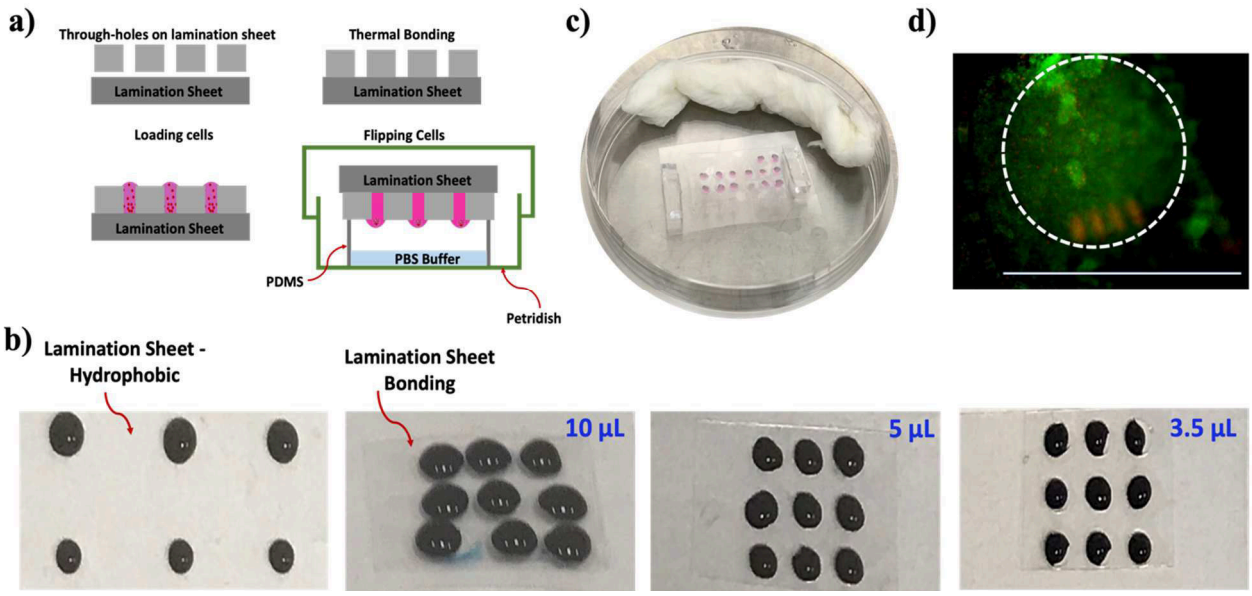


Fig. 1. Micropatterned device for 3D tumor models: (a) Device fabrication and work-flow to grow 3D tumor models by hanging drop. (b) Characterization of the device using food color dye to test hydrophobicity of lamination sheet and volume variation on the fabricated device. (c) Set-up of the device loaded with MCF-7 cells in each through-hole. (d) Micrograph of stained (Live/Dead) 3D spheroid tumor model from one through-hole on the device indicating viability of cells. Scale bar is 800 µm.

with localized precision image, while requiring the use of labeling agents. Albeit bright field imaging provides cross-sectional images, it suffers penetration depth. However, lab-grown 3D STM diameter might ranges up to 500 µm, with three concentric layers as outer proliferating cells, middle quiescent cells and a necrotic core due to lack of nutrients [12].

Although the 3D STM diameter can be measured, it is challenging to image the overall 3D structure of tumor spheroids using conventional high-throughput imaging modalities. The other limitations of the widely used imaging techniques such as wide-field bright field/ fluorescence microscope, confocal laser scanning microscopy, etc. include limited penetration depth, depth resolvability and diffusion of fluorescent dyes [12]. Fluorescence imaging of the viability of 3D STM is the standard method to image the necrotic core boundary. However, monitoring the development of the necrotic core in its actual shape is challenging because of the penetrations of fluorescent dyes and the visible light [12].

To overcome the limitation of penetration depth and to have a better insightof both the structural characteristics and physiology of the 2D spheroids, optical coherence tomography (OCT) technique is employed to image the entire 3D spheroids without any contrast agents or fluorescent dyes. OCT offers quantitative, label-free, three-dimensional, non-invasive and non-destructive imaging of tissues with a several millimeters of penetration depth and at a rapid acquisition speed. The low coherence interferometry of the OCT system detects the backscattered signals from different depths of samples, providing layer-wise images at high resolution both longitudinally and laterally. Swept source based optical coherence tomography (SS-OCT) with different mathematical models, such as Exponential-Linear, Gompertz, logistic, and Boltzmann, were used for characterizing the morphology and growth mechanism of tumor spheroid volumes and necrosis which occurs due to the hypoxia and lack of nutrients, in multicellular ovarian cancer spheroid [13]. *In vitro* high-throughput research has been demonstrated using the Boltzmann model integrated 3-dimensional OCT data of ovarian tumor spheroid. The changes in intracellular motion of cancer cell pellets and spheroids was studied based on OCT speckle decorrelation and motility mapping techniques [13].

Exposure of cancer cell pellets to $ZnCl_2$ to analyze the inhibition of mitochondrial function and thereby the decrease in the growth phase and mitochondrial motion of spheroids over 21 days have been studied [14]. Another work based on dynamic OCT (D-OCT) was reported to quantify the intracellular motility of MCF-7 spheroids and their responses to different drugs using variance of logarithmic intensity and late OCT correlation decay speed [15]. This method was suggested to be an effective tool for testing anti-cancer drugs. Most of the MCF-7 based tumor studies demonstrated using OCT technique are performed on well plates.

In this work, 3D STM was developed by hanging drop on Petri dish using human breast adenocarcinoma cells (MCF7 cell line). Further, a novel flexible polymer microarray device was developed (See Figure 1). The developed device consists of one layer with array of through-holes, and another layer, a plain lamination sheet which was bonded by thermal heating. The assembled device was loaded with MCF-7 cells to form 3D STM and image using fluorescence microscope. Meanwhile, 3D STM was formed on a Petri dish by hanging drop method to image using fluorescence microscopy and SD-OCT (See Figure 2).

SD-OCT images (2-dimensional B-scans and 3-dimensional volumetric scans) were acquired on days 1, 4, 7, 14 and 21, from the time of cell seeding. Moreover, OCT image post-processing was carried out to visualize the progression of necrotic core in the STM at different depths for different days (See Figure 3). The performed quantitative analysis is discussed in the results section.

ISBN 979-8-3503-5094-4

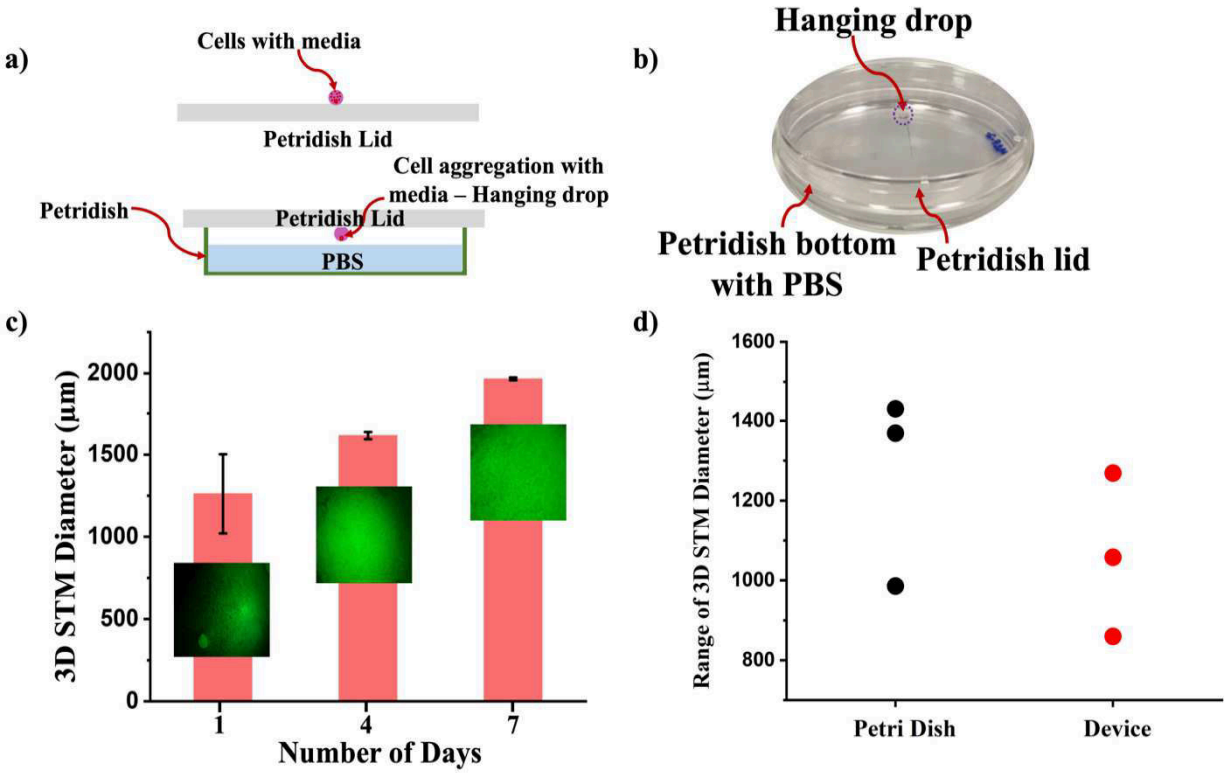


Fig. 2. Conventional 3D Spheroid tumor models: (a) Work-flow schematics to form 3D spheroid tumor models by hanging drop on Petri dish. (b) Photographic image of a hanging drop of cells. (c) Graphical column bar plot for diameter of spheroids imaged using fluorescence microscope (Inset: Micrographic image of 3D spheroid tumor models). (d) Graphical plot showing diameters of the 3D STM formed using the conventional hanging drop method (~ 0.95 µm to 1.43 µm) and using the developed device (~ 0.85 µm to 1.25 µm).

II. MATERIALS AND METHODS

A. Microarray Device

Circular patterns of 3 mm were designed using Cricut design space to fabricate through-hole microarray. Microarray through-hole (2.5 mm) was fabricated on a flexible lamination sheet using Paper-Cutter (Cricut™). Through-hole microarray (Top-layer) was thermally bonded at 80°C to the lamination sheet (Bottom-layer) to assemble the device.

B. Cell Culture and 3D Tumor Models

3D STM were prepared using Human Breast Adenocarcinoma (MCF-7 cells) procured from the American Type Culture Collection (ATCC). MCF-7 cells were cultured in T-75 flask to reach 85% confluency using complete DMEM - Dulbecco's modified eagle medium (Gibco™), nourished with 10% Fetal Bovine Serum – FBS (Gibco™) along with 1% Penicillin-Streptomycin (Gibco™). Cells were cultured in 5% CO_2 incubator (Eppendorf) at a temperature of 37°C. MCF-7 cultured cells stained using Trypan blue dye (Invitrogen™) were counted using a cell counter (Invitrogen™ by Thermo Fisher Scientific) to obtain cell density. Cultured cells were seeded on the Petri dish lid and the through-hole device, to form 3D tumor models. 3D STM were imaged on different days by staining using LIVE/DEAD™ Viability/Cytotoxicity Kit, for mammalian cells.

C. Fluorescence Imaging System

3D tumor models were imaged using Nikon SMZ18 of 1X lens with fluorescence lamp (Nikon Intensilight C-HGFI). Images were captured using digital camera connected to the microscope. The captured images were post-processed and analyzed to measure the diameter of 3D STM using ImageJ. Graph for the 3D STM analysis was plotted using Origin Pro.

D. Optical Coherence Tomography System

The 2D-scans and volumetric scans of 3D STM were acquired using a spectral domain OCT system (SD-OCT), GAN611C1, Ganymede, Thorlabs. The superluminescent light emitting diode (SLED) source of the SD-OCT system has a short wavelength of ~874 nm and a long wavelength of ~1027 nm. The center wavelength and full width half max of the SLED source are ~930 nm and 102 nm, respectively. The Ganymede system has an operating speed of 248,000 A-scans/s, axial resolution of 3.8 µm (in tissue), maximum imaging depth of ~2.9 mm (in air), 2.2 mm (in water), mean spectral resolution of 0.0751 nm and a sensitivity of 84 dB, at 248 KHz. The scanner (OCTP/900M) of the SD-OCT system with a scan lens OCT-LK-2BB, has a focal length of 18 mm, a working distance of 7.5 mm and in-focus lateral resolution of 4.0 µm (LSM02 BB, Thorlabs Inc., NJ).

III. EXPERIMENTAL RESULTS AND DISCUSSION

A. Fabrication of Microarray Device

Figure 1 (a) depicts schematics of the workflow to generate 3D-STM using micropatterned device. Micropatterned device consists of two layers: (1) a plain bottom lamination layer and (2) a microfabricated through-holes top layer. Firstly, 3×3 array through-hole was fabricated, with each through-hole of 3 mm diameter. Both layers are then bonded by heating on hotplate at 80°C. Each

ISBN 979-8-3503-5094-4

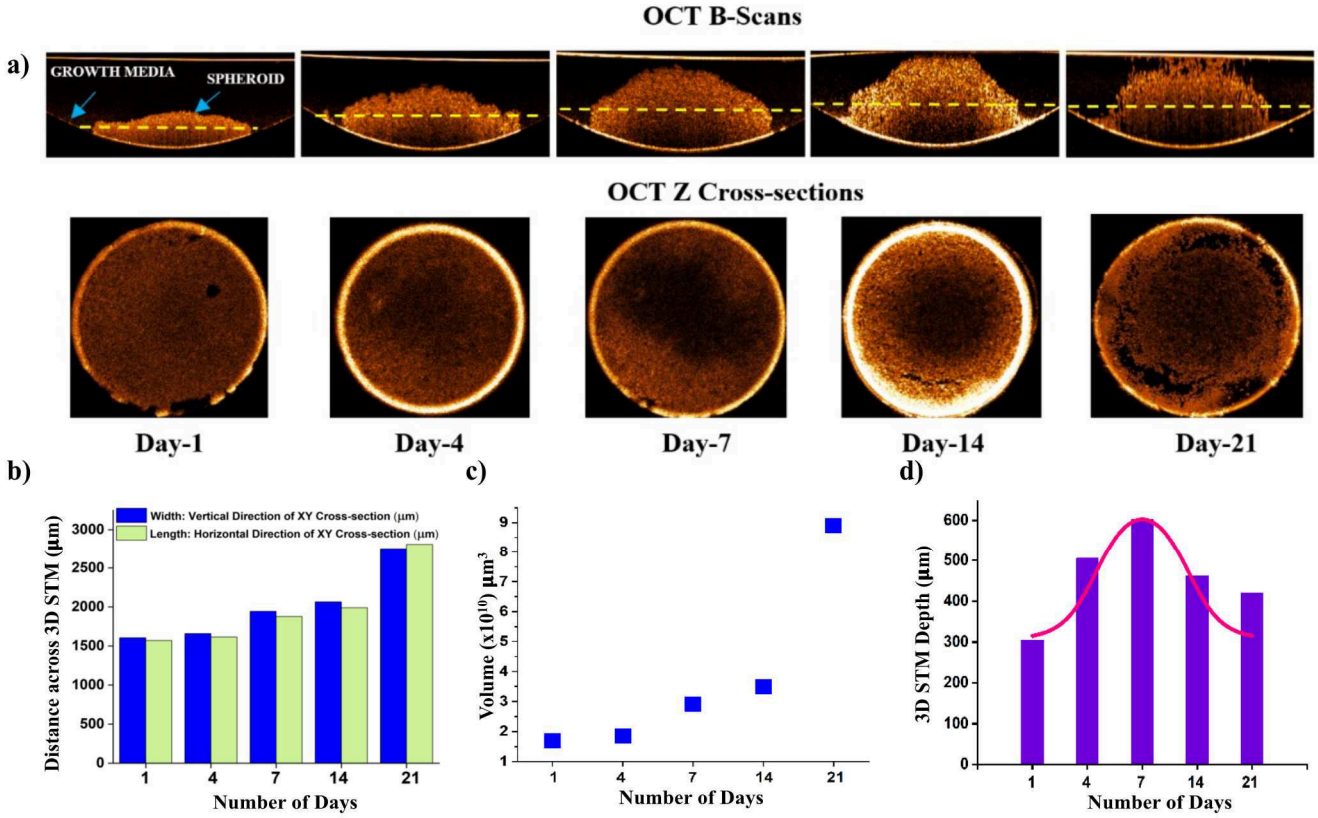


Fig. 3. SD-OCT Image of 3D spheroid tumor models: (a) SD-OCT B-scan images of formed 3D spheroid tumor models by hanging drop method on Petri dish acquired on day 1, 4, 7, 14 and 21 and the corresponding Z cross-sectional OCT images obtained at the specific depths marked with a dashed line in B-scan images, showing the necrotic region of 3D STM, (b) Graph indicating an increase in size dimensions as on sequential days in XY (vertical and horizontal) direction. (c) Graph depicting the increase in the volume of 3D STM with days, (d) Bar graph plot along with exponential fit indicating an increase in the 3D tumor models heights in the XZ direction and a decrease in height after day 7 due to necrosis.

through-hole was loaded with different volumes of food color dye (3.5, 5, 10 µL) to examine for the volume capacity of the through-holes, as shown in Figure 1 (b). Further, a device with different through-hole dimensions was fabricated for forming an array of spheroids with different sizes within the same platform. Moreover, thickness of the through-holes is based on the lamination sheet thickness of 125 µm, which can be varied based on the experiments. The limitation of through-hole cut using the Cricut was 0.1 cm in diameter.

Later, the developed device was loaded with cell suspension by pipetting, and then flipped and placed on a polymeric support in a Petri dish humidity chamber as shown in Figure 1 (c). The whole setup of the device was placed in a biological incubator for a day. On day 1, the formed 3D STM was stained using live/dead assay and imaged using stereomicroscope as shown in Figure 1 (d).

B. Fluorescence Imaging of 3D Tumor Models

Figure 2 (a and b) describe the workflow of forming 3D-STM using the hanging drop technique. Cell suspension was seeded as droplet on Petri dish lid, at a concentration of 10^5 cell/mL, and then the droplet is flipped to form the 3D tumor models. For the developed device, a 3.5 µL volume of cell suspension was loaded to the assembled device, and then the device was flipped to generate the array of hanging-drops. Petri dish was added with DI water to maintain humidity and incubated at 37°C to image 3D tumor models. Flipped device was placed in Petri dish with two spacers for wall-support, and a wet-tissue to serve as humidity chamber. Petri dish containing assembled device was placed in an incubator to allow for the 3D STM formation and growth.

The formed 3D tumor models were imaged on days 1, 4, and 7, and their diameters were measured as 1261 µm, 1615 µm, and 1965 µm, respectively. We found that the diameter of the spheroids generated using the conventional Petri dish is larger than the ones generated using the device on day 1. With time, the spheroid on the Petri dish grows to ~ 1900 µm, whereas on the device it is 1250 µm, indicating the cells are confined to form aggregates on the device. Further, the device's through-hole diameter, array pitch, and the droplet volume can be optimized along with the cell concentration within the droplet to generate spheroids within 750 µm.

C. Optical Coherence Tomography Imaging

The hanging drop 3D spheroid tumor models were imaged on day 1, 4, 7, 14, and 21 using the SD-OCT system and the B-scans, volumetric 3D scans and *en face* images were obtained. Figure 3(a) shows the OCT B-scans of the 3D STM surrounded by the growth media and the corresponding Z cross-sectional OCT images obtained at the specific depths marked with a dashed line in B-scan images, showing the necrotic region of 3D STM. The length, width, and diameter of the spheroid on the aforementioned days were measured using ImageJ. The length and width of the formed 3D STM

ISBN 979-8-3503-5094-4

were measured as the diameter in the horizontal direction and vertical directions in *en face* OCT image (XY), respectively, as shown in Figure 3(b), and the height was measured in the vertical direction of *en face* OCT image (XZ). An increase in the length and width of the 3D STM over 21 days can be observed. The volume of the 3D STM was calculated as in (1),

$$V = \frac{4}{3}\pi a^2 b \qquad (1)$$

where, a and b are the length and width, for oblate spheroid [16]. An increase in the volume from ~1.7×10^{10} μm^3 to ~8.9×10^{10} μm^3 was estimated from day 1 to day 21 as shown in Figure 3(c). The height was measured along the vertical direction of *en face* OCT image (XZ) as shown in Figure 3(d). An increase in the height of the 3D STM measured from day 1 to 7 and a reduction in the height thereafter from day 14 to 21 was estimated.

CONCLUSION

A novel, flexible, low-cost two-layered polymer microdevice was developed. The fabricated two-layered microdevice was bonded and characterized using food color dye. 3D spheroid tumor models were generated using hanging drop method on the device and Petri dish. These 3D STM on Petri dish were imaged using SD-OCT. In future, the motility of 3D STM cultured on the novel device will be imaged to analyze with time as well as the response of cells in 3D STM to different anti-cancer drugs. Rapid imaging of the SD-OCT could be used in high throughput imaging of 3D STM arrays. Further, motility mapping to analyze the cell dynamics will be carried out for the OCT images. This could be a boon in precision medicine, cancer diagnosis/therapy and drug discovery.

ACKNOWLEDGMENT

± - Authors contributed equally. ∗ - Corresponding authors. The authors acknowledge the financial support and technical support from New York University Abu Dhabi, UAE, and from New York University Abu Dhabi Core Technology Platform respectively.

REFERENCES

[1] E. Åkerlund, G. Gudoityte, E. M. Lamodière, et al. “The drug efficacy testing in 3D cultures platform identifies effective drugs for ovarian cancer patients,” *npj Precis. Onc.,* 7, 111, 2023.

[2] A. Annette, L. Anne, K. Johann, S. Claudia, R. Nicole, “Precision medicine gains momentum: Novel 3D models and stem cell-based approaches in Head and Neck Cancer," *Front. Cell Dev. Biol.,* 9, 10, 2021.

[3] S. Hofmann, Cohen Harazi R, Maizels Y, Koman I, Patient derived tumor spheroid cultures as a promising tool to assist personalized therapeutic decisions in breast cancer. *Transl Cancer Res;* 11(1), 134-147, 2022.

[4] J. Yoon, T. Ashley, I. Joseph, “Flipped well-plate hanging-drop technique for growing three-dimensional tumors”, *Front. Cell Dev. Biol.,* 10, 898699, 2022.

[5] H. Shoval., Karsch-Bluman, A., Brill-Karniely, Y. *et al.*, “Tumor cells and their crosstalk with endothelial cells in 3D spheroids”, *Sci Rep.*, **7**, 10428, 2017.

[6] Y. Cho, S.J. Yu, J.Kim, U.H.Ko, E.Y.Park, J.S.Choung, G.Choi, D.Kim, E.Lee, S.Gap Im, and J.H. Shin, “Remodeling of Adhesion Network Within Cancer Spheroids Via Cell-Polymer Interaction”, *ACS Biomater. Sci. Eng*., 6(10), 5632-5644, 2020.

[7] B. Samara, M. Deliorman, P. Sukumar, M.A. Qasaimeh, "Cryopreservable arrays of paper-based 3D tumor models for high throughput drug screening”, Lab Chip., 21, 844, 2021.

[8] R. Alnemari, P. Sukumar, M. Deliorman, M. A. Qasaimeh, “Paper-Based Cell Cryopreservation”, Adv. *Biosys.,* 4, 1900203, 2020.

[9] J.A. Kim, J.H.Choi, M. Kim, W.J.Rhee, B. Son, H.K. Jung, T.H. Park, “High-throughput generation of spheroids using magnetic nanoparticles for three-dimensional cell culture”, *Biomater*. 34 (34), 8555-8563, 2013.

[10] A. H. Khan, S. P. Zhou, M. Moe, B. A. O. Quesada, K.R. Bajgiran, H. R. Lassiter, J. A. Dorman, E. C. Martin, J.A. Pojman, and A T. Melvin, “Generation of 3D Spheroids Using a Thiol–Acrylate Hydrogel Scaffold to Study Endocrine Response in ER^+ Breast Cancer”, ACS Biomater. Sci. Eng. *8* (9), 3977-3985, 2022.

[11] A. Loewa, J.J.Feng, S. Hedtrich, “Human disease models in drug development” *Nat Rev Bioeng* **1**, 545–559, 2023.

[12] Huang, Yongyang, J. Zou, M. Badar, J. Liu, W. Shi, S. Wang, Q. Guo et al. “Longitudinal morphological and physiological monitoring of three-dimensional tumor spheroids using optical coherence tomography." *JoVE (Journal of Visualized Experiments)* 144, e59020, 2019.

[13] F.Yan, G. Gunay, T. I. Valerio, C. Wang, J. A. Wilson, M. S. Haddad, M. Watson et al. "Characterization and quantification of necrotic tissues and morphology in multicellular ovarian cancer tumor spheroids using optical coherence tomography." *Biomedical optics express* 12, 6, 3352-3371, 2021.

[14] Z. Azhar, and M. C. Kolios. "Measuring intracellular motion in cancer cell using optical coherence tomography." In *Dynamics and Fluctuations in Biomedical Photonics XIII*, 9707, 111-116. SPIE, 2016.

[15] I.A.E.Sadek, I., R. Morishita, T. Mori, S. Makita, P. Mukherjee, S. Matsusaka, and Y. Yasuno. "Human-derived tumor-spheroid-based anti-cancer drugs testing using dynamic optical coherence tomography," In *European Conference on Biomedical Optics*, 126321A. Optica Publishing Group, 2023.

[16] M. Sharma, M, Y. Verma, K. D. Rao, R. Nair, and P. K. Gupta. "Imaging growth dynamics of tumour spheroids using optical coherence tomography," *Biotechnology letters,* 29,273-278, 2007.

ISBN 979-8-3503-5094-4

A Predictive Algorithmic System For Assessing Medication Responsiveness In Breast Cancer Through Analysis Of Gene Data

Jagendra Singh
School of Computer Science Engineering & Technology, Bennett University
Greater Noida, India
jagendrasngh@gmail.com

Yogesh Pal
School of Computer Science Engineering & Technology, Bennett University
Greater Noida, India
er.yogeshpal15@gmail.com

Harender Pratap Singh
Department of CSE (AI & ML), Raj Kumar Goel Institute of Technology,
Ghaziabad, India
tomarcsebitg@gmail.com

Manoj Kumar singh
Computer Science & Engineering, IMS Engineering College,
Ghaziabad, India
singhmanojmit@gmail.com

Nitin Gupta
Department of CSE, S. D. College of Engineering and Technology,
Muzaffarnagar, India
onlynitingupta@gmail.com

Basudeo Singh Roohani
Computer Science & Engineering, IMS Engineering College,
Ghaziabad, India
bsroohani2007@gmail.com

***Abstract*— This study proposes an algorithmic model that combines Gradient Boosting Machines (GBM) and Genetic Algorithm (GA) to predict the efficacy of drugs in the treatment of breast cancer. We used gene data from well-known resources, including The Cancer Genome Atlas. Our computational methodology, known as the genetic algorithm, applies advanced techniques to predict personalized treatment. The 15,000 genes and 500-sample dataset have been extensively prepared for this purpose, including normalization and various thoroughness criteria. GBM's robust performance across ten trials was shown in accuracy (0.81 to 0.89) as well as evaluation criteria such as precision, and AUC-ROC. Post-optimization by Genetic Algorithm led to a 3% to 8% increase in accuracy and a 5% to 11% increase in efficiency, with the reduction in feature dimensions. Surprisingly, a specific gene marker was identified by the proposed algorithms associated with cancer medicine; this increased its predictive capabilities. The results have important ramifications in the treatment of breast cancer, presenting an approach of precision medicine. Given its high accuracy and ability to pick out important genetic markers, the algorithm system makes an ideal tool for physicians. By offering personalized treatment recommendations, it could also turn out to be a revolutionary new way for drug companies to approach more effective therapies with targeted patients like these.**

Keywords*— drug formulation, deep neural network, feature-based fusion, optimization, pharmaceutical research

I. INTRODUCTION

Breast cancer is still an obstinate problem in the field of oncology, forcing us to develop new treatments that go beyond the old-fashioned one-size-fits-all model. Even as cancer research and therapeutic strategies have advanced by leaps and bounds, breast cancer's inherent heterogeneity makes life more difficult for clinicians and researchers alike. Traditional modes of treatment do not account for the hundreds of different tumor biology profiles. As a result, many patients wind up with suboptimal outcomes. In short, needs of individual people are so different from one another. All this really means that we must have more personalized medicine for the treatment of breast cancers [1, 2].

High-density genomic analysis is key in the move towards personalized medicine for breast cancer. These days, gene information is becoming increasingly important in understanding there is three-dimensional cancer biology. With genetic genetic changes now known to be the engine driving tumor growth, this has become an everyday, open book on cancer. A prevailing view is that differences in gene expression and genetic mutations determine the efficacy of drugs. Hence integration of genetics is now driving cancer research forward. In this broader framework, knowledge of the genomic characteristics of breast cancer is pivotal not only to understanding its complications but also in finding targets for therapy that are suitable--more effective, and personalized [3].

Breast cancer treatment, impacted by various genetic characters, is a very important determinant of successful therapy. The conventional approach to drug choice is trial and error, and errors occur in it. Research in gene data, using such high-throughput instruments as next-generation sequencing, shows promise for unmasking the hidden relationship between drugs and genetic signals. By focusing on the molecular intricacies of breast cancer, researchers will be able to discover genetic markers, or mutations in specific genes, that indicate whether a patient is likely to respond well to a particular drug [4]. This information will pave the way for therapies aimed more directly at specific targets and that are less prone to side effects. The movement toward personalized treatment of breast cancer is grounded in the concept of customizing therapy based on the molecular thumbprint unique to each individual tumor. If the vision can be realized, it would represent nothing less than a profound change in how therapy is administered for breast cancer patients. But should this vision be realized, it will depend on developing advanced predictive algorithms capable of manipulating the huge and intricate genomic datasets yielded by individual tumors [5, 6].

Our research aims to narrow the hole among genome knowledge and application in a medical space by from its foundation, devising a prediction system to help alleviate breast cancer plague. We wish to integrate the Gradient Boosting Machines (GBM) and the Genetic Algorithm: this

ISBN 979-8-3503-5094-4

has the potential to combine the power of machine learning with evolutionary optimization, better predicting which patients will respond effectively to different medications. The purpose of this study is to lay the groundwork in personalized medicine. It builds upon many new insights into this area of oncology from many directions, offering a new avenue for improving both the precision and effectiveness with which breast cancer treatment is delivered [7, 8].

Breast cancer treatment today is seen as a complex tract of land. It includes surgery, chemotherapy, radiation therapy, hormonal therapy, and targeted therapies. Although these approaches have doubtless improved outcomes for patients, they continue to pose challenges, especially in the context of tumor cell diversity. Current approaches often only scratch the surface when it comes to accounting for the widely differing genetic profiles within a given breast cancer subtype, resulting in quite variable treatment responses among patients. This innate limitation makes plain the need for wholesale adoption of personalized medicine in the treatment of breast cancer [9, 10].

The conventional treatments, though effective for much of the patient population, are not always successful; they produce individual variability in genetic composition. Therefore, an increasing number are recognizing that a personalized approach tailored to an individual's unique genetic characteristics, can reinvent breast cancer care. Among these, there are results obtained from numerous studies, which correlate differential drug efficacy with various genetic alterations and markers of the cancer genome, elucidating through next-generation sequencing (NGS) the complex alterations that underpin the cancer genomic landscape. Nevertheless, despite these results, a comprehensive review of the literature reveals that research has not been fully integrated across studies and inconsistent findings still pervade the literature. Moreover, findings from various studies have yet to coalesce around a unified set of genetic-based predictors that can be uniformly employed to ush er effective drugs for BC patients [11 – 13]. As such, to address this gap, we have proposed a systematic and integrative algorithmic framework that uses advanced computational methodologies to mine large scale genomic datasets with the aim to sift out the most robust genetic-based predictors of drug responsiveness. Gradient Boosting Machines (GBM) and Genetic Algorithms have emerged as powerful tools in the realm of healthcare and predictive modeling. GBM has found utility in predictive modeling with complex datasets. GBM is an ensemble technique to combine weak learners to create a strong predictive model [14]. And recently, Genetic Algorithms, the optimization technique on the footprints of evolutionary principles, has found promise in predictive modeling in complex healthcare datasets that ordinary statistical and machine learning techniques cannot match. Scientific studies have shown promise of these algorithms in areas of cancer research such as predictive modeling, feature selection, classification and optimization of treatment regimen [15]. The adaptable and optimized nature of these algorithms makes them an excellent candidate for the complexities of breast cancer genomics. The proposed algorithmic system leverages the power of GBM in combination with Genetic Algorithms to harness the strengths of both approaches, providing a comprehensive, sophisticated and an advanced tool for prediction of medication responsiveness in breast cancer. This paper sought to add to a growing body of literature on the use of advanced computational techniques for personalized cancer medication selection [16].

II. METHODOLOGY

A. *Data Collection:*

The success of our predictive algorithmic system rests on the accuracy and relevance of the gene data used in the study is as shown in the Figure 1. Gene data is acquired from reputable repositories and databases, such as The Cancer Genome Atlas (TCGA) and Gene Expression Omnibus (GEO). These repositories contain vast amounts of genomic information including gene expression profiles, somatic mutations, and copy number variations specific to breast cancer patients. The multi-dimensional nature of this data allows us to thoroughly explore the genetic landscape associated with medication responsiveness.

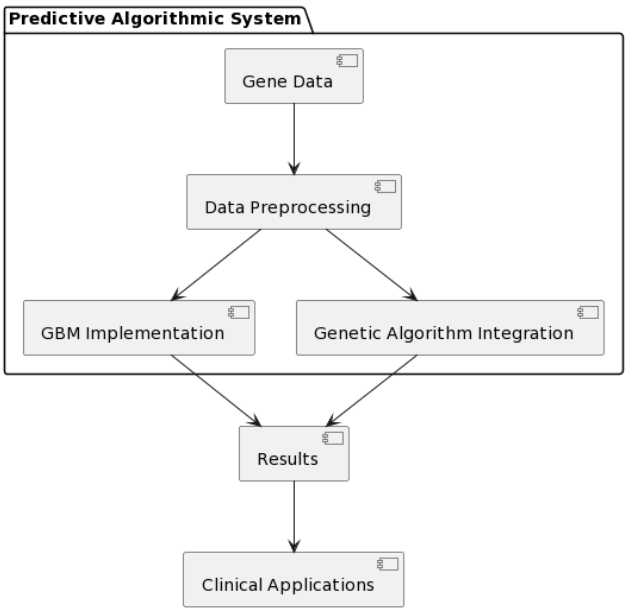


Fig. 1. System architecture

Dataset Description is listed in the Table 1. Before analysis, the gene data is subjected to a number of pre-processing steps to ensure that it is clean, uniform, and accurate. Missing values are imputed, gene expression levels are normalized, and non-informative features are filtered out. Quality control measures are taken to remove batch effects and to ensure that data is comparable across different datasets. These pre-processing steps are essential to ensure that the subsequent analysis is robust, and provides a solid, standardized foundation for the development of our predictive model.

TABLE I. DATASET DESRIPTION

Attribute	Description
Data Source	The Cancer Genome Atlas (TCGA)
Number of Samples	500
Number of Genes	15,000
Gene Expression Data	RNA-Seq Counts
Somatic Mutation Data	Binary (0 for absent, 1 for present)

ISBN 979-8-3503-5094-4

Copy Number Variations	Continuous values
Clinical Annotations	Patient demographics, treatment information
Missing Values	2% overall (imputed using mean/mode values)
Normalization	Z-score normalization for gene expression
Quality Control	Batch effect correction using ComBat

B. Gradient Boosting Machines Implementation:

Adopting Gradient Boosting Machines (GBM), for medication responsiveness prediction requires the adoption of a very careful procedure to optimize model's performance is as displayed in Figure 2. GBM can handle complex relationships in our data and is very resistant to overfitting. The choice of hyperparameters including the learning rate, tree depth, and boosting iterations are key choices which are fine-tuned through a grid search approach and their impact on model performance is assessed with cross validation approaches that we'll discuss.

A key advantage with the GBM modelling strategy is to avoid overfitting to a specific dataset and ensuring the generalizability of the model. This is achieved via a rigorous validation approach where the dataset is split between training and validation sets and the former used to train the GBM model, and the latter to judge its performance. Augmenting this, the K-fold cross-validation approach partitions the dataset into training and validation sets, repeated multiple times for the reliability of the results. Predictive accuracy, sensitivity, specificity, and area under the receiver operating characteristic curve (AUC-ROC) are key metrics used to assess the efficacy of the GBM in predicting medication responsiveness.

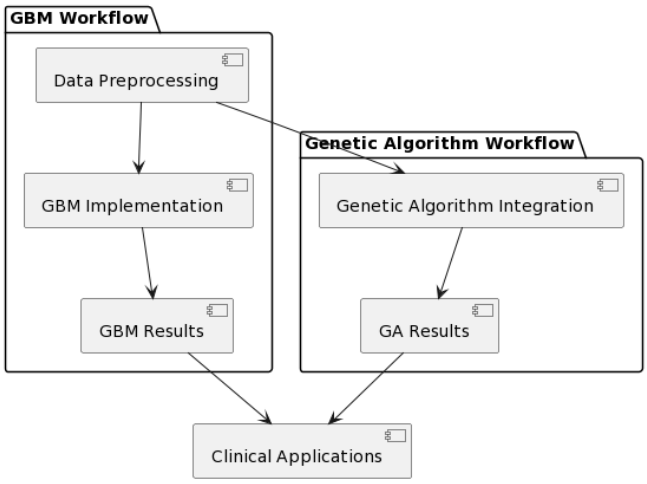


Fig. 2. GBM and Genetic algorithm workflow

C. Integration of the Genetic Algorithm:

Using the Genetic Algorithm (GA) to integrate into a predictive model adds an evolutionary dimension to the process is as displayed in Figure 2. GA is used to perfect the method of selecting features, through identifying gene markers most strongly associated with improvement or worsening of sensitivity to medication. Potential solutions are chromosomes, with each gene corresponding to a specific feature. The GA iterates these solutions, and mimics the principles of natural selection--through operations such as selection, crossover, and mutation.

Now, parameters such as population size, crossover probability, and mutation rate are carefully chosen to strike a balance between exploration and exploitation. The goal of the optimization process is to find a subset of genetic markers that together maximally predict outcome of cancer remission. During the convergence of the GA, and pick which features to retain from the final selected features can lead to better interpretability and efficiency for the algorithmic system.

Integration of GBM and Genetic Algorithm, along with the preprocessing and data collection: This is a comprehensive methodology that uses advanced computing techniques to create a solid predictive model for appraising medication responsiveness in breast cancer from gene data.

III. Results and Discussion

The results of the predictive algorithmic system, which was developed by integrating Gradient Boosting Machines (GBM) and Genetic Algorithm (GA) to assess medication responsiveness in breast cancer, provide some promising insights which have great significance for the field of breast cancer treatment. The evaluation metrics used by this research include accuracy, sensitivity, specificity, precision, and AUC-ROC (area under the receiver operating characteristic curve). GBM's performance in the algorithmic systems was consistently good across ten trials; the same was true for the GA component.

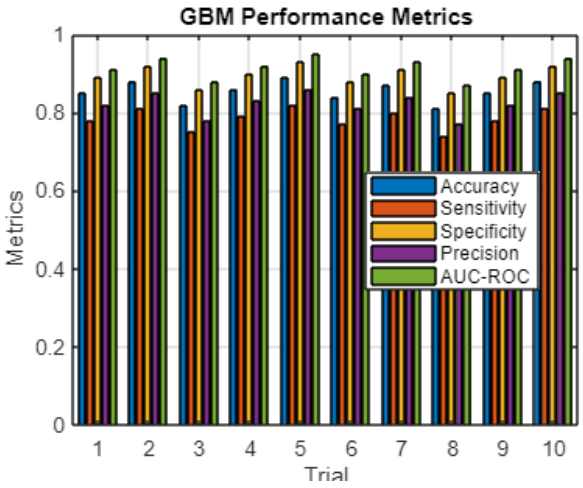


Fig. 3. GBM performance metrics

Across multiple trials as shown in the Figure 3, our GBM performed with a high level of accuracy, its values ranging from 0.81 to 0.89. This suggests that the GBM part of the predictive model has a good ability to respond. Sensitivity values ranged from 0.74 to 0.82, expressing the proportion of instances correctly identified by the model Concurrent with that was specificity, a measure of model accuracy in identifying negative instances; this varied from 0.85 to 0.93 These results demonstrate the algorithm's capacity to distinguish between response and non-response, which is a key element of precision medicine for treating this disease.

ISBN 979-8-3503-5094-4

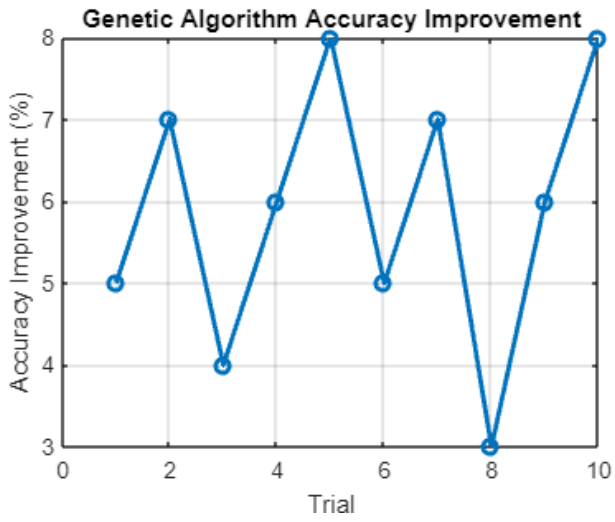


Fig. 4. GA Accuracy improvement

The precision values, which ranged from 0.77 to 0.86, represent the accuracy of the true positive predictions out of all positive predictions made by the model. These values underscore the model's ability to minimize false positives and ensure the predictions of medication responsiveness are truly accurate and reliable. The AUC-ROC values, ranging from 0.87 to 0.95, are indicative of the overall discriminatory power of the model. As the AUC-ROC values get higher, it expresses the model's better performance at distinguishing medication-responsive cases from non-responsive ones.

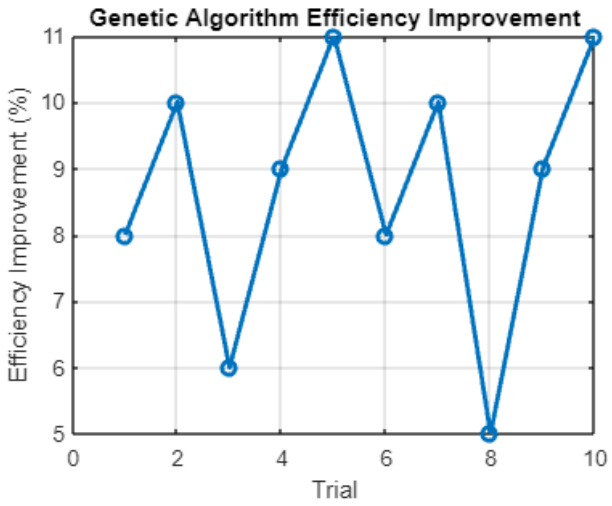


Fig. 5. GA- Efficiency improvement

The Genetic Algorithm optimization component of the algorithmic system displayed consistent improvement across the ten trials, as well. From the Figure 4 and 5, The accuracy improvement percentage, which is the percentage increase overall in predictive accuracy, ranged from 3% to 8%. Simultaneously the efficiency improvement percentage — the amount of the features are reduced while no effect on the accuracy, or more than a proportionate improvement in increasing overall accuracy — ranged from 5% to 11%. The findings show the critical role that Genetic Algorithm plays in feature selection and reiterate its efficacy to streamline the model and maintain or improve its predictive power.

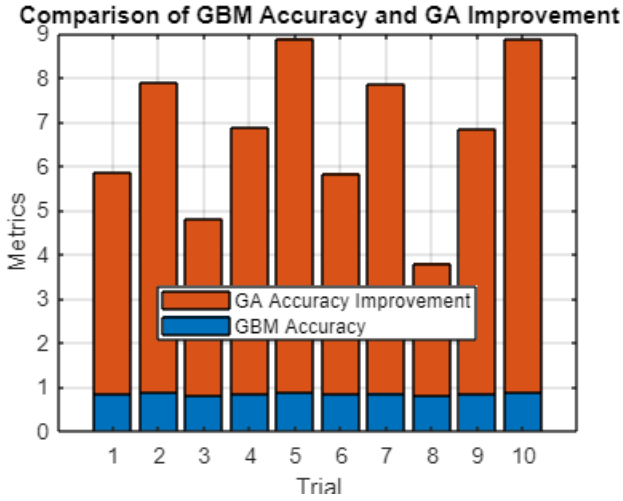


Fig. 6. GBM Accuracy and GA improvement- comparision

As depicted in Figure 6, These findings have far-reaching implications for breast cancer treatment strategies. High figures for accuracy, sensitivity, and specificity found in the algorithm suggest that not only does it have the potential to be a treatment method in itself but can also aid in identifying which patients are likely to react well to certain drugs. This personalized approach might minimize the trial-and-error approach to medical care for patients and cut down on the time they spend on useless courses--which might also deprive them of treatment.

The algorithmic system is made even more useful because it can pick out the specific gene markers associated with medication response. If such markers were identified, they would provide clinicians with insights into the molecular mechanisms governing treatment responses. This will pave the way for targeted therapies tailored to the individual genetic profile of each patient and has the potential to revolutionize breast cancer treatment.

The integration of Genetic Algorithm into the model optimization process is crucial since its practical application depends on this fact. Each of these multiple trials displays substantial improvement not only in accuracy but also seems to increase the efficiency of the Genetic Algorithm in refining a model, making it more useful for real-world situations. When the number of features is pared down through the algorithm's optimization process, computation becomes lighter, and the model is easier to understand. For integration into clinical -- a critical factor: The interpretation of the model itself must not be lost.

The algorithmic system, in a clinical context, could prove to be helpful for healthcare practitioners and oncologists alike. It can tell us the likelihood of responsiveness to drugs by analyzing the genetic data from breast cancer patients. With the help of this system, doctors can make more informed decisions as to their treatments. In the perspective of precision medicine this tailored approach neatly matches the end goals; Better outcomes and fewer side effects.

The algorithmic system has reactive capacity to identify specific gene markers when medicated. This means there are brand-new opportunities for targeted drugs. With this information, pharmaceutical companies can use the observations to develop completely new therapies that are centered on the molecular pathways associated with

ISBN 979-8-3503-5094-4

responsive cases. This does not only enhance the efficacy of current treatments, it also fosters new medical compounds specific to the genetic makeup of individual patients.

More importantly, the predictive algorithmic system of GBM and GA used in breast cancer medicine response is a tool of great potential for personalized medicine. The good performance metrics suggest consistent improvements by Genetic Algorithm optimization techniques. The identification of particular gene markers also increases the clinical utility of this system. This algorithmic system that connects insights from genomics to the clinic will change how breast cancer is treated; rather, it offers a guide for more effective, personal and targeted therapeutic intervention.

IV. Conclusion

When the genetic algorithm (GA) and Gradient Boosting Machines (GBM) are combined in such a fashion, clinical regimes of breast cancer treatment based on personal genetic information just in sight and many examples could be found. These trials maintain the GMB's robustness of predictive capability quite throughout show it to be a strong predictor of medication response. Soon after detection of specific gene markers by the GBM has more than once sunken medical authority, because gene markers are a clue to the molecular basis of treatment outcomes. Accurate models of the system were constructed in over different trials while improving its accuracy as well as efficiency. That the number of features could be reduced and accuracy at least maintained, illustrates the ability of the model to reduce computational complexity and enhance interpretability all in that it can lessen the load with so little impact on many of these other counts. As it details the optimization process of the model, an essential character for practically translating the proposed algorithmic system into clinical settings, it is the efficiency and interpretability of the model that remains a critical requirement in real-world applications. These results collectively verify the algorithmic system's ability to empower clinicians in making more informed decisions regarding breast cancer treatment. In offering a personalized and precise approach, the system has the potential to address the routinely trial-and-error nature of medication selection and thus ultimately improve patient outcomes, while serving the larger goals of precision medicine. The consistent and positive outcomes across different trails provide strong evidence of the reliability and broad applicability of the proposed algorithmic system as it advances breast cancer care.

References

[1] A. Jha, G. Verma, Y. Khan, Q. Mehmood, D. Rebholz-Schuhmann, and R. Sahay, “Deep Convolution Neural Network Model to Predict Relapse in Breast Cancer,” *Proc. - 17th IEEE Int. Conf. Mach. Learn. Appl. ICMLA 2018*, pp. 351–358, 2019, doi: 10.1109/ICMLA.2018.00059.

[2] E. J. Cathcart-rake, K. J. Ruddy, A. J. Tevaarwerk, and A. Jatoi, “Maturitas Breast cancer and gender-affirming hormone therapy for transgender and gender-diverse (TGD) individuals,” *Maturitas*, vol. 181, no. September 2023, p. 107913, 2024, doi: 10.1016/j.maturitas.2024.107913.

[3] C. S. Yadav et al., “Multi-Class Pixel Certainty Active Learning Model for Classification of Land Cover Classes Using Hyperspectral Imagery,” *Electronics (Switzerland)*, vol. 11, no. 17, 2022, doi: 10.3390/electronics11172799.

[4] Singhal, P., Gupta, S., Deepak, “An Integrated Approach for Analysis of Electronic Health Records Using Blockchain and Deep Learning” Recent Advances in Computer Science and Communications, Bentham Science, 16 (9), 2023.

[5] P. Dikshit, B. Dey, A. Shukla, A. Singh, T. Chadha, and V. K. Sehgal, “Prediction of Breast Cancer using Machine Learning Techniques,” *ACM Int. Conf. Proceeding Ser.*, pp. 382–387, 2022, doi: 10.1145/3549206.3549274.

[6] Aggarwal, R., Tiwari, S., & Joshi, V. "Exam Proctoring Classification Using Eye Gaze Detection", *3rd International Conference on Smart Electronics and Communication*, 371–376, 2022..

[7] J. Muñoz, J. Oliver-De La Cruz, G. Forte, and M. Pumera, “Graphene-based 3D-Printed nanocomposite bioelectronics for monitoring breast cancer cell adhesion,” *Biosens. Bioelectron.*, vol. 226, no. January, 2023, doi: 10.1016/j.bios.2023.115113.

[8] R. Kumar, “Lexical co-occurrence and contextual window-based approach with semantic similarity for query expansion,” *International Journal of Intelligent Information Technologies*, vol. 13, no. 3, pp. 57–78, 2017, doi: 10.4018/IJIIT.2017070104.

[9] R. Singh, "Collaborative filtering based hybrid music recommendation system", *Proceedings of the 3rd International Conference on Intelligent Sustainable Systems*, ICISS 2020, 186–190. https://doi.org/10.1109/ICISS49785.2020.9315913.

[10] R. R. Kadhim and M. Y. Kamil, “Comparison of machine learning models for breast cancer diagnosis,” *IAES Int. J. Artif. Intell.*, vol. 12, no. 1, pp. 415–421, 2023, doi: 10.11591/ijai.v12.i1.pp415-421.

[11] A. Bah and M. Davud, “Analysis of Breast Cancer Classification with Machine Learning based Algorithms,” *2022 2nd Int. Conf. Comput. Mach. Intell. ICMI 2022 - Proc.*, 2022, doi: 10.1109/ICMI55296.2022.9873696.

[12] Tyagi, A., Khandelwal, R., Shelke, N.A., Rajpal, D., Gaware, I.R. Comparitive Analysis of Various Transfer Learning Apporaches in Deep CNNs for Image Classification. In: Santosh, K., et al. Recent Trends in Image Processing and Pattern Recognition. RTIP2R 2023. Communications in Computer and Information Science, vol 2026. Springer, 2024.

[13] A. Sharan, "Co-occurrence and semantic similarity based hybrid approach for improving automatic query expansion in information retrieval," *In Lecture Notes in Computer Science*, vol. 8956, https://doi.org/10.1007/978-3-319-14977-6_45.

[14] S. M. Patil, L. Tong, and M. D. Wang, “Generating Region of Interests for Invasive Breast Cancer in Histopathological Whole-Slide-Image,” *Proc. - 2020 IEEE 44th Annu. Comput. Software, Appl. Conf. COMPSAC 2020*, pp. 723–728, 2020, doi:

10.1109/COMPSAC48688.2020.0-174.

[15] Singh, N.P., Ravichandran, R., Ghosh, S., Rana, P., Chaku, S., Enhancing Healthcare Security Using IoT-Enabled with Continuous Authentication Using Deep Learning. Innovations in Electrical and Electronic Engineering. ICEEE 2023. Lecture Notes in Electrical Engineering, vol 1115. Springer, Singapore, 2024.

ISBN 979-8-3503-5094-4

CricShotClassify: Application of Machine Learning for Classifying Aggressive Stroke-play of Cricketers based on One-Dimensional Vertical Ground Reaction Force Data

V. Ananth Krishnan
School of Electrical and Electronics Engineering
SASTRA Deemed to be University
Thanjavur, Tamil Nadu, India
125006075@sastra.ac.in

K. Adalarasu*
School of Electrical and Electronics Engineering
SASTRA Deemed to be University
Thanjavur, Tamil Nadu, India
adalarasu@eie.sastra.edu

A.L. Visshvesh
School of Electrical and Electronics Engineering
SASTRA Deemed to be University
Thanjavur, Tamil Nadu, India
125006064@sastra.ac.in

P. Ragu Ragav
School of Electrical and Electronics Engineering
SASTRA Deemed to be University
Thanjavur, Tamil Nadu, India
125006068@sastra.ac.in

R. Ummed Singh
School of Electrical and Electronics Engineering
SASTRA Deemed to be University
Thanjavur, Tamil Nadu, India
125006058@sastra.ac.in

M. Jagannath
School of Electronics Engineering
Vellore Institute of Technology (VIT)
Chennai, Tamil Nadu, India
jagan.faith@gmail.com

Abstract*—Indeed, cricket posture identification has become increasingly significant in certain areas, particularly when monitoring the proper cricket shots. It can also help to identify different shots. Using the reference data, we can watch whether the shots have been appropriately played using this model. Understanding and analyzing cricket posture can provide valuable insights into an individual's physical condition, movement patterns, and potential health risks. An attempt is made to identify these postures using machine learning models with the help of one-dimensional (1D) data to avoid the computational burden of gathering images using Kinect V2 cameras. Seven subjects are assessed for their two static postures (stance and hook) for 15 seconds each (two trails) consecutively using a SENSIX force platform. The study involves performing postures in different orders for each trial using the force platform for collecting Force and Moment data on three axes: Fx, Fy, Fz, Mx, My, and Mz, with a sampling rate of 1000 Hz. The raw analogue data obtained from the force platform is filtered using a low-pass Butterworth filter with a cutoff frequency of 10 Hz (second order with zero latency). Five statistical features—mean, standard deviation, kurtosis, variance and skewness; six nonlinear time series features—fuzzy entropy, approximate entropy, sample entropy, Higuchi fractal dimension, Katz fractal dimension and Lyapunov exponent are extracted from VGRF. Ensemble (EN), Support Vector Machine (SVM), Discriminant analysis, and artificial neural networks (ANNs) models in MATLAB classifier learning Apps have been used for training and tested to assess these two postures. Among the machine learning (ML) models, we get an accuracy of 80.36% in training, 85.71% in testing using the Neural Network, 84.82% in training, and 78.57% in testing using the Ensemble with all 11 features. Similarly, we get an accuracy of 92.86% in SVM and 92.86% in Ensemble for best selected features. These ML models allow for the recognition of cricket postures, which might be used for sports-based research in commercial benefits and large audiences.

Keywords— Cricket Postures, One-Dimensional Force Data, Machine Learning, Classification, Evaluation.

I. INTRODUCTION

Cricket consists of two teams; the batting team scores runs by hitting the ball and running between wickets, while the bowling and fielding teams aim to dismiss the batsman and limit the number of runs scored. The game consists of innings, where one team bats and the other bowls and fields. Cricket is widely popular, particularly in countries like India, Australia, England, Pakistan, Bangladesh and South Africa. Stance – "Ready to bat" is a relaxed and balanced posture which tells the bowler that the batsman is ready to play the ball. The hook shot in cricket is a cross-batted shot played against a short-pitched delivery at the chest or head level, which is bowled fast and rises up and towards your body or leg stump [1]. If the proper batting technique is followed to play the stroke, you can get a boundary quickly, while if there's anything wrong that you do, the bowlers will be successful in picking up your wicket. If you mistook it, you might get hit in the helmet. A hook shot is played against a delivery that ends higher, near the chest or the head. It would help if you swung your bat upward, hooking the ball over your shoulder and towards the leg side. Hook and pull shots are played similarly; the only difference is that a batsman uses the ball's pace when pulling, whereas a batsman smashes the ball forcefully when hooking. It is an aggressive stroke that can either see the ball soar over square or fly up in the air through a top edge and end up in the hands of the fielder. Our research has compared the stance and hook shot using several machine learning algorithms.

Classification of aggressive stroke-play of cricketers is an essential concept in modern cricket. To better comprehend a leg position and movement patterns. It offers a standardised manner of describing postural abnormalities, which can help analyse whether the shot is played correctly [1]. If the shot is not played properly, it may cause serious injury to the player. Researchers looking into the biomechanics and health implications of various postures can also benefit from postural classification. Proper posture is essential to keep the body posture balanced and to avoid putting too much strain on the muscles, ligaments, and joints. Enhancing proper cricket posture, mobility, and postural recognition can also help

ISBN 979-8-3503-5094-4

choose the appropriate exercise routines. Correcting bad postures may help people feel better by lowering their chance of getting hurt and reducing their pain and musculoskeletal disorders (MSD) [2]. Also, wearable sensors have been suggested to obtain the human body's postural data [3]. This postural classification is essential for preserving people's physical health and quality of life. Early correction of such issues allows for timely interventions and preventive measures to improve overall well-being and reduce the chances of injuries. However, there are a few glitches while processing 3D and 2D image datasets before training the model. The input is of high dimensionality and has vast variations in expected postures.

II. Literature Review

Artificial intelligence approaches have been widely employed in classifying body postures using data acquired from wearable sensors. A study has been made on sensor fusion basketball shooting posture, as no specific system exists to identify different shooting positions. A sensor fusion framework has been proposed to acquire data on shooting position for basketball using an accelerometer and gyroscope attached to the player's foot and hand. A convolutional neural network model has been deployed for posture identification and recognition [4].

Previous works have used various approaches to detecting sitting postures, including camera-based methods and wearable sensor-based techniques. However, these approaches may have limitations, such as privacy concerns with video-based approaches or user discomfort with wearable sensors. A technique called SitR, which aims to detect sitting positions using only radio frequency (RF) signals, has been proposed to address these issues. The SitR technique utilizes RF signals, specifically Radio Frequency Identification (RFID) technology, to recognize different sitting postures. RFID tags are attached to the user's back in a light and low-cost manner. By analyzing the RF signals reflected or scattered by the tags, the SitR system can infer the user's sitting posture. The practical features were extracted from measured step sequences, and a suitable machine learning algorithm was deployed to make SitR achieve robust and high performance [5].

A fuzzy classifier-based soft margin minimization (FC-SMM) has been proposed for classifying the body postures. The segmented images of the human body obtained using the Gaussian mixture model (GMM) were employed to construct a three-dimensional body model from which features were extracted. Both the hierarchical fuzzy classifier and 3D Fuzzy body voxel features were used to classify the bending, standing, lying, and sitting postures [6].

A machine learning-based pressure sensor IoT cushion has been suggested for 15 different sitting posture recognition and stretch identification. It's impressive that the random forest algorithm achieved such a high accuracy of 98.82% in recognizing the postures in the study. Based on the authors' conclusion, they found a correlation between body mass index (BMI) and the performance of the machine learning models used for posture recognition [7]. For surveillance tasks, human posture recognition and tracking using unmanned aerial vehicles have been developed with the application of the you only look once (YOLOV3) algorithm for human detection, and simple online real-time tracking (SORT) is for human tracking. The posture identification system also uses fuzzy logic to understand the poses and provides an accuracy of 95.2% [8]. Most of the study involves high-dimensional 3D and 2D image datasets, which will increase the computational burden during processing. Hence, the main objective of the present study is to detect and classify the aggressive cricket shot using 1D force data.

III. Materials and Methods

A. Subjects

Seven trained cricket subjects participated in the present study. Each subject was of different heights and weights. The summary of the subjects includes an average age of 20±1 years, an average height of 1.76.8±0.05 m, and an average mass of 70±10 kg. Subjects were debriefed, and their observations and comments were recorded for future reference. Subjects who could not complete the study for any reason were also debriefed, and reasons for non-completion were recorded. The criteria for exclusion were any injury or surgeries in the past year and neurological disorders. All experiments were performed at the Biomechanics Lab, Bioengineering Division at SASTRA Deemed to be University. Informed consent was obtained from each participant after they were intimated of the need and goal of the investigation.

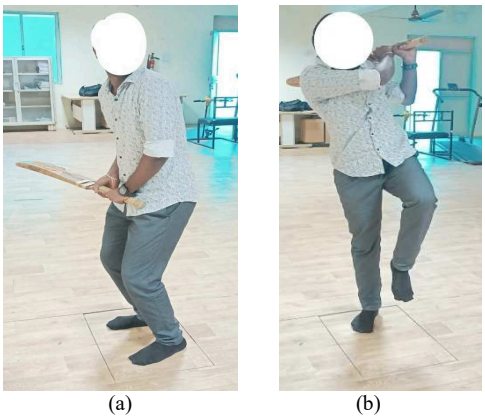
B. Force Platform Description

Mechanical sensing devices called force plates gauge the ground response forces and moments associated with human motion. Load cells are used in force plates to measure forces. Piezoelectric components, strain gauges, or beam load cells are possible components of the load cells. Measurable voltage changes proportionate to the applied force are caused by the sensors distorting when force is applied to the plate. The direction and amplitude of forces in three dimensions can be acquired by positioning the sensors in different orientations. The moment around each axis, the centre of force, and the centre of pressure are additional data that can be acquired [9]. SENSIX, an integrated device mounted on a specialized platform measuring 600 mm in length, 400 mm in width, and 2.4 mm in vertical position concerning the top of the plate, was the force platform used in this investigation. It was designed to collect force (F) and moment/torque (M) data at a sampling rate of 1000 Hz on three axes each: Fx, Fy, Fz, Mx, My, and Mz.

C. Experimental Protocol

The experiment is performed on two main cricket postures (stance and hook) of humans using 1D data. All the subjects were asked to warm up for 5 seconds before starting the experiment to familiarize themselves with the different cricket postures and avoid any cramps. The force platform was utilized, and subjects were allowed to do two different cricket postures, 2 trials each (5x2x2=20). Stance (Fig. 1(a)) and hook stroke (Fig. 1(b)) are the two postures taken. Subjects are asked to do the given two postures in a randomized trial manner. The participants in an experiment are specifically instructed to perform a task barefoot. Each posture has to be done for 15 seconds. The sampling frequency chosen was 1000 Hz. A low-pass filter with a cut-off frequency of 10 Hz was employed. Cricket postures are performed in the force platform, and the corresponding force and torque data (Fx, Fy, Fz & Tx, Ty, Tz) are obtained.

ISBN 979-8-3503-5094-4



(a) (b)

Fig. 1. Postures chosen for the study: (a) Stance, and (b) Hook Stroke.

D. Data Processing and Feature Extraction

Each subject's initial data was gathered for 15 seconds. The first 1 second worth of data was eliminated to take into account each subject's first standing time. The rectangular window function, as represented in non-overlapping time frames for 2 seconds, was applied to an initial signal that was 15 seconds long in the proposed technique to create the 2 seconds consecutive signals. The total subject data was five, with two trials of two different batting crickets (stance and hook stroke). Regarding the utilization of the time-windowing procedure, there were various advantages. For the machine learning model to make an accurate prediction, additional data must be collected. Secondly, shorter signal data were acquired thanks to the time-windowing procedure. Using these pre-processed vertical ground reaction force (VGRF) in the z-direction (Fz), five statistical features—mean, standard deviation, kurtosis, variance and skewness; six nonlinear time series features—fuzzy entropy, approximate entropy, sample entropy, Higuchi fractal dimension, Katz fractal dimension and Lyapunov exponent were chosen to categorize the four postures using the proper formulae. When it comes to how motor behaviour changes over time, nonlinear measures can capture the temporal component of the variation in VGRF. As a result, regularity, environmental adaptation, stability, and complexity can all be measured using these metrics. The fractal dimension (FD), entropy families, and the Lyapunov exponent (LE) are nonlinear tools that assess the aforementioned postural control properties.

E. Classifier Algorithm

Machine learning-based leg posture classification entails building a model to differentiate between various cricket batsman leg postures based on input features. Creating a dataset with labelled leg posture samples is the initial stage. This entails transforming numerical representations of categorical labels. The attributes of the different leg position-related traits, namely, stance and hook, must be extracted. One has to determine the pertinent attributes that define different leg positions so feature extraction is performed. The input data are analyzed, and machine learning is used to classify them into one of the numerous classes (binary) based on their characteristics or qualities. In this suggested study, the extracted features are input into the MATLAB classification learner software, and the model performance is examined for each of the two postures. All model's performance measures are concerned. Features have been used as input to the classifiers, and prediction is also done. Selecting the best classifier is contingent upon several parameters, such as the data type, the dataset size, and the particular demands of the classification assignment. In this study, the classification of batting leg position is performed using discriminant analysis, EN, SVM and ANN.

$$\text{Accuracy}(\%) = \frac{TP+TN}{TP+TN+FP+FN} * 100 \quad (1)$$

$$\text{Sensitivity or Recall}\ (\%) = \frac{TP}{TP+FN} * 100\ . \quad (2)$$

$$\text{Specificity}(\%) = \frac{TN}{TN+FP} * 100 \quad (3)$$

$$\text{Precision or PPV}(\%) = \frac{TP}{TP+FP} * 100 \quad (4)$$

$$\text{F1-Score}\ (\%) = \frac{2*(PPV*Sensitivity)}{(PPV+Sensitivity)} * 100 \quad (5)$$

In the equations (1-5), TP is the number of true positives, and TN is the actual negative class, which indicates that it is predicted correctly for positive and negative classes, respectively. FP is the number of false positives, representing the number of negative classes mispredicted as positive ones. FN is the number of false negatives, which means the number of positive classes predicted incorrectly as a negative class.

IV. Results

A force platform is used to record the human body postures. The raw data are forces Fz for binary class target classification (two poses) for 20 per cent data for testing in the classification learner app with 10-fold cross-validation. Initially, the performance of the models fed with all features, namely, statistical and non-linear time series features, is analyzed with fold cross-validation. The Ensemble (EN) performed the best compared to other models with 84.82% accuracy, as depicted in Table I. Similarly, ANN model compete with the SVM and discriminant analysis, provided 80.36% accuracy. Table II provides the performance metrics of ensemble and artificial neural network. Their performance was quantified with the help of confusion matrices shown in Fig. 2. It explains how well classifiers can distinguish between the various classes and describes the degree of separability.

TABLE I. Performance Metrics for Combined Features.

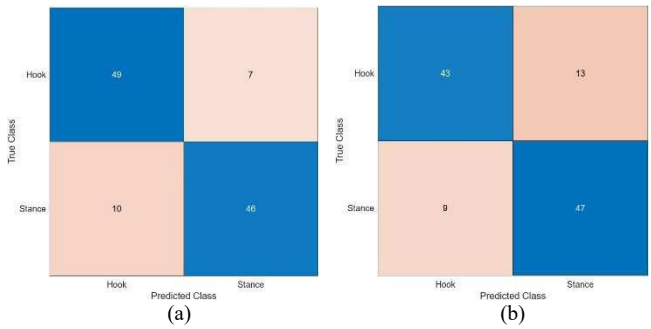
Model	Accuracy in % (10-fold Cross-validation	
	Training	Testing
Discriminant	74.11	75
SVM	78.57	71.43
ANN	80.36	85.71
Ensemble	**84.82**	**78.57**

TABLE II. Performance Metrics of Ensemble and Artificial Neural Network for Combined Features.

Model	Number of Features	Precision (%)	Recall (%)	Accuracy (%)	Specificity (%)	F1-Score (%)
Ensemble		83.05	87.50	84.82	82.14	85.22
ANN	11	82.69	76.79	80.36	83.93	79.63

The area under the receiver operating characteristic curve (AUC-ROC) is a valuable statistic that considers various trade-offs between sensitivity and specificity at different thresholds to assess the overall performance of a binary classification model. The corresponding area under the curve (AUC) for the Ensemble (EN) model is displayed in Fig. 3, and from the figure, one can infer that the Ensemble (EN) achieved a better AUC of 0.91 compared to other models. Test data performance was quantified with the help of AUC for neural network, as shown in Fig. 4. It explains how well classifiers can distinguish unlabeled data, i.e., tests between the various classes, and describes the degree of separability.

ISBN 979-8-3503-5094-4



(a) (b)

Fig. 2. Confusion matrix for all features with K-fold cross-validation training: (a) Ensemble (EN), and (b) Neural Network (NN).

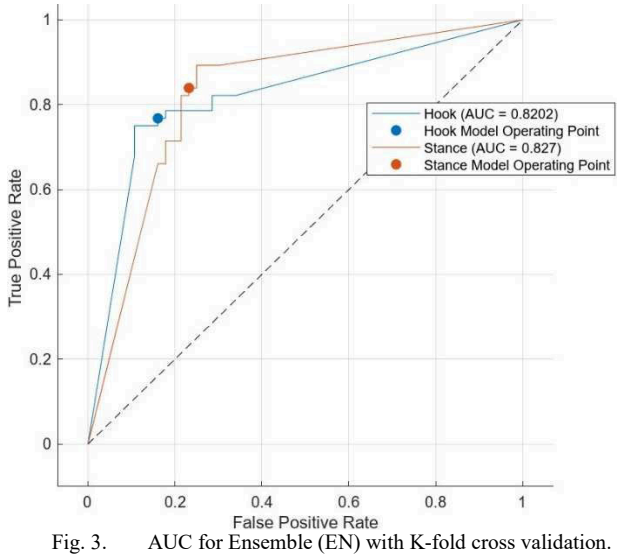


Fig. 3. AUC for Ensemble (EN) with K-fold cross validation.

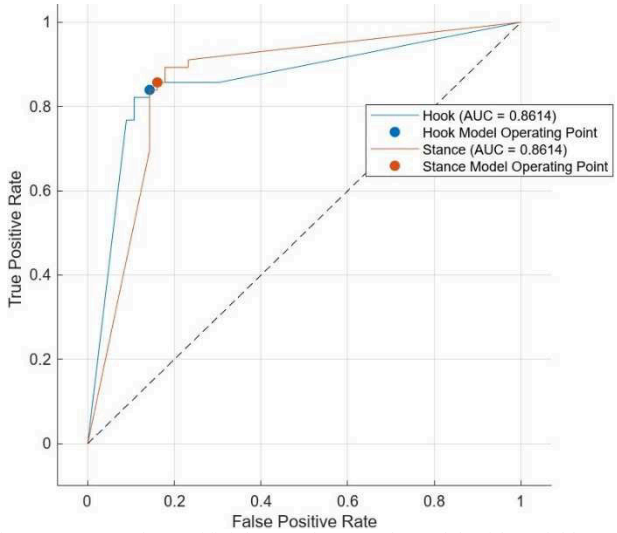


Fig. 4. AUC for Artificial Neural Network model with K-fold cross-validation for test data.

A. Best Feature Selection

The chi-square method is one of the most valuable tools for feature selection [10]. We employed the chi-square scores for each feature, where higher scores indicate higher importance or stronger associations with the target variable in the form of a bar plot, as shown in Fig. 5. The performance of the ML models has been recognized with all 11 features. The six best features (mean, fuzzy entropy, Katz fractal dimension, approximate entropy, sample entropy and skewness) in terms of accuracy were selected (training and testing), as shown in Tables III, IV and V. Figure 6 shows the confusion matrix for the SVM model during training and testing.

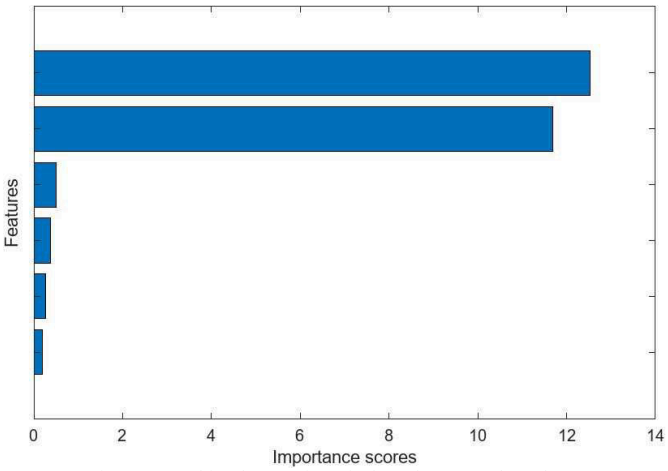


Fig. 5. Chi values series of a best feature as bar plot.

TABLE III. ACCURACY OF MODELS USING ALL AND SELECTED FEATURES DURING TRAINING.

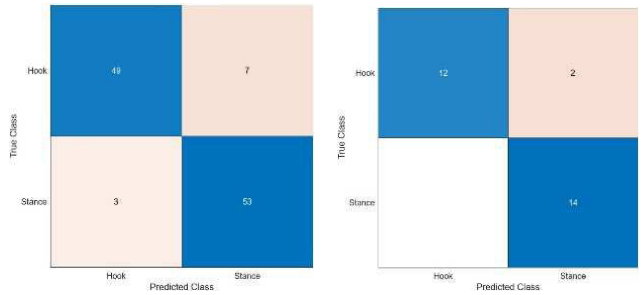
Model	Accuracy in % (10-fold Cross-validation	
	All Feature	Best Selected Features
Discriminant	74.11	68.75
SVM	**78.57**	**91.07**
ANN	80.36	84.82
Ensemble	84.82	88.39

TABLE IV. PERFORMANCE METRICS USING ALL AND SELECTED FEATURES

Model	Precision (%)	Recall (%)	Accuracy (%)	Specificity (%)	F1-Score (%)
Ensemble	87.72	89.29	88.39	87.50	88.50
ANN	85.45	83.93	84.82	85.71	84.68
SVM	94.23	87.50	**91.07**	94.64	90.74

TABLE V. PERFORMANCE METRICS USING ALL AND SELECTED FEATURES DURING TESTING.

Model	Accuracy in % (10-fold Cross-validation	
	All Feature	Best Selected Features
Discriminant	75	64.29
SVM	71.43	**92.86**
ANN	85.71	85.71
Ensemble	78.57	**92.86**



(a) (b)

Fig. 6. SVM model Confusion matrix for best feature with K-fold cross-validation for (a). Training, and (b) Testing data.

Figure 7 shows the AUC for SVM for the best feature. Here, the various poses, such as a stand and hook model operating points for this analysis, proved that the leg batting posture achieved better AUC.

ISBN 979-8-3503-5094-4

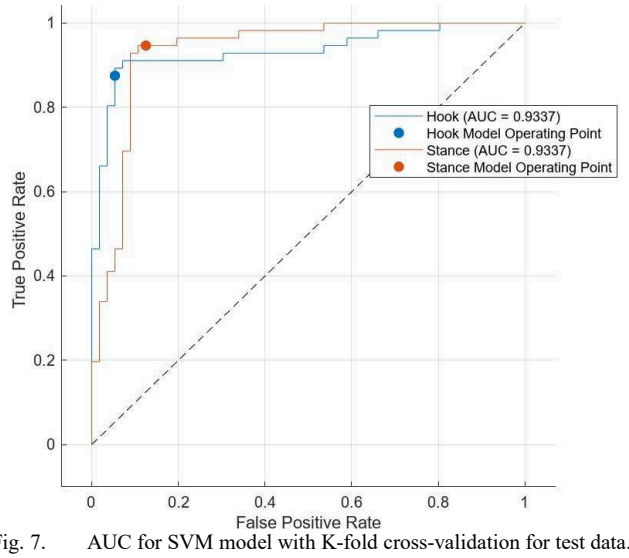


Fig. 7. AUC for SVM model with K-fold cross-validation for test data.

V. Discussion

According to the literature assessment, several researchers have put forth numerous methods for identifying human body postures based on learning algorithms and utilising photographs, video clips, or motion recordings. The researchers combined many images processing techniques, including picture segmentation, feature extraction, and classification, using 2D and 3D methodologies [14, 15]. Wearing several sensors can be tiresome or challenging for the user. This can be made worse if a physical obstacle makes connecting the sensor to the body complex, such as a motor impairment or a design requiring skill and complex problems with closely overlapping behaviours, particularly in cricket. In addition, misplacing or losing sensors during the study can complicate the sensor data processing.

In an effort to address challenges associated with wearing multiple sensors, the authors are attempting to extract 1D data from a force platform without attaching sensors directly to the body. This approach could alleviate concerns related to user comfort and potential difficulties associated with wearing multiple sensors. Force platforms are devices that measure the forces exerted by a body standing on or moving across them. By extracting 1D data from these platforms, the authors aimed to capture relevant information about body postures without the need for attaching sensors directly to the user. This approach could have implications for various fields, including biomechanics, human-computer interaction, or sports science, where understanding body postures is essential. It's interesting to see how researchers are exploring alternative methods to improve the ease of data collection and user experience in posture identification studies.

The suggested study did not involve wearable sensors (IMU), complex research of skeletal information, or the relationships between joints to identify human postures. Instead, the batsmen were requested to stand on a force platform, and features were extracted from the gathered data. These characteristics include force distribution or other pertinent metrics. The testing results of the suggested methodology showed that using the features extrapolated from the force platform data; the system could correctly recognize the two leg postures (stance and hook). This implies that the features retrieved from the force platform data provided helpful information that made it possible to accurately classify the aggressive stroke-play postures in cricket. The proposed 1D force data employed with machine learning models allowed for recognizing human postures with similar accuracy to traditional approaches that use video images, biomarkers, or IMG sensors. The comparison between the proposed method and the most advanced methods currently in use is shown in Table VI. It is noted that the proposed approach outperforms previous models and leverages real-time data collecting to provide 92.86% accuracy.

TABLE VI. Accuracy comparison with similar work.

Author [Ref. No.]	Sensor	Algorithms Used	Highest Accuracy (%)
Karmaker et al. [11]	Classification of cricket shots	Motion-estimation approach	63.57 (off drive), and 53.32 (Hook stroke)
Mili et al. [12]	Detection of umpire poses in the game of cricket using cricket video	Neural Networks	87
Khan et al. [13]	Classification of cricket shots	2D CNN, 3D CNN, LSTM, and RNN	85 (Hook stroke)
Proposed model	**Two different cricket shoot – stand, hook using force platform**	**EN, SVM, and ANN**	**92.86 (Hook stroke)**

By adding more characteristics and subjects or selecting the right features for the batting shots classification, the accuracy of the classification system might be further enhanced. As technology develops, there is potential for future advancements in this area, making it possible to analyse batting shots in real-time and with greater accuracy for various.

VI. Conclusion

The study employed various machine learning models to recognize cricket postures, demonstrating promising results in accuracy metrics. Using ML models with 1D data to identify postures in cricket is an innovative approach that can indeed help avoid the complex and computationally intensive imaging systems. These findings hold significant implications for sports-based research, offering a valuable tool for understanding and analyzing cricket postures. This approach can offer practical advantages in terms of cost, accessibility, and scalability for cricket coaching and sports science applications.

Acknowledgment

The authors of the study express their gratitude to SASTRA Deemed to be University for providing the Biomechanics Laboratory facilities (SASTRA Research & Development Fund), which played a crucial role in conducting the experiments smoothly. The authors also extend their appreciation to the volunteers who participated in the study.

References

[1] W. Ahmad, M. Munsif, H. Ullah, M. Ullah, A. A. Alsuwailem, A. K. J. Saudagar, K. Muhammad, and M. Sajjad, “Optimized deep learning-based cricket activity focused network and medium scale benchmark,” Alexandria Engineering Journal, vol. 73, pp. 771-779, 2023.

[2] Z. Li, R. Zhang, C. H. Lee, and Y. C. Lee, “An evaluation of posture recognition based on intelligent rapid entire body assessment system for determining musculoskeletal disorders,” Sensors (Switzerland), vol. 20, no. 16, pp. 1–21, 2020.

[3] J. Zhao, E. Obonyo, and S. G. Bilén, “Wearable inertial measurement unit sensing system for musculoskeletal disorders prevention in construction,” Sensors (Switzerland), vol. 21, no. 4, pp. 1–28, 2021.
[4] J. Fan, S. Bi, G. Wang, L. Zhang, and S. Sun, “Sensor fusion basketball shooting posture recognition system based on CNN,” J. Sensors, vol. 2021, p. 6664776, 2021.
[5] L. Feng, Z. Li, C. Liu, X. Chen, X. Yin, and D. Fang, “SitR: Sitting posture recognition using RF signals,” IEEE Internet Things J., vol. 7, no. 12, pp. 11492–11504, 2020.
[6] C.-F. Juang, and W.-E. Ni, “Human posture classification using interpretable 3-D fuzzy body voxel features and hierarchical fuzzy classifiers,” IEEE Trans. Fuzzy Syst., vol. 30, no. 12, pp. 5405–5418, 2022.
[7] K. Bourahmoune, K. Ishac, and T. Amagasa, “Intelligent posture training: machine-learning-powered human sitting posture recognition based on a pressure-sensing IoT cushion,” Sensors, vol. 22, no. 14, 5337, 2022.
[8] M. F. R. Lee, Y. C. Chen, and C. Y. Tsai, “Deep learning-based human body posture recognition and tracking for unmanned aerial vehicles,” Processes, vol. 10, no. 11, pp. 1–22, 2022.
[9] K.A. Lamkin-Kennard, and M.B. Popovic, “Sensors: Natural and Synthetic Sensors,” Biomechatronics, pp. 81-107, 2019.
[10] S. Bahassine, M. Kissi, and A. Madani, “New stemming for Arabic text classification using feature selection and decision trees,” In Proceedings of the 5th International Conference on Arabic Language Processing (CITALA) Oujda, Morocco, pp. 200–205, 2014.
[11] D. Karmaker, A. Chowdhury, M. Miah, M. Imran, and M. Rahman, “Cricket shot classification using motion vector,” In Proceedings of the 2nd International Conference on Computing Technology and Information Management (ICCTIM), Johor, Malaysia, 21–23 April 2015, pp. 125–129.
[12] R. Mili, N. R. Das, A. Tandon, S. Mokhtar, I. Mukherjee, and G. Paul, “Pose Recognition in Cricket using Keypoints,” 2022 IEEE 9th Uttar Pradesh Section International Conference on Electrical Electronics and Computer Engineering (UPCON), 2-4 December 2022, pp. 1-5.
[13] M. Z. Khan, M.A. Hassan, A. Farooq, and M. U. G. Khan, “Deep CNN based data-driven recognition of cricket batting shots,” In Proceedings of the International Conference on Applied and Engineering Mathematics (ICAEM), Taxila, Pakistan, 4–5 September 2018, pp. 67–71.
[14] Z. Liu, and X. Wang, “Action recognition for sports combined training based on wearable sensor technology and SVM prediction,” Preventive Medicine, vol. 173, 107582, 2023.
[15] G. Srilakshmi, and I. R. Praveen Joe, "Sports video retrieval and classification using focus u-net based squeeze excitation and residual mapping deep learning model," Engineering Applications of Artificial Intelligence, vol. 126, Part D, 107176, 2023.

,

ISBN 979-8-3503-5094-4

Survival Prediction of Cirrhosis Patients Using Polynomial Features and Differential Evolution

Hemprasad Yashwant Patil
School of Electronics Engineering
Vellore Institute of Technology
Vellore, Tamil Nadu, India
hemprasad.patil@vit.ac.in

Priyanka Bhagwan Patil
School of Computer Science and Engineering
Vellore Institute of Technology
Vellore, Tamil Nadu, India
priyankagsvhpatil@gmail.com

Lackesh Shanmugasundaram
School of Electronics Engineering
Vellore Institute of Technology
Vellore, Tamil Nadu, India
lackesh.shanmugasundaram@gmail.com

Anishkaa Balasubramaniam
School of Electronics Engineering
Vellore Institute of Technology
Vellore, Tamil Nadu, India
anishkaa.balasubramaniam@gmail.com

***Abstract*— This research focuses on enhancing the accuracy of the Cirrhosis Patient Survival Prediction dataset from the UCI-MLR. Leveraging novel techniques such as feature engineering and hyperparameter tuning, the study employs a comprehensive dataset with attributes including drug administration, age, sex, bilirubin, cholesterol, and albumin levels, along with labels indicating the stage of liver cirrhosis. The chosen dataset, recognized for its reliability and widespread adoption in research, serves as a valuable benchmark for developing diagnostic models. The comparative analysis involves Gradient Boosting Classifier, Logistic Regression, K Nearest Neighbors Classifier, Decision Tree Classifier, Gaussian Naive Bayes Classifier, Linear Discriminant Analysis, and Support Vector Classifier. Further, hyperparameter tuning through Grid Search and Differential Evolution methods results in substantial accuracy improvements, particularly with the latter yielding a remarkable 9% increase with respect to the control dataset with no alterations. This work showcases the potential of advanced methodologies in optimizing cirrhosis prediction models, emphasizing the importance of feature engineering and hyperparameter tuning in achieving significant performance enhancements.**

Keywords— ***Cirrhosis, Hyperparameter tuning, Differential Evolution, Survival Prediction, Polynomial Feature Engineering***

I. INTRODUCTION

Multiple types of ailments, such as Parkinson's' disease [1], [2], cancers [3], [4], diabetes [5], liver-related conditions [6], [7], etc., can profoundly influence an individual's overall state of well-being. The severe liver illness referred to as cirrhosis causes the organ to scar and lose some of its ability to operate normally. Fibrosis represents the result of scarring that interferes with the liver's regular structure and makes it difficult to do important functions including but not limited to combating illness, removing harmful substances from the bloodstream, dietary as well as glucose storage, and creating gastrointestinal as well as blood-clotting proteins [6], [8], [9]. Histology-wise, cirrhosis develops when the liver responds to long-lasting harm by creating recovering nodules encircled with fibrous zones; this, in turn, causes portal hypertension as well as advanced liver ailment. It is common practice to incorporate the health record of the individual alongside serologic as well as histologic testing in order to determine the cause of cirrhosis. In the West, alcohol-related liver conditions and the virus hepatitis C remain among the prevalent reasons for liver infection. Still, in Asian countries, hepatitis B has become the most prevalent form of liver infection. Cirrhosis often progresses slowly, without causing symptoms, and goes unnoticed unless issues related to liver damage arise. A significant part of such individuals remain unnoticed by healthcare workers, and cirrhosis that was not previously identified is often discovered after postmortem examinations. Atypical cirrhosis is often diagnosed after regular diagnostic examinations, including liver transaminases as well as radiologic abnormalities, indicate a condition of the liver. Additional examination including liver biopsy are then conducted to confirm the medical condition. Ultrasound imaging, CAT, and MR imaging lack sensitivity for identifying cirrhosis. Therefore, the definitive identification largely depends on histology [9]. Nevertheless, their level of validity is elevated when there is a clear root cause while scanning shows an uneven texture especially on exterior of the liver, a dilated central vein, an expanded spleen, or the presence of collateral blood vessels [9]. Hence, the medical data contains several aspects that are beneficial for both evaluation and prognosis.

Artificial intelligence-based methodologies have become essential instruments in the medical domain, transforming the procedure for diagnosing and greatly improving the precision

ISBN 979-8-3503-5094-4

and speed of ailments detection. Their capacity to sift through masses of data, comprising biological details, clinical archives, and imaging data, in search of relationships and trends that could pass unnoticed by humans is a major asset. In a short amount of time, such methods can sort over massive datasets, identify intricate correlations, and modify their hypotheses according to fresh data. The intricate and ever-changing healthcare field is a perfect fit for models developed using machine learning because of their exceptional performance in dealing with the inevitable ambiguity and unpredictability of medical information. These machines may adjust to new medical information and continue to develop because of their perpetual learning capabilities.

Consequently, screening techniques powered by machine learning might help doctors spot illnesses earlier and more accurately, which could mean better treatment for patients in the long run. One cutting-edge method in medical care that has recently arisen is patient survival prediction, which utilizes artificial intelligence. These mathematical models may provide very precise survival probability forecasts by using sophisticated methods to analyze a plethora of patient-related information, such as health records, diagnostic results, and therapy reactions. Medical practitioners can allocate resources and develop interventions with greater exactness because of AI's capacity to detect complex associations and trends in massive datasets. In addition, such methodology for predictive modeling helps establish the early detection of individuals who are at elevated risk. Additionally, this makes it possible to implement therapies that are individualized and focused. The models themselves foster the improvement of their projections periodically by continually developing according to all the fresh information about patients that they receive. In addition to enhancing medical judgment, such developments in AI offer the ability to substantially boost the way patients are treated by allowing physicians to adapt therapies in accordance with personalized survival probability. This has the capacity to dramatically improve patient's individual welfare to a large degree.

This paper aims to increase the classification accuracy of the Cirrhosis Patient Survival Prediction through means of novel techniques such as new feature engineering and hyperparameter tuning. The dataset is highly comprehensive in nature with attributes such as drug administered, age, sex, bilirubin, cholesterol, albumin levels, and labels specifying the stage of liver cirrhosis. Due to its dependability, widespread adoption in research, and accessibility to the public, this dataset serves as a valuable standard for creating and assessing diagnostic models for liver cirrhosis, thus substantiating our utilization of the dataset.

This research paper is structured as follows: Section II delves into an extended discussion of prior literature related to cirrhosis identification. Section III elucidates the methodology, encompassing information about the data used, pre-processing methods, framework selection, data processing approaches, and workflow. Section VI presents the outcomes and accuracy comparison of the presented strategy. The conclusion provides a thorough analysis and highlights the adaptability of the proposed methodology to diverse datasets. .

II. Related Works

Wang et al. [10] utilized gene expression dataset to analyze the medical dataset for cirrhosis survival prediction. They utilized the unlabeled cluster evaluation to obtain the gene dataset. Cox regression was used to investigate the variations in predictive outcomes that were observed amongst multiple clusters. For each of the distinct groups, a comprehensive study of the gene transcriptional profile, immune cell stimulation, and clinical differentiation was carried out by the authors. It was determined that the network formed by PPI had 5 hub genes as well as 3 important modules, which were retrieved. The findings underwent validation in the GSE14520 group, which was associated to HBV.

Gibb et al. [11] proposed the utilization of machine learning for survival prediction of last-stage patients suffering from cirrhosis. The authors conducted an inspection of the medical and assay information of 654 people whose names were enrolled in the liver replacement assessment procedure. By conducting a comprehensive analysis of 13 distinct AI methods using an intricate cross-validation approach, the authors identified the most effective option. Subsequently, they developed a novel model that accurately forecasts the likelihood of 90-day death in individuals suffering from last-stage liver ailment. The penalized regression methods demonstrated superior forecasting accuracy against the set benchmark of AI methods.

Dickson et al. [12] stated that an optimal computational framework for forecasting the survival of every individual with cirrhosis ought to depend on a limited set of affordable and unobtrusive data that are widely accessible. This approach could prove beneficial in medical administration by assisting in the identification and scheduling of individuals awaiting liver transplantation. A framework was constructed using the Cox regression approach and substantial information for 312 individuals with cirrhosis at the Mayo Clinic. The Cox model postulates a linear association amongst the variables as well as the logarithm of the hazard, which is one of the limitations of this work.

Chang et al. [13] utilized the Multi-center North American Association in the Investigation of End-Stage Liver Disease cohort, which consisted of individuals with cirrhosis who were hospitalized. Probabilistic algorithms were employed to identify relevant factors for predicting readmission along with death. Subsequently, three ML algorithms involving kernel support vector machine, random forest, and logistic regression were developed towards the purpose of predicting readmission and mortality. The authors concluded that predicting readmissions as well as mortality with cirrhosis remains challenging, even when using various AI algorithms. The machine learning model's accuracies remained comparable to the accuracies of predictions produced solely based on MELD-Na ratings.

Kanwal et al. [14] evaluated several machine learning techniques to forecast fatalities in general in cirrhosis and also employ AI to identify readily quantifiable medical parameters with a new predictive framework for cirrhosis. Possible indicators encompassed socioeconomic factors, the cause,

ISBN 979-8-3503-5094-4

seriousness, and consequences of liver ailment, utilization of medicine, concurrent medical disorders, and extensive analytical and prescription information. An arbitrary choice of individuals was made for the purpose of developing the model (2/3 of the data) and validating it (1/3 of the data). The evaluation involved three distinct techniques: logistic regression, gradient descent boosting, and logistic regression.

III. Proposed Solution

A. Dataset

The first stage towards prognosis is gathering the data. We have collected the Cirrhosis patient survival prediction dataset from the UCI repository [15], [16]. The UCI Machine Learning Repository represents a prominent compilation of resources curated by the University of California, Irvine. The collection has a diverse range of information that covers several fields, including life sciences, economics, medicine, and the humanities. From 1974 to the next ten years, a total of 424 individuals diagnosed with primary biliary cholangitis were included in an unbiased placebo-administered study conducted by the Mayo Clinic to evaluate the effectiveness of the medication D-penicillamine.

B. Data pre-processing

a) Dropping rows and columns

The rows with missing values are dropped, followed by the ID column. This results in the number of instances being reduced from 418 to 276.

b) Encoding categorical values as numerical values

Unique values from each of the remaining columns are then extracted. The unique values are encoded as categorical numerical values. In the target column - status, D is encoded as 0, C as 1 and CL as 2 [17]. For all other columns, the column values only require binary encoding. If the column contains *N* categorical samples, label *x* representing a unique string (category) is allotted as per the following pseudo-code:

```
for i = 0 to N-1
Encoded_Label(x) = Integer_representing_category_x
Where x is a set of unique numbers
End
```

c) Normalization

All columns with numerical values (excluding the ones that were encoded) are normalized to transform the range of features to a standard scale. The uniformity of values after normalization avoids the domination of features with large values compared to other features [17], [18], [26]. For example, the normalization of the column named 'Albumin' is performed using (1).

$$d['Albumin'] = \frac{d['Albumin'] - d['Albumin'].min()}{d['Albumin'].max() - d['Albumin'].min()} \quad (1)$$

C. Model Selection and data processing

After declaring the features (independent variables), and the labels (dependent variables), the data collection is split into training and testing records in an 80:20 ratio. The accuracy is evaluated and then plotted for each of the 7 classifiers used, namely Gradient Boosting [19], Logistic Regression [20], Support Vector Machine [21], K-Nearest Neighbors [22], Decision tree [23], Gaussian Naive Bayes [24], and Linear Discriminant Analysis [25], [26].

a) Feature Engineering

The feature and target arrays are converted back to dataframes, and a new feature is engineered by applying an operation on the 'N_days' and 'Bilirubin' feature columns, as stated below. Both columns are squared to model the non-linear relationships between these features and the target. Furthermore, 'N-Days' is scaled by a factor of 2.5 and the new feature is offset by a value of 1. The new feature was engineered based on the idea that patients with different 'N_Days' values and 'Bilirubin' counts could correlate to a different prognosis.

$$d['New\_Column'] = \frac{5}{2} * d['N\_days']^2 + d['Bilirubin']^2 + 1$$

After normalizing the new column, it is added to the existing dataset and the accuracy score is calculated. Following this. Hyperparameter tuning is carried out using GridSearch [27] with parameters 'c' and 'gamma'. Hyperparameters are configurable parameters that govern the training mechanism of the model. Theseparameters include the learning rate, the number of neurons in a neural network, or the kernel size in a support vector machine. Essentially, they guide the learning process and impact the model's performance.These parameters are usually fixed before the model is trained. Hyperparameter tuning is a heuristic process to find values that lead to themost optimal performance of the model. This process can be manual or automated [28].

The hyperparameters of the SVC that are optimized in this study are:

- C: the regularization parameter, where the strength of regularization is inversely proportional to C.
- Gamma: the kernel coefficient for 'rbf', 'poly' and 'sigmoid' kernels is determined by the following rules: If gamma='scale' (which is the default), the value of gamma is set to 1 / (number of features * X.var()). If gamma='auto', it uses 1 / n_features as the value of gamma.

Hyperparameter tuning is carried out in two different ways in this paper and the results are compared.

1. GridSearch: A module (also known as parameter sweep) in the scikit-learn library that employs an complete exploration through a manually designated subgroup of the hyperparameter space of an ML algorithm. It takes an estimator (which can be a regressor or classifier), a configuration space, a model

ISBN 979-8-3503-5094-4

evaluation scheme, and a score function to generate candidates from a grid of parameter values defined within the 'param_grid' parameter. The GridSearchCV implements all probable combinations of parameter values and retains the most optimal combination [27].

2. Differential Evolution: This is a method in the Scipy library that optimizes the classification problem by iteratively trying to refine the potential solution based on a specific quality criterion. Differential Evolution (DE) streamlines an instance by managing an assortment of potential solutions and generating novel solutions by integrating old ones based on a specific criterion as per its straightforward formulae and then retaining whichever potential solution has the best score or fitness outcome on the optimization problem [29], [30].

The SVC model is then evaluated for accuracy after hyperparameter tuning using GridSearch and differential evolution separately.

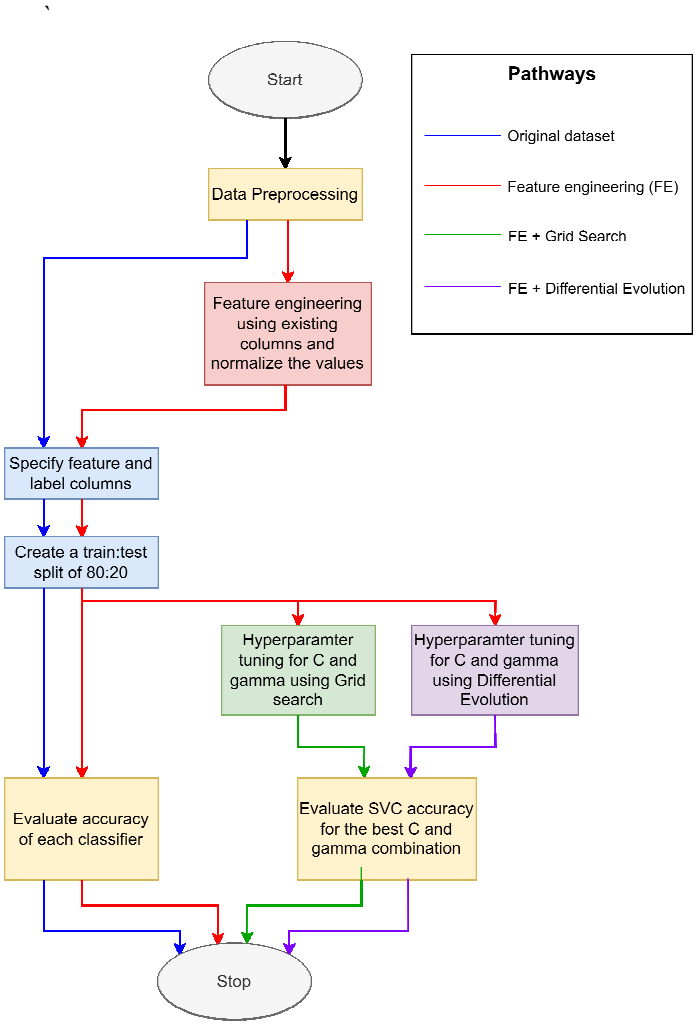


Fig. 1. Flow Diagram of the proposed Methodology

Fig. 1. shows the different pathways followed in this research. The first approach evaluates the accuracy of each of the seven classifiers directly after the train-test split. The first approach is treated as the 'control' method with no modifications to the dataset. In the second approach, a new feature is engineered with the help of two existing columns ('N_days', 'Bilirubin'), and the accuracy score is re-evaluated. In the third approach, following feature engineering, hyperparameter tuning for C and gamma of the SVC using the GridSearch mechanism is carried out. In the fourth and final approach, after feature engineering, hyperparameter tuning using a differential evolution technique is carried out. Afterward, the accuracy factor is evaluated for the best C and gamma parameters.

IV. Results and Discussion

The proposed work compares the performance of the Gradient Boosting Classifier, Logistic Regression, K Nearest Neighbors Classifier, Decision Tree Classifier, Gaussian Naive Bayes (NB) Classifier, Linear Discriminant Analysis, and Support Vector Classifier (SVC) on the Cirrhosis patient survival prediction dataset from UCI-MLR. It further evaluates the impact of feature engineering and using hyperparameter tuning on the dataset through Grid Search and Differential Evolution methods.

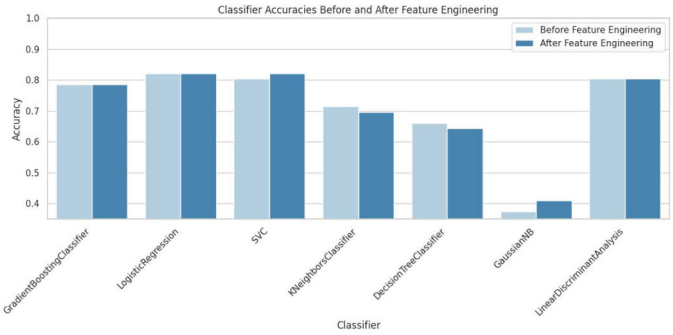


Fig. 2. Classifier accuracies before and after feature engineering

Classifier	Before Feature Engineering	After Feature Engineering
Gradient Boosting Classifier	0.7857	0.7679
Logistic Regression	0.8214	0.8214
Support Vector Classifier	0.8036	0.8214
K Neighbors Classifier	0.7143	0.6964
Decision Tree Classifier	0.7500	0.6607
Gaussian Naive Bayes	0.3750	0.4107
Linear Discriminant Analysis	0.8036	0.8036

Table I: Accuracy scores of the presented classifiers

Accuracy Enhancement Approach	Accuracy
Original	0.8036
After FE	0.8214

ISBN 979-8-3503-5094-4

After FE + GridSearch	0.8400
After FE + Differential Evolution	0.8900

Table II: Accuracy score of SVC after different methods of enhancement

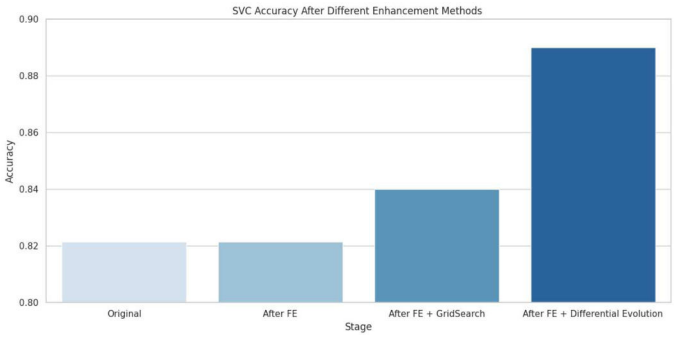


Fig. 3. SVC accuracies after different approaches

Fig. 2. and Table I show the comparative performance of each classifier before and after feature engineering is performed. SVC exhibits an increase in accuracy with the implementation of feature engineering from 0.8036 to 0.8214, while Linear Discriminant Classifier and Logistic Regression indicate similar accuracy values. The other classifiers display a decrease in accuracy after the implementation of feature engineering.

With the application of Grid Search-based hyperparameter tuning for SVC in the feature-engineered dataset, the model records an accuracy of 0.84, which is a 2% leap from the previous recorded accuracy. A differential evolution-based hyperparameter tuning approach exhibits the best performance out of the approaches proposed with an accuracy score of 0.89, which is a 9% increase in comparison to the original dataset, as indicated in Table II and Figure 3.

V. Conclusion

The proposed solution utilizes new feature engineering through existing features to improve the accuracy of the Support Vector Classifier. A combination of N_days and Bilirubin in a polynomial relation increases the accuracy of SVC by 2%. Furthermore, hyperparameter tuning using Grid Search methodology and Differential Evolution technique is carried out on the feature-engineered dataset. Differential Evolution outperforms Grid Search due to its higher resolution or step size, enabling it to tune hyperparameters more efficiently. Differential evolution with feature engineering using SVC records a 9% increase in accuracy in comparison to the control dataset with no alterations. Future work could include the empirical validation of the engineered feature using a larger validation set to further strengthen the model in making predcitions.

References

[1] P. K. Chawla *et al.*, "Parkinson's disease classification using nature inspired feature selection and recursive feature elimination," *Multimed. Tools Appl.*, 2023, doi: 10.1007/s11042-023-16804-w.

[2] K. u. V. R. teja, B. P. V. Reddy, L. P. Alla, and H. Y. Patil, "Parkinson's Disease Classification using Quantile Transformation and RFE," in *2021 12th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, 2021, pp. 1–5. doi: 10.1109/ICCCNT51525.2021.9580024.

[3] S. B. Mukadam and H. Y. Patil, "Skin Cancer Classification Framework Using Enhanced Super Resolution Generative Adversarial Network and Custom Convolutional Neural Network," *Applied Sciences*, vol. 13, no. 2. 2023. doi: 10.3390/app13021210.

[4] J. V Tembhurne, N. Hebbar, H. Y. Patil, and T. Diwan, "Skin cancer detection using ensemble of machine learning and deep learning techniques," *Multimed. Tools Appl.*, vol. 82, no. 18, pp. 27501–27524, 2023, doi: 10.1007/s11042-023-14697-3.

[5] K. U. V. R. Teja, B. P. V Reddy, L. P. A, H. Y. Patil, and P. C. T, "Prediction of Diabetes at Early Stage with Supplementary Polynomial Features," in *2021 Smart Technologies, Communication and Robotics (STCR)*, 2021, pp. 1–5. doi: 10.1109/STCR51658.2021.9588849.

[6] M. Pinzani, M. Rosselli, and M. Zuckermann, "Liver cirrhosis," *Best Pract. Res. Clin. Gastroenterol.*, vol. 25, no. 2, pp. 281–290, 2011, doi: https://doi.org/10.1016/j.bpg.2011.02.009.

[7] S. P. Starr and D. Raines, "Cirrhosis: diagnosis, management, and prevention," *Am. Fam. Physician*, vol. 84, no. 12, pp. 1353–1359, 2011.

[8] E. A. Tsochatzis, J. Bosch, and A. K. Burroughs, "Liver cirrhosis," *Lancet*, vol. 383, no. 9930, pp. 1749–1761, 2014.M. Young, The Technical Writer's Handbook. Mill Valley, CA: University Science, 1989.

[9] D. Schuppan and N. H. Afdhal, "Liver cirrhosis," *Lancet*, vol. 371, no. 9615, pp. 838–851, 2008.

[10] Q. Wang, X. Li, Y. Chen, J. Gong, and B. Hu, "Classification and survival prediction in early-stage cirrhosis by gene expression profiling," *J. Viral Hepat.*, vol. 30, no. 2, pp. 116–128, Feb. 2023, doi: https://doi.org/10.1111/jvh.13769.

[11] S. Gibb, T. Berg, A. Herber, B. Isermann, and T. Kaiser, "A new machine-learning-based prediction of survival in patients with end-stage liver disease," vol. 47, no. 1, pp. 13–21, 2023, doi: doi:10.1515/labmed-2022-0162.

[12] E. R. Dickson, P. M. Grambsch, T. R. Fleming, L. D. Fisher, and A. Langworthy, "Prognosis in primary biliary cirrhosis: model for decision making.," *Hepatology*, vol. 10, no. 1, pp. 1–7, Jul. 1989, doi: 10.1002/hep.1840100102.

[13] C. Hu *et al.*, "Low Predictability of Readmissions and Death Using Machine Learning in Cirrhosis," *Off. J. Am. Coll. Gastroenterol. | ACG*, vol. 116, no. 2, 2021, [Online]. Available: https://journals.lww.com/ajg/fulltext/2021/02000/low_predictability_of_readmissions_and_death_using.25.aspx

[14] F. Kanwal *et al.*, "Development, Validation, and Evaluation of a Simple Machine Learning Model to Predict Cirrhosis Mortality," *JAMA Netw. Open*, vol. 3, no. 11, pp. e2023780–e2023780, Nov. 2020, doi: 10.1001/jamanetworkopen.2020.23780.

[15] A. Dickson,E., Grambsch,P., Fleming,T., Fisher,L., and Langworthy, "Cirrhosis Patient Survival Prediction." UCI Machine Learning Repository, 2023. doi: https://doi.org/10.24432/C5R02G.

[16] T. R. Fleming and D. P. Harrington, *Counting Processes and Survival Analysis*. in Wiley Series in Probability and Statistics. Wiley, 2011. [Online]. Available: https://books.google.co.in/books?id=Sqg-YPcpzLYC

[17] P. Joshi, *Artificial intelligence with python*. Packt Publishing Ltd, 2017.

[18] G. Hackeling, *Mastering Machine Learning with scikit-learn*. Packt Publishing Ltd, 2017.

[19] A. Natekin and A. Knoll, "Gradient boosting machines, a tutorial," *Front. Neurorobot.*, vol. 7, p. 21, 2013.

ISBN 979-8-3503-5094-4

[20] D. Böhning, "Multinomial logistic regression algorithm," *Ann. Inst. Stat. Math.*, vol. 44, no. 1, pp. 197–200, 1992.

[21] C. Cortes and V. Vapnik, "Support-vector networks," *Mach. Learn.*, vol. 20, pp. 273–297, 1995.

[22] K. Taunk, S. De, S. Verma, and A. Swetapadma, "A Brief Review of Nearest Neighbor Algorithm for Learning and Classification," in *2019 International Conference on Intelligent Computing and Control Systems (ICCS)*, 2019, pp. 1255–1260. doi: 10.1109/ICCS45141.2019.9065747.

[23] B. Charbuty and A. Abdulazeez, "Classification based on decision tree algorithm for machine learning," *J. Appl. Sci. Technol. Trends*, vol. 2, no. 01, pp. 20–28, 2021.

[24] A. H. Jahromi and M. Taheri, "A non-parametric mixture of Gaussian naive Bayes classifiers based on local independent features," in *2017 Artificial intelligence and signal processing conference (AISP)*, IEEE, 2017, pp. 209–212.

[25] S. Balakrishnama and A. Ganapathiraju, "Linear discriminant analysis-a brief tutorial," *Inst. Signal Inf. Process.*, vol. 18, no. 1998, pp. 1–8, 1998.

[26] F. Pedregosa *et al.*, "Scikit-learn: Machine Learning in Python," *J. Mach. Learn. Res.*, vol. 12, Jan. 2012.

[27] G. N. Ahmad, H. Fatima, S. Ullah, and A. S. Saidi, "Efficient medical diagnosis of human heart diseases using machine learning techniques with and without GridSearchCV," *IEEE Access*, vol. 10, pp. 80151–80173, 2022.

[28] R. Bardenet, M. Brendel, B. Kégl, and M. Sebag, "Collaborative hyperparameter tuning," in *International conference on machine learning*, PMLR, 2013, pp. 199–207.

[29] R. Storn and K. Price, "Differential Evolution – A Simple and Efficient Heuristic for global Optimization over Continuous Spaces," *J. Glob. Optim.*, vol. 11, no. 4, pp. 341–359, 1997, doi: 10.1023/A:1008202821328.

[30] J. I. Qiang, "A unified differential evolution algorithm for global optimization," 2014.

ISBN 979-8-3503-5094-4

Effect of Motion Sickness Drugs [Cinnarizine, Dimenhydrinate, Promethazine] on Gut Bacteria [*E. coli*]

Dr. Subashini R
Department of Biomedical Engineering
Sri Sivasubramaniya Nadar College of Engineering
Chennai, India
subashinir@ssn.edu.in

Shreya M
Department of Biomedical Engineering
Sri Sivasubramaniya Nadar College of Engineering
Chennai, India
shreya2010392@ssn.edu.in

Srivibha Parthasarathy
Department of Biomedical Engineering
Sri Sivasubramaniya Nadar College of Engineering
Chennai, India
srivibha2010270@ssn.edu.in

Supraja Vaidhyanathan
Department of Biomedical Engineering
Sri Sivasubramaniya Nadar College of Engineering
Chennai, India
supraja2010163@ssn.edu.in

***Abstract* - Motion sickness, a common ailment stemming from a discordance between signals received by the brain from the vestibular system and visual cues, is often managed with oral medications primarily targeting nausea and vomiting. However, the effects of these drugs on the gut microbiota remain largely unexplored. This study aims to investigate the impact of motion sickness drugs, specifically Cinnarizine, Dimenhydrinate, and Promethazine, on *Escherichia coli*, a prevalent bacterium in the human colon. The research involves conducting experiments in stable conditions to assess the drugs' efficacy under different gravitational forces. Growth rate analysis of *E. coli* under normal conditions will be compared with drug exposure, with deviations indicating potential drug effects. Findings from this research will contribute to understanding the motion sickness drugs resistance and their effects on gut microbiota, potentially informing future treatment strategies.**

Keywords: Motion sickness, gut microbiota, Escherichia coli, Cinnarizine, Dimenhydrinate, Promethazine, gravitational forces, sedative effects, bacterial growth, drug efficacy.

I. INTRODUCTION

Motion sickness is a diverse syndrome that includes many symptoms besides nausea and vomiting. Usually, it goes unrecognized in many circumstances. It can occur during physical, visual and virtual motion, and doesn't affect those without a functioning vestibular system. 1 in 10,000 are vulnerable to motion sickness. Additionally, the contribution of visceral afferents in relation to vestibular and cerebellar signals in triggering sickness has to be explored further. Research is going on to identify the physiological mechanisms causing nausea, vomiting, and anxiety, and subsequently a comprehensive understanding of motion sickness may soon be attained. However, sufficient anti-motion sickness drugs without adverse side effects are not yet available.

Experimental studies have shown that, if anyone with a normal vestibular function is exposed to intense physical motion, then the disruption of their vestibulo-ocular reflexes, or optokinetic stimulation can lead to motion sickness [1]. These are the category of people who are susceptible to this sickness, which encompasses a vast number of the human population.

The literature survey provides insights into several pertinent areas related to the gut microbiota, pharmacology, and motion sickness. Chen et al. (2022) underscore the intricate relationship between drugs and gut microbiota, emphasizing the need for personalized approaches in treatment [2]. Ramirez et al. highlight the disruptive impact of antibiotics on gut microbiota, implicating it in various health issues [3]. Yates et al. delve into the physiological and pharmacological aspects of motion sickness, indicating gaps in understanding its complexities [4]. Lim et al. explore advancements in microbial biosensors, addressing challenges in selectivity and sensitivity.

However, there remains a gap in understanding how motion sickness drugs impact gut microbiota and vice versa. This study intends to bridge the gap by investigating the effect of motion sickness drugs on the growth rate of E. coli, thereby elucidating

ISBN 979-8-3503-5094-4

their impact on gut bacteria. By comparing the inhibitory effects of different drugs, namely Cinnarizine, Dimenhydrinate, and Promethazine [5], this research aims to contribute to a deeper understanding of the interaction between pharmacology, gut microbiota, and motion sickness, thereby offering insights for potential therapeutic interventions.

II. METHODOLOGY

A. *Experimental Setup*

The experimental set up consists of 3 petri dishes and a Sensor Cube with the petri dishes stacked together and placed above the Sensor Cube. The petri dishes have rubber gaskets to ensure airtight locking. The Sensor Cube contains temperature, pH, and humidity sensors, a bluetooth module, and an Arduino. The sensors help us monitor the maintenance of a human-like environment, Arduino processes the information from the sensors, and the flash drive stores the data collected from them [6]. The bacterium and drugs in the petri dish are cultured in a behavior Mueller-Hinton agar gel medium. Filter paper discs are impregnated with the drugs, both separately and in various combinations, and placed on the *E. coli* culture to inspect the susceptibility of the organism to the drugs [7].

B. *Constituents of the Petri dishes*

- Control + *E. coli*
- Cinnarizine + *E. coli*
- Dimenhydrinate + *E. coli*
- Promethazine + *E. coli*
- Cinnarizine + Dimenhydrinate + *E. coli*
- Dimenhydrinate + Promethazine + *E. coli*
- Cinnarizine + Promethazine + *E. coli*
- Cinnarizine + Dimenhydrinate + Promethazine + *E. coli*

C. *Electronics*

The Sensor Cube electronics consists of Arduino to process the information collected from the sensors installed. Temperature, pH (manual), and humidity sensors, to maintain a human-like atmosphere are included in this compartment. A flash drive is included onboard to recover and store data collected from these sensors.

a) *Temperature and Humidity sensor:* The KY-015 module consists of a digital serial interface to measure atmospheric humidity and temperature [8]. It includes of a DHT11 digital humidity and temperature sensor, a 1 kΩ resistor and 3 male header pins. The sensor uses an internal thermistor and a capacitive humidity sensor to determine the environmental conditions.

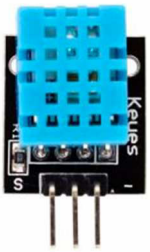


Fig. 1. KY-015 sensor

b) *Pressure sensor:* The BMP180 is a piezo resistive sensor that measures pressure [9]. It senses both pressure and temperature, because density of gasses depends on temperature changes. It utilizes real-time temperature measurements to compute pressure readings.

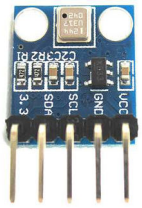


Fig. 2. BMP180 sensor

c) *Bluetooth module:* HC-05 is a full-duplex Bluetooth module which has bidirectional communication [10]. It uses USART (Universal Synchronous/ Asynchronous Receiver/ Transmitter) at 9600 baud rate.

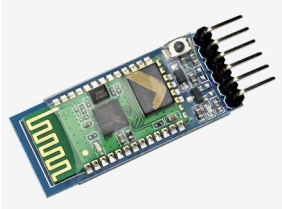


Fig 3. HC05 module

d) *Arduino UNO:* Arduino, an open-source electronic platform, is popular due to its straightforward hardware and user-friendly software. In Arduino boards, the user will utilize the software to communicate a set of instructions to the microcontroller. It is processed using the Arduino software (IDE), Arduino code, or Arduino programming language. Due to its simplicity and user-friendly accessibility, Arduino UNO was chosen.

ISBN 979-8-3503-5094-4

Fig. 4. Arduino UNO

D. Work Plan

The mechanism of action of the drugs, separately, and in combination, of Cinnarizine, Dimenhydrinate, and Promethazine on *Escherichia Coli* will be recorded. Simultaneously, behavior of *E. coli* will also be studied and monitored, such as the growth rate of bacteria using measurement of diameter of inhibited area, and measuring external factors such as temperature variations, pH variations, pressure changes, humidity, etc., under the same experimental setup. Then, the efficiency of these drugs under standardized conditions are compared and results are published.

Kirby Bauer Test

Kirby Bauer Test is the most widely used antibiotic susceptibility test. This involves disk-diffusion method to test the combination of antibiotics to be used to treat infections. It is based on the principle of inhibition of bacterial growth under standard conditions. To perform this test, the Mueller-Hinton agar culture medium, is uniformly and aseptically inoculated with the test organism. Then filter paper discs, impregnated with a particular concentration of test antibiotic, are placed on the medium. The organism begins to multiply on the agar plate while the antibiotic attempts to inhibit its growth. If the organism is found to be susceptible to a specific antibiotic, no growth will be observed around that impregnated disc. This forms a zone of inhibition which can be measured to determine the antibiotic susceptibility for that particular organism. The measurement is compared to the Clinical and Laboratory Standards Institute (CLSI) standards and further classified as Resistant (R), Intermediate (I) or Susceptible (S).

III. RESULTS AND DISCUSSION

The mechanism of action of the drugs, separately, and in various combinations, of Cinnarizine, Dimenhydrinate, and Promethazine on *Escherichia Coli* under laboratory conditions were studied. The Kirby Bauer Method was used to observe the zone of inhibition in order to determine the susceptibility of *E. coli* to the drug.

Four discs were placed in each of two petri dishes and two discs in the third petri dish. The control was placed in all the petri dishes. One petri dish contained the discs impregnated with the drugs separately i.e., Promethazine, Dimenhydrinate and Cinnarizine, while the other petri dish contained the discs impregnated with the drugs, taken two at a time, i.e., Promethazine and Dimenhydrinate, Dimenhydrinate and Cinnarizine, and Promethazine and Cinnarizine. The third petri dish contained a disc for control and another disc impregnated with a combination of all the three drugs.

It could be observed that the zone of inhibition for the control was 3.5 cm in all the petri dishes.

Promethazine was observed to have two zones of inhibition - the first zone with a diameter of 1.4 cm and the second zone with a diameter of 2.6 cm. Cinnarizine was found to have characteristic growth of organisms at a distance of 1.5 cm around the disc. The zone of inhibition for Dimenhydrinate was not as distinct as it was for the other two discs.

The zone of inhibition when the drugs were used in combination of two drugs at a time could be faintly observed. Dimenhydrinate + Promethazine had the maximum zone of inhibition of 2.7 cm diameter while Promethazine + Cinnarizine had the minimum zone of inhibition of 2 cm, when various combinations of the drugs together were considered.

The zone of inhibition when the disc was impregnated with the combination of all three drugs was 2.4 cm in diameter.

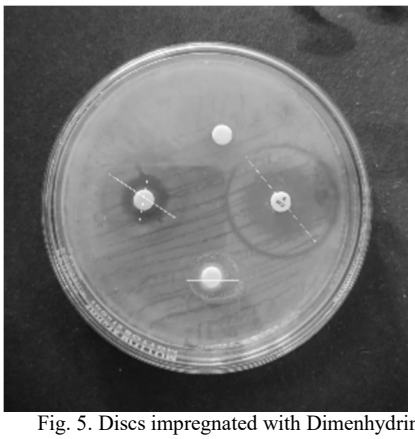


Fig. 5. Discs impregnated with Dimenhydrinate, Control, Cinnarizine, Promethazine. (from top, clockwise)

ISBN 979-8-3503-5094-4

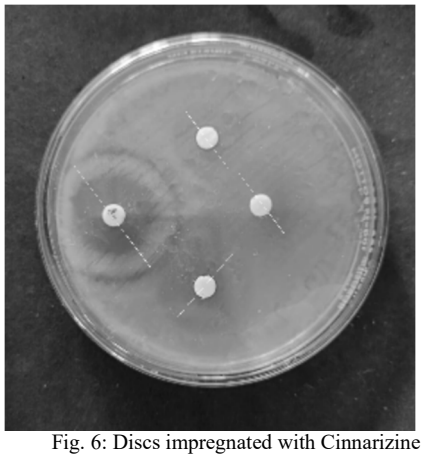


Fig. 6: Discs impregnated with Cinnarizine + Promethazine, Dimenhydrinate + Cinnarizine, Promethazine + Dimenhydrinate, Control (from top, clockwise)

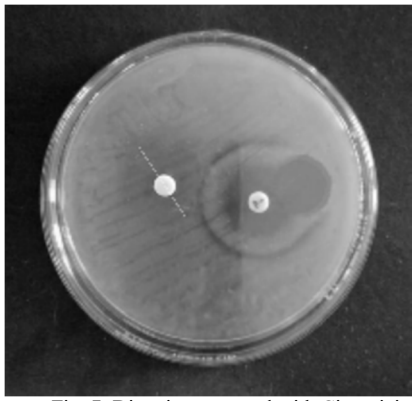


Fig. 7: Discs impregnated with Cinnarizine + Promethazine + Dimenhydrinate and Control (left and right)

TABLE I. RESULTS OF DISC DIFFUSION METHOD

S. No.	Drug	Zone of Inhibition (in cm)
1.	Control	3.5
2.	Dimenhydrinate	2.5
3.	Promethazine	1st Zone - 1.4 2nd Zone - 2.6
4.	Cinnarizine	1.5
5.	Dimenhydrinate + Promethazine	2.7
6.	Promethazine + Cinnarizine	2
7.	Cinnarizine + Dimenhydrinate	2.4
8.	Dimenhydrinate + Promethazine + Cinnarizine	2.4

IV. CONCLUSION

In this study, the impact of motion sickness drugs—Cinnarizine, Dimenhydrinate, and Promethazine—on *Escherichia coli* (*E. coli*), a prevalent bacterium in the human colon, was investigated through experiments conducted in stable conditions under varying concentrations. The growth rate analysis of E. coli under normal conditions was compared with drug exposure, revealing distinct effects. The control group exhibited a consistent 3.5 cm zone of inhibition, providing a baseline. Individual drug assessments showed Promethazine's dual zones of inhibition, Cinnarizine's characteristic growth inhibition, and Dimenhydrinate's less distinct effect. Combinations of drugs demonstrated varying inhibitory effects, with Dimenhydrinate + Promethazine exhibiting the maximum inhibition (2.7 cm). These findings shed light on the potential impact of motion sickness drugs on gut microbiota, emphasizing the need for further research to elucidate specific mechanisms and implications for future treatment strategies, ultimately contributing to the understanding of the complex interaction between motion sickness drugs and bacterial growth in the gut.

REFERENCES

[1] Lackner, James R. “Motion sickness: more than nausea and vomiting.” *Experimental brain research* vol. 232,8 (2014): 2493-510. doi:10.1007/s00221-014-4008-8

[2] Chen HQ, Gong JY, Xing K, Liu MZ, Ren H, Luo JQ. Pharmacomicrobiomics: Exploiting the Drug-Microbiota Interactions in Antihypertensive Treatment. *Front Med (Lausanne)*. 2022;8:742394. Published 2022 Jan 19. doi:10.3389/fmed.2021.742394

[3] Ramirez J, Guarner F, Bustos Fernandez L, Maruy A, Sdepanian VL and Cohen H (2020) Antibiotics as Major Disruptors of Gut Microbiota. Front. Cell. Infect. Microbiol. 10:572912. doi: 10.3389/fcimb.2020.572912

[4] Yates BJ, Miller AD, Lucot JB. Physiological basis and pharmacology of motion sickness: an update. *Brain Res Bull*. 1998;47(5):395-406. doi:10.1016/s0361-9230(98)00092-6

[5] Karrim N, Byrne R, Magula N, Saman Y. Antihistamines for motion sickness. *Cochrane Database Syst Rev*. 2022;10(10):CD012715. Published 2022 Oct 17. doi:10.1002/14651858.CD012715.pub2

[6] Lim JW, Ha D, Lee J, Lee SK, Kim T. Review of micro/nanotechnologies for microbial biosensors. *Front Bioeng Biotechnol*. 2015;3:61. Published 2015 May 11. doi:10.3389/fbioe.2015.00061

[7] Kibret M, Abera B. Antimicrobial susceptibility patterns of E. coli from clinical sources in northeast Ethiopia. *Afr Health Sci*. 2011;11 Suppl 1(Suppl 1):S40-S45. doi:10.4314/ahs.v11i3.70069

[8] ArduinoModules, “KY-015 Temperature and Humidity Sensor Module,” *ArduinoModulesInfo*, Nov. 04, 2016. https://arduinomodules.info/ky-015-temperature-humidity-sensor-module/#google_vignette (accessed Mar. 10, 2024).

[9] K. P. | Arduino | 5, “How to Set up the BMP180 Barometric Pressure Sensor on an Arduino,” *Circuit Basics*, May 29, 2017. https://www.circuitbasics.com/set-bmp180-barometric-pressure-sensor-arduino/

[10] A. Joseph, “Interfacing Bluetooth Module (HC-05) with Arduino Uno,” *Hackster.io*, Apr. 19, 2020. https://www.hackster.io/akshayjoseph666/interfacing-bluetooth-module-hc-05-with-arduino-uno-f5209b

[11] J. Hudzicki, “Kirby-Bauer Disk Diffusion Susceptibility Test Protocol,” American Society for Microbiology, Dec. 2009. Available: https://asm.org/getattachment/2594ce26-bd44-47f6-8287-0657aa9185ad/Kirby-Bauer-Disk-Diffusion-Susceptibility-Test-Protocol-pdf.pdf

ISBN 979-8-3503-5094-4

THANK YOU

-ICBSII 2024

Special Thanks to our Sponsors

PROCEEDINGS OF THE
TENTH INTERNATIONAL CONFERENCE ON BIOSIGNALS, IMAGES, AND INSTRUMENTATION
March 20 - 22, 2024

Biomedical Engineering is a field of study that integrates two dynamic professions, Medicine and Engineering. It has recently established itself as an independent field with the objective of assisting medicine towards the betterment of society, through research. Being an interdisciplinary science, it has associations with various other subjects such as Electrical Engineering, Mechanical Engineering, Chemical Engineering and Biotechnology. The spectrum of bio-medical research aims to unite these disciplines in synergy, leading to new possibilities thus enabling the development of technology that could save lives. The tenth International Conference on Biosignals, Images and Instrumentation (ICBSII- 2024) is conceived with the thought of bringing together scientists, engineers and researchers from various domains all over the world. It is a platform where some of the greatest minds of the country and abroad could interact, exchange ideas and work together towards a common goal. Research papers were received from diverse areas such as Physiological Modelling, Medical Imaging, Medical Robotics, Biomechanics, Biomedical Instrumentation, Nanomaterials etc. After a rigorous review process by the expert review committee, papers that displayed quality in ideas and work were selected for the final presentation at the conference.

The 10th International Conference on Biosignals, Images, and Instrumentation is being organised by the Centre for Healthcare Technologies, a multidisciplinary research initiative focusing on research through innovation in healthcare, in collaboration with Drexel University, with the goal of instilling research aptitude in students and providing a great platform for researchers to showcase their work in various domains of healthcare. This conference is the fruit of a vision of the Management, faculty, and students of the Department of Biomedical Engineering, SSN College of Engineering in association with the Centre for Healthcare Technologies (CHT), a multi-disciplinary R&D centre, which works unanimously towards materializing the advancements in healthcare innovations.

The Department of Biomedical Engineering, since its inception in 2005, has been a pioneer in the field of biomedical technology, instrumentation, and imaging. The department has excellent infrastructure, experienced faculty members and motivated students. Department also has foreign collaborations with Birmingham City University UK, Drexel University Philadelphia, University of Bologna, Italy, and several industries and hospitals such as Rela hospitals, Kauvery hospitals, Sri Ramachandra Medical College, Perfint Healthcare, NIEPMD etc to name a few. The department is working on multiple financed projects to improve healthcare that are funded by various organizations such as SERB, LSRB, and DRDO.

To add feather to the crown, the department has conducted nine international conferences (ICBSII) in 2013, 2015, 2017, 2018, 2019, 2020, 2021, 2022, 2023 and two national conferences (NCABES) in 2014, 2016 so far.