

# PROCEEDINGS OF 2021 IEEE SEVENTH INTERNATIONAL CONFERENCE ON BIOSIGNALS, IMAGES AND INSTRUMENTATION.

MARCH 25<sup>TH</sup> - 27<sup>TH</sup> 2021



DEPARTMENT OF BIOMEDICAL ENGINEERING



MADRAS SECTION



MADRAS CHAPTER



MADRAS CHAPTER

**Seventh International conference on Bio Signals, Images and Instrumentation 2021  
(ICBSII 21)**

Copyright ©2021 by IEEE

Copyright and Reprint Permission: Abstracting is permitted with credit to the source. Libraries are permitted to photocopy beyond the limit of U.S. copyright law for private use of patrons those articles in this volume that carry a code at the bottom of the first page, provided the per-copy fee indicated in the code is paid through Copyright Clearance Centre, 222 Rosewood Drive, Danvers, MA 01923.

For reprint or republication permission, email to IEEE Copyrights Manager at [pubs-permissions@ieee.org](mailto:pubs-permissions@ieee.org). All rights reserved. Copyright ©2021 by IEEE.

- For papers in which all authors are employed by the US government, the copyright notice is: U.S. Government work not protected by U.S. copyright
- For papers in which all authors are employed by a Crown government (UK, Canada, and Australia), the copyright notice is: 978-0-7381-4471-9/21/\$31.00 ©2021 Crown
- For papers in which all authors are employed by the European Union, the copyright notice is: 978-0-7381-4471-9/21/\$31.00 ©2021 European Union
- For all other papers the copyright notice is: 978-0-7381-4471-9/21/\$31.00 ©2021 IEEE

**ISBN: 978-0-7381-4471-9/21**

IEEE Meetings, Conferences & Events (MCE)

445 Hoes Lane

Piscataway, NJ 08854 USA

Fax: +1 732 981 1769

Email: [ieee-mce@ieee.org](mailto:ieee-mce@ieee.org)

Availability: Monday-Friday 08:00 - 16:30 Eastern Standard Time

IEEE: Advancing Technology for Humanity

# Proceedings of the

## 2021 IEEE Seventh International Conference on Biosignals, Images and Instrumentation

Department of Biomedical Engineering  
Sri Sivasubramaniya Nadar College of Engineering



In association with

**Centre for Healthcare Technologies**



**ICBSII - 2021**  
**(25<sup>th</sup> – 27<sup>th</sup> March 2021)**  
**Editorial Board**

**Chief Editor:**

**Dr. A. Kavitha, Prof & HoD/BME**

**Co- Editor:**

**Ms. B. Divya, AP/BME & Dr. S. Arun Karthick, ASP/BME**

# Contents

<b>Message from the Chief Patron .....</b>	<b>xv</b>
Dr. Kala Vijayakumar, President, Sri Sivasubramaniya Nadar Institutions	
<b>Message from the Patron .....</b>	<b>xvi</b>
Dr. V. E. Annamalai, Principal, Sri Sivasubramaniya Nadar College of Engineering	
<b>Convener’s Message... ..</b>	<b>xvii</b>
Dr. A. Kavitha, Professor & Head, BME	
<b>Message from the Coordinators.....</b>	<b>xviii</b>
Dr. S. Pravin Kumar, Associate. Prof/BME, Dr. K. Nirmala, Associate. Prof/BME	
<b>Conference Organizing Committee... ..</b>	<b>xx</b>
<b>Technical Advisory Committee.....</b>	<b>xxi</b>
<b>Review Committee.....</b>	<b>xxiii</b>
<b>Student Organizing Team .....</b>	<b>xxvii</b>

## Chief Guest

**Dr. Hasan Ayaz** .....xxx  
Drexel University ,USA

## Guests of Honour

**Dr. Catherine von Reyn**.....xxx  
Drexel University, USA

**Dr. Sriram Balasubramanian**.....xxx  
Drexel University, USA

**Dr. G. S. Bhuvaneshwar**.....xxxi  
SCTIMST, Trivandrum

**Dr. N. Kumarappan**.....xxxi  
Chairman, IEEE Madras Section

**Mr R. Balamurugan**.....xxxi  
Vice President, HCL Technologies, Chennai

## Keynote Speakers

<b>Dr. Jan G. Svec</b> .....	<b>xxxii</b>
Physiology and Acoustics of Human Voice	
<b>Dr. Barbara Zitova</b> .....	<b>xxxiii</b>
Medical imaging in computer aided diagnosis, focused on voice therapy	
<b>Dr. Sudhir Ganesan</b> .....	<b>xxxiv</b>
The Prospects Of Engineering In Medicine –A Spine Surgeon's Perspective	
<b>Dr. M. Murugappan</b> .....	<b>xxxv</b>
Affective Computing Applications in Healthcare	
<b>Dr. S. Kailash</b> .....	<b>xxxvi</b>
Brain stimulation methods in psychiatry	
<b>Dr. Hugo Gamboa</b> .....	<b>xxxvii</b>
Making Sense of Bio signals	
<b>Dr. Deepak Joshi</b> .....	<b>xxxviii</b>
Neuromuscular electrical stimulation: Benefits from Machine Learning	
<b>Ms. Jesin James</b> .....	<b>xxxix</b>
Languages and emotions: Towards speech technology development for under-resourced languages and secondary emotions	
<b>Dr. Justin Dauwels</b> .....	<b>xxxxx</b>
Artificial Intelligence for Applications in Neurology	

## Research Papers

### Session-I

<b>1. Detection of Autism Spectrum Disorder from EEG signals using pre-trained deep convolution neural networks</b>	2
Qaysar Mohi-ud-Din, A. K. Jayanthi, SRM Institute of Science and Technology, India	
<b>2. Cloud based analysis and classification of EEG signals to detect epileptic seizures</b>	7
Munyaradzi Charles Rushambwa, Mavis Gezimati, Dr. P. Govindaraj, Rajkumar Palaniappan, Dr Vikneswaran Vijean, Fizza Ghulam Nabi, Harare Institute of Technology, Zimbabwe	
<b>3. A Review on The Medical Applications of Electroencephalography (EEG)</b>	12
Tushar Kanti Bera, National Institute of Technology Durgapur, India	
<b>4. TSSummarize: A Visual Strategy to Summarize Biosignals</b>	18
João Rodrigues, Phillip Probst and Hugo Gamboa, Universidade Nova de Lisboa, Europe	
<b>5. Frame Work For EEG Based Emotion Recognition Based On Hybrid Neural Network</b>	24
Jancy Mariam Jose, J.Aravinth, Amrita Vishwa Vidyapeetham, India	
<b>6. Computational Approach to guide Mind Controlled Robotic Arm using BCI – A Review</b>	31
Anna Latha, Dr. E .Sathish, Dr. Florence Gnana Poovathy, Vellore Institute of Technology, India	
<b>7. A new lead system for improved recording of P-wave amplitude and its significance with existing optimal leads</b>	37
N. Prasanna Venkatesh, Arya Bhardwaj, J. Sivaraman, National Institute of Technology Rourkela, India	
<b>8. Automated EEG Analysis for Early Diagnosis of Epilepsy: A Comparative Study to Determine Relative Accuracy of Arithmetic and Huffman Coding Algorithms</b>	42
Anisha Kumar, Pratishtha Singh, Rajlakshmi Khawas, Vellore Institute of Technology, India	
<b>9. Automated Detection of Brain Abnormality using Deep-Learning-Scheme – A Study</b>	48
Seifedine Kadry, Yunyoung Nam, Hafiz Tayyab Rauf, Venkatesan Rajinikanth, Isah A. Lawal, Noroff University College, Norway	

## **Research Papers** **Session-II**

<b>1. Performance Metrics Analysis of Adaptive threshold Empirical Mode Decomposition Denoising method for suppression of noise in Lung sounds</b> B. Sangeetha, R.Periyasamy, National Institute of Technology, Trichy, India .....	54
<b>2. Study of P wave indices in sinus rhythm and tachycardia</b> Sivaraman Jayaraman, Arya Bhardwaj, National Institute of Technology Rourkela, India .....	60
<b>3. Sleep Disorder Diagnosis using EEG based Deep Learning Techniques</b> G.Hari Krishnan, T. Sudhakar, N. R. Krishnamoorthy, Sree Vidyanikethan Engineering College, India .....	66
<b>4. A Preliminary Study on Relief based Feature Ranking for Classification of Myoelectric Signals</b> Salman Mohd Khan, Mohd. Umair Alam, Abid Ali Khan, Aligarh Muslim University, India .....	70
<b>5. Deep Learning for Sleep Disorders: A Review</b> D.Hepsiba, R.Jane Preetha Princy, Karunya Institute of Technology and Sciences, India .....	75
<b>6. Arduino and LabVIEW based Heart Rate Detection System</b> Deepjyoti, Hemashree Bordoloi, Samar Jyoti Saikia, Assam Don Bosco University, India .....	81
<b>7. Analysis of Tremors in Parkinson’s Disease Using Accelerometer</b> Priscilla Dinkar Moyya, Niya Romy Markose, Mythili Asaithambi, Vellore Institute of Technology, India .....	86



## **Research Papers** **Session-III**

<b>1. A Comparative Study Of Various Multimodal Medical Image Fusion Techniques– A Review</b> Kanike Vijay, Anchula Sathish, R.G.M.C.E.T, India .....	92
<b>2. Brain Tumour Detection Using Deep Learning</b> R.Nitha Mohan , D.Abinaya, R.P. Aneesh, Avigyan Sinha, Ashwin G Singerji, Johns Hopkins University, USA .....	98
<b>3. Automated Detection of Schizophrenia from Brain MRI Slices using Optimized Deep-Features</b> Seifedine Kadry, David Taniar, Robertas Damaševičius, Noroff University College, Norway .....	103
<b>4. COVID-SEGNET: Diagnosis of Covid-19 Cases on Radiological Images using Mask R-CNN</b> Avik Kundu, Chandan Mishra, Saurabh Bilgaiyan, KIIT University, India .....	108
<b>5. Breast Density Classification in Mammogram Images</b> V.Vyshnavi, Devi Vijayan, R. Lavanya, Amrita School of Engineering, India .....	113
<b>6. Systematic study on diagnosis of lung disorders using machine learning and deep learning algorithms</b> R. Swathi Sri, Dr. A. Menaka Pushpa, VIT Chennai Campus, India .....	118
<b>7. Extraction of Tumour in Breast MRI using Joint Thresholding and Segmentation – A Study</b> Seifedine Kadry, Robertas Damaševičius, David Taniar, Noroff University College, Norway .....	126
<b>8. Machine Learning Approach to Study the Impact of Obesity on Autonomic Nervous System using Heart Rate Variability Features</b> Sr Rathod, U.M.Chaskar, C.Y.Patil, College of Engineering, India .....	131
<b>9. Image Character Recognition using Convolutional Neural Networks</b> Adith Narayan, Raja Muthalagu, Birla Institute of Technology and Science Pilani, UAE .....	136

# Research Papers

## Session-IV

<b>1. Comparison of Watershed and Local Center of Mass Techniques in Segmentation of Breast Mammogram</b>	142
Mohamed Yacin Sikkandar, College of Applied Medical Sciences, Majmaah University, Saudi Arabia	
<b>2. Prediction of cardiac rhythm based on heart rate variability features in sinus tachycardia conditions</b>	147
Sivaraman Jayaraman, B. Dhananjay, National Institute of Technology Rourkela, India	
<b>3. Modelling Of Electrocardiogram Using Autoregressive Moving Average Model And Linear Predictive Coefficient-a Comparative Study</b>	153
Rashmi Annamma George, Dr.R. Periyasamy, National Institute of Technology, India	
<b>4. Extraction of Abnormal Skin Lesion from Dermoscopy Image using VGG-SegNet</b>	157
Seifedine Kadry, David Taniar, Robertas Damaševičius, Noroff University College, Norway	
<b>5. Emotional Classification of EEG Signal using Image Encoding and Deep Learning</b>	162
K.A.Anjana, M.Ganesan, R.Lavanya, Amrita School of Engineering, India	
<b>6. Novel Method for Analyzing the Relation of miRNA Seed Length and the miRNA-mRNA Interaction Strength</b>	167
Binthiya Gabriel, Tessamma Thomas, Cochin University of Science and Technology, India	
<b>7. Texture Analysis Of Reaction Diffusion Level Set Evolution Of Multiple Sclerosis Lesions In Brain Mr Images</b>	172
A.Pandian, Udhayakumar Ganesan, SRM Valliammai Engineering College, India	
<b>8. Heart Rate Monitoring Using Electrical Impedance</b>	175
Yousef Alharbi, Anwar Alshrouf, Sofiene Mansouri, Prince Sattambin Abdulaziz University, Saudi Arabia	

## **Research Papers**

### **Session-V**

<b>1. Relationship between Glaucoma and Complexity Measures of the Electroretinogram</b> Marc Sarossy, Dinesh Kumar, Zhichao Wu, University of Melbourne, Australia .....	180
<b>2. Identification of Schizophrenia with using LSTM Recurrent Neural Network</b> Abinaya Sundari R, C. M. Sujatha, Anna University, India .....	184
<b>3. Segmentation of Spinal Canal using Active Contour Model</b> Gayathry S Warriar, Dr.K.S.Angel Viji, Noorul Islam Centre for Higher Education, India .....	190
<b>4. DCGAN based Pre-trained model for Image Reconstruction using ImageNet</b> Nandini Kumari, Shamama Anwar, Vandana Bhattacharjee, Birla Institute of Technology, India .....	194
<b>5. Classification of Dental X-ray Images Using Machine Learning</b> Sindu Divakaran, K. Vasanth , D. Suja , V.Swedha, Sathyabama Institute of Science and Technology , India .....	200
<b>6. U-Net Supported Segmentation of Ischemic-Stroke-Lesion from Brain MRI Slices</b> Seifedine Kadry, Robertas Damaševičius, David Taniar, Venkatesan Rajinikanth, Isah A. Lawal, St. Joseph's College of Engineering, India .....	203
<b>7. Drowsiness Detection System Using Deep Learning</b> Aneesh Rp, Avigyan Sinha, Saradha K Gopal, Johns Hopkins University, USA .....	208
<b>8. A Deep Learning Approach to classify the Honeybee Species and health Identification</b> M.Karthiga, S.Sountharajan, Nandhini SS, Bannari Amman Institute of Technology, India .....	214
<b>9. Estimation of the mean Metabolic Ratio For The Purpose Of Grading Of Glial Tumours Of The Brain</b> Gaurav Mishra, Datta Meghe Institute of Medical Sciences, India .....	221

# Research Papers

## Session-VI

<b>1. Retinal Vessel Segmentation with Slime-Mould-Optimization based Multi-Scale-Matched-Filter</b>	228
Seifedine Kadry, Venkatesan Rajinikanth, Robertas Damaševičius, Noroff University College, Norway	
<b>2. SITCOV-A Benchmark Database for COVID19 Analysis</b>	233
Uma Maheshwari M, R.Tamilselvi, M.Parisa Beham, N.Subanandhini, Sethu Institute of Technology, India	
<b>3. Positioning the 5-DOF Robotic Arm using Single Stage Deep CNN Mode</b>	239
Sreekar C, Sindhu V S, Rubin Bose S, Sathiesh Kumar V, Madras Institute of Technology , India	
<b>4. Malnutrition Detection Using Convolutional Neural Network</b>	245
Arun Raj Lakshminarayanan, Pavani B, Rajeswari V, A.Abdul Azeez Khan, K.Javubar Sathick, B. S. Abdur Rahman Crescent Institute of Science and Technology, India.	
<b>5. A Multiclass Skin Lesion classification approach using Transfer learning based convolutional Neural Network</b>	250
Siddalingaswamy P C, Cauvery K , Sameena Pathan, Noel D'souza, Manipal Institute of Technology, India	
<b>6. Eye tumor detection using deep learning</b>	256
Aneesh Rp, Avigyan Sinha, Nazneen. N. S, Johns Hopkins University, USA	
<b>7. Classification of Arrhythmia in Time Series ECG Signals Using Image Encoding and Convolutional Neural Networks</b>	261
Vandith Sreenivas K, Ganesan M, Lavanya R, Amrita School of Engineering, Coimbatore, India	
<b>8. Multimodal Emotion Recognition Using Different Fusion Techniques</b>	267
Gokul Subramanian, Kotapati Prathyusha, Niranjan Cholendiran, Noviya Balasubramanain, Aravinth J, Amrita School of Engineering, Coimbatore, India	
<b>9. Modelling and Simulation of High flux Haemodialyzer membranes of different porosities to identify the optimal Membrane Design</b>	273
Ahana Fatima Alex, R Vinoth, Ravishankar Dudhe, Manipal Academy of Higher Education, UAE	

# Research Papers

## Session-VII

<b>1. MicroErgo: A Concept for an Ergonomic Self-Assessment Tool</b>	João Rodrigues, Phillip Probst, Catia Cepeda, Federico Guede , Fern´andez, SaraSilva, PatriciaGamboa, Carlos Fujˆaoz Cl´audia R.Quaresma and HugoGamboa, , Faculdade de Ciˆencias e Tecnologia, Universidade Nova de Lisboa, Caparica, Portugal .....280
<b>2. Real time heart beat monitoring using computer vision</b>	Aneesh Raghavan Parameswaran, Arppana A R, N K Reshma, Gopika Raghu, Nita Mathew, Harsha R Nair, Regional Centre IHRD, India .....286
<b>3. Development of Reconfigurable Control Schemes for Epileptic Seizures</b>	Nambi Narayanan, Sutha Subbian, MIT Campus, Anna University, India .....292
<b>4. Non-invasive Estimation Of Salivary Biomarkers For The Diagnosis Of Diabetic Nephropathy</b>	Purnima Shanmugam, Atchaya. V, Avesha. R, Devi Lakshmi. A , Divya Bharathi. S, Dr.Prasanna. S, Jerusalem College of Engineering, India .....297
<b>5. Detection Of Estrus In Bovine Using Machine Learning</b>	Hemalatha. R.J, SonaShree.S.P , Thamizhvani T.R, Vijayabaskar.V , Vels Institute of Science, Technology and Advanced Studies .....309
<b>6. Length of Stay Prediction in Acute Intensive Care Unit in Cardiothoracic Surgery Patients</b>	Nafiseh Mollaei, Ana Rita Londral , Catia Cepeda, Salom´e Azevedo, LIBPhys-UNL Departamento de F´ısica, FCT Nova University of Lisbon, Portugal. .....314
<b>7. Breast-Cancer Detection using Thermal Images with Marine-Predators-Algorithm Selected Features</b>	Seifedine Kadry, Venkatesan Rajinikanth , David Taniar, Robertas Damaševičius, Hafiz Tayyab Rauf , Noroff University College, Kristiansand, Norway .....319

# Research Papers

## Session-VIII

<b>1. Development of a Smart Cervical Collar System for Feedback Control of Intracranial Pressure after Traumatic Head Injury</b>	
Frank Washko, William S. Edwards, Leslie A. Washko, Flinders University Tonsley, South Australia	
.....	326
<b>2. Data Acquisition System for Web-based Multi-modal Medical Data Repository</b>	
Munyaradzi Charles Rushambwa, Dr. Anirban Mukherjee, Maitreya Maity, Harare Institute of Technology Harare, Zimbabwe	
.....	332
<b>3. Design and Development of Non-invasive Auto mucus Removal Device By Acoustic Assisted Therapy</b>	
Anusha Bhavani K, L.Sujitha, P.Uvaraj, GEM Hospital and Research Institute, India	
.....	337
<b>4. Novel Design Approach For Treating Sinus Pain At Acupressure Points</b>	
Priya.L, Aarthi S, Preethi M, Lakshmi Prasanna K, Monisha T, , Vignan’s Foundation for Science, Technology and Research, India,	
.....	342
<b>5. Intelligent walking stick with static and dynamic obstacle detection technology for visually challenged people</b>	
Krishnakumar S, Bethanne J, Umashankar. G, Sathyabama Institute of Science and Technology, India	
.....	347
<b>6. Acoustic Based Separation of Blood Cells Using Microfluidic Device</b>	
Mohamed Yousuff, Pavithra devi. E, Swetha.C, Ramya. R, C.Abdul Hakeem College of Engineering and Technology, India	
.....	353

## Research Papers

### Session-IX

- 1. Apparent Diffusion Coefficient Values With Normalization In Correspondence With WHO Grading Of Glial Tumours Of The Brain**  
 Gaurav Mishra, Akhilesh Agrawal, Punit Fulzele, Ashutosh Bagde, Datta Meghe Institute of Medical Sciences, India  
 .....359
- 2. Osteositis: A Dataset for Exploration and Classification of Osteoarthritis**  
 Rajakumar K, M.Parisa Beham, R.Tamilselvi, D.Radhika, Sethu Institute of Technology, India  
 .....364
- 3. Voice Controlled Home Automation System**  
 Aneesh Rp, Nisha A, Ribil Mary Roy, Blessy Sabu, Sivaji College of Engineering and Technology , India  
 .....369
- 4. Automatic Anesthesia Control System**  
 Shelishiyah Raymond, Susmitha Edagottu, Lasngewhun Mawblei, Minhajul Ahmed, R&D Institute of Science and Technology, India  
 .....374
- 5. Design of a Medical Prototype Robot for Nurse Assistance**  
 Steeve Shibu Chempolil, Renie Melvinia, Sneha Saji, Karthik Raj V, Vel Tech Rangarajan Dr. Sagunthala R & D Institute of Science and Technology, India  
 .....379
- 6. A Survey on Diabetic Retinopathy Disease Detection and Classification using Deep Learning Techniques**  
 S Valarmathi, Vijayabhanu R, Avinashilingam Institute of Home Science and Higher Education for Women, India  
 .....384
- 7. Machine-Learning-Scheme to Detect Choroidal-Neovascularization in Retinal OCT Image**  
 Seifedine Kadry, Venkatesan Rajinikanth, Robertas Damaševičius, David Taniar, Hafiz Tayyab Rauf, Noroff University College, Kristiansand, Norway  
 .....388
- 8. An Optimization of Feature Selection for classification using Meta – Heuristic Whale Optimization Algorithm**  
 V. Yaraswini, Dr. Santhi Baskaran, Pondicherry Engineering College, India  
 .....393
- 9. Dental Plaque Segmentation in Digital Photographs using K-Means Clustering in L\*a\*b\* Color Space**  
 Dr.R.Karthika devi, D.Radhika, Sethu Institute of Technology, India  
 .....397

*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-round 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 which brings together healthcare and technology, the front runners in the modern world. I congratulate the Department of Biomedical Engineering for organizing the IEEE- EMBS sponsored International Conference on Biosignals, Images and Instrumentation, in association with the Centre for Healthcare Technologies of SSN. This flagship conference of the Department of BME will provide the participants and the students a unique opportunity to develop enriching perspectives by interacting with some of the renowned experts in these fields, from all over the world. I am certain that the talks by eminent scientists, researchers, clinicians and surgeons, and the papers presented will stimulate lively discussions to lay a strong foundation for further advanced research in these fields. 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.

Mrs. Kala Vijayakumar,  
Chief Patron,  
ICBSII – 2021

## From the Chief Patron

**Dr. V. E. Annamalai**

Principal, SSN College of Engineering



I am pleased that the department of Biomedical Engineering is organizing the IEEE- EMBS sponsored International Conference on Biosignals, Images and Instrumentation, in association with the Centre for Healthcare Technologies of SSN, in a manner befitting the stream.

Biomedical engineering is a multidisciplinary field integrating Engineering and healthcare. It focuses on the advances that improve human health and health care at all levels. The department's engagement in wide spectrum of activities with involvement of students and faculty along with strategic planning process has strengthened it.

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 Biomedical Engineering Department for structuring it to perfection and wish them all success.

Dr. V. E. Annamalai  
Patron,  
ICBSII – 2021.

# From the Convener

**Dr. A. Kavitha**

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



Education is a holistic endeavour, creating new paths with endless boundaries and priming minds to orient one to the world. That being said, it gives me immense pleasure to present the IEEE- EMBS sponsored International Conference on Biosignals, Images and Instrumentation (ICBSII 2021).

Biomedical engineering discipline is one which catalyzes interactions between biologists, physical scientists, and engineers to benefit medicine and human health. This serves society by conducting research that develops quantitative linkages across scales in the human body and uses that development to build new tools to improve human health. The outcomes of research assume a whole new level of importance and significance.

The department is frequently organizing workshops, seminars, project exhibitions and guest lectures on diverse concepts related to the core and interdisciplinary subjects in biomedical engineering to equip the students in gaining a comprehensive knowledge of the industrial requirements to the fullest.

The Department of Biomedical Engineering, in association with the Centre for Healthcare Technologies, a multidisciplinary research initiative, concentrating on research through innovation in healthcare, is organizing the 7<sup>th</sup> International Conference on Biosignals, Images and Instrumentation, hoping to instil research aptitude in students and provide a great platform for the researchers to showcase their work in various domains of Biomedical Engineering.

The Pandemic has taught us so much. Taking all positive notes from the yearlong experiences, we have organised this conference in virtual mode, inviting speakers and participants from all over the globe.

Come let us discuss Healthcare – And challenge anything that tries to disturb our wellbeing!

Dr. A. Kavitha  
Convener,  
ICBSII – 2021

# From the Coordinators

**Dr. S. Pravin Kumar**

Associate Professor,  
Department of Biomedical Engineering  
SSN College of Engineering



Today healthcare industry is a multi-trillion-dollar economy behemoth at a crossroads. Currently being weighed down by crushing costs, the industry is looking for ways to improve in nearly every imaginable area. That is where health tech comes in. Tech-infused tools are being integrated into every step of our healthcare experience. The importance of quality healthcare has risen multifold after the COVID-19 outbreak. The academic diaspora is much more enthusiastic now on healthcare technology development than ever before. We have witnessed new collaborative outcomes almost in all leading private and public universities, a thrust on opportunities for innovations and research efforts for improving the health state of the mankind.

Infusion of technology has also given much better opportunities to explore one's healthset, mindset and soulset like never before. Due to this, there is an increasing demand for quality tools for speech technology, human mechanics, neurological and wearable data. Our invited sessions and pre-conference workshop in the IEEE sponsored Seventh International Conference on Bio signals, Images and Instrumentation (ICBSII 2021) are therefore carefully structures to feature these talking points.

ICBSII 2021 to be held on March 25-27, 2021 in SSN College of Engineering is expected to serve as an exemplary platform for researchers, academicians and industry practitioners from all over the world to explore and discuss novel ideas. This conference provides an opportunity to collaborate with experts and scholars from various fields on topics like bio signals, medical images, medical data analytics and medical instrumentation. We have seen tremendous response this time, with papers received from almost all continents and speakers from USA, Europe, Australia, Middle East, and Singapore. I would like to thank Dr. Sriram Balasubraminan, Drexel University, USA, Dr. Hugo Gamboa, LIBPhys-UNL, Portugal and my mentor Dr. Jan G Svec, Palacky University, Olomouc for their support in connecting with the fellow scientists, industrialists and academicians in their respective countries.

I extend my sincere gratitude to the management of SSN College of Engineering, Mrs. Kala Vijayakumar-President, SSN Institutions and Dr. V. E. Annamalai- Principal, SSN College of Engineering for granting the department a delightful opportunity to organise this brilliant event that enables us to grow on a global level. I would like to thank our Head and organizing chair Dr. A. Kavitha, my colleague Dr. K. Nirmala in the organizing team and all my colleagues in the ICBSII-2021 coordinating committee for their support, constructive feedback, timely, tireless, and meticulous efforts in conducting this event. My special thanks to Ms. Divya for coordinating the pre-conference workshop. I am grateful to our sponsors IEEE Madras Section, EMBS Madras Chapter and Computer Society Madras Chapter. Finally, I 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.

**Dr. S. Pravin Kumar**  
Coordinator,  
ICBSII – 2021

## From the Coordinators

**Dr. K. Nirmala**

Associate Professor,  
Department of Biomedical Engineering,  
SSN College of Engineering



It gives me an immense pleasure to once again put our hands together in organizing the Department's annual extravaganza! The IEEE sponsored Seventh International Conference on Bio signals, Images and Instrumentation (ICBSII 2021) on March 25-27<sup>th</sup>, 2021 in Sri Sivasubramaniya Nadar College of Engineering. The purpose of this conference is to bring Researchers and Industry Practitioners together to present and discuss novel approaches and solutions. The outcomes in the field of Biological, Medical, Health Care, Pharmaceutical, Biotechnology, Bioinformatics, Computer Science, Information Technology and Communication to create synergy, support Interdisciplinary Research, and to exchange ideas and explore new avenues of collaborations. For this conference, research papers on the mentioned issues were received from India and Abroad. Papers received in this international conference were reviewed by the technical program committee members. The conference turned out to be a forum for scientists and researchers all over the world to share their ideas, experiences, findings, and conclusions of their work in due course of their scientific research.

I would like to thank the ICBSII-2021 organization and technical program committee for their support, constructive feedback, and timely proposal review. I extend my sincere gratitude to the management of SSN College of Engineering, Dr. Kala Vijayakumar, President, SSN Institutions and Dr. V.E. Annamalai, Principal, Sri Sivasubramaniya Nadar College of Engineering for granting the department a wondrous opportunity to organize and conduct this productive occasion helping us to take it forward to the global level. I would extend my sincere thanks to the sponsors IEEE Madras section, IEEE Computer society Madras chapter and IEEE EMBS madras chapter.

Finally, I would also like to take opportunity to thank all external reviewers and contributing authors for producing high quality papers to be presented at ICBSII 2021. I would also like to thank all the Faculties, Non-teaching staff and students of the biomedical department who have contributed towards organizing the occasion to make it a grand success!

**Dr. K. Nirmala**  
Coordinator,  
ICBSII – 2021

# Conference Organizing Committee

## Chief Patron

**Ms. Kala Vijayakumar, President, SSN Institutions**

## Patron

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

## Convener

**Dr. A. Kavitha, Professor and HOD/BME**

## Coordinators

**Dr. S. Pravin Kumar**

**Dr. K. Nirmala**

## Treasurer

**Dr. V. Mahesh**

**Dr. R. Subashini**

## Committee Members

**Dr. V. Mahesh**

**Dr. L. Suganthi**

**Dr. S. Bagyaraj**

**Dr. J. Vijay**

**Dr. S. Arun Karthick**

**Dr. B. Geethanjali**

**Ms. M. Dhanalakshmi**

**Ms. R. Nithya**

**Dr. R. Subhashini**

**Ms. B. Divya**

**Dr. Sachin Gaurishankar Sarate**

# Technical Advisory Committee

## International Advisory Committee Members

- **Dr. Sriram Balasubramanian**, Drexel University, USA
- **Dr. Kong Pui Wah**, NTU, Singapore
- **Dr. Teo Ee Chon**, NTU, Singapore
- **Dr. Mohamed Yacin Sikkandar**, Majmaah University, Saudi Arabia
- **Dr. M. Murugappan**, KCST, Kuwait
- **Dr. Eko Supriyanto**, Universiti Teknologi Malaysia, Malaysia
- **Dr. S. Arunachalam**, JIC, Kingdom of Saudi Arabia
- **Dr. Tinashe Mutswangwa**, Senior Lecturer., UCT, South Africa
- **Dr. Sekar Raju**, Xi'an jiaotong-Liverpool University, China
- **Dr. J. Jesu Christopher**, Senior Scientist, ASTRAZENCA, Cambridge, UK
- **Dr. S. Ramji**, Queen's University Belfast, UK
- **Dr. S. Ramesh**, University of Malaya, Malaysia
- **Dr. P. Sasikumar**, Higher Colleges of Technology, UAE

## National Advisory Committee Members

- **Dr. N. Kumarappan**, Chairman, IEEE Madras Section
- **Dr. D. Devaraj**, Secretary, IEEE Madras Section
- **Dr. Joseph Gladwin**, Treasurer, IEEE Madras Section
- **Dr. P. Sakthivel**, Chairman, IEEE Computer Society Madras Chapter
- **Dr. B. K. B. Jayanthi**, Chairman, IEEE Engineering in Medicine and Biology (EMBS) Madras Chapter
- **Dr. S. Ramakrishnan**, IIT Madras
- **Dr. S. Muttan**, Anna University
- **Dr. Renu John**, IIT Hyderabad
- **Dr. R. Periyasamy**, NIT Trichy
- **Mr. K. Mohanavelu**, Scientist E, DEBEL., DRDO, Bangalore
- **Mr. S. Sivagnanam**, Additional Industrial Advisor., Govt. of India, MSME

- **Dr. Niranjan D. Khambete**, Pune Govt. Hospital, Pune
- **Dr. G. Sudhir**, Orthopaedic Spine Surgeon., SRMC, Chennai
- **Dr. G. Kumaramanickavel**, Director of Research., Narayana Nethralaya, Bangalore
- **Dr. B. Minimol**, Gov. Model Engineering College, Kochi
- **Dr. C. M. Sujatha**, Anna University, Chennai



# Review Committee

## International List of Reviewers:

- **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, Universiti Teknologi Malaysia
- **Dr. Sriram Balasubramanian**, Associate Professor Drexel University, USA
- **Dr. M. Murugappan**, Associate Professor Kuwait College Of Science And Technology, Kuwait
- **Dr. S. Arunachalam**, Associate Professor Jic, Kingdom Of Saudi Arabia
- **Dr. Md. Mamun Bin Ibne Reaz**, Professor Universiti Kebangsaan Malaysia, Selangor Malaysia
- **Dr. Justin Dauwels**, Associate Professor, Nanyang Technological University, Singapore
- **Dr. Anita Singh**, Associate Professor, Widener University, Chester, PA
- **Dr. Alejandro Riera**, Asia Commercial Manager, Neuroelectrics Barcelona S.L.Barcelona Area, Spain
- **Dr. Dyah Ekashanti Octorina Dewi**, Senior Lecturer, Universiti Teknologi Malaysia, Malaysia

## IEEE Members:

- **Dr. P. Sakthivel**, Chairperson, CSI, Professor and Dean, University College of Engineering Kancheepuram.
- **Dr. K. B. Jayanthi**, Chairman, IEEE EMBS Madras Section, Professor, K.S. Rangasamy College of Technology, Tiruchengode.
- **Dr. O. Uma Maheswari**, Professor (Senior Grade), CEG Campus, Anna University, Chennai
- **Dr. Sanjay Singh**, Professor Manipal Institute of Technology, Manipal.
- **Dr. S. Prabakar**, Professor, Dr.NGP Institute of Technology, Coimbatore
- **Dr. V. Vijayan**, Professor, St. Joseph's College of Engineering, Chennai.
- **Dr. G. Kulanthaivel**, Professor, NITTTR, Chennai
- **Dr. B. Banurekha**, Associate Professor, PSG College of Technology, Coimbatore
- **Dr. M. Jagannath**, Associate Professor, VIT, Chennai
- **Dr. K. Adalarasu**, Associate Professor, Sastra, Deemed to be University, Thanjavur
- **Dr. Poornalatha G**, Associate Professor, Manipal Institute of Technology, Manipal
- **Dr. Malaya Kumar Nath**, Assistant Professor, NIT Puducherry, Karaikal
- **Dr. M. G. Sumithra**, Professor, KPR Institute of Engineering and Technology, Coimbatore
- **Dr. B. Balasubramaniam**, Professor, Sri Ramakrishna College of Engineering, Coimbatore
- **Dr. V. Ulagamuthalvi**, Associate Professor, Sathyabama Institute of Science and Technology, Chennai
- **Dr. P. Sivasankar**, Assistant Professor, NITTTR, Tharamani, Chennai
- **Dr. G. S. Uthaya Kumar**, Professor, St. Joseph Institute of Technology, Chennai
- **Dr. P. Sakthivel**, Professor, University College of Engineering Kancheepuram
- **Dr. K. B. Jayanthi**, Professor, K.S. Rangasamy College of Technology, Tiruchengode
- **Dr. T. Shanmuganatham**, Assistant Professor, Pondicherry University, Pondicherry
- **Dr. P. C. Siddhalinga Swamy**, Associate Professor, Manipal Institute of Technology, Manipal

## Indian Reviewers:

- **Dr. A. Kavitha**, Professor, SSN College of Engineering, Chennai
- **Dr. V. Mahesh**, Associate Professor, SSN College of Engineering, Chennai
- **Dr. S. Pravin Kumar**, Associate Professor, SSN College of Engineering, Chennai
- **Dr. L. Suganthi**, Associate Professor, SSN College of Engineering, Chennai
- **Dr. S. Bagyaraj**, Associate Professor, SSN College of Engineering, Chennai
- **Dr. J. Vijay**, Associate Professor, SSN College of Engineering, Chennai
- **Dr. S. Arun Karthick**, Associate Professor, SSN College of Engineering, Chennai
- **Dr. B. Geethanjali**, Associate Professor, SSN College of Engineering, Chennai
- **Dr. K. Nirmala**, Associate Professor, SSN College of Engineering, Chennai
- **Dr. R. Subashini**, Assistant Professor, SSN College of Engineering, Chennai
- **Dr. Sachin G Sarate**, Assistant Professor, SSN College of Engineering, Chennai
- **Dr. C. Vinoth Kumar**, Assistant Professor, SSN College of Engineering, Chennai
- **Dr. C. M. Sujatha**, Assistant Professor, Anna University, Chennai
- **Dr. G. Kavitha**, Assistant Professor (Senior Grade) MIT campus, Anna University
- **Dr. A. Mythili**, Associate Professor, VIT Vellore campus
- **Dr. K. Kamalanand**, Assistant Professor MIT campus, Anna University
- **Dr. B. Minimol**, Associate Professor, Model Engineering College, Kerala
- **Dr. R. Tamilselvi**, Professor, Sethu Institute of Technology, Madurai
- **Dr. A. K. Jayanthi**, Professor, SRM University, Chennai.
- **Dr. S. Poonguzhali**, Associate Professor, CEG Campus, Anna University
- **Dr. S. Sasikala**, Professor, CEG Campus, Anna University
- **Dr. K. Adalarsasu**, Associate Professor, Sastra University, Tanjavur
- **Dr. Ganesh Vaidhyanathan**, Principal SVCE, Chennai
- **Dr. M. G. Sumithra**, Professor Bannari Amman Institute of Technology, Sathyamangalam
- **Dr. T. R. Ganesh Babu**, Professor Muthayammal Engineering College, Rasipuram

- **Dr. M. C. Jobin Christ**, Professor, Rajalakshmi Engineering College, Chennai.
- **Dr. K. Vidhya**, Professor, Vel Tech, Chennai.
- **Dr. B. Padmapriya**, Associate Professor, PSG College of Technology, Coimbatore
- **Dr. Judith Justin**, Professor, Avinashilingam University, Coimbatore
- **Dr. M. Parisa**, Associate Professor, Sethu Institute of Technology, Virudhunager.
- **Dr. P. Hosanna Princye**, Associate Professor, SEA College of Engineering and Technology, Bangalore
- **Dr. K. Bommanna Raja**, Professor, KPR Institute of Engineering and Technology, Coimbatore.
- **Dr. O. Saraniya**, Associate Professor, Government College of Technology, Coimbatore
- **Dr. T. Jayashree**, Assistant Professor, CEG Campus, Anna University.
- **Dr. N. Venkateshwaran**, Professor, SSN College of Engineering, Chennai

# Student Organizing Team

## Registration:

- Sowmiya E C – Final Year PG
- Jeslin Libisha J – Final Year PG
- Sonali – Final Year UG
- Reshma C.A – Final Year, UG

## Conference Proceedings:

- Sowmiya E C – Final Year PG
- Jeslin Libisha J – Final Year PG
- Isa Bashir Salisu – Final Year PG

## Design:

- Radhika B – Final Year, UG
- Reshma C.A – Final Year, UG
- Rajkumar – Final Year, UG
- Ansar Ahamed – Final Year, UG

## Session Volunteers:

- Anjana A – Third Year, UG
- Rebecca Maria Gnanamuthu – Third Year, UG
- Crossny Snowlin - Third Year, UG
- Pooranima G – Third Year, UG
- Sai Kaviya Neharika – Third Year, UG
- Lokesh Kumar M –Third Year, UG
- Keerthana B –Third Year, UG
- Varsha Seshadhri –Third Year, UG
- Leenasri Ramaswamy –Third Year, UG
- Yuvasri – Third Year, UG
- Harshini Ramaswamy- Third Year, UG
- Prasadha Prabhu –Third Year, UG
- Sherwin Robert- Third Year, UG
- Divya Rajesh –Third Year, UG
- Hemalatha T- Third Year, UG
- Lohithaa J- Third Year, UG
- Sockalingam- Third Year, UG
- Sivagami Vishnukumar- Third Year, UG
- Manjula S- Third Year, UG

- Arthi V – Third Year, UG

### **Event Coordinators:**

- Jeslin Libisha J - Final Year PG
- Sowmiya E C - Final Year PG
- Isa Bashir Salisu - Final Year PG

**2021 IEEE Seventh International  
Conference on Bio signals, Images and  
Instrumentation**

**ICBSII 2021**

**KEYNOTE SPEAKERS' PROFILE**



**Dr. Hasan Ayaz**, is an Associate Professor at Drexel University, School of Biomedical Engineering, Science and Health Systems, Philadelphia, PA and in the Department of Psychology at the Drexel University College of Arts and Sciences, a core member of the Cognitive Neuroengineering and Quantitative Experimental Research Collaborative and with affiliations at the University of Pennsylvania and the Children's Hospital of Philadelphia. He received his BSc. in Electrical and Electronics Engineering at Boğaziçi University, Istanbul, Türkiye with high honors and MSc. and PhD degrees from Drexel University where he developed enabling software for functional Near Infrared Spectroscopy based brain monitoring and FDA approved medical devices. His research interests include neuroengineering in human computer interaction and neuroergonomics, as well as clinical and field applications of optical brain imaging.

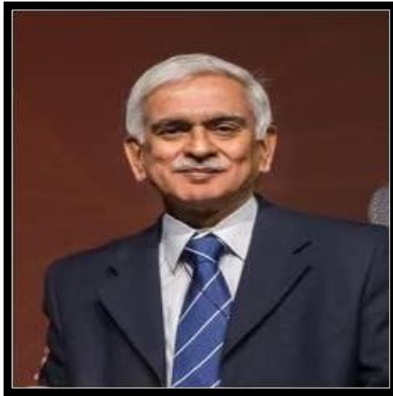


**Dr. Catherine von Reyn**, received her BS in Mechanical Engineering from the Georgia Institute of Technology ('Georgia Tech') and her PhD in Bioengineering from the University of Pennsylvania. She was a postdoctoral associate in Gwyneth Card's laboratory at the Janelia Research Campus of the Howard Hughes Medical Institute (HHMI) and is currently an assistant professor in the School of Biomedical Engineering, Science and Health Systems at Drexel University. Dr. von Reyn's research uses diverse, cutting edge techniques, including cell type-specific genetic engineering, whole-cell patch clamp in behaving animals, modeling, and detailed behavioral analysis to identify and characterize sensorimotor circuits. She applies these techniques to engineer de novo synaptic connections in sensorimotor circuits with the long-term goal of developing genetic strategies for re-engineering circuits in neurodegenerative and neurophysiological disorders.



**Dr. Sriram Balasubramaniam** is currently working as Associate Professor at School of Biomedical Engineering, Sciences and Health systems in Drexel University, USA. He focuses on developing novel biomedical engineering methods and advanced computational modeling tools to address the burden of debilitating pediatric skeletal deformities, traumatic injuries and orthopedic conditions. His research expertise also includes experimental biomechanics in the areas of pediatric head and neck injury, orthopedic biomechanics of the knee, thorax and spine, and knee biomechanics following ligament reconstructions.





Dr. G. S. Bhuvaneshwar is presently an independent Consultant in Medical Device design, development, testing and Quality management. He is an Adjunct Professor in the Department of Engineering Design at the Indian Institute of Technology, Madras. For the past years, he was Director of Innovation at Trivitron Healthcare Pvt, LTd., Chennai a leading company of Indian origin with a major focus in In vitro Diagnostics. Previously, he retired as Head of the Biomedical Technology Wing of Sree Chitra Tirunal Institute for Medical Sciences & Technology (SCTIMST), Trivandrum.



Dr. N. Kumarappan received the Graduate degree from Madurai Kamaraj University, Tamil Nadu, India in 1982, the Post-Graduate degree from Annamalai University, Annamalai Nagar, India in 1989 and the Ph.D. degree from CEG Anna University, Tamil Nadu, India in 2004 under QIP fellowship AICTE, India. He is the former head of the department and currently a Professor with the Department of Electrical Engineering, Faculty of Engineering and Technology, Annamalai University. He is having 33 years of experience to his credit as an educator and researcher. He has published more than 130 international journal and conference papers. International Journal Papers – 44, IEEE digital library – 36, International conference – 65, National conference – 9. According to Google scholar his citation is 1010, H index 16. He was the outstanding reviewer for the Elsevier international journal of electric power and energy system 2015. He is an IEEE Madras Section Chairman, IEEE CIS Madras Chapter Chair and a Coordinator for more than 100 IEEE Madras Section organized FDP, Workshop, tech meet and TISP programs ect.. He was the recipient of the IEEE-NNS Outstanding Paper Travel Grant Award, Australia 2002, IEEE-WCCI Outstanding Paper Travel Grant Award, Canada 2016, the IEEE PES Student Program Award, USA 2003 and IEEE CIS Madras Chapter Best Chapter Award 2017. He was also a recipient of best researcher award, Annamalai University, 2018. His current research interests include power system operation and control, electricity price forecasting, EHV transmission fault diagnosis, FACTS devices, power system reliability, artificial intelligence techniques, micro grid, distributed generation and smart grid.

Dr. Kumarappan is a Life Fellow of the Institution of Engineer's (India) and a Life Member of the Indian Society of Technical Education. He was the recipient of the IEEE Madras Section Motivation Award in 2010 and the Certificate of Appreciation in 2012, third rank in the chapter activities 2015. He was the organizer of a special session in IEEE WCCI 2010 Spain, IEEE WCCI 2016 Canada and IEEE WCCI 2018 Brazil. He is a Speaker, Reviewer and Session Chair for IEEE/IET-U.K. international conferences and other international journals.



Mr. Balamurugan, Vice President HCL Technologies. Chennai. He was the Head of Medical Devices Development group in HCL which offers design and development services to leading medical device companies across the world.

# Physiology and Acoustics of Human Voice

Dr. Jan G. Svec

## Abstract:

In this overview lecture, we will get familiar with the principles of voice production in human body. We will take a look at the vibrations of the vocal folds and at the primary sound of voice generated due to the modulation of the airflow caused by these vibrations. Further, we will show how the sound changes when passing through the acoustic cavities of the vocal tract to form the final quality of the human voice. Finally, we will investigate the basic properties of the acoustic spectrum of voice and show how the voice sound can be separated into its components.



**Dr. Jan G. Svec**

Associate Professor, Department of Biophysics, Palacky University Olomouc, Czech Republic.  
Associate Research Scientist, Voice and Hearing Centre, Prague, Czech Republic.

Jan G. Svec, Ph.D. is a Czech physicist performing basic research on production of human voice. He holds MSc degree in fine mechanics and optics and PhD degrees in biophysics as well as in medical sciences. He has worked as a research scientist at the Center for Communication Disorders, Medical Healthcom, Ltd. in Prague, the Czech Republic, at the National Center for Voice and Speech in Denver, CO, USA and at the University of Groningen, the Netherlands. Currently he is at the Palacky University Olomouc, the Czech Republic and serves as an associate research scientist at Voice and Hearing Centre Prague, the Czech Republic. He designed videokymography, a method for high-speed visualization of vocal-fold vibrations, which is being used for advanced diagnosis of voice disorders. His research interests include acoustics, biomechanics, voice measurement methodology, singing voice.

## Medical imaging in computer aided diagnosis, focused on voice therapy

**Dr. Barbara Zitova**

### **Abstract:**

Computer-aided diagnosis is the computer-based system that helps specialists to take proper decisions. Medical imaging is focused on information in images acquired by doctors and it evaluates and analyzes abnormalities. The novel software tool for voice disorder diagnosis, based on videokymography will be introduced which increases the objectivity of the vocal fold examination and enables to follow the therapy progress.



**Dr. Barbara Zitova**  
HOD, Department of Image  
Processing, ÚTIA AV ČR, Prague,  
Czech Republic.

Dr. Barbara Zitova received M.Sc. degree in 1995 and Ph.D. degree in 2000 under computer science from the Charles University, Prague, Czech Republic. Since 1995 she has been with the Institute of Information Theory and Automation, Academy of Sciences of the Czech Republic, Prague and appointed as the head of the department on the 1st January 2008. She has more than a 15-years' experience in digital image processing and gives tutorials on Image Processing and Pattern Recognition at the Czech Technical University. Jointly with J. Flusser she gives specialized graduate course on moment invariants and wavelets at the Charles University and the Czech Technical University. She is an author of book chapter in Invariants for Pattern Recognition and Classification (M.A. Rodrigues ed., World Scientific, 2000), co-author of a monograph "Flusser Jan, Suk Tomáš, Zitová Barbara: Moments and Moment Invariants in Pattern Recognition, Wiley & sons, (Chichester 2009)", and of 50 journal and conference papers on related topics. Her paper "Image Registration Methods: A Survey", Image and Vision Computing, vol. 21, pp. 977-1000, 2003, has become a major reference in image registration with more than 2000 citations.

## The Prospects Of Engineering In Medicine –A Spine Surgeon's Perspective

**Dr. Sudhir Ganesan**

### **Abstract:**

The realm of spine surgery in India began with Lord Krishna correcting spinal deformity of his devotee as early as 3500 BC. In western world, reports for correction of spinal deformity is documented in Edwin Smith Papyrus in 1700BC. Since then the progress in the field of spine has been phenomenal in the western world but there hasn't been anything indigenously phenomenal from India. At some point from our ancient Literature to today we have purposely removed our thinking hat and replace it with an in vogue hat and started following the crowd, unmindful of what the destiny might. Lack of trust in our literature or believing the rulers of our Land in the past, have something better that we don't, may have propelled us to this point. With time we can see the reinstallation of our ancient believes in turmeric, salt, clove and coal scientifically in the western world. So it's time to remove our in vogue hat and put our thinking cap on. If we do so, we can see innumerable opportunities to create, innovate and have our own mark, especially when the potential of biomedical engineering is rightly tapped and put to use in the field of Spine Surgery.



**Dr. Sudhir Ganesan**  
Ortho Spine Surgeon, SRIHER,  
Chennai.

Dr. Sudhir Ganesan completed his under graduation from Thoothukudi Government Medical College, Dr. MGR Medical University. He pursued his post-graduation in Orthopedic surgery in the prestigious Ganga Hospital, Coimbatore and was awarded the Dr. Balu Sankaran Gold Medal, for securing All India first rank in DNB Orthopaedics. His passion towards the field of spine surgery, lead him to pursue his FNB super speciality in Sir Ganga Ram Hospital, Delhi. He was awarded the best outgoing Fellow of FNB spine. To hone his skills further, he went as an observer in Spine Surgery to Queen's Medical Centre, Nottingham, UK. He was awarded the Tamil Nadu State Orthopaedic Association Foreign Fellowship Award 2017 which he utilized for his training in endoscopic spine surgery from St. Anna's Hospital, Herne, Germany. He was also the recipient of AO Asia Pacific Research Grant Award 2017 for his molecular project. He has been elected as an AO Spine Asia Pacific Delegate Council Member 2019-2020. His area of interests are Degenerative disorders, Trauma, deformities, Endoscopic Spine Surgery and Minimally Invasive Spine Surgeries. His Spine surgery count spans around 700 per year. He strongly believes in knowledge sharing and enlightening the young minds and has thus been the key note speaker and faculty in various conferences.

# Affective Computing Applications in Healthcare

**Dr. M. Murugappan**

## Abstract:

Affective computing or Artificial Emotional Intelligence (AEI) is the area of computer science, which mainly aims to measure, simulate, understand, and react to human emotions. This field of research utilizes information from several types of sensors to collect the user's emotional responses. Some of the most types of sensors used for understanding emotional responses are the camera (recording facial emotion expressions), the microphone (vocal emotional speech responses), biosensors (physiological changes), and gestures (body postural changes). Then advanced software and Artificial Algorithms process the information into actionable data. There has been copious academic research into applying affective computing in many areas, but the technology is still at a relatively nascent stage. This field of research has been widely applied in several areas such as marketing, economics, e-learning, e-governance, gaming, insurance, customer service, etc. In recent years, the application of affective computing in healthcare emerges in multiple areas, emotional impairment detection in neurological disorders, psychological counselling, developing a communication tool for autistic patients, emotional stress assessment, stress behavior analysis, clinical diagnosis, etc. The present talk is going to review the general state of the art in Affective Computing and its applications in medicine, challenges in developing assistive systems, and explore the potential applications used in clinical diagnosis. Finally, I will conclude the talk by presenting some experimental results related to affective computing applications in neurological disorders.



**Dr. M. Murugappan**

Associate Professor, Department of  
Electronics and Communication  
Engineering, Kuwait College of Science and  
Technology, Doha, Kuwait

M. Murugappan received the B.E (Electrical & Electronics Engg) (Distinction) from Madras University, India, M.E (Applied Electronics) from Anna University, India, and Ph.D. (Mechatronic Engineering) from University Malaysia Perlis, Malaysia, in 2002, 2006, and 2010, respectively. He is currently working as an Associate Professor in the Department of Electronics and Communication Engineering, Kuwait College of Science and Technology (KCST) (Private University), Kuwait since Feb 2016. He has gained more than 10 years of post-Ph.D. teaching and research experience from different countries (India, Malaysia, and Kuwait). He has received several research awards, medals, and certificates on excellent publications and research products. Recently, he has been included in the Top 2% of Scientist in the world in Experimental Psychology and Artificial Intelligence by Stanford University researchers. He has published more than 110 research articles in peer-reviewed conference proceedings/journals/book chapters. He has got a maximum citation of 3900 and the H index of 33 and i10 Index of 66 (Ref: Google Scholar citations).

## Brain stimulation methods in psychiatry

### Dr. S. Kailash

#### Abstract:

In the management of mental health disorders and in psychiatric research various brain stimulation methods are in day to day use. These brain stimulation methods use electrical currents or magnetic fields to modulate the neuronal firing of the brain and thereby produce the desired effect. This electric current or magnetic field is applied to the brain either transcranially or directly by surgical implantation of electrodes. The common transcranial methods of brain stimulation are Electroconvulsive therapy (ECT), Transcranial magnetic stimulation (TMS), Cranial electrical stimulation (CES), Transcranial direct current stimulation (tDCS), also called direct current polarization, Magnetic seizure therapy (MST). The important surgical methods of brain stimulation include Deep brain stimulation (DBS), Cortical brain stimulation (CBS), Vagus nerve stimulation (VNS). Salient features of these Brain stimulation techniques will be discussed in the presentation.



#### Dr. S. Kailash

Associate Professor,

Department of Psychiatry,

Chettinad Hospital and Research Institute

Qualification: M.D., D.N.B. Psychiatry;

Post-graduation from King George

medical university, Lucknow. Previously

worked at Centre for Addiction Medicine,

Adult and Geriatric psychiatry,

NIMHANS, Bangalore.

His area of interest includes Adult,

Geriatric and Addiction psychiatry,

Psychopharmacology. Recipient of

National young scholar award in

Psychiatry, Fellow of Indian Psychiatric

Society (FIPS); Member of National

academy of Medical Sciences (MNAMS)

Indian Psychiatric Society(IPS),

Convener - Spirituality and Mental health

Task Force 2020-21, Spirituality and

Yoga Committee, Member 2020-21.

Speaker in many National conferences.

Authored 8 indexed national and 17

international publications and 2 book

chapters.

## Making sense of Biosignal

**Dr. Hugo Gamboa**

### **Abstract:**

In this talk several types of biosignals will be covered explaining the challenges of collecting, processing extracting information and making sense of the physiological dynamics. The application of machine learning techniques to biosignals will also be covered to give a perspective on how to create automated decision mechanisms based on biosignals capture. During the several topics of the presentation, several research examples conducted at Nova University of Lisbon in collaboration with PLUX and Fraunhofer will be given.



### **Dr. Hugo Gamboa**

Associate Professor, Physics Department of the Sciences and Technology Faculty of the Universidade Nova de Lisboa.

Hugo Gamboa is an Associate Professor at the Physics Department of the Sciences and Technology Faculty of the Universidade Nova de Lisboa and member of LIBPhys-UNL. PhD in Electrical and Computer Engineering from Instituto Superior Técnico, University of Lisbon, He is a founder and President of PLUX, a technology-based innovative startup in the field of systems and wireless medical sensors, where he currently coordinates a multidisciplinary team that covers microelectronics, biosignal processing, and software development. He is also a Senior Scientist at Fraunhofer Portugal where he leads a team of 10 researchers and coordinates the Lisbon office of Fraunhofer Portugal.

## Neuromuscular electrical stimulation: Benefits from Machine Learning

**Dr. Deepak Joshi**

### **Abstract:**

Neuroprosthesis has shown promising potential in not only motor prosthesis but also in sensory and cognitive rehabilitation. It primarily includes the concept of neuroscience and biomedical engineering. Despite of pioneer success of invasive neuroprosthesis, the concerns related to surgical outcomes and associated complications limits the usefulness. Therefore, noninvasive methods like functional electrical stimulation and Transcranial Magnetic Stimulation have gain popularity in the past. The present keynote will briefly explain the introduction of electrical and magnetic stimulation and some preliminary simulation results in Spinal Cord injury patients. Further, the talk will address the present limitation with stimulation techniques and possible solutions with machine learning. The keynote will also discuss the challenges and future scope of such stimulations in neuromotor research.



**Dr. Deepak Joshi**

Assistant Professor, Centre for  
Biomedical Engineering, Indian Institute  
of Technology, Delhi.

Dr. Deepak Joshi received his PhD in biomedical engineering from Indian Institute of Technology (IIT) Delhi. He has been working in the area of neural prosthetic design and development for last ten years. During his PhD he developed a prototype lower limb prosthesis controlled by contra lateral limb which earned Invention award from Intellectual Ventures Asia. During his tenure at Department of electrical and computer engineering in National University of Singapore (NUS), he worked on development of artificial hand with integrated sensors to create an illusion of touch from human hand. This work demonstrated a significant impact on the social acceptance of upper limb prosthesis and was reported to be the most popular article in IEEE Transaction on neural system and rehabilitation engineering. His research work at Institute of Neuroscience (ION), Newcastle University in United Kingdom (UK) discovered that artificial proprioception can significantly improve the myoelectric control in upper limb amputee. During his postdoctoral at University of Oregon in United States of America (USA), he worked on integration of various sensing modalities to provide seamless transitions in lower limb prosthesis.



## Languages and emotions: Towards speech technology development for under-resourced languages and secondary emotions

**Ms. Jesin James**

### **Abstract:**

In the first part of this keynote address, the focus will be on emotional speech synthesis for healthcare robots. Many of these social robots can interact with humans using their voice. This research identified the type of voice that is needed for a healthcare robot interacting with humans, and synthesized voice for the healthcare robot. A perception test was conducted it was found that the participants could perceive high levels of empathy from a healthcare robot speaking with a synthesised voice containing the five secondary emotions. This study was the first of its kind analysing secondary emotions, which are much needed for future human-robot interaction applications.

Another aspect that will be addressed in this keynote speech will be speech technology development in under-resourced languages. The other languages that have limited resources for speech technology development are called under-resourced languages (Ex: Malayalam, te reo Maori, New Zealand English, Indian English). A classic example is how we have to alter the way we talk English so that Siri/Alexa understand us, because we speak Indian English and the technology is most often trained on American/British English. Our language knowledge and technological abilities can help in the development of resources for our under-resourced languages. Speech technology resource development conducted for Malayalam, te reo Maori and New Zealand English will be explained, along with the necessary next steps in the area.



**Ms. Jesin James**

Lecturer, Department of Electrical, Computer, and Software Engineering at the University of Auckland, New Zealand. Jesin is a Lecturer in the Department of Electrical, Computer, and Software Engineering at the University of Auckland, New Zealand. Jesin's main research areas are speech signal processing, under-resourced languages, machine learning and engineering education. During 2012-2014 she worked on developing a Text To Speech synthesis system in her mother tongue, Malayalam (language spoken in south India), with special emphasis on developing a prosodic model for the same. Further, during 2014-2016 she worked as a Lecturer in Electronics and Communication Engineering in India, during which she was involved in some student projects that have been published and presented to wider audience. In 2016, she joined University of Auckland as Ph.D scholar in the Computer systems engineering. She is currently working in the Healthcare Robot project (part of CARES, University of Auckland), with special focus on the development of the speech of the robot in New Zealand English. The aim is to improve the naturalness of the speech and incorporate empathy into the robot voice, to further improve its acceptance.

# Artificial Intelligence for Applications in Neurology

**Dr. Justin Dauwels**

## **Abstract:**

Many tasks in medicine still involve substantial manual work. In many cases there is strong potential for intelligent automation by Artificial Intelligence (AI), leading possibly to a reduction in costs and man-hours, while increasing the quality of clinical service. In this talk, we will consider applications of AI in the domain of neurology.

We are developing a low cost validated system to automatically interpret EEG via remote access. Diagnosis and management of neurological disorders rely on visual review of EEG data by specialized physicians. As the duration of EEG recordings ranges from 30 minutes to several weeks, the visual review is time consuming, and accounts for approximately 80% of total cost associated with EEG reading. Our system has the potential to reduce expenses associated with EEG testing and allows physicians to devote more quality time to their patients. One of the applications that we have explored so far is diagnosis of epilepsy of EEG. In this talk, we will show numerical results on large EEG datasets of epilepsy patients and healthy control subjects for multiple centers.



**Dr. Justin Dauwels**

Associate Professor, Circuits and Systems (CAS), Department of Microelectronics, Delft University of Technology, Netherlands

Dr. Justin Dauwels is an Associate Professor at the TU Delft (Circuits and Systems, Department of Microelectronics). He was an Associate Professor of the School of Electrical and Electronic Engineering at the Nanyang Technological University (NTU) in Singapore till the end of 2020. He was the Deputy Director of the ST Engineering – NTU corporate lab, which comprises 100+ PhD students, research staff and engineers, developing novel autonomous systems for airport operations and transportation. His research interests are in data analytics with applications to intelligent transportation systems, autonomous systems, and analysis of human behaviour and physiology. He obtained his PhD degree in electrical engineering at the Swiss Polytechnical Institute of Technology (ETH) in Zurich in December 2005. Moreover, he was a postdoctoral fellow at the RIKEN Brain Science Institute (2006-2007) and a research scientist at the Massachusetts Institute of Technology (2008-2010)

**2021 IEEE Seventh  
International Conference on  
Bio signals, Images and  
Instrumentation  
ICBSII 2021**

**SESSION I  
RESEARCH  
PAPERS**

# Detection of Autism Spectrm Disorder from EEG signals using pre-trained deep convolution neural networks

Qaysar Mohi-ud-Din  
*Dept. of Biomedical Engineering*  
 SRM Institute of Science and Technology, Kattankulathur  
 Chennai, India  
 qm6945@srmist.edu.in

A. K. Jayanthy  
*Dept. of Biomedical Engineering*  
 SRM Institute of Science and Technology, Kattankulathur  
 Chennai, India  
 jayanthk2@srmist.edu.in

**Abstract**-Electroencephalography (EEG) is a method of recording the electrical activity of the brain and it has been used to detect various neurological disorders which are associated with the abnormal brain electrical activity. Various diseases such as epilepsy and Alzheimer's disease are being detected using EEG. Autism spectrum disorder(ASD) which is a neurological disorder impairs the socialization, communication and behavior of the subjects suffering from it. Various studies have been conducted to find the EEG abnormalities in ASD subjects and normal controls, so that ASD can be detected using EEG. Various feature extraction techniques have been employed to extract features from EEG signals of the subject which can help in the classification of ASD and normal controls. Machine learning and deep learning networks have been used to classify normal and abnormal EEG waveforms and several studies have used Convolution neural networks (CNN) for the classification. In this study, we have used transfer learning approach to train the pre-trained CNNs, GoogLeNet and SqueezeNet for classifying ASD subjects and normal controls using their EEG signals. The accuracy achieved using the GoogLeNet and SqueezeNet were 75% and 82% respectively and the results obtained indicate that this method can assist in classifying ASD subjects and normal subjects using EEG signals.

**Keywords**- ASD, EEG, CNN, GoogLeNet, SqueezeNet

## I. INTRODUCTION

Electroencephalography (EEG) is a low cost, non-invasive and easy to use technique, used to record the electrical impulses produced by the brain. The brain waves can be classified into different bands of different frequency ranges and each frequency band corresponds to different mental states [1]. EEG signals are non-stationary, time varying signals and different methods have been applied to extract

features from EEG signals for classification [2] [3]. Different features have been extracted from EEG signals using different techniques for classification. Discrete wavelet analysis has also been performed on the EEG data in the past to classify EEG signals [4] [5].

EEG has been used to detect various neurological diseases such as epilepsy, Alzheimer's disease, sleep disorders etc. Machine learning techniques have been used for the automated classification of epilepsy using the EEG data [6] [7]. Pre-trained deep learning neural networks have been employed to classify motor imagery EEG [8]. A three dimensional CNN was used for the automated detection of seizures based on multi channel EEG data [9]. Channel-wise scalograms have been generated from EEG data and de-noising auto encoders employed for the epileptic seizure detection [10]. Brain waves were used to determine the sex of subjects using CNN and an accuracy of more than 80% was achieved in identifying that the brain waves are different in male and female subjects [11]. Convolution neural networks have also been used for the multi class motor imagery classification [12] and 13 layered CNN was used to detect seizure using EEG data [13].

Autism spectrum disorder are a group of complex neurodevelopment disorders, such as, Asperger syndrome, pervasive development disorders not otherwise specified, autism and other related disorders [14]. ASD subjects have deficits in socialization, communication and behavior [15] [16]. According to Center for Disease Control and Prevention (CDC), approximately 1 in 54 children aged 8 years suffer from ASD in USA [17]. ASD is a

complex disorder and its symptoms overlap with other psychiatric disorders. So, it is important to use the appropriate instruments and scales to diagnose subjects with ASD [18].

EEG abnormalities were found in gamma band between the autism subjects and control subjects [19]. Studies also reported frontal EEG power differences between ASD subjects and control subjects within first year after the birth [20]. Extraction of non linear features from EEG data was performed and used by statistical learning methods for the classification of ASD and non ASD subjects [21].

In this study, Scalograms, which are visual representations of wavelet transform of a signal and plotted as a function of time and frequency were channel wise generated from the EEG data. The scalograms were then used to train pretrained neural networks GoogLeNet and SqueezeNet to classify the ASD subjects and the normal controls. GoogLeNet deep CNN was proposed by researchers at Google and it was the winner at ILSVRC 2014 image classification challenge [22]. It has 22 neuron layers and has been used for computer vision tasks such as face detection and recognition. SqueezeNet deep CNN was released in 2016 and was developed by researchers at DeepScale, University of California, Berkeley and Stanford University. The SqueezeNet neural network is 18 layers deep [23].

## II. METHODS

The EEG data used in this study for the automated detection of ASD was obtained from King Abdul Aziz University, Saudi Arabia. The dataset consists of data from 13 ASD subjects and 4 normal controls. The 16 channel EEG data, sampled at 256 Hz, was recorded according to the international 10-20 electrode placement system. EEG data was then low pass filtered (50 Hz) and high pass filtered (0.5Hz). The data was segmented by taking 1000 data points per segment. A total of 4768 segments were formed and the data was labeled as "ASD" for autism subjects and "NOR" for control subjects. There were 2720 segments belonging to ASD subjects and 2048 belonging to control subjects. The workflow of this study has been given in Fig. 1.

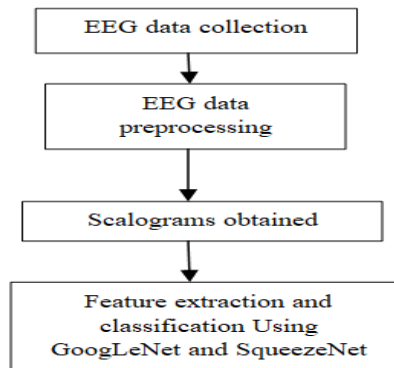


Fig. 1 Workflow of the study

A scalogram, which is a visual representation of wavelet transform of a signal and plotted as a function of time and frequency were channel wise generated from the EEG data and a sample scalogram is shown in Fig. 2.

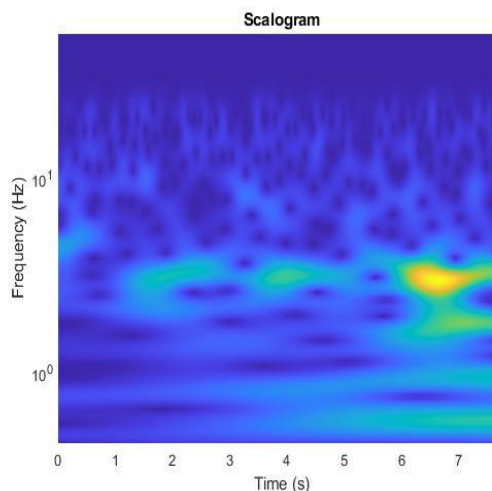


Fig. 2 A sample scalogram obtained from EEG waveform

Scalogram for signal with 1000 samples was generated as RGB images. The RGB image is an array of size 224x224x3. All the images were divided into training, validation and testing datasets. A total of 4768 images were formed in which 2720 belonged to ASD category and 2048 belonged to Normal category. 3338 images were used for training the networks, 476 images were used to validate the networks and 954 images were used to test the networks.

Pre-trained GoogLeNet neural network and SqueezeNet neural network were used for the

classification of ASD and control subjects using the generated RGB images. The input size of images for GoogLeNet should be  $224 \times 224 \times 3$ . The GoogLeNet network is pre-trained to classify images into 1000 categories and we have retrained the network. Each layer of the GoogLeNet acts as a filter. The earlier layers extract the common features of the images and the subsequent layers extract more specific features. The convolution layers extract the features and the final classification layer is used for classification. A dropout layer was used to prevent over-fitting. The training aims at achieving better accuracy and minimizing the loss function.

Another neural network called SqueezeNet was also trained to classify ASD subjects and control subjects. The input size of  $227 \times 227 \times 3$  is required for SqueezeNet and the images were therefore resized to  $227 \times 227 \times 3$ . The training, testing and validation samples sizes were same as used for GoogLeNet. The study was carried out using Matlab 2020a and Matlab toolboxes such as deep learning toolbox, signal processing toolbox and wavelet analyzer have been used in this study.

### III. RESULTS

The data was divided into training, validation and testing datasets. 3338 images were used to train the network. The validation dataset consisted of 476 images and was used to validate the network. The testing data was used to evaluate the network and 954 images were used for testing.

Stochastic gradient descent with momentum optimizer was used for training the neural networks. The minibatchsize was set to 10, the number of epochs was 20 and 4440 iterations were performed. The initial learning rate of 0.0001 was used. It took 806 minutes to train the neural network on a single CPU. The validation accuracy 75% was achieved by the network and the accuracy achieved using the testing data was 74%. The GoogLeNet training process is shown in Fig. 3.

After using GoogLeNet, we used SqueezeNet for classification of ASD and control subjects. The RGB images generated were used to train the neural network. The RGB images generated for GoogLeNet were resized into  $227 \times 227 \times 3$  format for training the SqueezeNet neural network. The number of training, validation and testing images were the same as used for GoogLeNet. The training set consisted of 3338 images. The network was validated using a validation set which consisted of 476 images.

After 10 epochs and 4995 iterations, the validation accuracy of 82.98% was achieved by the network. The accuracy achieved using the test data was 80.71% in classifying ASD and normal controls. The test set was used to evaluate the model and consisted of 954 images. The SqueezeNet neural network achieved an accuracy of 80.71% using the test set. The SqueezeNet training process is shown in Fig. 4.

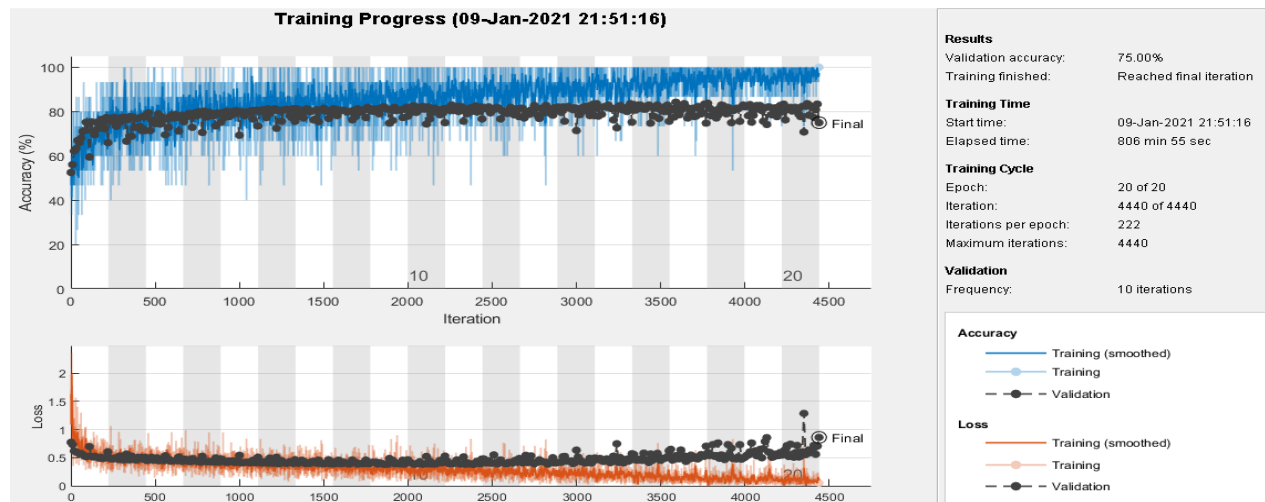


Fig. 3. GoogLeNet deep CNN training process

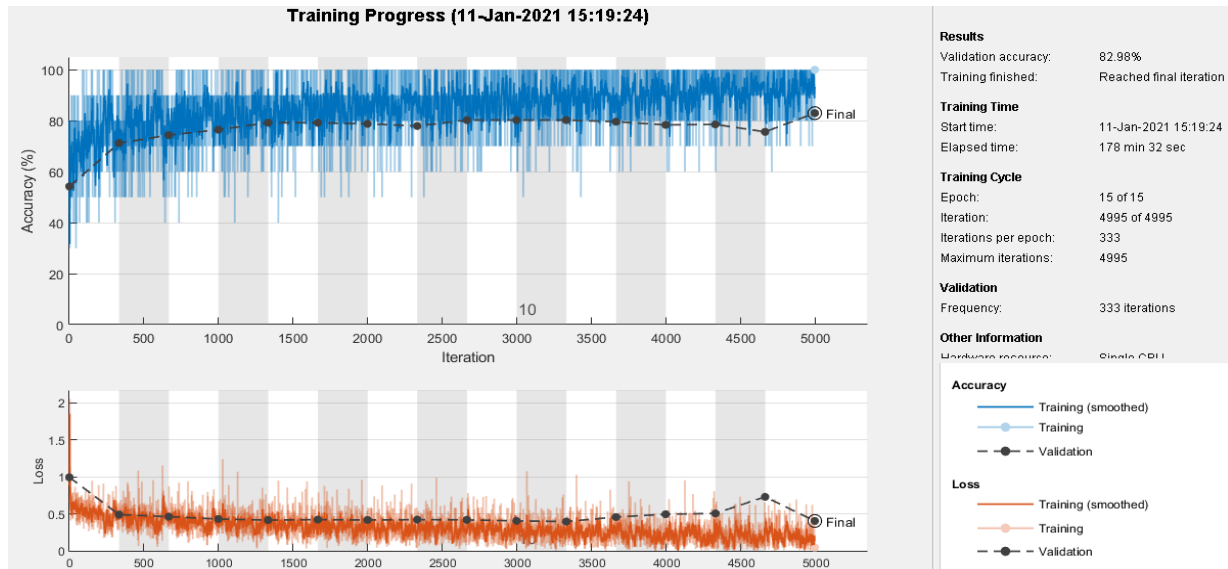


Fig. 4. SqueezeNet deep CNN Training process

#### IV. DISCUSSION

The severity and the symptoms of ASD vary and it is difficult to diagnose. The diagnosis is done by professionals and various tests and behavioral patterns are observed. Various studies have used EEG signals and used different feature extraction techniques for the detection of various neurological diseases, such as epilepsy, Alzheimer's disease, dementia, etc. Various studies have used Machine learning algorithms for classification of ASD subjects and normal subjects using EEG data. The classification accuracy of GoogLeNet CNN and SqueezeNet CNN has been shown in Fig. 5.

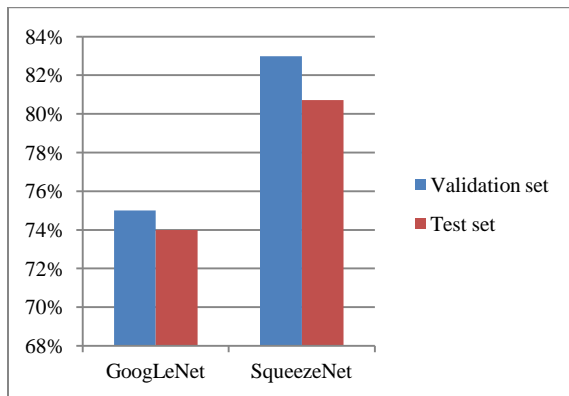


Fig. 5. Classification accuracy of GoogLeNet and SqueezeNet

In this study, transfer learning approach was adopted to use GoogLeNet and SqueezeNet, which are pre-trained neural networks for classification of images. We re-trained them using the RGB images generated from scalograms. The scalograms were generated from EEG data using the wavelet transform. The results indicated that the classification accuracy produced by SqueezeNet was higher than the classification accuracy obtained by GoogLeNet in classifying ASD. The same number of images were used to train both the neural networks but the time taken by SqueezeNet for training was much less as compared to GoogLeNet.

The GoogLeNet and SqueezeNet deep CNNs extract image features from the RGB images and then classify these RGB images using the generated features. The scalograms belonged to either "ASD" category or "NOR" category and the deep CNNs classified the obtained scalograms into these two categories. The results indicated that the SqueezeNet deep CNN is better at classifying the EEG signals of ASD subjects and normal subjects using the scalograms obtained from the EEG signals.

#### V. CONCLUSION

The transfer learning approach used in this study to classify ASD subjects and control subjects using EEG data achieved a good accuracy and the results indicate that this method can be used to

classify ASD subjects and normal controls. The limitations of this study are that this study was performed using a single CPU which led to larger training time. The dataset size also could be increased and several other pre-trained neural networks could be used to classify the ASD subjects and normal subjects.

## VI. ACKNOWLEDGEMENT

We thank Prof. Mohammed Alahddad (KAU University, Jeddah, Saudi Arabia) for providing the EEG data used in this study. We also thank SRM Institute of Science and Technology for providing us the facilities to conduct this research.

## VII. REFERENCES

- [1] M. Teplan, "Fundamentals of EEG measurement," *IEEE Measurement Science Review*, vol. 2, pp. 1-11, 2002.
- [2] S. Dodia, DR. Edla, A. Bablani, D. Ramesh, V. Kuppil, "An efficient EEG based deceit identification test using wavelet packet transform and linear discriminant analysis," *J Neurosci Methods*, vol. 15, pp. 31-40, Feb 2019.
- [3] P. Jahankhani, V. Kodogiannis, K. Revett, "EEG signal classification using wavelet feature extraction and neural networks," *IEEE John Vincent Atanasoff 2006 International Symposium on Modern Computing*, pp. 120-124, 2006.
- [4] HU. Amin, AS Malik, RF. Ahmad, N. Badruddin, N. Kamel, M. Hussain, WT. Chooi, "Feature extraction and classification for EEG signals using wavelet transform and machine learning techniques," *Australas Phys Eng Sci*, vol. 38, pp. 139-49, March 2015.
- [5] Lung, Chuin Cheong, Sudirman, Rubita, Hussin, Siti Suraya, "Feature extraction of EEG signal using wavelet transform for autism classification," *ARNP Journal of Engineering and Applied Sciences*, vol. 10, pp. 8533-8540, 2015.
- [6] B. Abbasi, DM. Goldenholz, "Machine learning applications in epilepsy," *Epilepsia*, pp. 2037-2047, Oct. 2019.
- [7] F. Rosenow, KM. Klein, HM. Hamer, "Non-invasive EEG evaluation in epilepsy diagnosis," *Expert Rev Neurother*, vol. 15, pp. 425-44, Apr. 2015.
- [8] G. Xu et al., "A deep transfer Convolutional neural network framework for EEG signal classification," in *IEEE Access*, vol. 7, pp. 112767-112776, 2019.
- [9] X. Wei, L. Zhou, Z. Chen, L. Zhang, Y. Zhou, "Automatic seizure detection using three-dimensional CNN based on multi-channel EEG," *BMC Med Inform Decis Mak*, vol. 18, pp. 0693-8, 2018.
- [10] Y. Yuan, G. Xun, K. Jia, A. Zhang, "A multi-context learning approach for EEG epileptic seizure detection," *BMC Syst Biol*, vol. 12, 2018.
- [11] MJAM van Putten, S. Olbrich, M. Arns, "Predicting sex from brain rhythms with deep learning," *Sci Rep*, vol. 8, 2018.
- [12] X. Lun, Z. Yu, T. Chen, F. Wang, Y. Hou, "A simplified CNN classification method for MI-EEG via the electrode pairs signals," *Front Hum Neurosci*, vol. 14, Sep. 2020.
- [13] U. Acharya, SL. Oh, Y. Hagiwara, JH. Tan, H. Adeli, H, "Deep convolutional neural network for the automated detection and diagnosis of seizure using EEG signals," *Computers in biology and medicine*, vol. 100, pp. 270-278, 2018.
- [14] H. Faras, N. Al Ateeqi, L. Tidmarsh, "Autism spectrum disorders," *Ann Saudi Med*, vol. 30, pp. 295-300, 2010.
- [15] L. Zwaigenbaum, JA. Brian, A. Ip A, "Early detection for autism spectrum disorder in young children," *Paediatr Child Health*, vol. 24, pp. 424-443, 2019.
- [16] C. Lord, M. Elsabbagh, G. Baird, J. Veenstra-Vanderweele, "Autism spectrum disorder," *Lancet*, vol. 11, pp. 508-520, Aug. 2018.
- [17] KA. Shaw, MJ. Maenner, J. Baio, A. Washington, DL. Christensen, LD. Wiggins, S. Pettygrove, JG. Andrews, T. White, CR. Rosenberg, JN. Constantino, RT. Fitzgerald, W. Zahorodny, J. Shenouda, JL. Daniels, A. Salinas, MS. Durkin, PM. Dietz, "Early Identification of Autism spectrum disorder among children aged 4 years - early Autism and developmental disabilities monitoring network, Six Sites, United States, 2016," *MMWR Surveill Summ*, vol. 69, pp. 1-11, March 2020.
- [18] C. Kupper, S. Stroth, N. Wolff, et al, "Identifying predictive features of autism spectrum disorders in a clinical sample of adolescents and adults using machine learnin," *Sci Rep*, Vol. 10, 2020.
- [19] G. Shou , MW. Mosconi, LE. Ethridge, JA. Sweeney, L. Ding, "Resting-state gamma-band EEG abnormalities in Autism," *Annu Int Conf IEEE Eng Med Biol Soc*, 2018, pp. 1915-1918, Jul. 2018.
- [20] LJ. Gabard-Durnam, C. Wilkinson, K. Kapur, K. et al, "Longitudinal EEG power in the first postnatal year differentiates autism outcomes," *Nat Commun*, vol. 10, 2019.
- [21] WJ. Bosl, H. Tager-Flusberg, CA. Nelson, "EEG analytics for early detection of autism spectrum disorder: a data-driven approach," *Sci Rep*, vol. 8, 2018.
- [22] Christian Szegedy, W. Liu, Y. Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, D. Erhan, V. Vanhoucke, Andrew Rabinovich, "Going deeper with convolutions," 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1-9, 2015.
- [23] FN. Iandola , MW. Moskewicz, K. Ashraf, S. Han, WJ. Dally, K. Keutzer, Squeezenet: alexnet-level accuracy with 50x fewer parameters and <1mb model size, arXiv preprint, 2016.



# Cloud based analysis and classification of EEG signals to detect epileptic seizures

1<sup>st</sup> Munyaradzi Charles Rushambwa  
*Electronic Engineering*  
*Harare Institute of Technology)*  
 Harare, Zimbabwe  
 mrushambwa@hit.ac.zw

2<sup>nd</sup> Mavis Gezimati  
*Electronic Engineering*  
*Harare Institute of Technology*  
 Harare, Zimbabwe  
 mgezimati@hit.ac.zw

5<sup>th</sup> Dr. Govindaraj P  
*Information Science Department*  
*AJ Institute of Engineering and Technology*  
 Manglore, India  
 govindaraj@ajiet.edu.in

4<sup>th</sup> Rajkumar Palaniappan  
*Department of Mechatronics Engineering*  
*College of Engineering,*  
*AMA International University*  
 Salmabad, Bahrain  
 r.palaniappan@amaiu.edu.bh

5<sup>th</sup> Dr Vikneswaran Vijejan  
*Mechatronic Engineering*  
*University of Malaysia Perlis*  
 Perlis, Malaysia  
 vicky.86max@gmail.com

6<sup>th</sup> Fizza Ghulam Nabi  
*Institute of quality and*  
*technology Management (IQTM)*  
*University of the Punjab*  
 Lahore, Pakistan  
 engr.fizza@yahoo.com

**Abstract**—Epileptic seizures are explained as the abnormal electrical activity occurring in the brain due to an internal or external triggering factors. EEG (Electroencephalograph) is used to record brain activity and can be used to detect the seizures before, during or after they occur. These signal characteristics, however differ from patient to patient due to the different emotional and physical wellbeing of the various individuals. In normal circumstances, anti-epileptic medication is used to treat patients but very few systems have been developed to manage and track the seizures. In most extreme and rare cases, some patients undergo invasive surgery to treat the seizures and this is common in seizures that are caused by tumours or physical brain damage. Non-invasive surface electrode EEG measurement gives an estimate of the seizure onset but more invasive intracranial electrocorticogram (ECoG) are required at times for precise localization of the epileptogenic zone. This project aims at designing and implementing a device that can be used to detect and monitor the attention and meditation values of a person in real time. The system measures the EEG waves of the brain, performs feature extraction, classification and sends the control command over wireless to a remote controller. The remote controller in turn issues commands with corresponding brain wave frequency and sends it to the cloud for remote analysis and classification.

**Index Terms**—cloud, electroencephalograph, epilepsy, non-invasive, wireless.

## I. INTRODUCTION

The EEG signals we mainly focused on alpha and beta wave, in which the beta waves represents the awakening/reasoning state, whilst alpha is associated with meditation and relaxation. EEG are extracted non-invasively from the brain by means of the electrodes and then they are sent to the cloud for monitoring in real time. The recommended range of operation of EEG waves according to the IEEE is shown in table 1 [1].

The recording of the electrical signal is done along the scalp and the variations of voltage resulting from ionic current flow within the neurons of the brain[2-3]. The human brain is

TABLE I  
NEURO-FEEDBACK WAVE RANGES

Brain Wave	Range
Delta	0-4 Hz
Theta	4-8 Hz
Alpha	8-13 Hz
Beta	13-30 Hz
Gamma	30-44 Hz

made of a billion nerve cells named neurons and when they interact, they emit a measurable electrical impulse which we can measure using EEG. The project aim is to design a system for epileptic diagnosis, monitoring and analysis by making sure that these objectives are achieved:

- Obtain real time EEG signal from the frontal lobe using the neurosky mindwave sensor.
- Establish communication between the microcontroller and the web interface through the GSM modem.
- Perform feature extraction using fast Fourier transform on the raw EEG signal to obtain percentile values of attention, meditation and eye movement.
- Determine maximum feature value for the extracted features.
- Create correlation plots between the 3 extracted features.
- Create alert mechanism using the react control command in the web interface.
- Display in real time physical location of the system, values of attention, meditation and eye movement as a percentage of their operating frequencies to the web interface.

## II. SYSTEM DESCRIPTION

EEG signal acquisition, digitization and transmission in figure 1 are done using the TGAM1 board, using a single

channel dry electrode [4-6]. A dry single sensor is used to pick up voltages from the brain using the ear as the reference. The resultant voltage is subtracted through common mode rejection to convert it into a single channel EEG.

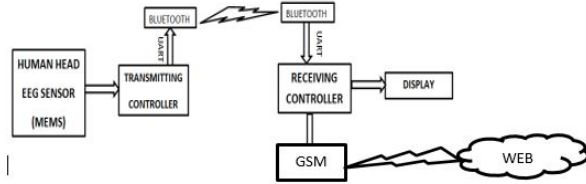


Fig. 1. Block Diagram of the Whole System.

The signal is then amplified at a gain of 2000 to enhance the faint EEG signals. The amplified signal is passed through digital low and high pass filters to get signals the signal in the range of 1-50Hz range, which is equivalent to the EEG frequency range[11]. The signals are then sampled at 512 hertz and FFT performed to separate the EEG rhythms for transmission. Figure 2 shows signal reception, feature extraction, display and control are done using an Arduino UNOR3 board.

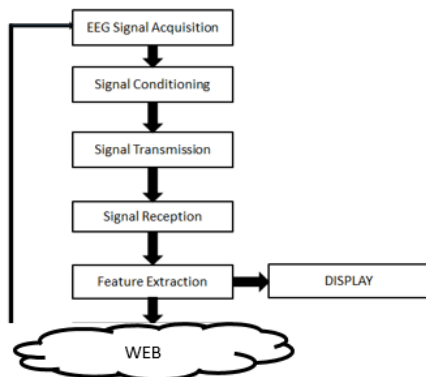


Fig. 2. Signal Flow Diagram.

HC05 Bluetooth module is used to automatically pair with the TGAM1 board and receive the transmitted digitized signal. The TGAM1 transmit data values encoded within ThinkGear packets as a serial stream of bytes via standard Bluetooth serial port profile (SPP) at a baud. Attention and meditation values were calculated and given on a scale between 0-100 as a percentage. A value between 40-60percent is considered a neutral value for both attention and meditation, which are given out as one-byte values that indicates the intensity of the user's level of mental focus which occurs during intense concentration and directed mental activity. Meditation value is an unsigned one-byte value that indicate the level of mental calmness or relaxation and given as a percentage in the range 0-100. Raw value data is delivered as an asynchronous serial stream of bytes from the TGAM1 to the receiving controller via Bluetooth.

### III. MATERIALS AND METHODOLOGY

#### A. EEG Sensor

The micro-electro-mechanical (MEMS) type human EEG acquisition unit is the TGAM1 (ThinkGear ASIIC Module) provided by Neurosky is shown in figure 3. The TGAM1 is a MEMS sensor embedded with a signal conditioning unit and a Bluetooth transceiver unit which communicates with the controller via UART protocol [3]. The TGAM1 type sensor consists of a reference electrode which is clipped on the ear of the subject and the frontal electrode which is placed on the forehead of the subject.

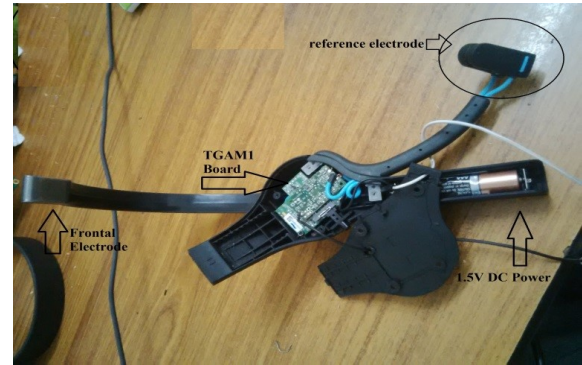


Fig. 3. EEG Sensor.

The sensor in figure 3 is used for signal processing and feature extraction of the raw digitized EEG signal through the TGAM1 board. After feature extraction the desired features (attention and meditation) are displayed on the web in real time to provide feedback of system action and analysis.

#### B. Signal Processing

The TGAM1 send information esteems encoded inside ThinkGear parcels as a sequential stream of bytes by means of standard Bluetooth sequential port profile (SPP) at a baud pace of 57600[7]. The ThinkGear information esteems that are removed from the encoded qualities are helpless sign quality, consideration and contemplation esteems. Consideration and contemplation esteems were determined and given on a scale between 0-100 as a rate. An incentive between 40-60percent is viewed as a nonpartisan incentive for both consideration and contemplation, which are given out as one byte esteems that shows the power of the client's degree of mental center which happens during extreme fixation and coordinated mental movement. Reflection esteem is an unsigned one-byte esteem that demonstrate the degree of mental tranquility or unwinding and given as a rate in the range 0-100.

Crude worth information is conveyed as a nonconcurrent sequential stream of bytes from the TGAM1 to the getting regulator by means of Bluetooth. The parcel comprises of a bundle header, payload and checksum. Parcel header is just 169 bytes and the relating information payload is a progression of bytes being sent. The payload checksum byte is utilized to check for the trustworthiness of the bundle's payload and

speaks to the aggregate of all bytes of the information payload. At the less than desirable end, the payload checksum is determined in programming by taking the most minimal 8 pieces of the whole and playing out one’s commendation opposite on the pieces[8]. The payload checksum is the contrasted and the checksum byte of the parcel and on the off chance that these match, at that point the bytes of the payload are parsed.

C. Connecting to the cloud

There are two methods to post data to the “Thingspeak”. They are the POST and the GET. While using POST the time taken was 1 sec. But by using GET it was reduced to 0.33 sec per field[9]. In POST we are using both httplib and urllib for sending data to server. So first it will get connected to thingspeak.com then it will be posting the data and after getting connected it will be posting the data to the respective channels. In GET we are directly using httplib and not using urllib. Instead of domain name IP address of site is given for establishing connection. There are 8 fields per channel and all of them can be simultaneously updated with a single value each. Therefore, at the rate of 1sec per reading it was taking nearly 40 mins for 2560 readings which was later reduced to 14 mins. It was taking 0.33 sec for reading. So total time taken was  $(.33*2560)/60=14$  minutes. [12-13] Figure 5 shows

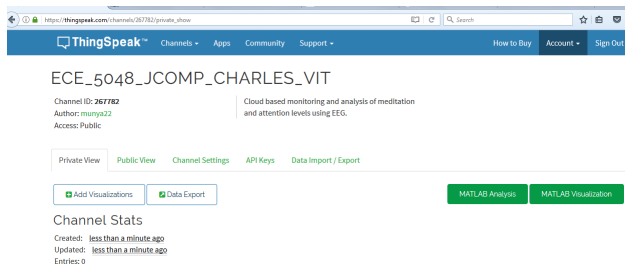


Fig. 4. Channel Status

the read and write API keys that are used to read and write to the cloud channel respectively. Programming interface keys empower you to compose information to a channel or read information from a private channel. Programming interface keys are auto-produced when you make another channel.

AT commands were used in code to activate the sensor and send data to the whe cloud by following these steps:

- Wake Seedstalker from sleep
- Power up GPRS Bee
- Pull data from sensors
- Connect to wireless network
- Connect to thingspeak
- Push data to thingspeak

the AT Commands used for pairing and configuring the bluetooth module to the wireless headset are shown in table 2.

ECE\_5048\_JCOMP\_CHARLES\_VIT

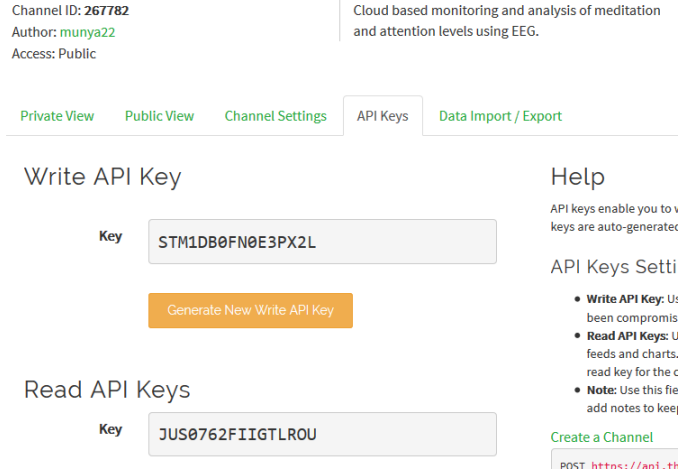


Fig. 5. API Key Generation

TABLE II  
NEURO-FEEDBACK WAVE RANGES

AT Command	Comment
AT	Checking module response
AT+NAME=SET PROJECT	Setting module name
AT+UART=57600,0,0	Setting baud rate
AT+ROLE=1	Setting module as master
AT+CMODE=0	Allowing connection to only bound address
AT+BIND=ODB5:2B:67B4	Used to automatically pair headset with Bluetooth module
AT+IAC=9E8B33	Setting default interpretation
AT+CLASS=0	Setting device class
AT+INQM=1,9,48	Setting query access patterns

IV. RESULTS

Figure 6 shows a 60 second variation plot of the attention and meditation values from the EEG sensor as they are used to determine the functioning of the human brain in real time , under different physical activities. The variation is used to track where there are high activities and low activities of attention and meditation values respectively.

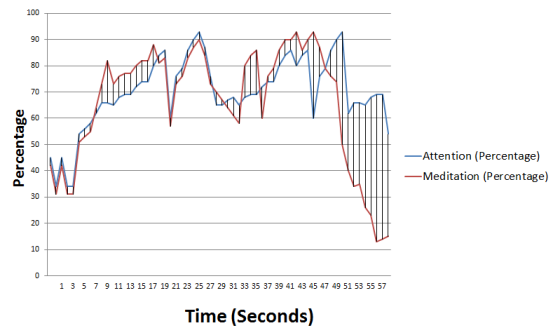


Fig. 6. 60 seconds Attention and Meditation Variation Test Plot

Figure 7 shows the corresponding serial monitor results of the real time values of the attention and meditation values

before they are sent to the cloud for analysis. A continuous irregular pattern of both meditation and attention is an indication of epileptic seizure.

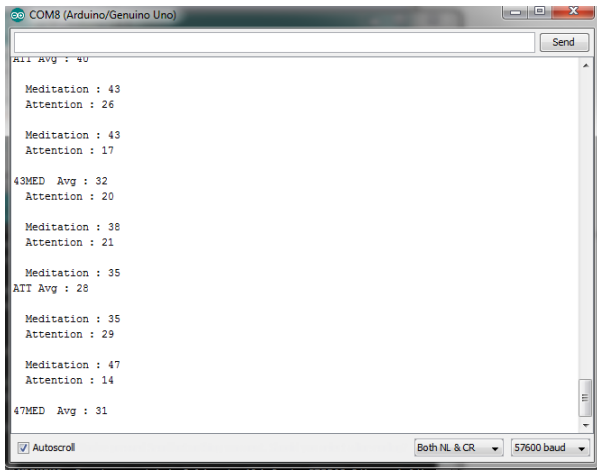


Fig. 7. Serial Monitor Output

Figure 8 shows a real time plot of the attention and meditation values plotted in real time from the cloud interface of thingspeak. A continuous irregular pattern of both meditation and attention is an indication of epileptic seizure.

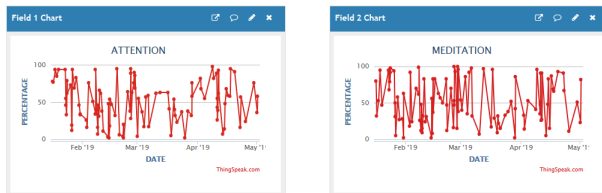


Fig. 8. a Attention values plotted in real time b Meditation values plotted in real time

Figure 9 shows eye movement values and the physical location of the system in real time. An irregular and non-continuous pattern of eye movement is an indication of an epileptic seizure. The location feature is used to show where the subject has been attacked at such that medical personnel can pin point the subject and attend to them. Figure 10 shows a

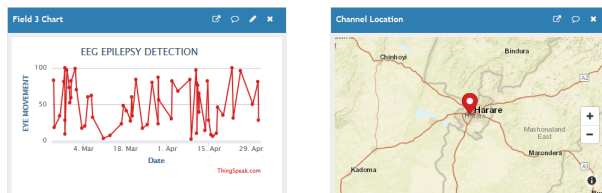


Fig. 9. a Eye movement values plotted in real time b Physical location of the system

2D plot comparison of Attention and Meditation values and a correlation plot of Attention and Eye Movement. A continuous irregular pattern of either meditation and attention or attention

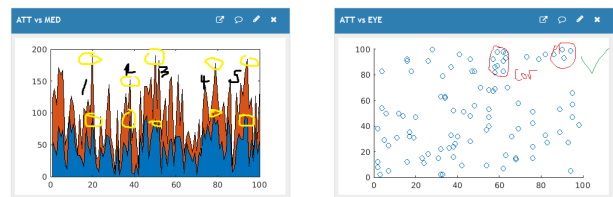


Fig. 10. a 2D plot comparison of Attention and Meditation values b Correlation plot of Attention and Eye Movement

vs eye movement is an indication of epileptic seizure. Figure 11 shows a Correlation plot of Attention and Meditation and a correlation plot of Meditation and Eye Movement. A continuous irregular pattern of either meditation and attention or ameditaion vs eye movement is an indication of epileptic seizure. Figure 12 shows a 2D plot comparison of Attention

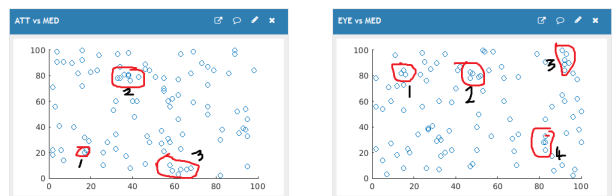


Fig. 11. a Correlation plot of Attention and Meditation b Correlation plot of Meditation and Eye Movement

and Eye movement values and a 2D plot comparison of Eye movement and Meditation values. A continuous irregular pattern of either meditation and eye movement or attention vs eye movement is an indication of epileptic seizure.

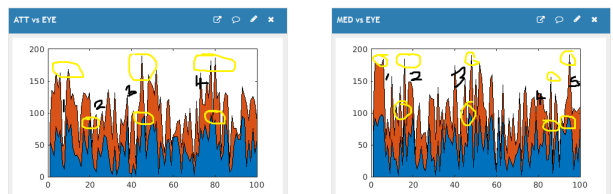


Fig. 12. a 2D plot comparison of Attention and Eye movement values b 2D plot comparison of Eye movement and Meditation values

Figure 13 shows gauge bar for Eye movement and an Indicator and Reaction control for epileptic detect.

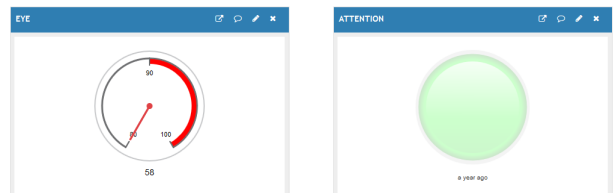


Fig. 13. a Gauge bar for Eye movement b Indicator and Reaction control for epileptic detect

## V. CONCLUSION AND DISCUSSION

### A. Analysis

A relatively low cost system to help the disabled can be used to assist the handicapped and physically challenged people to perform their daily function. Remote monitoring and analysis of eeg waves is possible using such a low cost system. A relatively low-cost system to detect early stages of epileptic seizures. Remote monitoring and analysis of EEG waves is possible using such a low-cost system. Results show that channel data classification using feature reduction performed better with 98 percent sensitivity. For high frequency analysis of frequencies between 60- 500Hz, the results show the same sensitivity yet less specificity when compared to the classification using lower frequency range of 0-30Hz. The results seen in this project show that PCA classifiers can be trained to classify the data as epileptic or nonepileptic with good accuracy. Even though training the classifiers took almost two hours, it was still noticeably less than other machine learning algorithms such as artificial neural networks. The accuracy of this algorithm can be improved. This project contributes to the detection of epilepsy by providing an automated classification method that allows the data to be sorted as epileptic or non-epileptic. While several commercially available software packages exist for assisting medical professionals in making this distinction, this method can be used to classify the data as either epileptic or non-epileptic based on currently available data. This allows the algorithm to be used for obtaining real time classification of data while recording. While, identification of epileptogenic zone will require the expert opinion of an epileptologist, the machine learning algorithm can be used to assist in signal analysis and classification.

### B. Future Work

The project is retrospective in nature and epilepsy detection is done in a number of different ways. There is need to carry out data collection for a number of subjects so as to have a comprehensive conclusion. There was need to take necessary precautions and record accurate measurements. It was also noted that there is need to get clearance from the relevant authorities when designing biomedical systems.

## VI. REFERENCES

- 1) M. C. Rushambwa and A. Mythili, "Impact assessment of mental subliminal activities on the human brain through neuro feedback analysis," 2017 Third International Conference on Biosignals, Images and Instrumentation (ICBSII), Chennai, 2017, pp. 1-6
- 2) Alzheimer's Association. 2015 Alzheimer's Disease Facts and Figures. Alzheimer's and Dementia 2015;11(3)332+.
- 3) Dario Assante, Claudio Fornaro, "Involving graduating engineers in applying a commercial brain-computer interface to motorized wheelchair driving", Global Engineering Education Conference (EDUCON) 2015 IEEE, pp. 446-452, 2015, ISSN 2165-9567
- 4) H. Ke, D. Chen, X. Li, Y. Tang, T. Shah and R. Ranjan, "Towards Brain Big Data Classification: Epileptic EEG Identification With a Lightweight VGGNet on Global MIC," in IEEE Access, vol. 6, pp. 14722-14733, 2018,
- 5) İşik, H., Sezer, E. Diagnosis of Epilepsy from Electroencephalography Signals Using Multilayer Perceptron and Elman Artificial Neural Networks and Wavelet Transform. J Med Syst 36, 1–13 (2012).
- 6) J. Katona, D. Peter, T. Ujbanyi, A. Kovari, "Control of incoming calls by a windows phone based brain computer interface", Computational Intelligence and Informatics (CINTI) 2014 IEEE 15th International Symposium on, pp. 121-125, 2014.
- 7) K.Crowley, A.Sliney, I.Pitt, D.Murphy," Evaluating a Brain-Computer Interface to Categorise Human Emotional Response", Proc. 10th IEEE Int. Conf. Adv. Learn. Technol., pp. 276-278, 2010.
- 8) Kimmi Aswini1, Y. Aruna Suhasini2, K. Rama Rao, Home Appliances Control Using Brain Wave Sensor by EEG, International Journal of Advanced Research in Computer Science and Software Engineering, Volume 5, Issue 10, October-2015 ISSN: 2277 128X
- 9) M. Hosseini, A. Hajisami and D. Pompili, "Real-Time Epileptic Seizure Detection from EEG Signals via Random Subspace Ensemble Learning," 2016 IEEE International Conference on Autonomic Computing (ICAC), Wurzburg, 2016, pp. 209-218
- 10) M. Hosseini, H. Soltanian-Zadeh, K. Elisevich and D. Pompili, "Cloud-based deep learning of big EEG data for epileptic seizure prediction," 2016 IEEE Global Conference on Signal and Information Processing (GlobalSIP), Washington, DC, 2016, pp. 1151-1155
- 11) Rangayyan, Rangaraj M. Biomedical signal analysis. Vol. 33. John Wiley and Sons, 2015.
- 12) S. Sareen, S. K. Sood and S. K. Gupta, "A Cloud-Based Seizure Alert System for Epileptic Patients That Uses Higher-Order Statistics," in Computing in Science and Engineering, vol. 18, no. 5, pp. 56-67, Sept.-Oct. 2016
- 13) S. Vaid, P. Singh and C. Kaur, "EEG Signal Analysis for BCI Interface: A Review," 2015 Fifth International Conference on Advanced Computing and Communication Technologies, Haryana, 2015, pp. 143-147. doi: 10.1109/ACCT.2015.72
- 14) Sareen, S., Sood, S.K. Gupta, S.K. An Automatic Prediction of Epileptic Seizures Using Cloud Computing and Wireless Sensor Networks. J Med Syst 40, 226 (2016).

# A Review on The Medical Applications of Electroencephalography (EEG)

Tushar Kanti Bera<sup>1,2</sup>

<sup>1</sup>Department of Electrical Engineering, National Institute of Technology Durgapur (NITDgp)

Mahatma Gandhi Avenue, A-Zone, Durgapur, West Bengal 713209, INDIA

<sup>2</sup>Centre for Biomedical Engineering and Assistive Technology (BEAT), National Institute of Technology Durgapur (NITDgp)

Mahatma Gandhi Avenue, A-Zone, Durgapur, West Bengal 713209, INDIA

Email: [tkbera77@gmail.com](mailto:tkbera77@gmail.com)

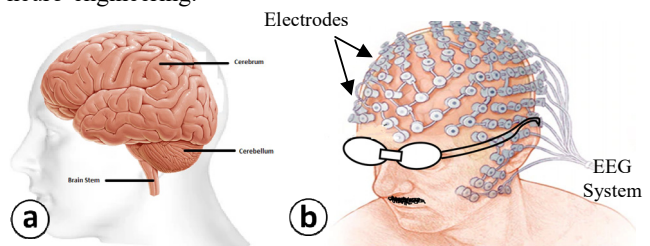
**Abstract**—The human-brain is a complex-organ developed with nerve cells which are known as the neurons. The human body functions are controlled by the electrochemical activity of the nerve cells which in turn generates electric signals. The functions and activities of the human body parts and organs are disturbed by the abnormal brain functions created by unusual brain electrical activity. The brain electricity could be monitored as electrical signals with an electrical signal measurement process called electroencephalography (EEG). The function and the health of the brain can be studied through EEG signals-analysis which are acquired noninvasively or invasively as the electrical signature of the brain physiology. Therefore, several brain-diseases, brain disorders, and brain injuries can be diagnosed by analyzing the information extracted from the EEG signals. The paper reviews the application of the EEG in medical science and summarises the present status and future possibilities of the EEG technology in neuroscience and neuro-engineering. The papers reviewed the EEG application for major brain disorders such as Parkinson's disease, epilepsy, Alzheimer's disease, brain tumours, and stroke.

**Keywords** - Human brain, neurons, synapse, brain electricity, electroencephalography (EEG), EEG measurement, brain disease, brain disease diagnosis with EEG

## I. INTRODUCTION

The brain of the human body [1-2] is a complex-organ (Fig. 1a) developed with about a hundred billion nerve cells, called neurons [1-2]. The brain is a part of the nervous-system [3] which includes the brain, the spinal-cord and a very large network developed with a huge number of nerve cells. The functioning of the human brain depends on the electrochemical activity [4] of the nerve cells which in turn generates electric signals. The depolarization and repolarization in neurons [5-6] which are produced due to the flow of the  $Ca^{++}$ ,  $Na^{+}$  and  $K^{+}$  ions, are the origin of the development of the electric potentials and the signal transmission at the synapse [1-2, 5-6]. Abnormal brain activity, which changes the function and activity of the human body parts and organs, appear as the unusual brain electrical activity [7]. The brain electricity could be monitored noninvasively (or invasively) as electrical signals called electroencephalograms [8-11] which could be obtained from an electrical signal measurement process (Fig. 1b) called electroencephalography (EEG) [11-17]. EEG is performed using a high gain amplifier

circuit [17-21] interfaced with an array of surface electrodes called EEG electrodes [21-24] attached to the human head-scalp with a number of possible electrode configurations. Voltage amplification and acquisition at different position of the brain produces different EEG signals such as gamma-wave ( $\gamma$ ), beta-wave ( $\beta$ ), alpha-wave ( $\alpha$ ), theta-wave ( $\theta$ ) and delta-wave ( $\delta$ ) [21, 25]. Studying these EEG waves the brain functioning and the brain health can be studied as these waves are the electrical signature of the brain physiology. Therefore, EEG could be studied to extract the information for the diagnosis of several brain-diseases [26], brain-disorders [27], and brain-injuries [28]. Utilizing EEG signals, a number of brain diseases and disorders could be detected such as epilepsy [29], Alzheimer's disease [30], Parkinson's disease [31], brain tumors [32], and stroke [33]. Moreover utilizing the information extracted from the EEG signals other physiological states such as stress, depression, emotions, mental fatigue or other mind status can also be studied. In this direction, the paper represents a short review on the applications of the EEG technology in medical, clinical and other applications. The papers reviewed the EEG application for few major brain disorders along with the application of EEG to for the non-medical and non-clinical applications. The paper also tries to summarize the present status and future possibilities of the EEG technology in neuroscience and neuro-engineering.



**Figure-1:** Human brain and EEG acquisition (a) Human brain-anatomy showing three major parts of the brain: the cerebrum, the cerebellum and the brainstem (b) EEG recording with EEG electrodes.

## II. MATERIALS AND METHODS

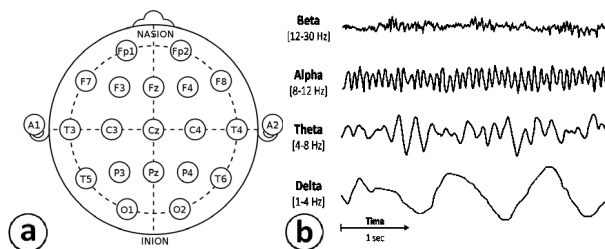
### A. Human Brain Anatomy

The human brain which is enclosed by the skull has three major parts: the cerebrum, the cerebellum and the brainstem

[1-2]. Among these three parts, the cerebrum, which is considered as the brain's largest part, is composed of right-hemisphere and left-hemisphere. The cerebrum is responsible to perform the brain functions like touch interpretation, vision analysis and hearing. It also acts for controlling the speech, learning, reasoning, emotions, as well as the movement control. The cerebellum which is located below the cerebrum controls the muscle movements, body posture, and balance. The third part, named the brainstem acts as connecting part between the cerebrum and cerebellum to the spinal cord and controls heart rate, breathing, body-temperature. It also works to control the digestion process, swallowing, vomiting, sneezing, coughing, and wake and sleep cycles.

### B. Electroencephalography (EEG)

The EEG is a test that measures electrical activity inside the brain. In EEG procedure an electronic instrumentation called EEG-instrumentation [17-21] is used to acquire the brain electrical signal. The EEG-instrumentation is interfaced with the brain through small conducting discs called EEG-electrodes [21-24] which are attached to the scalp (Fig. 2a) which is the outer part of the human head developed with the layers of skin and subcutaneous tissue. The brain cells which are active all the time, even during our asleep, communicate via electrical impulses which collectively appear as the electrical signal called electroencephalogram. Electroencephalograms are found as the time varying signal (Fig. 2b) which can be analysed to extract many important information related to brain anatomy, physiology and the brain disorder and brain diseases.



**Figure-2:** Positions/locations of the EEG-electrodes for the international 10-20 system for electroencephalography-recording, (b) EEG signals with their different bandwidth,

### C. EEG Examination Procedure and Precautions

The EEG procedure may be noninvasive or invasive [34]. In noninvasive EEG the electrodes are put on the scalp-surface whereas the noninvasive EEG procedure puts the electrodes either on the brain-surface or even inside the brain with some needle electrodes. In invasive EEG [35-36] the electrodes are inserted into the brain through skull and either placed directly on the brain surface [37-38] or inserted further inside the brain [38-39]. In the noninvasive EEG, the procedure is conducted in such a way that the patients could not feel any discomfort [34] and return to the normal life very soon after the EEG procedure. The noninvasive EEG procedures are safe and painless. Even the invasive EEG is also conducted with some medical sedation so that the patients can not feel any discomfort. The EEG sensors or the surface

electrodes neither apply anything into the brain nor pass on anything to the patient to feel discomfort or pain [34]. The duty of the electrodes to interface the brain/scalp with the EEG recording instruments and to acquire the brain electrical signals and sent them to the PC.

### D. EEG Electrodes and Electrode Placement

In EEG procedure, a number of EEG electrodes are attached to the head-scalp with a standard electrode placement protocol such as international standard 10–20 system [40-42]. The 10–20 system which is an internationally recognized EEG-acquisition process to place the EEG signal recording sensors to determine the locations of the electrodes on the scalp surface. The 10–20 protocol based electrode placement system for EEG recording has been proposed and developed to maintain a standard testing procedure so that the clinical test results obtained from the EEG procedures could not only be reproduced but also could be effectively analysed and compared with some standard data sets to extract conclusive information and research finding. The 10–20 electrode-array system is basically based on the relationship between the location/position of the EEG-electrodes attached to the head-scalp and the underlying area of cerebral-cortex. In the 10-20 system, the electrodes are placed on the scalp surface in such a way that the actual distances between two adjacent electrodes are either 10% or 20% of the total front–back or right–left distance of the skull and hence "10" and "20" in the "10–20 system" actually refer to inter electrode distance as a fraction of the total distances. Therefore, during an EEG experiment, the technician would place the EEG electrodes after measuring the dimension of the head and making the markings on the scalp surface to fix the location of the electrodes to be placed. As the patient's head generally contain lot of hairs the electrodes are placed individually with the help of a certain type of adhesive through an array of surface electrodes are placed on the scalp as a head net arrangement housed on an elastic cap.

### E. EEG Instrumentation

During the EEG procedure, the patient is generally asked to relax and be in a comfortable position throughout the test. Movement of the head is not expected as the motional artefacts [43-44] may come in the signals. Sometimes, the doctors or the clinicians may ask the patient to open and close the eyes during the test [34] to study the effect of opening and closing of the eyes on the EEG signals. If required, the technician even may ask the patient perform few simple tasks like performing mathematical calculations, reading books, looking at the pictures, taking the breaths or even looking at a flashing light. The electroencephalography is found as the time varying signal which provides much important information related to brain anatomy, physiology and hence the EEG can be utilized to study the brain disorder and brain diseases. EEG signals are low frequency (1 Hz to 60 Hz) [45] low amplitude signal. The amplitude of the EEG signals of a normal subject in the awake condition is about 10–100  $\mu\text{V}$  [46] when measured on the scalp. But if we measure on the surface

on the brain (subdural measurement) it is increased up to the range of mV. The electrodes are used to sense the brain potentials available on the scalp surface and to send it to the instrumentation through the connecting wires which may be coaxial cables or else. The brain signals sensed by the electrodes are passed through the instrument amplifiers, filters, digitizers and then are saved in a personal computer. Therefore, for the EEG instrumentation could be developed with an amplification circuit along with a low pass filter. Basically the EEG instrumentation is developed with amplifiers, filters and data acquisition system [47].

#### F. EEG Acquisition

EEG represents the signature of the brain electrical signal generated by the nerve cells. The Brain electricity is generated to communicate with the nerve cells to communicate with all other organs and body parts by sending the signals through nerves. To correlate the body motion with the EEG recorded, the live video may be taken [34] by the clinicians during the entire EEG procedures. The video is used by the doctors who are trained for EEG interpretation to diagnose the health problem and start to treat the condition. The EEG can be recorded for a short term (few minutes or few hours) or long terms (for several hours or even for several days) which is called the Ambulatory EEG (aEEG) [48-49] monitoring. The aEEG procedure is used for recording the brain electrical activity for several days. The aEEG provides the EEG information or the information about the brain electrical activity which allows the doctors and the clinicians to capture the abnormal brain activities which occurs occasionally [50]. Therefore, the aEEG has the ability to enhance the chances of capturing the brain electrical activity during seizure which is totally unpredicted in most of the cases.

#### G. Applications of EEG

EEG represents the signature of the brain electrical signal generated by the neurons. The Brain electricity is generated to communicate with all the neurons to collectively control all the human body function to keep the body alive and healthy. Therefore, the EEG can not only be utilized for the anatomy and physiology studies on the brain and also to study and diagnose numerous brain-disease and brain-disorder. Through the routine or regular short term EEG is used to study and diagnose several neural diseases but the EEG can also be used for longer time to collect the information which occasionally appears in some special brain disorders. Long term EEG is also found sometimes essential for ICU patients, patients in coma. Even the EEG can be used to confirm brain death in someone in a persistent coma.

#### H. Brain Disorder and Brain Diseases

The brain which is a part of the nervous-system is the control centre of the living body. The nervous system [3] includes the brain, spinal cord and a peripheral network of nerves constructed with neurons. Together with the brain, the nervous-system [3] communicates with the entire body through its communication by sending the electrical signal and

receiving their responses through a particular biological circuit called nervous system generates. If due to some reason, the brain function or structure are disturbed or altered then the brain is called to have a disease called brain disease [26]. The brain disease may be created due to either any infection or any abnormal physiological changes in the body such as formation of tumour or stroke. When the brain fails to work in a normal healthy order, the physiological status is called the brain disorder [27]. Malfunctioning of brain or the brain-disorders may include an abnormal-conditions or disabilities of the brain that affect our body functions. Brain damage followed by a brain disorder or brain disease may also occur due to a trauma which may be either a mental trauma or a physical trauma. When the brain suffers from any disease it fails to function properly and hence it affects the health status of many body functions such as memory, sensation, and reasoning ability, personality, seizures etc. Among a huge number of brain disease or brain disorder, a few brain diseases can be found as Brain Malformations, Cerebellar Disorder, Alzheimer's Disease, Brain Tumors, Dementia, Concussion, Creutzfeldt-Jakob Disease, Degenerative Nerve Diseases,.

### III. RESULTS AND DISCUSSIONS

Among a number of brain diseases many of them could be detected by EEG. The following section has reviewed the application of EEG for detecting few of the brain diseases.

#### A. EEG Applications for Brain Disease Detection

Utilizing EEG signals, a number of brain diseases and disorders could be detected such as epilepsy [29], Alzheimer's disease [30], Parkinson's disease [31], brain-tumors [32], and stroke [33]. Brain injuries may be created either by the accidental impact or even be happened by blunt-trauma. The trauma can damage brain-tissue, nerves and neurons and hence it can affect the brain function. Examples of brain-injuries may include 1) blood-clots, 2) hematomas, 3) cerebral-edema, contusions, or 4) bruising of brain-tissue, or 5) swelling inside the skull, 6) concussions, 7) strokes, or else. When a brain injury happens, a number of symptoms may occur such as vomiting, nausea, bleeding from the ear, speech-difficulty, problems with concentration, memory loss, numbness, paralysis, etc.

Albert *et al.*, 2016 [51] applied the EEG to study the traumatic Brain Injury (TBI). Traumatic Brain Injury (TBI) [51-52] is known as one of the significant causes of death and after the post- accident disabilities [51]. The authors tried to contribute to develop an automatic, fast and reliable technique to process the EEG data to diagnose the TBI. In the proposed study, the authors diagnosed the TBI using a discriminant analysis process based on the quantitative EEG (qEEG) features-extraction from EEG data recorded. The authors utilized an EEG system called EmerEEG [51] which was found as an integrated portable decision support system and the artifacts were removed automatically and then the EEG features are extracted. The proposed algorithm was developed utilizing a model extracted from clinically-labelled EEG signals for TBI diagnosis. The authors reported that the



system was found able to diagnose the TBI using the developed algorithms with 79.1% accuracy in removing artefacts, and 87.85% accuracy in TBI-diagnosis. The authors anticipated that the proposed technique is found as capable of conducting an early diagnosis of TBI and the malfunctioning of the brain. A brain tumour is an abnormal tissue growth inside the brain which may be converted into cancer in the later stage. Before the invention of the modern conventional imaging techniques such as CT scanning, MRI, it was difficult to localize a brain tumours or other lesions [53]. EEG can be considered as a reliable tool in detecting the lesions within the superficial portions of cerebral hemispheres. The localizing the lesions situated inside the deeper regions such as posterior fossa the EEG based detections faces some limitations. Though the brain tumor can be studied with CT scan and MRI but EEG can also provide a lot of informations regarding the brain tumor [32, 54]. There will be many symptoms for the brain-tumors [55]. The seizures may be found as one of the symptoms for brain tumor [56]. The abnormalities found in the EEG signals obtained from a patient with brain-tumors may depend on the nature of the tumor and also at what stage the tumor belongs to. EEG variations which are observed for a patient with brain tumors are generally caused for the disturbances produced due to the tumor and hence the localization of the tumor is difficult is misleading from the EEG signal though advanced EEG techniques are also being studied for brain source localization [57].

#### B. Cognitive Load Prediction.

The cognitive load is the amount of the working-memory which has been used for performing different tasks given. Studying of the tasks associated cognitive-load is very important for several applications. The measurement and assessment of cognitive load are very useful for designing of the instructional materials for students, for monitoring the mental well-being of person with important and complex tasks. EEG can be utilized for cognitive load prediction [58-59]. Friedman *et al.* (2019) [58] tried to use the EEG to understand and analyze the cognitive work-load of subjects undergoing the intelligence-tests. The authors applied the machine learning (ML) technique by training the classic ML models using EEG, signal-complexity [58] and network connectivity. The research findings obtained by the authors demonstrated that using the features extracted from EEG waves, cognitive load can be well understood.

#### C. EEG for mental state and emotion detection

Tandle *et al.* (2018) [60] applied EEG techniques to explain the mental state and emotion detection. They analyzed the EEG signals to try to understand the musical-stimulus-effect on the human-brain. Authors studied the features extracted from the EEG waves and tried to correlate with the human emotions associated with different brain-areas. The authors used the ML-algorithms to conduct their investigations to classify and validate the developed models. They concluded that EEG could be an effective procedure to assess the influence of the musical-stimulus on the human brain.

Depression is a mental health disorder which is a group of conditions associated with the elevation or lowering of a person's mood. The depression is generally characterized by persistently unusual mood or mental state associated with the loss of interest in activities which can cause a significant impairment in daily life. At present, the depression is the one of the key concerns on the human health in the world. Cai *et al.* (2018) [61] studied EEG to diagnose and analyze the depression and conducted a research on a psychophysiological database. The data based of 92 depressed-patients and 121 normal human-subjects were utilized from the database. Using a pervasive three-electrode (PTE) EEG acquisition system [61] the EEG waveforms were collected from of all the 213 subjects under resting-state as well as with the sound-stimulation. The EEG-waveforms were denoised with the Finite Impulse-Response filter (FIRF) by combining with the Kalman derivation formula, Discrete Wavelet Transformation, and an Adaptive Predictor Filter. After that, the linear and non-linear features were extracted from the EEG-waveforms. The information obtained from the depressed participants were distinguished from that obtained under normal controls using the Support Vector Machine (SVM), Classification Trees, K-Nearest Neighbor (KNN), and Artificial Neural Network (ANN) were used to distinguish. Results demonstrated that in the detection of depression, the KNN technique appeared more accurate for the EEG waveforms collected through PTE-EEG acquisition system.

#### D. EEG for Criminal Forensic

During the recent past, the electroencephalography-based brain-computer-interface (BCI) has been found as an emerging research studies in neuro-science, neuro-engineering and rehabilitation. The BCI allow a human brain to be interfaced with the computer through the brain signal (without the normal neurophysiologic pathways). Thus the BCI is found a medium through which the brain can communicate to a computer directly using the brain signals probed by the suitable sensor and instrumentation interfaced with the PC. Mohanchandra *et al.* (2015) [62] studied EEG and found that the brain electrical signals are very effective to establish an interface between the human-brain and a personal computer. The authors studied the possibility of the BCI in forensic investigation. They studied EEG to use it as a lie-detection tool which has been adopted for investigating and solving the criminal cases. Saini *et al.* (2019) [63] studied the information extracted from the EEG data for lie detection using support vector machine (SVM). Compared to the other state-of-the-art feature extraction methods reported earlier, the proposed technique of features extraction with customized SVM showed better accuracy than. The proposed method of hybrid features distinguished the guilty and innocent subjects with the classification accuracy of 99.44%.

#### E. EEG for Fatigue Detection

Driving fatigue is an important issue as it is often found as the cause of traffic accidents [64]. Detecting the fatigue state of the drivers by monitoring their psychological parameters is

an effective and promising technique in preventing the traffic-accident [64]. Jing *et al.* (2020) [64] tried to study the driver's field driving fatigue test by monitoring the driver's real-time EEG signals. The EEG waveforms of the drivers were obtained by basically in three typical states: awake, critical, and fatigue. The EEG signal was analysed using nonlinear and linear methods. The authors used a EEGLAB toolbox in MATLAB with a nonlinear method and the EEG-eigenvalues were collected. The EEG power characteristic values were calculated to assess the trend of EEG signals. The study provided the foundation in analyzing the driver's fatigue states to help in designing of the driver warning-system in low-voltage and hypoxia plateau environment.

#### F. Fetal electroencephalography

Adverse perinatal events which affect the cerebral-functions not only work as the major origin for neonatal mortality and morbidity but also for long-term neurologic problem [65]. The intrapartum fetal-EEG is recorded to study the electrical activities of the fetal-brain to assess the fetal central nervous system (CNS) during labor [65]. Jing *et al.* (2000) [65] described a technique to monitor fetal condition using the fetal EEG and its real-time spectral analysis during labor. Authors studied the fetal-EEG analysis for fourteen pregnant women subjects who were found with uncomplicated term pregnancies. The fetal-EEG signals were studied with two suction-cup electrodes placed on the fetal-scalp. The electrodes were located at the occipitoparietal or parietal region of the brain and the real-time spectral-analyses were conducted for assessing the frequency and amplitude of the fetal-EEG waves. During the fetal-EEG acquisition studies two fundamental EEG patterns were identified: 1) high-voltage slow-activity and 2) low-voltage fast-activity. The authors mentioned that the spectral edge frequency (SEF) could be considered to be an excellent index of cyclic-EEG activities. Authors concluded that the proposed study may be useful to enable us for getting a rapid alertness to the cerebral hypoxia for prompt intervention to decrease the risk for birth asphyxia and subsequent brain damage.

The EEG has been found applicable also for a number of brain disorders as discussed above as well as for other medical studies such as Schizophrenia studies [66], Covid-19 studies [67] cognitive workload studies [68] and also could be studied in many other applications in future.

#### IV. CONCLUSIONS

Utilizing EEG signals, a number of brain diseases and disorders could be detected such as Alzheimer's disease, epilepsy, Parkinson's disease, brain cancers brain tumors, and stroke. Also, EEG can be efficiently utilized for cognitive load assessment, brain computer interface for physically disabled person, lie detection, mental health studies, and fatigue analysis and so on. Though the functional-neuroimaging techniques, like positron emission tomography (PET), single-photon emission computed tomography (SPECT), and functional MRI (fMRI), can provide tomographic representation of the brain along with its physiologic

information but fail to provide the high temporal resolution like EEG. Furthermore, EEG can provide the continuous or ambulatory monitoring facility of the brain which is extremely useful for studying a number of brain disease and disorders.

#### V. REFERENCES

- [1] Haines, D. E. (2004). Neuroanatomy: an atlas of structures, sections, and systems (Vol. 153, No. 2004). Lippincott Williams & Wilkins.
- [2] Marieb, E. N., & Hoehn, K. (2007). Human anatomy & physiology. Pearson education.
- [3] Mai, J. K., & Paxinos, G. (Eds.). (2011). The human nervous system. Academic press.
- [4] Frackowiak, R. S. (2004). Human brain function. Elsevier.
- [5] Chen, I., & Lui, F. (2019). Neuroanatomy, Neuron Action Potential.
- [6] Lodish, H., Berk, A., Zipursky, S. L., Matsudaira, P., Baltimore, D., & Darnell, J. (2000). The action potential and conduction of electric impulses. In Molecular Cell Biology. 4th edition. WH Freeman.
- [7] Duffy, F. H. (Ed.). (2013). Topographic mapping of brain electrical activity. Butterworth-Heinemann.
- [8] Ngo, D., Ombao, H., Sun, Y., Genton, M. G., Wu, J., Srinivasan, R., Cramer, S. An exploratory data analysis of electroencephalograms using the functional boxplots approach. *Frontiers in neurosc.*, 9, 282, 2015.
- [9] Henry, C. E., & Greulich, W. W. (1944). Electroencephalograms of normal children. *Monographs of the Society for Research in Child Development*, 9(3), i-71.
- [10] Mattson, R. H., Pratt, K. L., & Calverley, J. R. (1965). Electroencephalograms of epileptics following sleep deprivation. *Archives of Neurology*, 13(3), 310-315.
- [11] Malmivuo, J., & Plonsey, R. (1995). Bioelectromagnetism: principles and applications of bioelectric and biomagnetic fields. Oxford University Press, USA.
- [12] Light, G. A., Williams, L. E., Minow, F., Sprock, J., Rissling, A., Sharp, R., ... & Braff, D. L. (2010). Electroencephalography (EEG) and event - related potentials (ERPs) with human participants. *Current protocols in neuroscience*, 52(1), 6-25..
- [13] Britton, J. W., Frey, L. C., Hopp, J. L., Korb, P., Koubeissi, M. Z., Lievens, W. E., ... & Electroencephalography, E. L. S. (2016). An introductory text and atlas of normal and abnormal findings in adults, children, and infants. American Epilepsy Society, Chicago.
- [14] Bronzino, J. D. (1995). Principles of electroencephalography. The biomedical engineering handbook, 1.
- [15] Read, G. L., & Innis, I. J. (2017). Electroencephalography (Eeg). The international encyclopedia of communication research methods, 1-18.
- [16] Biasucci, A., Franceschiello, B., & Murray, M. M. (2019). Electroencephalography. *Current Biology*, 29(3), R80-R85.
- [17] Bera, T. K., Choudhary, S., Maiti, T., & Barnwal, S. P. (2020, March). A Low-Cost Electroencephalography (EEG) Instrumentation for Epileptic Seizure Detection. In *Journal of Physics: Conference Series* (Vol. 1495, No. 1, p. 012038). IOP Publishing.
- [18] Dias, N. S., Carmo, J. P., Mendes, P. M., & Correia, J. H. (2012). Wireless instrumentation system based on dry electrodes for acquiring EEG signals. *Medical engineering & physics*, 34(7), 972-981.
- [19] Allen, P. J. (2009). EEG instrumentation and safety. In *EEG-fMRI* (pp. 115-133). Springer, Berlin, Heidelberg.
- [20] Teplan, M. (2002). Fundamentals of EEG measurement. *Measurement science review*, 2(2), 1-11.
- [21] Bera, T. K. (2015). Noninvasive electromagnetic methods for brain monitoring: a technical review. In *Brain-Computer Interfaces* (pp. 51-95). Springer, Cham.
- [22] Rush, S., & Driscoll, D. A. (1969). EEG electrode sensitivity-an application of reciprocity. *Ieee T Bio-Med Eng.*, (1), 15-22.
- [23] Taheri, B. A., Knight, R. T., & Smith, R. L. (1994). A dry electrode for EEG recording. *Electroencephalogr Clin Neurophysiol*, 90(5), 376-383.
- [24] Oostenveld, R., & Praamstra, P. (2001). The five percent electrode system for high-resolution EEG and ERP measurements. *Clinical neurophysiology*, 112(4), 713-719.
- [25] Mai, G., Minnett, J. W., & Wang, W. S. Y. (2016). Delta, theta, beta, and gamma brain oscillations index levels of auditory sentence processing. *Neuroimage*, 133, 516-528.

- [26] Zigmond, M. J., Coyle, J. T., & Rowland, L. P. (Eds.). (2014). *Neurobiology of brain disorders: biological basis of neurological and psychiatric disorders*. Elsevier.
- [27] Brain, R. (1933). *Diseases of the nervous system*. Oxford University Press.
- [28] Carney, N., Totten, A. M., O'Reilly, C., Ullman, J. S., Hawryluk, G. W., Bell, M. J., ... & Rubiano, A. M. (2017). Guidelines for the management of severe traumatic brain injury. *Neurosurgery*, 80(1), 6-15.
- [29] Li, Y., Liu, Y., Cui, W. G., Guo, Y. Z., Huang, H., & Hu, Z. Y. (2020). Epileptic seizure detection in EEG signals using a unified temporal-spectral squeeze-and-excitation network. *IEEE Trans Neural Syst Rehabil Eng.*, 28(4), 782-794.
- [30] Pineda, A. M., Ramos, F. M., Betting, L. E., & Campanharo, A. S. (2020). Quantile graphs for EEG-based diagnosis of Alzheimer's disease. *Plos one*, 15(6), e0231169.
- [31] Anjum, M. F., Dasgupta, S., Mudumbai, R., Singh, A., Cavanagh, J. F., & Narayanan, N. S. (2020). Linear predictive coding distinguishes spectral EEG features of Parkinson's disease. *Parkinsonism & Related Disorders*, 79, 79-85.
- [32] Liu, Y., Huang, Y. X., Zhang, X., Qi, W., Guo, J., Hu, Y., ... & Su, H. (2020). Deep C-LSTM Neural Network for Epileptic Seizure and Tumor Detection Using High-Dimension EEG Signals. *IEEE Access*, 8, 37495-37504.
- [33] Chang, R., Reddy, R. P., Sudadi, S., Balzer, J., Crammond, D. J., Anetakis, K., & Thirumala, P. D. (2020). Diagnostic accuracy of various EEG changes during carotid endarterectomy to detect 30-day perioperative stroke: A systematic review. *Clinical Neurophysiology*, 131(7), 1508-1516.
- [34] Internet Article, EEG (electroencephalogram), Link <https://www.mayoclinic.org/tests-procedures/eeeg/about/pac-20393875>, Access Date: Oct. 07, 2020.
- [35] Blume, W. T., Whiting, S. E., & Girvin, J. P. (1991). Epilepsy surgery in the posterior cortex. *Annals of Neurology: Official Journal of the American Neurological Association and the Child Neurology Society*, 29(6), 638-645.
- [36] Taussig, D., Montavont, A., & Isnard, J. (2015). Invasive EEG explorations. *Neurophysiologie Clinique/Clinical Neurophysiology*, 45(1), 113-119.
- [37] Dümpelmann, M., Fell, J., Wellmer, J., Urbach, H., & Elger, C. E. (2009). 3D source localization derived from subdural strip and grid electrodes: a simulation study. *Clinical Neurophysiology*, 120(6), 1061-1069.
- [38] Im, C., & Seo, J. M. (2016). A review of electrodes for the electrical brain signal recording. *Biomedical Engineering Letters*, 6(3), 104-112.
- [39] Chu, N. S., Wu, C. L., Tseng, T. S., & Kuo, L. L. (1991). Sphenoidal EEG recording using acupuncture needle electrode in complex partial seizure. *Electroencephalogr Clin Neurophysiol*, 79(2), 119-126.
- [40] Herwig, U., Satrapi, P., & Schönfeldt-Lecuona, C. (2003). Using the international 10-20 EEG system for positioning of transcranial magnetic stimulation. *Brain topography*, 16(2), 95-99.
- [41] Homan, R. W., Herman, J., & Purdy, P. (1987). Cerebral location of international 10-20 system electrode placement. *Electroencephalogr Clin Neurophysiol*, 66(4), 376-382.
- [42] Homan, R. W. (1988). The 10-20 electrode system and cerebral location. *American Journal of EEG Technology*, 28(4), 269-279.
- [43] Butkevičiūtė, E., Bikulčienė, L., Sidekierskienė, T., Blažauskas, T., Maskeliūnas, R., Damaševičius, R., & Wei, W. (2019). Removal of movement artefact for mobile EEG analysis in sports exercises. *IEEE Access*, 7, 7206-7217.
- [44] Tandle, A., Jog, N., D'cunha, P., & Chheta, M. (2016). Classification of artefacts in EEG signal recordings and EOG artefact removal using EOG subtraction. *Commun Appl Electron*, 4, 12-9.
- [45] Amo, C., De Santiago, L., Barea, R., López-Dorado, A., & Boquete, L. (2017). Analysis of gamma-band activity from human EEG using empirical mode decomposition. *Sensors*, 17(5), 989.
- [46] Blinowska, K., & Durka, P. (2006). *Electroencephalography (ceg)*. Wiley encyclopedia of biomedical engineering.
- [47] AA1: Velarde-Reyes, E., Márquez-Bocalandro, Y., & Martínez-Montes, E. (2011). EEG Wireless Recording and Multidimensional Analysis. *Google Scholar*, 1-16.
- [48] Olson, D. M. (2001). Success of ambulatory EEG in children. *Journal of Clinical Neurophysiology*, 18(2), 158-161.
- [49] Kandler, R., Ponnusamy, A., & Wragg, C. (2017). Video ambulatory EEG: a good alternative to inpatient video telemetry?. *Seizure*, 47, 66-70.
- [50] van der Ree, M., & Wijnberg, I. (2012). A review on epilepsy in the horse and the potential of Ambulatory EEG as a diagnostic tool. *Veterinary Quarterly*, 32(3-4), 159-167.
- [51] Albert, B., Zhang, J., Noyvirt, A., Setchi, R., Sjaheim, H., Velikova, S., & Strisland, F. (2016, July). Automatic EEG processing for the early diagnosis of traumatic brain injury. In 2016 World Automation Congress (WAC) (pp. 1-6). IEEE.
- [52] Rapp, P. E., Keyser, D. O., Albano, A., Hernandez, R., Gibson, D. B., Zambon, R. A., ... & Nichols, A. S. (2015). Traumatic brain injury detection using electrophysiological methods. *Frontiers in human neuroscience*, 9, 11.
- [53] Internet article, EEG in Brain Tumours, Medscape, LLC, 395 Hudson Street, 3rd Floor, New York, NY 10014 Webpage: <https://emedicine.medscape.com/article/1137982-overview>, Date of Access: Oct. 6, 2020.
- [54] Liu, Y., Huang, Y. X., Zhang, X., Qi, W., Guo, J., Hu, Y., ... & Su, H. (2020). Deep C-LSTM Neural Network for Epileptic Seizure and Tumor Detection Using High-Dimension EEG Signals. *IEEE Access*, 8, 37495-37504.
- [55] Lovely, M. P. (2004, November). Symptom management of brain tumor patients. In *Seminars in oncology nursing* (Vol. 20, No. 4, pp. 273-283)..
- [56] Vecht, C. J., & Van Breemen, M. (2006). Optimizing therapy of seizures in patients with brain tumors. *Neurology*, 67(12 Sup 4), S10-13.
- [57] Jatoi, M. A., Kamel, N., Malik, A. S., & Faye, I. (2014). EEG based brain source localization comparison of sLORETA and eLORETA. *Australasian physical & engineering sciences in medicine*, 37(4), 713-21.
- [58] Friedman, N., Fekete, T., Gal, Y. A. K., & Shriki, O. (2019). EEG-based Prediction of Cognitive Load in Intelligence Tests. *Frontiers in human neuroscience*, 13, 191.
- [59] Saitis, C., Parvez, M. Z., & Kalimeri, K. (2018). Cognitive load assessment from EEG and peripheral biosignals for the design of visually impaired mobility aids. *Wireless Communications and Mobile Computing*, 2018.
- [60] Tandle, A. L., Joshi, M. S., Dharmadhikari, A. S., & Jaiswal, S. V. (2018). Mental state and emotion detection from musically stimulated EEG. *Brain informatics*, 5(2), 14.
- [61] Cai, H., Han, J., Chen, Y., Sha, X., Wang, Z., Hu, B., ... & Gutknecht, J. (2018). A pervasive approach to EEG-based depression detection. *Complexity*, 2018.
- [62] Mohanchandra, K. (2015, January). Criminal forensic: An application to EEG. In 2015 Recent and Emerging trends in Computer and Computational Sciences (RETCOMP) (pp. 18-21). IEEE.
- [63] Saini, N., Bhardwaj, S., & Agarwal, R. (2019). Classification of EEG signals using hybrid combination of features for lie detection. *Neural Computing and Applications*, 1-11.
- [64] Jing, D., Liu, D., Zhang, S., & Guo, Z. (2020). Fatigue Driving Detection Method Based on EEG Analysis in Low-voltage and Hypoxia Plateau Environment. *Int. J. of Transportation Science and Technology*.
- [65] Thaler, I., Boldes, R., & Timor-Tritsch, I. (2000). Real-time spectral analysis of the fetal EEG: a new approach to monitoring sleep states and fetal condition during labor. *Pediatric research*, 48(3), 340-345.
- [66] Namazi, H. (2020). Information-based classification of electroencephalography (EEG) signals for healthy adolescents and adolescents with symptoms of Schizophrenia. *Fluctuation and Noise Letters*, 2050033.
- [67] Assenza, G., Lanzone, J., Ricci, L., Boscarino, M., Tombini, M., Galimberti, C. A., ... & Mecarelli, O. (2020). Electroencephalography at the time of Covid-19 pandemic in Italy. *Neurological Sciences*, 41(8), 1999-2004.
- [68] Iqbal, M. U., Srinivasan, B., & Srinivasan, R. (2020). Dynamic assessment of control room operator's cognitive workload using Electroencephalography (EEG). *Computers & Chemical Engineering*, 141, 106726.

# TSSummarize: A Visual Strategy to Summarize Biosignals

João Rodrigues\*, Phillip Probst\*, and Hugo Gamboa\*

\*LIBPhys (Laboratory for Instrumentation, Biomedical Engineering and Radiation Physics),  
Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, Caparica, Portugal

**Abstract**—Visual tools enhance the human ability to detect structures found on time series. Medical doctors and data-scientists rely on their visual abilities to perform time series analysis. A visual tool that would summarize several sources of information of time series would be of great value and is not yet provided in the literature. This work proposes a novel unsupervised visual strategy to summarize a time series and compact several layers of information. The strategy extracts information from the Self-Similarity Matrix (SSM). This data source is able to segment the time series, detect events and show relationships between subsequences. The visual strategy has been tested on several use-cases from the medical domain, proving to be type agnostic, intuitive and compact.

**Index Terms**—time series, summarize, visualization, biosignals, segmentation, events, unsupervised

## I. INTRODUCTION

Time series are complex sequences of samples that are characterized by having several structures worth identifying, such as *patterns*, *events*, *subsequences*, among others. The interpretation of these structures on time series can be further enhanced by leveraging visualisation tools and profiting of the human eye. When analysing a time series visually, one is trying to identify relevant structures that give insights on these series and allow for their characterization. With an increasing complexity of the time series, this task turns to be harder. Having tools that could facilitate and enhance this ability of searching for relevant structures on time series are of utmost importance. Current visual tools are able to provide limited information and are not scalable to the multitude of structures time series have. [1].

In this work, we study how to design a strategy that could summarize the information available on a time series and expose it in a meaningful way. This process is inspired by methodologies that exist in other scenarios for statistical analysis, such as the available methods from the pandas python library: *pandas.profile()* and *pandas.describe()*. These methods are able to provide a summarization of the dataset that is given as input. A method that could provide this type of summary for time series, maintaining the graphical information that time series have would be interesting to have.

This work gives a new insight over the concept of time series summarization. A reasoning over which strategies can be utilized to develop this methodology, as well as which information is important to extract and used in order to summarize a time series. Additionally, it proposes a graphical

representation of this information and delivers it in an intuitive and meaningful way.

The manuscript is divided in the following sections: (1) The definition of the problem will be discussed, considering what information is reliable and which methods can be used to retrieve this information, (2) visual concepts about the way data can be displayed, (3) related work, indicating which research groups have already contributed on this topic, (4) proposal of a solution, (5) demonstration of empirical examples with time series from the biomedical domain and (6) overall conclusions and future work.

### A. Problem definition

Time series are sequences of real-valued and timely ordered numbers that carry information. From this simplistic definition, the information that time series carry can vary, as well as depending on the user, the needed source of information can change as well. Due to the high variability in the contexts from which time series can stem from, a generalized structuring of these is not an obvious endeavour. The major problem lies in understanding the significant structures in the series and based on these make a decision on what is critical information to be visualized. The process of identifying significant structures within time series is referred to as data mining.

Data mining tasks on time series involve several topics, such as time series segmentation, event detection, time series representation, time series classification, similarity search, among others [2], [3]. From these strategies, it is possible to get hints regarding what kind of information is typically of interest and *significant* for users:

- Segmentation - Identification and division of differentiated segments on time series. These segments are usually well defined having specific properties that differ in some way from preceding and subsequent sections.
- Events - Relevant instances on time series. These can represent anomalies, peaks, changes on the signal's properties, which can represent transition areas or indicate the beginning of cyclic segments of the time series.
- Similarity - Relationship between samples or subsequences of the time series. Identifying which samples or segments are most similar, or where these sections are repeated over the time series.
- Characterization - Features of the signal are an important source of information because of their descriptive value, as well as statistical information about the time series,

such as the histogram, which reveals the print of the time series.

### B. Proposed Strategy

The aforementioned elements are an initial source of inspiration to understand which information is of utmost interest in most problems. A strategy that could include these elements into a compact yet meaningful and intuitive form of visualization would be of great interest in the data-science community. An ideal strategy, comprising this source of information, should have the following properties:

- Domain Agnostic - Should be adequate for any domain of time series and be of general purpose.
- Dimension Safe - Be able to provide information on the time series regardless of its dimensionality and should be expressed in the time scale that is adequate for the user's need.
- Intuitive - The visualization should provide intuitive information even for complex time series.
- Compact - The design of the graph should be compact and concise, so that the user can have all the information on one single area, without having to scroll for multiple pages of information.
- Qualitative/Quantitative - The information provided, even if mostly qualitative, should be demonstrated in a quantifiable manner, so that the user can have a notion of dimensions, amplitudes and other quantities of the time series.
- Unsupervised - Be able to extract the necessary information without any previous knowledge of the data structure of the time series.

Such a visualisation scheme would be advantageous in many ways, such as: (1) identification of representative subsequences of time series, (2) quick overview of the structure of a time series, (3) quick identification of areas and events of interest, (4) use as image descriptor.

## II. RELATED WORK

### A. Time Series Summary

Very few strategies are found to make compact and meaningful representation of time series. The works that can be highlighted refer to time *snippets* and time series *bitmaps* [1], [4]. The first highlights the limitation of current methods in providing a satisfactory solution to time series *summarization*. It proposes a method that is able to segment the  $k$  most *representative* sub sequences of a time series, and use these elements as the summary. This strategy answers several of the discussed demands aforementioned in Section I-A, namely the segmentation and similarity. Regarding the time series *bitmap* representation, the strategy is able to provide a coded bitmap with information on cluster, anomaly and other regularities on data collection. These bitmaps were used as folder icons, and also answer several of the aforementioned characteristics, such as *similarity* and *events*. An example of both strategies can be seen on Figure 1.

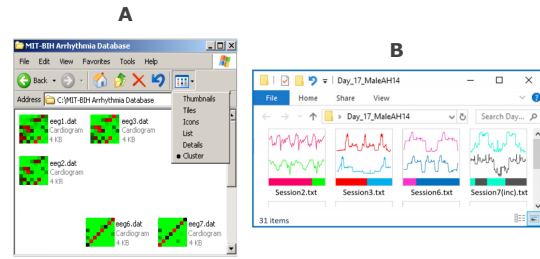


Fig. 1: Strategies for time series summary found on the literature. These images are taken from the works from [1], [4]

Time series *shapelets* are also a method that could provide interesting results. However, the strategy is *supervised*, and the point of the proposed method is to have *no apriori* knowledge about the structure of the data, except the time scale in which the summarization is performed.

Other interesting strategies provide a transformation of time series into text and could be used for time series summarization, but are not able to suitably summarize a time series from the textual representation [5], [6].

### B. Visual Inspiration

Strategies that are typically used to present information in a compact way are found in several domains. In text analysis, for instance, the relationship between repeating sequences is illustrated with arc diagrams [4], [7]. These show where repeating sequences occur in a very concise way. This has a range of applications that include, for example text and DNA sequence analysis.

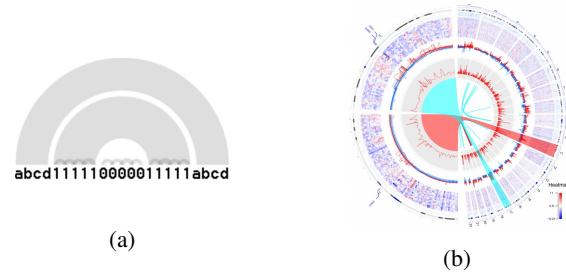


Fig. 2: A - Diagram for string association. This image is taken from the works from [7]; B - Circular plot by OmicCircos. Several layers (Circular tracks) identify genome position, expression heatmaps, correlation between expression and CNV, among other features. The image is taken from the works from Ying Hu, et al. [8].

One domain that has a particular relevance in data visualization is genomics. Graphical genome maps are found to concatenate a significant amount of information in a very compact way. Genome features and sequence characteristics are assessed with this visual strategy. An example can be found on Figure 2b. This visualization strategy can provide increasing circular layers of information. Although we are used to look at time series from left to right, a circular

representation can have benefits to concatenate the information we want to include.

In the musical domain, strategies have also been developed that summarize audio time series with segmentation techniques. One of the strategies that is common to be used involves detecting novelty instances on a similarity matrix representation of the audio signal, called *Self-Similarity Matrix* (SSM). This data structure provides a significant range of information that can be used to retrieve structural information, such as block and periodic structures [9], [10], [11], [12]. This method will inspire our visualization strategy, which will be explained further.

### III. METHOD

As mentioned in Section I-A, a summary of the time series should include several of the structural information available. This section will present the process to reach this information and how to translate it into a summarised version.

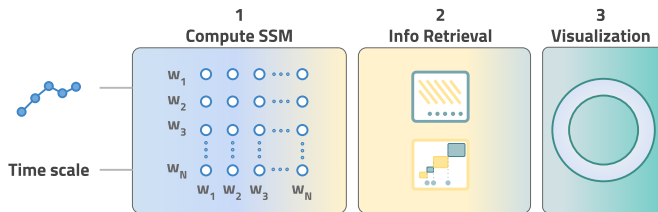


Fig. 3: The stages of the summarization process, from time series to visual summary.

#### A. Information Retrieval

In this work, the process of information retrieval requires the usage of a feature-based self-similarity matrix of the time series, the *SSM*.

The process to compute the *SSM* is shown in Figure 4. First, the method requires the input of the time series to be summarized, then features are extracted to unveil several properties that describe the morphological behaviour over time. The features were extracted with the *TSFEL* library, by means of a sliding window. All features were extracted, except wavelet related [13]. Building the feature matrix ( $F$ ), we are able to compute the *SSM* with the following equation:

$$SSM = F^T \hat{F} \quad (1)$$

The product between  $F$  and itself provides a rich visual information, such as blocks and diagonals, based on similarity scores from each time window. Higher similarity scores mean that the corresponding time windows are similar [11], [12]. Blocks indicate areas of homogeneous morphology on the time series. Changes between block structures along the diagonal are able to segment the time series in different homogeneous areas. The off-diagonals provide periodic information. Whenever off-diagonals are visible on the *SSM*, it reflects the presence of cycles or pattern repetition [9], [10]. From these structures, we are able to retrieve information to build a summary of the time series.

Several methods have to be used to extract the needed information: (1) Block transition information is extracted using the strategy proposed by Foote *et al.*, which is to convolve a square shaped kernel with checkerboard pattern over the *SSM* diagonal. The result is the novelty function (*novaf*) [14]. (2) Sliding a segmented subsequence over the *SSM* provides a distance function between all subsequence segmented. (3) Periodic and relevant events are extracted by searching the minimum peaks of the similarity function (*sf*), which is the column-wise sum of the *SSM*.

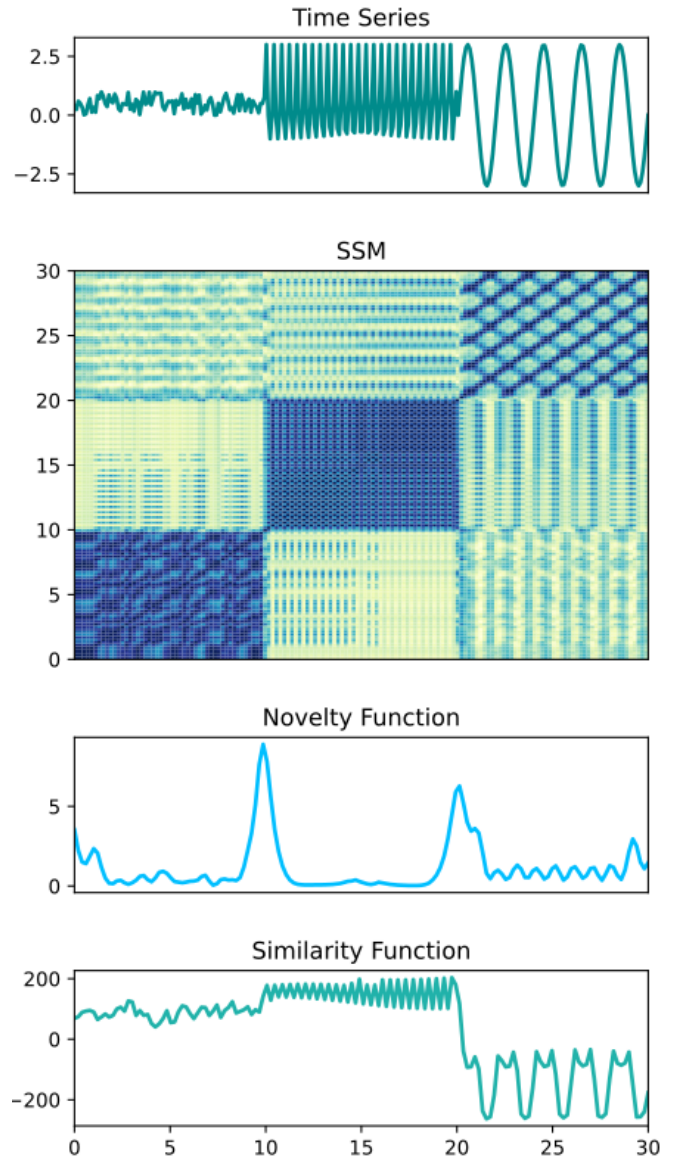


Fig. 5: SSM representation of the time series depicted on the first graph. The segments differ substantially, which is visible by the two main peaks of the *novaf*. Off-diagonals are represented on the third segment, and periods are visible on the *sf*.

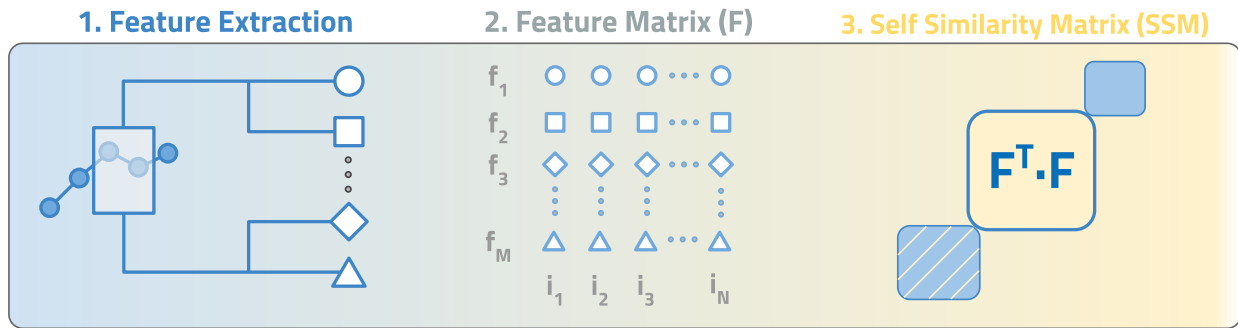


Fig. 4: Method to obtain the SSM. The stages are (1) Feature extraction based on a sliding window with a size defined by the time scale; (2) Building the feature matrix, with each feature occupying a row and (3) compute the SSM based on the dot product between the transpose  $F$  and itself.

**B. Visual Strategy**

The proposed solution for the visual summary of the time series is to create a circular plot that includes the information regarding: (1) Segmentation of the time series into differentiated subsequences, (2) represent the distance between the segmented subsequences (3) highlight relevant events inside the time series and (4) show the distribution of amplitude values of each segment on the center of the image.

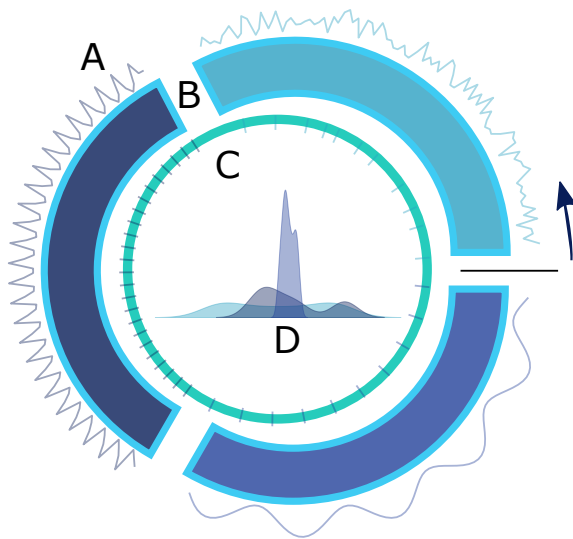


Fig. 6: Visual summary of a time series. Layer A and B highlight the 3 segments automatically detected (novaf). Layer C shows the cyclic information and graph (sf) D provides the histogram of each segment.

An example with a synthetic time series is presented on Figure 5. The time series has 3 major structures, the first is noise, the second and third are periodic signals with different frequencies and shape.

From the computed functions (novaf and sf), we are able to create the summary of the time series depicted in Figure 6. This Figure shows different layers of information, being: (A)

the **signal layer**, where the segmented time series is computed, (B) **similarity layer**, where the colored sectioning of the time series is computed, being the color related with the similarity with other segments, (C) **the event layer**, where significant events are annotated and (D) the density value distribution of each time series segment. The bar and arrow indicate the start and direction of the plot, respectively.

**IV. DEMONSTRATION**

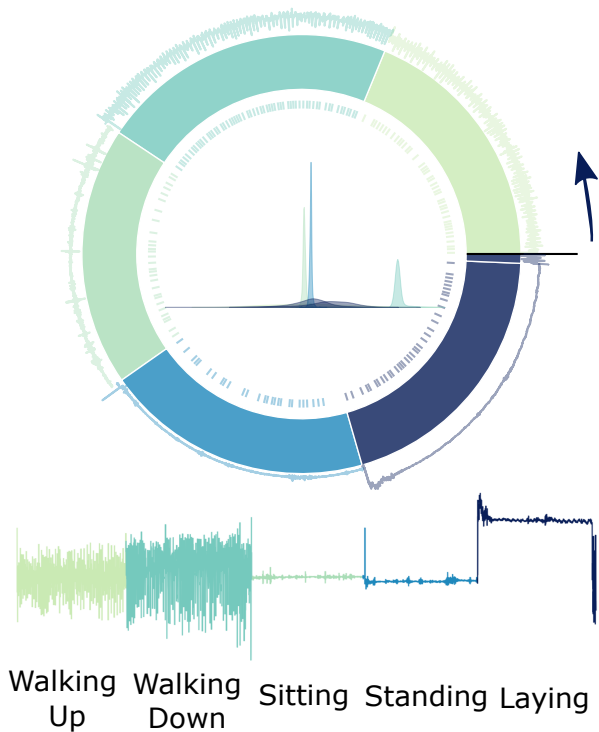


Fig. 7: Visual summary of use-case 1.

The visual strategy has been employed on several use-cases to present its ability to provide a summarized version of a biomedical time series, considering the factors mentioned on Section I-B. The use-cases are: (1) accelerometer data used for human activity recognition tasks [15], (2) blood

volume pressure (BVP) data, acquired to evaluate the changes in the BVP signal when posture changes occur [16] and electromyographic data (EMG), acquired in the context of hand posture classification [17].

#### A. Use-case 1: Human Activity Time Series

The signal was acquired at 50 Hz, and has a duration of 240 seconds. The data is divided in segments of activities that the subject was performing. The SSM was computed with a time scale of 5 seconds, being able to segment the signal and show on the summary the highlighted segments. The color of the segments highlight the similarity between walking activities, while the difference with *Laying*. Events signaled show occurrences on the signal but not related with important information. The steps of walking activities are not able to be detected, due to the time scale used.

#### B. Use-case 2: BVP Dataset

The BVP signal is periodic. It was acquired at a sampling rate of 140 Hz. The signal has several anomalies, which are present due to bad sensor connection. Again, the summary shows the segments and their relationship, and is able to provide the cyclic information from the BVP signal. The anomalies are detected and highlighted with a much different color than the clean areas of the BVP signal.

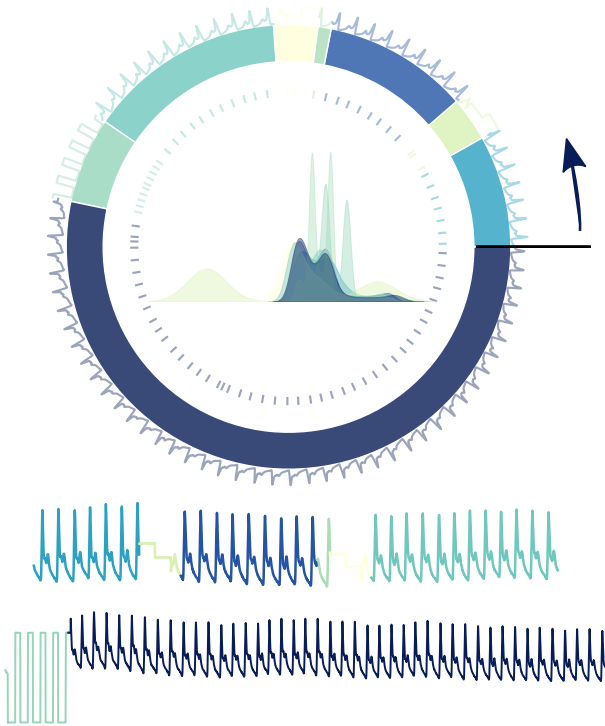


Fig. 8: Visual summary of use-case 2.

#### C. Use-case 3: EMG Dataset

The EMG signal was acquired with a sampling frequency of 200 Hz. The main exercise was to segment the active segments of the EMG from the inactive ones. The method was able to detect all activations, but missed to detect the inactivation

of the third muscular contraction. The colors also show the similarity between the segments.

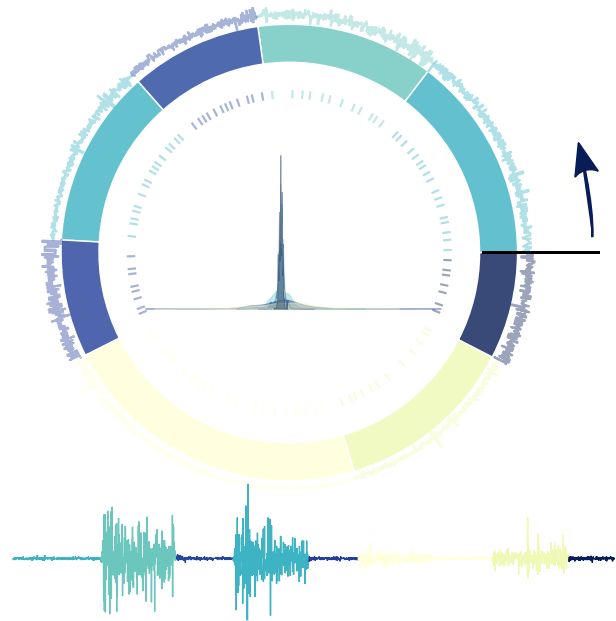


Fig. 9: Visual summary of use-case 3.

#### D. Overview

This first approach to summarize a time series in a multi-layered set of information shows great potential. The segmentation into categories, linked by color is an important feature that highlights differences along the signal. The presence of relevant events, such as periodic events is possible and clear to understand. These can be important to understand frequency rate differences along the signal. The centered histogram shows the *fingerprint* of each segment, and gives an additional distinction level between them.

Regarding the main expected features of this visual tool, it was able to demonstrate its agnosticism to types of time series, as well as its capabilities in making a time series more intuitive, while in a compact format. The process is completely unsupervised, relying only on a time scale value, that guides the search process.

Several improvements have to be made, namely in what regards the visualization of relevant events and similarity between the segmented subsequences. Events are often showed without having a clear relevance, although providing the cyclic information when occurring.

The distance measure used to show similarity between segments should provide more reliable results, e.g. in Figure 8 shows the square wave to be similar with the BVP signal that precedes it, which is not valid.

The method still has to further explore its ability to represent longer time series, as well as how to provide quantitative information more clearly.



## V. CONCLUSION AND FUTURE WORK

In this work, we have presented a novel strategy to summarize multiple sources of information carried by a time series into a compact yet intuitive visual graph. This strategy differs from the ones stated on literature because of the quantity of information it provides and the design inspired on *genomic maps*. Although improvements have to be made, this first instance shows promising results, even in using the SSM as the sole source of information.

In a future work, several improvements can be made, namely regarding the distance strategy used to measure the similarity between segmented subsequences and events detected that are other than the cyclic type. Further investigation should be made regarding the ability of this strategy to represent time series with higher dimensionality. Quantitative information is still limited and should be accessible, even in non-interactive format.

The circular strategy provides the possibility to expand the graph into more layers of information. This would particularly be useful for a multi-scale approach, in which multiple layers of segmentation are performed. Future strategies should also comply with the possibility of have multivariate records. Other sources of information should also be studied to be added as a new layer on the circular plot.

## ACKNOWLEDGMENT

This work was partly supported by Fundação para a Ciência e Tecnologia, under projects OPERATOR (ref. 04/SI/2019) and PREVOCUPAI (DSAIPA/AI/0105/2019), and Ph.D. grant PD/BDE/142816/2018.

## REFERENCES

- [1] S. Imani, F. Madrid, W. Ding, E. S. Crouter, and E. Keogh, "Introducing time series snippets: a new primitive for summarizing long time series," *Data Min. Knowl. Discov.*, pp. 1713–1743, 2020.
- [2] D. Gunopulos and G. Das, "Time series similarity measures and time series indexing (abstract only)," in *Proceedings of the 2001 ACM SIGMOD International Conference on Management of Data*, ser. SIGMOD '01. New York, NY, USA: Association for Computing Machinery, 2001, p. 624. [Online]. Available: <https://doi.org/10.1145/375663.375808>
- [3] T. chung Fu, "A review on time series data mining," *Engineering Applications of Artificial Intelligence*, vol. 24, no. 1, pp. 164 – 181, 2011. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0952197610001727>
- [4] N. Kumar, V. Lolla, E. Keogh, S. Lonardi, and C. Ratanamahatana, "Time-series bitmaps: a practical visualization tool for working with large time series databases," 04 2005.
- [5] J. Rodrigues, D. Folgado, D. Belo, and H. Gamboa, "Ssts: A syntactic tool for pattern search on time series," *Information Processing Management*, vol. 56, no. 1, pp. 61 – 76, 2019. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0306457318302577>
- [6] E. Keogh, J. Lin, and A. Fu, "Hot sax: efficiently finding the most unusual time series subsequence," in *Fifth IEEE International Conference on Data Mining (ICDM'05)*, 2005, pp. 8 pp.–.
- [7] M. Wattenberg, "Arc diagrams: visualizing structure in strings," in *IEEE Symposium on Information Visualization, 2002. INFOVIS 2002.*, 2002, pp. 110–116.
- [8] Y. Hu, C. Yan, C.-H. Hsu, Q.-R. Chen, K. Niu, G. A. Komatsoulis, and D. Meerzaman, "Omiccircos: A simple-to-use r package for the circular visualization of multidimensional omics data," *Cancer Informatics*, vol. 13, p. CIN.S13495, 2014, pMID: 24526832. [Online]. Available: <https://doi.org/10.4137/CIN.S13495>
- [9] M. Müller, *Fundamentals of Music Processing*. Springer Verlag, 2015.

- [10] M. Müller and F. Zalkow, "FMP notebooks: Educational material for teaching and learning fundamentals of music processing," in *Proceedings of the International Conference on Music Information Retrieval (ISMIR)*, Delft, The Netherlands, November 2019.
- [11] J. Paulus, M. Müller, and A. Klapuri, "Audio-based music structure analysis," in *Proceedings of the International Conference on Music Information Retrieval (ISMIR)*, Utrecht, The Netherlands, 2010, pp. 625–636.
- [12] J. P. Bello, P. Grosche, M. Müller, and R. J. Weiss, "Content-based methods for knowledge discovery in music," in *Springer Handbook on Systematic Musicology*, R. Bader, Ed. Springer, Berlin, Heidelberg, 2018, pp. 823–840.
- [13] M. Barandas, D. Folgado, L. Fernandes, S. Santos, M. Abreu, P. Bota, H. Liu, T. Schultz, and H. Gamboa, "TSFEL: Time series feature extraction library," *SoftwareX*, vol. 11, p. 100456, 2020. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S2352711020300017>
- [14] J. Foote, "Automatic audio segmentation using a measure of audio novelty," in *IEEE International Conference on Multi-Media and Expo*, 2000, pp. 452–455.
- [15] D. Anguita, A. Ghio, L. Oneto, X. Parra, and J. L. Reyes-Ortiz, "Human activity recognition on smartphones using a multiclass hardware-friendly support vector machine," in *Ambient Assisted Living and Home Care*, J. Bravo, R. Hervás, and M. Rodríguez, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 216–223.
- [16] A. Bhogal and A. Mani, "Pattern analysis of oxygen saturation variability in healthy individuals: Entropy of pulse oximetry signals carries information about mean oxygen saturation," *Frontiers in Physiology*, vol. 8, p. 555, 08 2017.
- [17] S. Lobov, N. Krilova, I. Kastalskiy, V. Kazantsev, and V. Makarov, "Latent factors limiting the performance of semg-interfaces," *Sensors*, vol. 18, p. 1122, 04 2018.

# Frame Work For EEG Based Emotion Recognition Based On Hybrid Neural Network

Jancy Mariam Jose and Aravinth.J

Department of Electronics and Communication Engineering

Amrita School of Engineering, Coimbatore

Amrita Vishwa Vidyapeetham, India

Email Id : jancym777j@gmail.com, j\_aravinth@cb.amrita.edu

**Abstract**—In recent years, there were many attempts to classify human emotions based on corporeal signals including ECG, EEG, EMG. EEG based emotion classification is more accurate because it cannot be tainted by subjects' will. The recent development in CNNs has made it easier to systematically extract features from EEG easily. But again, the traditional CNNs fail to comprehend the multi-channel aspect of EEG. In this work, a simple and efficient pre-processing method by considering baseline signals is proposed to enhance the accuracy of recognition and we proposed a hybrid neural network which combines Recurrent Neural Network (RNN) and Convolutional Neural Network (CNN) to identify human emotions by extracting spatial and temporal features from raw EEG stream effectively. In CNN, the 1D EEG sequence is then efficiently converted into a 2D frame structure. In order to extract the inter-channel connection between physically adjacent EEG signals, the CNN module is used, and to extract the contextual information, the LSTM module is used. Using this logic, we were able to create a deep learning model which predicts arousal and valence emotions with 86.98% and 85.82% accuracy respectively.

**Keywords**—EEG, CNN, RNN, LSTM, Spatial-temporal, SFV, TFV.

## I. INTRODUCTION

Emotion is a powerful feeling that arises from one's circumstances, mood, or interactions with others that play an important role in everyday life. Biological states associated with the nervous system in the medical field are emotions caused by neuro-physiological changes that are in tandem with perceptions, feelings, reactions associated with behavioural, and a degree of satisfaction or dissatisfaction.

Facial expressions, voice, eye blinking, and physiological signals can detect human emotions. The first three methods, however, are vulnerable to the participants' subjective factors, i.e., participants may intentionally conceal their feelings. Although physiological signals are spontaneously produced by the human body, such as Electrooculography (EOG), Electroencephalograms (EEG), Blood Volume Pressure (BVP) [23]. Consequently, in recording actual human emotional states, the corporeal signals are more accurate. The EEG signals come out directly from the human brain than other physiological signals, so the changes in EEG signals represent the human emotional states changes. Many researchers aim to research human emotion by means of EEG signals for this purpose [[14],[16]]. Many systematic studies using machine learning already exist to distinguish the emotions using EEG signals. The standard techniques based on machine learning have proven successful in classifying

emotion states, but the lack of these suggests that people who research in this area must make multiple trials to design and find different characteristics from EEG signals which can point to certain emotions. The estimation of such characteristics will cost more time. A wide range of methods of extracting/ selecting features has been proposed in past years [1]. Fourier Transform (FT), Wavelet Transform (WT), and Power Spectral Density (PSD) [2] are the most common forms of feature extraction methods Chena and Zhang used a comparison approach between several models to extract the features and multiple different classifiers of machine learning which led to the discovery that the dynamics in the nonlinear characteristics contribute to greater efficiency [3].

Deep Learning (DL) based techniques have gained broad attention in recent years and some of the DL-based methods obtained competitive accuracy, since they had a humongous success in 2D data. Tabar, Yousef, and Ugur proposed a CNN-based method to extract EEG signal time, frequency, and position information characteristics, and a Stacked Auto Encoder (SAE) was used to boost the accuracy of the classification [4]. EEG-based recognition using movement intention, Zhang, Dalin, et al. suggested parallel and cascade convolutional recurrent neural network and obtained a great performance [5]. Li suggested a processing technique prior to feeding data in to model, that transforms the multiple channel EEG data into a corresponding 2-Dimensional frame like representation and combines RNN and CNN to classify emotional features at the test point [6]. Most EEG-based emotion recognition using CNN-based approaches continues to rely to a large extent on complicated methods for prior processing and specially engineered features, for example translating raw EEG signals into images [[4],[7],[8],[24]], which do not use deep learning skills. The studies based on recognition of emotions use the EEG directly considering the baseline, to solve this problem inexpensively and a simple approach for pre-processing was proposed. The baseline signals were considered and eventually converted a one-dimensional raw chain like EEG signal into two-dimensional frames such as sequences.

The rules to be followed when translating 1D vectors into 2D frames are: signals are always adjacent to the frame from physically adjacent channels so that the information such as spatial can be preserved after conversion. Next, the 1D vector and 2D frame are provided as the input to a hybrid DL architecture that combines the CNN and RNN, which is used in a single framework to perform tasks of emotion recognition. Primarily CNN extracts the spatial

characteristics from the input, and RNN extracts the temporal characteristics from the EEG series. Once the CNN and RNN processing is completed, the fusion of spatial and temporal characteristics is performed by the method of feature fusion. The methods described above are applied to the DEAP Dataset.

The rest of the paper is organised as follows: Section II deals with the related work. Section III describing our proposed method. Results and Discussion is described in the section IV and finally the Conclusion in section V.

## II. RELATED WORK

Two simple models such as the discrete model and the dimensional model, are historically categorized by emotion [9]. Discrete emotion theories suggest that the presence of distinct emotions is characterized by synchronized patterns of response in morphological expressions, and neural characteristics. The basic six emotions are fear, disgust, rage, sadness, happiness and surprise, defined by a distinct model to represent emotion [10][11]. Dimensional theories, however, claim that emotion is continuous, not discrete, and multiple consecutive values may define it. The popular dimensional theory is based on arousal, dominance, and valence. The human gratification from positive to negative is expressed by valence and the level of passive to active excitement is represented by arousal and dominance varies from a feeling of helplessness and vulnerability (without control) to a feeling of empowerment. Due to its simplicity and ability to explain emotions well, the three-dimensional space model is commonly used [12]. The DEAP Dataset, which tracks EEG signals consisting of thirty-two channels and peripheral signals comprising of eight channels of thirty-two subjects while they watch forty, one min videos, was proposed by Sander Koelstra et al. Extracted features of GSR, BVP, patterns of respiration, temperature of skin, EEG, EOG and EMG are included in the dataset. The signal noise of human physiological variations in their work makes a single-trial classification difficult. The DEAP Dataset is used for this work since they made the dataset publicly available [13].

A model based on emotion recognition using EEG signals with LSTM-RNN with 4-fold cross-validation was suggested by S.Alhagry et al. The advantage of the LSTM approach which introduces a high average precision compared to conventional methods and the limitation of their work observed was it uses a great deal of money and time to draw a conclusion. The precision obtained was 85.45% for valence emotion and 85.65% obtained for emotion in arousal band [15]. H.chao suggested a model based on ensemble DBN with Glia Chains (GC) and used a 10-fold cross-validation technique. An ensemble DL model combining parallel Deep Belief Networks-GCs and a restricted Boltzmann machine was proposed for the best detection of emotion. The model obtained 76.83% valence accuracy and 75.92% arousal accuracy [17].

Y.H.Kwon et.al. proposed a two-dimensional CNN fusion-based EEG model for emotion recognition system with the CNN fusion approach leaves one-out cross-validation. Their model improves human emotional

classification accuracy using the CNN model but their model is not suitable for multi-level feature extraction of EEG because the single CNN is not able to extract the features. The model obtained an accuracy of 80.46% for valence and 76.56% for arousal [18]. R.Sharma et al. proposed a Bi-LSTM 10 fold cross-validation technique for efficient recognition of human emotion in high speed but their proposed algorithm is highly data-dependent. The model obtained an accuracy of 84.16% for valence and 85.21% for arousal [19]. S.Siddharth et al. proposed a model based on SVM with 10 fold cross-validation for the description and model of important regions of brain which corresponds to various states that are versatile and capable of extracting features from different datasets. One of the main limitation with this model was the electrode placement causes issues while capturing frontal images of the subject and the model obtained an accuracy of 62.49% for valence, 62.17% for arousal, and 61.84% for dominance [20]. W.Liu, J.L Qiu et al. proposed a Multimodal recognition of emotion with a specific type of cross-validation using deep canonical correlation analysis using the VGG-16 network. Their model has greater mutual details, suggesting that processes of DCCA transformation retain information related to emotions and discard irrelevant information. It is hard for the model to represent the time synchronization between different features of the modality. The model obtained 78.99% valence accuracy and 79.23% arousal accuracy [21]. The work done by D. Song, X. Li, P. Zhang on CNN and RNN with transformation based on wavelet features obtained an accuracy of 72.06% for valence and arousal obtained an accuracy of 74.12% [26]. S. Tripathi, S. Acharya et al proposed a CNN with 101 extracted features in each channel obtained about 81.41%, 73.36% for valence and arousal respectively [27].

A novel DL architecture based on the combination of CNN and RNN in parallel for extracting spatial-temporal information is proposed in this work to address the above problems. Where the CNN extracts the spatial information and RNN extracts the temporal information.

## III. SYSTEM DESIGN

We have extensively researched on the existing DL models on recognizing emotions based on multiple channel EEG signal and we have come to the inexcusable conclusion that baseline signal plays a vital role in precise feature selection which in turn results in a robust classification model that can classify valence or arousal emotion in the provided EEG signal. None of the DL architecture can correctly identify both spatial and temporal features. In the proposed work used a parallel CNN-RNN (LSTM) architecture to identify both spatial and temporal features. DEAP Dataset was used to conduct the research which contains 60sec trial data and 3sec of 32 subjects baseline data. 40 one-minute-long clips are included in the emotional music videos and people participated were to rate each video on a scale of 1-10 based on liking, arousal, valence, and dominance. Two binary classification concerns, namely arousal, and valence can now be viewed as the task at hand.

In this session we introduce the DEAP Dataset analysis and followed by the pre-processing step, 1D EEG signal transformation to 2D EEG frames, and finally about the parallel convolutional recurrent neural network. Also, the detailed diagram of the model is being shown in Fig.1.

Secondly, we add these matrices variable wise and measure the mean value.

$$\text{Base Mean} = \frac{\sum_1^N \text{mat}_i}{N} \quad (1)$$

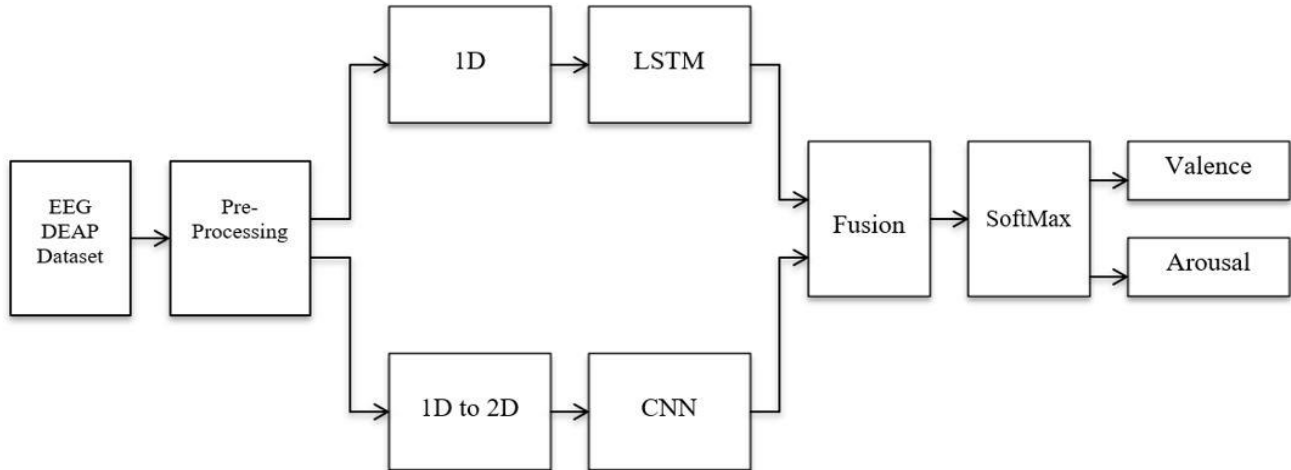


Fig. 1. Proposed model architecture for emotion recognition using EEG

#### A. EEG Deap Dataset

The DEAP Dataset registers thirty-two channel EEG data and eight-channel of peripheral natured signals. It is gathered from thirty-two participants when they were made to watch forty 1-minute videos. The dataset we used (DEAP) offers an already pre-processed signal version. At 512Hz the signals were sampled and then subsequently sampled to 128 Hz. A bandpass filter is applied to eliminate the artifact. The blind source separation technique is used to remove eye artifacts, such as to make it easier to analyse independent components. A 3sec long signal is recorded that is the baseline in a relaxed state and also a 60-sec signal which is during stimulation is included in each EEG signal sample.

Each DEAP Dataset signal may be split into sixty segments with a one-second sliding window which in turn contains one hundred and twenty-eight sampling points. There are forty signals in the dataset for each subject, which gives us 2400 EEG samples (Forty trials per sixty segments) for each subject. When downloaded, the size of the dataset was 5Gb and the size of the pre-processed dataset was about 8Gb.

#### B. Pre-processing

The pretrial information is used to assess the variations between experimental signals and baseline signals that are registered when the subjects are stimulated to significantly improve the efficiency of emotion recognition. First of all, pre-trial signals are taken out from all channels (No of electrodes) and the pre-trial signals is cut into small parts of a similar length and acquired a matrix  $N (C \times L)$ . Where  $C$  is the channel number, length is  $L$  and  $N$  part number.

After the above steps a matrix  $(C \times L)$  is obtained which is called the BASE MEAN MATRIX. Next, the raw EEG signals are segmented to  $M (C \times L)$  matrices known as RAW EEG matrix and find the base removed matrix by taking the difference between raw EEG matrix and base mean matrix. Finally, all of these base extracted matrices are merged into a large matrix of the same size as raw EEG signals.

#### C. One Dimensional EEG Signals Transformation to Two Dimensional EEG Frames

To capture EEG signals, the EEG based Brain Control Interface device uses a special headset with several electrodes. The International Framework 10-20 is a globally accepted approach for the classification and application of the electrode position on scalp and the cerebral cortex underlying that area [22]. “10” and “20” apply to the fact that either ten percent or twenty percent of the overall front to back and/or right to left distance of the skull is the real distance between the neighbouring electrodes [22]. In the DEAP Dataset, the internationally accepted 10-20 system of test electrodes is used to form a matrix. Every electrode is in adjacency to several other electrodes that record EEG signals of a particular region in the brain, whereas the chain-like one-dimensional EEG data vector components are confined to two neighbours. The arrangement of electrodes on the scalp is shown in Fig.2. According to the map showing the electrode placement, the 1-dimensional EEG signals are transformed into 2-dimensional frames and Fig.3 denotes the two-dimensional data frame  $f_t$  of the one-dimensional data vector  $v_t$  at time index  $t$ .

Zero value is used to denote unused signals in different channels in the DEAP Dataset that do not influence the neural network. The pre-processed one-dimensional vector sequences  $[V_t, V_{t+1}, \dots, V_{t+L}]$  are transformed to 2D data frame sequences  $[f_t, f_{t+1}, \dots, f_{t+L}]$  based on the electrode arrangement. Using Z-score normalization (2), after conversion the data frame is normalized over the nonzero elements.

$$Z = \frac{X - \mu}{\sigma} \quad (2)$$

Where;

$x$  represents the non-zero element from a certain location in the frame.

$\mu$  represents mean of all non-zero elements.

$\sigma$  represents the standard deviation of the elements.

Subsequently, to segment the streaming 2D frames into individual frame classes, the sliding window method is applied.

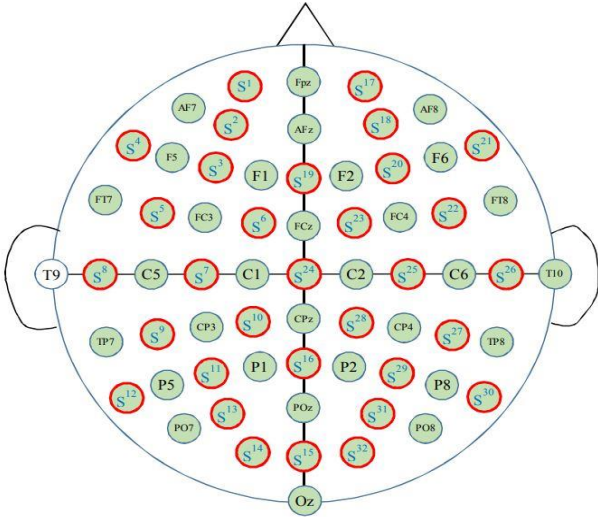


Fig. 2. Electrode distribution on scalp

$$f_t = \begin{bmatrix} 0 & 0 & 0 & S_t^1 & 0 & S_t^{17} & 0 & 0 & 0 \\ 0 & 0 & 0 & S_t^2 & 0 & S_t^{18} & 0 & 0 & 0 \\ S_t^4 & 0 & S_t^3 & 0 & S_t^{19} & 0 & S_t^{20} & 0 & S_t^{21} \\ 0 & S_t^5 & 0 & S_t^6 & 0 & S_t^{23} & 0 & S_t^{22} & 0 \\ S_t^8 & 0 & S_t^7 & 0 & S_t^{24} & 0 & S_t^{25} & 0 & S_t^{26} \\ 0 & S_t^9 & 0 & S_t^{10} & 0 & S_t^{28} & 0 & S_t^{27} & 0 \\ S_t^{12} & 0 & S_t^{11} & 0 & S_t^{16} & 0 & S_t^{29} & 0 & S_t^{30} \\ 0 & 0 & 0 & S_t^{13} & 0 & S_t^{31} & 0 & 0 & 0 \\ 0 & 0 & 0 & S_t^{14} & S_t^{15} & S_t^{32} & 0 & 0 & 0 \end{bmatrix}$$

Fig. 3. 2D matrix form

#### D. Parallel Convolutional Recurrent Neural Network

A deep learning model called ‘‘Parallel Convolutional Recurrent Neural Network’’ to distinguish emotion states, which is the combination of two forms of deep learning structures such as CNN and RNN. CNN and RNN merge

their strong ability to derive spatial and temporal characteristics, respectively. To gain correlation between cross channel and to remove characters from two dimensional frames, the CNN unit helps. The RNN structure is generally known to efficiently extract time dependent characteristics. The CNN and RNN are accompanied by a process of function fusion, which is used to gradually combine the extracted characteristics for final emotion recognition.

Deap dataset has pre-processed files for every 32 subjects. Every corresponding file has 40 rows pointing to the number of videos watched by the subject. Each row has 16,384 entries which have readings from all 32 channels. So, each channel has 512 entries which again is split into the four characteristic waves of EEG namely alpha, beta, theta, delta which have significantly different frequencies. Our deep learning model can figure out the most relevant features, i.e., the relevant patterns in the occurrence of crest and trough. When we have ample training data, the CNN-RNN can efficiently pull out such patterns and assign probability values for those to create a model. These feature weights can then identify features from unknown data and make predictions.

Three continuous two-dimensional convolutional layers in the CNN section have the same kernel size, which is 4x4 for extraction of spatial features. Generally, CNN’s kernel size is 3x3 and it is commonly used in computer vision field, but in our model a 4x4 filter is used because the two-dimensional EEG frame signals are sparse. Thus, the 4x4 filter will get more channel correlations than the 3x3 kernel. Also, zero-padding is used in each convolutional layer to avoid information missing on the edge of the input data.

There are 32 feature maps in the first convolutional layer and doubled the features for the remaining convolutional layers such as 64 and 128. In classical CNN architectures, a convolution layer is always followed up by pooling layer, which is not required in this model because the size of the data frame is very small in this EEG recognition task than that used in the computer vision area, so the use of pooling layer will result in more loss of data. A batch normalization (BN) operation is also followed by each convolution operation to boost the model training. The output of the three convolutional layers is combined together after the three convolutional layers a depth concatenate operation is used and thus the spatial characteristics are mapped into a large cube. To reduce the cube size into 13x9x9 we used 13x1x1 convolutional kernels and finally after shrinking the cube size we flattened it into a spatial feature vector (SFV).

The input of the CNN section is a pre-processed two-dimensional data frame. The RNN section is accountable for the extraction of temporal characteristics so the one-dimensional data vectors are not translated to sequences of 2D frames. The final input segment of CNN is represented as;

$$S_j = [f_t, f_{(t+1)}, \dots, f_{(t+s-1)}] \in R^{(S \times h \times w)} \quad (3)$$

Where;

$S$  = data frames

$j$  = position

Each of the segment is then given to 2D-CNN and obtained the SFV as follows:

$$SFV_j = conv2D(s_j), SFV_j \in R^{1053} \quad (4)$$

Long Short-Term Memory unit is used for constructing two stacked layers of RNN. The second RNN layer's input is the first RNN layer's output. At the current time stage  $t$ , the LSTM unit's hidden state of the first layer is denoted as  $h_t$ , and  $h_{t-1}$  is the previous time level hidden state ( $t-1$ ).

The output of the LSTM unit is the hidden state. Therefore, the second LSTM layer input sequence is the first LSTM layer's hidden state sequence. Since we concentrate primarily on the identification of emotion at the segment level rather than the level of the time stage, the output of the final time step is given to the next completely connected layer. RNN component is used to obtain the temporal attributes so the 1D EEG signal is not transformed into two-dimensional frames.

The input of the RNN part is:

$$R_j = [v_t, v_{t+1}, \dots, v_{t+s-1}] \quad (5)$$

Where;

$V_t$  = the time step  $t$  vector

$S$  = size of the window

$j$  = position in 1D array

Final time step hidden state:

$$h_{(t+s-1)} = LSTM(R_j), h_{(t+s-1)} \in R^d \quad (6)$$

Where;

$d$  = LSTM unit hidden state size

To improve the temporal information representation capability, a fully connected layer is added before and after LSTM layers. Final Temporal Feature Vector (TFV $_j$ ) of segment  $R_j$  is therefore represented as:

$$TFV_j = FC(h_{(t+s-1)}), TFV_j \in R^l \quad (7)$$

Where;

$l$  = last fully connected layer size.

The spatial and temporal features are merged together and obtained a single spatial-temporal feature vector and finally to predict the human emotion a softmax layer accepts it as input;

$$P_j = softmax(SFV_j, TFV_j), P_j \in R^n \quad (8)$$

Where;

$n$  = represents the class quantity.

The fused output from CNN and RNN is fed into a logit SoftMax layer with cross-entropy since our classes to be predicted are mutually exclusive. It then calculates the probability for this discrete classification task, i.e., predicting valence or arousal.

$$P(y_i|x_i; W) = \frac{e^{fy_i}}{\sum_j e^{fj}} \quad (9)$$

Where;

$W$ = input weight matrix

$j$ = set of weights of class  $j$

$i$ =  $i$  th position in column

$x$ = set of features

Dropout operation is applied as a form of regularization after the fully connected layers in the LSTM portion to avoid the cause of overfitting.

Table I

Developed model description

LAYERS	INPUT SHAPE
Conv_1	4x4x32
Conv_2	4x4x64
Conv_3	4x4x128
Conv_4	1x1x13
Flatten	9x9x13 cube into 1x1053 vector
Batch Normalization	1x1053
Dropout Regularization	1x1053
RNN FC in	1x1024
RNN FC out	1x1024

#### IV. RESULTS AND DISCUSSION

A hybrid deep learning model has been used in this work to identify emotion states. In which the spatial and temporal features of the EEG signal are established. The pre-processed EEG signals of the DEAP Dataset is used and the baseline signal provided in the dataset is removed from the original signal to make the EEG signal more skewed to emotions. The pre-processing step done in this work is the removal of the baseline signal from the original signal. The output after the pre-processing step is the baseline removed EEG signal in 1D format. This 1D signal is directly given to the LSTM module because in the LSTM module we are extracting the temporal features. For extracting the spatial features, we need a CNN network. The 1D signals can't be directly given to the CNN module so the conversion of 1D signal to 2D frame is done for extracting spatial features.

The 2D data frame is given to the convolutional layer to extract the spatial features. Continuous three two-dimensional convolutional layers with a filter size of 4x4 along with zero padding is used. 32,64 and 128 features are mapped in each convolution layer and it is followed by batch normalization. Batch normalization is commonly used to accelerate the model training. The output from each convolution layer is combined or fused together by depth concatenation operation. These spatial features are then mapped into a large cube format. Then to shrink the size of the cube to 13x9x9 we used 13x1x1 convolutional kernels. Finally, the spatial information in 13x9x9 matrix format is flattened to a single spatial feature vector.

RNN is responsible for extraction of temporal data, such that the one-dimensional data vector is not translated into a series of 2D frames. For building two stacked RNN layers, Long Short-Term Memory Unit is used. The input to the second RNN layer is the first RNN layer's output. The parallelly extracted spatial and temporal characteristics are

merged together to obtain a spatial-temporal feature vector and to predict human emotion as SoftMax layer receives it.

At first, the experiment is carried out to find the accuracy of the model without doing the pre-processing step. The pre-processing step consists of accounting for the presence of a baseline signal present in EEG signals with respect to a subject. The baseline signal provided in the DEAP dataset is then effectively compared and removed from the EEG signal and it is fed into the DL architecture. The accuracy obtained without performing this pre-processing step was 57.15% for valence and 57.58% for arousal. Which indicated the model performance was poor when we didn't consider the importance of the baseline signal. Then, we conducted the experiment by doing the pre-processing step. Then the model performance was substantially increased. The model obtained accuracy of 85.82% for valence and 86.98 for arousal. The increased accuracy of the model indicates that the baseline signal plays an important role and the table 3 shows that our model performance is better than some of the existing models.

Table II  
Model average accuracy

State	Without Baseline Removal	With Baseline Removal
Valence	57.15%	85.82%
Arousal	57.58%	86.98%

Table III

Performance analysis of model performance with existing studies

Sl. No:	Description	Valence (%)	Arousal (%)
1	Relevance vector classifier decision fusion and EEG graph-theoretic features for automatic affective state characterization [25]	69	67
2	Emotion recognition from multi-channel EEG data through convolutional recurrent neural network [26]	72.06	74.12
3	Using deep and convolutional neural networks for accurate emotion classification on DEAP Dataset [27]	81.41	73.36
4	Recognition of Emotions Using Multichannel EEG Data and DBN-GC-Based Ensemble Deep Learning Framework [28]	76.83	75.92
5	Electroencephalography Based Fusion Two-Dimensional (2D)-Convolution Neural Networks (CNN) Model for Emotion Recognition System [29]	80.46	76.56
6	Our model	<b>85.82</b>	<b>86.98</b>

#### IV. CONCLUSION

In this work, a superior emotion recognition model based on multi-channel EEG signals is developed. A hybrid deep learning model in this approach each designed specifically to extract the spatial and temporal features within the EEG signal. The Convolutional DL network is employed in parallel with Recurrent DL network in which the former is best suited to obtain the spatial features while the later concentrates on the temporal features. Furthermore a 5-fold cross validation for each subject is done to obtain the best model possible.

DL architecture makes it easy for feature selection without much domain knowledge. when the importance of baseline signal was incorporate and removed that from the signal that is to be fed into the model there was a significant increase in the models prediction accuracy. Also, for the best performance of CNN, the 1D signal in DEAP Dataset is converted into 2D signal with respect to the electrode placement on scalp which correlates with the selection of channel features. Various weight, learning rate adjustments and several other hyperparameter tuning to get the best prediction accuracy was performed. Thus, it can be said that the hybrid model can be used for several real time applications including health monitoring tasks, BCI tasks etc. Also, there are several areas to improve pertaining to the memory management while training the model considering the size of the dataset.

#### REFERENCES

- [1] Jenke, Robert, Angelika Peer and Martin Buss. "Feature extraction and selection for emotion recognition from EEG." *IEEE Transactions on Affective Computing* 5.3 (2014): 327- 339.
- [2] Alarcao, Soraia M and Manuel J. Fonseca. "Emotions recognition using EEG signals: a survey." *IEEE Transactions on Affective Computing* (2017).
- [3] Chen, Peng, and Jianhua Zhang. "Performance Comparison of Machine Learning Algorithms for EEG-Signal-Based Emotion Recognition." *International Conference on Artificial Neural Networks* Springer, Cham, 2017.
- [4] Tabar, Yousef Rezaei, and Ugur Halici. "A novel deep learning approach for classification of EEG motor imagery signals." *Journal of neural engineering* 14.1 (2016): 016003.
- [5] Dalin Zhang, Lina Yao, Xiang Zhang, Sen Wang, Weitong Chen and Robert Boots. "EEG-based Intention Recognition from SpatioTemporal Representations via Cascade and Parallel Convolutional Recurrent Neural Networks." *arXiv preprint arXiv:1708.06578* (2017).
- [6] Xiang Li, Dawei Song, Peng Zhang, Guangliang Yu, Yuexian Hou and Bin Hu. "Emotion recognition from multi-channel EEG data through Convolutional Recurrent Neural Network." *Bioinformatics and Biomedicine (BIBM)*, 2016 IEEE International Conference on. IEEE, 2016.
- [7] Li, Youjun, Jiajin Huang, Haiyan Zhou, Ning Zhong. "Human Emotion Recognition with Electroencephalographic Multidimensional Features by Hybrid Deep Neural Networks." *Applied Sciences* 7.10 (2017): 1060.
- [8] Liu, Wei, Wei-Long Zheng, and Bao-Liang Lu. "Emotion recognition using multimodal deep learning." *International Conference on Neural Information Processing*. Springer International Publishing, 2016
- [9] A.Al-Nafjan, Manar Hosny, Yousef Al-Ohali. "Review and classification of emotion recognition based on eeg brain-computer interface system research: a systematic review", *Appl. Sci.* 7 (12) (2017) 1239.
- [10] E.L. Broek, Ubiquitous emotion-aware computing, *Pers. Ubiquitous Comput.* 17 (1) (2013) 53–67.

- [11] P. Ekman, An argument for basic emotions, *Cogn. Emot.* 6 (3–4) (1992) 169–200.
- [12] G.K. Verma, U.S. Tiwary, Affect representation and recognition in 3d continuous valence–arousal–dominance space, *Multimedia Tools Appl.* 76 (2) (2017) 2159–2183.
- [13] Sander Koelstra, Mohammad Soleymani, Ashkan Yazdani, Touradj Ebrahimi, Anton Nijholt. “DEAP: A Database for Emotion Analysis Using Physiological Signals”, *IEEE Transactions on Affective Computing*, Vol. 3, No. 1, January–March 2012
- [14] J. Aravindh and S. Valarmathy, “Multi classifier-based score level fusion of multi-modal biometric recognition and its application to remote biometrics authentication,” *Imaging Sci. Journal* vol. 64, no. 1, 2016.
- [15] S. Alhagry, A.A. Fahmy, R.A. El-Khoribi, “Emotion recognition based on eeg using lstm recurrent neural network”, *Emotion* 8 (10) (2017) 355–358.
- [16] Amritha Varshini S and Aravindh J. “Matcher Performance-Based Score Level Fusion Schemes for Multi-modal Biometric Authentication System” 2020 6th International Conference on Advanced Computing and Communication Systems, ICACCS 2020, 9074446, pp. 79-85.
- [17] H. Chao, H. Zhi, L. Dong, Y. Liu, J. Dauwels, “Recognition of emotions using multichannel eeg data and dbn-gc-based ensemble deep learning framework”, *Intell. Neurosci.* 2018 (2018).
- [18] Y.-H. Kwon, S.-B. Shin, S.-D. Kim, “Electroencephalography based fusion twodimensional (2d)-convolution neural networks (cnn) model for emotion recognition system”, *Sensors* 18 (5) (2018).
- [19] R. Sharma, R.B. Pachori, P. Sircar, “Automated emotion recognition based on higher order statistics and deep learning algorithm”, *Biomed. Signal Process. Control* 58 (2020) 101867.
- [20] S. Siddharth, T. Jung, T.J. Sejnowski, “Utilizing deep learning towards multimodal biosensing and vision-based affective computing”, *IEEE Trans. Affect. Comput.* (2019) 1.
- [21] W. Liu, J.-L. Qiu, W.-L. Zheng, B.-L. Lu, “Multimodal emotion recognition using deep canonical correlation analysis”, 2019, arXiv:1908.05349 .
- [22] Jungchan Cho and Hyoseok Hwang, “Spatio-Temporal Representation of an Electroencephalogram for Emotion Recognition Using a Three-Dimensional Convolutional Neural Network”.
- [23] S. Veni and S. Thushara, “Multimodal approach to emotion recognition for enhancing human machine interaction - a survey,” *Int. J. Adv. Sci. Eng. Inf. Technol.*, vol. 7, no. 4, pp. 1428–1433, 2017
- [24] Saiharsha B, Abel Lesle A, B. Diwakar, Karthika R, Ganesan M. "Evaluating Performance of Deep Learning Architectures for Image Classification," 2020 5th International Conference on Communication and Electronics Systems (ICCES), Coimbatore, India, 2020, pp. 917-922, DOI: 10.1109/ICCES48766.2020.9137884
- [25] R. Gupta, and T. H. Falk, “Relevance vector classifier decision fusion and EEG graph-theoretic features for automatic affective state characterization,” *Neurocomputing*, vol. 174, pp. 875-884, Jan. 2016.
- [26] X. Li, D. Song, P. Zhang, G. Yu, Y. Hou, and B. Hu, “Emotion recognition from multi-channel EEG data through convolutional recurrent neural network,” *IEEE Int. Conf. Bioinform. Biomed. (BIBM)*, Guangdong, China, 2016, pp. 352–359.
- [27] S. Tripathi, S. Acharya, R. D. Sharma, S. Mittal, and S. Bhattacharya, “Using deep and convolutional neural networks for accurate emotion classification on DEAP Dataset,” *AAAI Conf. Innovative Appl.*, San Francisco, California, USA, 2017, pp. 4746–4752.
- [28] Hao Chao , Huilai Zhi , Liang Dong , and Yongli Liu. “Recognition of Emotions Using Multichannel EEG Data and DBN-GC-Based Ensemble Deep Learning Framework” *Hindawi Computational Intelligence and Neuroscience Volume 2018*, Article ID 9750904, 11 pages.
- [29] Yea Hoon Kwon, Sae-Byuk Shin and Shin-Dug Kim, “Electroencephalography Based Fusion Two-Dimensional (2D)-Convolution Neural Networks (CNN) Model for Emotion Recognition System” *Sensors* 2018, 18, 1383; doi:10.3390/s18051383



# Computational Approach to guide Mind Controlled Robotic Arm using BCI – A Review

Anna Latha M  
School of Electronics Engineering  
Vellore Institute of Technology Chennai,  
India  
annalatha.bme@gmail.com

Dr. Sathish E.  
School of Electronics Engineering  
Vellore Institute of Technology Chennai,  
India  
sathish.e@vit.ac.in

Dr. Florence Gnana Poovathy  
School of Electronics Engineering  
Vellore Institute of Technology Chennai,  
India  
florence.poovathy@vit.ac.in

**Abstract**— This paper reports a review of computational approach to guide mind controlled robotic arm using Brain Computer Interface (BCI). BCI controlled robotic arm provides an individual to use brain signals, Electroencephalogram (EEG) or Electrooculogram (EOG) for their physical activity. For assisting physically challenged people robotic arm can be a major rehabilitation device, which consists of five stages, such as Data acquisition, signal preprocessing, feature extraction, classification and interfacing robotic arm. This method is clinically important since it paves a way for making a fully dependent patient into a partially dependent person, which in turn supports their mental health and stability.

**Keywords**—BCI, EEG, EOG, Feature Extraction, Robotic arm.

## I. INTRODUCTION

The people with paralysis or movement disorders are struggling very much for moving their hands and doing their necessary daily activities, such as moving the objects and pick up any kind of objects, it should be a challenge for persons having paralysis. Electroencephalogram was a non-invasive approach to analyzing the asynchronous biological activities of neurons firing inner region of the brain having scalp electrodes. Now a days, this kind of robotic arm was developed for the reason to help the peoples to contact with their external environment. It includes grasping an object and orienting the palm. Upper extremities used for our day to day interactions, works related with external environment in and around us.

In traditional day's keyboard, mouse or joystick are used interfaces between the human and the physical environment. But now a day's human machine interfaces are available like Robotic arm. The previous tools were suite for person having residual motor abilities. Brain Computer Interfacing with robotic arm technique was the best approach for recognition of patient's intention and doing works with external environments. The aim non-invasive BCI is used for translating patients mind thoughts into control signals, for outer applications without any kind of surgeries in brain.

## II. Literatures on Robotic arm

Robotic arm is an device which has multidisciplinary research areas which consists of signal acquisition, processing, feature extraction, classification and mechanical parts integration with BCI.

### a. Data Acquisition

Olivier Valentin et al [1] presented the paper for the data acquisition, signal modification and signal changes of weak signals. Especially the device concentrates on electroencephalography signals. They used the device dubbed CochIEEG. The features of this CochIEEG has 4 kHz sampling rate, supports for active EEG electrodes, great potential for data transmission with the help of Wi-Fi module and low noise signal acquisitions. It is pocket size, affordable

and lightweight. So it is utilized for real time applications and wearable devices.

Vojkan Mihajlović et al [2] described the improvement of the wearable, wireless and daily life EEG applications. Recording human brain electrical activity has great support in helping us identify the brain functioning, preventing cognitive decline, mental disorders and increase our life quality. Non-invasive EEG surface electrode was the dominant for brain dynamics studies and real-time contacts of persons with external environment. For better performance EEG systems has to be converted from stationary wired systems used in the clinical practice now days, to wireless, wearable, comfortable lifestyle solutions and convenient that provide signals with high quality. The electrical activity of brain is mainly used for the brain signal analysis, brain signal acquisition, feedback generation and brain signal generation. They provide some of the aspects that need to improve the research to handle EEG artifacts such as end-user driven development, development of sophisticated approaches, standardization and sharing of EEG data.

Ali Shahidi et al [3] investigated the effects of low noise environment to acquire the scalp EEG. In Rustrel, France the (LSBB) low-noise underground laboratory is used to provide scalp EEG with the help of ultra-shielded capsule. They got clean EEG from 3 volunteers that can be acquired through LSBB capsule. Using this scalp EEG, they can detect the better mental activity with the great level of  $\beta$ -band activities. Then, the energy ratio of counting relaxed  $\beta$ -band was calculated with S-transform and the output is related with capsule and hospital. Final results of these LSBB capsule studies demonstrate the establishment of low-noise novel Electroencephalography benchmarks.

Marcelo Alejandro et al [4] implemented the acquisition of a multichannel EEG with the help of DDRL (Digital Driven Right Leg) and SE (Single Ended) amplifiers. This way of data acquisition reduces the number of parts. Initially front end amplifier is used. It will provide poor CMRR (Common-Mode Rejection Ratio). But this scheme DDRL and SE amplifiers overcomes the poor CMRR and reduce the power line interference at any EMI environments.

W.J.R. Dunseath et al [5] described the data detection and complete system with high resolution Electroencephalogram. The main parts of the system are rigid helmets, preamplifier EEG electrodes, software controlled gain amplifiers, optical coupling with 486 PC, digitizing and control differential instrumentation. Imbalanced and/or increased electrode impedances cause signal distortion and coupling to external electromagnetic fields. It will be reduced with the help of miniaturized pre-amplifiers mounted on separate electrodes. This type of system used to provide best performance of electrical activities compared to recent systems.

Alan Francisco Perez Vidal et al [6] presented the BCI used to controlling the robotic arm by getting brain signal with the help of visual stimuli. Here the brain signals are acquired through EEG electrodes. After that reduction of noise, signal characteristics extraction and classification of signal is

processed. Emotiv EPOC hardware used to detect the brain signals. OpenViBE software was used for the signal processing. Matlab, Arduino board, 2 servo motors are controlled to operate 2 joints with a five degrees of freedom robot commanded by evoked potential P300 type. Neuron signal from visual stimulus when a person needs to concentrate on specified image from computer screen. In order to find out the noise influence the experiments were recorded with and without visual and auditory noise. The obtained results efficiency is 100% for the both hearing noise and noiseless conditions. For visual efficiency 50% was reached. For servo motor control without noise efficiency 100%.

Shang-Lin Wu et al [7] presented a work HCI-Human computer interface) systems were controlled by EOG signals. The EOG signal may help human computer interface users can able to short-out most of the inconvenience and physical problems in daily life. For monitoring eye movements there is no effective multidirectional classification method is available. This paper describes classification method using wireless EOG based HCI device for identifying eye movements. The parts of the device include EOG signal classification algorithm, wireless EOG signal acquisition components and wet electrodes. There are 8 kinds of directions of the eye movements are down, down-right, down-left, up, up-left, up-right, right and left. The EOG classification algorithm is used to extract the features from the electrical signals from the eight directions and blinking. This 8 directions of eye movements will provide the real life conditions. This is the effective method to identify the eye movements.

Ludovico Minati et al [8] reported the hybrid control of a robotic arm by various bio signals such as EEG, EOG, EMG and head movement, which are acquired through consumer grade wearable device. It is possible to acquire above mentioned bio signals via wearable devices. This device is used to provide practical bio signal based interface for assistive robots under stereoscopic vision guidance the robotic arm performed automatically with 2 levels of computer approach. First level is robot ram steering combines proportional and discrete aspects. Second level the user needs to issue a single command. After these two level completions the user can automatically grasp the object.

## b. Artifact Removal

Saleha Khatun et al [9] provided the comparative study of ocular artifact removal for single channel EEG using wavelet based unsupervised algorithm. EEG was the approach for analyzing the neuronal signals of human brain. During EEG acquisition come artifacts such as eye blinking will corrupt the neuronal signals. So removal of OA is the major consideration of multiple channel EEG systems. In this paper they removed the OA from single channel EEG. Here the unsupervised wavelet transforms decomposition algorithms used. Seven set of EEG data was analyzed. There are 2 types wavelet transform methods applied. One is Stationary Wavelet Transform, and another one is Discrete Wavelet Transform. The OA artifacts were removed with the help of four basic WT functions, such as, *coif3*, *bior4*, *haar* and *sym3* with statistical threshold-ST. and universal threshold-UT. For removing Ocular artifacts 5 performance matrices were used: mutual information, time-frequency analysis, correlation coefficients and normalized mean square error. The wavelet transform can be the best technique for unsupervised ocular artifact removing from EEG with single channel for real life applications.

M. Chavez et al [10] described the artifact removal from single channel EEG with surrogate based algorithm. From single channel EEG there are two types of artifacts identified,

such as ocular artifacts and muscular artifacts. These artifacts were automatically removed with the surrogate based artifact removal (SuBAR) approach. Canonical correlation analysis and wavelet thresholding combined technique was the traditional approach to remove the artifacts. This surrogate based approach provides the relative error 4 to 5 times lower than previous techniques. This approach is a promising solution for sleep stage scoring, anesthesia monitoring and ambulatory health care systems.

Yuan Zou et al [11] reported automatic approach for the artifact-related independent components for removal of noise in EEG data. For EEG recording some artifacts also removed. In traditional approach the artifacts were removed by ICA (independent component analysis) technique, Since it can decompose signals into related independent components. This approach has lot of restrictions. So, this paper implemented the automatic algorithm for the detection of general artifacts. It has two parts:1) To identify the artifacts due to physiological origins – Event related feature 2) To identify non-biological artifacts. This algorithm was applied 10 subjects. It can effectively identify, separate and remove both non biological and physiological artifacts. This approach also used to enhance the signal quality.

W. De Clercq et al [12] presented a paper on removal of artifact and background in EEG for multichannel signals. They developed a method for modeling the common dynamics in multi-channel signals. In this approach the muscle artifacts were removed by comparing PCA the common dynamics approach provides better results.

Ruhi Mahajan et al [13] described eye blink artifact removal of EEG data with Kurtosis modified multiscale sample entropy (mMSE) and wavelet ICA. This mMSE and kurtosis was used to automatically identify the artifacts and denoise it using biorthogonal wavelet decomposition. ICA decomposed data the 95%, and two-sided confidence interval of the mean was needed to identify the threshold for mMSE, Kurtosis was used to determine the eye-blink related components. After that, the improved performance of the results showed in the reconstructed EEG data using the un-supervised approach in the form of correlation coefficient, mutual information and spectral coherence. The proposed algorithm was tested with 12 channel EEG system. The average sensitivity of the proposed system was 90% and 98% was the average specificity for the proposed work.

S. R. Sreeja et al [14] reported the removal of EB (Eye blink) artifacts from EEG data with sparsity. For using BCI systems the removal of EB artifacts are very important. There are lot of artifact removal technique was available. Such as, thresholding, ICA, wavelet transformation etc., these techniques are very much expensive and results in information loss. So these techniques are unsustainable for BCI system. The paper reported sparsity approach for EB artifact removal. There are two methods:1)nk-SVD based artifact removal method 2)MCA(Morphological component Analysis). Both methods were implemented both real and synthetic EEG data. This sparsity method removes the EB artifacts.

Yan Liu et al [15] implemented the robust muscle artifact approach for the few channel EEG. They implemented Fast Multivariate Empirical Mode Decomposition (FMEMD) and CCA approach. This FMEMD technique initially decomposes the EEG input into variety of multivariate (IMF) Intrinsic Mode Functions. Next all the multivariate IMFs are preprocessed by CCA. Finally, the sources with low auto – correlations are identified and removed.

## c. Feature extraction

Xinyang Li et al [16] represented the discriminative OA

correction for feature – learning in data analysis (EEG). Using ICA approach, the data or signal may loss. To avoid this kind of loss novel discriminative OA correction approach implemented. Using this approach, the artifact was automatically removed from EEG data without any extra ocular movement measurements. 68 subjects were evaluated. This approach is used to remove the artifacts and improve the discriminative power of a classifier.

Juan Andrés Mucarquer et al [17] presented the work on removal of artifacts improved with an array of EEG. Ensemble empirical mode decomposition with canonical correlation analysis (EEMD-CCA) is the efficient method for removing EEG artifacts. In order to increase the performance, gain they include EMG array with EEMD-CCA. More number of electrodes (128) was used. So it may be expensive. For inexpensive enhancement need to use a few electrodes.

Hamza Baali et al [18] described the transform based feature extraction technique for motor-imagery task classification. This transform is a signal dependent orthogonal transform, named as (LP-SVD) linear prediction singular value decomposition for extracting the features of EEG signal. The extracted feature was then classified using a logistic tree based model classifier. This LP-SVD approach has two methods. 1) (AAR) adaptive autoregressive 2) (DCT) Discrete cosines transform. The accuracy of LP-SVD approach is 67.35%. Then the extracted features will be classified with logistic tree-based model classifier. Accuracy of the classifier was 81.38%.

Yongkoo Park et al [19] represented the novel approach for the extraction of features on the EEG channel pairs.. Initially the preprocessing stage each pair of EEG channel will be demixed using 2D rotation matrix. It will reduce the artifacts between the channel pairs. After that, the correlation coefficient will be enhanced for cognitive task classification. They follow the criteria for optimization. Minimum within-class correlation coefficient distance, maximum Fisher ratio and maximum inter-class correlation coefficient distance. Fisher ratio is the ratio of inter-class correlation coefficient distance to within-class correlation coefficient distance. The output will provide the best classification performance.

Raja Majid et al [20] implemented the Deep learning ensembles and optimal feature selection approach for emotion recognition from EEG sensors. They recognize the emotional states of the subjects by analyzing the features of EEG signals. 21 healthy subjects were recorded with fourteen channel electroencephalography machine. Initially the patients asked to watch 4 types of images with following stimuli, such as, calm, happy, scared or sad). Next to the preprocessing stage Hjorth parameters such as mobility, complexity and activity were measured. Then the feature will be classified through Support Vector Machine (SVM), KNN, LDA, etc., for emotion classification.

Liwei Cheng et al [21] implemented the motor-imagery EEG feature extraction approach with energy PCA-DBN Expansion of DBN is Deep Belief-Networks. For rehabilitation (MI-EEG) motor imagery electroencephalography reflects the major attention. Combination of PCA and DBN is called PCA-DBN. Initially 2nd order moment was used to identify time domain of MI EEG. Then PCA was used to identify time domain interval and identify the PC feature points. Finally, classification task was completed by softmax classifier. The classification accuracy was 97.69%.

Tianwei Shi et al [22] described the feature extraction technique for BCI-EEG based on motor imagery. The feature extraction algorithms used here are CSP-Common Spatial Pattern, AAR-Adaptive Auto-Regressive and illustrates order accumulation, band energy feasibility and sample entropy for

characteristics of motor imagery classification. Next, the extracted features ease classified using common space classifier, Bayesian classifier and LDA.

#### d. Classification Algorithms

InanGuler et al [23] proposed the approach for EEG signal classification using multiclass SVM. The PNN (probabilistic neural network) and multilayer perceptron neural network also tested for classification. For feature extraction Lyapunov exponents and wavelet coefficients are extracted. This Lyapunov exponents and wavelet coefficients will represent the EEG signals. Finally the extracted features was classified using PNN and multiclass SVM and it has higher accuracies.

Pawel Andrzej et al [24] described the classification of motor imagery induced EEG patterns for handling uncertainty effects in BCI system design with Interval Type-2 Fuzzy Logic. There are many uncertainties are taken place during BCT. For reducing this uncertainty (T2fl) Type-2 Fuzzy Logic Approach was used. It has two - fold: 1) Interval T2FL system (IT2FLS) deals with inter-session and within-session 2) Fuzzy classifier used in both on-line and off-line EEG classification studies. For further enhancement many BCI classifiers were used, such as, kernel Fisher discriminant, support vector machines, conventional type-1 FLS, and linear discriminant analysis.

Muhammad Hamza Bhatt et al [25] introduced the neural network and optimal feature selection approach for soft computing based EEG classification. For noise removal filter bank and band pass filter was used. But these techniques will also remove some of the useful information. So avoiding this soft computing technique was used. For motor imagery based EEG classification sub band common spatial patterns with sequential backward floating selection was proposed in this paper. Using filter bank the signal was decomposed into sub bands for feature extraction. LDA was applied for training two datasets ie) EEG signals and open BCI dataset acquired by EmotivEpoC. The accuracy for the open BCI was 93.05% and the accuracy for the EEG signal was 85.00%. Finally, neural network based classification method was applied for classification.

Yu Zhang, Guoxu Zhou et al [26] implemented the Sparse Bayesian classification approach of EEG for BCI. Cross validation (CV) was used to determine the effectiveness of regularization it depends upon the selection of regularization parameters. CV imposes 2 main limitations on Brain computer interfaces are from the subject required more amount of training data set and for the classifier the calibrating time was long. These are the disadvantages of CV. They introduce a Sparse Bayesian classification approach by exploiting Laplace priors, that is SBLaplace for classification of EEG. Under Bayesian evidence the sparse discriminant vector was learned. In this approach all the features can automatically estimate from training data set with-out need of cross validation. This SBLaplace algorithm provides better overall performance compared to all other algorithms.

Qin Jiang et al [27] described the classification approach for the online motor-imagery EEG was clustering and adaptive CSP. They proposed online and fast generated Eigen decomposition approach by RLS-CSP. RLS-CSP expanded to Recursive Least Squares updates of the CSP filter coefficients. The classifier used here was ISCDC-Incremental Self-Training Classification algorithm based on Density Clustering based on density-clustering. This approach was feasible to develop the real-time performance for the on-line BCI systems.

#### e. BCI control of robotic arm

Lei Wu et al [28] developed the smart phone based Human robot interfaces for persons with disabilities. Human robot interfaces were needed to use for complex ADL-activities of daily living. The same way smart phones were packed with different kind of sensors, such as precise touch sensors, accelerometers and gyroscopes. Then it was interfaced with robot control. The developed BCI with smart phones provide low cost, intuitiveness, and environmental adaptability.

Christopher G. et al [29] presented a paper on BCI control in a virtual reality and applications for the IoT. The combination of VR systems and IoT we can offer real time control of physical environment. They demonstrated the validity of control the unity based virtual reality environment and commercial IoT device via direct neural interfacing.

Xiaogang Chen et al [30] represented the BCI and computer vision was controlled by high level robotic arm with combination of augmented reality. In this paper they realized a robotic arm by combine AR-augmented reality, SSVEP (steady-state visual evoked potential) brain computer interface and computer vision. There are 4 commands SSVEP-BCI is designed for the subject to select the specified object with the help of robotic arm. After selecting the object, the computer vision would provide the color and location of the objects. The robotic arm automatically picked up and placed with the objects. 12 participants were tested. The  $93.96 \pm 5.05\%$  was the classification accuracy of the system

Ji-HoonJeong et al [31] implemented the approach of Multidirectional CNN-BiLSTM Network with EEG (data) Signals for BCI robotic arm systems. Here the EEG can be acquired for imagery and movement execution. 15 subjects participated. MDCBN (multidirectional convolution neural network-bidirectional long short term memory network) based deep learning framework was proposed. NRMSE (normalized root mean square error) and CC (correlation coefficient) was used to measure the decoding performance with 6 directions in 3-D space. It has been applied for real world BCI controlled robotic arm.

Yiliang Liu et al [32] proposed a approach of dual arm robot working manipulation tasks based on motor-imagery tele-operation. Spatial pattern approach was used to identify the filtered EEG signals. For the reference commands human intentions have been recognized and classified. The classification based on the phenomena of event-related desynchronization/synchronization in the task space for robot, so the object manipulates the task guided by the subjects mind should be achieved.

Nicholas Cheng et al [33] presented the BCI-SRG (Brain Computer Interface based Soft Robotic Glove) for stroke rehabilitation. 11 subjects were separated into 2 sub-groups BCI-SRG (or) SRG. Each person undergoes 120 minute interventions per session. There are two therapies provided. 1)30-minute standard arm therapy 2) 90-minute experimental therapy. It consists of 18 sessions over 6 weeks. A final result of this approach provides sustained improvements and elicited perceptions of motor-movements.

Wei He et al [34] described a wireless BMI and BCI system for wearable robots. Initially the recorded EEG data will be transmitted to the computer and it will be interfaced with robotic arm via a Bluetooth. For effective processing noise removed, feature extracted and classified data will be used for the processing. Some feedback signal also acquired such as arm accelerated speed, joint angle and angular speed. Here noise removal wavelet transform was used. For feature extraction common spatial pattern algorithm was applied. For classification LDA algorithm used.

Antonio Frisoli et al [35] proposed a paper on a new gaze

independent upper limb exoskeleton for rehabilitation in the real-time tasks. This work based on Kinect based vision system. It was used to track 3D objects and eye tracking system for object selection. The robot will be used to assist the kinematic parameters such as, jerk, acceleration and speed. The system was evaluated with 3 healthy subjects and 4 chronic stroke patients. Classification error rate was  $89.4 \pm 5.0\%$ . This approach will be helpful for the earliest stage of recovery for stroke patients.

Patrick Ofner et al [36] developed a classification of movement imaginations of the robotic arm in two planes with non-invasive decoding method. Using EEG, they classified vertical and horizontal imagined rhythmic movement of the right-arm in healthy persons. Then it will be decoded with higher correlation. Here the imagined movements classified of one limb in 2 different movement planes. Here the classifier works based on decoding approach.

Chin-Teng Lin et al [37] described the BCI assistive system based on a wireless multi-functional SSVEP approach. This is a noninvasive, multi-functional assistive system and wireless which integrates steady state VEP based BCI and a robotic arm for assisting patients to feed themselves. There are 3 other functions: i) video entertainment; ii) video calling; and iii) active interaction also integrated. Accuracy achieved to this device was 90.91%.

### III CONCLUSION

Most of the paralyzed and partially disabled persons suffering for doing their day to day activities, where mind controlled robotic arms will be a boon for their assistance and support. Mind controlled robotic arm deals with EEG or EOG signal acquisition through single channel, multi-channel or 8 channel systems using scalp electrodes or wet electrodes. The acquired signals have ocular and muscular artifacts which has to be removed by suitable methodologies. Artifact removal approach such as wavelet transform, filter bank, band pass filter, surrogate based algorithm, ICA, MSE and robust muscle artifact removal approach, etc. were performed in literature for removal of such artifacts. Different types of feature extraction techniques were discussed like DCT, Correlation coefficient approach, etc., After that the extracted signals are classified using SVM, PNN, IT2FLS approach, LDA, SBLaplace transform, etc., Finally the classified output was interfaced with robotic arm through Bluetooth, Wi-Fi module, etc., Brain computer interface (BCI) approach was a gifted device for the identifying of subject intention and for interacting with external environments. Non-invasive BCIs used to transmits the subject thoughts, decoded by brain electrical activity into control signal, for external uses without a need of any brain surgery which intern supports the person who needs assistance. This technology can make a fully dependent person to a partially independent person, with which they can improve their physical and mental health.

### REFERENCES

- [1] Olivier Valentin et al, "Validation and Benchmarking of a Wearable EEG Acquisition Platform for Real-World Applications" *IEEE Transactions on Biomedical Circuits and Systems*, (Volume:13, Issue: 1, 2019), doi:10.1109/TBCAS.2018.2876240.
- [2] Vojkan Mihajlović et al, "Wearable, Wireless EEG Solutions in Daily Life Applications: What are we Missing?," *IEEE Journal of Biomedical and Health Informatics*, Volume: 19, Issue: 1, Jan. 2015, doi: 10.1109/JBHI.2014.2328317

- [3] Ali Shahidi et al, "Scalp EEG Acquisition in a Low-Noise Environment: A Quantitative Assessment", *IEEE Transactions on Biomedical Engineering* (Volume: 58, Issue: 8, Aug. 2011) doi:10.1002/adhm.201801084.
- [4] Marcelo Alejandro et al "A Multichannel EEG Acquisition Scheme Based on Single Ended Amplifiers and Digital DRL", *IEEE Transactions on Biomedical Circuits and Systems* (Volume: 6, Issue: 6, Dec. 2012)doi:10.1109/TBCAS.2012.2190733.
- [5] W.J.R. Dunseath et al, "Multichannel PC-based data-acquisition system for high-resolution EEG", *IEEE Transactions on Biomedical Engineering* (Volume: 42, Issue: 12, 1995) doi:10.1109/10.476129
- [6] Alan Francisco Perez Vidal et al, "Development of a Brain-Computer Interface Based on Visual Stimuli for the Movement of a Robot Joints", *IEEE Latin America Transactions* (Volume: 14, Issue: 2, 2016) doi:10.1109/TLA.2016.7437182
- [7] Shang-Lin Wu et al, "Controlling a Human-Computer Interface System With a Novel Classification Method that Uses Electrooculography Signals", *IEEE Transactions on Biomedical Engineering* (Volume: 60, Issue: 8, 2013) doi:10.1109/TBME.2013.2248154
- [8] Ludovico Minati et al, "Hybrid Control of a Vision-Guided Robot Arm by EOG, EMG, EEG Bio signals and Head Movement Acquired via a Consumer-Grade Wearable Device", *IEEE Access* (Volume: 4, Issue 2016) doi:10.1109/ACCESS.2017.2647851
- [9] SalehaKhatun et al, "Comparative Study of Wavelet-Based Unsupervised Ocular Artifact Removal Techniques for Single-Channel EEG Data", *IEEE Journal of Translational Engineering in Health and Medicine* (Volume: 4, issue: 2016) doi:10.1109/JTEHM.2016.2544298.
- [10] M. Chavez et al, "Surrogate-Based Artifact Removal From Single-Channel EEG", *IEEE Transactions on Neural Systems and Rehabilitation Engineering* (Volume: 26, Issue: 3, 2018) doi:10.1109/TNSRE.2018.2794184
- [11] Yuan Zou et al, "Automatic Identification of Artifact-Related Independent Components for Artifact Removal in EEG Recordings" :*IEEE Journal of Biomedical and Health Informatics* (Volume: 20, Issue: 1, 2016)doi:10.1109/JBHI.2014.2370646
- [12] W. De Clercq et al , "Modeling common dynamics in multichannel signals with applications to artifact and background removal in EEG recordings", *IEEE Transactions on Biomedical Engineering* ( Volume: 52, Issue: 12, 2005)doi:10.1109/TBME.2005.857669
- [13] Ruhi Mahajan et al, "Unsupervised Eye Blink Artifact Denoising of EEG Data with Modified Multiscale Sample Entropy, Kurtosis, and Wavelet-ICA", *IEEE Journal of Biomedical and Health Informatics* (Volume: 19, Issue: 1, 2015) doi:10.1109/JBHI.2014.2333010
- [14] S. R. Sreeja et al, "Removal of Eye Blink Artifacts from EEG Signals Using Sparsity", *IEEE Journal of Biomedical and Health Informatics* (Volume: 22, Issue: 2018) doi:10.1109/JBHI.2017.2771783
- [15] Yan Liu et al, "An Efficient and Robust Muscle Artifact Removal Method for Few-Channel EEG", *IEEE Access* (Volume: 7, Issue:03 December 2019) doi:10.1109/ACCESS.2019.2957401
- [16] Xinyang Li et al, "Discriminative Ocular Artifact Correction for Feature Learning in EEG Analysis", *IEEE Transactions on Biomedical Engineering* (Volume: 64, Issue: 8, Aug. 2017)doi:10.1109/TBME.2016.2628958
- [17] Juan Andrés Mucarquer et al, "17. Improving EEG Muscle Artifact Removal With an EMG Array", *IEEE Transactions on Instrumentation and Measurement* (Volume: 69, Issue: 3, March 2020) doi:10.1109/TIM.2019.2906967
- [18] Hamza Baali et al, "A Transform-Based Feature Extraction Approach for Motor Imagery Tasks Classification", *IEEE Journal of Translational Engineering in Health and Medicine* (Volume: 3, Issue: 16 October 2015) doi: 10.1109/JTEHM.2015.2485261
- [19] Yongkoo Park et al, "A Novel EEG Correlation Coefficient Feature Extraction Approach Based on Demixing EEG Channel Pairs for Cognitive Task Classification", *IEEE Access* (Volume: 8, Issue: 08 May 2020) doi:10.1109/ACCESS.2020.2993318
- [20] Raja Majid et al, "Optimal Feature Selection and Deep Learning Ensembles Method for Emotion Recognition From Human Brain EEG Sensors", *IEEE Access*(Volume: 5, Issue: 31 July 2017) doi:10.1109/ACCESS.2017.2724555
- [21] Liwei Cheng et al, "A Motor Imagery EEG Feature Extraction Method Based on Energy Principal Component Analysis and Deep Belief Networks", *IEEE Access*(Volume: 8, Issue:23 January 2020) doi:10.1109/ACCESS.2020.2969054
- [22] Tianwei Shi et al, "Feature Extraction of Brain-Computer Interface Electroencephalogram Based on Motor Imagery", *IEEE Sensors Journal*(Volume: 20, Issue: 20, 2020) doi:10.1109/JSEN.2019.2939343
- [23] InanGuler et al, "Multiclass Support Vector Machines for EEG-Signals Classification", *IEEE Transactions on Information Technology in Biomedicine* (Volume: 11, Issue: 2, 2007) doi:10.1109/TITB.2006.879600
- [24] Pawel Andrzej et al, "Designing an Interval Type-2 Fuzzy Logic System for Handling Uncertainty Effects in Brain-Computer Interface Classification of Motor Imagery Induced EEG Patterns", *IEEE Transactions on Fuzzy Systems*(Volume: 25, Issue: 1, 2017) doi: 10.1109/TFUZZ.2016.2637934
- [25] Muhammad Hamza Bhatt et al, "Soft Computing-Based EEG Classification by Optimal Feature Selection and Neural Networks", *IEEE Transactions on Industrial Informatics* (Volume: 15, Issue: 2019) doi:10.1109/TII.2019.2925624
- [26] Yu Zhang;Guoxu Zhou et al, "Sparse Bayesian Classification of EEG for Brain-Computer Interface", *IEEE Transactions on Neural Networks and Learning Systems* ( Volume: 27, Issue: 11, Nov. 2016) doi: 10.1109/TNNLS.2015.2476656
- [27] Qin Jiang et al, "An Adaptive CSP and Clustering Classification for Online Motor Imagery EEG", *IEEE Access* (Volume: 8, Issue: 2020) doi:10.1109/ACCESS.2020.3016700
- [28] Lei Wu et al, "Development of Smartphone-Based Human-Robot Interfaces for Individuals With Disabilities", *IEEE Robotics and Automation Letters* (Volume: 5, Issue: 4, 2020) doi: 10.1109/LRA.2020.3010453
- [29] Christopher G. et al, "Brain-Computer Interface Control in a Virtual Reality Environment and Applications for the Internet of Things", *IEEE Access*(Volume: 6, Issue: 27 Feb 2018) doi:10.1109/ACCESS.2018.2809453
- [30] Xiaogang Chen et al, "Combination of Augmented

- Reality Based Brain- Computer Interface and Computer Vision for High-Level Control of a Robotic Arm”, *IEEE Transactions on Neural Systems and Rehabilitation Engineering* (Volume: 28, Issue: 12, Dec. 2020) doi: 10.1109/TNSRE.2020.3038209
- [31] Ji-HoonJeong et al, “Brain-Controlled Robotic Arm System Based on Multi-Directional CNN-BiLSTM Network Using EEG Signals”, *IEEE Transactions on Neural Systems and Rehabilitation Engineering* (Volume: 28, Issue: 2020) doi:10.1109/TNSRE.2020.2981659
- [32] Yiliang Liu et al, “Motor-Imagery-Based Teleoperation of a Dual-Arm Robot Performing Manipulation Tasks”, *IEEE Transactions on Cognitive and Developmental Systems* (Volume: 11, Issue: 3, 2019) doi: 10.1109/TCDS.2018.2875052
- [33] Yiliang Liu et al, “Brain-Computer Interface-Based Soft Robotic Glove Rehabilitation for Stroke”, *IEEE Transactions on Biomedical Engineering* (Volume: 67, Issue: 12, 2020) doi: 10.1109/TBME.2020.2984003
- [34] Wei He et al, “A Wireless BCI and BMI System for Wearable Robots”, *IEEE Transactions on Systems, Man, and Cybernetics: Systems*(Volume: 46, Issue: 7, July 2016) doi: 10.1109/TSMC.2015.2506618
- [35] Antonio Frisoli et al, “A New Gaze-BCI-Driven Control of an Upper Limb Exoskeleton for Rehabilitation in Real-World Tasks”, *IEEE Transactions on Systems, Man, and Cybernetics*, (Volume: 42, Issue: 6, 2012) doi: 10.1109/TSMCC.2012.2226444
- [36] Patrick Ofner et al, “Using a Noninvasive Decoding Method to Classify Rhythmic Movement Imaginations of the Arm in Two Planes”, *IEEE Transactions on Biomedical Engineering* (Volume: 62, Issue: 3, 2015) doi: 10.1109/TBME.2014.2377023
- [37] Chin-Teng Lin et al, “A Wireless Multifunctional SSVEP-Based Brain-Computer Interface Assistive System”, *IEEE Transactions on Cognitive and Developmental Systems* (Volume: 11, Issue: 3, 2019) doi:10.1109/TCDS.2018.2820153.

# A new lead system for improved recording of P-wave amplitude and its significance with existing optimal leads

N. Prasanna Venkatesh, Arya Bhardwaj, J. Sivaraman\* *Member, IEEE*, B. Dhananjay

Bio-signals and Medical Instrumentation Laboratory

Department of Biotechnology and Medical Engineering

National Institute of Technology Rourkela, Odisha, India.

[519bm6012@nitrrkl.ac.in](mailto:519bm6012@nitrrkl.ac.in); [219bm1413@nitrrkl.ac.in](mailto:219bm1413@nitrrkl.ac.in); [jsiva@nitrrkl.ac.in](mailto:jsiva@nitrrkl.ac.in); [518bm1004@nitrrkl.ac.in](mailto:518bm1004@nitrrkl.ac.in)

**Abstract**— Improving P-wave amplitude is of clinical significance in the diagnosis of atrial arrhythmias. In this study, a new lead system is developed to improve atrial activity and compared the results with the existing optimal lead systems. The developed new lead system focuses on P-wave amplitude. The study involves 20 healthy male volunteers of mean age  $25 \pm 2.81$ . The existing lead systems adopted in this study are the Standard Limb Lead (SLL), P-lead, Modified Limb Lead (MLL), and the proposed new lead system. Only Lead-II values of SLL and MLL were taken into consideration. ECGs were recorded using Mindray Beneheart R12 ECG machine supported by the Glasgow algorithm. The new lead system showed an increased mean P-wave amplitude of  $169 \pm 67 \mu\text{V}$  in L-I and  $181 \pm 67 \mu\text{V}$  in L-II which was significantly greater than the  $129 \pm 46 \mu\text{V}$  in Lead-II of SLL, ( $p < 0.05$ ). The new lead system has the advantage of improving P-wave amplitude with minimal leads in less muscle area and also supports ambulatory recording.

**Keywords**— Atrial arrhythmias, Electrocardiogram, New lead system, Optimal leads, P-wave amplitude.

## I. INTRODUCTION

Electrocardiogram (ECG) is the recording of the electrical activity of the heart. Currently available gold standard method for recording ECG is the standard 12-lead electrode placement system. The fundamental challenge of this method is to detect the atrial ECG components, which are obscured by greater ventricular activity. Evaluating atrial ECG has a greater significance in detecting atrial related diseases. The atrial ECG components, mainly includes P-wave, which has a vital role in the detection of atrial arrhythmias like Atrial Fibrillation (AF) and assessing AtrioVentricular (AV) conduction mechanism [1]. The incidence of AF increases with increasing age and results in substantial morbidity with a 2.6 to 4.5 fold risk of stroke and associated with 50% to 90% of increased risk of death [2]. Karantoumanis et al. [3] observed that delay in P-wave onset and prolonged duration of waveforms are entrenched markers for predicting AF repetition in patients with normal sinus rhythm after the first incidence. Recording of P-waves is conventionally difficult, mainly during ambulation [4]. P-waves from standard 12 lead recording systems are highly difficult to distinguish from electrostatic noise due to incorrect skin preparation, motion artifacts, and electromyographic noise [5]. Generally, atrial activity is obscured by greater ventricular activity with its dominant peaks, which motivates to develop an ECG lead system that can acquire improved atrial amplitude for enhanced diagnosis of atrial arrhythmias [6]. Besides, improved atrial activity contributes more to clinical aspects

for diagnosis and treatment of atrial fibrillatory conditions [7].

The Standard Limb Lead (SLL) electrode placement for recording the ECG was first framed by Einthoven [8]. Chest leads and unipolar lead electrode placements were later established by Wilson et al. [9] and further standardized [10]. To enhance the atrial information, several modifications were done in the standard 12-lead placement by replacing the electrodes in close proximity to the atria [11]. Sir Lewis, in 1910, was the pioneer to develop an optimized placement of bipolar leads and acquired increased amplitude of atrial activity by placing the electrodes in the right second-costochondral junction and right fourth-intercostal space on AF [12]. Petrenas et al. [13] developed a lead system by modifying the Lewis lead by replacing the first electrode to the manubrium sterni, followed by changing the second electrode from the fourth to fifth right intercostal space to reduce motion artifacts by arm movement to monitor atrial arrhythmias during ambulation. The amplitude of the P-wave obtained is three times more than that of the original Lewis lead, and it was named as modified Lewis lead. Sivaraman et al. [14,15] introduced a Modified Limb Lead (MLL) system that stands alone in unmasking the atrial repolarization phase (Ta-wave) in PR segment, which is generally obscured by the QRS complex. The MLL system is designed by modifying the limb electrodes to thorax by placing the right arm electrode to the third right intercostal space, left arm and left leg electrodes to the fifth right intercostal space adjacent to one another across the mid-clavicular line and concluded that the duration of Ta-wave is longer than P-wave. The consequent study concluded that Ta wave duration is three times longer in AV block patients. Kennedy et al. [16] developed a P-lead by placing the electrodes on the right sternal clavicular junction and the midpoint of the left costal margin inline with seventh intercostal space, which outperformed by resulting increased P-wave RMS from all other analyzed lead systems that are considered.

A comparative study [17] between SLL and MLL reports that SLL lacks significant information on atrial ECG components. In this study, a new lead system is proposed to improve P-wave amplitude during normal atrial activation time by placing the electrodes on less muscle area in the thorax. Then comparative study was done for new lead system with the existing leads such as SLL, MLL, and P-lead in terms of amplitude and duration of P-wave, R-wave, and PR-Interval (PRI) to compare the signal strength.

\*Corresponding author: Dr. J. Sivaraman, Bio-signals and Medical Instrumentation Laboratory, National Institute of Technology Rourkela.

## II. METHODS

### A. Study Population

The study included 20 healthy male volunteers with a mean age of  $25 \pm 2.81$  years. All volunteers gave informed consent for their participation in this study. None of them were having any clinical abnormality affecting the cardiovascular system.

### B. New Lead System

The electrode placements of the new lead system proposed in this study is shown in Fig. 1. The bipolar lead L-I is obtained by placing electrode 1 on manubrium just below the suprasternal notch at the upper sternal end, and electrode 2 is placed in the right seventh clavicular junction near costal margin. L-II is obtained from electrode 1 and electrode 3. Electrode 3 is placed on the left seventh clavicular junction near the costal margin. L-III is formed between electrodes 2 and 3. The right leg electrode is the reference electrode positioned over the right ankle. The polarity is negative for electrode 1, and the polarity of electrode 2 and electrode 3 is positive; with this, the ECG was recorded. The unipolar chest leads from the standard 12-lead system are not included in this study.

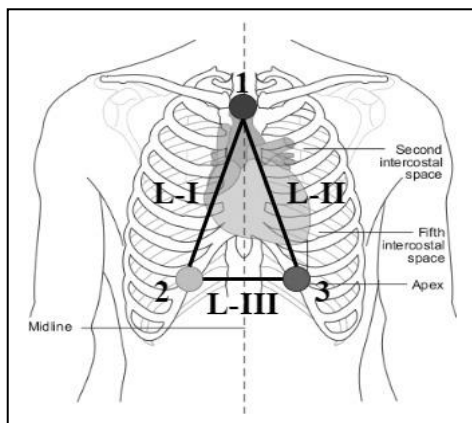


Fig. 1. Electrode placement of New Lead System.

### C. Data Acquisition

The ECG data were recorded using Mindray Beneheart R12, a 12-channel electrocardiograph machine supported with the Glasgow algorithm [18]. It operates at the sampling rate of 1 KHz, and the frequency response is between 0.05 Hz to 150 Hz. Data were recorded in four different lead configurations (SLL, MLL, P-lead, and new lead) for 10 seconds each in supine rest position at standard paper speed of 25 mm/s and gain of 10 mm/mV. Muscle artifact filter of 25 Hz and baseline drift removal of 0.56 Hz was applied. The limb electrodes, surface electrodes, and suction electrodes were used for recording along with silver chloride (AgCl) electrolyte. Other than SLL, all other lead systems were recorded by placing the electrodes on torso. Only lead II values of SLL and MLL were considered in this study, which have greater P-wave amplitude as discussed in previous studies [1], other than the new lead system. The waveform morphologies that are obtained from different lead systems are shown in Fig. 2.

### D. Statistical Analysis

The data given in mean  $\pm$  standard deviation (SD). Statistical significance (P-value) of the acquired data was analyzed using the Student t-test. The data were distributed normally, and tested by using the Shapiro-Wilkinson W test. The Origin Pro, version 2020b of Origin Lab Corporations, was used for statistical analysis in this study.

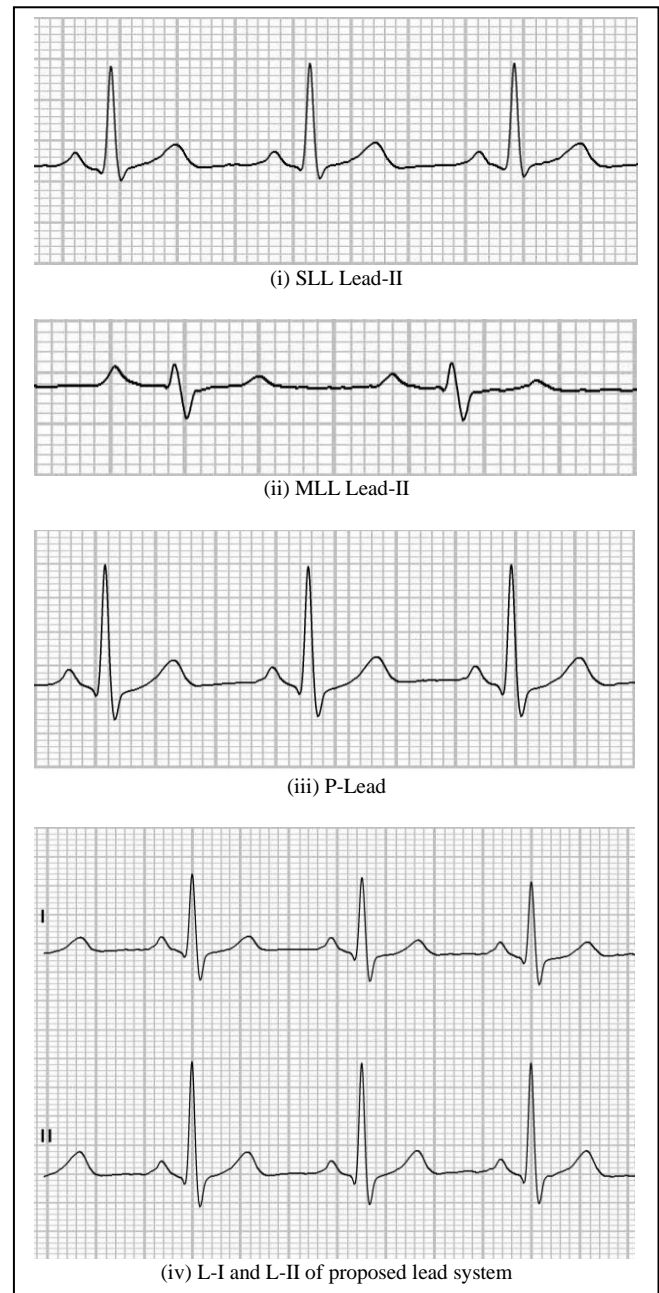


Fig. 2. (i) Lead-II of SLL, (ii) Lead-II of MLL, (iii) P-lead, and (iv) Lead-I and II of new lead system at paper speed of 25 mm/sec and 10 mm/mV.

## III. RESULTS

The atrial ECG parameters like amplitude and duration of P-wave and PRI are plotted in box and whisker plot. In addition, R-wave peak amplitude is also plotted. The plot of mean  $\pm$  SD for P-wave amplitude between three bipolar and three unipolar leads of four different lead systems are also studied. The P-wave amplitude variations in SLL, P-lead, MLL, and new leads L-I and L-II are shown in Fig. 3.



TABLE 1. AMPLITUDE AND TEMPORAL CHANGES IN DIFFERENT LEADS AND ITS SIGNIFICANCE

Measurements	SLL-II	MLL-II	P-Lead	New Lead-L-I	New Lead-L-II	P-value*
<b>P-wave amplitude (µV)</b>	129 ± 46.61	105 ± 40.62	177 ± 65.13	169 ± 67.70	181 ± 67.64	<0.05 #SLL-II & MLL-II are NS
<b>R-wave amplitude (µV)</b>	1205 ± 498.26	616 ± 373.01	1999 ± 645.96	1285 ± 584.80	2010 ± 672.51	<0.05 #SLL-II & New lead L-I are NS
<b>Heart Rate (bpm)</b>	81 ± 13.56	76 ± 10.88	76 ± 13.32	73 ± 9.95	73 ± 9.95	NA
<b>P-wave Duration (ms)</b>	96 ± 10.31	89 ± 12.91	92 ± 16.19	96 ± 11.70	97 ± 11.22	>0.05
<b>PR-Interval (ms)</b>	139 ± 16.08	139 ± 21.61	129 ± 21.19	136 ± 19.74	136 ± 20.23	>0.05

Values represented in mean ± SD; NS – not significant; NA – not applicable; \*P-value between SLL and other optimal leads.

The minimum, median and maximum values for P-wave amplitude in SLL, P-lead, and MLL are 60 µV, 112 µV, and 218 µV; 88 µV, 162 µV, and 297 µV; and 20 µV, 111 µV and 178 µV. The minimum, median and maximum values of P wave amplitude for the new leads L-I are 79 µV, 167 µV, and 291 µV; and for L-II are 43 µV, 175 µV, and 296 µV. The amplitude values of P-wave between SLL and the newly designed leads L-I and L-II were statistically significant with  $p < 0.05$ .

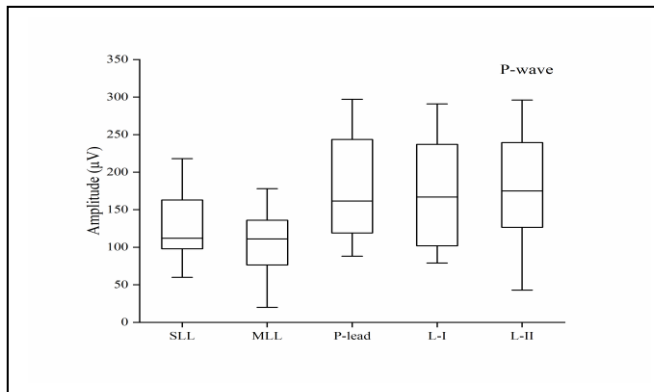


Fig. 3. Box and whisker plot for P-wave amplitude in different leads.

The P-Wave Duration (PWD) plot in different lead systems is shown in Fig. 4. The values of PWD are expressed in minimum, median and maximum for SLL, MLL, and P-lead as 76 ms, 98.5 ms, and 118 ms; 68 ms, 91 ms, and 113 ms; and 42 ms, 97 ms, and 114 ms, respectively. The minimum, median and maximum values of PWD for the new leads L-I and L-II are 74 ms, 97.5 ms, and 111 ms; and 74 ms, 100 ms, and 110 ms. The duration of P-wave in new leads were not statistically significant with the existing leads,  $p > 0.05$ .

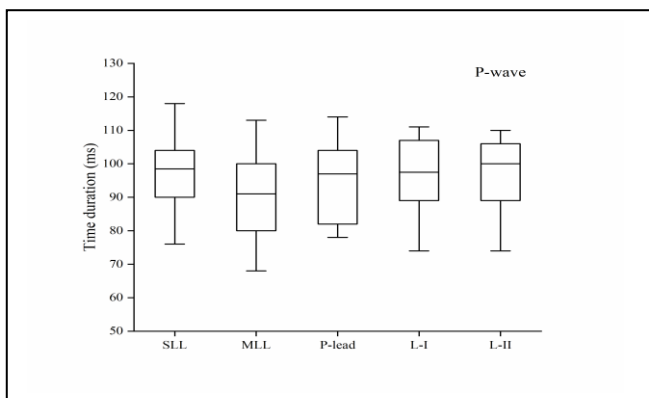


Fig. 4. Box and whisker plot for P-wave duration in different leads.

Fig. 5 shows the PRI values in different lead systems. The minimum, median and maximum values of PRI for SLL, MLL, and P-lead are 105 ms, 141 ms, and 173 ms; 90 ms, 144 ms, and 171 ms; and 92 ms, 129.5 ms, and 162 ms, respectively. For new lead L-I are 94 ms, 141 ms, and 166 ms; and L-II are 96ms, 143 ms, and 166 ms. The PRI values between the different leads have no significant changes and are not statistically significant,  $p > 0.05$ .

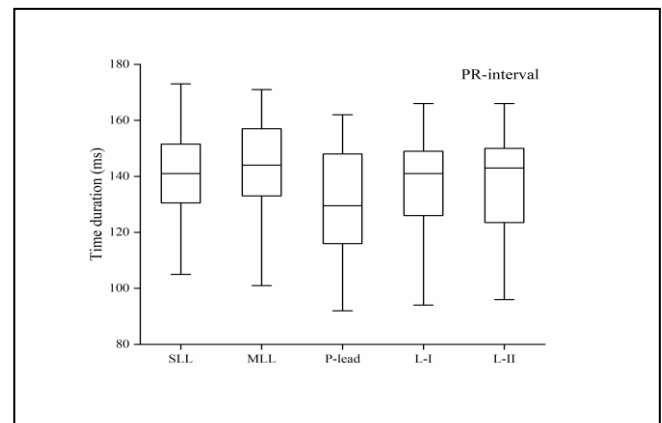


Fig. 5. Box and whisker plot for PR-interval in different leads.

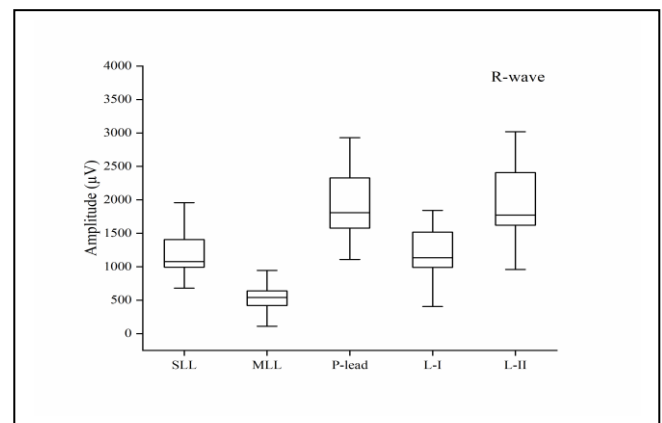


Fig. 6. Box and whisker plot of R-wave amplitude in different leads involved.

Fig. 6 shows the amplitude of R-wave peaks in different leads under this study. The minimum, median and maximum values of R-wave amplitude in SLL, MLL, and P-lead are 75 µV, 1076 µV and 2555 µV; 110 µV, 540 µV, and 1690 µV; and 1108 µV, 1809 µV, and 3850 µV, respectively. The minimum, median and maximum values of R-peak amplitude for the new leads L-I are 407 µV, 1134 µV, and 2875 µV; and L-II are 959 µV, 1773 µV, and 3854 µV,

respectively. The R-wave peak amplitude in L-II was larger than the R-peak in SLL, and statistically significant,  $p < 0.05$ . The low R-peak in MLL was noted, a unique characteristic of MLL lead system that enhances the atrial activity comparable to ventricles for improved analysis. The R-peak amplitude in SLL and new lead L-I were not statistically significant,  $p > 0.05$ .

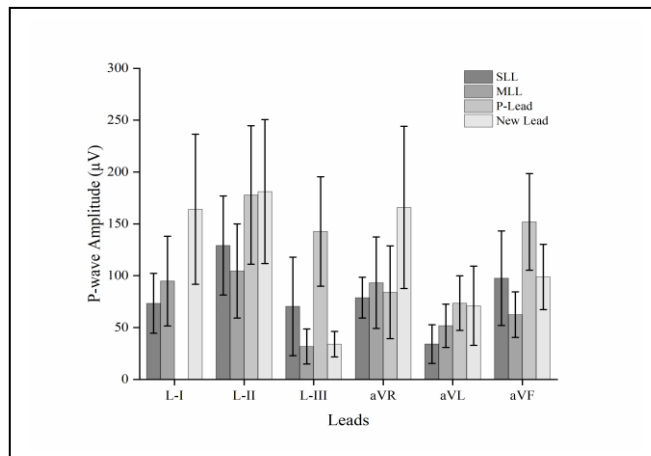


Fig. 7. P wave amplitudes for different lead system.

Fig. 7 represents the amplitude of P-wave from the studied lead systems in L-I, L-II, L-III, aVR, aVL, aVF. P-wave amplitude is plotted in terms of mean  $\pm$  SD. Absolute values of lead aVR is considered. The amplitude of P-wave has higher values in major leads of new lead system. P-lead was a single lead system and has no values in lead L-I. The maximum amplitude value of the P-wave obtained was  $181 \pm 70 \mu\text{V}$  from the lead L-II of the new lead system. The second and third maximum values are  $177 \pm 67 \mu\text{V}$  and  $169 \pm 68 \mu\text{V}$  from the P-lead and L-I of the new lead system.

#### IV. DISCUSSION

The new lead system proposed in this study, is parallel to the direction of electrical conduction vector of the heart. The new leads L-I is parallel to the electrical conduction in atria. So, greater P-wave amplitude is achieved in L-I of new lead system compared to the lead II of SLL. Moreover, the new lead L-II is parallel to the heart's electrical mean vector. Thus, a greater deflection of ECG waves in terms of peak amplitude was observed. The amplitude and temporal changes in different lead systems and their significance values were tabulated in Table. 1. The new lead L-II outperforms in terms of P-wave and R-wave amplitude when compared to existing optimal leads. The Ta-wave was observed in few volunteers for the new leads of this study, represented by a small depression following P-wave. Previously, MLL lead system proposed by Sivaraman et al. [14] unmasked the Ta-wave in sinus rhythm volunteers. The P-lead proposed by Kennedy et al. [16] noted greater P-wave amplitude, and it has the second-highest amplitude in this study. The L-I and L-II of the new lead system outperformed in P-wave amplitude compared to lead-II of SLL, and their difference is statistically significant with  $p < 0.05$ , which can be the future advantage of new lead system to diagnose the atrial diseases. Reduced R-wave amplitude was noted in MLL ECG recording of this study,

which is consistent with previous finding [19]. The differences observed between P-wave and R-wave amplitudes in different recording positions in SLL, MLL, P-lead, and new leads L-I and L-II were majorly due to the false positive QRS axis shift in the frontal plane due to different electrode positioning [20]. No significant changes were noted for PWD in all the lead systems, which is in line with the previous study in sinus rhythm and sinus tachycardia [21]. The PRI observed in this work depicted no changes, and similar results have been shown in [22-24]. The study results of P-wave amplitude in SLL are in accordance with the previous work [25].

The newly proposed lead system of this study, improved P-wave amplitude with minimal leads in maximal possible ways, and also supports prolonged recording of ECG during ambulatory purposes. The electrodes are placed in less muscle area, which reduces EMG noise and artifacts. Both L-I and L-II were considered for analysis as they have the greater acquisition of atrial activity compared to other leads. L-III of this lead system is not considered for the study, which has lesser amplitude as it is perpendicular to the heart's electrical mean vector.

#### V. CONCLUSION

The new lead system proposed in this study outperformed the other lead system in terms of higher P-wave amplitude in two different leads L-I and L-II. The amplitude of P-wave from the new leads is significantly greater than lead-II of SLL. As the new lead system focuses more on amplitude aspects of atria in its maximal leads, it is named as Atrial Lead (AL) System with atrial leads. Also, during normal atrial activation, the selection of optimized lead placement provides greater signal strength. This study considered only male volunteers, further study is required to validate it in female volunteers. Also, further comparative study of the new lead with more existing optimal lead systems in different heart conditions is yet to be established.

#### ACKNOWLEDGMENT

The authors acknowledge the support from the Ministry of Education, Government of India. The present study was supported by financial grants from the Science Engineering Research Board (SERB), Department of Science and Technology (DST), Government of India (EEQ/2019/000148).

#### REFERENCES

- [1] R. L. Lux and R. Greg, "New leads for P wave detection and arrhythmia classification," *J. Electrocardiol.*, vol. 37, pp. 80, October 2004.
- [2] E. J. Benjamin et al., "Impact of atrial fibrillation on the risk of death: the Framingham Heart Study," *Circulation*, vol. 98, pp. 946-952, September 1998.
- [3] I. Karantoumanis et al., "Atrial conduction time associated predictors of recurrent atrial fibrillation," *Int. J. Cardiovasc. Imag.*, pp. 1-11, January 2021.
- [4] F. Portet, "P wave detector with PP rhythm tracking: Evaluation in different arrhythmia contexts," *Physiol. Meas.*, vol. 29, pp. 141, January 2008.
- [5] J. E. Waktare et al., "Optimum lead positioning for recording bipolar atrial electrocardiograms during sinus rhythm and atrial fibrillation," *Clin. Cardiol.*, vol. 21, pp. 825-830, November 1998.

- [6] A. L. M. Bakker et al., "The Lewis lead: making recognition of P waves easy during wide QRS complex tachycardia," *Circulation*, vol. 119, pp. 592-593, June 2009.
- [7] P. G. Platonov, V. D. A. Corino, M. Seifert, F. Holmqvist, L. Sörnmo, "Atrial fibrillatory rate in the clinical context: natural course and prediction of intervention outcome," *EP Europace*, vol. 16, pp. iv110-iv119, November 2014.
- [8] W. Einthoven, "The different forms of the human electrocardiogram and their signification," *The Lancet*, vol. 179, pp. 853-861, March 1912.
- [9] F. N. Wilson, F. D. Johnston, A. G. MacLeod, P. S. Barker, "Electrocardiograms that represent the potential variations of a single electrode," *Am. Heart. J.*, vol. 9, pp. 447-471, April 1934.
- [10] A. Barnes, H. E. B. Pardee, P. D. White, F. N. Wilson, C. C. Wolfert, "Standardization of precordial leads - Supplementary report," *Am. Heart. J.*, vol. 15, pp. 235-239, February 1938.
- [11] Z. Ihara, A. van Oosterom, V. Jacquemet, R. Hoekema, "Adaptation of the standard 12-lead electrocardiogram system dedicated to the analysis of atrial fibrillation," *J. Electrocardiol.*, vol. 40, pp. 68.e1-8, January 2007.
- [12] N. Goldschager, M. J. Goldman, "Principles of Clinical Electrocardiography," 1989.
- [13] A. Petrėnas, V. Marozas, G. Jaruševičius, and L. Sörnmo, "Modified lewis ECG lead system for ambulatory monitoring of atrial arrhythmias," *J. Electrocardiol.*, vol. 48, pp. 157-163, March 2015.
- [14] J. Sivaraman, G. Uma, S. Venkatesan, M. Umapathy, and M. S. Ravi, "Unmasking of atrial repolarization waves using a simple modified limb lead system," *Anatol. J. Cardiol.*, vol. 15, pp. 605-610, August 2015.
- [15] J. Sivaraman, G. Uma, S. Venkatesan, M. Umapathy, and V. E. Dhandapani, "Normal limits of ECG measurements related to atrial activity using a modified limb lead system," *Anatol. J. Cardiol.*, vol. 15, pp. 2-6, January 2015.
- [16] A. Kennedy, D. D. Finlay, D. Guldenring, R. R. Bond, and J. McLaughlin, "Detecting the elusive P-wave: a new ECG lead to improve the recording of atrial activity," *IEEE Trans. on Biomed. Eng.*, vol. 63, pp. 243-249, February 2016.
- [17] S. Karimulla, J. Sivaraman, "The role and significance of atrial ECG components in standard and modified lead systems," *Electronic Systems and Intelligent Computing (ESIC 2020)*, vol. 686, pp. 347-355, September 2020.
- [18] P. W. Macfarlane et al., "Methodology of ECG interpretation in the Glasgow program," *Meth. Inf. Med.*, vol. 29, pp. 354-361, 1990.
- [19] J. Sivaraman, G. Uma, and M. Umapathy, "A modified chest leads for minimization of ventricular activity in electrocardiograms," *International Conference On Biomedical Engineering (ICoBE 2012)*, pp. 79-82, IEEE, February 2012.
- [20] J. Sivaraman, S. Venkatesan, R. Periyasamy, J. Joseph, and S. M. Ravi, "Modified limb lead ECG system effects on electrocardiographic wave amplitudes and frontal plane axis in sinus rhythm subjects," *Anatol. J. Cardiol.*, vol. 17, pp. 46-54, January 2017.
- [21] B. Dhananjay, J. Sivaraman, "The role of heart rate variability in atrial ECG components of normal sinus rhythm and sinus tachycardia subjects," *Advances in Intelligent Systems and Computing (AISC 2020)*, vol. 1171, pp. 637-644, August 2020.
- [22] J. Sivaraman, G. Uma, P. Langley, M. Umapathy, S. Venkatesan, G. Palanikumar, "A study on stability analysis of atrial repolarization variability using ARX model in sinus rhythm and atrial tachycardia ECGs," *Comput. Meth. Prog. Bio.*, vol. 137, pp. 341-351, December 2016.
- [23] A. Jyothsana, B. Arya, J. Sivaraman, "Stability analysis on the effects of heart rate variability and premature activation of atrial ECG dynamics using ARMAX model," *Phys. Eng. Sci. Med.*, vol. 43, pp. 1361-1370, November 2020.
- [24] R. John, and J. Sivaraman, "Effects of sinus rhythm on atrial ECG components using modified limb lead system," In 2017, 4<sup>th</sup> International Conference on Signal Processing, Computing and Control (ISPC 2017), pp. 527-530, IEEE, September 2017.
- [25] P. W. Macfarlane et al., "Normal limits of the electrocardiogram in Indians," *J. Electrocardiol.*, vol. 48, pp.652-668, July 2015.

# Automated EEG Analysis for Early Diagnosis of Epilepsy: A Comparative Study to Determine Relative Accuracy of Arithmetic and Huffman Coding Algorithms

Anisha Kumar  
School of Electronics Engineering  
Vellore Institute of Technology  
Vellore 632014, India  
[anisha.k.kumar@gmail.com](mailto:anisha.k.kumar@gmail.com)

Pratishtha Singh  
School of Electronics Engineering  
Vellore Institute of Technology  
Vellore 632014, India  
[pratishthasinghh@gmail.com](mailto:pratishthasinghh@gmail.com)

Rajlakshmi Khawas  
School of Electronics Engineering  
Vellore Institute of Technology  
Vellore 632014, India  
[rajlakshmi.khawas14@gmail.com](mailto:rajlakshmi.khawas14@gmail.com)

Priscilla Dinkar Moyya  
School of Electronics Engineering  
Vellore Institute of Technology  
Vellore 632014, India  
[priscilladinkar@gmail.com](mailto:priscilladinkar@gmail.com)

Mythili Asaithambi  
School of Electronics Engineering  
Vellore Institute of Technology  
Vellore 632014, India  
[mythili.asaithambi@vit.ac.in](mailto:mythili.asaithambi@vit.ac.in)

**Abstract**— Epilepsy is a prevalent neurological disorder typically characterized by recurrent seizure activity and detected using an electroencephalogram (EEG). The manual inspection of EEG however is a challenging and slow process that is susceptible to visual errors and variability amongst subjects. Hence, significant efforts have been made towards developing algorithms for automated epilepsy diagnosis and detection. The present study focuses on comparing two algorithms employing arithmetic encoding and Huffman encoding to separate epileptic signals from seizure-free (normal) samples. The proposed diagnostic technique comprises three major steps. In the first step, discrete wavelet transform (DWT) is used to decompose the EEG signal into detail and approximation coefficients. The second step involves computation of compression ratios using encoding techniques to convert the significant coefficients into bitstreams. Finally, the compression vector set is normalized and fed to a machine learning classifier that identifies seizure activity from normal, seizure free signals. The study utilizes the standard database for epilepsy as provided by the University of Bonn in order to validate the results against prior benchmarks. The proposed methodology with arithmetic encoding algorithm achieved 100% accuracy and the classification results vary from 30.6% to 100% respectively in case of Huffman encoding. Hence, a computer aided diagnostic (CAD) technique employing DWT along with arithmetic encoding and machine learning algorithms would form a robust diagnostic system in early-stage epilepsy diagnosis.

**Keywords**— Arithmetic encoding, Computer aided diagnostic (CAD), Discrete wavelet transform (DWT), Electroencephalogram (EEG), Epilepsy, Huffman encoding

## I. INTRODUCTION

Epilepsy is a chronic non-communicable brain disorder generally characterized by recurrent seizures and sometimes accompanied by the loss of consciousness [1]. Recurrent seizures occur due to abnormality in the neuronal activities of the brain and can affect mood, sensing and/or movement in the human body [2]. According to statistics reported by the World Health Organization, the disease currently affects about 50 million people worldwide and an estimate of 5 million people are diagnosed with epilepsy each year. The number of people affected with epilepsy is significantly higher in developing countries with an estimate of 139 per 1000 cases

reported annually [3]. The disease is diagnosed by the visual inspection of an EEG while treatment usually involves administration of anti-seizure medication along with surgery in some instances.

The electroencephalogram (EEG) reflects the electrophysiological conditions of the brain at a given instant of time and thus can be used to detect epileptic seizure activity [4]. The low cost of an EEG analysis makes it a widely adopted tool for clinical diagnosis involving brain activity. As EEGs can contain various pathological as well as physiological information, they are utilized globally to detect and evaluate the treatment and progress of epileptic patients. For the purpose for epileptic seizure studies, EEG signals are evaluated for three types of activity: (i) normal EEG activity recorded at eyes open and eye closed intervals of healthy subjects; (ii) seizure free/inter-ictal activity containing small spikes and/or subclinical seizures occurring between two epileptic episodes; (iii) and seizure/ictal EEG activity represented by sudden spikes in the EEG.

Principally, EEG recordings collected from subjects can be quite lengthy and hence contain a significant amount of data. A visual inspection of EEG recordings thus can be slow and time consuming. Moreover, it is also prone to manual errors and is dependent on inter-observer variability [5]. To combat this limitation, several automated techniques have been reported to help in computer aided diagnosis of epileptic seizures. Computer aided diagnostic (CAD) techniques usually involve extraction of frequency and/or time domain features from signals. Machine learning classifiers are further utilized to characterize seizure activity from a set of seizure free or normal EEG recordings. One such method is to count standard deviations between various linear and non-linear techniques and then apply thresholding to differentiate seizure from normal signal [6].

Another widely used method is the application of mathematical transformations. In earlier studies, Fourier transform was used for feature extraction but studies reported decomposing EEG using wavelet transforms is more feasible as it provided reduction in computational time without compromising the accuracy. Previously, many studies have reported the use of wavelet transforms for EEG signals however, most of these studies have utilized either advanced

wavelet transforms or complex wavelet transform along with nonlinear feature extraction techniques and classifiers [4], [7]. These methods usually require additional processing time due to the heavy load and are not feasible for 1D signal decomposition. Additionally, there were few studies reported that classified normal and seizure EEG activity but the studies reporting classification of inter-ictal seizure free intervals have been limited. The use of machine learning classifiers, such as fuzzy classifiers, Gaussian mixture model, support vector machine (SVM), artificial neural networks (ANN), k-nearest neighbour (k-NN) and Naïve Bayes classifier have also been reported.

It has been shown previously that frequency as well as time-based methods prove to be useful for extracting features in EEG signals, and the attributes obtained can be utilized further as classifier inputs. In 2014, a wavelet based fuzzy approximate entropy (fApEn) algorithm was proposed [8]. The EEG was decomposed into sub-bands using DWT and fApEn was computed for each sub-band as a means to rate the chaotic behaviour of EEG recordings. Support vector machine (SVM) was used as a classification technique and highest classification accuracy was reported by the authors. Finally, the work presented by Amin et al., [5] first utilized DWT with arithmetic coding to extract features from EEG signals. Four different machine learning classifiers were utilized and the authors reported with 100% classification accuracy.

There is limited research employing the use of arithmetic coding as a feature extraction technique and the utility of Huffman coding has not yet been established. The present study aims to fulfil this gap by laying an analytic comparison between the two methodologies in detecting seizure activity from seizure free EEG recordings and seizure free activity from healthy EEG signals. In this work, discrete wavelet transform was employed on the EEG signals to obtain coefficients for different decomposition levels. In present work, a 10-fold cross validation technique was employed to evaluate classifier performance. Finally, the results of the two encoding methodologies were compared based on several parameters and the superior alternative was stated.

## II. MATERIALS & METHODS

### A. Datasets

The EEG data used throughout this analysis is extracted from a publicly available database at the University of Bonn [9]. The database consists of five sub datasets labelled A, B, C, D and E. Each dataset is 23.6s in length and contains 100 single channel EEG segments with 4097 samples per channel. The scalp EEG recording obtained with eyes open (EO) and eyes closed for five healthy subjects are labelled respectively as sets A and B. Datasets C, D and E are EEG recordings of five epileptic patients with different spatial locations. Dataset C contains EEG readings from the hippocampal formation in the hemisphere opposite to the epileptogenic zone while dataset D contains recordings from the epileptogenic zone. Both datasets C and D were recorded during seizure-free intervals. Dataset E contains recordings of the hippocampal focus at the time of seizure. The data was obtained using 12-bit analogue-to-digital converters operating with a sampling frequency of 173.61 Hz. All five data sets were collected using 128 channels and an average common reference. Channels depicting pathological activities were neglected while computing the reference. Further, artifacts involving eye movement characteristics were removed from data sets A and

B. Finally, a bandpass filter with cut off frequencies 0.55 to 40 Hz was applied to filter any noise artifacts.

### B. Discrete Wavelet Transform (DWT)

The DWT of a given signal is determined by the inner product of the signal  $[x(t)]$  and the scaled version of the wavelet function  $[\Psi_{a,b}(t)]$  as given in Eq. 2. DWT is represented mathematically as given in Eq. (1):

$$W_{\Psi}X(a, b) = \langle x, \Psi_{a,b} \rangle \quad (1)$$

$$\Psi_{a,b} = |\alpha|^{-\frac{1}{2}} \Psi\left(\frac{t-b}{a}\right) \quad (2)$$

In Eq. (1), 'a' and 'b' are defined as scale and translation parameters respectively and represented as  $[a, b \in \mathbb{R}]$ . The scaling parameter 'a' compresses or dilates the wavelet function while translation parameter 'b' alters its location. Correlation of signal  $[x(t)]$  with scaled version of wavelet function  $[\Psi_{a,b}(t)]$  yields high or low frequency components respectively. The wavelet transform as identified at discrete scales ( $a_j=2^j$ ) and locations ( $b_{j,k}=2^j k$ ) is defined as the DWT. When applied to a signal, DWT decomposes it into detail and approximation coefficients at distinct scales. Information contained by low frequency components are provided by higher scales and lower scales contain information for high frequency components. The number of levels of decomposition required depends on the relation between the sampling frequency and the bandwidth of the signal [10].

### C. Huffman Coding

The Huffman coding algorithm is a systematic data compression technique that yields prefix-free instantaneous codes. The resulting code-word is uniquely decodable and can be mapped to individual source symbol. Each symbol is arranged in descending order of probabilities and considered as a unique node. The last two probabilities, i.e., two nodes with the least probabilities are joined to construct a new node and this process is repeated till a single node with probability 1 is obtained. Binary codes 0 and 1 are assigned to each node starting from the last to the first and the final code-word is given by the sequence of 0's and 1's. An order reversal during the generation of the code-word generally causes the final result to be non-instantaneous.

### D. Arithmetic Coding

Arithmetic coding is another data compression algorithm where each symbol of the source data reduces the size of the interval in accordance with its occurrence probability. The source data is mapped to produce non-block codes in a way that allows reconstruction of original data from the code generated. Hence, each symbol and its code cannot be compared individually as unlike Huffman coding, this method does not execute individual conversion of source data into the code-word. Arithmetic coding is performed on the sequence as a whole and a single arithmetic code-word is generated to represent an interval of real numbers between 0 and 1. An increase in the number of source symbols results in smaller intervals that require additional bits to denote them [11].

## III. METHODOLOGY

As outlined in the block diagram of the proposed CAD technique as represented in Fig. 1, the following four steps majorly constitute the proposed CAD technique.

### 1. Decomposition of EEG Signals

- Wavelet decomposition via the discrete wavelet transform
  - Thresholding coefficients to maintain reconstructed signal energy ~99% when rounded to the nearest integer
2. Feature Extraction
    - Encoding significant DWT coefficients using arithmetic coding
    - Encoding significant DWT coefficients using Huffman coding
    - Computation of compression features (compression ratios) for each coding technique
    - Normalization of extracted features
  3. Feature Classification
    - Classifying extracted features using linear and non-linear machine learning classifiers
    - Evaluation of classifier performance using k-fold cross validation
  4. Comparison of Encoding Techniques
    - Computation of accuracy, sensitivity and specificity for each method
    - Determining relative efficiency based on computational speed and results of EEG classification.

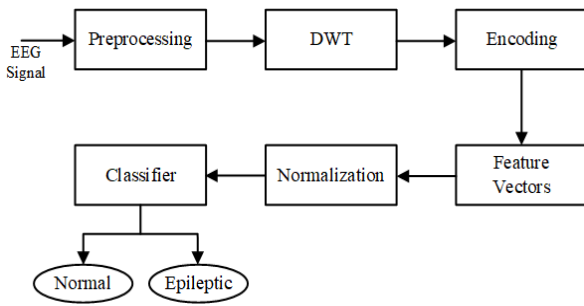


Fig. 1: Block diagram of major steps for epileptic seizure classification as outlined in the proposed CAD technique.

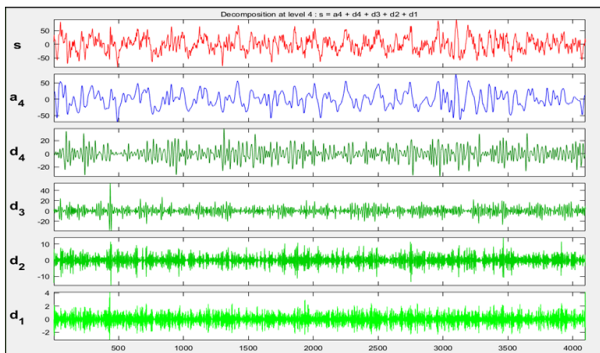


Fig. 2: Example of DWT decomposition of an epileptic EEG signal at level 4. The signal  $s$  is decomposed into detail coefficients  $d_1, d_2, d_3,$  and  $d_4$  and approximation coefficient  $a_4$ .

Initially, DWT is used with Daubechies wavelet to decompose the EEG signal  $x[n]$  into level 4 detail and approximation coefficients as shown in Fig. 2. The db4 wavelet was selected in this analysis as its orthogonal property and efficient filter implementation make it the best fit for decomposing EEG signals, especially when epileptic EEGs are considered. DWT utilizes a series of high pass and low pass filters  $g(n)$  and  $h(n)$  respectively. High pass filter

$g(n)$  represents the mother wavelet while its mirror versions are represented by the low pass filter  $h(n)$ . The cutoff frequencies for the filters are set to approximately one-fourth of the sampling frequencies used for EEG signals. Approximation ( $A_i$ ) and detail ( $D_i$ ) coefficients are obtained in the first level of DWT decomposition, when input signals are filtered through  $h(n)$  and  $g(n)$  respectively. Mathematically,  $A_i$  and  $D_i$  at the  $i^{\text{th}}$  level are defined as follows:

$$A_i = \frac{1}{\sqrt{M}} \sum x(n) \cdot \varphi_{j,k(n)} \quad (3)$$

where:  $\varphi_{j,k(n)} = 2^{-\frac{j}{2}} h(2^{-j}n - k)$  represents the scaling function, and

$$D_i = \frac{1}{\sqrt{M}} \sum x(n) \cdot \Psi_{j,k(n)} \quad (4)$$

where:  $\Psi_{j,k(n)} = 2^{-\frac{j}{2}} g(2^{-j}n - k)$  represents the wavelet function. Here,  $n = 0, 1, 2, \dots, M-1$ ;  $j = 0, 1, 2, \dots, J-1$ ;  $k = 0, 1, 2, \dots, 2^j-1$ ;  $J = 4$ ; where  $M$  denotes the length of the discrete EEG signal  $x[n]$ .

DWT coefficients  $D_{jk}$  are given by detailed coefficients  $D_{i=0,1,2,3}$  and approximation coefficient  $A_{i=3}$  condensed by neglecting non-significant coefficients when a specific threshold value ( $\alpha$ ) is implemented.  $D_{jk}$  can then be defined as follows:

$$\widehat{D}_{jk} = \begin{cases} D_{jk}, & |D_{jk}| \geq \alpha \\ 0, & |D_{jk}| < \alpha \end{cases} \quad (5)$$

The implementation of a threshold value ensures that the reconstructed EEG signal retains significant energy, i.e., energy retained >99%.

$$\text{energy}(E) = \frac{100 \times \|X'\|_2^2}{\|X\|_2^2} > 99\% \quad (6)$$

where ( $X'$ ) represents the reconstructed signal and the original signal is denoted by ( $X$ ).

The calculation of the threshold ( $\alpha$ ) is done by considering the last level of detail coefficients vector for the computation of standard deviation of noise [12] and defined as:

$$\alpha = \hat{\sigma} \sqrt{2 \log N} \quad (7)$$

where ' $N$ ' is the number of wavelet coefficients present in the last level of the detailed coefficients.

As per the median absolute deviation,

$$\hat{\sigma} = \frac{\text{median}(|-D_{jk}|)}{0.6745} \quad (8)$$

The denominator represents the scaling factor, which is dependent on the distribution of coefficients  $\sim D_{jk}$ , and equals 0.6745 in case of normally distributed data.  $\sim D_{jk}$  represents the wavelet coefficients present in the last level of the detailed coefficients.

The threshold value is altered if the energy criterion stated in Eq. (6) is not satisfied. This ensures that the quality of the reconstructed signal is retained even after eliminating non-significant coefficients. The DWT coefficients are subjected to thresholding criteria and rounded up to the nearest integers to be further denoted as  $\overline{D}_{jk}$ .

The second step in the proposed methodology includes encoding of the rounded DWT coefficients  $\overline{D}_{jk}$  into bitstreams by applying the two encoding techniques, i.e. arithmetic coding and Huffman coding. Arithmetic coding can be used in two ways to compress data, static and adaptive. In this study, static arithmetic coding has been used since the number

of wavelet coefficients and their occurrence probability is known. The size of the DWT coefficients is reduced which leads to EEG signal compression. The output bits' stream of the encoders gives the compression future computed as follows:

$$CR = \frac{\text{Original Signal Size } (X)}{\text{Compressed Signal Size } (X'')} \quad (9)$$

Next, the extracted CR features are normalized to zero mean and unit variance as follows:

$$\hat{x}_i = \frac{x_i - x_m}{\sigma} \quad (10)$$

Where  $i = 1, 2, \dots, L$  and  $L$  denotes the number of instances in a particular feature  $x$ .  $x_m$  and  $\sigma$  represent mean and standard deviation of  $x_i$  respectively and  $\hat{x}_i$  is the normalized value of the feature vector. These normalized features are then fed as inputs to the machine learning classifiers. In order to illustrate the effectiveness of the proposed methodology, three machine learning classifiers are employed. Support vector machine (SVM) with radial basis function (RBF) is used as a non-linear classifier while K-nearest neighbor (k-NN) and kernel Naïve Bayes are employed as linear classifiers. Finally, a 10-fold cross validation method is employed to validate classifier performances.

Performance of each classifier was determined using accuracy, sensitivity and specificity as [13]:

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \times 100\% \quad (11)$$

$$\text{Sensitivity} = \frac{TP}{TP+FN} \times 100\% \quad (12)$$

$$\text{Specificity} = \frac{TN}{TN+FP} \times 100\% \quad (13)$$

Where  $TP$ - True Positive,  $TN$ -True Negative,  $FP$ -False Positive,  $FN$ -False Negative.

Averages of accuracy, sensitivity and specificity along with time computation characteristics are computed and estimated results are compared for accuracy of the coding algorithms.

#### A. Cross Validation of Classifiers

Cross validation is essentially a resampling procedure employed for the evaluation of machine learning samples with a limited data set. Commonly, k-fold cross validation (CV) is used for evaluating the efficiency of different classification algorithms [14]. The method is widely used to claim superiority of one classification method over the word. The parameter k determines the number of disjoint folds a dataset is split into. At each fold, 1 set is reserved as test while k-1 folds are used as training sets for the classifier. This is repeated k times and each k-fold is used exactly once as a test set. A major advantage of the k-fold CV technique is that each sampled in the dataset is utilized for training, while each fold is used for validation once. By increasing the number of folds, i.e., the value of k, variance in resulting estimates can be reduced. Assuming a rational value for k (e.g., 5-fold or 10-fold), computational costs can be kept to a minimum.

## IV. RESULTS & DISCUSSION

In this section the proposed CAD technique, followed by the results of k-fold cross validation have been described. The results of the classifier have been compared with reference to two feature extraction algorithms Huffman coding and arithmetic coding and three machine learning techniques such as SVM, k-NN and Naïve Bayes. As outlined in TABLE I,

five classification cases were considered based on the five EEG sub datasets. Case selection was performed to smooth comparison with subject to previous studies.

TABLE I. CASES FOR EXPERIMENTAL CLASSIFICATION

Case	Datasets Used	Specification
A-E	A and E	Normal and ictal (seizure) segments
AB-E	AB and E	Normal and ictal (seizure) segments
C-E	C and E	Inter-ictal (seizure free) and ictal (seizure) segments
CD-E	CD and E	Inter-ictal (seizure free) and ictal (seizure) segments
ABCD-E	ABCD and E	Normal, inter-ictal (seizure free) and ictal (seizure) segments

100 EEG segments were used from each of the five datasets A, B, C, D and E. To obtain wavelet coefficients, 4th level DWT data decomposition was performed and non-significant coefficients were neglected. To ensure quality retention of reconstructed signals, only signals meeting the energy criteria >99% were considered from each of the five datasets (A-E). TABLE II shows us the quality of retained signals from each dataset and TABLE III represents the number of signals retained. This was done to establish a standard balanced criterion for neglecting certain wavelet coefficients over all datasets.

TABLE II. RECONSTRUCTED SIGNAL QUALITY FOR EEG SIGNALS FROM ALL DATASETS

Set	Energy%	Standard Deviation (SD)	Mean Square Error (MSE)	Peak Signal-to-Noise Ratio (PSNR) (dB)
A	99.97	0.12	0.01	58.51
B	99.98	0.02	0.00	72.43
C	99.96	0.06	0.00	64.41
D	99.97	0.12	0.01	58.51
E	99.96	0.10	0.01	59.70
AB	99.99	0.08	0.01	61.32
CD	99.96	0.06	0.00	64.82
ABCD	99.97	0.07	0.01	62.55

The retained significant wavelet coefficients were then utilized to compute feature vectors, i.e., compression ratios (CR) via the feature extraction techniques. Three machine learning classifiers SVM, k-NN and Naïve Bayes were trained and tested for all datasets. 10-fold cross validation technique was used to validate the results of the classifiers. Final feature sets were divided into 10 uniform subsets, out of which nine subsets were utilized to train the classifiers while one subset was reserved for testing. Average accuracy, sensitivity and specificity were then computed to determine ultimate classifier performances.

Fig. 3 and Fig. 4 represents the box plot of all datasets to indicate the distribution of feature values obtained in case of arithmetic coding and Huffman coding respectively. The number of correctly classified EEGs with seizure activity by the presented CAD technique were represented as TP; the number of correctly identified normal segments were represented as TN; the number of normal EEG segments

incorrectly classified as seizure were referred to as FP; and the number of seizure segments missing from the classification results were labelled under FN.

TABLE III. NUMBER OF REJECTED SIGNALS FROM EACH DATASET (A-E)

Set	Total Signals	Rejected Signals	Accepted Signals
A	100	6	94
B	100	6	94
C	100	13	87
D	100	20	80
E	100	51	49
Total Accepted Signals:			404

TABLE IV. COMPARISON OF CLASSIFICATION ACCURACIES FOR ALL DATASETS BY ARITHMETIC AND HUFFMAN CODING ALGORITHMS

Set	Accuracy (%)					
	<i>SVM with RBF</i>		<i>k-NN (k=5)</i>		<i>Kernel Naïve Bayes</i>	
	A.C.	H.C.	A.C.	H.C.	A.C.	H.C.
A-E	100	98	100	98	100	51
AB-E	100	100	100	100	100	79.6
C-E	100	98	100	98	100	30.6
CD-E	100	100	100	100	100	42.9
ABCD-E	100	100	100	100	100	85.7

<sup>a</sup>. A.C. - Arithmetic Coding, H.C. - Huffman Coding

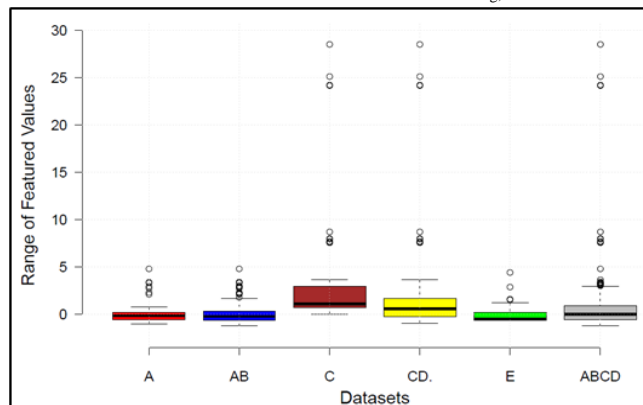


Fig. 3: Box Plot of feature values when Arithmetic Coding is used for feature extraction

The CAD technique proposed in the current study achieved perfect classification results (100% accuracy) across all three machine learning classifiers for cases A-E, AB-E, C-E, CD-E and ABCD-E when arithmetic coding was employed for feature extraction. The accuracy results with Huffman coding stretched across a wider range. In case of SVM fine Gaussian classifier and weighted k-NN method, Huffman algorithm achieved perfect classification (100% accuracy) for datasets AB-E, CD-E and ABCD-E, while accuracy declined to 98% for datasets A-E and C-E. Accuracy results declined as low as 30.6% for dataset C-E with case A-E at 51%, AB-E at 79.6%, CD-E at 42.9%, and finally ABCD-E at 85.7% when Naïve Bayes algorithm was used in conjunction with Huffman coding. Additionally, both arithmetic coding and Huffman coding methods achieved

100% sensitivity and specificity for all three classifiers across all the datasets. However, the sensitivity for the dataset AB-E, C-E and CD-E is observed to be deviated with 1 and 2% in Huffman algorithm. Similarly, the specificity is 99% for AB-E in Kernel Naïve Bayes with Huffman algorithm.

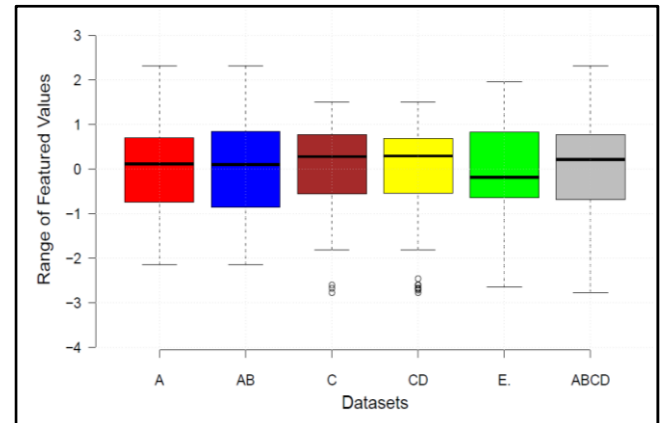


Fig. 4: Box Plot of feature values when Huffman Coding is used for feature extraction

The proposed approach is based on the applicability of DWT for the decomposition of non-stationary signals. Features are extracted from significant DWT coefficients with the application of Huffman coding and arithmetic coding techniques and classification is performed using linear and non-linear classifiers. The accuracy of the two coding techniques is evaluated using five different cases for classification. A number of prior studies have reported DWT to be the most appropriate method to study non stationary rapid events in EEG signals, especially when epileptic data is considered [4]. However, there is limited literature involving feature extraction with static arithmetic coding and to our knowledge, the use of Huffman coding with DWT has not yet been studied. This study provides a comparison between the two techniques and determines the superior alternative to obtain perfect classification results. In TABLE IV, the accuracy of the arithmetic coding and Huffman coding for three different classifiers are listed and results of the present study are reported in comparison to previously studied automated seizure detection techniques in TABLE V.

In order to compare the computational efficiency of the two coding alternatives, time computations were calculated for each of the five classification cases. Feature extraction from a segment of 23.5 seconds required 0.0102 seconds in case of Huffman coding and 0.0181 seconds for arithmetic coding which is significantly improved than the time-frequency analysis methods proposed in previous studies. The classification algorithms SVM, k-NN and Naïve Bayes reported 6.5~9.95 s, 0.05~0.08 s and 0.06~1.02 s of testing time respectively. Thus, for the accurate detection of a seizure from say, a 24 hours continuous recording of an epileptic patient, the proposed CAD method will require 28 minutes 15 seconds via Huffman coding and 28 minutes 29 seconds via arithmetic coding. Thus, the CAD approach presented in this study is exceptionally faster than analysis methods reported and empirical mode decomposition (EMD) algorithms proposed earlier [15]. Further, the proposed method is also slightly faster than a similar DWT and arithmetic coding approach presented by [5] which required 30 minutes of processing time for a 24 hours continuous recording. Hence, the CAD technique presented in this study is exceptionally



efficient in terms of computational speeds and thus can be implemented as a diagnostic tool in clinical environments.

TABLE V. ACCURACY COMPARISON OF PRESENT WORK WITH PRIOR STUDIES

Author(s)	Classifier(s)	Dataset	Accuracy%
Kumar et al.,[8]	ANN, SVM	A-E, B-E, C-E, D-E, ABCD-E	100,92.5,100,95, 94,99
Acharya et al.,[16]	Gaussian Mixture Model (GMM)	Normal, inter-ictal and ictal	99
Song et al.,[17]	Extreme learning machine (ELM)	A-E	94.20
Lee et al.,[18]	Weighted fuzzy membership function (NEWFM)	A-E	98.17
Fu et al.,[19]	SVM (RBF)	A-E	99.12
Dhiman et al.,[20]	SVM	A-E	100
Present Study	SVM (RBF), k-NN and Naïve Bayes	A-E, AB-E, C-E, CD-E, ABCD-E	(98, 98, 51), (100, 100, 79.6), (98, 98, 30.6), (100, 100, 42.9), (100, 100, 85.7)
	SVM (RBF), k-NN and Naïve Bayes	A-E, AB-E, C-E, CD-E, ABCD-E	100,100,100,100 ,100

However, this study has some limitations. Firstly, the results vouch for a relatively small data set. To obtain similar accuracy and precision in larger datasets, higher values of K can be considered for k-fold cross validation technique. Additionally, feature selection methods like Principal Component Analysis (PCA), Fisher's Discriminant Ratio can be used to optimize the extracted features. Furthermore, the algorithm can be optimized for real time clinical testing in order to additionally reduce processing time.

## V. CONCLUSION

The current study presents a robust CAD algorithm to detect the presence of seizure activity in seizure free EEG recordings and healthy EEG signals. The study employs DWT as a preliminary wavelet transformation step followed by arithmetic encoding and Huffman encoding individually to extract significant EEG features. Three machine learning classifiers SVM, weighted k-NN and kernel Naïve Bayes algorithm are employed to segregate seizure (ictal) signals from its seizure-free (inter-ictal) and normal EEG counterparts. All three classifiers achieve perfect accuracy when arithmetic coding is used in the feature extraction step. Hence this proposed CAD technique could be useful for diagnostic purposes.

## REFERENCES

- [1] C. E. Stafstrom and L. Carmant, "Seizures and Epilepsy : An Overview for Neuroscientists," pp. 1–18, 2015.
- [2] L. Chisci et al., "Real-time epileptic seizure prediction using AR models and support vector machines," *IEEE Trans. Biomed. Eng.*, vol. 57, no. 5, pp. 1124–1132, 2010, doi: 10.1109/TBME.2009.2038990.
- [3] C. Officers, "EPILEPSY : A manual for Medical and Clinical Officers in Africa and Clinical Officers."
- [4] U. R. Acharya, S. Vinitha Sree, G. Swapna, R. J. Martis, and J. S. Suri, "Automated EEG analysis of epilepsy: A review," *Knowledge-Based Syst.*, vol. 45, pp. 147–165, Jun. 2013, doi: 10.1016/j.knosys.2013.02.014.
- [5] H. U. Amin, M. Z. Yusoff, and R. F. Ahmad, "A novel approach based on wavelet analysis and arithmetic coding for automated detection and diagnosis of epileptic seizure in EEG signals using machine learning techniques," *Biomed. Signal Process. Control*, vol. 56, p. 101707, 2020, doi: https://doi.org/10.1016/j.bspc.2019.101707.
- [6] K. K. Jerger et al., "Early seizure detection.," *J. Clin. Neurophysiol. Off. Publ. Am. Electroencephalogr. Soc.*, vol. 18, no. 3, pp. 259–268, May 2001, doi: 10.1097/00004691-200105000-00005.
- [7] Y. Ech-Choudany, D. Scida, M. Assarar, J. Landré, B. Bellach, and F. Morain-Nicolier, "Dissimilarity-based time–frequency distributions as features for epileptic EEG signal classification," *Biomed. Signal Process. Control*, vol. 64, no. September 2018, p. 102268, 2021, doi: 10.1016/j.bspc.2020.102268.
- [8] Y. Kumar, M. L. Dewal, and R. S. Anand, "Epileptic seizures detection in EEG using DWT-based ApEn and artificial neural network," *Signal, Image Video Process.*, vol. 8, no. 7, pp. 1323–1334, 2014, doi: 10.1007/s11760-012-0362-9.
- [9] R. G. Andrzejak, K. Lehnertz, F. Mormann, C. Rieke, P. David, and C. E. Elger, "Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state," *Phys. Rev. E - Stat. Physics, Plasmas, Fluids, Relat. Interdiscip. Top.*, vol. 64, no. 6, p. 8, 2001, doi: 10.1103/PhysRevE.64.061907.
- [10] A. Grossmann and J. Morlet, "Decomposition of Hardy Functions Into," *SIAM J. Math. Anal.*, vol. 15, no. 4, 1984.
- [11] R. E. W. profile imageRichard E. W. Rafael Corsino González profile imageRafael C. Gonzalez, *Digital Image Processing Using MATLAB: AND " Mathworks, MATLAB Sim SV 07 "*. Prentice Hall PressOne Lake Street Upper Saddle River, NJ, United States, 2007.
- [12] D. L. Donoho and J. M. Johnstone, "Ideal spatial adaptation by wavelet shrinkage," *Biometrika*, vol. 81, no. 3, pp. 425–455, 1994, doi: 10.1093/biomet/81.3.425.
- [13] F. Pereira, T. Mitchell, and M. Botvinick, "Machine learning classifiers and fMRI: A tutorial overview," *Neuroimage*, vol. 45, no. 1, pp. S199–S209, Mar. 2009, doi: 10.1016/j.neuroimage.2008.11.007.
- [14] T.-T. Wong, "Performance evaluation of classification algorithms by k-fold and leave-one-out cross validation," *Pattern Recognit.*, vol. 48, no. 9, pp. 2839–2846, Sep. 2015, doi: 10.1016/j.patcog.2015.03.009.
- [15] S. M. S. Alam and M. I. H. Bhuiyan, "Detection of Seizure and Epilepsy Using Higher Order Statistics in the EMD Domain," *IEEE J. Biomed. Heal. Informatics*, vol. 17, no. 2, pp. 312–318, Mar. 2013, doi: 10.1109/JBHI.2012.2237409.
- [16] U. R. Acharya, F. Molinari, S. V. Sree, S. Chattopadhyay, K. H. Ng, and J. S. Suri, "Automated diagnosis of epileptic EEG using entropies," *Biomed. Signal Process. Control*, vol. 7, no. 4, pp. 401–408, 2012, doi: 10.1016/j.bspc.2011.07.007.
- [17] Y. Song and J. Zhang, "Automatic recognition of epileptic EEG patterns via Extreme Learning Machine and multiresolution feature extraction," *Expert Syst. Appl.*, vol. 40, no. 14, pp. 5477–5489, 2013, doi: 10.1016/j.eswa.2013.04.025.
- [18] S. H. Lee, J. S. Lim, J. K. Kim, J. Yang, and Y. Lee, "Classification of normal and epileptic seizure EEG signals using wavelet transform, phase-space reconstruction, and Euclidean distance," *Comput. Methods Programs Biomed.*, vol. 116, no. 1, pp. 10–25, 2014, doi: 10.1016/j.cmpb.2014.04.012.
- [19] K. Fu, J. Qu, Y. Chai, and Y. Dong, "Classification of seizure based on the time-frequency image of EEG signals using HHT and SVM," *Biomed. Signal Process. Control*, vol. 13, no. 1, pp. 15–22, 2014, doi: 10.1016/j.bspc.2014.03.007.
- [20] R. Dhiman, J. S. Saini, and Priyanka, "Genetic algorithms tuned expert model for detection of epileptic seizures from EEG signatures," *Appl. Soft Comput. J.*, vol. 19, pp. 8–17, 2014, doi: 10.1016/j.asoc.2014.01.029.

# Automated Detection of Brain Abnormality using Deep-Learning-Scheme: A Study

Seifedine Kadry

*Faculty of Applied Computing and  
Technology, Noroff University College,  
Kristiansand, Norway.*  
email: skadry@gmail.com

Yunyoung Nam

*Department of Computer Science and  
Engineering, Soonchunhyang  
University, Asan-si, South Korea.*  
email: ynam@sch.ac.kr

Hafiz Tayyab Rauf

*Centre for Smart Systems, AI and  
Cybersecurity, Staffordshire University,  
Stoke-on-Trent, United Kingdom*  
email: hafiztayyabrauf093@gmail.com

Venkatesan Rajinikanth

*Department of Electronics and  
Instrumentation Engineering,  
St. Joseph's College of Engineering  
Chennai 600119, India.*  
email: v.rajinikanth@ieee.org

Isah A. Lawal

*Faculty of Applied Computing and  
Technology, Noroff University College,  
Kristiansand, Norway*  
email: Isah.Lawal@noroff.no

**Abstract**— Brain is the vital organ in human physiology; which is conscientious for sensory signal handling and judgment making. The irregularity in brain severely influence entire decision making procedure and the unrecognized and untreated defect will lead to various harsh conditions. This research aims to implement pre-trained Deep-Learning-Scheme (DLS) to classify the brain MRI slices using a multi-class classifier. In this research, the brain MRI slices with classes; normal, stroke, Low-Grade-Glioma (LGG) and High-Grade-Glioma (HGG) are considered for the experimental study. In this work every test picture is resized into 224x224x3 pixels and these imagery are then considered to validate the performance of DLS, such as VGG16, VGG19 and ResNet50 using different classifiers. The attained classification accuracy of every DLS with classifiers, SoftMax, SVM-RBF and SVM-Cubic are presented and discussed.

**Keywords**—Brain abnormality, MRI slices, Deep-Learning, Multi-Class classification, Accuracy.

## I. INTRODUCTION

In recent days, due to a variety of physical and environmental conditions, the occurrence rate of infectious and chronic diseases are gradually rising and appropriate detection and treatment implementation is essential to reduce the disease impact [1,2]. In the current era, modern disease detection and treatment implementation methods are existing to reduce the medical stress.

In humans, the abnormality in brain will considerably affect the whole health condition and hence it is considered as one of the chief emergency. From former works, it can be observed that the major brain abnormality includes the Ischemic-Stroke (IS) and the Brain-Tumour (BT) and these abnormalities can be efficiently diagnosed with a chosen bio-signal/bio-imaging scheme. The bio-signal based assessment needs a single/multi channel sensor array to collect the brain signals and later a chosen signal processing method needs to be implemented to detect the abnormality in the recorded EEG. Further, the bio-signal is very complex to evaluate compared to the bio-imaging approach. Hence, bio-images are widely considered to evaluate the brain abnormality compared to the EEG [3-5].

Radiology based imaging procedure is widely recommended to screen the brain abnormality and compared to the computed-tomography, the MRI is a widely used radiological imaging scheme due to its multimodality and 3D

nature [6]. Hence, in most of the brain abnormality detection process, MRI slices of a chosen imaging modality are considered to achieve a better detection.

The prior research related to the brain abnormality detection can be found in [7,8]. Most of the earlier works implemented a segmentation task or the classification approach to identify the irregularity using the 2D slices of brain MRI. The work related to the heuristic algorithm based thresholding and segmentation can be found in [9-11]. The machine-learning scheme based classification is available in [12] and the Deep-Learning-Scheme (DLS) with a two-class and multi-class classification is presented in [13-15]. Every earlier technique, shows that a number of detection procedures are proposed and executed separately for the IS and BT and this work plans to execute a integrated detection practice with a chosen DLS.

For the BT, a specified grading scheme is available to categorize the images based on the dimension and orientation of the BT. A group of BT known as the glioma is a harsh condition in which the tumour originates in glial-cells of the brain/spine. Further, the glioma is one of the common abnormalities and 80% of the glioma is malignant. Based on the size and its location, it can be classified as the Low-Grade/High-Grade (LGG/HGG) class.

In this work, the necessary images for the experimental study is attained from the ISLES2015 [16] and BRATS2015 [17] dataset. Every image is available in the form of 3D reconstructed picture and hence a 3D to 2D conversion is executed using ITK-Snap [18,19] and after the extraction, a resizing process is executed to get a picture with a dimension of 224x224x3 pixels.

Here, Flair modality MRI pictures are used for the assessment. The main advantage of the considered datasets is, the examination pictures are obtainable without the skull fragment and hence the skull-stripping practice is not required for these images. Further, the visibility of the IS and BT sections are good in the chosen pictures and hence better detection accuracy can be achieved.

In this research, a DLS with a multi-class classification is implemented to classify the considered Flair modality slices into different classes, such as normal, stroke, LGG and HGG. The various initial processing procedures considered in this work involves; collection of the essential images from the chosen database, conversion of 3D to 2D and resizing it

to a chosen dimension based on the DLS and implementing the essential classifier and validation.

Compared to the earlier studies, the key part of the proposed work is the implementation of a multi-class classification using the VGG16, VGG19 and ResNet50. The attained result is then compared and validated with the help of classifiers, such as SoftMax, SVM-RBF and SVM-Cubic.

The enduring sections of this work are implemented as follows; Section 2 and 3 depicts the context and methodology of this work, Section 4 and 5 presents the experimental results and conclusions, respectively.

## II. CONTEXT

In earlier works, large quantity of Machine-Learning (ML) and DLS schemes is planned to detect the disease condition in 2D brain MRI of chosen modality [10-15].

The brain images of ISLES [16] and BRATS [17] are widely used to calculate the performance of the developed automated schemes because of its multi modality nature. Further, these dataset images are having clinical grade images without the skull section; which minimizes the computational complexity during the assessment.

Table I summarises the chosen earlier works implemented using the ISLES and BRATS datasets.

TABLE I. SUMMARY OF EARLIER WORKS RELATED TO THE BRAIN ABNORMALITY DETECTION

Reference	Methodology implemented	Dataset
Revanth et al. [5]	Instigation of stroke wound is presented using CT/MR images	ISLES2015
Rajinikanth and Satapathy [9]	Addition of thresholding and segmentation is employed to examine the stroke lesion	ISLES2015
Rajinikanth et al. [10]	Segmentation of the stroke lesion is achieved using the watershed algorithm	ISLES2015
Dey et al. [12]	A machine learning based brain abnormality assessment is discussed	ISLES2015 and BRATS2015
Rajinikanth et al. [14]	DLS based brain image classification with two-class classifier is discussed	BRATS2015
Lin et al. [20]	Hybrid image processing scheme is employed to mine the abnormal brain section	ISLES2015 and BRATS2015
Kadry et al. [21]	Mining of tumour section from brain MRI is presented	BRATS2015

In literature, a quantity of earlier works are implemented a chosen approach to examine the abnormality from the brain MRI slices and the essential details can be found in [1-5]. In this work, the objective is to implement an automated detection system to classify the considered brain MRI slices into different classes using a chosen DLS with various classifiers.

## III. METHODOLOGY

The main contribution of this work is to implement a DLS with a multi-class classifier for the brain MRI classification. The various stages existing in the proposed scheme can be found in Figure 1.

Primarily, the vital imagery is accumulated from ISLES2015 and BRATS2015 dataset and the collected images are then converted into 2D slices via the ITK-Snap. The extorted 2D slices are then resized into 224x224x3 pixel dimension picture and then these pictures are used to train and test the DLS classifier considered in the proposed work. Finally, a multi-class classifier is employed to classify the 2D brain MRI into Normal/Stroke/LGG/HGG class with improved accuracy.

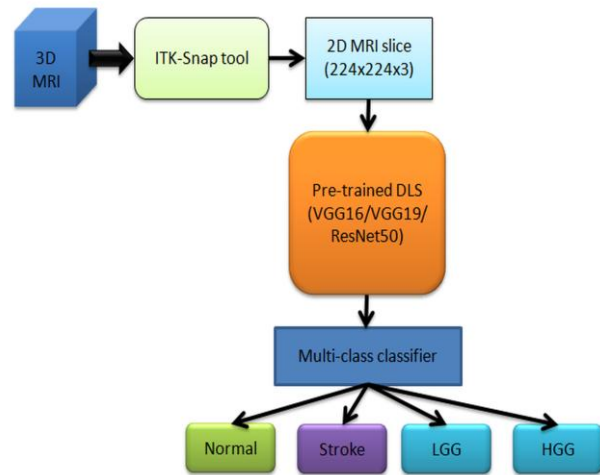


Fig.1. Brain MRI assessment with DLS and multi-class classification

### A. Image database

The performance of the developed disease detection system is to be confirmed with clinical grade images. The availability of clinical level MRI is very less and hence, in this work the brain images of ISLES2015 and BRATS2015 are considered for the assessment. The information concerning the imagery used in this research can be found in Table II. This table shows that the total number of images considered is 500 per category in which 400 images (80%) are considered for training and 100 pictures (20%) are considered for the validation. All these images are resized into 224x224x3 pixels before evaluation. The sample test images adopted in this work can be found in Figure 2. The main merit of this dataset is, every picture is free from the skull section and helps to achieve better classification accuracy.

TABLE II. THE TEST IMAGE INFORMATION FOR THE CONSIDERED IMAGE DATASET

Image	Database	Dimension	Training images	Validation images
Normal	Mixed	224x224x3	500	100
Stroke	ISLES2015	224x224x3	500	100
LGG	BRATS2015	224x224x3	500	100
HGG		224x224x3	500	100

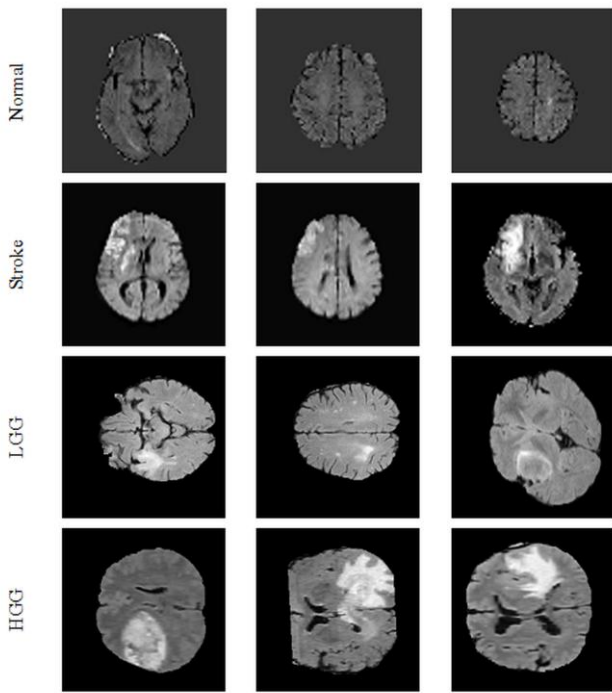


Fig.2. Trial brain MRI considered in this research

**B. Pre-trained Deep-Learning-Scheme**

The earlier works on the DLS confirms that the pre-trained schemes are very efficient in attaining the better result on a class of image datasets [13-15]. The chief benefit of the pre-trained DLS is; its performance is already confirmed on a range of datasets and only the initial training is essential to implement the existing scheme for the considered dataset. In this scheme, pre-trained DLS such as VGG16, VGG19 and ResNet50 are employed and the essential information on these networks can be found in [21].

The proposed DLS is implemented using MATLAB software and the result attained is presented in the form of confusion matrix. Initially, SoftMax classifier is used to categorize the assessment imagery into four classes and after getting the essential result, other techniques, such as SVM-RBF and SVM-Cubic are implemented instead the SoftMax.

**C. Performance evaluation and Validation**

Merit of this idea is estimated based on the result attained in the confusion matrix. In this work, the overall categorization accurateness is computed to confirm the performance of the chosen DLS using the implemented classifier. In this work, 100 numbers of test images is every group is considered for the assessment and the finest outcome of the five-fold cross validation is considered for performance confirmation.

**IV. RESULT AND DISCUSSION**

This division of research shows the outcome achieved using a workstation; Intel i5 2.5GHz processor with 16GB RAM and 2GB VRAM set with MATLAB®. In this work, initially the proposed experiment s implemented using VGG16 and later, it is executed with VGG19 and ResNet50.

The VGG16 with the SoftMax unit is initially considered to classify the brain MRI slices using the SoftMax classifier and the attained results, such as convergence and confusion

matrix are depicted in Figure 3 and 4 respectively. The perfectly trained DLS helps to achieve better classification accuracy as depicted in Fig 3 and 4. Form figure 3, it can be observed that the accuracy and loss function achieved with respect to number of operations are better with the VGG16 and this scheme helped to achieve an overall accuracy of 92% with the SoftMax classifier. Similar procedure is repeated with classifiers, such as SVM-RBF and SVM-Cubic and the corresponding result is depicted in Figure 5 and 6 respectively. From these figures, it can be noted that the overall accuracy in SVM-RBF (95.8%) is better compared to SVM-Cubic (91%) for the multi-class classification and this procedure is then repeated using VGG19 and ResNet50.

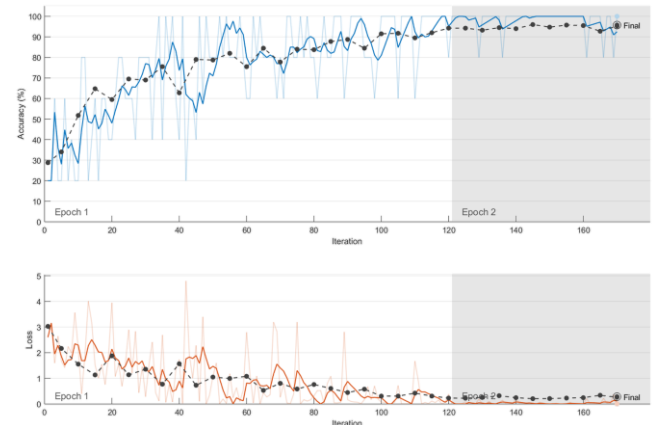


Fig.3. Training and testing traces achieved for VGG16 with SoftMax

		Confusion Matrix					
		HGG	LGG	Normal	Stroke		
Output Class	HGG	93 23.3%	2 0.5%	0 0.0%	0 0.0%	97.9%	2.1%
	LGG	7 1.8%	92 23.0%	13 3.3%	0 0.0%	82.1%	17.9%
	Normal	0 0.0%	0 0.0%	83 20.8%	0 0.0%	100%	0.0%
	Stroke	0 0.0%	6 1.5%	4 1.0%	100 25.0%	90.9%	9.1%
		93.0%	92.0%	83.0%	100%	92.0%	8.0%
		7.0%	8.0%	17.0%	0.0%		
		HGG	LGG	Normal	Stroke	Target Class	

Fig.4. Multi-class confusion matrix result achieved for VGG16 with SoftMax

The result attained with the implemented DLS for ISLES2015 and BRATS2015 are depicted in Table III. The computed accuracy for every approach is then compared against each other and the graphical representation of this comparison is presented in Figure 7. From Table III and Fig 7; it can be noted that the VGG19 with SVM-Cubic classifier helped to achieve a better overall accuracy (96%) for the considered dataset.

**Confusion Matrix**

Output Class	HGG	97 24.3%	2 0.5%	0 0.0%	0 0.0%	98.0% 2.0%
	LGG	3 0.8%	89 22.3%	3 0.8%	0 0.0%	93.7% 6.3%
	Normal	0 0.0%	2 0.5%	97 24.3%	0 0.0%	98.0% 2.0%
	Stroke	0 0.0%	7 1.8%	0 0.0%	100 25.0%	93.5% 6.5%
		97.0% 3.0%	89.0% 11.0%	97.0% 3.0%	100% 0.0%	95.8% 4.2%
	Target Class					
	HGG	LGG	Normal	Stroke		

Fig.5. Multi-class confusion matrix attained with SVM-RBF

**Confusion Matrix**

Output Class	HGG	93 23.3%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
	LGG	3 0.8%	74 18.5%	0 0.0%	0 0.0%	96.1% 3.9%
	Normal	4 1.0%	25 6.3%	97 24.3%	0 0.0%	77.0% 23.0%
	Stroke	0 0.0%	1 0.3%	3 0.8%	100 25.0%	96.2% 3.8%
		93.0% 7.0%	74.0% 26.0%	97.0% 3.0%	100% 0.0%	91.0% 9.0%
	Target Class					
	HGG	LGG	Normal	Stroke		

Fig.6. Multi-class confusion matrix attained with SVM-Cubic

TABLE III. OVERALL ACCURACY ATTAINED WITH CHOSEN CLASSIFIERS

Architecture	Classifier	Accuracy (%)
VGG16	SoftMax	92.0
	SVM-RBF	95.8
	SVM-Cubic	91.0
VGG19	SoftMax	93.6
	SVM-RBF	94.8
	SVM-Cubic	96.0
ResNet50	SoftMax	90.2
	SVM-RBF	92.6
	SVM-Cubic	91.8

In this work, the detection of the brain abnormality is presented using the Flair modality MRI slices and in future, the other modalities, like T1, T2 and T1C are to be considered for the assesment. Moreover, the merit of DLS can be improved by employing other classifiers.

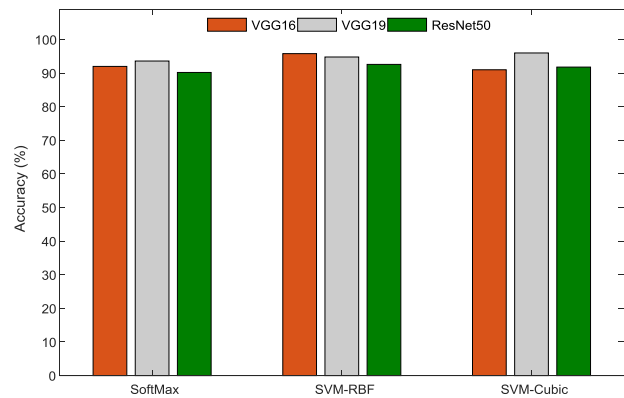


Fig.7. Performance evaluation of chosen classifiers

V. CONCLUSION

The investigation of various brain abnormalities using brain MRI slices of ISLES2015 and BRATS2015 is presented in this work. This work employs the DLS approaches, such as VGG16, VGG19 and ResNet50 for the experimental investigation. The multi-class categorization is then implemented using the classifiers, such as SoftMax, SVM-RBF and SVM-Cubic and the performance is computed based on the overall accuracy attained with every scheme. The outcome of this research confirmed that the overall accuracy of VGG19 with SVM-Cubic (96%) is better compared to other techniques. This accuracy can be improved in future by uniting the considered deep-features with handcrafted-features.

REFERENCES

- [1] E. Priya, and V. Rajinikanth, Signal and Image Processing Techniques for the Development of Intelligent Healthcare Systems. Springer, 2020.
- [2] V. Rajinikanth, E. Priya, H. Lin, and F. Lin, Hybrid Image Processing Methods for Medical Image Examination. CRC Press, 2021.
- [3] S.L. Fernandes, U. J. Tanik, V. Rajinikanth, and K. A. Karthik, "A reliable framework for accurate brain image examination and treatment planning based on early diagnosis support for clinicians," Neural Computing and Applications, vol. 32, no. 20, pp. 15897-15908, 2020. <https://doi.org/10.1007/s00521-019-04369-5>.
- [4] S. Kadry, V. Rajinikanth, N.S.M. Raja, D. J. Hemanth, N. M. S. Hannon, and A. N. J. Raj. "Evaluation of brain tumor using brain MRI with modified-moth-flame algorithm and Kapur's thresholding: a study," Evolutionary Intelligence, pp. 1-11, 2021. <https://doi.org/10.1007/s12065-020-00539-w>.
- [5] K. Revanth, N.S.M. Raja, and V. Rajinikanth, "Computational investigation of stroke lesion segmentation from Flair/DW modality MRI," In 2018 Fourth International Conference on Biosignals, Images and Instrumentation (ICBSII), pp. 206-212. IEEE, 2018. DOI: 10.1109/ICBSII.2018.8524617.
- [6] O. Maier, M. Wilms, and H. Handels, "Image features for brain lesion segmentation using random forests," In BrainLes 2015, pp. 119-130. Springer, Cham, 2015. [https://doi.org/10.1007/978-3-319-30858-6\\_11](https://doi.org/10.1007/978-3-319-30858-6_11).
- [7] M. A. Khan, I. Ashraf, M. Alhaisoni et al., "Multimodal brain tumor classification using deep learning and robust feature selection: A machine learning application for radiologists," Diagnostics, vol. 10, no. 8, pp.565, 2020. <https://doi.org/10.3390/diagnostics10080565>.
- [8] Q. Ke, J. Zhang, W. Wei, R. Damaševičius, and M. Woźniak, "Adaptive independent subspace analysis of brain magnetic resonance imaging data," IEEE Access, vol. 7, pp.12252-12261, 2019. DOI: 10.1109/ACCESS.2019.2893496.
- [9] V. Rajinikanth, and S. C. Satapathy, "Segmentation of ischemic stroke lesion in brain MRI based on social group optimization and Fuzzy-Tsallis entropy," Arabian Journal for Science and Engineering,

- vol. 43, no. 8, pp. 4365-4378, 2018. <https://doi.org/10.1007/s13369-017-3053-6>.
- [10] V. Rajinikanth, K. P. Thanaraj, S. C. Satapathy, S. L. Fernandes, and N. Dey, "Shannon's entropy and watershed algorithm based technique to inspect ischemic stroke wound," *Smart intelligent computing and applications*, vol.105, pp. 23-31. 2019. [https://doi.org/10.1007/978-981-13-1927-3\\_3](https://doi.org/10.1007/978-981-13-1927-3_3).
- [11] V. Rajinikanth, S. C. Satapathy, N. Dey, and H. Lin, "Evaluation of ischemic stroke region from CT/MR images using hybrid image processing techniques," In *Intelligent multidimensional data and image processing*, pp. 194-219. IGI Global, 2018. DOI: 10.4018/978-1-5225-5246-8.ch007.
- [12] N. Dey et al., "Social-Group-Optimization based tumor evaluation tool for clinical brain MRI of Flair/diffusion-weighted modality," *Biocybernetics and Biomedical Engineering*, vol. 39, no. 3, pp. 843-856, 2019. <https://doi.org/10.1016/j.bbe.2019.07.005>.
- [13] M.A. Khan, I. Ashraf, M. Alhaisoni, R. Damaševičius, R. Scherer, A. Rehman, and S.A.C. Bukhari, "Multimodal brain tumor classification using deep learning and robust feature selection: A machine learning application for radiologists," *Diagnostics*, vol. 10, no. 8, pp. 565, 2020.
- [14] V. Rajinikanth, A.N.J. Raj, K.P. Thanaraj, and G.R. Naik., "A customized VGG19 network with concatenation of deep and handcrafted features for brain tumor detection," *Applied Sciences* 10, no. 10, pp.3429, 2020. <https://doi.org/10.3390/app10103429>.
- [15] M. A. Khan, I. Ashraf, M. Alhaisoni et al., "Multimodal brain tumor classification using deep learning and robust feature selection: A machine learning application for radiologists," *Diagnostics*, vol. 10, no. 8, pp.565, 2020. <https://doi.org/10.3390/diagnostics10080565>.
- [16] O. Maier, B.H. Menze, J. V. D. Gablentz et al., "ISLES 2015-A public evaluation benchmark for ischemic stroke lesion segmentation from multispectral MRI," *Medical image analysis*, vol 35, pp.250-269, 2017. <https://doi.org/10.1016/j.media.2016.07.009>.
- [17] B.H. Menze, A. Jakab, S. Bauer, J. Kalpathy-Cramer, K. Farahani, J. Kirby, Y. Burren et al., "The multimodal brain tumor image segmentation benchmark (BRATS)," *IEEE transactions on medical imaging*, vol. 34, no. 10, pp.1993-2024, 2014. DOI: 10.1109/TMI.2014.2377694.
- [18] P.A. Yushkevich, J. Piven, H. C. Hazlett, R. G. Smith, S. Ho, J. C. Gee, and G. Gerig, "User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability," *Neuroimage*, vol.31, no.3, pp. 1116-1128, 2006. <https://doi.org/10.1016/j.neuroimage.2006.01.015>.
- [19] <http://www.itksnap.org/pmwiki/pmwiki.php?n=Main.Publications>
- [20] D. Lin, V. Rajinikanth, and H. Lin., "Hybrid Image Processing-Based Examination of 2D Brain MRI Slices to Detect Brain Tumor/Stroke Section: A Study," In *Signal and Image Processing Techniques for the Development of Intelligent Healthcare Systems*, pp. 29-49. Springer, Singapore, 2021. [https://doi.org/10.1007/978-981-15-6141-2\\_2](https://doi.org/10.1007/978-981-15-6141-2_2).
- [21] S. Kadry, V. Rajinikanth, N.S.M. Raja, D. J. Hemanth, N. M. S. Hannon, and A. N. J. Raj. "Evaluation of brain tumor using brain MRI with modified-moth-flame algorithm and Kapur's thresholding: a study," *Evolutionary Intelligence*, pp. 1-11, 2021. <https://doi.org/10.1007/s12065-020-00539-w>.
- [22] S. Ahuja et al., "Deep transfer learning-based automated detection of COVID-19 from lung CT scan slices," *Applied Intelligence*, vol. 51, no. 1, pp.571-585, 2021.

**2021 IEEE Seventh  
International Conference on  
Bio signals, Images and  
Instrumentation  
ICBSII 2021**

**SESSION II  
RESEARCH  
PAPERS**

# Performance Metrics Analysis of Adaptive threshold Empirical Mode Decomposition Denoising method for suppression of noise in Lung sounds

B.Sangeetha  
Research Scholar  
Department of ICE  
National Institute of Technology  
Trichy-620015, Tamilnadu, India  
[sangeetha27may@gmail.com](mailto:sangeetha27may@gmail.com)

R.Periyasamy  
Assistant Professor  
Department of ICE  
National Institute of Technology  
Trichy-620015, Tamilnadu, India  
[periyasamyr@nitt.edu](mailto:periyasamyr@nitt.edu)

**Abstract**— Chest auscultation is a non-invasive and widely used tool for the detection of pulmonary disease. Though it is a powerful method of pulmonary function test, it relies challenges in different issues that bound its diagnostic ability. Mainly in motion artifacts, other environmental noise, and heart sound interference, hence contaminates the original content of the lung sound. This study proposes an adaptive threshold Empirical Mode Decomposition (aEMD) technique to denoise the lung sound signals thus improving the signal quality for the detection of respiratory pathologies. The proposed scheme is refined to offset maximum noise suppression against preserving the reliability of the lung sound signal. With this algorithm, the maximum Signal to Noise ratio (SNR) of 43.89 dB, Correlation Coefficient of 0.995, and minimum Root Mean Square Error (RMSE) of 0.00122 is achieved. Further, it can be implemented in real-time to assist medical experts doctors to make clear interpretations of the respiratory sound-related disorder.

**Keyword**—Asthma, Wheeze sound, lung sound denoising, Empirical Mode Decomposition

## I. INTRODUCTION

Asthma is a heterogeneous disease, usually considered chronic airway inflammation. Wheezing, shortness of breath, chest tightness, and cough that fluctuate over time and in intensity are the major symptoms of asthma [1]. The major symptom of asthma is wheezing so it becomes essential to analyze wheeze sound for asthma diagnosis [2]. Auscultation and acoustic analysis of lung sounds are the most important method used to diagnose lung abnormalities. The rapid fluctuations of gas pressure or the oscillations of solid tissues are the main cause for lung sounds [3]. Lung sounds that are recorded during auscultation are contaminated with heart sound (HS), background noise, thoracic tissues effects, and measurement noise [4]. Although modern stethoscopes can help in hearing the sounds more clearly, heart sounds still impede the respiratory sounds that hinder the potential of respiratory sound analysis [5]. Various signal denoising algorithms have been proposed to improve the performance of the diagnostic tool. M. T. Pourazad, et.al

first locate HS segments by employing multi-resolution decomposition of the wavelet transform coefficients, and calculate the missing data via a 2D interpolation in the time-frequency (TF) domain [6]. Thato Tsalaiile and Saeid Sanei used Adaptive Line Enhancement adaptive filters with Least-Mean Square (LMS) adaptation algorithm for the segregation of heart sound from lung sound [7]. Foad Ghaderi et.al localizes heart sound segments using Singular Spectrum Analysis to construct a data-driven filter in denoising the lung sounds [8]. Noman Qaid Al-Naggar used Least Mean Square (LMS) Algorithm for the Adaptive noise cancellation technique in removing of heart sound from lung sounds [9]. Nersisson and Mathew M. Noel proposed a superior adaptive noise cancellation technique based on a hybrid Nelder-Mead (H-NM) optimization algorithm to segregate heart sounds from lung sounds [10]. Vedansh Thakkar explained Noise Cancellation using Least Mean Square Algorithm [11]. Meng Fei, et.al proposed a wavelet de-noising algorithm for lung sound signals analysis. The locations of lung sound parts are observed by the mean values of autocorrelation coefficients. The noises between lung sound parts are filtered [12]. Nishi Shahnaj Haider et.al proposed Savitzky-Golay Filter for Denoising Lung Sound [13]. Recently Fei Meng et.al shows that noise suppression in respiratory sound signals was performed by modified de-noising algorithms [14]. Further Ashok Mondal et.al evaluated the empirical mode decomposition technique to localize the heart sound from lung sound. Also, Fast Fourier Transform (FFT) and inverse FFT are used to predict the missing sample and reconstruction of the original signal [15]. Even though different methods are employed in denoising lung sounds still there is a limitation in estimating the performance of a filter in signal-to-noise ratio, Correlation coefficient, and Root mean square error. This study proposes an adaptive-based Empirical Mode Decomposition technique to suppress unwanted noise by maintaining the reliability of lung sound. Further, the quality of the processed lung sound signal was validated quantitatively by considering the performance metrics parameter.



Section II describes Materials and Methods. Evaluation results are described in section III. Section IV and V describes the discussion and conclusion.

## II. MATERIALS AND METHOD

### Data Acquisition

The significant part of this research is the systematic collection of wheeze sound samples with reliable ground truth. The database includes recorded wheeze sound from repositories like the RALE database. The lung sounds are recorded from elderly patients from age 45 to age 60 that include mild wheeze, expiratory wheeze, monophonic, and polyphonic wheeze sound signals. The signals are recorded in 32 bits with a sampling frequency of 44100 Hz. The lung sounds are acquired at the right anterior region and the left posterior region of the chest. Lung auscultation is done in a sitting position using a single-channel Littmann 3200 Electronic Stethoscope in .wav file format. As the wheezing sound is contaminated with heart sound (HS), environmental sounds, and measurement noise, denoising is a preliminary step in pre-processing.

### Implementation of adaptive threshold Empirical Mode Decomposition

The adaptive threshold Empirical mode decomposition is a data-driven algorithm established by N.Huang et.al in the year 1996. This is the most appropriate adaptive algorithm offered for the time-frequency analysis of a signal to overcome the limitations of conventional techniques like Discrete Fourier Transform (DFT), Short-Time Fourier Transform (STFT), and Wavelet Transform (WT). The EMD technique decomposes the complex data into a small number of mono-component signals called intrinsic mode functions (IMF) based on statistical properties of the signal. The length of the IMFs and the original signal is the same since the signal is decomposed in the time domain, it is known that other unwanted components emphasize the existence of interference in the signal. Thus adaptive threshold EMD technique finds a better solution to remove the interference from the signal. An IMF is chosen in two ways

1. The number of zero-crossings must be equal to the number of extrema or the difference should be one in the whole data set.
2. The envelope is obtained on local minima and the local maxima.

The flow diagram of adaptive threshold EMD algorithm is shown in Fig.1.

The extraction of IMFs through the shifting process from lung sound signals  $x(n)$  is summarized as follows: Decompose the data set into IMFs  $x_n(t)$  and residue  $r(t)$ , so that signal is represented as

$$x(t) = \sum_n x_n(t) + r(t)$$

1. Identifying all local extrema in the lung sound signal.
2. The upper and lower envelope is obtained by connecting all the local maxima and local minima using the cubic spline interpolation technique.
3. The local mean values of the upper and lower envelope ' $m_1(t)$ ' is calculated.
4. Calculate  $h_1(t)$  by finding the difference between data and  $m_1(t)$ .
 
$$h_1(t) = x(t) - m_1(t)$$
5. Check whether  $h_1(t)$  is an IMF or not, if not assume ' $x(t)$ ' as ' $h_1(t)$ ', and repeat the steps (1) to (4).
6. If  $h_1(t)$  is an IMF1 store  $h_1(t)$  as  $c_1(t)$ , Subtract the IMF1 from input signal  $x(t)$  and obtain the residue  $r_1(t)$ .
 
$$r_1(t) = x(t) - c_1(t)$$
7. If  $r_1(t)$  is not the residue repeat the above all steps by treating  $r_1(t)$  as  $x(t)$  until it becomes the residue.
8. The above process ends after not more than one extremum in the residue  $r(t)$ .

Thus we can extract the original signal using adaptive threshold EMD by decomposing the original signal into IMFs and its residue.

### Performance metrics

#### Signal to Noise Ratio

The signal-to-noise ratio (SNR) in decibels of a signal, is computed by taking the ratio of summation of squared magnitude of the filtered signal to the summation of squared magnitude of the noise signal or it is the ratio of the power of the original signal to the power of a noisy signal and is given by

$$SNR \text{ in dB} = \frac{\text{Power of Signal}}{\text{Power of Noise}}$$

#### Correlation Coefficient

The Correlation Coefficient (COR) is the measure of the similarity between two signals. The ideal value of COR is 1, this specifies that the reconstructed signal exactly replicates the original signal. The mathematical expression of COR is given by

$$COR = \frac{\sum_{i=0}^n (x_i - \bar{x})(D_i - \bar{D})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})(D_i - \bar{D})}}$$

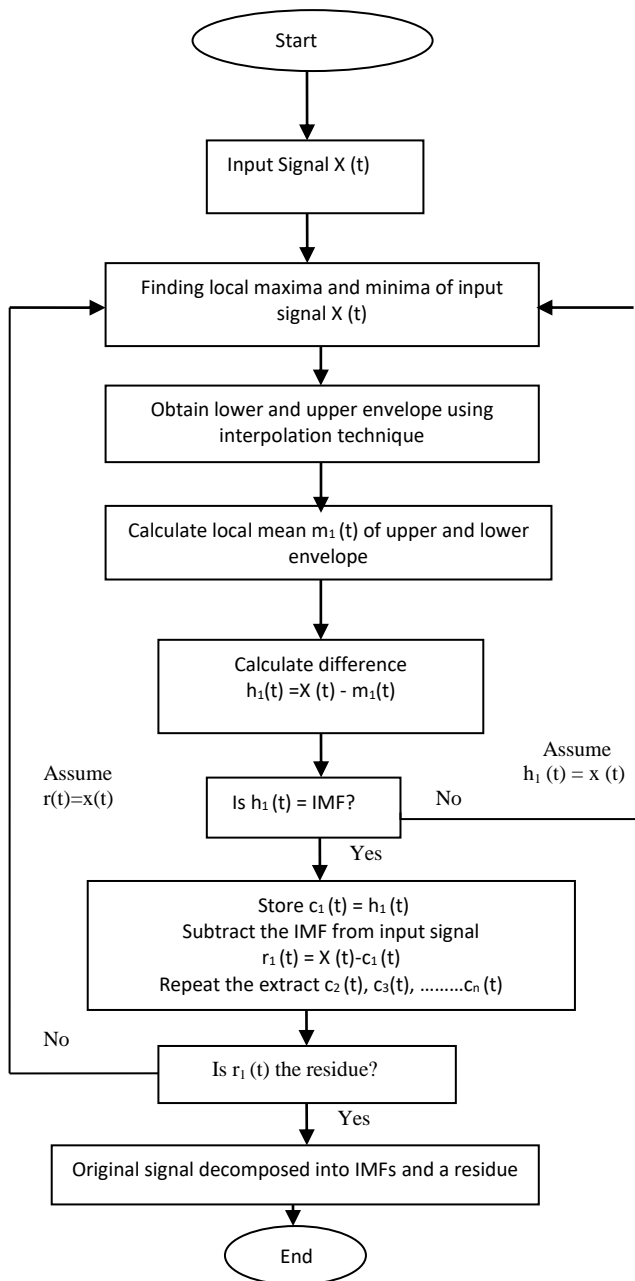


Fig. 1. Flow diagram of the adaptive threshold EMD algorithm

#### Root Mean Square Error (RMSE)

The root-mean-square error (RMSE) is used to measure the differences between the estimated values and the observed values. It is the square root of the average of squared errors. Here  $\hat{x}_i$  are estimated values,  $x_i$  is observed values and  $n$  is the number of observations.

$$RMSE = \sqrt{\frac{\sum_{i=0}^n (\hat{x}_i - x_i)^2}{n}}$$

### III. RESULTS

The validation of the proposed method requires a balance of the suppression of unwanted noise while maintaining the reliability of the lung sound signal. Even though signal processing techniques are used to enhance signal quality but they should not result in an altered noisy signal. To properly estimate the performance of the proposed algorithm, objective signal analyses are done. Several performance metrics and quality measures are projected in literature but it needs the true signal information or statistical estimate [17]. But this is not suitable for existing applications, since lung sound signals have two characteristics like it can be estimated over a population but each subject has unique traits that should be sensibly assessed. Therefore it is vital to suppress the noise without hampering the original content of the lung sound signal. The objective metrics chosen to evaluate the performance of the filter are the Signal to noise (SNR) ratio, Root Mean Square Error (RMSE), and Correlation Coefficient (COR). Denoising algorithm is implemented on different types of wheeze sounds like monophonic, polyphonic, mild expiratory wheeze to calculate the performance metrics.

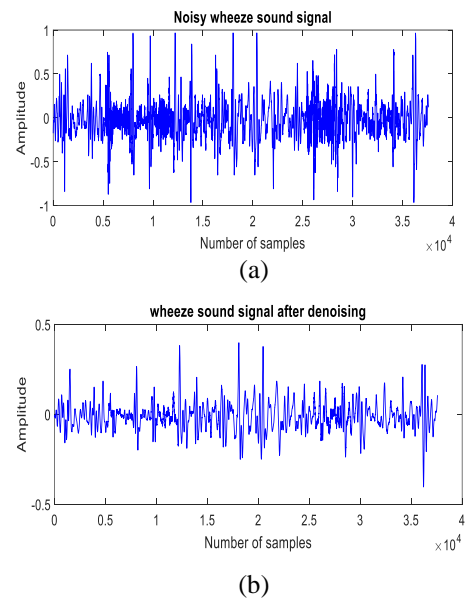


Fig. 2. (a) Noisy wheeze sound signal  
(b) Denoised Wheeze sound signal

Fig.2 (a) shows a noisy wheeze sound signal, (b) shows a denoised wheeze sound signal. It is clear from Fig. 2 (a) that noisy signals of higher amplitude are eliminated thus preserving the original content of wheeze sound. Fig. 3 (a) shows the single-sided amplitude spectrum of noisy wheeze sound signal and Fig. 3 (b) shows the single-sided amplitude spectrum of denoised wheeze sound signal. It is observed from the above figures that the amplitude of the noisy signal has been suppressed after denoising. Similarly higher frequency components have

been suppressed to the desired frequency. Table 1 shows the SNR, RMSE, and COR values are obtained from the proposed denoising algorithm. The SNR value varies from a minimum value of 19.97dB to a maximum of 43.89 dB. It fluctuates between these two values because of contamination in the wheeze sound signal at the time of lung auscultation. RMSE values are maximum for subject 5 of 0.09921 and subject 6 has minimum RMSE values is 0.0012. COR value for subject 6 is 0.995 that specifies that the filtered signal replicates the original wheeze sound signal. The highest SNR value of 43.89 dB, least RMSE of 0.0012, and high COR value of 0.9921 signifies that the proposed algorithm effectively denoises the lung sounds.

TABLE 1. PERFORMANCE METRICS OF PROPOSED ALGORITHM FOR LIMITED SUBJECT DATA

Lung Sound signal	Signal to Noise Ratio in dB (SNR)	Root Mean Square Error (RMSE)	Correlation Coefficient (COR)
LS1	23.22	0.06082	0.867
LS2	19.97	0.08335	0.963
LS3	40.39	0.00394	0.986
LS4	37.87	0.00576	0.955
LS5	36.85	0.09921	0.988
LS6	43.89	0.00122	0.995

Fig. 4 (a) Noisy wheeze sound signal (b) Plot of IMF1-IMF15. The adaptive threshold EMD algorithm generates IMF signal from IMF1 to IMF 15, in this output IMF 15 is considered as residue as there is no more than one extreme. Based on the IMF, the threshold value of the noisy signal is identified and the segments from lung sounds are removed without distorting the original content of a signal. According to aEMD technique the IMFs are generated from 1 to 15. In IMF1 and IMF2 there is more number of local maxima and minima value. In IMF3 and IMF4 the extremes are certainly reduced. From IMF5 to IMF15 the extremes are fully reduced and in final IMF15 there is negative envelope which implies no IMF can be generated. IMF15 is considered as residue and it is a stopping criterion for an EMD technique.

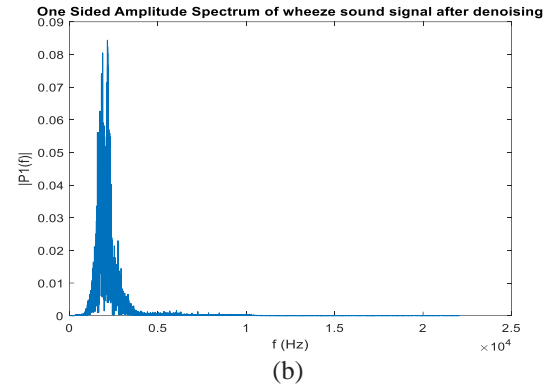
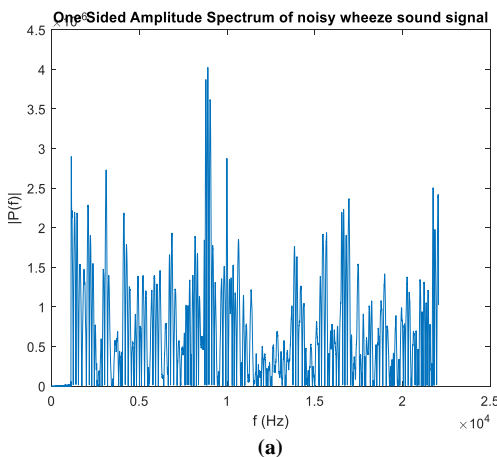


Fig. 3. (a) single-sided amplitude spectrum of noisy wheeze a sound signal (b) single-sided amplitude spectrum of noise free wheeze sound signal

#### IV. DISCUSSION

In this study, the suppression of interference from lung sound signals is evaluated by considering an adaptive empirical mode decomposition technique. In this technique, the original signal is split into intrinsic mode functions, and subsequently filtering operation is performed. Therefore the proposed technique is suitable for denoising the non-stationary and non-linear and signal. Many denoising techniques are available for noise cancellation from lung sound but still, there are certain drawbacks in the existing algorithm. Ziva novice et.al [18] proposed a quasi-periodic signal modeling for denoising lung sound but they attain a signal to deviation ratio of 13.22 dB. This was comparatively low as compared to the proposed algorithm. Meng Fei, et.al [12] proposed a wavelet-based de-noising algorithm for lung sound signals in which the original sounds are decomposed into seven layers. However, the challenging part of this algorithm is selecting the threshold values based on the optimization rule. Achmad Rizal et.al. proposed an EMD and Hjorth descriptor for denoising and classification [19]. Even though the author used many data the performance metrics is not done. Many denoising techniques have been used for processing lung sound signals. Even though they produce promising results, still there is a deficiency in terms of performance metrics, algorithm complexity, choosing several data sets for pre-processing, and dependency of former knowledge of wheeze sound signal. Table 2 shows the assessment of the proposed method with previous work. The proposed algorithm gives the solution for the above limitations and generates the maximum Signal to Noise Ratio (SNR) of 43.89 dB, minimum Root Mean Square Error (RMSE) of 0.00122, and correlation coefficient (COR) of 0.995. It is observed that the proposed algorithm is performing effectively in denoising the lung sound signals. The limitations of this proposed work is that the identification of threshold for is done manually and in future this can be done automatically.

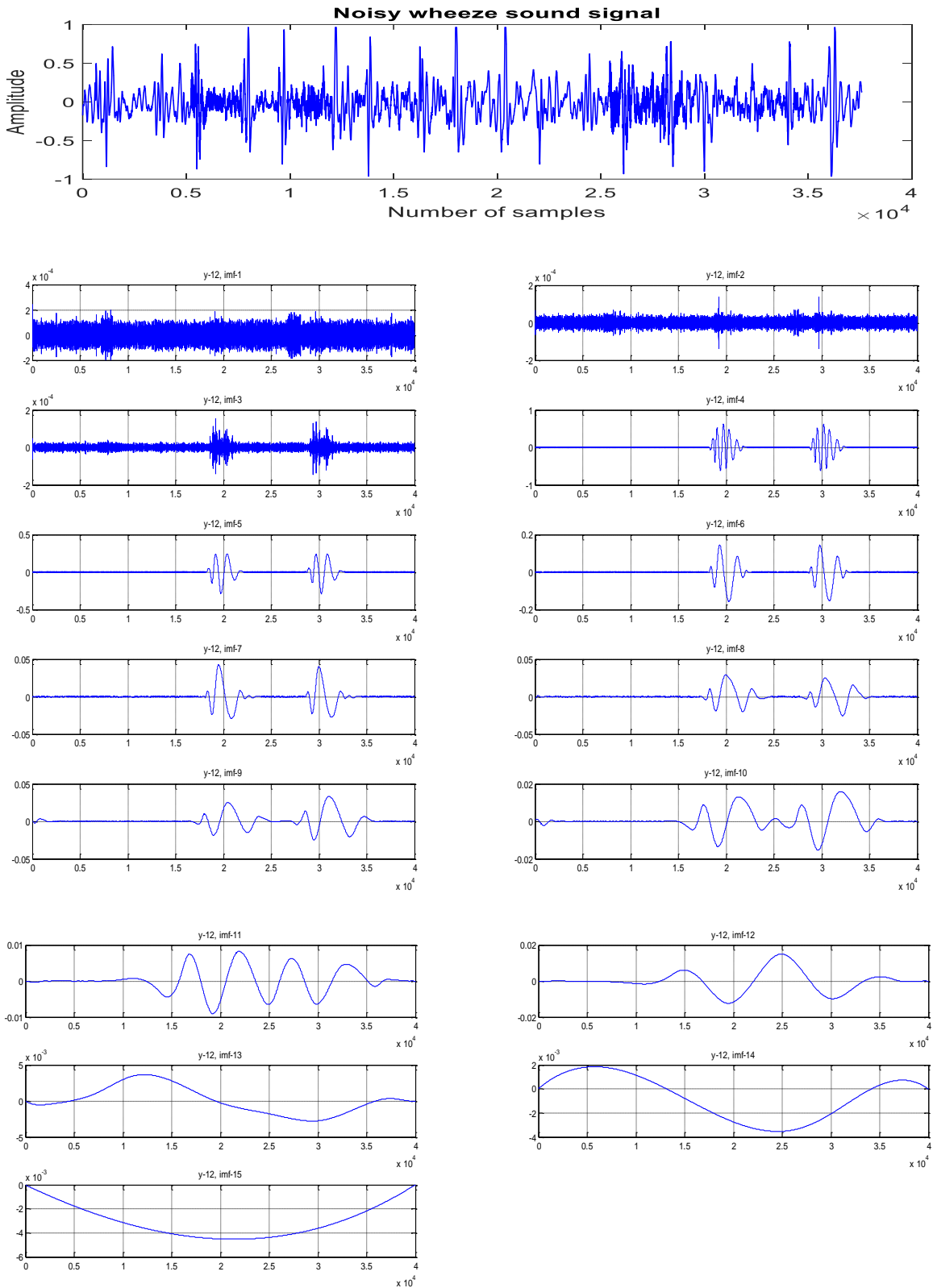


Fig. 4 (a) Noisy wheeze sound signal (b) Plot of IMF 1 to IMF 15

TABLE 2. COMPARISON WITH PREVIOUS WORK

Author & year	Number of data used	Filtering technique	Limitations	Evaluation metrics
Zivanovic et.al & 2013	20	Quasi periodic signal modeling	Restricted to low flow rate	SNR=13.22 dB
Achmad Rizal, 2017	81	EMD & Hjorth descriptor	Mode Mixing problem	-
Meng Fei et.al & 2019	11	FIR BPF, wavelet filter and adaptive filter	Selection of threshold is tedious	No metrics
Proposed method	30	Adaptive EMD based method	-	SNR=43.89 dB RMSE=0.00122 COR=0.995

## V.CONCLUSION

This work presents an exclusive approach in filtering wheeze sound signals using adaptive threshold empirical mode decomposition (EMD). This method suppresses the unwanted noise however preserving the reliability of lung sound signal. The proposed algorithm generates the maximum Signal to Noise Ratio (SNR) of 43.89 dB, minimum Root Mean Square Error (RMSE) of 0.00122, and correlation coefficient (COR) of 0.995. By integrating the proposed method with lung sound acquisition protocol, this method can be extended for real-time implementation. The integrity of the proposed algorithm is adaptive and can be used in denoising non-linear and non-stationary signals. This study is limited to wheeze sounds alone and promising results can be obtained while extending this algorithm implementation to all other lung sounds for better diagnosis.

## References

1. E. D. Bateman, S.S. Hurd, P.J. Barnes, J. Bousquet, J.M. Drazen, M. FitzGerald, P. Gibson, K. Ohta, P. O'Byrne, S.E. Pedersen, E. Pizzichini, S.D. Sullivan, S.E. Wenzel, and H.J. Zar., "Global strategy for asthma management and prevention: GINA executive summary," *Eur. Respiration J.*, vol. 31, pp. 143–78, Jan. 2008.
2. N. Meslier, G. Charbonneau, and J. L. Racineux, "Wheezes," *Eur. Respiration J.*, vol. 8, pp. 1942–1948, Nov. 1995.
3. Jessica Shank Coviello, Wolters Kluwer Lippincott Williams & Wilkins, "Auscultation Skills Breath & Heart Sounds" Fifth Edition
4. Ruban Nersisson and Mathew M. Noel, "Heart sound and lung sound separation algorithms: a review", *Journal of Medical Engineering & Technology*, 2016
5. Raymond LH Murphy Jr MD MPH, "In Defense of the Stethoscope", *RESPIRATORY CARE*, MARCH 2008, VOL 53 NO 3
6. M. T. Pourazad, Z. Moussavi, G. Thomas, "Heart sound cancellation from lung sound recordings using time-frequency filtering", *Med Biol Eng Comput* (2006) 44: 216–225
7. Thato Tsalaile and Saeid Sanei , "Separation of Heart Sound Signal from Lung Sound Signal by Adaptive Line Enhancement", *15th European Signal Processing Conference (EUSIPCO 2007)*, Poznan, Poland, September 3-7, 2007
8. Foad Ghaderi, Hamid R. Mohseni, and Saeid Sanei, "Localizing Heart Sounds in Respiratory Signals Using Singular Spectrum Analysis", *IEEE Transactions on Biomedical Engineering*, VOL. 58, NO. 12, DECEMBER 2011
9. Noman Qaid Al-Naggar, " A new method of lung sounds filtering using modulated least mean square—Adaptive noise cancellation", *Journal of Biomedical Science and Engineering*, 2013, 6, 869-876
10. Ruban Nersisson , Mathew M. Noel, "Hybrid Nelder-Mead search based optimal Least Mean Square algorithms for heart and lung sound separation" *Engineering Science and Technology, an International Journal* 20 (2017) 1054–1065
11. Vedansh Thakkar, "Noise Cancellation using Least Mean Square Algorithm", *IOSR Journal of Electronics and Communication Engineering (IOSR-JECE)*
12. Meng Fei, Cai Maolin, Wang Yixuan , Yang Liman, Shi Yan, Shen Dongkai, "A new type of wavelet de-noising algorithm for lung sound signals" 2018 *IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*
13. Nishi Shahnaj Haider, R. Periyasamy, Deepak Joshi, Bikesh Singh, "Savitzky-Golay Filter for Denoising Lung Sound", *Brazilian Archives of Biology and Technology* · November 2018
14. Fei Meng, Yixuan Wang, Yan Shi1, Hongmei Zhao, "A kind of integrated serial algorithms for noise reduction and characteristics expanding in respiratory sound", *International Journal of Biological Sciences* 2019; 15(9): 1921-1932
15. Ashok Mondal , Poulami Banerjee, Ajay Somkuwar "Enhancement of lung sounds based on empirical mode decomposition and Fourier transform algorithm" *Computer Methods and Programs in Biomedicine* 139 2017, pp. 119–136
16. N.E. Huang, Z. Shen, S.R. Long, M.C. Wu, H.H. Shih, Q. Zheng, et al., "The empirical mode decomposition and Hilbert spectrum for nonlinear and non-stationary time series analysis", *Proceedings of Royal Society of London A* 454(1998)903–995.
17. P. T. C. A. Quackenbush and R. Schuyler Barnwell, *Objective Measures of Speech Quality*, 1st ed. Englewood Cliffs, NJ, USA: Prentice-Hall, 1998.
18. M. Zivanovic, M. Gonzalez-Izal, "Quasi-periodic modeling for heart sound localization and suppression in lung sounds", *Biomed. Signal Process. Control* 8 (6) (2013) 586–595.
19. Achmad Rizal, Risanuri Hidayat, and Hanung Adi Nugroho, "Lung Sound Classification Using Empirical Mode Decomposition and the Hjorth Descriptor", *American Journal of Applied Sciences* , 2017, 14 (1): 166.173

# Study of P wave indices in sinus rhythm and tachycardia

Arya Bhardwaj, J. Sivaraman\* *Member, IEEE*  
 Bio-signals and Medical Instrumentation Laboratory  
 Department of Biotechnology and Medical Engineering  
 National Institute of Technology Rourkela, Odisha, India  
[219bm1413@nitrkl.ac.in](mailto:219bm1413@nitrkl.ac.in); [jsiva@nitrkl.ac.in](mailto:jsiva@nitrkl.ac.in)

**ABSTRACT - P Wave Indices (PWI) have been proved to be a potential marker in determination of various heart conditions. This study shows the extended implications drawn from these indices for determination of tachycardia and encompassing arrhythmias. ECGs were recorded from 50 Sinus Rhythm (SR) and 30 Sinus Tachycardia (ST) volunteers using EDAN PC ECG machine at the paper speed of 25 mm/s and gain of 10 mm/mV. Higher P wave axis ( $^{\circ}$ ) and amplitude ( $\mu$ V) values were found in ST ( $62 \pm 15.38$ ;  $73 \pm 15.32$ ) than SR ( $49 \pm 17.90$ ;  $63 \pm 7.93$ ) and differed significantly between the groups with  $P < 0.05$ . P wave duration (ms) (SR; ST:  $96 \pm 9.18$ ;  $96 \pm 10.01$ ) and PR Interval (ms) (SR; ST:  $138 \pm 17.60$ ;  $135 \pm 18.25$ ) did not differ significantly with  $P > 0.05$  between the groups. However, P wave Dispersion (PD) were higher in SR ( $20 \pm 11.32$  ms) compared to ST ( $13 \pm 10.01$  ms). In ST condition, higher heart rate lead to an increased mean P wave axis and amplitude compared to SR. Higher PPI variability was observed within the limits in SR. PD for both the groups was within the normal limits. This study highlights the variation in PWI in different heart conditions. Further clinical analysis is needed to study the prevalence of Atrial Fibrillation (AF) in tachycardia conditions.**

**Keywords – Atrial fibrillation, Electrocardiogram, P wave indices, P wave dispersion, Sinus rhythm, Sinus tachycardia**

## I. INTRODUCTION

Electrocardiogram (ECG) is the electrical and muscular activity of heart representing the atrial and ventricular components. Atrial components include atrial depolarization, P wave, signifying the generation of impulse from SA node until both the atria depolarizes. Similarly, ventricular components include ventricle depolarization (QRS complex) and repolarization wave (T wave) representing the depolarization and repolarization of ventricles respectively. ECG provides a wide scope to understand the internal mechanism of heart and diagnose the diseased conditions. Various studies have been carried out on ventricular and atrial components for predetermination of diseases [1-3]. Left atrial abnormalities and Atrial Fibrillation (AF) can be determined by the variabilities observed in the P wave morphology [4-6].

Sinus Tachycardia (ST) is the condition of increased heart rate ( $>100$  bpm) for higher cardiac output. It is the compensatory response of heart governed by the autonomic nervous system to modulate the cardiac electrophysiology and arrhythmia during rest, exercise or stress conditions.

However, the electrocardiographic features i.e. P wave, QRS complex, T wave and their characteristics remain normal. As the sustained ventricular tachycardia leads to the morphological changes in ventricle components such as widening of QRS complex [7] similarly various studies have noted that persistent atrial tachycardia leads to the induction of severe AF [8] and respective changes in P wave morphology [5]. Prospective factor for the conversion of tachycardia to AF is the irregular tachycardia cycle length. Tachycardia leads to atrial dilatation in turn affecting the cardiac membrane potential leading to AF [9]. It is also noted that elimination of atrial tachycardia has decreased the occurrence of sustained AF from paroxysmal AF [10].

Various studies have shown the importance of determining P wave characteristics associated with the prevalence of diseased conditions [11]. P Wave Indices (PWI) represent the atrial depolarization electrophysiology and morphological parameters such as P wave duration, amplitude, area, P wave axis, dispersion, terminal force. PWI have been employed to differentiate the diverse abnormal ECG waveforms from normal.

P Wave Duration (PWD) and P wave Dispersion (PD) are the potential markers for various cardiac disorders [12, 13]. PWD is measured as the duration between the beginning of P wave or first time when wave crosses the isoelectric line ( $P_{on}$ ) until P wave ends ( $P_{off}$ ). PD is the difference between the longest and shortest PWD recorded from multiple different surface ECG leads [13]. Prolonged PWD and PD are the important predictors of AF [3]. Prolongation of PWD  $>120$  ms signifies the occurrence of left atrial abnormalities [14]. A study [15] noted that shorter PWD is also associated with the occurrence of AF. Moreover, a study [16] noted that interatrial block patients having PWD between 110-140 ms attains Sinus Rhythm (SR); however, when it increases beyond 140-200 ms may lead to the development of AF in patients. Thus, longer PWD leads to the increased occurrence of prolonged diseases. Further, [17] noted the normal limits of PD as  $29 \pm 9$  ms, and stated that higher values of  $PD \geq 40$  ms indicates the heterogeneous electrical activity of heart generating the re-entrant pathways in atria which can lead to atrial tachyarrhythmia especially AF. A study [11] stated the significance of  $P_{on}$ ,  $P_{off}$ ,  $P_k$  (P wave peak) for determining the P wave duration, amplitude, and asymmetry. P wave asymmetry computed as ratio between duration of P wave

\*Corresponding author: Dr. J. Sivaraman, Bio-signals and Medical Instrumentation Laboratory, National Institute of Technology Rourkela.

terminal to the initial portion ( $P_{ter}/P_{ini}$ ) which may indicate the affected chamber of atria.  $P_{area}/P_{duration}$  is another PWI stated in [4], where  $P_{area}$  calculated as the product of P peak amplitude \* PWD, which gives the better outlook of the proportion of P wave variability in terms of duration and amplitude for different heart conditions.

P wave axis indicates the electrical mean vector within the atria. Angle of the depolarization electrical vector to the ECG lead I determines the amplitude of the waves. Hence, during any abnormal conditions the deviation of P wave axis occurs from its normal limit ( $0^{\circ}$ - $75^{\circ}$ ) in turn affecting the amplitudes. PR Interval (PRI) is also the potential marker which represents the atrial APD. It includes the atrial depolarization, P wave and PR segment representing the delay at AV node to conduct the electrical impulse further to ventricles with the normal value within 200 ms. In contrary to PWD, short PRI is associated with mortality [18]. They also stated that level of contribution of P wave to PRI determines the prognosis of cardiac diseases.

Therefore, the understanding of deviation from the normal limit values of PWI plays a major role for determination of abnormal cardiac conditions. The normal limits of PWI are influenced by various factors other than pathological conditions such as age and sex. The effect of age may lead to left and right axis shift within the normal limits of PWI. Hence, a diagnostic study inclusive of external and internal parameters increase the probability for clinicians to generate true positive and true negative. In this study, PWI such as P wave axis, amplitude, duration, dispersion, PP Interval (PPI) and PRI were evaluated in SR and ST for determining the variations in atrial features among groups for early detection.

## II. METHODS

### A. Study population

ECGs were recorded from 50 SR (mean  $\pm$  SD;  $35 \pm 18.42$ ) and 30 ST ( $19 \pm 3.93$ ) volunteers. No SR volunteers having cardiovascular co-morbidity were included in this study. ST ECG was recorded from volunteers after physical activity and heart rate  $\geq 100$  bpm were considered in the study.

### B. Data acquisition

EDAN PC-1010 machine was used to record ECG for limb leads I - aVF at the standard paper speed 25 mm/s and voltage gain 10 mm/mV. Further, ECGs were magnified at variable speed and gain for better analysis and minimizing intra and inter observer variability. P wave features were labelled for a single beat SR ECG at paper speed of 200 mm/s and 100 mm/mV.  $P_{on}$  is the point when the P wave crossed the isoelectric line for the first positive deflection,  $P_{off}$  is the point when P wave ends and combines with the isoelectric line,  $P_{amp}$  is the point when P wave reach its maximum amplitude.

### C. Statistical analysis

All the values are given in mean  $\pm$  SD. Kolmogorov Smirnov test was performed to check the normality. Pearson's

correlation, r value, was calculated to check the correlation for data and 2 tailed Student's T-test to determine the significant differences for unpaired data.  $P < 0.05$  is considered as the level of significance. Outlier values for all the data were not considered in the statistical analysis to prevent the error which may have caused due to the intra or inter observer variability.

## III. RESULTS

Fig. 1. represents the single beat ECG of a SR volunteer indicating  $P_{on}$ ,  $P_{off}$ ,  $P_{amp}$ , and  $P_{dur}$  as P wave features.

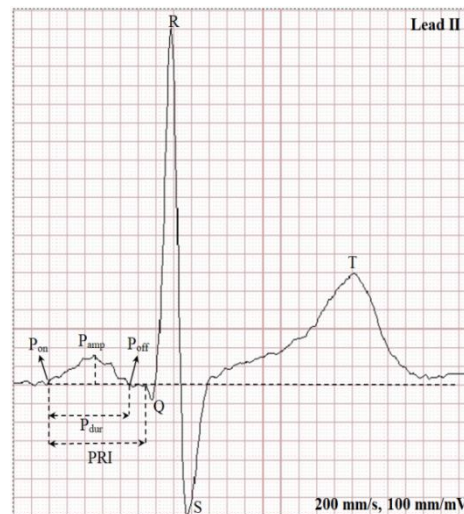


Fig. 1. P wave features in sinus rhythm single beat ECG.

Table I shows the P wave amplitude for each lead and P wave axis in SR and ST groups. Significant differences  $P < 0.05$  were found between the amplitude values of SR and ST volunteers. Maximum and minimum value indicates the range of variability for a lead and higher range was observed for SR volunteers. However, ST volunteers have higher mean amplitude and P wave axis than in SR. Fig. 2. represents the P wave amplitudes for all leads between the SR and ST groups. Higher amplitude values were noted for ST (except lead I) compared to SR.

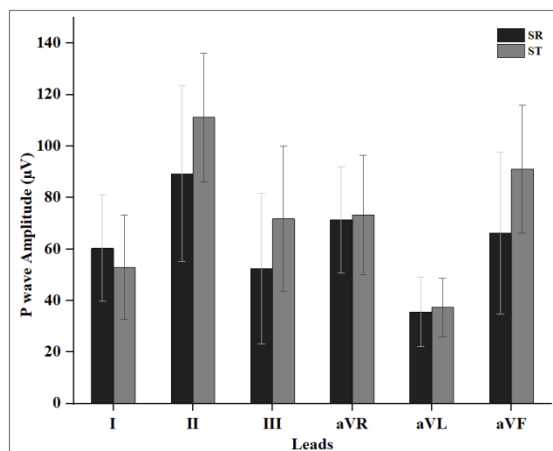


Fig. 2. P wave amplitudes for each lead in SR and ST groups.

TABLE I. P WAVE AMPLITUDE AND AXIS IN SR AND ST GROUPS

Measurements	Limb Leads	SR (n = 50)			ST (n = 25)			P value
		Mean	SD	Min; Max	Mean	SD	Min; Max	
P wave amplitude (μV)	I	60	20.64	23; 125	53	20.22	12;87	< 0.05
	II	89	34.16	17;161	111	25	63;150	< 0.05
	III	52	29.24	15;133	72	28.29	20;143	< 0.05
	aVR	71	20.61	32;142	73	23.24	5;115	< 0.05
	aVL	36	13.46	17;82	37	11.33	17;70	< 0.05
	aVF	66	31.46	15;133	91	24.90	50;145	< 0.05
P wave axis (°)	-	49	17.90	5;90	62	15.38	21;90	< 0.05

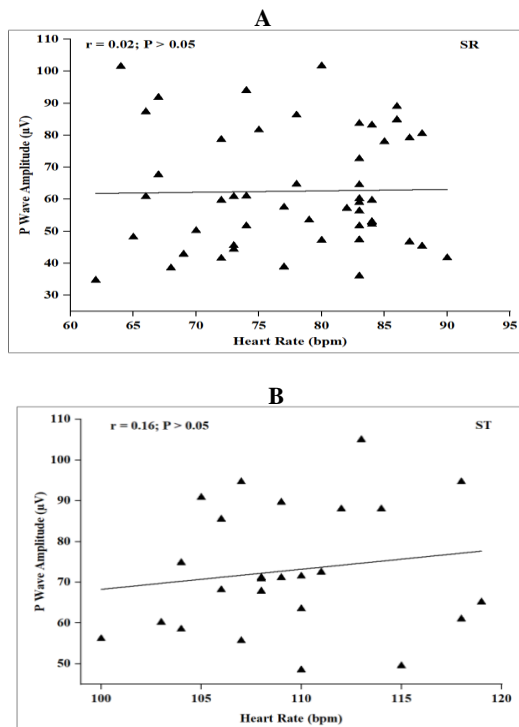


Fig. 3. Relationship of heart rate with P wave amplitude in (A) SR (n = 50) and (B) ST (n = 25).

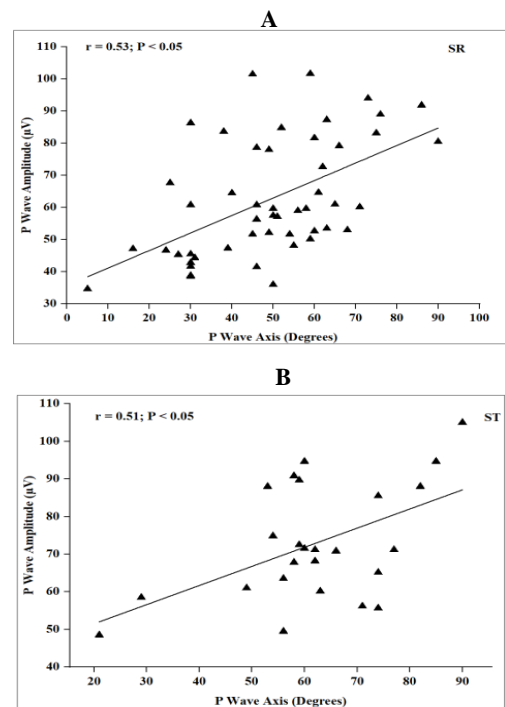


Fig. 4. Relationship of P wave axis with P wave amplitude in (A) SR (n = 50) and (B) ST (n = 25).

Heart rate vs. P wave amplitude given in Fig. 3. indicates the positive correlation between the two variables with higher positive slope value for ST ( $r = 0.16$ ) than in SR ( $r = 0.02$ ). Correlation of P wave axis with P wave amplitude is shown in Fig. 4. for SR ( $r = 0.53$ ) and ST ( $r = 0.51$ ). Fig. 5. shows the heart rate varying with respect to P wave axis in SR ( $r = 0.21$ ) and ST ( $r = 0.21$ ) groups. Table II represents the temporal aspects of atria. Heart rate, PP interval, PD differed significantly between the groups,  $P < 0.05$ , however the PWD and PRI values did not differ significantly with  $P > 0.05$  between the groups. Fig. 6. refers the heart rate vs. PWD (SR; ST:  $r = 0.09; 0.01$ ), PD ( $r = -0.16; 0.43$ ) and PRI ( $r = -0.19; 0.05$ ). It is noted that heart rate correlated well with PD in ST compared to SR.

#### IV. DISCUSSIONS

Various studies have shown the importance of PWI in detection of cardiac diseases [3,4]. These PWI represents the morphological and electrical features of P wave. Morphology of P wave is influenced by various factors: (1) Location of SA node which determines right atrial electric impulse vector, (2) structure and dimensions of both atria, (3) location of inter-atrial conduction pathway, which determines the left atrial depolarization vector [5]. During disease conditions, the above given factors tend to modify, leading to changes in P wave morphology. Various studies have used PWI for the prediction of atrial enlargement, and AF; and found that prolonged PWD is a consistent potential marker [14, 19]. Tachycardia leads to the induction of AF [9], hence early detection of persistent or



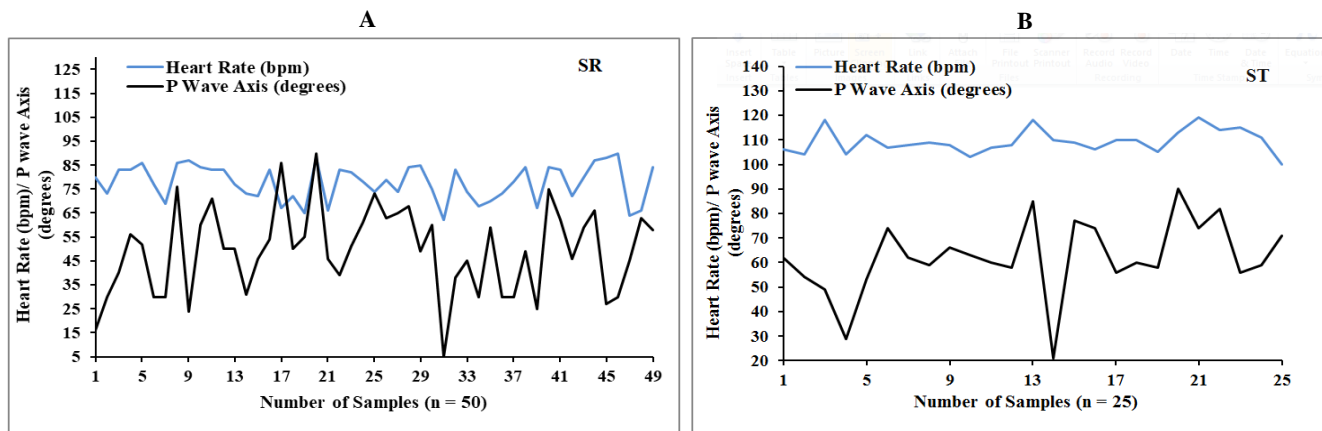


Fig. 5. Variations of P wave axis with respect to heart rate in each volunteer of (A) SR and (B) ST.

TABLE II. HEART RATE AND P WAVE FEATURES IN SR AND ST VOLUNTEERS

Measurements	SR (n=50)			ST (n=25)			P value*
	Mean	SD	Min; Max	Mean	SD	Min; Max	
Heart rate (bpm)	78	7.53	62; 90	109	4.84	100; 119	< 0.05
PP interval (ms)	778	78.91	667; 968	550	24.08	504; 600	< 0.05
P wave duration (ms)	96	9.18	72; 116	96	10.01	77; 119	> 0.05
P wave dispersion (ms)	20	11.32	2; 48	13	10.01	3; 45	< 0.05
PR interval (ms)	138	17.60	107; 181	135	18.25	101; 164	> 0.05

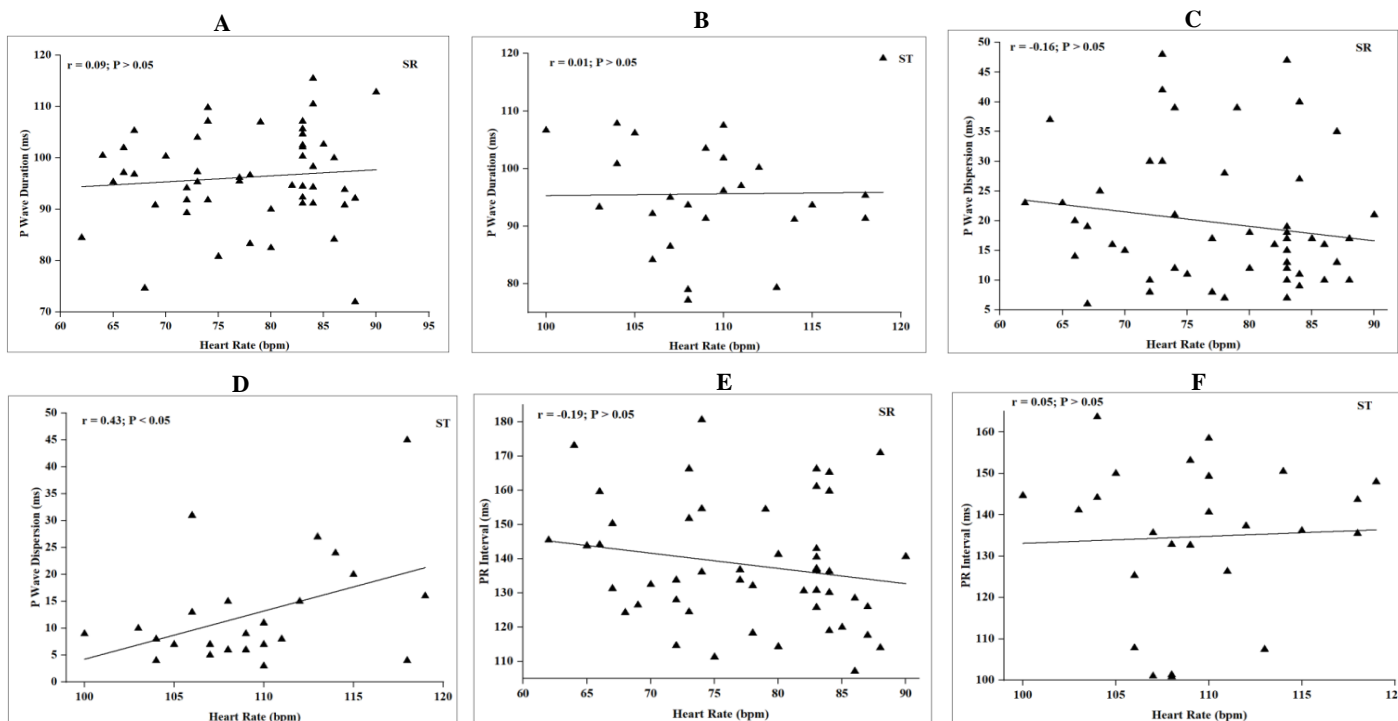


Fig. 6. Correlation of heart rate with P wave duration (A, B), P wave dispersion (C, D) PR Interval (E, F) in SR and ST groups.

inappropriate ST is of significance to prevent the occurrence of AF. In this study, PWI were analyzed and compared among SR and ST. Higher P wave amplitude in ST volunteers than SR were noted which indicates that increase in heart rate, significantly increase the P wave amplitude [20]. It is also noted that, P wave axis is larger in ST than SR. Hence, higher heart rate leads to an increased P wave axis, making atrial

depolarization electric mean vector more parallel to lead II; thus, increasing the amplitude values. Moreover, higher variability of P amplitude is noted in SR compared to ST. It can be reasoned by the study [21-23] that healthy heart is more dynamic leading to higher heart rate variability compared to diseased heart. From all the recorded leads, lead II showed highest variable range of amplitude in SR,

which is in consistent with [24]. In this study, different age group of NSR volunteers were studied and negative trend line was noted for P wave axis with the advancing age, which highlights that age has an effect on the P wave axis and morphology [25]. A study [26] improved the atrial ECG components using modified lead system to determine the P wave amplitude and durations in sinus rhythm. Moreover, it was noted that the modified placement of limb leads on human torso led to an increase in atrial amplitudes in par with the ventricular components [27] as the lead II axis is parallel to the electrical mean vector of atria. Hence, an optimal lead system to enhance the atrial components will provide detailed information of PWI in different heart conditions.

ST volunteers having higher heart rate >100 bpm lead to short PPI, similarly for SR volunteers, having lower heart rate < 90 bpm lead to longer PPI. Hence, heart rate and PPI are inversely correlated with  $r = -0.9$ . It was noted that PWD and PRI did not differ significantly between the groups. However, the minimum and maximum values have a higher range for SR compared to ST volunteers, indicating the high variability in healthy heart conditions [22]. PD for ST is lesser than the SR and both the values are within the 40 ms with outliers below 48 ms for both the groups. It is noted from the previous studies that, PD value increases significantly in atrial tachycardia condition [28], however in this study, the ST volunteers have lesser PD may be due to the effect of physiological changes. Further in depth studies are needed to validate the lesser value of PD in ST conditions. Moreover, heart rate correlated strongly with PD in ST signifying the dependency of higher heart rhythm with PWD.

## V. CONCLUSION

P wave indices are the clinical markers for the predisposition of AF and other atrial related disorders. Variability of PWI beyond the normal limits is of great significance in determining the abnormal conditions. SR and ST have been analyzed for P wave duration, amplitude, axis and dispersion to understand the physiological variabilities. Hence, a better diagnosis can be made for determining the pathological conditions with higher true positive detection. Better determination of the PWI requires the analysis of wider range of age group for future studies.

## ACKNOWLEDGMENT

The authors acknowledge the support from the Ministry of Education, Government of India. The present study was supported by financial grants from the Science Engineering Research Board (SERB), Department of Science and Technology (DST), Government of India (EEQ/2019/000148).

## REFERENCES

- [1] P. Brugada, J. Brugada, L. Mont, J. Smeets, and E. W. Andries, "A new approach to the differential diagnosis of a regular tachycardia with a wide QRS complex," *Circulation*, vol. 83, pp. 1649–1659, May 1991.
- [2] A. S. Vink et al. "Determination and interpretation of the QT interval," *Circulation*, vol. 138, pp. 2345–2358, August 2018.
- [3] J. W. Magnani, M. A. Williamson, K. M. Monahan, P. T. Ellinor, and E. J. Benjamin, "P wave indices: current status and future directions in epidemiology, clinical and research applications," *Circ. Arrhythmia. Electrophysiol.*, vol. 2, pp. 72–79, February 2009.
- [4] M. U. Rasmussen et al. "P-wave indices as predictors of atrial fibrillation," *Ann. Noninvasive Electrocardiol.*, vol. 25, pp. e12751, September 2020.
- [5] P. G. Platonov, "P-wave morphology: underlying mechanisms and clinical implications," *Ann. Noninvasive Electrocardiol.*, vol. 17, pp. 161–169, July 2012.
- [6] D. M. German, M. M. Kabir, T. A. Dewland, C. A. Henrikson, and L. G. Tereshchenko, "Atrial fibrillation predictors: Importance of the electrocardiogram," *Ann. Noninvasive Electrocardiol.*, vol. 21, pp. 20–29 November 2016.
- [7] H. J. Wellens, "Ventricular tachycardia: diagnosis of broad QRS complex tachycardia," *Heart*, vol. 86, pp. 579–585, November 2001.
- [8] E. N. Prystowsky, "Tachycardia-induced tachycardia: a mechanism of initiation of atrial fibrillation," in *Atrial Arrhythmias: State of the Art*, Futura Publishing Co, 1995, pp. 81–95.
- [9] P. Delise, L. Corò, P. Scipione, and M. Fantinel, "Tachycardia induced atrial fibrillation: What incidence? How to diagnose and treat it?," in *Cardiac Arrhythmias*, Springer, 1997, pp. 18–23.
- [10] P. Jais et al., "A focal source of atrial fibrillation treated by discrete radiofrequency ablation," *Circulation*, vol. 95, pp. 572–576, February 1997.
- [11] V. Martínez, R. Alcaraz, and J. J. Rieta, "Study on the P-wave feature time course as early predictors of paroxysmal atrial fibrillation," *Physiol. Meas.*, vol. 33, pp. 1959–1974, November 2012.
- [12] G. Piccirillo et al., "Oscillatory behavior of P wave duration and PR interval in experimental congestive heart failure: a preliminary study," *Physiol. Meas.*, vol. 39, pp. 035010, March 2018.
- [13] P. E. Dilaveris, and J. E. Gialafos, "P wave dispersion: A novel predictor of paroxysmal atrial fibrillation," *Ann. Noninvasive Electrocardiol.*, vol. 6, pp. 159–165, April 2001.
- [14] R. Andlauer, "Influence of left atrial size on P-wave morphology: differential effects of dilation and hypertrophy," *Europace*, vol. 20, pp. iii36–iii44, November 2018.
- [15] I. C. Chang et al., "Shorter minimum p-wave duration is associated with paroxysmal lone atrial fibrillation," *J. Electrocardiol.*, vol. 47, pp. 106–112, October 2013.
- [16] Y. K. Agarwal, W. S. Aronow, J. A. Levy, and D. H. Spodick, "Association of interatrial block with development of atrial fibrillation," *Am. J. Cardiol.*, vol. 91, pp. 882, April 2003.
- [17] A. R. Pérez-Riera et al., "P-wave dispersion: an update," *Indian Pacing Electrophysiol. J.*, vol. 16, pp. 126–133, October 2016.
- [18] E. Z. Soliman, M. Cammarata, and Y. Li, "Explaining the inconsistent associations of PR interval with mortality: the role of P-duration contribution to the length of PR interval," *Heart Rhythm*, vol. 11, pp. 93–98, January 2014.
- [19] F. Holmqvist, J. Carlson, J. E. P. Waktare, and P. G. Platonov, "Noninvasive evidence of shortened atrial refractoriness during sinus rhythm in patients with paroxysmal atrial fibrillation," *Pacing Clin. Electrophysiol.*, vol. 32, pp. 302–307, March 2009.
- [20] M. E. Field et al., "P-wave amplitude and PR changes in patients with inappropriate sinus tachycardia: Findings supportive of a central mechanism," *J. Am. Heart Assoc.*, vol. 7, pp. e008528, April 2018.
- [21] J. Sivaraman et al., "A study on stability analysis of atrial repolarization variability using ARX model in sinus rhythm and atrial tachycardia ECGs," *Comput. Methods Programs Biomed.*, vol. 137, pp. 341–351, December 2016.
- [22] J. Arumughan, A. Bhardwaj, and J. Sivaraman, "Stability analysis on the effects of heart rate variability and premature activation of atrial ECG dynamics using ARMAX model," *Phys. Eng. Sci. Med.*, vol. 43, pp. 1361–1370, December 2020.
- [23] B. Dhananjay, and J. Sivaraman, "The role of heart rate variability in atrial ECG components of normal sinus rhythm and sinus tachycardia subjects," *Intelligent System Design. Advances in Intelligent Systems and Computing*, Springer, Singapore, vol. 1171, pp. 637–644, January 2021.
- [24] P. W. Macfarlane et al., "Normal limits of the electrocardiogram in Indians," *J. Electrocardiol.*, vol. 48, pp. 652–668, May 2015.
- [25] R. Havmoller et al., "Age-related changes in P wave morphology in healthy group," *BMC Cardiovasc. Disor.*, vol. 7:22, July 2007.

- [26] R. John and J. Sivaraman, "Effects of sinus rhythm on atrial ECG components using modified limb lead system," 2017 4th International Conference on Signal Processing, Computing and Control (ISPPCC), Solan, India, pp. 527–530, September 2017.
- [27] J. Sivaraman, S. Venkatesan, R. Periyasamy, and J. Justin, M. S. Ravi, "Modified limb lead ECG system effects on electrocardiographic wave amplitudes and frontal plane axis in sinus rhythm subjects," *Anatol. J. Cardiol.*, vol. 17, pp. 46–54, January 2017.
- [28] O. Ozdemir et al., "Does P-wave dispersion predict the atrial fibrillation occurrence after direct-current shock therapy?," *Angiology*, vol. 57, pp. 93–98, January 2006.

# Sleep Disorder Diagnosis using EEG based Deep Learning Techniques

T. Sudhakar  
 Department of Biomedical Engineering  
 Sathyabama Institute of Science &  
 Technology  
 Chennai, India

Bethanney Janney  
 Department of Biomedical Engineering  
 Sathyabama Institute of Science &  
 Technology  
 Chennai, India

G. Hari Krishnan  
 Department of Electrical and  
 Electronics Engineering  
 Sree Vidyanikethan Engineering  
 College  
 Tirupati, India  
 haris\_eee@yahoo.com

M. Pradeema  
 Department of Biomedical Engineering  
 Sathyabama Institute of Science &  
 Technology  
 Chennai, India

N. R. Krishnamoorthy  
 Department of Electronics and  
 Instrumentation  
 Sathyabama Institute of Science &  
 Technology  
 Chennai, India

J. P. Raghavi  
 Department of Biomedical Engineering  
 Sathyabama Institute of Science &  
 Technology  
 Chennai, India

**Abstract**—This Proposed system detects the sleep disorder through the EEG signals using by deep learning techniques (Alex net, Google net) in which EEG signals are used as inputs to a deep convolution network to solve visual recognition tasks. Electroencephalograph (EEG) based on sleep stage analysis is helpful for detect the sleep disorder. thirty-layer CNN model is designed to automatically detect the sleep disorder using EEG signals We Obtained accuracy to received output. Obtained good performance even with a smaller number of normal and sleep disorder data sets.

**Keywords**—EEG signals, Neural network, Sleep disorder, Alex net, Google net

## I. INTRODUCTION

The Electroencephalogram (EEG) signals are recorded over the scalp using multiple numbers of electrodes to obtain brain electrical signals. In existing methods EEG signals are recorded by extracting the patterns of EEG during sleeping condition and analysed using discrete wavelet transform (DWT) [1] & [2]. Statistical parameters like variance, SD, Energy etc are used to calculate sub band coefficients to identify the sleep disorders (T.V.K.H Rao et al and Omer Turk et al.,). The disadvantage is the inability to detect the different sleep disorders and less efficient and natural [3] & [4]. The proposed method data sets were well trained and analysed by using developed algorithms to determine the level of success of neural networks [4] & [5]. This project proposed to detect the sleep disorder from EEG signals by using deep learning neural networks [6] & [7].

The aim of our proposed system is to detect the sleep disorder from the Electroencephalogram (EEG) using Deep learning neural networks [8], [9] & [10]. By Alex Net, Google Net [11] & [12]. In existing methods detection of difference in sleep disorders was not possible and few existing methods are very less efficient and natural [13] & [14]. In the proposed method we use 3 types of disorder e.g. epilepsy normal EEG signals, sleep disorder. Detected by using 2 different types of architecture which each has more than 7 hidden layers [15], [16] & [17]. We compare theoretical parameters of all the architecture involved and detect the sleep disorders which belong to data sets [18], [19] & [20]. Data augmentation is used in order to avoid overfitting [21], [22] & [23]. In future this processed method will be used as the hardware technique [24] & [25]. In the EEG machine, MATLAB software is installed, which

automatically detects the type of disorder the patient is suffering [26], [27] & [28].

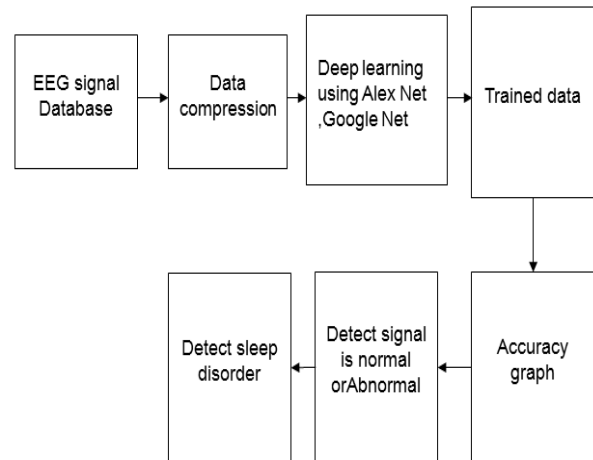


Figure.1 Proposed system main block diagram



Figure.1 Google net architecture

### 1.1 PREPROCESSING:

#### Google Net Graph and Confusion Matrix:

1) Google Net Graph:

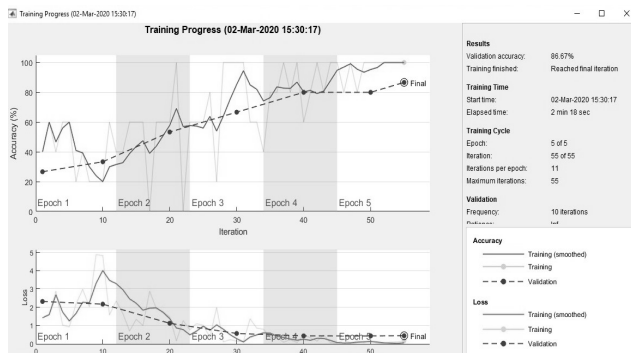


Fig.3. Google Net Training progress-(5,5)

When the maximum epoch size :5 and maximum epoch:5. Figure 3 shows that accuracy level of Google Net is 86.67%.

Confusion Matrix:

Table 1 Confusion Matrix-(5,5)

Max epoch size:5 Max epoch:5	Epilepsy	Normal	Sleep disorder
Epilepsy	5	0	0
Normal	0	5	2
Sleep disorder	0	0	3

The Table 1 shows confusion matrix for epilepsy,normal EEG, sleep disorder for epilepsy it shows 5,normal EEG it shows 5,sleep disorder it shows 3.

**RESULT AND DISCUSSION**

The objective of the proposed method is to detect the respective disorder of the given input data between different sleep disorders [29], [30] & [31]. The sleep disorder through the EEG signals were analysed using AlexNet and Google Net architectures. Finally we conclude that AlexNet has high accuracy [32].

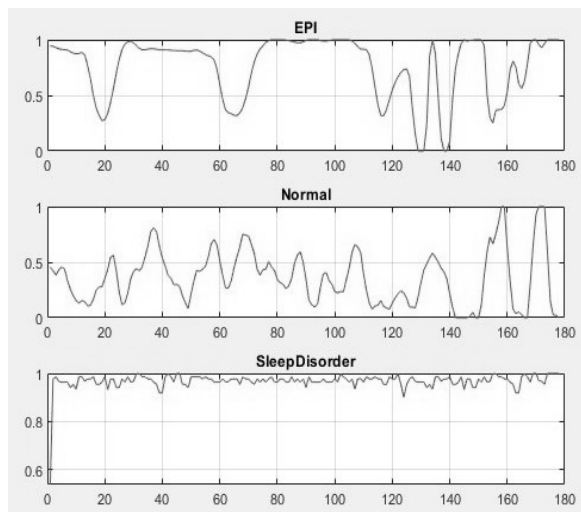


Figure 3 Alex Net signal plot

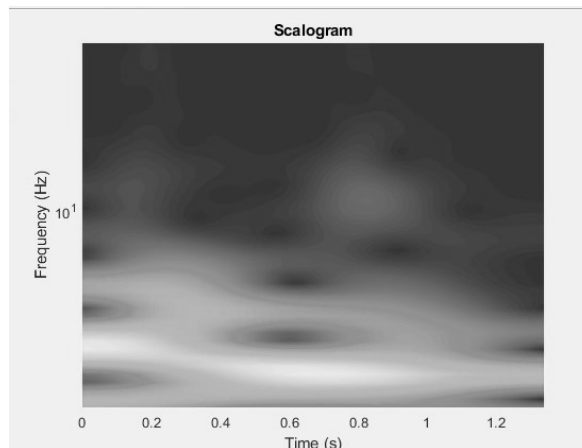


Figure. 4 Alex Net scalogram

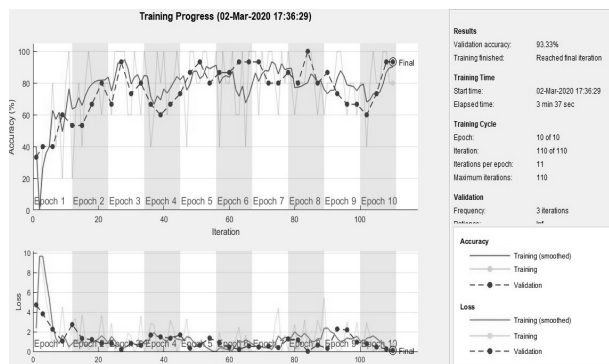


Figure. 5 Alex Net Training Progress

Confusion matrix:

Table 2 Confusion Matrix

Max epoch size:5 Max epoch:10	Epilepsy	Normal	Sleep disorder
Epilepsy	5	0	0
Normal	3	2	0
Sleep disorder	0	0	5

The Table 2 shows confusion matrix for epilepsy,normal EEG, sleep disorder for epilepsy it shows 5,normal EEG it shows 2,sleep disorder it shows 5.

This method has two architectures Alex Net , Google Net to train the data and detect the sleep disorder using EEG signals. This processed method will be used to diagnose the sleep disorder.

## SUMMARY AND CONCLUSION

In this project, the attempts to detect the sleep disorder by using EEG signals with deep learning neural networks. For this purpose, the database and real time data collected from the hospital. This data was taken as standard data for healthy subjects and particular sleep disorders. The architecture of Alex Net and Google Net was used to predict the sleep disorder by using EEG signals. We described our work in predictive modelling in sleep disorder using a deep learning convolutional neural network (CNN).

Evidently Alex Net performs better than Google Net as deep learning method detection of sleep disorder. The level of AlexNet accuracy is 93.33%.

### Acknowledgement:

The authors of this research work were thankful to the Department of Biomedical Engineering, Sathyabama University and Centre for Energy, Sree Vidyanikethan Engineering College for providing necessary support in all means.

## REFERENCES

- [1] Aboalayon, K.A.I., Faezipour, M., Almuhammadi, W.S., Moslehpour, S.: Sleep stage classification using EEG signal analysis: a comprehensive survey and new investigation. *Entropy* 18(9), 272–303 (2016).
- [2] A deep learning architecture to detect events in EEG signals during sleep Stanislas Chambon<sup>1,2,3</sup>, Valentin Thorey<sup>2</sup>, Pierrick J. Arnal<sup>2</sup>, Emmanuel Mignot<sup>1</sup>, Alexandre Gramfort<sup>3,4,5</sup>. 2016.p.426-440.
- [3] Automatic Sleep Staging Based on Deep Neural Network Using Single Channel EEG Yongfeng Huang ,Yujuan Zhang ,Cairong Ya.pp. 63-73.12,June 2019 .vol.1027.
- [4] Classification of Sleep Disorders Based on EEG Signals by using Feature Extraction Techniques with KNN Classifier T.V.K.H. Rao and Dayanand Vishwanath Dhongade..2012;105-114.
- [5] Deep convolutional neural network for the automated detection and diagnosis of seizure using EEG signal open ShuLihOhaYuki HagiwaraaJen Hong Tana Hojjat Adelid. *comput. Boil* 2018, 100, 270-278.
- [6] DeepSleep: Convolutional Neural Networks for Predictive Modelling of Human Sleep Time-Signals Sarun Paisarnsrimsuk and Sergio A. Alvarez and Carolina Ruiz.2002,26 ,1041-1054.
- [7] Epilepsy Detection by Using Scalogram Based Convolutional Neural Network from EEG Signals ÖmerTürk 1,\* and Mehmet Siraç Özerdem.2005, 46 ,470-472.
- [8] Epilepsy Detection by Using Scalogram Based Convolutional NeuralNetwork from EEG Signals Songyun Xie, Yabing Li ,Xinzhou Xie, Wei Wang,Xu Duan. 2000;101(23):215-220.
- [9] Detecting Sleep Disorders Based on EEG Signals by Using Discrete Wavelet Transform T.V.K.H. Rao and DhongadeDayanand Vishwanath.978-4244-1815-2,2008.
- [10] Fraiwan, L., Lweesy, K., Khasawneh, N., Wenz, H., Dickhaus, H.: Automated sleep stage identification system based on time-frequency analysis of a single EEG channel and random forest classifier. *Comput.Methods Programs Biomed.*108(1), 10–19 (2012).
- [11] Guru Anand, V., Hari Krishna, G., Mohandass, G., Hemalatha, R.J., Sundaram, S. “Predicting grade of prostate cancer using image analysis software,” *Proceedings of the 2nd International Conference on Trendz in Informtion Sciences and Computing, TISC-2010*, 5714621, pp. 122–124, 2010.
- [12] Hassan, A.R., Bashar, S.K., Bhuiyan, M.I.H.: Automatic classification of sleep stages from single-channel electroencephalogram. In: *India Conference 2015*, pp. 1–6. IEEE, New Delhi, India (2015)
- [13] Hinton, G., Deng, L., Yu, D., Dahl, G.E., Mohamed, A.R., Jaitly, N., Senior, A., Vanhoucke, V., Nguyen, P., Sainath, T.N.: Deep neural networks for acoustic modelling in speech recognition: the shared views of four research groups. *IEEE Sig. Process. Mag.*29(6),82-97(2012)
- [14] Hari Krishnan, G.H., Ananda Natarajan, R., Nanda, A. Impact of upper limb joint fluid variation on inflammatory diseases diagnosis. *Journal of Electrical Engineering and Technology*, 9(6), 2114–2117, 2014.
- [15] Hari Krishnan, G., Hemalatha, R.J., Umashankar, G., Ahmed, N., Nayak, S.R., “Development of Magnetic Control System for Electric Wheel Chair Using Tongue”, *Advances in Intelligent Systems and Computing* 308 AISC (VOLUME 1), pp.635-641, 2015.
- [16] Herrera, L.J., Mora, A.M., Fernandes, C., Migotina, D., Guillen, A., Rosa, A.C.: Symbolic representation of the EEG for sleep stage classification. In: *International Conference on Intelligent Systems Design and Applications* 2011, vol. 60, pp. 253–258. IEEE, Cordoba, Spain (2011).
- [17] Hari Krishnan, G., Umashankar, G., Abraham, S., “Cerebrovascular disorder diagnosis using MR angiography”, *Biomedical Research (India)*, 27(3), pp.773-775 2016.
- [18] Hari Krishnan, G., Abhinaya, N., Hemalatha, R.J., Mohandass, G., “Hardware implementation for feedback control based health monitoring and drug delivery”, *Biomedicine (India)*, 37(1), pp.123-126, 2017.
- [19] Hemalatha, R.J., Krishnan, G.H., Umashankar, G., Abraham, S. “Computerized breast cancer detection system”, *Biosciences Biotechnology Research Asia* 11(2), pp.907-910, 2014.
- [20] Ilangovan, N., Hari Krishnan, G. “Wheel chair movement control using human input: Comparative study approach”, *Research Journal of Pharmaceutical, Biological and Chemical Sciences*, 6(3), pp.568-570, 2015.
- [21] Krishnan, G.H, Natarajan, R.A., Nanda, A. Comparative study of rheumatoid arthritis diagnosis using two different methods, *Biomedical and Pharmacology Journal* 7(1), 379–382, 2014.
- [22] Krishnan, G.H, Nanda, A., Natarajan, A. Synovial fluid density measurement for diagnosis of arthritis, *Biomedical and Pharmacology Journal* 7(1), 221–224, 2014.
- [23] Krishnan, G.H., Natarajan, R.A., Nanda, A. “Microcontroller based non invasive diagnosis of knee joint diseases”, *International Conference on Information Communication and Embedded Systems, ICICES 2014*, 7034178, 2015.
- [24] Mohandass, G., Ananda Natarajan, R., Hari Krishnan, G. Comparative analysis of optical coherence tomography retinal

- image using multidimensional and cluster methods. *Biomedical Research (India)*, 26(2), pp.273-285, 2015.
- [25] Nagarjuna Reddy, A., Hari Krishnan, G., Raghuram, D., "Real time patient health monitoring using raspberry PI, *Research Journal of Pharmaceutical, Biological and Chemical Sciences*, 7(6), pp.570-575, 2016.
- [26] Pan, S.T., Kuo, C.E., Zeng, J.H., Liang, S.F.: A transition-constrained discrete hidden markov model for automatic sleep staging. *BioMed.Eng. Online* 11(1), 1–19 (2012).
- [27] Rechtschaffen, A.Q., Kales, A.A.: A manual of standardized terminology techniques and scoring system for sleep stages in human subjects. *Psychiatry Clin.Neurosci.*26(6), 6-44 (1968).
- [28] Ronzhina, M., Janoušek, O., Kolářová, J., Nováková, M., Honzík, P., Provazník, I.: Sleep scoring using artificial neural networks. *Sleep Med. Rev.* 16(3), 251–263 (2012).
- [29] Schmidhuber, J.: Deep learning in neural networks: an overview. *Neural Network.*61, 85–117 (2014).
- [30] Schulz, H.: Phasic or transient? Comment on the terminology of the AASM manual for the scoring of sleep and associated events. *J. Clin. Sleep Med. JCSM Official Publ. Am. Acad. Sleep Med.* 3(7), 7-52 (2007).
- [31] SleepEEGNet: Automated sleep stage scoring with sequence to sequence deep learning approach Sajad MousaviID1\*, Fatemeh Afghah1, U. Rajendra Acharya2,3,4. 2002,253,1087-1100.
- [32] Usage of Deep Learning in Epileptic Seizure Detection Through EEG Signal Nalini Singh ,Satchidananda Dehuri.August 2, 2018, vol.551.Pp.219-228.

# A Preliminary Study on ReliefF based Feature Ranking for Classification of Myoelectric Signals

Salman Mohd Khan  
 Dept. of Mechanical Engineering  
 Aligarh Muslim University  
 Aligarh, India  
[salmanmkhan225@gmail.com](mailto:salmanmkhan225@gmail.com),  
 0000-0003-4662-1315

Mohd. Umair Alam  
 Dept. of Electrical Engineering  
 Aligarh Muslim University  
 Aligarh, India  
[alamumair080@gmail.com](mailto:alamumair080@gmail.com)

Abid Ali Khan  
 Dept. of Mechanical Engineering  
 Aligarh Muslim University  
 Aligarh, India  
[abid.khan.me@amu.ac.in](mailto:abid.khan.me@amu.ac.in),  
[abidalikhan@gmail.com](mailto:abidalikhan@gmail.com)

Omar Farooq  
 Dept. of Electronics Engineering  
 Aligarh Muslim University  
 Aligarh, India  
[omar.farooq@amu.ac.in](mailto:omar.farooq@amu.ac.in)

**Abstract**— Myoelectric pattern recognition is an important technique in the design of prosthetic devices. The two important parameters in the working of the myoelectric pattern recognition technique are feature extraction and feature ranking. The computation of excess feature extraction results in redundancy. In this study; 12 time domain features were extracted. To optimize the number of features for the classification of phases of pick and place task, a ReliefF feature ranking algorithm is used. The subjectwise classification was performed using the k-NN classifier. From the classification outcomes; the evaluation of the features was performed using 5 levels effective chart. It was found that Auto Regressive coefficient and Willison Amplitude were the most effective features, followed by the Waveform Length, Root Mean Square and Myopulse Percentage Rate. The tabulation of features in the designed chart will ease the processing time and procedure in the design of the myoelectric pattern recognition process.

**Keywords**—*Electromyography, Feature extraction, Feature selection, ReliefF*

## I. INTRODUCTION

Myoelectric pattern recognition (MPR) technique plays a pivotal role in the design of assistive devices to mimic the functionalities of the natural limb [1]. Myoelectric Prosthetic devices are one such assistive devices used for upper limb amputations. The advantage associated with the MPR technique is its ability to compile a large amount of information and process it to perform different activities of daily living.

The MPR techniques use electrical signals from the human muscles by the technique called Electromyography (EMGs) [2]. These signals are used as the inputs for processes involved in the MPR. The MPR technique is a compilation of a number of sub-processes. These sub-processes include feature extraction, feature selection, feature ranking and classification [3].

The feature extraction converts raw EMG data acquired from the muscles to useful information by mathematical computation [4]. The features extracted are classed into 3 sections: time domain features, frequency domain features and time frequency domain features [3]. Different researchers have computed different features for the extraction of useful information. Usually; authors prefer to use time domain features for the MPR process for upper

limb devices [5]–[7]. The feature selection and feature ranking processes are the algorithm processes required for the selection of the optimal features from the various features computed in a study [8], [9]. This allows reduced processing in the designed algorithm. Lastly; the selected features are used as the inputs for the classification process. The classification processes use different outputs such as hand gesture classification, grip force classification and grasp type classification [10]–[12].

Previous studies have largely focused on the feature extraction and classification processes in the MPR techniques [12]–[14]. One of the reasons is that these processes play a pivotal role in real time processing. However; the feature selection and feature ranking process is also an essential process, especially to identify redundant features. Among the two processes; the feature ranking process is a relatively lesser used technique. However; in recent studies, it has been found the feature ranking demonstrates a clear perspective in the feature selection process. One such technique is the ReliefF technique [15]. This technique is previously used in the feature ranking process for the electroencephalography technique [16].

In this study; the application of ReliefF has been evaluated for the feature selection process. The ranked features are used for the classification of different phases of the pick and place task. The different phases of this task are defined as the reach-to-grasp phase, grasping phase and reaching away phase. For classification; different classification schemes have been used in previous studies. This includes linear discriminant analysis, neural networking, support vector machines and k-nearest neighbors, etc. [17], [18]. In this study; k-NN has been used as a classifier owing to its high recognition rate in the previous studies [17], [19].

The following sections of the study are designed as follows: Section 2 details the experimental details and methods used in the study with the analysis procedure has been incorporated in section 3. The results have been discussed in Section 4, followed by the discussion in Section 5. Lastly; the conclusion has been made in Section 6.

## II. EXPERIMENT AND METHODS

The EMG signals were acquired from 2 muscles of the forearm. These muscles were: flexor carpi radialis and extensor carpi radialis. The recordings were performed using bipolar electrodes (M/s Biometrics Ltd., UK)



positioned according to the Surface Electromyography for the Non-Invasive Assessment of Muscles (SENIAM). The ground electrode was positioned behind the earlobe. The sampling rate of the EMG DAQ was 1024Hz. The recording was performed using a Datalink (Biometric Ltd.) software as shown in Fig. 1.

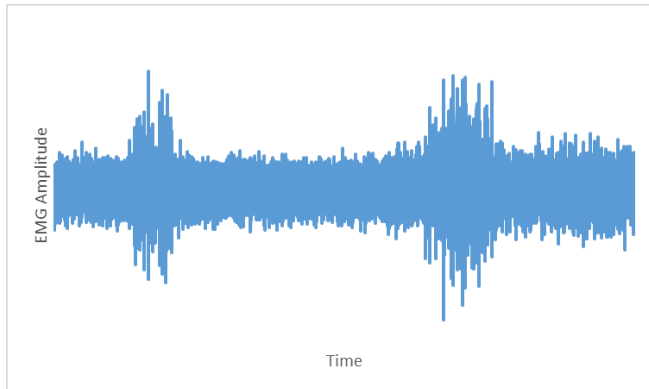


Fig. 1: Sample EMG Recording during Pick and Place task.

A preliminary experiment was performed for the pick and place task. 4 able-bodied, right-handed, male subjects were asked to perform a simple pick and place task. Before the experiment; all subjects were asked to sign a written consent.

During the experiment; the subjects were seated on an adjustable chair. The subjects were asked to grasp and lift a gripper. The gripper was designed such that it was capable to detect the fingertip grasping force for different levels of weights. The sampling rate of the gripper was 90Hz. The span of the task was 12-15 seconds. The subjects were asked to make multiple repetitions of the task.

### III. ANALYSIS

On the basis of the EMG data and force data acquired during the pick and place activity; the whole span was divided into 3 phases. The division of the phases was decided by detecting the changes in gripper force. The end of reach to grasp and start of grasping phase was detected based on any increment in force from zero value. Similarly; when the hand was removed from the gripper after the grasp and lift gesture; the force value changed to zero, indicating the start of phase 3.

The raw EMG data was subjected to pattern recognition processing. The pattern recognition in this study involved data segmentation, feature extraction, feature ranking and classification. The overlapped data segmentation was performed for a segment length of 250 ms and an overlap of 150ms. For each segment; different time domain features were extracted. These time domain features and their definitions in this study are listed in Table 1.

The feature extracted from the EMG signals were then up-sampled to equate the data points of EMG and force signals. For the feature ranking of different features; a relatively new technique, called ReliefF is used. The ReliefF technique is a statistical technique which computes statistical values in terms of weights [20]. This algorithm computes feature weights for all the feature values. The range of these weights varies from -1 to +1. The features with a weight value closer to +1 are selected as the ones with better

TABLE 1: List of features used in this study.

S. No.	Features	Definition
1.	Auto-Regressive Coefficient (AR)	Auto Regressive coefficients are computed using a prediction model in which the EMG sample is considered as a combination of previous signals and errors.
2.	Waveform Length (WL)[21]	WL is related to the fluctuations of a signal when the muscle is active. .
3.	Mean Absolute Value (MAV) [21]	MAV is an estimate of the standard deviation of the signal and is often used for proportional control of prostheses.
4.	Root Mean Square (RMS)[3]	RMS is the quadratic mean of the signal
5.	Modified Mean Absolute Value (Mod MAV) [21]	Mod MAV is an extension of MAV, where a Hanning window is applied to the signal to improve the robustness of the feature.
6.	Willison amplitude (WAMP)[3]	WAMP estimates the number of active motor units, which is an indicator of the level of muscle contraction.
7.	Myopulse percentage rate (MPR)	MPR is calculated from the average value of myopulse output.
8.	Slope Sign Index (SSI) [21]	SSI is defined as the number of times that the slope of the EMG waveform changes sign within an analysis window.
9.	Histogram (HIST) [22]	This feature provides information about the frequency with which the EMG signal reaches various amplitudes.
10.	Variance (VAR) [22]	The Variance of EMG expresses the power of the EMG signal as a useable feature.
11.	Integrated EMG (Int EMG) [3]	Integrated EMG features is defined as the area under the curve from the absolute values of EMG signals.
12.	Difference of mean absolute value (Diff MAV) [21]	It is defined as the difference of two mean absolute values.

TABLE 2: List of features ranked on the basis of ReliefF for each Subject.

	Subject 1	Subject 2	Subject 3	Subject 4
Features	AR	AR	AR	AR
	RMS	WAMP	WAMP	WAMP
	WAMP	RMS	MPR	SSI
	MPR	Int EMG	Int EMG	MPR
	SSI	WL	WL	VAR
	HIST	Mod MAV	RMS	WL
	MAV	SSI	Mod MAV	Mod MAV
	WL	VAR	SSI	Int EMG
	VAR	MPR	VAR	RMS
	VAR	HIST	HIST	HIST
	Int EMG	MAV	MAV	Diff. MAV
	Diff. MAV	Diff. MAV	Diff. MAV	MAV

influence than a feature with a lesser weight value. In this way; different features extracted are ranked. In this study; subject-specific feature ranking is calculated. The ranking for different subjects is listed in Table 2.

The final step of the whole process was classification. The ranked features were subjected to the classification process by using the k-NN classifier. K-NN classifier uses Euclidean distance as a measure relative to the nearest neighbors for the assignment of an event in a given class [1].

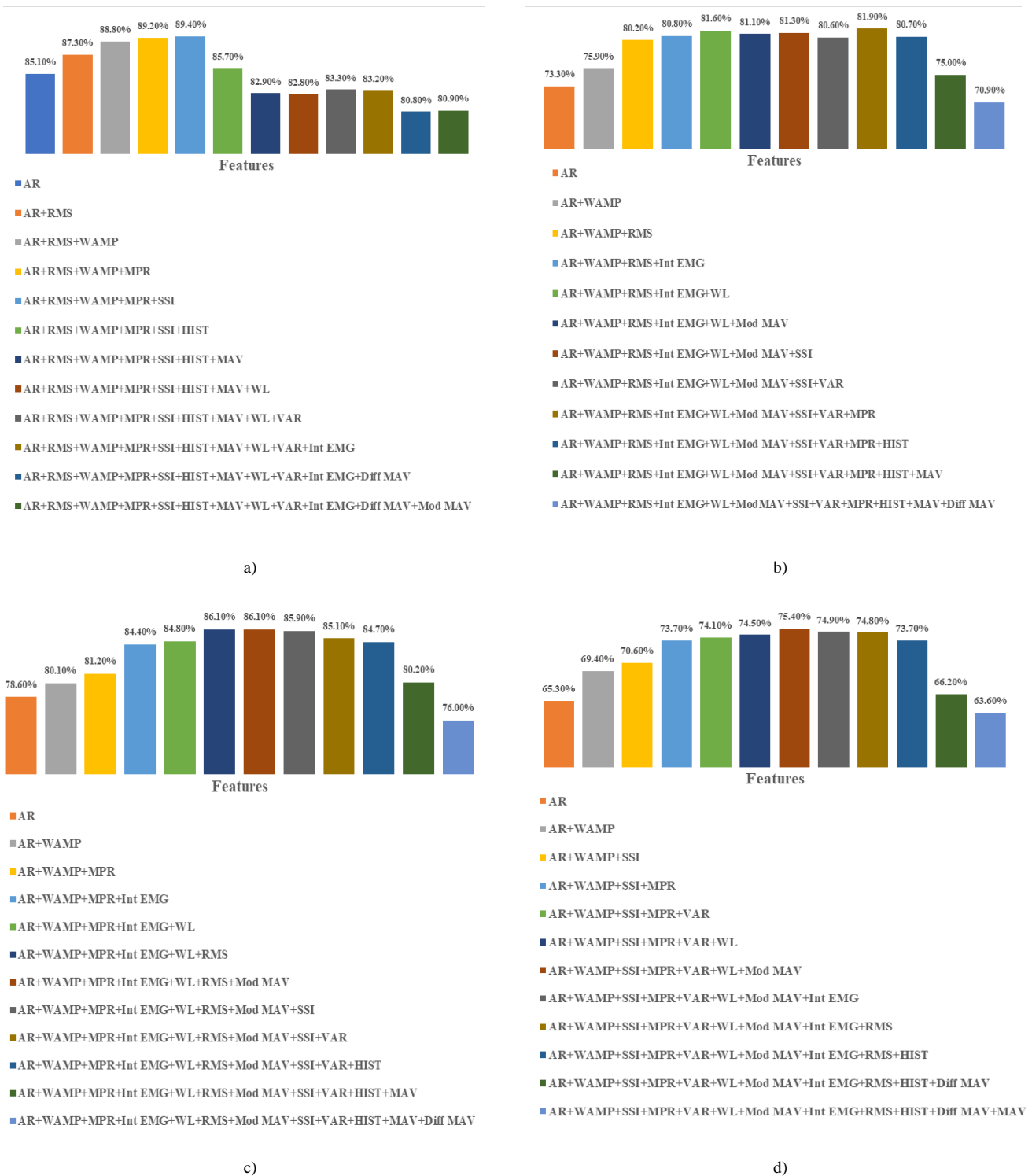


Fig. 2: Classification Accuracy of 4 subjects with different feature sets as input. A) Subject 1 Classification for different Feature Combinations; b) Subject 2 Classification for different Feature Combinations; c) Subject 3 Classification for different Feature Combinations; d) Subject 4 Classification for different Feature Combinations.

The k value in this study was assigned as the 5 to ensure a better likelihood of the classification of the features. The classification was a subject-specific process. The

classification was performed for feature sets designed on the basis of the ranks. For each subject; the first classification was performed by using the highest ranked feature as the

input. For the next classification; the 2<sup>nd</sup> ranked feature was added to the top ranked feature as the input. In this way; sequential classification was performed for different feature combinations designed on the basis of the rank of the features. Thus; for each step of the classification; 2 feature sets were added as the input for the classification of the phases. The classification process was performed with 5 cross fold validation.

#### IV. RESULTS

The classification outcomes by using different feature sets for each subject are shown in Fig. 2. The highest classification accuracy achieved by using different feature sets was 89.4 for subject 1. This classification accuracy was obtained for a feature set of AR, RMS, WAMP, MPR and SSI. Similarly; the highest classification accuracy achieved for subject 2, subject 3 and subject 4 were found 81.30%, 86.10% and 75.4% respectively.

#### V. DISCUSSION

The feature ranking process using the ReliefF technique helped us to select optimal feature combinations. The feature sets for which the highest classification accuracy was achieved in 4 subject cases are shown in Table 3. The selection of feature set was dependent on the fact that the classification accuracy must have increased by at least 0.5% when a new features are added to the previous feature set.

Table 3: Optimal Feature Sets obtained for different subjects and their classification accuracy

Subject Number	Optimal Feature Set	Classification Accuracy
Subject 1	AR+RMS+WAMP+MPR+SSI	89.40%
Subject 2	AR+WAMP+RMS+Int EMG+WL	81.60%
Subject 3	AR+WAMP+MPR+Int EMG+WL+RMS	86.10%
Subject 4	AR+WAMP+SSI+MPR+VAR+WL + Mod MAV	75.40%

For instance; the highest classification accuracy was found to be 81.9%. However; the optimal feature set shown in Table 3 resulted in classification accuracy 81.6%. The selection of the feature set was based on the fact the increment in classification accuracy must be sufficient enough to make conclusive outcomes. The result shown suggests that the AR coefficient feature and Willison Amplitude features were part of the feature set in all cases. Thus; these features were found to be important features in detecting 3 different phases of the pick and place task. Another set of features such as WL, RMS and MPR were found to be the part optimal feature set in cases of 3 subjects. This group of features is also expected to improve the classification accuracy. The influential role of these features in obtaining better classification accuracy of the upper limb activities can be confirmed from the previous studies as well [13].

Thus; on the basis of the information of the optimal feature sets; the 5 levels effective chart is drawn for different

features. This designates the importance of different features. The features which were repetitive for all the subjects were club as extremely effective features while features that were not the part of optimal feature set were clubbed as not-effective features shown in Table 4. Thus; it is found that features such as MAV, HIST and Diff MAV were found ineffective on the basis of the chart.

Table 4: Classes assigned for different features based on different optimal feature sets

	Extremely Effective	Very Effective	Moderately Effective	Slightly Effective	Not Effective
Features	AR	RMS	SSI	VAR	MAV
	WAMP	MPR	Int EMG	Mod MAV	HIST
		WL			Diff MAV

#### VI. CONCLUSION

This study helped us to obtain high classification accuracy for the selection of the phases. The classification accuracy was achieved using a feature set obtained by the feature ranking process by the ReliefF algorithm.

With the optimal feature sets obtained, a 5 level effectiveness chart is designed, in which AR coefficient and WAMP were found as the most effective features, followed by RMS, MPR and WL. In the future; the conclusions drawn from this study will be validated by using the importance of the feature sets. Besides this; this study included data of two muscles of the forearm. To improve the classification accuracy; more number of muscles can be selected for EMG data acquisition.

#### REFERENCES

- [1] S. Abbaspour, M. Lindén, H. Gholamhosseini, A. Naber, and M. Ortiz-Catalan, "Evaluation of surface EMG-based recognition algorithms for decoding hand movements," *Med. Biol. Eng. Comput.*, vol. 58, no. 1, pp. 83–100, 2020, doi: 10.1007/s11517-019-02073-z.
- [2] J. L. Bethausen *et al.*, "Limb Position Tolerant Pattern Recognition for Myoelectric Prosthesis Control with Adaptive Sparse Representations from Extreme Learning," *IEEE Trans. Biomed. Eng.*, vol. 65, no. 4, pp. 770–778, 2018, doi: 10.1109/TBME.2017.2719400.
- [3] S. M. Khan, A. A. Khan, and O. Farooq, "Selection of features and classifiers for EMG-EEG-Based upper limb assistive devices - A review," *IEEE Rev. Biomed. Eng.*, vol. 13, no. 1, pp. 248–260, 2020, doi: 10.1109/RBME.2019.2950897.
- [4] D. Joshi, S. Atreya, A. S. Arora, and S. Anand, "Trends in EMG based Prosthetic Hand Development: A Review," *Indian J. Biomech.*, no. March, pp. 228–232, 2009.
- [5] D. Tkach, H. Huang, and T. A. Kuiken, "Study of stability of time-domain features for electromyographic pattern recognition," *J. Neuroeng. Rehabil.*, vol. 7, no. 1, pp. 1–13, 2010, doi: 10.1186/1743-0003-7-21.
- [6] A. D. Bellingegni *et al.*, "NLR, MLP, SVM, and LDA: a comparative analysis on EMG data from people with trans-radial amputation," *J. Neuroeng. Rehabil.*, pp. 1–16, 2017, doi: 10.1186/s12984-017-0290-6.
- [7] A. Krasoulis, S. Vijayakumar, and K. Nazarpour, "Evaluation of regression methods for the continuous decoding of finger movement from surface EMG and accelerometry," in *International IEEE/EMBS Conference on Neural Engineering, NER*, 2015, vol. 2015-July, pp. 631–634, doi: 10.1109/NER.2015.7146702.

- [8] J. Camargo and A. Young, "Feature Selection and Non-Linear Classifiers: Effects on Simultaneous Motion Recognition in Upper Limb," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 27, no. 4, pp. 743–750, 2019, doi: 10.1109/TNSRE.2019.2903986.
- [9] J. Too, A. R. Abdullah, and N. M. Saad, "Study of EMG Feature Selection for Hand Motions Classification," *Int. J. Hum. Technol. Interact.*, vol. 3, no. April, pp. 19–24, 2019, [Online]. Available: <http://journal.utem.edu.my/index.php/ijhati/article/view/5094>.
- [10] M. Zia, S. O. Gilani, A. Waris, I. K. Niazi, E. N. Kamavuako, and A. Denmark, "A Novel Approach for Classification of Hand Movements using Surface EMG Signals," *2017 IEEE Int. Symp. Signal Process. Inf. Technol.*, pp. 265–269, 2017.
- [11] A. H. Al-Timemy, G. Bugmann, J. Escudero, and N. Outram, "Classification of finger movements for the dexterous hand prosthesis control with surface electromyography," *IEEE J. Biomed. Heal. Informatics*, vol. 17, no. 3, pp. 608–618, 2013.
- [12] F. Leone *et al.*, "Simultaneous sEMG classification of hand/wrist gestures and forces," *Front. Neurobot.*, vol. 13, no. June, pp. 1–15, 2019, doi: 10.3389/fnbot.2019.00042.
- [13] A. Krasoulis, S. Vijayakumar, and K. Nazarpour, "Multi-grip classification-based prosthesis control with two EMG-IMU sensors," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 28, no. 2, pp. 508–518, 2020, doi: 10.1101/579367.
- [14] S. M. Khan, A. A. Khan, and F. Omar, "A Neural Network Classification of sEMG Signals for Estimation of Force While Gripping," in *Advances in Engineering Design*, 2019, pp. 585–593.
- [15] R. J. Urbanowicz, M. Meeker, W. La Cava, R. S. Olson, and J. H. Moore, "Relief-based feature selection: Introduction and review," *Journal of Biomedical Informatics*, vol. 85. Academic Press Inc., pp. 189–203, Sep. 01, 2018, doi: 10.1016/j.jbi.2018.07.014.
- [16] J. Zhang, M. Chen, S. Zhao, S. Hu, Z. Shi, and Y. Cao, "RelieFF-based EEG sensor selection methods for emotion recognition," *Sensors (Switzerland)*, vol. 16, no. 10, pp. 1–15, 2016, doi: 10.3390/s16101558.
- [17] Y. Paul, V. Goyal, and R. A. Jaswal, "Comparative analysis between SVM & KNN classifier for EMG signal classification on elementary time domain features," in *4th IEEE International Conference on Signal Processing, Computing and Control, ISPC 2017*, 2018, pp. 169–175.
- [18] P. Geethanjali, "Comparative study of PCA in classification of multichannel EMG signals," *Australas. Phys. Eng. Sci. Med.*, vol. 38, no. 2, pp. 331–343, 2015, doi: 10.1007/s13246-015-0343-8.
- [19] P. Phukpattaranont, S. Thongpanja, K. Anam, A. Al-Jumaili, and C. Limsakul, "Evaluation of feature extraction techniques and classifiers for finger movement recognition using surface electromyography signal," *Med. Biol. Eng. Comput.*, vol. 56, no. 12, pp. 2259–2271, 2018, doi: 10.1007/s11517-018-1857-5.
- [20] T. Tuncer and F. Ertam, "Neighborhood component analysis and relief based survival recognition methods for Hepatocellular carcinoma," *Phys. A Stat. Mech. its Appl.*, vol. 540, p. 123143, 2020, doi: 10.1016/j.physa.2019.123143.
- [21] A. Phinyomark, P. Phukpattaranont, and C. Limsakul, "Feature reduction and selection for EMG signal classification," *Expert Syst. Appl.*, vol. 39, no. 8, pp. 7420–7431, 2012.
- [22] C. Spiewak, M. Islam, A. Zaman, and M. H. Rahman, "A Comprehensive Study on EMG Feature Extraction and Classifiers," vol. 1, no. 1, pp. 1–10, 2018.

# Deep Learning for Sleep Disorders: A Review

Hepsiba. D

*Department of Biomedical Engineering  
Karunya Institute of Technology and Sciences,  
Coimbatore, Tamil Nadu, India  
[hepsiba@karunya.edu](mailto:hepsiba@karunya.edu)*

R.Jane Preetha Princy

*Department of Biomedical Engineering  
Karunya Institute of Technology and Sciences,  
Coimbatore, Tamil Nadu, India  
[janepreetha@karunya.edu](mailto:janepreetha@karunya.edu)*

L. D. Vijay Anand

*Department of Robotics Engineering  
Karunya Institute of Technology and Sciences,  
Coimbatore, Tamil Nadu, India  
[vijayanand@karunya.edu](mailto:vijayanand@karunya.edu)*

**Abstract**—Deep Learning is a revolutionary technology that has the potential to enhance the health care industry. Deep Learning interprets enormous numbers of unstructured data that humans would typically take decades to comprehend and process. This article discusses how Deep Learning is applied to predict different types of sleep disorders. It compares the various approaches made by different researchers, including the signals used, different algorithms, processing techniques, as well as their advantages and drawbacks. Sleep is an essential component. It is as much important as food, drinking and respiring, necessary for the normal physical and mental health maintenance. A multitude of serious medical conditions such as heart failure and diabetes have been linked to sleep deprivation and excessive sleep. As a result, it is important to monitor sleep using more feasible methods, in order to diagnose and distinguish sleep disorders. While there are several difficulties ahead, self-monitoring the sleeping routine could help to resolve this issue.

**Keywords**—Deep Learning, predictive analysis, sleep disorders, mental health.

## I. INTRODUCTION

Sleep is a significant determinant of mental wellbeing and happiness. Troubled sleep is a common complaint all over the world. According to the American Sleep Association (AMA), sleep disorders affect 50 to 70 million adults in the United States of America [1]. Insomnia, sleep apnea, restless legs syndrome, and narcolepsy are the widely known forms of sleep disorders. The most predominant manners of sleep are Rapid Eye movement (REM), sleep and Non-REM sleep. REM sleep first begins about one and a half hours after falling asleep. The eyes switch rapidly behind shut eyelids from side to side. The respiration is quicker; the pulse and heart rate rise to waking levels. Most of the dreams take place during REM sleep, while some could occur in Non-REM sleep too [2]. The body passes through all stages of Non-REM sleep before entering into the REM. Each of the stage lasts for 5-15 minutes.

The individual would be either awake or in a state of very light sleep at the beginning of Non-REM [3] sleep. The second stage is marked by a deeper sleep, where the body's temperature and pulse rate decreases. Slow wave sleep, also known as delta sleep, is the third cycle. While the muscles are relaxed, the flow of blood to the muscles increases, and the body restores and builds up tissues. Hormones are released, and energy is recharged [4].

An individual having difficulty in sleeping is said to have insomnia. Frequent nocturnal awakenings or extended period of wakefulness, prolonged sleep latency are taken as the evidences of insomnia [5]. Sleep apnea is an under-diagnosed illness, with a lot of negative outcomes on patient's prosperity and wellbeing [6]. Daytime drowsiness, loud wheezing, and restlessness are the outcomes of sleep apnea. Polysomnography (PSG) is used to identify sleep apnea. Sleep apnea, if left untreated, might cause high blood pressure, depression, heart failure, and possibly death. [7]. Restless Leg Syndrome (RLS) is a neurological condition that is most prominent in women and older adults. It is mostly associated with pregnancy, Uremia and Iron deficiency. The clinical observation of RLS [8] follows a pattern which is similar to the cardiac rhythm and Periodic Limb Movements [9]. Narcolepsy is a chronic sleep disorder which results in excess daytime sleepiness. There are two distinct groups, Type-1 narcolepsy accompanies an unforeseen loss of muscle tone that causes shortcoming and makes you incapable to control your muscles (cataplexy), and Type-2 narcolepsy is without cataplexy [10]. Artificial intelligence could significantly improve sleep research; improve advance disorder diagnosis and treatments. It could provide high-precision outcomes and offer personalized monitoring and care [11].

## II. RELATED WORK

Jarchi [12] and colleagues proposed a Deep Learning algorithm for detecting sleep disorders. The study examines respiratory and movement-related sleep disturbances using ECG and EMG signals. To extract the features from the ECG, signal processing techniques are used. The derived features from ECG and EMG signals are integrated into a Deep Learning architecture using Synchro-squeezed Wavelet Transform (SSWT). The dataset contains records from 40 subjects, with ten of each type of subject: healthy subjects, Restless Leg Syndrome (RLS) subjects, OSA subjects, and both RLS and OSA subjects. The efficiency of various algorithms is measured, and a mean accuracy of 72 percent is achieved. The research also seeks to involve a wider range of biosignals.

Yadollahi [13] et al. designed a tracheal movement-based Deep Learning model for estimating sleep apnea. In addition to PSG, a wearable device mounted over the subject's suprasternal notch records tracheal movements and could be used to detect sleep apnea. This study, which was conducted at Toronto Sleep Laboratory in Canada, included 69 participants. To determine

the events and compute the Apnea Hypopnea Index, supervised Deep Learning classifiers are developed and applied. The accuracy of the proposed was 84 percent. In future, the research plans to use this approach in home-based monitoring systems.

Lee [14] devised a Deep Learning method for detecting sleep apnea by analysing ECG data. The six Deep Learning approaches used are Deep Neural Network (DNN), One Dimensional (1D) CNN, Two Dimensional (2D) CNN, RNN, LSTM, and GRU. The PSG dataset from 86 patients is used, and 63,441 events from 69 subjects were used for training phase and 11,670 events from 17 subjects were used for testing phase. The efficiency of 1D CNN and GRU was found to be superior to that of other algorithms. They are suitable for automatic identification using ECG and other physiological time series signals. Limitations such as a distorted ECG signal and a limited number of data samples could be solved.

Applying Deep Learning methods, Hafezi et al [15] demonstrated the severity of sleep apnea from respiratory associated swings. A lightweight, energy- and cost-efficient portable accelerometer-based Wearable monitoring system mounted to the trachea is used to estimate sleep apnea. The gestures are captured. A Deep Learning model is designed to extract features from respiration, and to determine the Apnea Hypopnea Index (AHI), which is the standard indicator for sleep apnea. It is then compared to the gold standard, polysomnography. Three Deep Learning models are used; Convolutional Neural Network (CNN), Recurrent Neural Network (RNN) with long short term memory (LSTM), and CNN+LSTM and their performances are compared. The CNN+LSTM architecture outperformed other models due to its strong association with PSG-derived AHI values. A few drawbacks, such as an unbalanced dataset, poor accelerometer signal quality, and spikes triggered by movement, could be overcome.

The implementation of Deep Learning for the purpose of applying facial depth to diagnose obstructive sleep apnea (OSA) was postulated by Islam et al.[16]. In comparison to a 2D colour picture, the depth map of human facial scans provides additional information about facial morphology. The subjects who went to Genesis Sleep care provided the 3D scans and sleep data. The ECU Human Research Ethics Committee approved this study, which included 39 males and 30 females. For transfer learning and facial recognition, they used VGG Face, Pose aware model (PAMs) VGG19 model, and PAMs-Alex model. Among the models, VGG Face performed the best by delivering an accuracy of around 69 percent. Further, the study aims to employ this model on larger datasets.

Kim et al. [17] used Deep Learning to build an interactive method for diagnosing sleep disorders. The EOG signal, which is an estimation of electrical activity in the eye to indicate sleep disturbances, is utilized in the study. Wake status, sleep status (from S1 to S4), and REM are the six scores used to diagnose the disorders. The physionet cap sleep database is used for analysis. Advanced Deep Learning methods are used to determine GRU in this automated model of sleep disorder prediction. Its immersive process of gathering input from doctors and practitioners also allows for the correction of errors and increased prediction accuracy.

A Deep Learning approach on insomnia Diagnosis was done by Shahin et al. [18]. At the Interdisciplinary Sleep Center of Charité University Hospital in Berlin, data from 41 controls and 42 insomnia patients were approved for PSG assessment. Deep Learning is applied to separate insomnia subjects from controls with no sleep problems using 57 extracted EEG attributes from two EEG passages. Stage based and stage independent EEG data obtained from insomnia influenced sleep stage and normal sleep respectively. The classifiers DNN and DNN-HMM were used. The findings demonstrated that Deep Learning would benefit in the diagnosis of insomnia.

Pathinarupothi [19] et al. adopted Deep Learning techniques to estimate the sleep apnea extremity using instantaneous heart rate (IHR). IHR is the equivalent of each minute's inter-beat time. The Deep Learning method LSTM-RNN (Long short term memory recurrent neural network) is used. Computing in cardiology sleep apnea challenge is taken as database. It is comprised of 8-hour ECG readings collected from 35 sleep apnea patients. A physionet toolkit converts ECG data into equivalent IHR readings. The findings show that the IHR function, which has a strong feature length, feature selection, and the intellectual capacity of LSTM+RNN to spot out high-dimensional patterns, could detect sleep apnea with exceptional precision.

Applying Deep Learning architecture, Xia [20] et al. conducted an EOG-based sleep analysis. This research examines at 20 people who are between the ages of 25 and 34. The sleep phases are classified using a Deep Belief Network (DBN) and a combination of DBN and Hidden Markov's Model (HMM). Among the two, DBN-HMM surpassed the other, achieving an accuracy of 83.3 percent. Future studies would focus on extracting features in the time domain along with combining EOG signals with other signals to better distinguish between different sleep stages.

Giri [21] used EOG, EEG, and EMG signals to classify sleep phases. The data was collected from physionet, and it was from a single night of sleep monitoring. Various processing techniques are used in the analysis, including the relative power feature extraction method, median, standard deviation, correlation, kurtosis, and correlation. Different classifiers are used, including SVM, Neural Networks, Classification Tree, KNN, and Naive Bayes. The best performing algorithm among them is the neural network, which has a high accuracy score and a short computation time. Future research will concentrate on using more complex neural network architecture, such as CNN.

Tan [22] et al. developed a Deep Learning framework that could detect sleep spindles. The EEG signals of 30 young people aged 20 to 23 are studied. The crowdsourcing method is used to identify sleep spindles from EEG samples. The Deep Belief Network's efficiency is compared to that of other classifiers such as Decision Tree, KNN, and SVM. While all of the classifiers did better, DBN outperformed the others by a factor of three. Prior applying for clinical supervision, the study should be tested on a wider population.

TABLE I  
KEY FINDINGS AND INTERPRETATION

S.No	Author(s)	Signals used	Algorithms used	Processing Techniques	Advantages	Disadvantages
1	Jarchi [12] et al. (2020)	ECG and EMG	Multilayer Perceptron, SVM, LSVM, Random Forest, KNN, XGB and Autokeras	SSWT Transform	Accurate and reliable.	Lacks wider range of bio-signals.
2	Yadollahi [13] et al. (2020)	Tracheal movements	CNN+LSTM+ Fully connected Layer	Band pass Filter	Convenient, high precision and reliable model.	The study was performed in a controlled environment.
3	Lee[14] et al. (2019)	ECG	DNN, 1D CNN, 2D CNN,RNN, LSTM, and GRU	FIR bandpass filter	Automatic detection of Sleep Apnea events	Lack of elaborated study of Sleep Apnea.
4	Hafezi[15] et al. (2019)	Accelerometer-based portable system for recording tracheal movements.	CNN, RNN+LSTM, CNN+LSTM	Butterworth bandpass filter	Increased accuracy and robustness.	Imbalanced dataset and movement artifacts.
5	Islam[16] et al. (2018)	Depth map of human facial scans	VGG Face, PAMs-VGG19 and PAMs-AlexNet	Image processing is done in Artec studio and Meshlab.	Identifies sleep apnea by just scanning the face.	Working on limited dataset is seen as an drawback.
6	Kim[17] et al. (2018)	EOG	GRU	T hour-long window is shifted as by T/10 hours	Automatic and less time consuming.	The model misinterprets and is unreliable
7	Shahin[18] et al. (2017)	EEG	DNN, DNN-HMM	Butterworth Filters	A deeper perception of insomnia's severity	Lacked scope of other sleep disorders
8	Pathinarupothi[19]et al. (2017)	Instantaneous heart rate	LSTM-RNN	ECG signals are converted to IHR by using scikit-learn library operations.	Large-scale and economic sleep apnea detection is possible.	The framework has not yet been applied in hospitals.
9	Xia [20] et al. (2015)	EOG	DBN and DBN-HMM	DBN extracts features by itself.	Time-saving and higher-scoring results.	Failed to effectively distinguish other sleep stages.
10	Giri [21](2015)	EOG, EMG, EEG	SVM, Neural Network, KNN, classification trees and Naïve Bayes	Relative power feature extraction method.	Good performance and efficiency.	The classifiers are unfit to automatically retrieve potential features.
11	Tan [22] et al. (2015)	EEG	Decision Tree, KNN, SVM, and DBN	Power spectrum density	Potential capacity to learn the sleep spindle's raw inner characteristics.	Study has be evaluated with wider range of population.

### III. DEEP LEARNING METHODS

Deep Learning [23] is a branch of Machine Learning that knows, forecasts, and categorizes data from unsupervised sources. Data scientists who are entrusted with gathering, processing, and deciphering vast volumes of data will find Deep Learning to be incredibly useful; it makes the process much quicker and simpler. Deep Learning networks have several layers,[26] which adds more parameters and makes it easier to perform more complex functions. The description of some of the most common Deep Learning algorithms is done.

#### A. Convolutional Neural Network

Convolution Neural Network performs the task of image classification, recognition of images, object detection etc. The CNN[24] model first trains and test the image one by one then send it through multiple convolution layers with kernels then to pooling process. From there it passes through fully connected layers and softmax function is applied to classify the image between 0 and 1 probabilistic values. The convolutional layer splits the image into  $h \times w \times d$ . Reduction of excess parameters in an image is done by pooling layer. The matrix is converted to a vector and fed to a completely connected layer, where the features are combined to construct a model, and the output is classified using the softmax activation function. In [14] 1D and 2D CNN models were used to automatically predict sleep apnea. 1D model had 3 kernels of the sizes  $50 \times 1$ ,  $30 \times 1$  and  $10 \times 1$  with  $1 \times 2$  pooling layer. 2D CNN model had  $50 \times 2$ ,  $30 \times 2$  and  $10 \times 2$  kernels with  $2 \times 2$  max pooling regions and an extra convolutional layer than 1 D model.

#### B. Recurrent Neural Network

Recurrent Neural Network [25] recollects the things which they have learnt from previous inputs, and also they remember things that they learn during the training. RNN's are capable of taking in more than one input and producing one or more outputs. They are not only influenced by input weights applied on them but also by hidden state vectors which represents previous inputs and outputs. RNN is appropriate for sequential data and proved to be efficient for time series as it has good memory as well. LSTM and GRU models were used within RNN in [14]. Three layers of RNN with 60, 80, and 120 memory cells were used in each LSTM and GRU. This model is recognized as an effective model for automatic identification of sleep apnea.

#### C. LSTM

Long Short Term Memory (LSTM) is a type of RNN [29]. It consists of Feedback network. It is restricted from processing single data points (such as images), rather it processes sequence of data such as video or speech. The unit of LSTM consists of the following, a cell, input gate, output gate and forget gate. The progression of data in the cell is monitored by gates, and the cell recollects the details over time. By inserting memory blocks in the recurrent hidden layer, the LSTM model is used to solve the problem of vanishing and bursting gradient in [19]. This 3-layer LSTM architecture consists of 61 input-layer neurons, two cell-

hidden-layer memory blocks and two neurons depicting 2 groups in the output-layer.

#### D. GRU

Gated Recurrent Unit (GRU)[26] is a gating mechanism which is more like LSTM but lacks a few parameters like output gate. GRU solves the vanishing gradient problem. It possesses two gates, an update gate and a reset gate, that determines what data should be sent to the output. They are best suited for filtering inappropriate information in prediction, and in storing prolonged information. They could perform skillfully well in complex situations with careful training. In [14] GRU is employed in minimizing the number of calculations per training period. It gave a robust performance and found to be a suitable model for automatic detection of SA from ECG signals.

### IV. PROPOSED APPROACH

In the existing studies, a patch had been placed over the suprasternal notch. It records the neck and respiratory movements and identifies the possibilities of sleep apnea. The patch has a 3D accelerometer to record the movements in X, Y and Z dimension. Although the above said method achieved high accuracy in estimating AHI, the size of the patch seems to be inconvenient, and may cause hindrance to the subject.

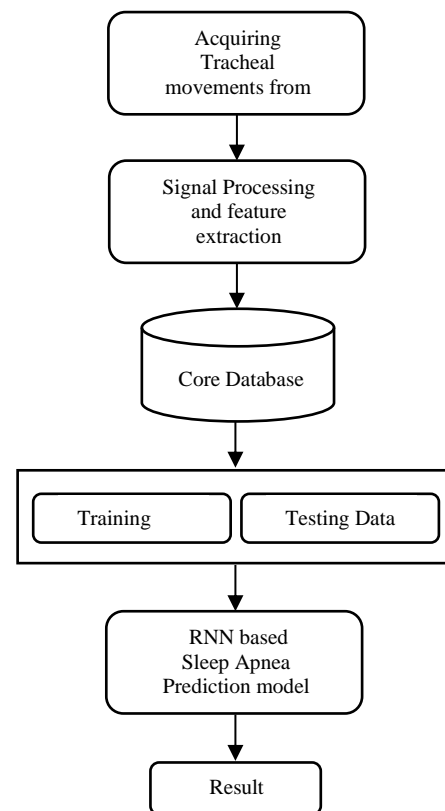


Fig. 1. Proposed Approach

Instead of conventional sensors, we put forth the idea of using MEMS sensor. MEMS sensor are tiny, less power consuming



and more sensitive to variations. The application of Tri-Axial Piezoelectric MEMS sensor could possibly enhance the quality of signal acquisition with reduced artifacts [28]. It has a total of 16 distributed sensing elements in  $7 \times 3.5$  mm area cross section. This model would offer high sensitivity and linearity compared to the traditional accelerometer sensor. The acquired movements are stored as a time-series data. It is then processed, and necessary morphological features are extracted. The robust Deep Learning algorithm, Recurrent Neural Network is applied to get better prediction of sleep apnea.

#### V. CONCLUSION AND FUTURE WORK

Deep Learning technology identifies and classifies sleep stages effectively. The sleep monitoring systems presents detailed information about the subject's quality of sleep, thereby assisting the physicians to offer appropriate remedies and treatment. The future work aims at developing the proposed model, which could possibly revolutionize sleep related research. The MEMS sensor-based approach could detect even the slightest change in events. Moreover, it could be applied for remote and home based monitoring systems. MEMS sensors are inexpensive to manufacture, making them accessible to people of all socioeconomic backgrounds. MEMS should be used to effectively acquire other signals such as EEG, ECG, PCG, and so on, in addition to tracheal movements.

#### REFERENCES

- [1] Rita brooks, "How to diagnose & treat the 5 most common sleep disorders" July 24th, 2017, <https://www.aastweb.org/blog/how-to-diagnose-treat-the-5-most-common-sleep-disorders>
- [2] John Peever and Patrick M. Fuller, "The Biology of REM Sleep" NCBI, March 2018, doi: 10.1016/j.cub.2015.11.011
- [3] Mohamed H. Al-Meer; MD Abdullah Al Mamun, "Deep Learning in Classifying Sleep Stages", Thirteenth International Conference on Digital Information Management (ICDIM), IEEE, September 2019, DOI: 10.1109/ICDIM.2018.8846973.
- [4] Zeman, Adam & Reading, Paul. (2005). The science of sleep. Clinical medicine (London, England). 5. 97-101. 10.7861/clinmedicine.5-2-97.
- [5] Thomas Roth, "Insomnia: Definition, Prevalence, Etiology, and Consequences", Vol:3(5 Suppl), Aug 15, 2007. <https://www.ncbi.nlm.nih.gov/pmc/journals/368/>
- [6] Osman, Amal M et al. "Obstructive sleep apnea: current perspectives." *Nature and science of sleep* vol. 10 21-34. 23 Jan. 2018, doi:10.2147/NSS.S124657
- [7] Kevin K. Motamedi, Andrew C. McClary, and Ronald G. Amedee, "Obstructive Sleep Apnea: A Growing Problem" Vol 9(3), pg: 149-153, 2009 <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3096276/>
- [8] Kwatra, Vishal et al. "Differential Diagnosis and Treatment of Restless Legs Syndrome: A Literature Review." *Cureus* vol. 10,9 e3297. 13 Sep. 2018, doi:10.7759/cureus.3297
- [9] Shiyi Guo, Jinsha Huang, Haiyang Jiang, Chao Han, Jie Li, Xiaoyun Xu, Guoxin Zhang, Zhicheng Lin, Nian Xiong, and Tao Wang, "Restless Legs Syndrome: From Pathophysiology to Clinical Diagnosis and Management", doi: 10.3389/fnagi.2017.00171 June, 2, 2017, <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5454050/>
- [10] Gbolagade Sunmaila Akintomide and Hugh Rickards, "Narcolepsy: a review", doi: 10.2147/NDT.S23624, September 2008, <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3173034/>
- [11] Kent, J. (2020, March 2). Artificial Intelligence May Boost Sleep Disorder Treatment, Diagnosis. HealthITAnalytics. <https://healthitanalytics.com/news/artificial-intelligence-may-boost-sleep-disorder-treatment-diagnosis>
- [12] Jarchi, D., Andreu-Perez, J., Kiani, M., Vysata, O., Kuchynka, J., Prochazka, A., & Sanei, S. (2020). Recognition of Patient Groups with Sleep Related Disorders using Bio-signal Processing and Deep Learning. *Sensors*, 20(9), 2594. <https://doi.org/10.3390/s20092594>
- [13] Hafezi, M., Montazeri, N., Saha, S., Zhu, K., Gavrilovic, B., Yadollahi, A., & Taati, B. (2020). Sleep Apnea Severity Estimation From Tracheal Movements Using a Deep Learning Model. *IEEE Access*, 8, 22641-22649. <https://doi.org/10.1109/access.2020.2969227>
- [14] Urtnasan Erdenebayar, Yoon JiKim, Jong-UkPark, Eun YeonJoo, Kyoung-JoungLee, "Deep learning approaches for automatic detection of sleep apnea events from an electrocardiogram" *Computer Methods and Programs in Biomedicine*, Volume 180, October 2019.
- [15] Maziar Hafezi, Nasim Montazeri, Kaiyin Zhu, Hisham Alshaer, Azadeh Yadollahi, Babak Taati, "Sleep Apnea Severity Estimation from Respiratory Related Movements Using Deep Learning" 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), October 2019
- [16] Syed MS Islam, Hassan Mahmood, Adel Ali Al-Jumaily, Scott Claxton. "Deep Learning of Facial Depth Maps for Obstructive Sleep Apnea Prediction", 2018 International Conference on Machine Learning and Data Engineering, IEEE, 2019, DOI 10.1109/iCMLDE.2018.00036.
- [17] Woonghee Lee and Younghoon Kim, "Interactive Sleep Stage Labelling Tool For Diagnosing Sleep Disorder Using Deep Learning", Annual International Conference of the IEEE Engineering in Medicine and Biology Society, IEEE, July 2018
- [18] Mostafa Shahin, Beena Ahmed, Sana Tmar-Ben Hamida, Fathima L. Mulaffer, Martin Glos, and Thomas Penzel. "Deep Learning and Insomnia: Assisting Clinicians with their Diagnosis" *Biomedical and Health Informatics*, IEEE, 2019, DOI 10.1109/BHJ.2017.2650199
- [19] Rahul K. Pathinarupothi, Vinaykumar R., Ekanath Rangan, Gopalakrishnan E., and Soman K. P. "Instantaneous Heart Rate as a Robust Feature for Sleep Apnea Severity Detection using Deep Learning" *IEEE EMBS International Conference on Biomedical & Health Informatics (BHI)*, April 2017, **Doi**:10.1109/BHI.2017.7897263
- [20] Bin Xia et al., "Electrooculogram based sleep stage classification using deep belief network," 2015 International Joint Conference on Neural Networks (IJCNN), Killarney, Ireland, 2015, pp. 1-5, doi: 10.1109/IJCNN.2015.7280775.
- [21] E. P. Giri, A. M. Arymurthy, M. I. Fanany and S. K. Wijaya, "Sleep stages classification using shallow classifiers," 2015 International Conference on Advanced Computer Science and Information Systems (ICACSIS), Depok, Indonesia, 2015, pp. 297-301, doi: 10.1109/ICACSIS.2015.7415162.
- [22] D. Tan, R. Zhao, J. Sun and W. Qin, "Sleep spindle detection using deep learning: A validation study based on crowdsourcing," 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Milan, Italy, 2015, pp. 2828-2831, doi: 10.1109/EMBC.2015.7318980.
- [23] A. Shrestha and A. Mahmood, "Review of Deep Learning Algorithms and Architectures," in *IEEE Access*, vol. 7, pp. 53040-53065, 2019, doi: 10.1109/ACCESS.2019.2912200.
- [24] S. Albawi, T. A. Mohammed and S. Al-Zawi, "Understanding of a convolutional neural network," 2017 International Conference on Engineering and Technology (ICET), Antalya, 2017, pp. 1-6, doi: 10.1109/ICEngTechnol.2017.8308186.
- [25] N. M. Rezk, M. Purnaprajna, T. Nordström and Z. Ul-Abdin, "Recurrent Neural Networks: An Embedded Computing Perspective," in *IEEE Access*, vol. 8, pp. 57967-57996, 2020, doi: 10.1109/ACCESS.2020.2982416.
- [26] J. Kim, J. Kim, H. L. Thi Thu and H. Kim, "Long Short Term Memory Recurrent Neural Network Classifier for Intrusion Detection," 2016 International Conference on Platform Technology and Service (PlatCon), Jeju, 2016, pp. 1-5, doi: 10.1109/PlatCon.2016.7456805.
- [27] R. Fu, Z. Zhang and L. Li, "Using LSTM and GRU neural network methods for traffic flow prediction," 2016 31st Youth Academic Annual Conference of Chinese Association of Automation (YAC), Wuhan, 2016, pp. 324-328, doi: 10.1109/YAC.2016.7804912.

- [28] Liu, Y., Hu, B., Cai, Y., Liu, W., Tovstopyat, A., & Sun, C. (2021). A Novel Tri-Axial Piezoelectric MEMS Accelerometer with Folded Beams. *Sensors*, 21(2), 453. <https://doi.org/10.3390/s21020453>

# Arduino and LabVIEW based Heart Rate Detection System

Deepjyoti Mahanta<sup>1</sup>  
Assam Don Bosco  
University  
Guwahati  
deepjyoti94@yahoo.com

Hemashree Bordoloi<sup>2</sup>  
Assam Don Bosco  
University  
Guwahati  
hemashree.bordoloi@dbuniversity.ac.in

Samar Jyoti Saikia<sup>3</sup>  
Assam Don Bosco  
University  
Guwahati  
samar\_saikia@yahoo.com

**Abstract**— This Paper contains the whole process of ECG signals Acquisition from ECG Sensor to its analysis using LabVIEW and Biomedical Workbench. The Sensor data of ECG has the amplification, filtering and conversion of analog ECG information to digital by using Arduino Uno. The acquisition part deal with acquiring the hardware information to analyzablefile format into PC. Here 6-channel ADC in Arduino Uno with LabVIEW interface is employed for conversion. Now the acquired ECG information is processed and analyzed with biomedical workbench that gives the various features of ECG signal processing. This system is very simple to implement and cost effective.

**Keywords**— LabVIEW, ECG Signal Processing Tools, ECG Analysis, Biomedical Workbench, Arduino.

## I. INTRODUCTION

Electrocardiogram (ECG) is employed to measure and monitor the heart electrical activities throughly from many years. These electrical details are used to diagnosis the heart conditions These electrical details are used to diagnosis the heart conditions of the subject From centuries to till now there is many advanced hardware and software package tools are developed for Electrocardiogram signal acquisition and analysis [1].The ECG signal is the graphical representation of heart electrical activities within the form of voltage and current generated throughout the cardio muscles contraction and relaxation. The generated voltage/current is extremely little in magnitude and these can be measure from the body skin surface by placing the appropriate ECG electrode. These cardio signals frequency is in between 0.05 Hertz to 100 Hertz.[9] [10]

The electrocardiogram tests in hospitals are increasing with time. However modern ECGs produce digital output, but still plain paper is in use to record the ECG information. Typically ECG information of patient become necessary to transfer at another distance place for analysis and paper based data is simply too much time consuming and also tough to have record of patient information for long period. So it is requirement of present time to have the information in digital form in numerous analysable file formats [2][11] [12][13].

We described the entire ECG information acquisition method from hardware to further signal processing software package tools in computer system. Hardware having the many stages from ECG signals amplification, filtering, conditioning to analog to digital conversion and software packages having the real time plotting of ECG signal in LabVIEW and save the plotted information for further usable digital file formats like .txt, .tdms, .tdm, .xlsx for required time duration. The additional Biomedical workbench uses the files to analysis of

the ECG data; this is having ECG feature extraction and Heart

Rate Variability Analyzer tools because the core requirement of recent study for analysis. The main objective of the project work is to design the ECG system which could facilitate the researchers and doctors to acquire and analyse the ECG information thoroughly with simple and cost effective tools in very less time.[14] [15] [16] [17] [18]

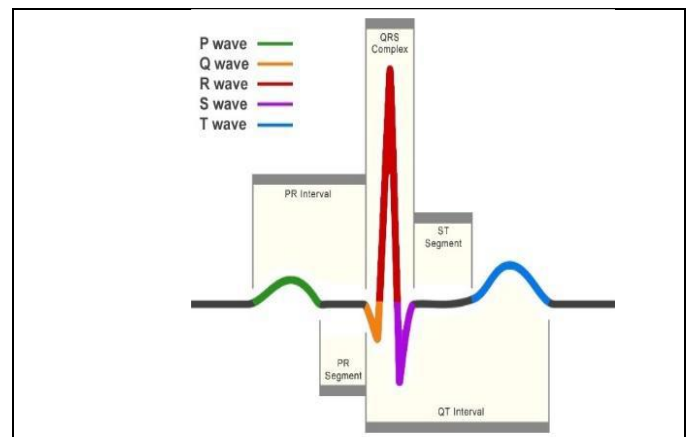


Fig.1. Schematic representation of normal ECG

## II. SYSTEM DESIGN OVERVIEW

This Complete ECG System will be understood by simple block diagram having different stages. The Block Diagram of the system is shown in fig 2.The physiological parameter ie; Electrocardiogram (ECG) of the subject are measured using ECG AD8232 sensors and Ag-AgCl electrodes and their output is processed through a microcontroller. The Microcontroller actually performs the function of a Data Acquisition and simply serves the signals to LabVIEW. The Signal is processed through the LabVIEW Virtual Instruments (VI). The output is displayed in the front panel of LabVIEW.

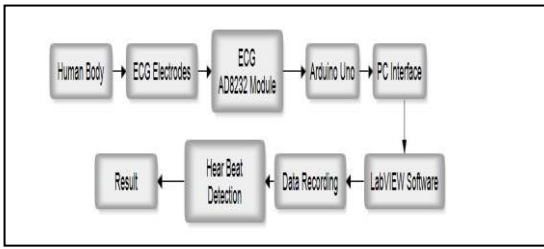


Fig.2. System Model of the proposed work



Fig.4: ECG Sensor with Ag-AgCl electrodes and 3 Pin connector cable

### III. HARDWARE REQUIRED

#### A. Arduino Uno

As we know that the Arduino Uno is one of the most commonly used microcontroller board based on the ATmega328. It has some unique features compare to other microcontroller like it has 14 digital input/output pins - 6 analog inputs, a 16 MHz oscillator, a USB connection, a power jack and a reset button, out of which 6 pins is used as PWM outputs [6].



Fig.3. Arduino Uno

In this Experiment Arduino Uno is simply used as a Data Acquisition to communicate the input signals to the LabVIEW. The Signals are extract and analyse using the LabVIEW Virtual Instruments

#### B. ECG Sensor

AD8232 ECG Sensor Module [3] is employed to measure the electrical activity of the heart. The AD8232 is an integrated signal conditioning Sensor for ECG measurement and related applications. It measures the Electro Cardiogram readings of the subject through three electrodes connected to the Subject's body [4] [7].

### IV. SOFTWARE PLATFORM IMPLEMENTATION

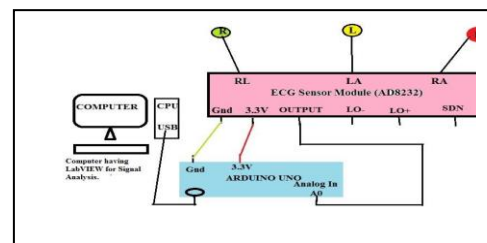
#### A. LabVIEW

LabVIEW - *Laboratory Virtual Instrument Engineering Workbench*, is an interactive programming environment within which program is created using a graphical notation. The programming language employed in LabVIEW is a dataflow programming language within which the execution is decided by the structure of a graphical block diagram. The programmer can connect different function-nodes by drawing wires or lines. These wires or lines propagate data as variables and a node can execute when all its input data is obtainable. This graphical is incredibly advantageous because it allows the nonprogrammers to form programs by dragging and dropping virtual representations of lab equipment in LabVIEW. One among of the extra features of LabVIEW is that, it includes extensive support for interfacing various devices, instruments and even cameras [8].

### V. IMPLEMENTATION OF THE PROPOSED SYSTEMS

#### A. Extraction and Analysis of ECG Signal

The objective of the work is to extract the ECG signal from different subjects and the raw ECG signal fed into the ECG sensor module which filters and amplifies it in presence of noisy conditions. The output of the sensor module is fed into a computer having LabVIEW using Arduino uno. Using LabVIEW, proper analysis of the extracted signal is done. Various real time data from different subjects are taken in order to analyse them by comparing them with an ideal data. The Schematic diagram of the system is shown in Fig5.



82 Fig.5: Schematic diagram of the System for analysis of ECG signal Using

### B. Electrode Placement

Einthoven's triangle is an imaginary formation of three limb leads in an exceedingly triangle employed in electrocardiography, formed by the two shoulders and also the pubis.. It is named after Willem Einthoven, who theorized its existence. It is a very Popular theorem.

Fig 6 shows the Placement of Electrodes using Einthoven's Triangle.

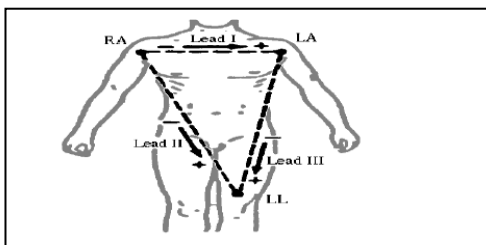


Fig.6: Placement of Electrodes

### C. Signal Conditioning and Processing using LabVIEW

After the signal is extracted from a particular subject using Silver-Silver Chloride electrode, the signal is interfaced to the Computer with LabVIEW using Arduino Uno where it will be further processed. Fig.7 shows the Flowchart for the observation of the heart rate. The normal heart rate is 60-100 bpm. This value is considered as the standard value for comparison. In this work of detecting the heart beat, if the heart beat is below 60 then it is shown as the Bradycardia with the help of an indicator in the LabVIEW. And if the heart beat is greater than 100 bpm then it is detected as tachycardia.

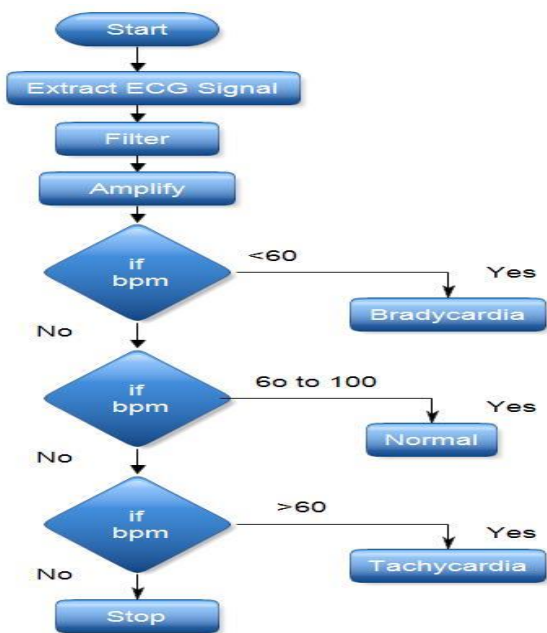


Fig.7: Algorithm for the detection of Heart Rate.

Fig.8 shows the back panel of LabVIEW for the detection of heart rate. Fig.9 shows the front panel of LabVIEW for the detection of heart rate.

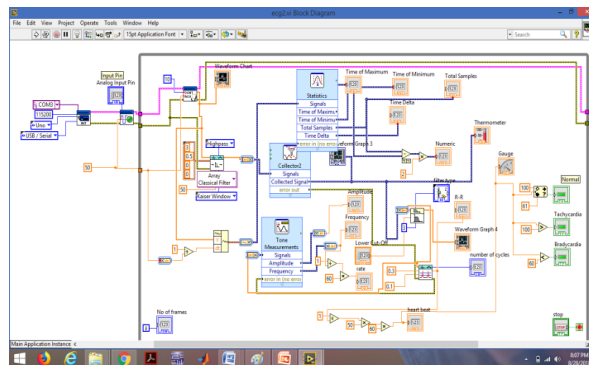


Fig.8: The back panel of LabVIEW for the detection of Heart Rate.



Fig.9: The front panel of LabVIEW for the detection of Heart Rate.

### D. Experimental Results

An experimental set-up has been shown in the figure 10 while extracting the ECG signal from a subject.



Fig.10: Experimental Set-Up.

The ECG signals that we have extracted from various subjects with the help of electrodes and then processed them in LabVIEW using Arduino are as follows:



Fig.11: Subject-1, Female, Age: 32.

**Comment 1:** From the extracted ECG signal as shown in Fig11, we have come to a point that the subject-1 having tachycardia as it has irregular heart rhythm which is greater than 100.



Fig.12. Subject-2, Female, Age: 76

**Comment 2:** From the extracted ECG signal as shown in Fig12, we have come to a point that the subject-2 having Normal as it has regular heart rhythm of 93 which is in between 60 and 100

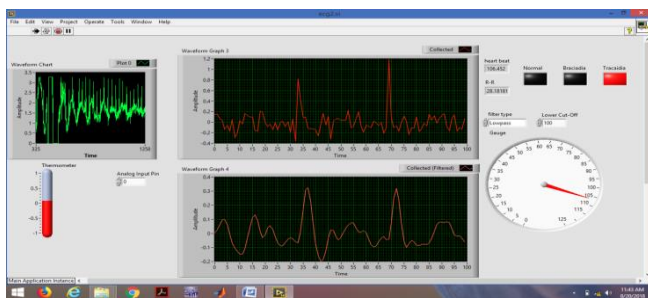


Fig.13. Subject-3, Male, Age: 30

**Comment 3:** From the extracted ECG signal as shown in Fig13, we have come to a point that the subject-3 having tachycardia as it has irregular heart rhythm which is greater than 100.



Fig.14. Subject-4, Male, Age: 18

**Comment 4:** From the extracted ECG signal as shown in Fig14, we have come to a point that the subject-4 having Normal as it has regular heart rhythm of 83 which is in between 60 and 100.

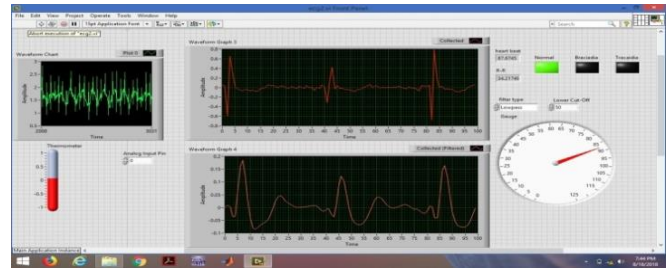


Fig.15. Subject-5, Male, Age: 20

**Comment 5:** From the extracted of ECG signal as shown in Fig15, we have come to a point that the subject-5 having Normal as it has regular heart rhythm of 87 which is in between 60 and 100.

The ECG signal has been extracted by means of Ag-AgCl electrodes by placing it around the chest of the subject and it has been processed in the LabVIEW to detect the heart rate in beats per minute. Front panel visualizes the detection of heart rate.

## VI. CONCLUSION

Hardware and LabVIEW software both together creates real time ECG waveform. This real time electrocardiogram waveform hold on a digital file at required time duration and sampling rate. Biomedical workbench of National Instrumentation have very efficient tools for ECG signal processing, feature extraction and vital sign (heart rate variability) analysis. All the analysis techniques described above are very advanced and extensively utilized by researchers This complete process from getting signal to its heart rate variability analysis is very easy and can be used for self-diagnosis.

## REFERENCES

- [1] Raja Brij Bhushan Et.al, "Development of simultaneous quantitative ECG system," National Conference on Biomedical Engineering, April 21-22, 2000, pp. 39-51
- [2]. Braunwald E. (Editor), Heart Disease: A Textbook of Cardiovascular Medicine, 5<sup>th</sup> Edition, p. 108, Philadelphia, W.B. Saunders Co., 1997. ISBN 0-7216-5666-8.

[4] Joyanta Kumar Roy Et.al, "The Wearable Electronic Rescue System for Home Alone Elderly- Labview & Arduino Evaluation," *IOSR-JECE, Volume Eight, Issue -six (Nov. - Dec. 2013), PP 50-55.*

[5] M.M Et.al , Development of Cost Effective ECG Data Acquisition System for Clinical Applications using LabVIEW 2014 IEEE 10<sup>th</sup> International Colloquium on Signal Processing & its Applications (CSPA2014), 7 - 9 March. 2014, Kuala Lumpur, Malaysia

[6]<https://components101.com/microcontrollers/arduino-uno>.

[7] <https://learn.sparkfun.com/tutorials/ad8232-heart-rate-monitor-hookup-guide>.

[8] <http://www.ni.com/academic/students/learn-labview/execution-structures/>.

[9] Apoorv Gautam and Maninder Kaur, "ECG Analysis using Continuous Wavelet Transforms (CWT)", *IOSR Journal of Engineering*, Apr.2012, Vol.2 (4), 632-635.

[10] Nagendra H, S. Mukherjee and Vinod Kumar, "Application of Wavelet Techniques in ECG Signal Processing: An Overview", *International Journal of Engineering Science and Technology (IJEST)*, October 2011, Vol.3, No.10, 7432-7443.

[11] Adam Josko and Remigiusz J. Rak, "Effective Simulation of Signals for Testing ECG Analyzer", *IEEE Transactions on Instrumentation and Measurement*, June 2005, Vol.54, No.3, 1019-1024.

[12] K.V.L. Narayana, A. Bhujanga Rao," Wavelet based QRS detection in ECG using MATLAB", *Innovative Systems Design and Engineering*, 2011, Vol. 2, No 7, 60- 69.

[13] B. Anuradha, K. Suresh Kumar and V.C. Veera Reddy, "Classification of Cardiac signals using Time Domain Methods", *ARNP Journal of Engineering and Applied Sciences*, June 2008, Vol.3, No.3, 7-12.

[14] C. Saritha, V. Sukanya, and Y. Narsimha Murthy," ECG Signal Analysis Using Wavelet Transforms", *Bulg.J.Phys.*35, 2008, 68-77.

[15] Rajiv Ranjan, V.K Giri, "A Unified Approach of ECG Signal Analysis", *International Journal of Soft Computing and Engineering (IJSCE)*, July 2012, Volume-2, Issue-3, 5-10.

[16] A. Muthuchudar, Lt. Dr. S. Santosh Baboo, "A Study of the Processes Involved in ECG Signal Analysis", *International Journal of Scientific and Research Publications*, March 2013, Volume 3, Issue 3, 1-5.

[17] S. Karpagachelvi, Dr. M. Arthanari, M. Sivakumar, "ECG Extraction Techniques- A Survey Approach", *International Journal of Computer Science and Information Security*, April 2010, Vol .8.No .1, 76-80.

[18] A.K.M Fazlul Haque, "Improved Detection of ECG Features Using Wavelet for Emergency Medical Application", *IJAITI*, Mar/Apr 2012, Vol. 1, No. 2.

# Analysis of Tremors in Parkinson's Disease Using Accelerometer

Niya Romy Markose  
School of Electronics Engineering  
Vellore Institute of Technology  
Vellore 632014, India  
[niyaromyarkose@gmail.com](mailto:niyaromyarkose@gmail.com)

Priscilla Dinkar Moyya  
School of Electronics Engineering  
Vellore Institute of Technology  
Vellore 632014, India  
[priscilladinkar@gmail.com](mailto:priscilladinkar@gmail.com)

Mythili Asaithambi  
School of Electronics Engineering  
Vellore Institute of Technology  
Vellore 632014, India  
[mythili.asaithambi@vit.ac.in](mailto:mythili.asaithambi@vit.ac.in)

**Abstract**— Parkinson's Disease is one that affects the brain and symptoms include shaking, stiffness, and difficulty to walk. Tremor is the most identifiable symptom of this disease and it affects nearly 80% of the patients with Parkinson's Disease. This prototype was designed to observe and quantify the tremor signal from Parkinson's disease patients. The prototype is based on Arduino Uno programming and interfacing, and the ADXL335 tri-axial accelerometer is used as a sensor. The resting tremor signal was acquired in the form of acceleration using the sensor accelerometer from fingertip, wrist and forearm of the patient. The Arduino processed the data which was transferred to MATLAB for further processing. The resting tremor was observed in terms of amplitude and spectral density. For the three parts considered, the amplitude values of acceleration were ranging from 40 dB/Hz- 80 dB/Hz and spectral density were observed and compared. Hence this basic prototype could be useful and developed further to assist the Parkinson's Disease patients.

**Keywords**— Accelerometer, ADXL335, Arduino Uno, Parkinson's disease, Tremor

## I. INTRODUCTION

Parkinson's disease is a disorder that affects the central nervous system which in turn has a deteriorating effect on the movements of a person [1]. According to the US Census Bureau, there were 680,000 Parkinson's disease patients in US, and it is estimated to rise to 1,300,000 in the year 2030 [2]. Although, the chance of developing Parkinson's disease increases with age, there are people who are less than 50 years of age who suffer from Parkinson's disease [3]. Symptoms of this disease develop when the dopamine level drops due to nerve cell damage and this the neurotransmitter that is involved is important for movement and coordination [4]. Symptoms include tremors, slow movement, and difficulty to speak [5]. The tremor usually starts in one hand.

Tremor as it is known, is the involuntary shaking of body parts, and it is one of the most observable symptoms that Parkinson's disease patients can develop [6]. Nearly 80% of the patients develop tremor. This tremor holds much significant medical or physiological information which can be considered for therapy [7]. Two important points to be considered are Firstly, a good tremor assessment is important to decide the differential diagnosis of Parkinson's disease [8]. Secondly, tremor assessment is another method which enables us to measure the severity of the disease and hence decide upon the treatments [9], [10]. It is important to understand that the tremor signals are highly variable in terms of frequency and distribution [11]. It varies from person to person, and it is different in different contexts [12].

The research done in this field discusses about studying the tremor using IMU (Inertial Measurement Units), administering the dopamine dosage to patients and

observing the changes that take place [6]. The severity of the disease can be observed by IMU by taking tremor signals. Parkinson's disease tremor can be used to discriminate between the postural tremor and resting tremor in terms of band limited energy and amplitude [13]. Some use this wearable device to acquire tremor in order to screen patients for the onset of the diseases.

The main objective of the project is to develop a wearable device that could acquire and analyze the tremor from Parkinson's disease affected patients as done previously [14]. This prototype is a combination of ADXL335 accelerometer and Arduino Uno. MATLAB is used to observe and process the acceleration signals acquired as in the previous research [15]–[17]. The device is a wearable, and the device is in the form of a glove. The acceleration reading was taken from three parts: - fingertip, wrist and the forearm. The code that detects acceleration from these parts was uploaded to the Arduino. The acceleration of three directions (X-Axis, Y-Axis and Z-Axis) were acquired, and the direction with the highest variation of the acceleration data values were selected. These values could be observed in the Arduino IDE. These values were taken and were uploaded to the MATLAB online in the form of a data file. The MATLAB program was used to observe the amplitude graph and the periodogram. Above procedures were repeated several times to decrease the interobservity.

## II. MATERIALS & METHODS

The main idea was to generate a device that acquires signals and observes them, so using a user interface or display were not quite necessary. A sensor which measures the motion and a processor that converts the signals observable in the personal computer are the most important components.

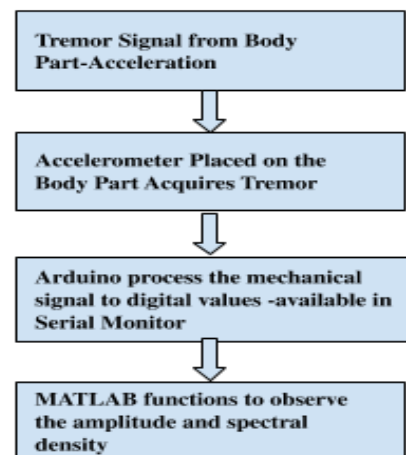


Fig. 1. Flowchart of the process involved



As shown in Fig. 1, the design of the prototype includes Arduino Uno, interfaced to ADXL 335 accelerometer. MATLAB is used to process the tremor signal to observe its amplitude and spectral density. The flowchart depicts the processes involved in acquiring the signals.

#### A. ADXL335 Accelerometer

ADXL335 sensor is a low power tri-axial accelerometer. It measures both static as well as dynamic acceleration. It has six pins, GND stands for Power Ground, and  $V_{cc}$  is a pin where the Regulated 5V DC power supply is to be provided. X\_OUT, Y\_OUT and Z\_OUT provide the acceleration values in the X, Y and Z axis. ST stands for Self-Test is an input pin. Additionally, it has a 3.3V voltage regulator. The programming and integration into the device are found to be compatible. This sensor is used to acquire the tremor from patients. The complete pin diagram is as shown in Fig. 2 [18].



Fig. 2. Pin diagram of ADXL335

#### B. Arduino Uno

Arduino Uno is an Atmega380 Chip based microcontroller. Applications of Arduino Uno include building electronics projects, it has 6 analog input pins, 14 pins which provide digital pins, 16 MHz ceramic crystal resonator, USB-B port, an ICSP header, power jack and reset button. The Atmega380 chip is the brain of the Arduino, it has 32K Flash Memory and 2K RAM [19].

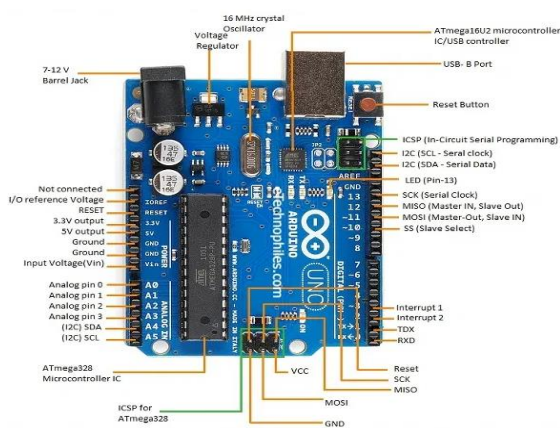


Fig. 3. Pin diagram of Arduino Uno

Pin diagram of Arduino Uno is shown in Fig. 3. Analog pins acquire the analog signals in the range 0-5. Digital pins can be used as input or output pins for digital signals. The resonator used to generate clock signals. Universal Serial Bus (USB) is used to connect the microcontroller board to the computer. The power jack supplies the microcontroller

with power. The In-Circuit Serial Programming (ICSP) has the ability of the microcontroller to be programmed without taking out of its circuitry. This ability is attributed to the presence of ICSP header on the Arduino board. Reset pin resets the microcontroller to the initial stage.

#### C. Hardware and Software Requirements

Arduino Uno, Accelerometer, Breadboard and Jumper Wire are required to design this device. Arduino IDE and MATLAB ONLINE 2020 are required to program the Arduino and process the tremor signal respectively.

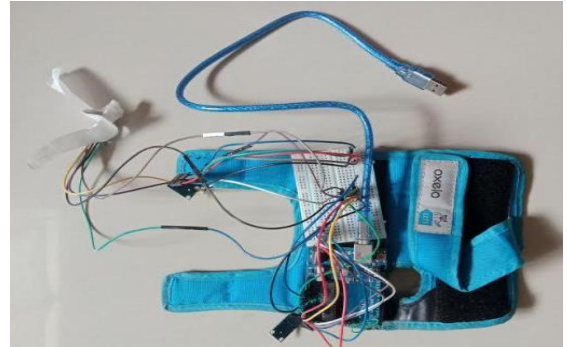


Fig. 4. The wearable Prototype

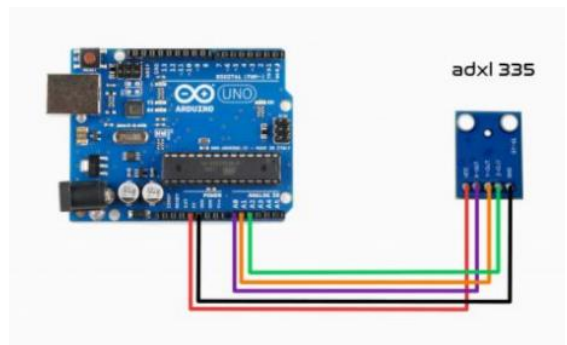


Fig. 5. Circuit connection between Arduino Uno and ADXL335

The  $V_{cc}$  of the sensor is connected to 5V of the Arduino Uno. The X-OUT, Y-OUT and Z-OUT pins of the sensor ADXL335 were connected to analog pins A0, A1 and A2 of the Arduino Uno. The ground pin of the Arduino Uno was connected to the Ground pin of the Sensor. The designed wearable prototype is shown in Fig. 4 and the circuit connection between Arduino Uno and ADXL335 is as shown in Fig. 5 respectively.

#### D. Acquiring the Tremor Signal

The glove has to be placed, worn in a stable and appropriate position so that USB cable can be attached to the laptop as shown in Fig. 6.

Also, the sensors have to be placed in appropriate position so that signal can be acquired with ease, and can be used with less processing. Using three accelerometers the acceleration values in  $m/s^2$  are acquired. Uploading the file into the MATLAB will help us to process the tremor value in order to observe the amplitude of the tremor signal in visual form like amplitude graphs and periodogram, in which the spectral density was observed in the form of graphs. The sampling frequency was chosen as 25Hz.

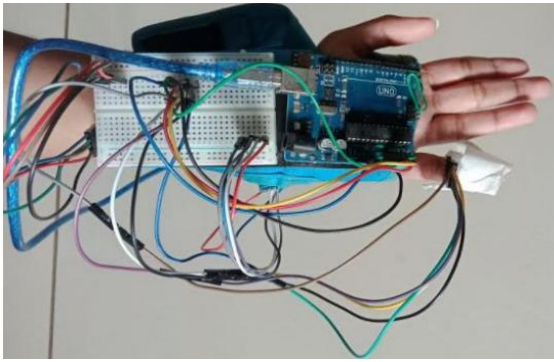


Fig. 6. Placing the sensor

### III. RESULTS & DISCUSSION

To verify the working of the wearable prototype, the device was tested by acquiring the signal, processed using software and observed, to get the following results.

The ADXL335 sensor was first placed on the fingertip of the left hand. The Fig. 7 shows the amplitude signal and the finger has the highest degree of freedom among all three parts and shows the highest amplitude ( $0.14 \text{ m/s}^2 - 0.2 \text{ m/s}^2$ ) among all the three parts. The power spectral density has given power in the range of  $40 \text{ dB/Hz} - 80 \text{ dB/Hz}$  and shown in Fig. 8.

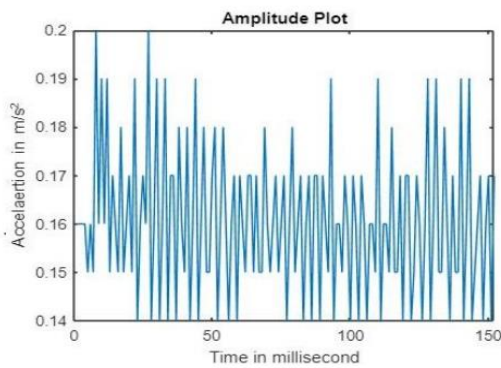


Fig. 7. Amplitude graph of the tremor from the fingertip

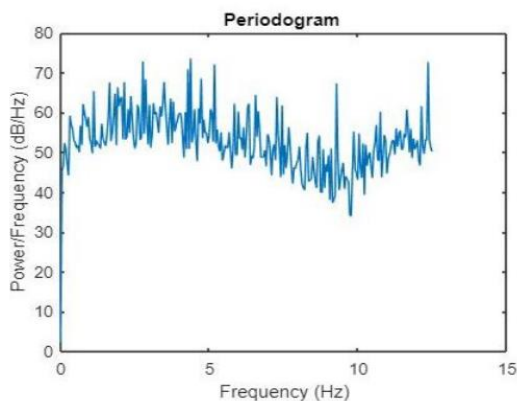


Fig. 8. Welch Periodogram of tremor from the fingertip

Fig. 9 shows the amplitude, while the Fig. 10 shows the periodogram plot when the accelerometer was placed on the wrist. The wrist tremors have the next highest range of amplitudes ( $0.14 \text{ m/s}^2 - 0.17 \text{ m/s}^2$ ) as shown in the Fig. 9. The periodogram has values in the range of  $50 \text{ dB/Hz} -$

$80 \text{ dB/Hz}$ , as observed from Fig. 10. The Fig. 11 and 12 shows the readings from the accelerometer placed on the forearm beneath the elbow joint.

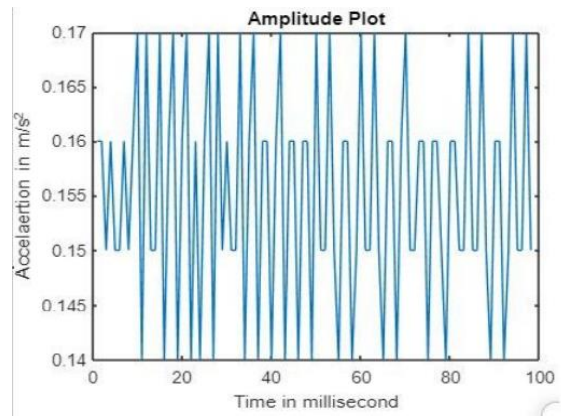


Fig. 9. Amplitude graph of the tremor from the wrist

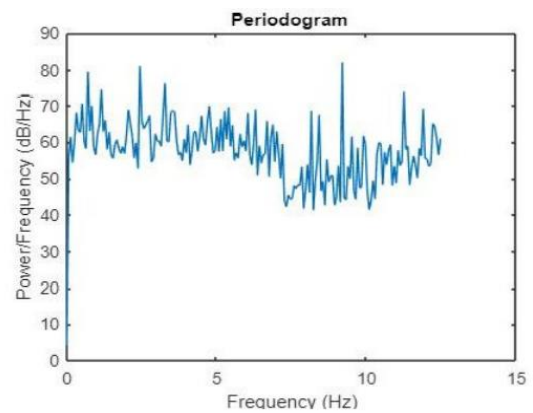


Fig. 10. Welch Periodogram of tremor from the wrist

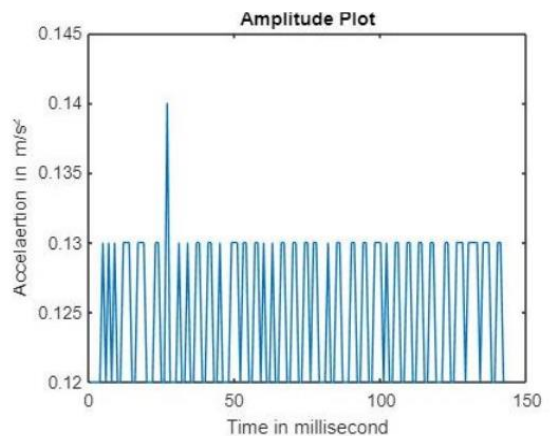


Fig. 11. Amplitude graph of the tremor from the forearm

Since the forearm has limited or constrained movements the amplitude of the tremor is less compared to the other two cases, in the range of  $0.12 \text{ m/s}^2 - 0.13 \text{ m/s}^2$ , one spike of amplitude of  $0.14 \text{ m/s}^2$ . Similarly, Periodogram has power distribution in the amplitude range of  $60 \text{ dB/Hz} - 80 \text{ dB/Hz}$  with occasional spikes for the forearm.

The amplitude plots and periodogram in all the three axis directions are considered and shown in Fig. 13 and 14 respectively. Since the movements of the finger in all three have different degrees of freedom, the acceleration values differ. In the X axis, the amplitude of the tremor was in the

range  $0.09 \text{ m/s}^2 - 0.11 \text{ m/s}^2$ . The Y-axis acceleration had a value in the range  $0.12 \text{ m/s}^2 - 0.14 \text{ m/s}^2$ . Z axis acceleration had the amplitude  $0.14 \text{ m/s}^2 - 0.17 \text{ m/s}^2$ . These can be seen from the Fig. 13. However, the power distribution is the same for all three graphs. These can be observed in the Fig. 14.

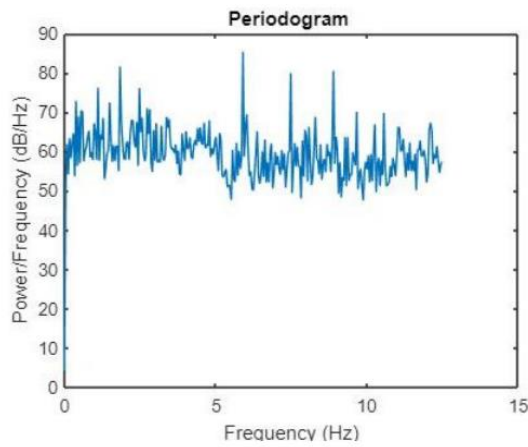


Fig. 12. Welch Periodogram of tremor from the forearm

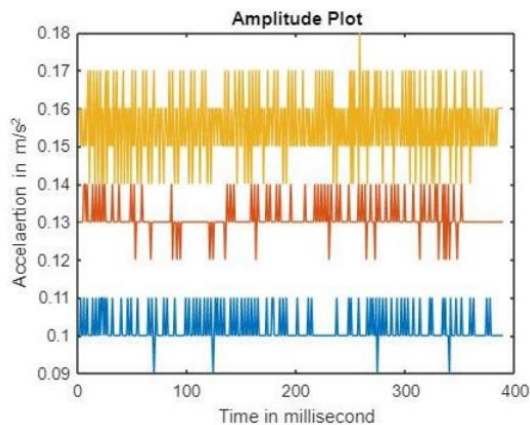


Fig. 13. Amplitude graph of the tremor from wrist, all three-axis considered

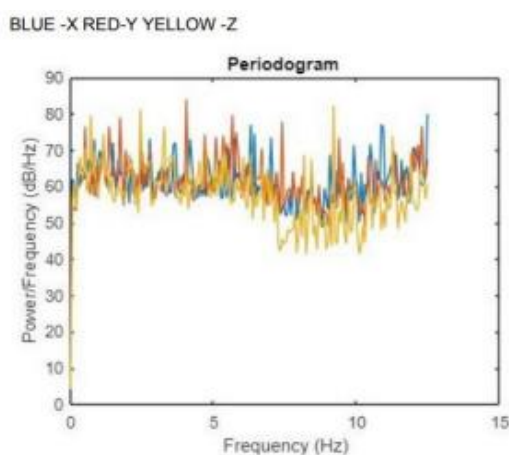


Fig. 14. Welch Periodogram of tremor from Wrist, all three-axis considered

The amplitude of the result obtained by Edwards et.al., [20] falls in the range of  $0.12 \text{ m/s}^2$  to  $-0.2 \text{ m/s}^2$ . For resting tremor, the amplitude falls in the range of  $0.144 \text{ m/s}^2$  to  $-0.156 \text{ m/s}^2$ . But the variation in the value can be

attributed to the fact that the real PD patient's condition was not analyzed. The power spectral density was similar in terms of amplitude to the research addressed [21]. However, there was no single dominant peak is observed, and the variation in the waveform can be attributed to the increased sample size used. Hence, this device can be used to analyze Parkinson's disease tremor accurately with appropriate sampling and filtering of the signals that achieved better results.

This prototype has an advantage especially when the large number of PD patients makes it difficult for the medical practitioners to diagnose them individually. Furthermore, this prototype provides an opportunity to acquire and study the tremors, when patients are involved in daily activities and outside the clinical environments. Also, this prototype provides an opportunity to provide telerehabilitation (involving telemedicine) since the tremor signals can be sent to physicians via cloud.

However, the prototype has certain limitations. The elderly people might find it difficult to wear without medical assistance. This is because the sensors used are not only fragile, but their placement is of much importance to acquire tremor signals appropriately. For continuous monitoring, this prototype might appear inappropriate as it is placed on the arm itself, and it hampers many of the daily activities. Finally, the elderly might find the size of the prototype to be huge enough for a comfortable use. Since the prototype is in its initial stages of development, certain aspects of it are to be improved to overcome the limitations.

Further, to make this prototype user friendly, the size of the microcontroller can be reduced. One way to achieve this could be using Arduino pro mini instead of Arduino Uno. The glove is to be designed such that the sensors and the Arduino is fixed, so there would be no difficulties in the placement of the sensors. Secondly, to avoid the usage of breadboard, the wires of the sensors should be soldered properly and the wires should be secured within the gloves to avoid mishap and variations. Furthermore, USB cable can be replaced by Bluetooth module and advanced Wi-Fi modules.

#### IV. CONCLUSION

This work presents a method to acquire and analyze the resting tremor from Parkinson's disease (PD) patients. This prototype has an accelerometer and Arduino Uno to acquire the tremor signal. The Arduino was connected to the PC via USB. Both the accelerometer and Arduino Uno were mounted on the glove to get the signal from fingertip, wrist and forearm. The amplitude and the periodogram of the tremors from were analyzed with amplitude plots and periodograms using MATLAB IDE. The prototype presents a basic medical wearable device that could be used to study the tremors from the PD patients and assist them with the right treatment.

#### REFERENCES

- [1] R. LeMoyné, C. Coroian, and T. Mastroianni, "Quantification of Parkinson's disease characteristics using wireless accelerometers," in 2009 ICME International Conference on Complex Medical Engineering, 2009, pp. 1–5, doi: 10.1109/ICME.2009.4906657.

- [2] C. Marras et al., "Prevalence of Parkinson's disease across North America," *npj Park. Dis.*, vol. 4, no. 1, pp. 1–7, 2018, doi: 10.1038/s41531-018-0058-0.
- [3] D. H. Romero and G. E. Stelmach, "Changes in postural control with aging and Parkinson's disease," *IEEE Eng. Med. Biol. Mag.*, vol. 22, no. 2, pp. 27–31, 2003, doi: 10.1109/MEMB.2003.1195692.
- [4] H. Badem, A. Caliskan, A. Basturk, and M. E. Yuksel, "Classification and diagnosis of the parkinson disease by stacked autoencoder," in 2016 National Conference on Electrical, Electronics and Biomedical Engineering (ELECO), 2016, pp. 499–502.
- [5] V. Majhi, S. Paul, G. Saha, and J. K. Verma, "Sensor based Detection of Parkinson's Disease Motor Symptoms," in 2020 International Conference on Computational Performance Evaluation (ComPE), 2020, pp. 553–557, doi: 10.1109/ComPE49325.2020.9200051.
- [6] M. Bhat, S. Inamdar, D. Kulkarni, G. Kulkarni, and R. Shriram, "Parkinson's disease prediction based on hand tremor analysis," in 2017 International Conference on Communication and Signal Processing (ICCSP), 2017, pp. 625–629, doi: 10.1109/ICCSP.2017.8286433.
- [7] R. Lemoyne, T. Mastroianni, M. Cozza, C. Corioan, and W. Grundfest, "Implementation of an iPhone for characterizing Parkinson's disease tremor through a wireless accelerometer application," *Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. IEEE Eng. Med. Biol. Soc. Annu. Int. Conf.*, vol. 2010, pp. 4954–4958, 2010, doi: 10.1109/ieems.2010.5627240.
- [8] H. Zhang, X. Chen, W.-Y. Lin, W.-C. Chou, and M.-Y. Lee, "A novel accelerometer-based method for the real-time assessment of Parkinson's tremor," in 2014 IEEE International Conference on Communication Problem-solving, 2014, pp. 87–90, doi: 10.1109/ICCPS.2014.7062224.
- [9] A. Zhang et al., "Automated Tremor Detection in Parkinson's Disease Using Accelerometer Signals," in 2018 IEEE/ACM International Conference on Connected Health: Applications, Systems and Engineering Technologies (CHASE), 2018, pp. 13–14, doi: 10.1145/3278576.3278582.
- [10] N. Kostikis, D. Hristu-Varsakelis, M. Arnaoutoglou, and C. Kotsavasiloglou, "A Smartphone-Based Tool for Assessing Parkinsonian Hand Tremor," *IEEE J. Biomed. Heal. informatics*, vol. 19, no. 6, pp. 1835–1842, Nov. 2015, doi: 10.1109/JBHI.2015.2471093.
- [11] J. A. Stamford, P. N. Schmidt, and K. E. Friedl, "What Engineering Technology Could Do for Quality of Life in Parkinson's Disease: A Review of Current Needs and Opportunities," *IEEE J. Biomed. Heal. informatics*, vol. 19, no. 6, pp. 1862–1872, Nov. 2015, doi: 10.1109/JBHI.2015.2464354.
- [12] A. K. Vijay, K. Sangeetha, A. A. Shibani, and M. Pranitha, "Tremomarker Tremor Detection for Diagnosis in a Non-Clinical Approach Using IoT," in 2018 Fourth International Conference on Biosignals, Images and Instrumentation (ICBSII), 2018, pp. 206–212, doi: 10.1109/ICBSII.2018.8524816.
- [13] D. Pan, R. Dhall, A. Lieberman, and D. B. Petitti, "A mobile cloud-based Parkinson's disease assessment system for home-based monitoring," *JMIR mHealth uHealth*, vol. 3, no. 1, p. e29, Mar. 2015, doi: 10.2196/mhealth.3956.
- [14] A. de Oliveira Andrade et al., "Task-Specific Tremor Quantification in a Clinical Setting for Parkinson's Disease," *J. Med. Biol. Eng.*, vol. 40, no. 6, pp. 821–850, 2020, doi: 10.1007/s40846-020-00576-x.
- [15] M. Bravo et al., "A system for finger tremor quantification in patients with Parkinson's disease," *Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. IEEE Eng. Med. Biol. Soc. Annu. Int. Conf.*, vol. 2017, pp. 3549–3552, Jul. 2017, doi: 10.1109/EMBC.2017.8037623.
- [16] G. N. McKay, T. P. Harrigan, and J. R. Brašić, "A low-cost quantitative continuous measurement of movements in the extremities of people with Parkinson's disease," *MethodsX*, vol. 6, pp. 169–189, 2019, doi: <https://doi.org/10.1016/j.mex.2018.12.017>.
- [17] C. Bhavana, J. Gopal, P. Raghavendra, K. M. Vanitha, and V. Talasila, "Techniques of measurement for Parkinson's tremor highlighting advantages of embedded IMU over EMG," in 2016 International Conference on Recent Trends in Information Technology (ICRTIT), 2016, pp. 1–5, doi: 10.1109/ICRTIT.2016.7569560.
- [18] D. Sueaseenak, N. Namjirachot, and K. Sukkit, "Accelerometer-based angle measurement system with application in hospital bed," in 2015 8th Biomedical Engineering International Conference (BMEiCON), 2015, pp. 1–4, doi: 10.1109/BMEiCON.2015.7399568.
- [19] Y. A. Badamasi, "The working principle of an Arduino," in 2014 11th International Conference on Electronics, Computer and Computation (ICECCO), 2014, pp. 1–4, doi: 10.1109/ICECCO.2014.6997578.
- [20] R. Edwards and A. Beuter, "Using time domain characteristics to discriminate physiologic and parkinsonian tremors," *J. Clin. Neurophysiol. Off. Publ. Am. Electroencephalogr. Soc.*, vol. 17, no. 1, pp. 87–100, Jan. 2000, doi: 10.1097/00004691-200001000-00009.
- [21] M. F. Dirx, H. Zach, B. R. Bloem, M. Hallett, and R. C. Helmich, "The nature of postural tremor in Parkinson disease," *Neurology*, vol. 90, no. 13, pp. e1095–e1103, Mar. 2018, doi: 10.1212/WNL.0000000000005215.

**2021 IEEE Seventh  
International Conference on  
Bio signals, Images and  
Instrumentation  
ICBSII 2021**

**SESSION III  
RESEARCH  
PAPERS**

## A COMPARATIVE STUDY OF VARIOUS MULTIMODAL MEDICAL IMAGE FUSION TECHNIQUES– A REVIEW

1<sup>st</sup> KANIKE VIJAY KUMAR  
 Research Scholar  
 Dept., of ECE  
 R.G.M.C.E.T  
 Affiliated to JNTUA Ananthapuramu  
 AP, India  
[vijay.kanike1987@gmail.com](mailto:vijay.kanike1987@gmail.com)

2<sup>nd</sup> ANCHULA SATHISH  
 Professor & HOD  
 Department of ECE  
 R.G.M.C.E.T  
 Affiliated to JNTUA Ananthapuramu  
 AP, India  
[sathishanchula@gmail.com](mailto:sathishanchula@gmail.com)

**Abstract:** The main objective of image fusion for multimodal medical images is to retrieve valuable information by combining multiple images obtained from various sources into a single image suitable for better diagnosis. In this paper, a detailed survey on various existing medical image fusion algorithms, with a comparative discussion is presented. Image fusion algorithms available in the current literature are categorized into various methods known as (1) morphological methods, (2) human value system operator based methods, (3) sub-band decomposition methods, (4) neural network based methods, and (5) fuzzy logic based methods. This research concludes that even though there exists a few open-ended creative and logical difficulties, the fusion of medical images in many combinations assists in utilizing medical image fusion for medicinal diagnostics and examination. There is tremendous progress in the fields of deep learning, artificial intelligence and bio-inspired optimization techniques. Effective utilization of these techniques can be used to further improve the efficiency of image fusion algorithms.

### 1. Introduction

With the recent advancements in the field of technology, digital image processing systems have turned into a reality in developing the number of fields, for example, machine vision and medical imaging, remote sensing [1] and military applications. The consequence of the utilization of these strategies is an awesome increase of the amount of data available. To extract all the valuable information from the source images and to reduce the increasing volume of data, a powerful method is used in image processing called image fusion. The main aim of image fusion is to produce new images that are more appropriate for human/machine perception.

Image fusion is the process of combining multiple images of a scene into a single composite image that contains all the important features from each input image. The subsequent fused image delivers extra reliable and exact information about a scene than any of the individual source images. The fusion of redundant and complementary data increases the accuracy and decreases the overall uncertainty of multimodal medical images. Medical image fusion algorithms involve two stages of processes. They are: (a) Image registration and (b) Fusion of significant features from the registered images. A basic image fusion model is presented in Figure 1.

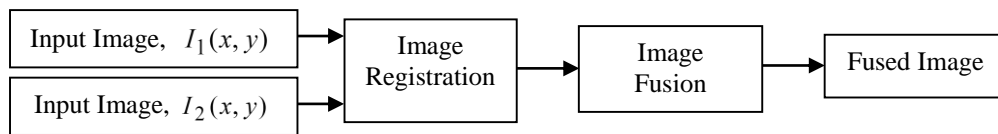
Information attained in the clinical track of events gained from two source images is generally complementary. Appropriate incorporation of valuable information achieved from the individual images is frequently chosen. The initial phase in this fusion procedure is to bring the modalities associated with spatial alignment and this happens with a technique called registration [2, 3].

The registration of the images requires a method to correct the spatial misalignment between the different image data sets that often involve compensation of variability resulting from scale changes, translations, and rotations. After registration, a fusion step is required for the integrated display of the data involved. In general, fusion scheme should satisfy the following requirements [4]:

- It should recognize the most important features in the source images and transfer them without loss of detail into the fused image.
- The image fusion process should not introduce any artifacts or inconsistencies which mislead or divert a human observer for further processing tasks.
- It should be robust, reliable, and suppress the irrelevant parts of the image and noise.

Multimodal medical images [5, 6, 7, 8, 9, 10] provide different types of information: Computed Tomography (CT) image provides details of dense structures, such as bones, Magnetic Resonance Image (MRI) affords information about pathological soft tissues, Magnetic Resonance Angiography (MRA) easily detects abnormalities in the brain, X-ray detects fractures and abnormalities in bone positions, Vibro-Acoustography (VA), provides the depth and thickness of the disease object, and Positron Emission Tomography (PET) and Single Photon Emission Computed Tomography (SPECT) both give functional and metabolic information, e.g. of the brain.

Multimodal medical image pairs such as CT-MRI [11, 12, 13, 14], CT-SPECT [15], PET-CT [16], MRI-PET [17], MRI-CT-PET [18], MRI-SPECT [19], Ultrasound-Xrays [20, 21] are fused to extract additional clinical information. Therefore, we can say that all relevant information cannot be provided by a single image. Thus, to acquire all complete information in a single composite image called a fused image, multimodal medical image fusion is fundamentally necessary. This fused information is very useful for physicians in diagnosing abnormalities. The advantage of a medical image fusion application is to decrease the difficulty in detecting different diseases [22, 23, 24].



**Fig. 1.** Block diagram of a basic image fusion model

## 2. Image fusion techniques and imaging modalities

Image fusion integrates different modality images to provide complete information of the image content, to increase interpretation capabilities and to generate more reliable outcomes. There are several advantages of combining multi-modal images, including refined geometric corrections, complementing data for enhanced classification, and enhancing features for investigations and so forth. Image Fusion is widely applied and proven its efficiency in different research areas, such as computer vision, multimedia analysis, biomedical research, and material sciences. Image fusion techniques are categorized into five different groups [4]: based on the type of domain used, based on an acquisition of input images, based on the level of processing, based on the activity level and based on the information content of the image. They are described as follows:

### 2.1. Domain Type

Images can be processed in the spatial domain (pixels in the image), in a transformed domain (change of frequencies in the data), or a combination.

**2.1.1 Spatial Domain:** These methods are specifically dealt with image pixels. These types of algorithms are known as spatial domain fusion algorithms. The following are the fusion-based techniques that fall under the spatial domain, for example, Averaging, Principal Component Analysis (PCA), Brovey method, and Intensity Hue Saturation (IHS).

**2.1.2 Transform Domain:** The basic disadvantage of the spatial domain method is that it produces spatial distortions in the fused image. In the Transform domain method the image is first converted into the frequency domain by finding the Fourier transform of the image. The Fourier transform is defined in both spatial frequency and spatial location while the spatial domain is defined only on the spatial frequency domain. In the transform domain technique, Fourier transform is applied to all the input images. Then, the inverse Fourier transform is performed to obtain the final fused image. Most popular transforms are Discrete Fourier Transform (DFT), Discrete Cosine Transform (DCT), Discrete Hadamard Transform (DHT), Discrete Wavelet Transform (DWT), Redundant Wavelet Transform (RWT), Contourlet Transform, Curvelet Transform and Singular Value Decomposition (SVD).

**2.1.3 Hybrid Domain:** In the hybrid domain image fusion algorithms, either a fusion of spatial and transform domains or a fusion of two transform domains is used for fusion.

### 2.2. Input Image Acquired

Images are acquired and fused in different ways such as in the form of multi-sensor, multi-temporal, multi-focus, multimodal and multi-view.

**2.2.1 Multi-sensor Image Fusion:** This type of fusion incorporates the source images taken by various sensors.

**2.2.2 Multi-temporal Image Fusion:** This type of fusion joins the images taken at various conditions with a specific end goal to blend accurate images of articles that were not snapped in anticipated time.

**2.2.3 Multi-focus Image Fusion:** This type of image fusion contracts with image scenes brought repetitively with various central lengths and in this, the supplementary information of source images is fused.

**2.2.4 Multimodal Image Fusion:** This type of fusion integrates both supplementary and complementary information of source images.

**2.2.5 Multi-view Image Fusion:** This incorporates the images from a similar methodology taken at a time from various perspectives.

### 2.3. Level of Processing


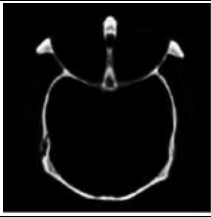
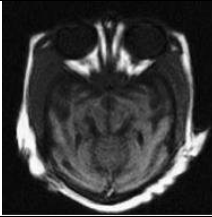
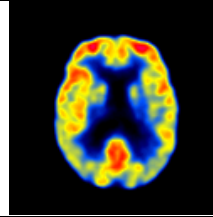
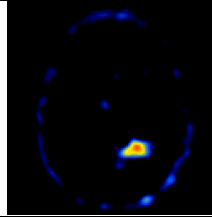
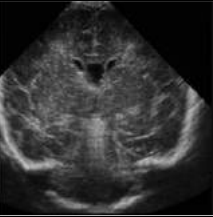
Images are fused in different levels of processing such as pixel, feature, and decision. The details of each level of processing are given below.

**2.3.1 Pixel-level Fusion:** Pixel-level methods fuse two images by transforming one of the images in such a way to make it the most similar to the other concerning for to some objective measure of similarity.

**2.3.2 Feature-level Fusion:** Fusion at feature level initially extracts objects of interest for different image modalities, e.g., utilizing segmentation strategies. These comparative articles (e.g., regions) from different modality images are then fused for extra appraisal utilizing factual methodologies.

**2.3.3 Decision-level Fusion:** Information extraction by decision-level fusion uses the interpreted/labeled data where the input images are independently handled. Choice tenets are connected to consolidate the data to strengthen basic translation and to give a superior comprehension of the watched objects. The key benefit of this method is that the higher-level representations make multi-modality fusion more robust and reliable.

**Table 1** Comparison of various multimodal medical images [3]

	<b>X-ray</b>	<b>CT</b>	<b>MRI</b>	<b>PET</b>	<b>SPECT</b>	<b>Ultrasound</b>
<b>Modality</b>						
<b>Characteristics</b>	<ul style="list-style-type: none"> <li>• Detect fractures and abnormalities in bone positions of the human body</li> </ul>	<ul style="list-style-type: none"> <li>• Provides details regarding dense structures, such as bones</li> <li>• Excellent at delineating bones</li> </ul>	<ul style="list-style-type: none"> <li>• Provides information about pathological soft tissues</li> <li>• Widely utilized imaging modalities in medical examinations in secret clinical sets</li> </ul>	<ul style="list-style-type: none"> <li>• It is one of the scan which provide functional information about the human brain</li> <li>• Permits to record the deviations in the healthy brain activity and also the symptoms of several diseases</li> </ul>	<ul style="list-style-type: none"> <li>• One of the non-obtrusive based strategy where cross-sectional images of radiotracer inside the human body are organized</li> </ul>	<ul style="list-style-type: none"> <li>• It is a technique based on sound waves that possess high temporal frequency</li> <li>• Proficient in generating quantitative and qualitative diagnostic information</li> </ul>
<b>Application Examples</b>	<ul style="list-style-type: none"> <li>• Widely used for breast cancer assessment</li> </ul>	<ul style="list-style-type: none"> <li>• Assistance in surgical planning, training, and guidance</li> <li>• Diagnosis of brain</li> <li>• Head and neck cancer diagnosis</li> <li>• Tumor detection</li> </ul>	<ul style="list-style-type: none"> <li>• Lung/Liver diagnosis and breast cancer assessment</li> <li>• Specimen shape, color, and structure is extracted</li> </ul>	<ul style="list-style-type: none"> <li>• Cancer treatments</li> <li>• Gynecological cancer diagnosis</li> <li>• Size detection of unrefined tumor</li> <li>• Esophageal cancer diagnosis</li> </ul>	<ul style="list-style-type: none"> <li>• Head and neck cancer diagnosis</li> <li>• Cancer treatment of Bone, Vulnar, Breast, and Lungs</li> <li>• Diagnosis of Liver</li> <li>• Detection of tumors</li> </ul>	<ul style="list-style-type: none"> <li>• Treatment of Prostate cancer</li> <li>• Detection of breast cancer, Image fusion</li> <li>• Diagnosis of the Liver tumor and Esophageal cancer</li> </ul>
<b>Advantages</b>	<ul style="list-style-type: none"> <li>• Detect fractures and abnormalities in bone positions of the human body</li> </ul>	<ul style="list-style-type: none"> <li>• Relative short scan time and high spatial imaging resolutions</li> <li>• Detection of even subtle differences between body tissues</li> <li>• High penetration depth</li> </ul>	<ul style="list-style-type: none"> <li>• Higher resolution</li> <li>• Capable of showing anatomical details</li> <li>• Does not involve any exposure to radiation</li> </ul>	<ul style="list-style-type: none"> <li>• High sensitivity</li> <li>• High penetration depth</li> <li>• One of the widely used clinical scan</li> </ul>	<ul style="list-style-type: none"> <li>• High sensitivity</li> <li>• High penetration depth</li> <li>• One of the widely used clinical scan</li> </ul>	<ul style="list-style-type: none"> <li>• High spatial resolution</li> <li>• Low cost</li> <li>• Clinical interpretation</li> <li>• Broadly accessible</li> <li>• Simple to utilize</li> <li>• Free of Radiation</li> </ul>
<b>Disadvantages</b>	<ul style="list-style-type: none"> <li>• X-ray does not contain information about the depth and thickness of the objects</li> </ul>	<ul style="list-style-type: none"> <li>• Constrained affectability</li> <li>• Radiation</li> <li>• Tissue non-specificity</li> <li>• Poor soft-tissue contrast</li> </ul>	<ul style="list-style-type: none"> <li>• Difficult to use for patients with metallic devices, such as pacemakers</li> </ul>	<ul style="list-style-type: none"> <li>• Radiation</li> <li>• High cost and most expensive technique</li> </ul>	<ul style="list-style-type: none"> <li>• Blurring effects are produced</li> <li>• Inadequate spatial resolution</li> <li>• Radiation</li> </ul>	<ul style="list-style-type: none"> <li>• Operator dependent</li> <li>• Imaging is restricted to the vascular compartment</li> <li>• Challenging image of bone and lungs</li> </ul>
<b>Radiation Source &amp; Type</b>	<ul style="list-style-type: none"> <li>• X-rays (ionizing)</li> </ul>	<ul style="list-style-type: none"> <li>• X-rays (ionizing)</li> </ul>	<ul style="list-style-type: none"> <li>• Electric &amp; Magnetic Fields (Non-ionizing)</li> </ul>	<ul style="list-style-type: none"> <li>• Positron (ionizing)</li> </ul>	<ul style="list-style-type: none"> <li>• Photons (ionizing)</li> </ul>	<ul style="list-style-type: none"> <li>• Sound waves (Non-ionizing)</li> </ul>
<b>Cost</b>	<ul style="list-style-type: none"> <li>• Low cost</li> </ul>	<ul style="list-style-type: none"> <li>• Intermediate cost</li> </ul>	<ul style="list-style-type: none"> <li>• Intermediate cost</li> </ul>	<ul style="list-style-type: none"> <li>• High cost</li> </ul>	<ul style="list-style-type: none"> <li>• High cost</li> </ul>	<ul style="list-style-type: none"> <li>• Low cost</li> </ul>



## 2.4. Algorithms using Human Value System (HVS) Operator's model

Human value System operator models (or) Knowledge-based systems permits to backtrack the new emblematic image portrayal in a very much characterized implies with a specific end goal to create elective translations. This model permits information relating to the fusion of image, geometric and spatial properties that interestingly characterize anatomical structures. Information about the idea of the image and the restorative space are utilized to aid the segmentation/acknowledgment [34, 35] errand by:

- Hypothesis age - delivering an enthusiastic present-day world.
- Knowledge-based coordination of the regions made by low-level segmentation.
- Permitting procured data around an image and the space to permit a more exact depiction of anatomical structures to be accomplished iteratively.

There is a progression of utilizations where the domain-dependent learning is appropriate for image fusion, for instance, segmentation, small scale calcification finding, tissue characterization, cerebrum analysis, classifier fusion, breast cancer tumor discovery, and depiction and acknowledgment of anatomical mind question. The advantage of HVS is the capacity to benchmark the images with the known human vision principles. While the disadvantage is the impediment forced by human choice in images that are inclined to huge pixel power changeability. A few Human Value System based techniques [36, 37, 38] for multimodal medical image fusion are available in the current literature. Some of the algorithms are reviewed as follows.

Li et al. [36] displayed a cerebrum show and an entire framework for the outline and the acknowledgment of anatomical mind protests in an arrangement of parallel CT-examines. The analyses are carried on 10 patients and the following points are concluded: The framework can naturally segment and name 25 diverse 3-D mind-objects from a full arrangement of 14 CT-examines. Segmentation mistakes are kept low when contrasted with manual division by the radiologist. However, providing quantitative estimations for decay assessment to the doctor, for example, different width parameters of the ventricular framework isn't given properly.

Yang et al. [42] proposed a multiscale geometric examination device, contourlet transforms for multimodal restorative image fusion of CT and MRI images. It offers favourable circumstances of directionality, confinement, anisotropy and multiscale which cannot be accomplished by wavelet transform. However, contourlet transform flops in recognizing curved visual subtle elements in a fused image.

Since medical images have a few entities and curved shapes, it is smarter to utilize a different transform for identifying these in MR and CT medical images. CT images are for the most part connected with conceiving thick structures, like bones. Then again, MR images are utilized to speak to the morphology of soft tissues along these lines, they are rich in points of interest. Ali et al. [43] proposed the fusion of MR and CT images utilizing the curvelet transform where curved visual subtle elements are better identified. Since the two modalities, MR and CT are

complementary; merging both images, provide as much information as possible. However, these wavelets have an absence of shift-invariance which results in extreme blocking artifacts in fused images.

Yang et al. [44] introduced another technique for medicinal image fusion, utilizing a shift-invariant multiscale wavelet (SIMW) system. The deterioration system is accomplished by leaving the down sampling administrators of a morphological Haar wavelet. A trial given because of genuine medicinal images demonstrates the proposed strategy that enhances the nature of the fused image fundamentally. Singh et al. [45] likewise proposed a fusion calculation to join sets of multispectral magnetic resonance images, for example, T1, T2 and Proton Density mind images utilizing Redundant Discrete Wavelet Transform (RDWT), a shift-invariant wavelet. Analyses on the Brain Web database demonstrate this fusion method conserves both edge and component data [46]. Tirupal et al. [47, 48, 49, 50] applied Undecimated Discrete Wavelet Transform (UDWT) to multimodal medical image fusion along with modified spatial frequency (MSF), contrast visibility (CV) and moments. However, if there is no decimation then these algorithms have more computational complexity.

Kavitha et al. [51] proposed image fusion in view of Integer Wavelet Transform (IWT) and Neuro-Fuzzy. The anatomical and the functional images are disintegrated utilizing IWT. The wavelet coefficients are then combined utilizing a neuro-fuzzy approach. At that point, Inverse Integer Wavelet Transform (IIWT) is smeared to the fused coefficients to get the fused Image. In IWT, there is no loss of data through forward and inverse transform. The triangular membership function is utilized for the fusion of two images. However, this transform undergoes discontinuities such as edges and delineations in fused images.

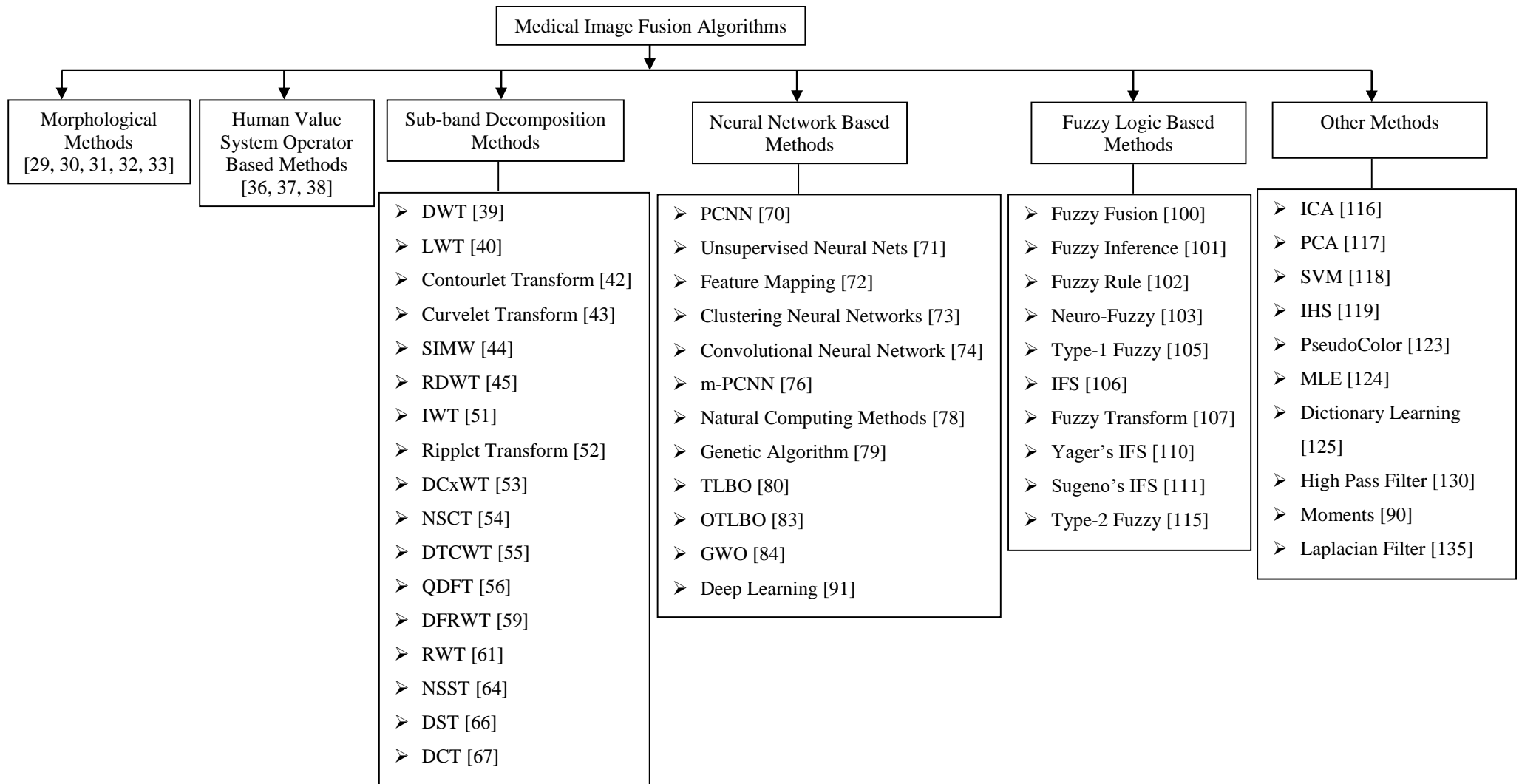
Poor directionality, shift sensitivity, absence of phase data of genuine esteemed wavelet transforms roused to utilize complex wavelet transform for fusion. Singh et al. [53] utilized Daubechies complex wavelet transform (DCxWT) and applied the properties accessibility of phase data, shift-invariance and multiscale edge data for multimodal medical images. The DCxWT has the accompanying favourable circumstances:

- Idealized reconstruction property.
- Redundancy is not present.
- The number of computations in DCxWT is same as that of DWT.
- Symmetric in nature. This property makes it simple to deal with edge focuses amid the signal reconstruction.

The above properties and advantages make the DCxWT suitable for medical image fusion.

Yang et al. [54] introduced a novel multimodal medical image fusion strategy in view of Non-Subsampled Contourlet Transform (NSCT) and Log-Gabor energy. The potential advantages of this algorithm are:

- It is a multiresolution, multidirectional, and shift-invariance.
- Phase congruency and Log-Gabor vitality are utilized to safeguard more helpful data in the combined image to enhance the nature of the fused image.
- The proposed algorithm conveys improved performance regardless of whether the source images are noisy.



*Fig. 2. Image fusion algorithms available in the current literature*

## Conclusion

Multimodal medical image fusion is a technique wherein, two multimodal medical images are combined to enhance the nature of the output image for better treatment and precise diagnosis. The main purpose of using this multimodal medical image fusion is to combine multiple medical source images to get an accurate output image for better diagnosis. Several algorithms are available in the literature for this purpose. Examples of different types of multimodal medical image fusion are presented in Figure 3. In this paper, various multimodal medical image fusion algorithms available in the current literature are reviewed and the comparison of these algorithms is given in Tables 3 and 5. In this work, various objective criteria are considered in finding the quality of the output fused image. A detailed comparison is made. These are highlighted in this work.

## References

- [1] Gopi Krishna, E., Tirupal, T.: 'Performance of Image Fusion Techniques for Satellite Images', *International Journal for Technological Research in Engineering*, 2015, 2(12), pp. 3184–3188
- [2] Maintz, J.B.A., Viergever, M.A.: 'A survey of medical image registration', *Medical Image Analysis*, 1998, 2, (1), pp. 1–36
- [3] El-Zahraa, F., El-Gamal, A., Elmogy, M., et al.: 'Current trends in medical image registration and fusion', *Egyptian Informatics Journal*, 2015, pp. 1–26.
- [4] Mitchell, H.B.: 'Image Fusion theories, techniques and applications' (Springer-Verlag, Berlin, 2010)
- [5] James, A.P., Dasarathy, B.V.: 'Medical image fusion: a survey of the state of the art', *Information Fusion*, 2014, 19, (1), pp. 4–19
- [6] Luca Piras., Giorgio Giacinto.: 'Information Fusion in Content Based Image Retrieval: A Comprehensive Overview', *Information Fusion*, 2017, doi: 10.1016/j.inffus.2017.01.003
- [7] Ayush Dogra., Bhawna Goyal., Sunil Agrawal.: 'From Multi-scale Decomposition to Non-multi-scale Decomposition Methods: A Comprehensive Survey of Image Fusion Techniques and its Applications', *IEEE Access*, 2017, doi:10.1109/ACCESS.2017.2735865
- [8] Qiang Zhang., Yi Liu., Rick S.B., Jungong Han., Dacheng Tao.: 'Sparse Representation based Multi-sensor Image Fusion for Multi-focus and Multi-modality Images: A Review', *Information Fusion*, 2017, doi: 10.1016/j.inffus.2017.05.006
- [9] Jing-jing Zong., Tian-shuang Qiu.: 'Medical image fusion based on sparse representation of classified image patches', *Biomedical Signal Processing and Control*, 2017, 34, pp. 195–205
- [10] Bikash Meher., Sanjay Agrawal., Rutuparna Panda., Ajith Abraham.: 'A survey on region based image fusion methods', *Information Fusion*, 2018, doi: https://doi.org/10.1016/j.inffus.2018.07.010
- [11] Marshall, S., Matsopoulos, G., Brunt, J.: 'Multiresolution morphological fusion of MR and CT images of the human brain'. *Proc. of IEE in Vision, Image and Signal Processing*, 1995, 141, pp. 1–5
- [12] Yu Liu., Xun Chen., Rabab K.W., Jane Wang, Z.: 'Medical Image Fusion via Convolutional Sparsity Based Morphological Component Analysis', *IEEE Signal Processing Letters*, 2019, 26(3)
- [13] Ali, F.E., El-Dokany, I.M., Saad, A.A., et al.: 'Curvelet fusion of MR and CT images', *Progress in Electromagnetic Research C*, 2008, 3, pp. 215–224
- [14] Mishra, H.O.S., Bhatnagar, S.: 'MRI and CT image fusion based on wavelet transform', *Int. Journal of Information and Computation Technology*, 2014, 4, (1), pp. 47–52
- [15] Lin, K.P., Yao, W.J.: 'A SPECT-CT image fusion technique for diagnosis of head-neck cancer'. *IEEE Annual Conference on Engineering in Medicine and Biology Society*, 1995, 1, pp. 377–378
- [16] Shangli, C., Junmin, H., Zhongwei, L.: 'Medical image of PET/CT weighted fusion based on wavelet transform'. *IEEE Int. Conference on Bioinformatics and Biomedical Engineering*, 2008, pp. 2523–2525
- [17] Daneshvar, S., Ghassemian, H.: 'MRI and PET image fusion by combining IHS and retina-inspired models', *Information Fusion*, 2010, 11, (2), pp. 114–123
- [18] Megalooikonomou, V., Kontos, D.: 'Medical data fusion for telemedicine', *IEEE Engineering in Medicine and Biology Magazine*, 2007, 26, (5), pp. 36–42
- [19] Barra, V., Boire, J.Y.: 'A general framework for the fusion of anatomical and functional medical images', *Neuro Image*, 2001, 13, (3), pp. 410–424
- [20] Holupka, E., Kaplan, I., Burdette, E., et al.: 'Ultrasound image fusion for external beam radiotherapy for prostate cancer', *Int. Journal of Radiation Oncology Biology Physics*, 1996, 35, (5), pp. 975–984

# Brain Tumour Detection Using Deep Learning

Nitha Mohan R<sup>\*</sup>, Abinaya D<sup>\*</sup>, Aneesh R P<sup>§</sup>, Avigyan Sinha<sup>@</sup>, Ashwin G Singerji<sup>\*</sup>

<sup>\*</sup> Mar Ephraem College of Engineering & Technology, Marthandam,

<sup>§</sup>Regional centre, IHRD, Thiruvananthapuram, India(aneeshprakkulam@gmail.com)

<sup>@</sup>Johns Hopkins University, Baltimore, USA, (avigyan15@gmail.com)

**ABSTRACT:** The motivation behind this study is to detect brain tumour and provide better treatment for the sufferings. The abnormal growths of cells in the brain are called tumours and cancer is a term used to represent malignant tumours. Usually CT or MRI scans are used for the detection of cancer regions in the brain. Positron Emission Tomography, Cerebral Arteriogram, Lumbar Puncture, Molecular testing are also used for brain tumour detection. In this study, MRI scan images are taken to analyse the disease condition. Objective this research works are i) identify the abnormal image ii) segment tumour region. Density of the tumour can be estimated from the segmented mask and it will help in therapy. Deep learning technique is employed to detect abnormality from MRI images. Multi level thresholding is applied to segment the tumour region. Number of malignant pixels gives the density of the affected region.

**Index Terms:** Medical Image Processing, Brain tumour, MRI, Artificial neural network, CNN, , Keras

## I. INTRODUCTION

The early detection and treatment of brain tumour helps in early diagnosis which aids in reducing mortality rate. Image processing has been widespread in recent years and it has been an inevitable part in the medical field also. The abnormal growth of cells in the brain causes brain tumour. Brain tumour is also referred to as intracranial neoplasm. The two types of tumours are malignant and benign tumours. Standard MRI sequences are generally used to differentiate between different types of brain tumours based on visual qualities and contrast texture analysis of the soft tissue. More than 120 classes of brain tumours are known to be classified in four levels according to the level malignancy by the World Health Organization (WHO) [1].

All types of brain tumours evoke some symptoms based on the affected region of the brain. The major symptoms may include headaches, seizures, vision problems, vomiting, mental changes, memory lapses, balance losing etc [2]. Incidence of brain tumours are due to genetics, ionizing radiation mobile phones, extremely low frequency magnetic fields, chemicals, head trauma and injury, immune factors like viruses, allergies, infections, etc[3]. The malignant tumours, also known as cancerous tumours, are of two types - primary tumours, which start from the brain, and secondary tumours, which originate somewhere and spread to the brain. The risk factors for brain tumour are exposure to vinyl chloride, neurofibromatosis, ionising radiations and so on. The various diagnostic methods are computed tomography, magnetic resonance imaging, tissue biopsy etc.

Better treatments are now available for brain tumours. There is a chance of focal neurological deficits, such as motor deficit, aphasia or visual field defects in the treatment. Side effects can be avoided by measuring tumour size and time to tumour progression (TTP)[4]. Estimation of density of affected areas can give a better measurement in therapy.

Deep learning is a machine learning technique that instructs computers what to do as a human think and do in a scenario. In deep learning, a computer model is able to do classification tasks from images, sound or text. Sometimes human level performance is being exceeded by deep learning techniques. One of the most popular neural networks is an artificial neural network that has a collection of simulated neurons. Each neuron acts as a node and by links each node is connected to other nodes[5].

The aim of this paper is to build a system that would help in cancer detection from MRI images through the convolution neural network. The proposed method was tested and compared with the existing classification techniques to determine the accuracy of the proposed method.

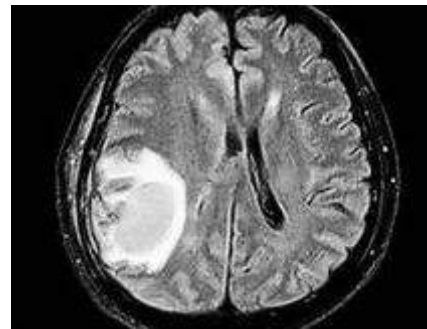


Figure 1: MRI image

## II. RELATED WORKS

Image segmentation and classification is one of the major tasks in machine learning and it is widespread in clinical diagnosis also. Mircea Gurbin, Mihaela Lascu, and Dan Lascu et al. [6] proposed a method consisting of Continuous Wavelet Transform (CWT), Discrete Wavelet Transform (DWT) and Support Vector Machine (SVM). It uses different levels of wavelets, and by training, the cancerous and non-cancerous tumours can be identified. The computation time is longer for the proposed method.

Somasundaram S. and Gobinath R. et al. [7] explains the present status of detection and segmentation of tumour through deep learning models. For deeper segmentation, 3D based CNN, ANN and SVM is used. Damodharan S. and Raghavan D. et al.[8] address segmentation of pathological tissues (Tumor), normal tissues (White Matter (WM) and Gray Matter (GM)) and fluid (Cerebrospinal Fluid (CSF)), extraction of the relevant features from each segmented tissues and classification of the tumor images with Neural Network (NN).

G. Hemanth, M. Janardhan and L. Sujihelen et al.[9] states that with the appropriate use of data mining classification technique early detection of tumour is made possible. It uses an automatic segmentation method based on CNN. Reema Mathew A. and Dr. Babu Anto P. et al. [10] stated that by the segmentation of MRI tumour region can be identified. With the help of radiological evaluations, the size and location of tumours can be identified. Here the segmentation is done manually and it is time consuming. The pre processing is done using anisotropic diffusion filters. The segmentation and classification is done using support vector machines. Wei Chen, Xu Qiao, Boqiang Liu, Xianying Qi, Rui Wang and Xiaoya Wang et al. [11] propose a novel method based on the features of separated local squares. Super pixel segmentation, feature extraction and segmentation model construction are done on this proposed method for brain tumour segmentation.

### III. PROPOSED METHOD

System architecture of the proposed system is shown in figure 1. The components are image acquisition, pre-processing, segmentation, feature extraction and classification.

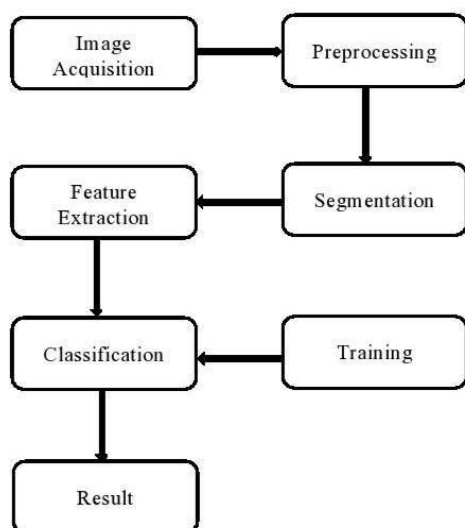


Figure 2 : System architecture of proposed method

#### A. Image Acquisition

Different bio-medical image records are available for the study of brain tumour detection. Conventional methods are Computer Tomography (CT) and Magnetic Resonance Imaging (MRI). Positron Emission Tomography, Cerebral Arteriogram, Lumbar Puncture, Molecular testing are also used for brain tumour

detection. But these are expensive. MRI is working with the principle that both the magnetic field and radio waves can create an image of the interior of the human body by detecting the water molecule present. Portable and miniaturised MRI machines are developed now to avoid the complexity of conventional scanning methods. MRI has a better resolution and contains rich information. The MRI dataset from the kaggle uploaded by Navoneel Chakrabarty has been used here[12].It contains 98 normal brain images and 155 abnormal images. In this dataset, 'yes' means tumour images and 'no' means healthy images. The augmentation process is also applied here to increase the number of samples. Augmentation step contains a rotation range of 10 degrees, width shift range of 0.1, height shift range of 0.1, brightness range of (0.3,1.0), horizontal and vertical flip. A total of 2530 images were selected from the augmented data. The final dataset contains 980 normal and 1550 abnormal images.

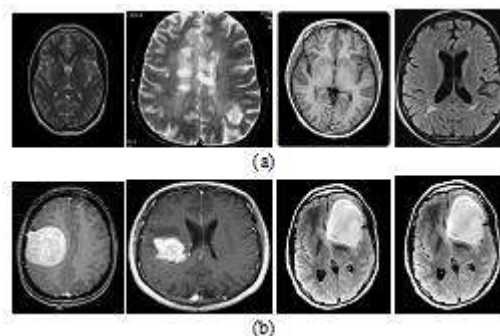


Figure 3 : Brain MRI dataset (a) Normal (b) Tumour

#### B. Pre processing

The aim of the pre-processing step is preparing the brain images for further processing [13].This process mainly depends on the data acquisition device which has its own intrinsic parameters. Gray scale or 2D conversion is needed, if the raw data is in 3D. Median filtering is best suited for biomedical images to avoid noise. The dataset contains images in different resolutions. As part of the augmentation process, each image is rotated and scaled to a standard format. Histogram equalisation helps to enhance the image quality. Contrast limited adaptive histogram equalisation algorithm is applied to enhance the images.

#### C. Image Segmentation

In this step a digital image is partitioned into multiple segments. A particular region of the image is being separated from the background. This step is very for feature extraction. Thresholding and morphological operations (erosion, dilation, opening) are the simple steps to segment disease. But in the brain tumour images, the segmentation process at this level will not give the details of tumour regions. The healthy images also have a similar intensity that resembles the tumour region. So the segmentation process can be used to separate the skull of the brain. This Region of Interest (ROI) contains the tumour. OTSU based thresholding algorithm gives a segmented mask of the skull[14].

Active contour method draws the boundary of the enclosed region. Second stage of segmentation can also be applied to the ROI to prepare the mask of tumour region.

This method may not give good results in healthy images. This segmented image can be used to study the features of tumour region , which will help in the density estimation.

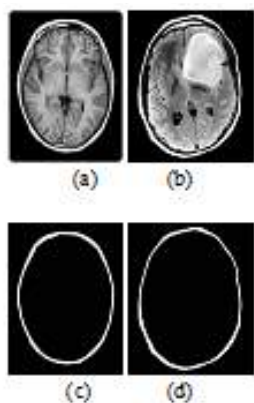


Figure 4 : Skull Segmentation (a) Normal Input image(b) Abnormal Input image (c) Normal Segmented image (d) Abnormal segmented image

D. Feature Extraction

Computing the actual features can be analysed to illustrate the behaviour or symptom of the disease. The classification is mainly influenced with the feature selection. Common features are asymmetry, diameter, and border irregularity[15].

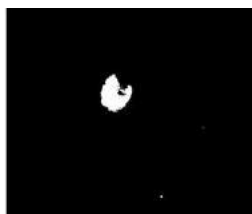


Figure 5 : Segmented tumour region using multiple thresholding

E. Classification

Many machine learning approaches are being implemented in disease detection from brain images. Artificial neural networks can be used here to classify , if the features are extracted in an order.[16].An ANN classifier assumes one feature that is not related to any other feature.

Deep learning techniques will be effective here to classify tumour image without segmentation. A deep neural network can be created with Convolutional neural network algorithm[17]. General architecture of convolutional neural networks is shown in figure 6. In deep learning , the feature is extracted from the entire image automatically. Convolution in the CNN architecture performs this operation. Number of feature maps increases with the increase in CONV layer. Reduction of dimension is required to initiate training. Pooling layer down samples the feature dimension. Fully connected layers manipulate the score of each label. Softmax layers prepare the model with feature and class score.

The CNN architecture is slightly modified in its dimension for training the brain tumour images. The modified model architecture is listed in table 1.

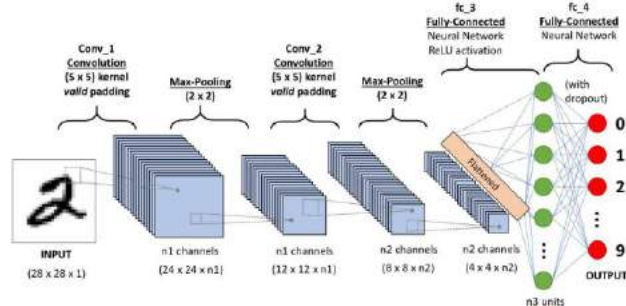


Figure 6 : General architecture of CNN

Model: "BrainTumorDetectionModel"

Layer (type)	Output Shape	Param #
input_1 (InputLayer)	[(None, 240, 240, 3)]	0
zero_padding2d (ZeroPadding2D)	(None, 244, 244, 3)	0
conv0 (Conv2D)	(None, 238, 238, 32)	4736
bn0 (BatchNormalization)	(None, 238, 238, 32)	128
relu0 (Activation)	(None, 238, 238, 32)	0
max_pool0 (MaxPooling2D)	(None, 59, 59, 32)	0
max_pool1 (MaxPooling2D)	(None, 14, 14, 32)	0
flatten (Flatten)	(None, 6272)	0
fc (Dense)	(None, 1)	6273

Total params: 11,137  
 Trainable params: 11,073  
 Non-trainable params: 64

Table 1: Modified Model architecture

The model is compiled in Keras with ‘Adam’ optimizer and ‘Binary cross entropy’ loss. Default learning rate of 0.001 is used. The model is trained with batch size of 32 for 24 epochs. For the test images our trained model gives an accuracy of 95.6%.

For those images classified as having brain tumour, the tumour location is detected using a combination of multilevel thresholding, morphological operations and contour extraction.

$$g(x, y) = \begin{cases} 1, & f(x, y) > T \\ 0, & f(x, y) \leq T \end{cases} \dots\dots (1)$$

Where T is the mean of Maximum to Minimum intensities of the image. Morphological Open function is used to segment the regions. Contours of all regions are plotted and the region with maximum area contains tumour region.

The Density of the tumour area can be estimated using Gaussian kernel distribution.

$$f(x) = \frac{1}{n\sigma\sqrt{2\pi}} \sum_{i=0}^n e^{-\frac{1}{2}\left(\frac{x_i-x}{\sigma}\right)^2} \dots\dots\dots (2)$$

IV. RESULTS AND DISCUSSIONS

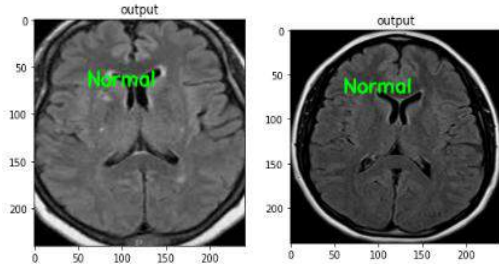
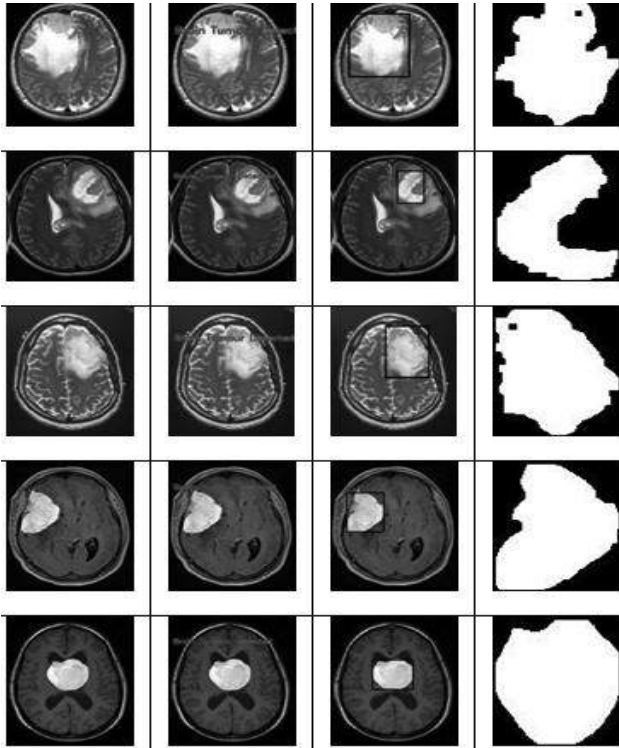


Figure 7 : Normal Images Results



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

Figure 8 : Tumour detection results : (a) Input Image (b) Abnormality Detection (c) Tumour region detection (d) tumour mask for density estimation

Objective of the proposed system is to classify malignant brain tumour from the MRI images. 253 MRI images were collected from Kaggle dataset. The count of the data is insufficient for modelling a deep neural network. So 2530 images have been created with augmentation technique. The extracted cropped images are then resized to (240, 240) resolution. Keras (with Tensorflow backend) framework is chosen creating the model. Two types of segmentation at different level are implemented to analyse the performance of the system. Segmentation was done before and after classification. From the performance analysis, segmentation after the classification gives better result.

This algorithm is faster in execution for normal MRI images. If it identifies the abnormal images, it goes to the next step ,ie : segmentation.

ROC curve shows the relation between sensitivity and specificity .

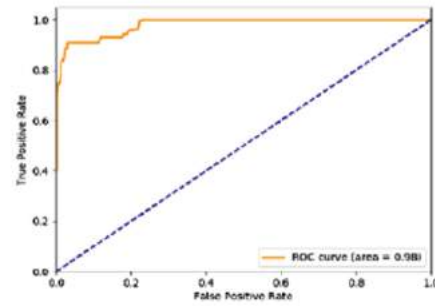


Figure 9 : ROC plot of Normal cases

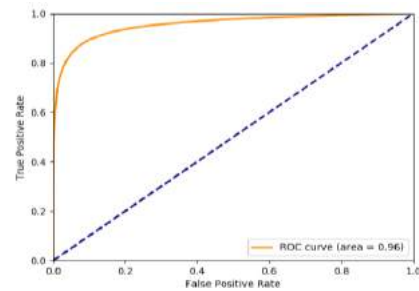


Figure 10 : ROC plot of Abnormal cases

$$Accuracy = \frac{(TP+TN)}{(TP+TN+FP+FN)} = 0.98$$

$$Sensitivity = \frac{TP}{TP+FN} = 0.97$$

$$Specificity = \frac{TN}{TN+FP} = 0.99$$

Table 2: Performance Analysis

Algorithm	Over all Accuracy
Nandpuru[18]	96.77%
El-Dahshan[19]	97%
Ibrahim[20]	96.33%
Rajini[21]	90%
Proposed Method	98%

IV. CONCLUSION AND FUTURE SCOPE

This paper provides a new method for detecting brain tumour by deep learning method. The early detection of cancer helps timely and effective treatment. Kaggle dataset contains good quality of MRI images for research purposes. Different segmentation algorithms were experimented. From this , multilevel thresholding and OTSU thresholding are the best methods for the dataset. Convolutional Neural Network with modified

approach helped to get a result with accuracy 98%. Density estimation method is also proposed using Gaussian kernel distribution.

This system can be improved to support the web interface. Detection of different diseases can be identified from the MRI images. Apart from the density some other parameters can also estimated for therapeutic purposes.

#### REFERENCES

- [1] David N. Louis, Arie Perry, et al. , “The 2016 World Health Organization Classification of Tumors of the Central Nervous System: a summary” , Acta Neuropathol , Springer may 2016
- [2] Pär Salander, A Tommy Bergenheim, Katarina Hamberg, Roger Henriksson, Pathways from symptoms to medical care: a descriptive study of symptom development and obstacles to early diagnosis in brain tumour patients, Family Practice, Volume 16, Issue 2, April 1999, Pages 143–148,
- [3] McKinney PA ,”Brain tumours: incidence, survival, and aetiology”,Journal of Neurology, Neurosurgery & Psychiatry 2004;75:ii12-ii17.
- [4] Heimans, J., Taphoorn, M. Impact of brain tumour treatment on quality of life. J Neurol 249, 955–960 (2002)
- [5] Malavika Suresh, et al. “Real-Time Hand Gesture Recognition Using Deep Learning”, International Journal of Innovations and Implementations in Engineering(ISSN 2454-3489), 2019, vol 1
- [6] M. Gurbină, M. Lascu and D. Lascu, “Tumor Detection and Classification of MRI Brain Image using Different Wavelet Transforms and Support Vector Machines”, 42nd International Conference on Telecommunications and Signal Processing (TSP), Budapest, Hungary, 2019
- [7] Somasundaram S and Gobinath R, “Early Brain Tumour Prediction using an Enhancement Feature Extraction Technique and Deep Neural Networks”, International Journal of Innovative Technology and Exploring Engineering (IJITEE), ISSN: 2278-3075, Volume-8, Issue-10S, August 2019
- [8] Damodharan S and Raghavan D, “Combining Tissue Segmentation and Neural Network for Brain Tumor Detection”, The International Arab Journal of Information Technology, Vol. 12, No.1, January 2015
- [9] G. Hemanth, M. Janardhan and L. Sujihelen, “Design and Implementing Brain Tumor Detection Using Machine Learning Approach”, 3rd International Conference on Trends in Electronics and Informatics (ICOEI), Tirunelveli, India, 2019
- [10] A. R. Mathew and P. B. Anto, “Tumor detection and classification of MRI brain image using wavelet transform and SVM”, International Conference on Signal Processing and Communication (ICSPC), Coimbatore, 2017
- [11] W. Chen, X. Qiao, B. Liu, X. Qi, R. Wang and X. Wang, “Automatic brain tumor segmentation based on features of separated local square”, Chinese Automation Congress (CAC), Jinan, 2017
- [12] Navoneel Chakrabarty, “Brain MRI Images for Brain Tumor Detection Dataset”, Kaggle , April 2019
- [13] S. Poornachandra and C. Naveena, "Pre-processing of MR Images for Efficient Quantitative Image Analysis Using Deep Learning Techniques," 2017 International Conference on Recent Advances in Electronics and Communication Technology (ICRAECT), Bangalore, 2017, pp. 191-195, doi: 10.1109/ICRAECT.2017.43.
- [14] Mohammed Thanveersha N., et al. “Automatic Brain Hemorrhage Detection Using Artificial Neural Network”, International Journal of Innovations and Implementations in Engineering(ISSN 2454- 3489), 2019, vol 1
- [15] Soumya R S , et al. “Advanced Earlier Melanoma Detection Algorithm Using Colour Correlogram”, 2016 International Conference on Communication Systems and Networks (ComNet) | 21-23 July 2016 | Trivandrum
- [16] J. A. Akhila, C, Markose , et al. "Feature extraction and classification of Dementia with neural network," 2017 International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICT), Kerala, India, 2017, pp. 1446-1450
- [17] Avigyan Sinha, Aneesh R. P., “Real Time Facial Emotion Recognition using Deep Learning”, International Journal of Innovations and Implementations in Engineering(ISSN 2454-3489), 2019, vol 1
- [18] H. B. Nandpuru, S. S. Salankar, and V. R. Bora, “MRI brain cancer classification using support vector machine,” in Proc. IEEE Students’ Conf. Electr., Electron. Comput. Sci., Mar. 2014, pp. 1–6.
- [19] E.-S.-A. El-Dahshan, T. Hosny, and A.-B.-M. Salem, “Hybrid intelligent techniques for MRI brain images classification,” Digit. Signal Process., vol. 20, no. 2, pp. 433–441, Mar. 2010.
- [20] W. H. Ibrahim, A. A. A. Osman, and Y. I. Mohamed, “MRI brain image classification using neural networks,” in Proc. Int. Conf. Comput., Electr. Electron. Eng. (ICCEEE), Aug. 2013, pp. 253–258.
- [21] N. H. Rajini and R. Bhavani, “Classification of MRI brain images using knearest neighbor and artificial neural network,” in Proc. Int. Conf. Recent Trends Inf. Technol. (ICRTIT), Jun. 2011, pp. 563–568



# Automated Detection of Schizophrenia from Brain MRI Slices using Optimized Deep-Features

Seifedine Kadry

*Faculty of Applied Computing and  
Technology, Noroff University College,  
Kristiansand, Norway*  
email: skadry@gmail.com

Venkatesan Rajinikanth

*Department of Electronics and  
Instrumentation Engineering  
St. Joseph's College of Engineering  
Chennai 600119, India*  
email: v.rajinikanth@ieee.org

David Taniar

*Faculty of Information Technology  
Monash University  
Clayton, Victoria 3800, Australia*  
email: David.Taniar@monash.edu

Robertas Damaševičius

*Faculty of Applied Mathematics  
Silesian University of Technology  
44-100 Gliwice, Poland*  
email: robertas.damasevicius@polsl.pl

**Abstract**— In humans, the incident rate of mental illness is gradually rising due to various causes. Schizophrenia is one of the chronic mental illness and its happening rate also rising in the current era. The patient with Schizophrenia will experience a confused mental condition and a timely recognition and treatment is essential to reduce the risk. The proposed work aims to implement a methodology to support the automated detection of Schizophrenia from the brain MRI slices of T1 modality (T1W). The assessment of brain MRI is executed using the pre-trained VGG16 system and the deep-features extracted are optimized with the Slime-Mould-Algorithm (SMA) and the reduced features are then considered to train, test and validate the binary classifiers employed in this work. This research is implemented using 500 images of each case (healthy/abnormal) and the attained result with the SVM-Cubic is superior compared to other classifiers considered in the automated disease detection system.

**Keywords**—Mental illness, Schizophrenia, Brain MRI, VGG16, Slime-Mould-Algorithm, SVM-Cubic.

## I. INTRODUCTION

Brain is one of the vital organs in human physiology and the illness in brain will severely affect the common activities of human. The illness in the brain will severely affect the decision making and the control activities and hence, these disorders will be treated with higher priority [1-3]. Schizophrenia is one of persistent and harsh brain disorder and efficient detection and treatment planning is essential to reduce the severity level [4-7].

The report of World-Health-Organisation (WHO) authenticates that, the occurrence rate of this disorder is rising globally and around 20 million active cases are reported [8]. The other research on Schizophrenia occurrence rate also confirmed its harshness and incidence rate [9,10].

Schizophrenia is persistent brain disorder; which influence the person's ability related with common activities, such as sensing, judgment, and unmistakably behaviour [11]. Further, this disorder may cause; hallucination, illusion, and exceptionally chaotic thinking, which will disintegrate the usual performance of the individual. The cause of this disorder may be of hereditary motive or the environmental trigger. Usually, the indication of early signs of schizophrenia may visible from the age group 16 to 30 and the happening speed schizophrenia in men is large contrast to women. The warning sign of this disorder includes; (i) Hopelessness, (ii) Antagonism, (iii) Weakening of personal cleanliness, (iv) Insensitive gaze, etc. There are no

appropriate clinical handling procedures for the individual with harsh mental disorder and if the disease is detected in its early stage; an appropriate treatment can be implemented to cure the patient.

Because of its harshness, a range of diagnostic approaches are recommended to control its impact in humans. The common detection methodology includes; (i) Detection using a multi-channel EEG and (ii) brain MRI assisted diagnostic. The assessment schizophrenia using the EEG is reported in many earlier studies and these methods implemented the signal assisted and well as the signal to image supported automated detection schemes to efficiently detect this disorder using machine-learning approaches. The assessment of the EEG is quite complex and a necessary pre-processing methods are essential to treat the signals due to its non-linear nature. Along with the EEG, MRI supported schemes are also implemented to examine this illness with better accuracy and the earlier works related to this research can be found in [4-7,10].

This research aims to develop an automated examination system to detect the destructive brain irregularities from the brain MRI with better accuracy. In this work, a pre-trained VGG16 scheme is implemented to classify the brain MRI slices into healthy/abnormal (schizophrenia) class with the help of a binary classifier. The essential test images for this experimental investigation are attained from [12,13]. This dataset consist a clinical grade 3D brain MRI of T1 modality (T1W) and for this research only the axial slices are considered.

Initially, a 3D to 2D conversion is executed using ITK-Snap [14,15] and the extracted images are then resized to the required image dimension. The extracted brain MRI slices are having an initial dimension of 176x256x3 pixels and the resizing is implemented to get an image with size of 224x224x3 pixels. All these images (axial view) are associated with the skull section and for the assessment the images are considered with the skull section. In this work 500 images of normal/abnormal cases are considered for experimental investigation. The original and the augmented images are considered to pre-train the VGG16 architecture and the outcome of the features are then reduced using the Slime-Mould-Algorithm (SMA) and the optimised features are then considered to activate the binary classifier employed in this work. For the classification task, 70% images (350 numbers) are considered for training the classifier and 30% images (150 numbers) are considered for validation. In this work, a 10-fold cross validation is employed and the best classification result attained is considered as the final result.

The proposed work is implemented using the traditional VGG16 with fully-connected-layer (FCL) and the SMA optimizer and every approach is separately examined with the chosen classifiers. The experimental result attained with the proposed technique confirms that the SVM-Cubic with SMA optimized features helped to achieve a better result compared to other methods implemented in this research.

The upcoming regions are prearranged as follows; Section 2 depicts the methodology, Section 3 and 4 presents' experimental results and conclusion of this research.

## II. METHODOLOGY

This part presents the methodology executed in this work and the various stages involved in this work is presented in Figure 1.

The essential test images are collected from the benchmark dataset and in order to reduce the computational complexity; the 3D to 2D image conversion is implemented using the ITK-Snap tool. The 2D brain MRI slice (axial view) is then resized to 224x224x3 pixels and the resized image (original and augmented) is then considered to train the VGG16 scheme, till it learns to extract the essential disease feature from the test images. When the VGG16 is completely trained with the proposed dataset, then 70% of test images are considered for the training process and then extracted features are considered to train and test the classifiers. For the validation process, 30% of images are considered and the results attained with a binary classifier are then considered and the final result.

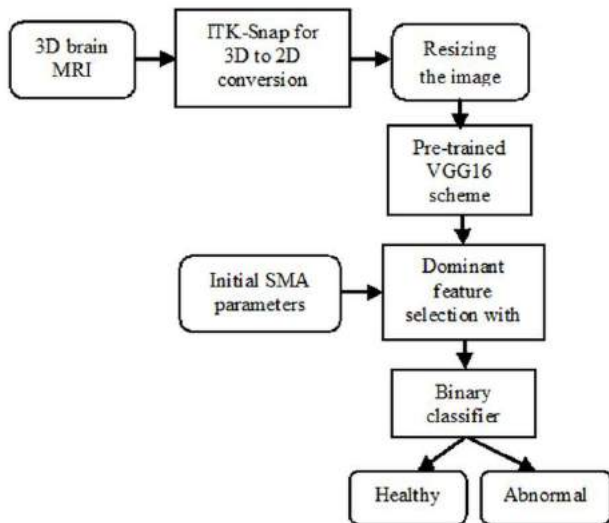


Fig.1. Proposed Schizophrenia examination scheme using the brain MRI slice

In this research, the detection of Schizophrenia is implemented using; (i) Pre-trained VGG16 with appropriate FCL with 50% dropout to reduce 4096 features into 1024 features and (ii) Pre-trained VGG16 with SMA based feature selection process to select the dominant feature from 4096 feature. The results of these two methods are separately evaluated using the chosen binary classifiers and based on the attained values; the performance of the VGG16 on the considered database is confirmed.

### A. Image database

The abnormality in brain is normally examined using the imaging technique due to its simplicity and the abnormality in medical images can be easily checked and confirmed with a visual check by an experienced doctors.

Hence, most of the brain disorders are examined with imaging procedures and this work considered the 3D brain MRI of the Schizophrenia existing at [12,13]. This database consist 99 patient's information recorded with T1W modality and in this work only 500 image slices from healthy/abnormal case are considered for the experimental investigation.

In this dataset, the raw image existing is in 3D form and hence a 3D to 2D conversion is implemented to extract the axial view brain MRI for the experimental evaluation. This image is associated with the skull section and every extracted slice is available with a size of 176x256x3 pixels. The VGG16 needs only a prescribed dimension images and hence, every image of this database is resized to 224x224x3 pixels and the resized image is then considered for assessment.

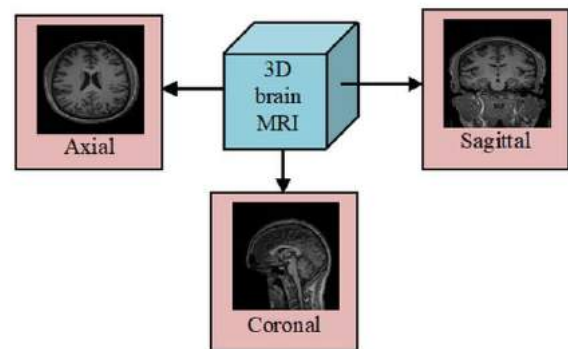


Fig.2. Various 2D slices extracted form the 3D brain MRI

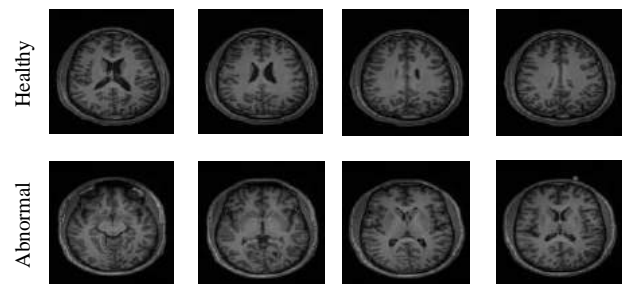


Fig.3. Sample test images collected from the considered image database

TABLE I. TEST IMAGES CONSIDERED IN THIS RESEARCH

Class	Image dimension		Number of images		
	Original	Resized	Total	Training	Testing
Healthy	176x256x3	224x224x3	500	350	150
Abnormal	176x256x3	224x224x3	500	350	150

Figure 2 depicts the extraction of different 2D MRI slices and the axial view is considered for evaluation. The sample test image of healthy and abnormal class is depicted in Figure 3 and the numbers of images considered to evaluate the performance of VGG16 are presented in Table I.

### B. VGG16 Architecture

Deep-Learning (DL) supported disease detection is attracted the researchers due to its diagnostic accuracy and clinical significance [16-18]. The earlier works on the DL system confirms that this method helps to achieve a better result on a number of medical image modalities and support the automated detection of the disease with enhanced results. Most of the DL scheme will help to classify the existing test image dataset into a binary (two-class) or multi-class category which is then confirmed based on the attained values of the Image Performance Measures (IPM). The success of the implemented DL depends on; (i) Chosen scheme, (ii) Feature reduction process, and (iii) Classifier. Most of the pre-trained DL system is having the inbuilt SoftMax classifier which offers a satisfactory result on most of the image cases. Based on the need, the SoftMax can be replaced with other classifier units existing in the literature. Figure 4 depicts the conventional VGG16 scheme widely adopted for medical image examination jobs.

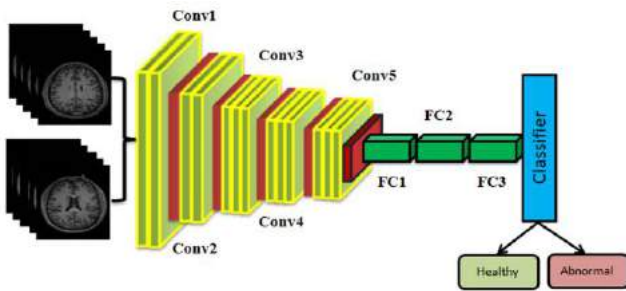


Fig.4. Pre-trained VGG16 with traditional architecture

### C. Slime-Mould-Algorithm based Feature Selection

The performance of the DL unit mainly depends on the extracted features by the dense-layers and in VGG16, the initial layer will help to get  $1 \times 1 \times 4096$  features and all the features are ranked based on their values and then sufficient 50% dropout is implemented in every layer to get a final 1D feature vector of dimension  $1 \times 1 \times 1024$  features. In this work, instead of employing the 50% dropout, the dominant features are selected using the Slime-Mould-Algorithm (SMA) based feature selection methodology. The structure of the proposed methodology is depicted in Figure 5 and the optimally selected feature is then considered to train and validate the classifier with a 10-fold cross validation.

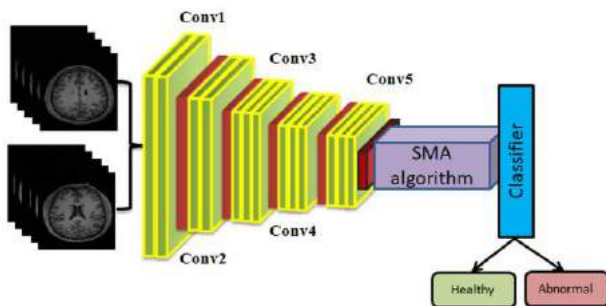


Fig.5. VGG16 with SMA based dominant feature selection

The SMA is a recently invented approach based on the food searching mechanism existing in a single cell Slime-Mould [19]. Due to its merit, it was widely adopted by the researchers to find optimal solutions for a variety of problems. Compared to recent optimization techniques, the

SMA initial parameters are less and this includes; agent's size, search dimension, upper and lower-bound, search style and the terminating condition. The fundamental code for the SMA is existing in [20]. The necessary information and the arithmetic explanation of SMA are available in [20,21]. In this work, the following algorithm values are initiated; agent size=25, search dimension=2, upper/lower-bound=assigned based on the lower and upper values of attained p-value with Student's t-test, number of iterations=3000 and stopping criteria = maximal Cartesian distance between the features.

The implementation of heuristic algorithm supported feature selection is clearly discussed in earlier work [22]. In this work, the aim of this task is to identify optimal features (n-numbers) from 4096 features and this process is depicted in Figure 6.

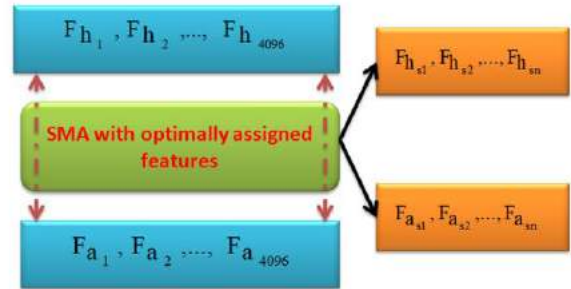


Fig.6. SMA supported feature reduction

Let the extracted features from VGG16 is depicted as in Eqn. (1) and (2);

$$\text{Healthy class} = F_{h_1}, F_{h_2}, \dots, F_{h_{4096}} \quad (1)$$

$$\text{Abnormal class} = F_{h_1}, F_{h_2}, \dots, F_{h_{4096}} \quad (2)$$

$$\text{Selected Healthy class} = F_{h_{s1}}, F_{h_{s2}}, \dots, F_{h_{s614}} \quad (3)$$

$$\text{Selected Abnormal class} = F_{h_{s1}}, F_{h_{s2}}, \dots, F_{h_{s614}} \quad (4)$$

When the SMA optimization task is implemented then it helps to reduce the  $1 \times 1 \times 4096$  features into  $1 \times 1 \times 614$  features (n) as in Eqns. (3) and (4). These features are then considered to train and validate the classifiers employed in this work.

### D. Classifier Implementation

The performance of the disease diagnosis scheme relies main on the classifier employed to categorize the considered images into two/multi class. In this work, the inbuilt SoftMax classifier is initially considered and later its performance is validated with other techniques, such as K-Nearest Neighbour (KNN), SVM-Linear kernel (SVM-L), SVM-RBF kernel (SVM-RBF) and SVM-Cubic and the related information on these units can be found in [23]. In this work, a 10-fold cross validation is employed and the best result among the trials is chosen as the final result.

### E. Validation

The aim of every research work is to implement a suitable automated disease diagnostic system. The success of this system needs to be tested and confirmed before using it in real-time. The common evaluation procedure is computing the Image-Performance-Values (IPV) during the assessment and based on these values, the merit of the proposed scheme is to be confirmed. The evaluation of IPV

based on the testing images is widely considered and in this work 150 images (30% of database) are adopted to compute the IPV.

The initial IPV's include; True-Positive (TP), False-Positive (FP), False-Negative (FN), and True-Negative (TN), further the other measures, such as Accuracy (ACC), Precision (PRE), Sensitivity (SEN), Specificity (SPE), F1-Score (F1S) and Negative-Predictive-Value (NPV) are also measured to validate the merit of VGG16 [1-3].

### III. RESULT AND DISCUSSION

The experimental results are obtained using the workstation; Intel i5 2.5GHz processor with 16GB RAM and 2GB VRAM set with MATLAB®.

Initially, the considered VGG16 architecture is trained with Schizophrenia dataset (including original and augmented images). After the essential training (after getting better training accuracy), the proposed scheme is tested with 150 images and the attained performance measures are recorded. In this work, a 10-fold cross validation is implemented and the best IPV attained is selected as the final outcome of the scheme.

Figure 7 depicts the experimental outcome of the VGG16 with the SMA selected features and the result confirms the better categorization accuracy with SVM-Cubic classifier. Similar procedure is executed using the learned features and the SMA optimized features and the corresponding results are depicted in Table II and Table III. From these tables, it can be confirmed that the proposed scheme helps to achieve a better result with SMA optimized features.

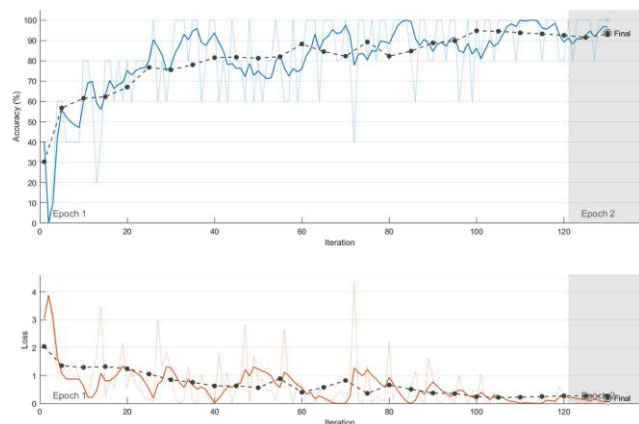


Fig.7. The training and testing accuracy attained for VGG16 with SMA optimization

The results attained in Table III confirm that the classifier performance is good with optimal-features compared to learned-features. Further, in both the cases, the classifier accuracy of SoftMax is better compared to KNN and SVM-L. For the learned-features case, the SVM-RBF provided the better accuracy and for optimal-features case the SVM-Cubic is provided enhanced results. Further, the proposed approach helps to achieve a better IPV compared to the traditional VGG16 with the learned-feature vector. The sample results wrongly predicted during the SVM-Cubic classifier is presented in Figure 8. In Figure 8, the Glyph-plot for the IPV's attained with optimal-features are depicted and from this figure, it can be confirmed that SVM-Cubic provided better accuracy (94.33%)

TABLE II. THE INITIAL OUTCOME ATTAINED WITH THE BINARY CLASSIFIER

Approach		TP	FN	TN	FP
Learned features (1x1x1024)	SoftMax	136	14	132	18
	KNN	135	15	130	20
	SVM-L	136	14	131	19
	SVM-RBF	138	12	134	16
	SVM-Cubic	134	16	137	13
Optimal features (1x1x614)	SoftMax	137	13	142	8
	KNN	140	10	138	12
	SVM-L	141	9	140	10
	SVM-RBF	138	12	139	11
	SVM-Cubic	141	9	142	8

TABLE III. THE PEST PERFORMANCE VALUES COMPUTED WITH PROPOSED CLASSIFIER

Approach		ACC (%)	PRE (%)	SEN (%)	SPE (%)	F1S (%)	NPV (%)
Learned features (1x1x1024)	SoftMax	89.33	88.31	90.67	88.00	89.47	90.41
	KNN	88.33	87.10	90.00	86.67	88.52	89.66
	SVM-L	89.00	87.74	90.67	87.33	89.18	90.34
	SVM-RBF	90.67	89.61	92.00	89.33	90.79	91.78
	SVM-Cubic	90.33	91.16	89.33	91.33	90.23	89.54
Optimal features (1x1x614)	SoftMax	93.00	94.48	91.33	94.67	92.88	91.61
	KNN	92.67	92.11	93.33	92.00	92.71	93.24
	SVM-L	93.67	93.38	94.00	93.33	93.69	93.96
	SVM-RBF	92.33	92.62	92.00	92.67	92.31	92.05
	SVM-Cubic	94.33	94.630	94.00	94.67	94.31	94.04

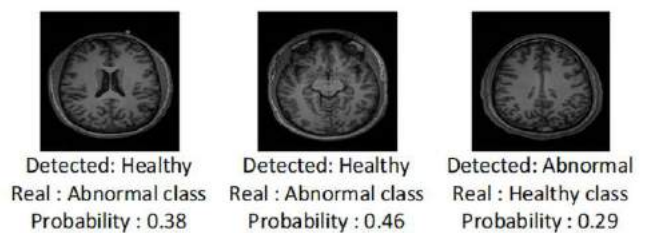


Fig.8. Classifier outcome with sample images belongs to FN and FP

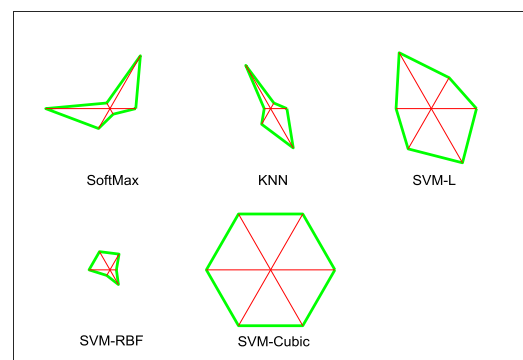


Fig.9. Glyph-plot of IPV's attained with SMA optimized features

The proposed scheme employed the VGG16 based assessment and in future, other deep-learning schemes can be considered to examine the brain MRI database. Further, the performance of SMA can be compared and validated against the recently developed Red-Fox-Algorithm [24].

#### IV. CONCLUSION

In this work an automated practice is proposed to detect the Schizophrenia from the axial-view brain MRI slices of T1W modality. This work employed the well known deep-learning scheme; VGG16 with a binary classification. The classification for the considered test images are executed using the learned-features (1x1x1024) and SMA optimized features (1x1x614). In this work, a 10-fold cross validation is implemented and the best value of attained result is adopted for validation. The experimental outcome of this technique confirms that the classification accuracy of VGG-Cubic is better compared approaches. The proposed scheme is efficient in evaluating the disease using the brain MRI and in future, it can be considered to examine the clinically collected real time images.

#### REFERENCES

- [1] M.A. Khan, I. Ashraf, M. Alhaisoni, R. Damaševičius, R. Scherer, A. Rehman, and S.A.C. Bukhari, "Multimodal brain tumor classification using deep learning and robust feature selection: A machine learning application for radiologists," *Diagnostics*, vol. 10, no. 8, pp. 565, 2020.
- [2] V. Rajinikanth, A.N.J. Raj, K.P. Thanaraj, and G.R. Naik., "A customized VGG19 network with concatenation of deep and handcrafted features for brain tumor detection," *Applied Sciences* 10, no. 10, pp.3429, 2020. <https://doi.org/10.3390/app10103429>.
- [3] M. A. Khan, I. Ashraf, M. Alhaisoni et al., "Multimodal brain tumor classification using deep learning and robust feature selection: A machine learning application for radiologists," *Diagnostics*, vol. 10, no. 8, pp.565, 2020. <https://doi.org/10.3390/diagnostics10080565>.
- [4] M.E. Shenton, C.C. Dickey, M. Frumin, and R.W. McCarley, "A review of MRI findings in schizophrenia," *Schizophrenia research*, vol. 49, no. 1-2, pp.1-52, 2001.
- [5] J. Oh, B-L. Oh, K-U. Lee, J-H. Chae, and K. Yun, "Identifying Schizophrenia Using Structural MRI With a Deep Learning Algorithm," *Frontiers in psychiatry*, vol. 11, pp.16, 2020.
- [6] R.G. Steen, C. Mull, R. McClure, R.M. Hamer, and J.A. Lieberman, "Brain volume in first-episode schizophrenia: systematic review and meta-analysis of magnetic resonance imaging studies," *The British Journal of Psychiatry*, vol. 188, no. 6, pp. 510-518, 2006.
- [7] V. Jahmunah, S.L. Oh, V. Rajinikanth, E.J. Ciaccio, K.H. Cheong, N. Arunkumar, and U. R. Acharya, "Automated detection of schizophrenia using nonlinear signal processing methods," *Artificial intelligence in medicine*, vol. 100, pp.101698, 2019. <https://doi.org/10.1016/j.artmed.2019.07.006>.
- [8] <https://www.who.int/news-room/fact-sheets/detail/schizophrenia>
- [9] C. Devia et al., "EEG classification during scene free-viewing for schizophrenia detection," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 27, no. 6, pp.1193-1199, 2019. DOI: 10.1109/TNSRE.2019.2913799.
- [10] S. Arunmozhi, N.S.M. Raja, V. Rajinikanth, K. Aparna, and V. Vallinayagam, "Schizophrenia Detection using Brain MRI—A Study with Watershed Algorithm," In 2020 International Conference on System, Computation, Automation and Networking (ICSCAN), pp. 1-4. IEEE, 2020. DOI: 10.1109/ICSCAN49426.2020.9262281.
- [11] <https://www.mayoclinic.org/diseases-conditions/schizophrenia/symptoms-causes/syc-20354443>
- [12] G. Repovš, and D. M. Barch, "Working memory related brain network connectivity in individuals with schizophrenia and their siblings," *Frontiers in human neuroscience*, vol. 6, pp.137, 2012. DOI: 10.3389/fnhum.2012.00137.
- [13] <https://openneuro.org/datasets/ds000115/versions/00001>
- [14] P.A. Yushkevich, J. Piven, H. C. Hazlett, R. G. Smith, S. Ho, J. C. Gee, and G. Gerig, "User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability," *Neuroimage*, vol.31, no.3, pp. 1116-1128, 2006. <https://doi.org/10.1016/j.neuroimage.2006.01.015>.
- [15] <http://www.itksnap.org/pmwiki/pmwiki.php?n=Main.Publications>
- [16] M. Nisa, J. H. Shah, S. Kanwal, M. Raza, M. A. Khan, R. Damaševičius, and T. Blažauskas, "Hybrid malware classification method using segmentation-based fractal texture analysis and deep convolution neural network features," *Applied Sciences*, vol. 10, no. 14, pp.4966, 2020.
- [17] S.A. Roseline, S. Geetha, S. Kadry, and Y. Nam, "Intelligent Vision-Based Malware Detection and Classification Using Deep Random Forest Paradigm," *IEEE Access*, vol. 8, pp.206303-206324, 2020.
- [18] N. Dey, Yu-Dong Zhang, V. Rajinikanth, R. Pugalenthii, and N.S.M. Raja, "Customized VGG19 architecture for pneumonia detection in chest X-rays," *Pattern Recognition Letters*, vol.143, pp.67-74, 2021. <https://doi.org/10.1016/j.patrec.2020.12.010>.
- [19] S. Li, H. Chen, M. Wang, A. A. Heidari, and S. Mirjalili, "Slime mould algorithm: A new method for stochastic optimization," *Future Generation Computer Systems*, vol. 111, pp.300-323, 2020.
- [20] <https://in.mathworks.com/matlabcentral/fileexchange/76619-slime-mould-algorithm-sma-a-method-for-optimization>
- [21] M. Premkumar et al., "MOSMA: Multi-Objective Slime Mould Algorithm Based on Elitist Non-Dominated Sorting," *IEEE Access*, vol.9, pp. 3229 – 3248, 2020. DOI: 10.1109/ACCESS.2020.3047936.
- [22] A. Bakiya, K. Kamalanand, V. Rajinikanth, R.S. Nayak, and S. Kadry, "Deep neural network assisted diagnosis of time-frequency transformed electromyograms," *Multimedia Tools and Applications*, vol. 79, no. 15, pp.11051-11067, 2020.
- [23] M.A. Khan et al., "Computer-aided gastrointestinal diseases analysis from wireless capsule endoscopy: A framework of best features selection," *IEEE Access*, vol. 8, pp.132850-132859, 2020. DOI: 10.1109/ACCESS.2020.3010448.
- [24] D. Połap, and M. Woźniak, "Red fox optimization algorithm," *Expert Systems with Applications*, vol.166 pp.114107, 2021. <https://doi.org/10.1016/j.eswa.2020.114107>.

# COVID-SEGNET: Diagnosis of Covid-19 Cases on Radiological Images using Mask R-CNN

Avik Kundu  
School of Computer Engineering  
KIIT University  
Bhubaneswar, India  
avikkundu8d@gmail.com

Chandan Mishra  
School of Computer Engineering  
KIIT University  
Bhubaneswar, India  
chandanapmishra2017@gmail.com

Saurabh Bilgaiyan  
School of Computer Engineering  
KIIT University  
Bhubaneswar, India  
saurabhbilgaiyan01@gmail.com

**Abstract**—The novel coronavirus (COVID-2019) pandemic has caused a catastrophic effect on the health and global economy. Early screening and diagnosis of COVID-19 pneumonia are considered to be the critical step to stop the further spread of the virus. The most common standard for confirming the virus relies on RT-PCR tests. This method generates false-negative results if there is limited viral load. Recent radiological findings suggest that the distinct distribution of ground-glass opacities (GGOs), which are found on certain parts of lungs, can determine the status of the infection among patients. As a complement to RT-PCR, Computed tomography (CT) can be used for diagnosing COVID-19. In this study, the authors have described a Mask R-CNN (region-based convolution neural network) approach for the detection of the ground glass opacities (GGOs) in chest CT images of COVID-19 infected persons. The proposed approach provides an accuracy of 98.25% during instance segmentation. Therefore, the authors believe this proposed method will aid health professionals to fasten the screening and validation of the initial assessment towards COVID-19 patients.

**Index Terms**—Coronavirus (COVID-19); Instance segmentation; Mask R-CNN; Computer Vision; Chest CT images

## I. INTRODUCTION

The outbreak of the COVID-19 pandemic in Wuhan, China has put tremendous load over healthcare facilities throughout the world. The limited availability of kits for diagnosis and proper ventilation systems for treatment, has caused healthcare systems of several countries to collapse. Several countries have been forced to incorporate nation-wide lock downs and curfews to curb the spread of the virus. The most popular standard for virus validation is based on RT-PCR tests. But this strategy generates false negatives if the viral load is not sufficient. Alternatively, Radiography Scans, Chest Radiography or Computed Tomography Imaging are considered by radiologists to classify based on certain specific visual markers.

In this paper, the authors have used Mask R-CNN approach to identify the ground glass opacities by performing instance segmentation on chest CT images [1]. The main contributions of the authors are the identification of the locations of the GGOs found typically with a peripheral and sub-pleural distribution, which is the main CT feature of viral infection [2]. In the proposed study, the authors were able to achieve an accuracy of 98.25% during instance segmentation, through which one can infer that the model would perform well even during classification.

## II. IMAGE SEGMENTATION

Image segmentation is a crucial component in various systems involving visual data. It requires the division of digital images into different segments. In a large variety of applications it has a critical function, like analysing medical images (e.g., finding tumours in the brain), video surveillance, augmented reality, autonomous vehicles etc.

Recently, deep learning models are being implemented in various computer vision applications and research related to image segmentation have gained popularity.

Image segmentation can be further classified into 2 types:

- **Semantic segmentation:** Semantic segmentation process involves marking pixels with a variety of entities (eg., balloon, vehicle, tree, human, animal) for all image pixels.
- **Instance segmentation:** Instance segmentation [3] is a further identification and delineation of each object of interest in the image.

As an example, consider an image consisting of balloons **Fig. 1**, there will be 4 fundamental steps to be followed for Instance Segmentation. First is classification which means one needs to specify the model that there is a balloon in this image. Then comes semantic segmentation in which the model will detect all balloon pixels. After performing semantic segmentation object localization is performed i.e, there are 7 balloons in this image **Fig. 1** at these coordinates (overlapping of images is also taken into consideration). Last step consists of Instance segmentation of the 7 balloons i.e, the model will specify the pixels at which balloons are present.

## III. GROUND GLASS OPACITIES (GGOs)

The SARS-CoV-2 virus causes typical fever, cough, myalgia and dyspnoea but sometimes lead to severe pneumonia mostly among senior citizens. These in-turn causes extreme damage to lungs. In majority cases 80% shows mild symptoms, 14% show symptoms of pneumonia, 5% suffer from organ failure (mainly respiratory organs like lungs) and for 2% of the cases, it turns to be fatal. Covid-19 severely damages the lungs and the alveoli (which is responsible for transferring oxygen to the blood vessels) [4]. **Fig. 2** consists of 4 CT images of lungs

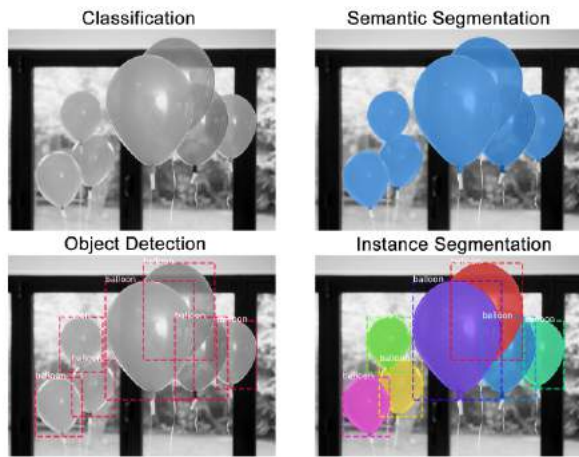


Fig. 1. Various steps to be undergone to perform Image Segmentation

of COVID-19 infected patients. The arrows on these images point to the regions where there is high possibility of the presence of GGOs.

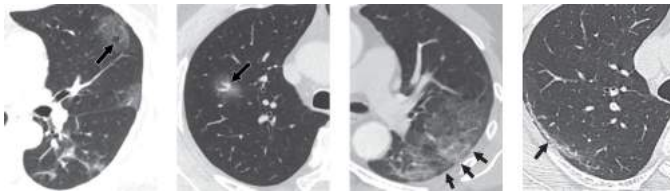


Fig. 2. CT images of lungs pointing at the GGOs [5]

The CT scans of lungs of the affected patients showed a common manifestation i.e., the existence of multifocal nodular ground-glass opacities (GGOs), typically showing sub-pleural and peripheral scattering [6]. GGO could be identified by observing white-flecked patterns seen on lung CT scans. It has been confirmed by medical experts that GGOs are not found in healthy lungs or exposure to air pollution or smoking. Specifically ground-glass opacities can be caused by various other lung diseases but there's a distinct distribution i.e., preference of certain parts of lungs in case of COVID-19 affected patients [5].

#### IV. CHEST CT DATASET

In this study, CT scans were provided by local hospitals situated in Moscow, Russia dated from 1st March, 2020 to 25th April, 2020 [7]. The dataset can be downloaded from the link provided: [https://mosmed.ai/datasets/covid19\\_1110](https://mosmed.ai/datasets/covid19_1110). The dataset of 1110 studies contains findings of chest CT scans affected by COVID-19 Fig. 3

The dataset was stored in NifTi format and compressed using Gzip. Some parts of the study were annotated by the medical experts. The GGOs were selected as +ve pixels for each image on the corresponding binary pixel mask during the time of annotation of the dataset. MedSeg®, a web-based annotation tool named was used in creation of binary masks.

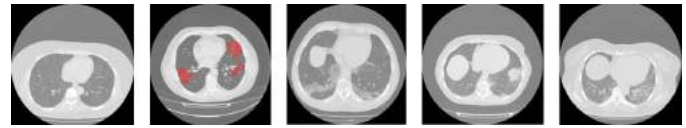


Fig. 3. Datasets of CT images for (first image) normal case and (rest of images) COVID-19 cases [7].

The extracted images from the benchmarked dataset were not in the desired format that the model was designed to accept. Since the benchmark dataset provides the images in the form of .nii files for the Mask R-CNN, the authors had to convert them into the desired format which consists of a .png file and .json file which stores the x and y coordinates of the segmentation regions along with image name and unique id. The conversion of images to the desired format had an accuracy of 99.99%. The next step involved, dividing the total images between training and validation following the ratio of 4:1.

#### V. PROPOSED WORKFLOW

Fig. 4 displays the images obtained from the proposed model's intermediate stages. The chest CT scan images are passed into the neural network of Convolution, which helps to assess the edges. Since they got a feature map from CNN it can now be fed to the subsequent layers. The next layer of the model is the Region Proposal Networks abbreviated as RPN. In order to generate "proposals" for region where GGOs tend to exist a small network is fit after convolution feature map i.e. the output by the last convolution layer.

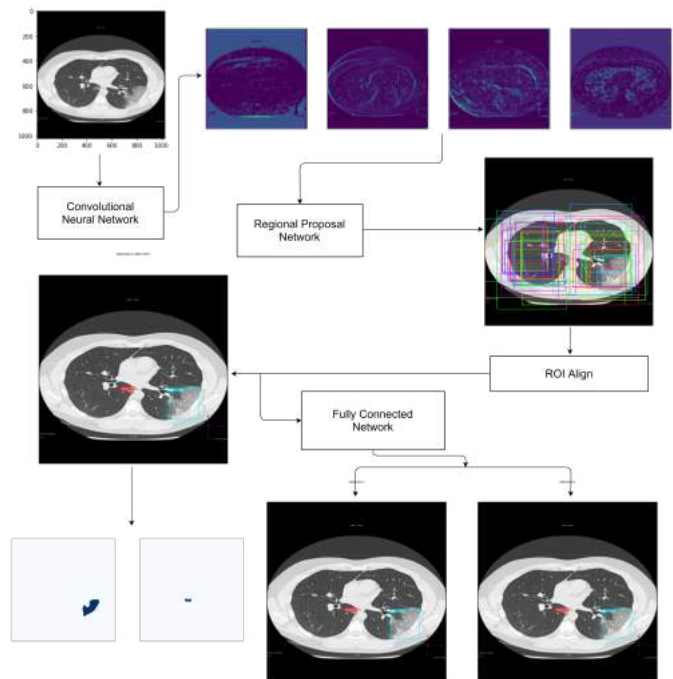


Fig. 4. Output generated from each module of Mask R-CNN

The feature map then enters the RPN for getting the probable locations of the ground glass opacities. It uses anchors which are a set of boxes with predefined locations and according to the inputted image it scales itself. Individual anchors are assigned to the ground-truth classes and bounding boxes.

It generates 2 outputs for each anchor. The first output is the detection of classes and the second is the Bounding Box Refinement. To be more specific it encapsulates the region of interest in the case the GGOs. After finding the probable regions perform RoI alignment, the feature map is sampled at different points in RoI Align's case and bilinear interpolation is applied to get a precise picture of what the pixel will be at 2.93. Then the fully connected layer is called, which is used to generate masks for the regions selected by the RoI classifier.

## VI. METHODOLOGY

### A. Mask RCNN

In modern technology, segmentation of images has a major contribution towards the world of Computer Vision. These segmentations are performed using techniques of Deep learning. R-CNN (Region-Based Convolution Neural Networks) is the most notable Network, based on which, Fast Region with CNN (Fast R-CNN) [8] and Mask R-CNN have been proposed. Mask R-CNN [1] is an efficient general framework for detection, localization, and instance segmentation of objects in natural images. From Fig. 5 different modules of Mask R-CNN can be visualized.

Mask R-CNN consists of 2 stages :

- The principal stage checks the picture and creates proposals (areas liable to contain an item).
- The subsequent stage characterizes the recommendations and produces bounding boxes and covers.

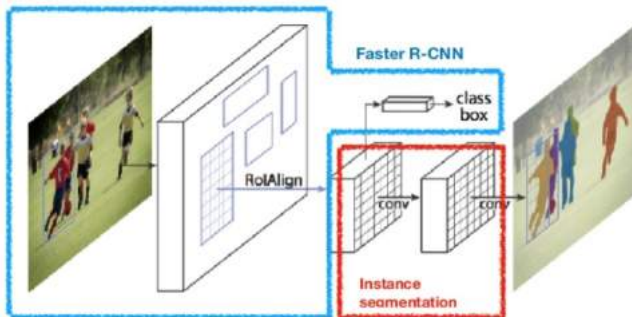


Fig. 5. Mask R-CNN framework for Instance Segmentation

Instance segmentation is difficult because of its needs. It requires accurate detection of all entities in an image along with segmentation of each instance precisely. To overcome this difficulty, the authors have used Mask R-CNN which is an add-on to Faster R-CNN. It follows the same framework but the detection branch and mask branch run in parallel to each other

i.e., the FCN semantic segmentation [9], segmentation jobs and arrangement, classifications are all parallelly executed. Mask R-CNN has received 3 major changes as compared to Faster R-CNN i.e., the use of Feature Pyramid Networks (FPN), replacement of RoI pooling with RoI align [10]. Use of RoI align increased the accuracy of mask R-CNN from 10% to 50%. The loss function of the Mask R-CNN model can be represented as:

$$L = L_{cls} + L_{box} + L_{mask} \quad (1)$$

The loss is similar to Faster R-CNN for classification and box regression. A sigmoid per pixel is added to each map. The map loss is then described as an average loss of binary cross entropy. Mask R-CNN's multi-task loss function incorporates the loss of mask labelling ( $L_{cls}$ ), localization ( $L_{box}$ ) and segmentation ( $L_{mask}$ ). Mask loss ( $L$ ) is only defined for the ground truth class. It decouples class prediction and mask generation, and thus empirically better results and model becomes easier to train.

### B. Training Method

For training the Mask R-CNN model, random weights were initialized. ResNet-101 feature pyramid network model was chosen as the backbone of the model. The learning rate was set to  $10^{-3}$  and the learning momentum was set to 0.9. The size of the images had to be scaled down to 640x480 to work efficiently. The batch size was set to 2 and the model was trained for 30 epochs on a Tesla K80 with 12GB GDDR5 VRAM.

The weights obtained during the model training process, were continuously saved after each epoch. Since Early Stop was implemented, the model automatically stopped the training process once the accuracy did not improve for more than 3 epochs. The final weight was re-loaded in place of the random weights during the validation process. After successful completion, an accuracy of 98.25% was achieved Fig. 6.

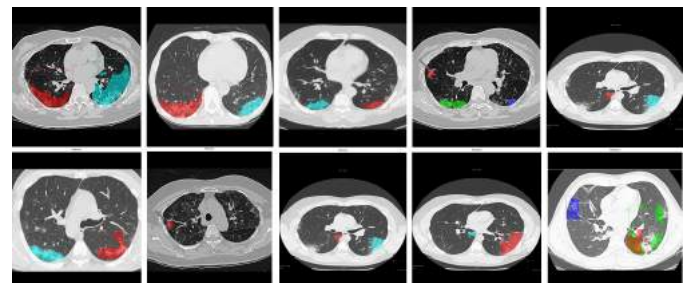


Fig. 6. Output validation images obtained after performing Mask-RCNN segmentation on chest CT images



## VII. EXPERIMENTAL RESULTS

In Fig. 7 the ground glass opacities found on the chest CT image, can be clearly identified. The instance segmentation performed by the Mask R-CNN model, had perfectly masked the GGOs present on the image. Thus, by looking at the output image, a medical practitioner can be fully sure that the patient is suffering from the virus. Thus, the authors have inferred that Mask R-CNN can significantly improve the accuracy of image segmentation.

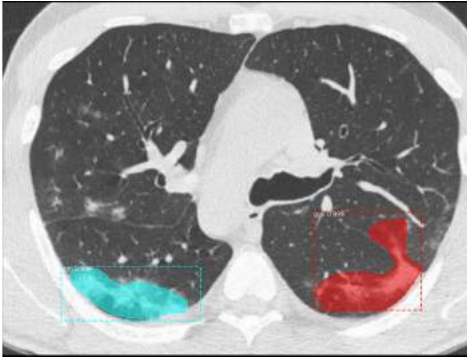


Fig. 7. Final Prediction of the GGOs found on the chest CT of COVID-19 patient

### A. Evaluation Criteria

Several evaluation criteria like region overlap and boundary similarity are proposed to evaluate image segmentation results [11]. The authors have selected Dice coefficient (DICE) [12] to evaluate the overlapping of prediction and the ground-truth.

The Dice indices are derived from the following formula:

$$DICE = \frac{2|\alpha \cap \beta|}{|\alpha| + |\beta|} \quad (2)$$

In this case, the Sorensen index, given the two masks ( $\alpha$ ) and ( $\beta$ ), is twice the number of pixels common to both masks, divided by the total of each mask's number of pixels. An average Sørensen–Dice index score of **0.764(GGO)** was achieved by the backbone of the proposed model.

### B. Related Works

Most of the research works related to COVID-19 virus, were based on the classification of COVID-19 and normal patients. Till now, no study has been conducted regarding instance segmentation for detection of ground glass opacities, which forms a visible unique pattern on the CT images of the COVID-19 infected patients. In Table 1, the authors have compared various Models developed by other researchers in the field of Classification of COVID-19 X-Ray/CT Images.

### C. Statistical Analysis

In Fig. 8 X-Axis reflects Epochs and Y-Axis reflects the value of loss. There was an overall decrease in loss of the model, which signifies that the model can predict the correct class. They got an **average precision (AP=50) of 0.700** from the precision recall curve Fig. 9, where X-Axis reflects recall

TABLE I  
CLASSIFICATION PERFORMANCE OF DEEP LEARNING MODEL PROPOSED AFTER RE-TRAINING.

Method	Accuracy %	COVID Images	Normal Images
M-Inception [13]	82.90	195	258
UNet+3D Deep Network [14]	90.80	313	229
<b>Deep CNN ResNet-50 [15]</b>	<b>98.00</b>	<b>50</b>	<b>50</b>
ResNet [16]	86.70	219	224
DarkCovidNet [17]	87.02	125	500
COVIDX-Net [18]	90.00	25	25
COVID-Net [19]	92.40	53	5526
ResNet50+ SVM [20]	95.38	25	25
VGG-19 [21]	93.48	224	700
DRE-Net [22]	86.00	777	708

and precision is expressed by Y-Axis. In case of precision-recall graph, the curve close to the PRC for a perfect test have a better performance level as compared to those close to the baseline.

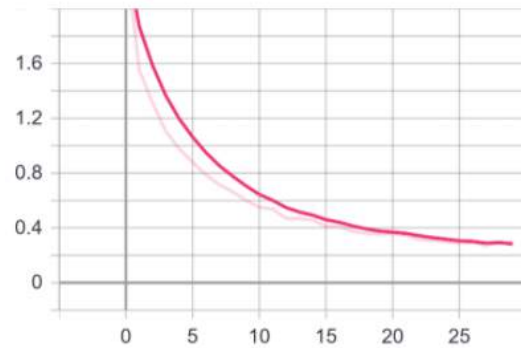


Fig. 8. Overall loss of the model

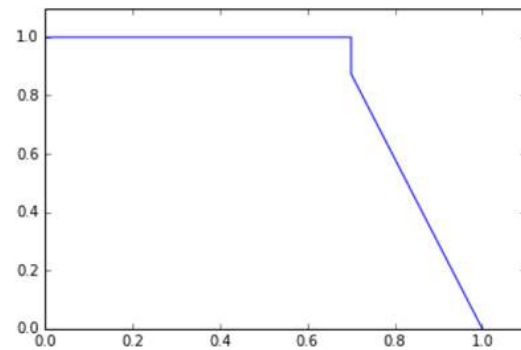


Fig. 9. PR-curve of Mask R-CNN

In Fig. 10 X-Axis represents Epochs and Y-Axis represents loss value of different segments of the model. One can observe a decreasing trend in the loss values of RPN layer which means the model can precisely locate the expected regions. Similarly, mask loss penalizes wrong per-pixel binary classifications which in this case, shows a steady decrease in penalty. The distance between the true box parameters is reflected by bounding box loss values i.e. the coordinates of the box

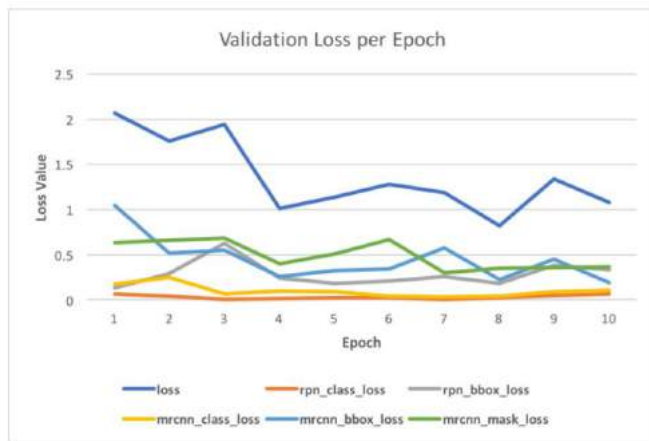


Fig. 10. There is an overall decrease in validation loss

locations. `rpn_bbox_loss` represents how accurate a model is at locating objects with an image and `mrcnn_bbox_loss` represents how good a model is at precisely predicting the areas inside an image corresponding to the different objects present. A decreasing trend in these graphs Fig. 10 show the model is well trained and is not overfitted. However the fluctuations in the validation loss may occur either due to high learning rate or due to small size of validation set.

### VIII. CONCLUSION AND FUTURE WORK

COVID-19 disease is continuing to spread despite the precautions taken by the people and government regulations. In this study, the authors have proposed COVID-SEGNET, Mask R-CNN-based approach for the detection of ground glass opacities (GGOs) in chest CT images of infected persons. The appearance of GGOs and a single lesion signifies that the disease was in its primitive phase. The automated approach can perform classification and instance segmentation without manual intervention with a 98.25% accuracy.

One of the current limitations in this field of study is the scarcity of publicly available datasets of COVID-19 CT images. For the segmentation of GGO regions the authors have used several features but those are not adequate for the efficiency of the model. The future work might involve using the weights of a model that is trained on chest CT images. Furthermore, they would also try using ResNet50 as the backbone of the Mask R-CNN instead of ResNet101 to limit the variance of the model. In addition, it is also important to remember that input from clinicians in real-time on the improvement of ROI detection in the scan would further enhance model output at the time of training and inferencing. Lastly, pre-trained models, specific to Chest Infections could be used to limit over-fitting for instance segmentation during pre-processing [23].

### REFERENCES

- [1] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask r-cnn," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 2961–2969.
- [2] C. Hani, N. H. Trieu, I. Saab, S. Dangeard, S. Bennani, G. Chassagnon, and M.-P. Revel, "Covid-19 pneumonia: a review of typical ct findings and differential diagnosis," *Diagnostic and interventional imaging*, 2020.
- [3] Y. Li, H. Qi, J. Dai, X. Ji, and Y. Wei, "Fully convolutional instance-aware semantic segmentation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 2359–2367.
- [4] H. Meng, R. Xiong, R. He, W. Lin, B. Hao, L. Zhang, Z. Lu, X. Shen, T. Fan, W. Jiang *et al.*, "Ct imaging and clinical course of asymptomatic cases with covid-19 pneumonia at admission in wuhan, china," *Journal of Infection*, 2020.
- [5] S. Zhou, Y. Wang, T. Zhu, and L. Xia, "Ct features of coronavirus disease 2019 (covid-19) pneumonia in 62 patients in wuhan, china," *American Journal of Roentgenology*, vol. 214, no. 6, pp. 1287–1294, 2020.
- [6] Y. Wang, C. Dong, Y. Hu, C. Li, Q. Ren, X. Zhang, H. Shi, and M. Zhou, "Temporal changes of ct findings in 90 patients with covid-19 pneumonia: a longitudinal study," *Radiology*, p. 200843, 2020.
- [7] S. Morozov, A. Andreychenko, N. Pavlov, A. Vladzymirskyy, N. Ledikhova, V. Gomboleviskiy, I. A. Blokhin, P. Gelezhe, A. Gonchar, and V. Y. Chernina, "Mosmeddata: Chest ct scans with covid-19 related findings dataset," *arXiv preprint arXiv:2005.06465*, 2020.
- [8] J. Redmon and A. Farhadi, "Yolo9000: better, faster, stronger," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 7263–7271.
- [9] T.-Y. Lin, P. Dollár, R. Girshick, K. He, B. Hariharan, and S. Belongie, "Feature pyramid networks for object detection," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 2117–2125.
- [10] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 3431–3440.
- [11] A. A. Taha and A. Hanbury, "Metrics for evaluating 3d medical image segmentation: analysis, selection, and tool," *BMC medical imaging*, vol. 15, no. 1, p. 29, 2015.
- [12] L. R. Dice, "Measures of the amount of ecologic association between species," *Ecology*, vol. 26, no. 3, pp. 297–302, 1945.
- [13] S. Wang, B. Kang, J. Ma, X. Zeng, M. Xiao, J. Guo, M. Cai, J. Yang, Y. Li, X. Meng *et al.*, "A deep learning algorithm using ct images to screen for corona virus disease (covid-19)," *MedRxiv*, 2020.
- [14] C. Zheng, X. Deng, Q. Fu, Q. Zhou, J. Feng, H. Ma, W. Liu, and X. Wang, "Deep learning-based detection for covid-19 from chest ct using weak label," *medRxiv*, 2020.
- [15] A. Narin, C. Kaya, and Z. Pamuk, "Automatic detection of coronavirus disease (covid-19) using x-ray images and deep convolutional neural networks," *arXiv preprint arXiv:2003.10849*, 2020.
- [16] X. Xu, X. Jiang, C. Ma, P. Du, X. Li, S. Lv, L. Yu, Y. Chen, J. Su, G. Lang *et al.*, "Deep learning system to screen coronavirus disease 2019 pneumonia. arxiv 2020," *arXiv preprint arXiv:2002.09334*.
- [17] 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*, p. 103792, 2020.
- [18] 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.
- [19] L. Wang and A. Wong, "Covid-net: A tailored deep convolutional neural network design for detection of covid-19 cases from chest x-ray images," *arXiv preprint arXiv:2003.09871*, 2020.
- [20] P. K. Sethy and S. K. Behera, "Detection of coronavirus disease (covid-19) based on deep features," *Preprints*, vol. 2020030300, p. 2020, 2020.
- [21] 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*, p. 1, 2020.
- [22] Y. Song, S. Zheng, L. Li, X. Zhang, X. Zhang, Z. Huang, J. Chen, H. Zhao, Y. Jie, R. Wang *et al.*, "Deep learning enables accurate diagnosis of novel coronavirus (covid-19) with ct images," *medRxiv*, 2020.
- [23] S. Reji, E. Earley, and M. Basak, "Brain tumor segmentation."

# Breast Density Classification in Mammogram Images

Vyshnavi V<sup>[1]</sup>

Department of Electronics and  
Communication Engineering,  
Amrita School of Engineering,  
Amrita Vishwa Vidyapeetham,  
Coimbatore, India

Email ID:

cb.en.p2bme19012@cb.students.a  
mrita.edu<sup>[1]</sup>

Devi Vijayan<sup>[2]</sup>

Department of Electronics and  
Communication Engineering,  
Amrita School of Engineering,  
Amrita Vishwa Vidyapeetham,  
Coimbatore, India

Email ID: v\_devi@cb.amrita.edu<sup>[2]</sup>

R. Lavanya<sup>[3]</sup>

Department of Electronics and  
Communication Engineering,  
Amrita School of Engineering,  
Amrita Vishwa Vidyapeetham,  
Coimbatore, India

Email ID:

r\_lavanya@cb.amrita.edu<sup>[3]</sup>

**Abstract** - Breast cancer is the common type of cancer in the world, which is most common among women. It is found that there is a correlation between the breast cancer and breast density, hence there is a need for developing a method that identifies the breast density so as to reduce morbidity and mortality. The aim of this work is to develop a Computer Aided Design (CAD) system that classifies the breast density according to the Breast Imaging-Reporting and Data System (BI-RADS) standard with the help of the digital mammographic images. To this end, we propose an effective feature descriptor, namely Locally Encoded Transform feature histogram (LETRIST) for capturing the essential characteristics that discriminates across the density categories. Proposed descriptor evaluated using Support Vector Machine (SVM) on the Mammographic Image Analysis Society (MIAS) database demonstrates the efficiency of proposed system.

## I. INTRODUCTION

Breast cancer is a type of cancer in which the cells present in the breast tissue behaves abnormally and starts to multiply rapidly which is uncontrollable, and this result in the formation of lumps or mass. In 2019, 268, 600 invasive breast cancers were diagnosed among women, approximately 41, 760 women were dead from breast cancer [1]. The other reason for breast cancer is breast density, women with high breast density likely tends to have breast cancer any sooner or later. Women with high breast density need to be diagnosed as early as possible, so that they can get some kind of treatment before it becomes cancerous. It is necessary to find a solution for this, so that there will be reduced morbidity and mortality. Designing a computer aided design (CAD) system for the detection of breast cancer is the best option to assist radiologist and physicians to diagnose it at an early stage [19]. Breast density is categorized into four classes according to the Breast Imaging-Reporting and Data System (BI-RADS). The four classes of BI-RADS are shown is Fig 1. Classification of breast density based on BI-RADS standard is as follows:

- BI-RADS I: Almost entirely fatty breast (0–25%).
- BI-RADS II: Some fibro-glandular tissue (26–50%).
- BI-RADS III: Heterogeneously dense breast (51–75%).
- BI-RADS IV: Extremely dense breast (76–100%).

Visual judgment of mammograms to estimate breast density is a crucial task, so CAD systems developed can aid the radiologists for detecting the breast cancer.

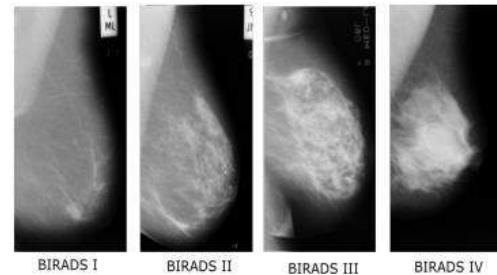


Fig 1: BI-RADS Classes

The proposed methodology includes pre-processing, segmentation, feature extraction, dimensionality reduction and classification of digital mammogram images. Pre-processing of the digital mammogram images are done at two stages, Stage 1: removing the noises encountered during acquisition and removal of artifacts such as labels and tags; Stage 2: contrast enhancement. Segmentation is an important step in mammogram image processing; this is because presence of pectoral muscles affects the classification performance of the classifier due to its intensity which is almost same as that of the denser area or the mass so, the pectoral muscle is removed before proceeding with the next step. Breast region is extracted after the removal of the pectoral muscle, which will be used for feature extraction process. The major contribution of the work is in analysing the variation across density categories using effective feature descriptor namely Locally Encoded Transform feature histogram (LETRIST). To exploit the correlation

between the features to extract the reduced dimensional discriminative feature vector, principal component analysis (PCA) was employed. The remaining sections cover the literature survey, methodology, experimental results and the conclusion of the proposed project work.

## II. LITERATURE SURVEY

Recent literatures related to mammograms pre-processing, segmentation, feature extraction and classification inspired in designing a computer aided design (CAD) system for the breast density classification using mammogram images. Chen et al. [2] has used fuzzy C algorithm for finding the region of interest (ROI), pre-processing includes median filtering. Extracted the statistical features; the extracted features are given to the KNN classifier and decision tree classifier. Vaidehi et al. [3] has done the artifact and pectoral muscle removal. The Haralick texture features are extracted from the interleaved 9x9 block using candidate block splitting. The extracted features are classified into dense, glandular, fatty using the support vector machine (SVM). Jen et al. [4] has employed global equalization transformation to stretch the image intensity to full range of 0 – 255. The uniform and the gaussian type noises are removed using the mean filter.

The breast density is classified according to the BI-RADS category using Uniform Local Directional Pattern (ULDLP) texture descriptor by Moreno et al. [5]. Both linear SVM (LSVM) and non-linear SVM (NLSVM) are used for decision making. Wener et al. [6] has performed reduction of the original image size so as to reduce the computation time. To identify the dense and non-dense breasts variance filter was used. Mean filtering and Contrast Limited Adaptive Histogram Equalization (CLAHE) was performed to remove noise and enhance the contrast of the image. Alhelou et al. [7] has used IRMA dataset as input and extracted a bag of features such as Local Binary Pattern (LBP), Gray Level Co-occurrence Matrix (GLCM), Gray Level Run Length (GLRL) and Wavelet Transform (WT). The Support Vector Machine (SVM) using Sequential Minimal Optimization (SMO) has outperformed and obtained an accuracy of 59%. CAD system was designed by Vidivelli et al. [8] which removes the noises, labels and artifacts automatically by morphological operations. Rampun et al. [9] and the other authors has attempted using a new variant of Local Binary Pattern (LBP) known as the Local Ternary Pattern (LTP). Extracted the LTP texture feature from the Fibro-Glandular Disk (FGD) region rather than extracting it from the entire breast region. When these features were fed into the SVM classifier, a ten-fold cross validation was performed and achieved 82.33% accuracy considering only the FGD region.

Breast density classification using LQP with various neighbourhood topologies was performed by Rampun et al. [10]. The input image was pre-processed and the LQP texture feature was extracted from the FGD region of the breast. A multi-resolution and multi-topology approach along with Local Septenary Pattern (LSP) was carried out by Rampun et al. [11]. Used dimensionality reduction technique to select the dominant

pattern to reduce the overlapping texture information. For both MIAS and InBreast database, the Multi-LSP ellipse topology achieved the highest classification accuracy of 73.6%. Haipeng Li et al. [12] has used multi-fractal spectrum and histogram analysis to classify the digital mammograms. The breast region segmentation was performed after the pre-processing and the sub-regions were obtained from it so as to capture the fibro-glandular texture features. The feature vectors are extracted from the multi-fractal spectrum to calculate the breast density and to classify it into fatty/dense category. George et al. [13] employed contour growing, extracted Elliptical Local Binary Pattern (ELBP) descriptors and Local Directional Patterns (LDP) from the ROI's of FGD & whole breast region to compare the classification results.

The ROI's obtained from the FGD region attained highest accuracy than the whole breast region for MIAS database. Rampun et al. [14] removed the salt and pepper noise using median filter. The breast boundary was estimated using the region based active contour and the restricted contour growing was incorporated to obtain the edge information for pectoral muscle boundary estimation. The obtained dominant patterns were fed into SVM classifier for classification. Tiecheng Song et al. [15] proposed a new texture descriptor called LETRIST and used it to classify texture images of different texture datasets. The main advantage of using this feature is it is rotation invariant, insensitive to noise and yet discriminant. The LETRIST recognition rate for Outex and KTH-TIPS datasets are 100% and 99% respectively. John et al. [16] proposed LETRIST for face recognition and achieved an accuracy of 99%. Compared with LBP and LTP, LETRIST shows very strong anti-noise ability.

## III. METHODOLOGY

The proposed methodology consists of the pre-processing, segmentation, feature extraction, dimensionality reduction and classification. Fig. 2 shows the work flow of the proposed methodology.

### A. Database:

The proposed methodology is evaluated using the publically available dataset, Mammographic Image Analysis Society (MIAS). The MIAS dataset consist of 322 digital mammogram images which are labelled into three classes (fatty, glandular and dense) according to BI-RADS standard.

### B. Pre-Processing:

The pre-processing is the basic and very important step in image processing [21]. The images are pre-processed so that if any kind of noise present in it can be removed to view the image clearly [22]. Contrast enhancement is also performed to adjust the contrast of the image so that even the precise details can be observed. The artifacts are also removed in the pre-processing stage which might produce less classification accuracy.

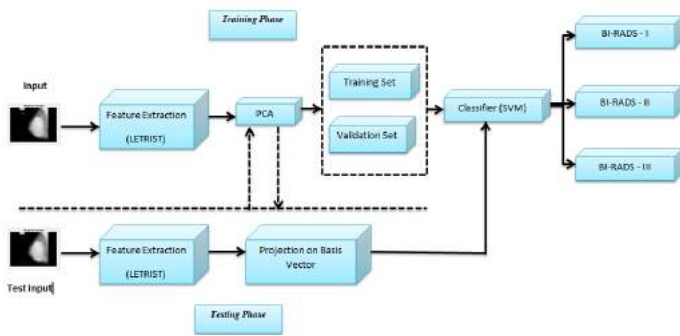


Fig 2: Proposed Methodology

### Median Filtering:

Median filtering, a non-linear filtering technique [3] is used for removal of impulse noise present in the image, it resembles sprinkled black and white dots that preserves the edge information. The algorithm of this technique as follows: A window is present say, 3x3 (default) sized window, and this window slides throughout the image pixel by pixel. The pixel values are arranged in ascending order and then each centre value of the window is replaced with the median value (middle value) of the neighbourhood pixels.

### Contrast Limited Adaptive Histogram Equalization (CLAHE):

Contrast Limited Adaptive Histogram Equalization (CLAHE) is an image processing technique used to improve the contrast of the image thereby reducing the noise amplification [12]. This calculates the contrast transfer function for each and every tile present in an image. And finally, to eliminate the boundaries induced artificially bilinear interpolation is used. In this project CLAHE is used to enhance the contrast of the gray scale image.

#### C. Segmentation:

Image segmentation is performed to segment the region of interest (ROI) or object of interest [22]. This reduces the computation time because further image processing is concentrated only on segmented ROI of the image.

### Pectoral Muscle Removal:

Pectoral muscle is visible in the Medio-Lateral Oblique (MLO) view of the mammographic images. It is present in as triangular shaped region at one side of both the left and right mammogram images. Removal of pectoral muscle is very important because, the intensity of pectoral muscle is as bright as the intensity of the denser region of the breast. And this might lead to misclassification of breast density. Hence, in this proposed methodology the pectoral muscle is removed by gray threshold and morphological operations [4]. The gray threshold ( $t$ ) is performed to differentiate the foreground from the background region. Based on the threshold value the pixel values greater than the thresholds ( $t$ ) are assigned as 1 and the pixels values below thresholds ( $t$ ) are assigned as 0 followed by morphological operation.

### Breast Region Segmentation:

The breast region consists of the fatty or glandular or dense region, fibro-glandular disk region (FGD) and the nipple area. Segmentation of the breast region plays a major role in identification of the breast density. The breast region is segmented with the help of boundary detection technique [5] after the detection of the breast boundary the breast region will be extracted leaving behind the background region.

#### D. Feature Extraction:

### Local Binary Pattern (LBP):

Local binary pattern (LBP) is a texture feature which gives the texture information of the input image. LBP is a local feature and it extracts the texture features from the local neighbourhood of the centre pixel. Let us consider the default LBP feature extraction algorithm in which the radius,  $R$  is 1 and the number of neighbourhood,  $n$  is 8. In this case there will be a 3x3 matrix in which the centre pixel value must be considered as the threshold value, depending upon that threshold value ( $t$ ), the neighbourhood pixel values will be assigned as 1 and 0. The pixel value greater than the  $t$  will be 1 and the pixel value less than  $t$  will be assigned as 0. Now, the LBP for each pixel in an image is calculated and the LBP histogram or LBP feature vector will be constructed.

### Locally Encoded Transform feature histogram (LETRIST):

Locally encoded transform feature histogram (LETRIST) explicitly encodes the combined information within an image across feature and scale space. The following steps are involved in obtaining the LETRIST texture feature extraction:

1. *Extremum Filtering*: The extremum filtering is performed to get the useful information and to extract the rotation invariant local features.
2. *Feature Transformation*: The features obtained from the extremum filtering are transformed with the help of the linear and non-linear operators.
3. *Scalar Quantization*: Scalar quantization is performed to quantize the transformed features into discrete texture codes.
4. *Cross-Scale Joint Coding and Image Representation*: The final histogram is obtained by aggregating all the texture codes.

The LETRIST texture descriptor at the final image representation stage consists of 413 dimensional image descriptors. The main advantage of this LETRIST texture feature is that it is discriminative, insensitive to noise and computationally efficient.

#### E. Dimensionality Reduction:

Dimensionality reduction technique is used for the higher dimension feature vector for reducing the dimension space into low dimension vector which retains all the important properties

and the information of the higher dimension input feature vector [18]. Principal component analysis (PCA) is a technique applied to larger dataset to get the strong features and eliminate the variations. In this way the dimension of the larger input dataset will be reduced, which in turn minimizes the classification computation time. The LETRIST texture feature is dimensionally reduced from 413 feature vectors to 50 feature vectors, which in turn reduces the computation time.

#### F. Classification:

Classification in image processing is done after analysing various numerical properties of the input images extracted features and the categorizing it into different classes [17]. The classification algorithm has got two processing steps called training and testing. Support Vector Machine (SVM) was employed for classification.

## IV. RESULTS AND DISCUSSIONS

The proposed methodology is validated on publically available database such as MIAS. In this experiment 208 normal mammogram image of MIAS dataset is used in which 76 images – dense, 65 images – glandular and 66 images – fatty for three class classification, for two class classification the glandular images are labelled as dense images, wherein there will be 66 images for fatty and 144 dense images respectively. The input mammogram image is shown in Fig 3(a). The images are pre-processed by removing the labels followed by the removal of the unwanted noises that are present in the image which is performed by the 5x5 median filters (Fig 3(b)). Here, CLAHE is used for contrast enhancement of the image and the result is shown in Fig 3(c). In the proposed methodology the pectoral muscle is removed by gray thresholding and morphological operations. The result of artifact removal (label/tag) is shown in Fig 4(a). The mammogram images after the removal of pectoral muscle is shown in Fig 4(b). The breast region is extracted by the region growing method which is shown in Fig 4(c). The texture features including LBP and LETRIST are extracted from the breast area. The LBP texture feature consists of 59 feature vectors for each image and the LETRIST consists of 413 feature vectors. To reduce the dimension of the obtained feature vector, PCA dimensionality reduction algorithm is performed. The LETRIST 413 feature vectors are reduced to 50 feature vectors for each and every sample used in this proposed work. Finally, the extracted texture features are fed in to support vector machine (SVM). Parameter optimization and performance assessment was performed using 10 fold cross validation. Grid search was employed for parameter tuning to avoid over fitting.

The combination of LETRIST and SVM-polynomial kernel classifier achieved an accuracy of 98.6% for two class classification and 78.5% for three class classification. For the same dataset, LBP obtained an accuracy of 97.6% for two class classification and 71.4% for three class classification. Experimental results shows, the proposed descriptor outperforms LBP by 7.2% for three class classification. As the

number of classes increases, the variation within the data increases and the proposed descriptor effectively captures those variations for three class problem. Table 1 and 2 shows the comparison of the proposed methodology results with the state-of-the-art techniques for two classes and three class classification of MIAS dataset. The confusion matrix (CM) for LBP is shown in Table 3 & 4 and confusion matrix for LETRIST 2 class & 3 classes is shown in Table 5 & 6 respectively.

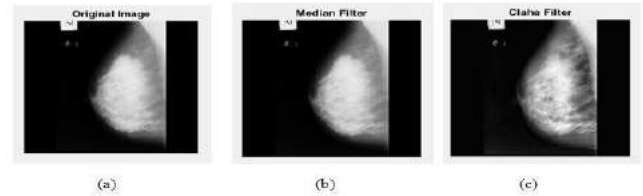


Fig 3: Pre-processing Result (a) original image, (b) median filter, (c) CLAHE filter

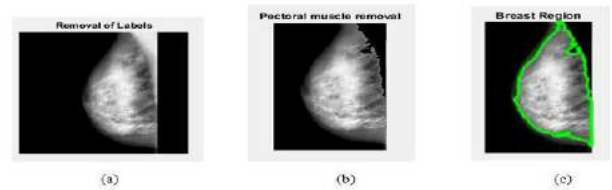


Fig 4: Segmentation Result (a) Label removal (b) Pectoral muscle removal (c) Breast region extraction

TABLE 1: COMPARISON RESULTS -TWO CLASS CLASSIFICATION

Features	Classifier	Accuracy	Reference
Local Ternary Pattern (LTP)	SVM	82.33%	Rampun et al. [9]
Local Septenary Pattern (LSP)	SVM	86.9%	Rampun et al. [11]
Multi-fractal spectrum	Chi-square statistic classifier	90.1%	Li et al. [12]
<b>Locally encoded transform feature histogram (LETRIST)</b>	<b>SVM</b>	<b>98.6%</b>	<b>Proposed work</b>

TABLE 2: COMPARISON RESULTS -THREE CLASS CLASSIFICATION

Features	Classifier	Accuracy	Reference
Statistical features	SVM-SMO	59.12%	Alhelou et al. [7]
Multi-fractal spectrum	Chi-square statistic classifier	75.53%	Li et al. [12]
Local Septenary Pattern (LSP)	SVM	72.1%	Rampun et al. [11]
<b>Locally encoded transform feature histogram (LETRIST)</b>	<b>SVM</b>	<b>78.5%</b>	<b>Proposed work</b>

TABLE 3: CM OF LBP FOR TWO CLASSES

ACTUAL CLASS \ PREDICTED CLASS	DENSE	FATTY
	DENSE	27
FATTY	1	14

TABLE 4: CM OF LBP FOR THREE CLASSES

PREDICTED CLASS \ ACTUAL CLASS	DENSE	GLANDULAR	FATTY
DENSE	12	2	0
GLANDULAR	0	13	1
FATTY	8	1	5

TABLE 5: CM OF LETRIST FOR TWO CLASSES

PREDICTED CLASS \ ACTUAL CLASS	DENSE	FATTY
DENSE	50	0
FATTY	1	22

TABLE 6: CM OF LETRIST FOR THREE CLASSES

PREDICTED CLASS \ ACTUAL CLASS	DENSE	GLANDULAR	FATTY
DENSE	12	2	2
GLANDULAR	0	9	2
FATTY	1	2	12

## V. CONCLUSION

Breast cancer is impacting women all over the world. It is necessary to diagnose it at an early stage [20]. In this, breast density plays as a major risk indicator for breast cancer. The proposed work addresses the development of Computer Aided Design (CAD) system for density classification. Proposed locally encoded transform feature histogram (LETRIST) feature vector effectively captures the variation within the data, obtaining a maximum accuracy of 78.5% for three class and 98.6% for two classes. The performance can be further improved by extracting fibro-glandular disk region instead of the whole breast area and employing various feature fusion techniques.

## REFERENCES

- [1] American Cancer Society. Breast Cancer Facts & Figures 2019-2020. Atlanta: American Cancer Society, Inc. 2019.
- [2] Z. Chen, E. Denton and R. Zwiggelaar, "Local feature based mammographic tissue pattern modelling and breast density classification," 2011 4th International Conference on Biomedical Engineering and Informatics (BMEI), Shanghai, 2011, pp. 351-355.
- [3] K. Vaidehi and T. S. Subashini, "Automatic classification and retrieval of mammographic tissue density using texture features," 2015 IEEE 9th International Conference on Intelligent Systems and Control (ISCO), Coimbatore, 2015, pp. 1-6.
- [4] Chun-Chu Jen, Shyr-Shen Yu, Automatic detection of abnormal mammograms in mammographic images, Expert Systems with Applications, Volume 42, Issue 6, 2015, Pages 3048-3055, ISSN 0957-4174.
- [5] Mohamed Abdel-Nasser, Hatem A. Rashwan, Domenec Puig, Antonio Moreno, Analysis of tissue abnormality and breast density in mammographic images using a uniform local directional pattern, Expert Systems with Applications, Volume 42, Issue 24, 2015, Pages 9499-9511, ISSN 0957-4174.
- [6] Wener Borges de Sampaio, Aristofanes Corr ́ea Silva, Anselmo Cardoso de Paiva, Marcelo ˆ Gattass, Detection of masses in mammograms with

adaptation to breast density using genetic algorithm, phylogenetic trees, LBP and SVM, Expert Systems with Applications, Volume 42, Issue 22, 2015, Pages 8911-8928, ISSN 0957-4174.

- [7] M. Alhelou, M. Deriche and L. Ghouti, "Breast Density Classification Using a Bag of Features and an SVM Classifier," 2017 9th IEEE-GCC Conference and Exhibition (GCCCE), Manama, 2017, pp. 1-9.
- [8] S. S. Devi and S. Vidivelli, "Classification of breast tissue density in digital mammograms," 2017 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS), Coimbatore, 2017, pp. 1-7.
- [9] Rampun A., Morrow P., Scotney B., Winder J. (2017) Breast Density Classification Using Local Ternary Patterns in Mammograms. In: Karray F., Campilho A., Cheriet F. (eds) Image Analysis and Recognition. ICIAR 2017. Lecture Notes in Computer Science, vol 10317, Springer.
- [10] Rampun, A.; Scotney, B.W.; Morrow, P.J.; Wang, H.; Winder, J. Breast Density Classification Using Local Quinary Patterns with Various Neighbourhood Topologies. J. Imaging 2018, pp. 4-14.
- [11] Rampun, B. W. Scotney, P. J. Morrow and H. Wang, "Breast Density Classification using Local Septenary Patterns: A Multi-resolution and Multi-topology Approach," 2019 IEEE 32nd International Symposium on Computer-Based Medical Systems (CBMS), Cordoba, Spain, 2019, pp. 646-651.
- [12] H. Li, R. Mukundan and S. Boyd, "Breast Density Classification Using Multifractal Spectrum with Histogram Analysis," 2019 International Conference on Image and Vision Computing New Zealand (IVCNZ), Dunedin, New Zealand, 2019, pp. 1-6.
- [13] George, M.; Zwiggelaar, R. Comparative Study on Local Binary Patterns for Mammographic Density and Risk Scoring. J. Imaging 2019, pp. 5-24.
- [14] Andrik Rampun, Philip J. Morrow, Bryan W. Scotney, Hui Wang, Breast density classification in mammograms: An investigation of encoding techniques in binary-based local patterns, Computers in Biology and Medicine, Volume 122, 2020, 103842, ISSN 0010-4825.
- [15] T. Song, H. Li, F. Meng, Q. Wu and J. Cai, "LETRIST: Locally Encoded Transform Feature Histogram for Rotation-Invariant Texture Classification," in IEEE Transactions on Circuits and Systems for Video Technology, vol. 28, no. 7, pp. 1565-1579, July 2018.
- [16] Soldara, John and Dodson, C.T.J. and Scharcanski, Jacob, Face recognition based on texture information and geodesic distance approximations between multivariate normal distributions, Measurement Science and Technology, Volume 29, Issue 11, article id. 114001 (2018).
- [17] K. Kiruthika, Devi Vijayan and Lavanya. R, "Retrieval Driven Classification for Mammographic Masses, International Conference on Communication and Signal Processing, April 4-6, 2019.
- [18] Reshma.R, V.Sowmya, K.P.Soman "Dimensionality Reduction using Band Selection Technique for Kernel based Hyperspectral Image Classification", 6th International Conference On Advances In Computing & Communications, ICACC 2016, pp.396 – 402, September 2016.
- [19] Karthika, R., and Latha Parameswaran. "An automated vision-based algorithm for out of context detection in images." *International Journal of Signal and Imaging Systems Engineering* 11.1 (2018): 1-8.
- [20] Dharani, V., and R. Lavanya. "Improved microaneurysm detection in fundus images for diagnosis of diabetic retinopathy." *International Symposium on Signal Processing and Intelligent Recognition Systems*. Springer, Cham, 2017.
- [21] Anupama, M. A., V. Sowmya, and K. P. Soman. "Breast cancer classification using capsule network with preprocessed histology images." *2019 International Conference on Communication and Signal Processing (ICCSP)*. IEEE, 2019.
- [22] Ram, Parvathy, and T. R. Swapna. "A Study on Various Quantification Algorithms for Diabetic Retinopathy and Diabetic Maculopathy Grading." *Computational Vision and Bio Inspired Computing*. Springer, Cham, 2018. 412-421.

# Systematic Study on Diagnosis of Lung Disorders using Machine Learning and Deep Learning Algorithms

R.Swathi Sri  
*Master's in School of Computer  
 Science and Engineering  
 VIT Chennai Campus  
 Chennai, India  
 swathisri.r2020@vitstudent.ac.in*

Dr. A. Menaka Pushpa  
*Assistant Professor, School of Computer  
 Science and Engineering  
 VIT Chennai Campus  
 Chennai, India  
 menakapushpa.a@vit.ac.in*

**Abstract**—Presently, lung infection is severe to humans that leads to death if left untreated, and Tracking down a disease on the dot is a way we get out of a hock. Recurrently, the eleventh-hour tracking of bugs or incertitude in prognosis margins to voluminous hefty blow-offs. Deep learning branches out a scheme to cut and run of it. Deep learning models work on medical images to detect the type of lung disease. The fashioning of a gadget aids the medical technicians to put a finger on the type and also sub-type of the syndrome without chaos by postulating the prototype as input to it, to enforce the primitive and definite therapy. Grail is to define a common system that categorizes all possible lung diseases like cancer, TB, Pneumonia, and COVID from CT or CXR images.

**Keywords**—Deep learning, CT, CXR, lung disease.

## I. INTRODUCTION

There are several sorts of diseases comprehending with the lungs such as tuberculosis, pneumonia, Chronic obstructive pulmonary disease (COPD), covid-19, and so on. Consistently, such cases are misdiagnosed. Despite the methodologies like bronchoscopy would help in diagnosing the syndrome, it would not guide a certain set of victims accounting to their health conditions. When such ailments are misdiagnosed, it would breed very severe outcomes that allot demerits for ages. Uncertainties would surely provoke mishaps. Hence, there's a primary requisite for a paradigm that emends the cited issues in diagnosing.

Heretofore, the CAD tool is used for ferreting out disease. It takes input as CT, CXR images, and tries to discern a nodule. It also undergoes three stages (1). It classifies the kind of disease from presumed types i.e., it pin-downs TB, Pneumonia images under their respective classes. (2). It confines the clustering of nodules in the images in the detecting phase. It turns-out to produce images that have bounded the diseased region. (3). Each pixel of images is split into their classes i.e., the pixels of lesion and pixels of vigor.

In the recent past, AI also works on images and discerns the class of disease. Scientists have created a Neural Network based on the principles of AI that predicts in good accuracy. In this survey paper, we will discuss different Machine Learning (ML) and Deep Learning (DL) models the collaborate themselves for the identification of lung disease. Adverse to it, the model is necessitated to get equipped with diverse data sets to brace up the certainty of it. Incongruously, data augmentation can be put into effect to accustom the model. components, incorporating the applicable criteria that follow. DL has been transpired as one of the breakthrough technologies of the recent year and

evolving in ML and computer vision. DL is a revival of the convolutional neural network (CNN) with the ability to learn and make intelligent decisions, which is key to apply deep learning networks in medical images.

## II. METHODS

### 1) Literature search strategy:

The search of the overall survey was done in IEEE Xplore by using the following keywords: “lung disease”, “TB”, “Covid-19”, “Pneumonia” with input type as medical images as well as segmented data. Added to this, we particularly concentrated on a research article that worked on DL and ML algorithms to classify lung disease.

### 2) Study selection and eligibility criteria:

Studies that met the following requirements were included:

- Reach papers that are new.
- Medical images or image texture as input.
- Enough information that talks about accuracy as the performance metric.
- Research papers for lung disease using a DL and ML algorithm.

Lung cancer was left out because most recent articles just focused on it briefly.

### 3) Data extraction:

On each deserving work, we independently extricated its following data: the algorithm suited, prophesy, dataset source and scale, augmentation, and image type. We've focused on how segmentation techniques work and how approaches are used to achieve accuracy. Without segmentation, the result shown in the recent past has a high level of accuracy.

## III. TECHNIQUES IN DEEP LEARNING AND MACHINE LEARNING

### A. Convolutional Neural Network (CNN):

CNN is the prevailing classification algorithm (both binary and multi-class) that has become a prior model in various domains like computer vision, radiology. CNN model can be built on supervised and unsupervised techniques. Some of the layers are convolution, pooling, and fully connected which are called building blocks of CNN [10,19]. A variant of CNN is DCNN. It is the same as CNN but after the last max-pooling layer, the dropout function is used to prevent overfitting of the model. It is a classification



algorithm (multiclass) and the accuracy of DCNN tumbled down in uncertain cases.

### B. Bayesian Optimization Algorithm (BOA):

A Bayesian optimization algorithm is an optimization function applied on CNN to return optimized classification model hyperparameters. BOA's goal is to identify validation errors set by using the Bayesian formula during model

training. The function is correlated after the SoftMax layer. B-CNN (CNN with Bayesian Optimization Algorithm) validates results produced by the SoftMax layer by filtering confusion cases on calculating variance to increase accuracy. BOA works on the Bayesian theorem (i.e., calculates the probability of ongoing event) and tedious [14,24].

TABLE I. DATA SET DESCRIPTION OF EXISTING WORKS

Prognosticate	Year	Reference	Input type	Image Type	Segmentation	Dataset	Source of DataSet
Covid19	2020	[7]	Image	Chest 3D Image	Lobe segmentation - 3D-UNet	4657 (936-Normal, 2406 - ILD, 1315 - Covid19)	Own collected dataset from different hospitals
	2020	[11]	Image	X-Ray, Ultrasound, CT Scan	No	60,823 - X-Ray, 746 - CT Scan, Ultrasound - 911	National Institute of Health (NIH) Chest X-Ray dataset, COVID-CT Dataset, POCOVID-Net dataset, X-Ray COVID-19 dataset
	2020	[5]	Image	Posterior-Anterior (PA) and Anterior-Posterior (AP) XR Chest X-Ray	Yes (U-Net semantic segmentation algorithm)	79500	HM hospitals covid19, BIMCV covid19, Actualmed Set (ACT), China Set-The Shenzhen Set, The Montgomery Set, ChestX-ray8, CheXpert, MIMIC-CXR
	2020	[19]	Image	Posteroanterior (PA) Chest X-Ray	No	33,231	CheXpert, ChestX-ray8, GitHub for COVID-19 X-rays
	2020	[15]	Image	Posteroanterior (PA) Chest X-Ray	No	5000	Own collected dataset from publicly available dataset
	2020	[2]	Image	Posterior-anterior (PA) and anterior-posterior (AP) Chest X-Ray	No	1394	Kaggle, NIH Chest X-rays, J. P. Cohen, P. Morrison, and L. Dao, "Covid-19 image data collection,"
Pneumonia	2021	[8]	Image	Chest X-Ray	No	Viral pneumonia - 5,977, Non-viral pneumonia - 18,619, Healthy controls - 18,774	X-VIRAL and X-COVID
	2020	[10]	Image	Posterior-to-anterior (AP)/anterior-to-posterior (PA) image of chest X-ray	No	423-COVID-19, 1485-Viral Pneumonia, and 1579 - Normal	Kaggle databases, Italian Society of Medical and Interventional Radiology (SIRM) COVID-19 DATABASE, Novel Corona Virus 2019 Dataset, COVID-19 positive chest x-ray images from different Articles, COVID-19 Chest imaging at thread reader
	2020	[12]	Images, Video, clinical and demographic information	Lung Window Images, High Attenuation Images, Low Attenuation Images and chest CT scans Video	Text segmentation in Chinese words	Normal - 450 and pneumonia - 450	Private dataset from the Radiology Department of The First Affiliated Hospital of Army Medical University and hospital PACS (Picture Archiving and Communication Systems)
	2020	[23]	Image	Chest CT	Hand-crafted lung segmentation algorithm	Healthy: 73282 COVID-19: 32301, H1N1: 1981, CAP: 11172	PACS, Shanghai Public Health Clinical Center, Fudan University medical history
	2020	[22]	Image	CT	Pneumonia infection Mask (upper right) and the lung MASK (bottom right) using VB-Net toolkit	4,982	Tongji Hospital of Huazhong University of Science and Technology, Shanghai Public Health Clinical Center of Fudan University, Xiangya Hospital of Central South University, the Hospital of Jilin University, Ruijin Hospital of Shanghai Jiao Tong University School of Medicine, Hangzhou First People's Hospital of Zhejiang University, the Beijing Chaoyang

							Hospital of Capital Medical University, and Sichuan University West China Hospital.
	2020	[1]	Image	CT	No	5856	University of California, San Diego, Kermany et al., 2018
	2019	[4]	Image	CT	No	2000 (lobar pneumonia, bronchial pneumonia, interstitial pneumonia and bronchiolitis)	Own Data set collected from Hospitals
TB	2020	[18]	Image	Chest X-Ray	No	Normal-1,583, bacterial and viral pneumonia - 4,273	RSNA, Pediatric pneumonia, Indiana dataset, Shenzhen CXR collection
	2020	[16]	Data	TB Strain and genome sequencing	No	144 sputum	Through Belarus TB Portal <a href="http://tuberculosis.by">http://tuberculosis.by</a> . The Republican Scientific and Practical Center of Pulmonology and Tuberculosis of Ministry of Health of Belarus
	2013	[17]	Image	Conventional posteroanterior chest radiographs	Graph cut segmentation method	Total: 753, Normal: 420, Abnormal: 333	Health department in the United States, and a set collected by Shenzhen Hospital, China
	2015	[6]	Image	Posterior anterior chest radiograph (CXR)	Lung segmentation, texture feature extraction and pixel classification	917	Urban health center in Lusaka, Zambia.
	2020	[24]	Image	Chest X-Ray	No	800 images are taken in which 394 are TB positive and rest are normal.	Montgomery and Shenzhen
	2020	[20]	Image	Chest X-Ray	U-net models	3500 TB infected and 3500 normal	ChexNet, DenseNet201
COPD	2017	[9]	Image	Posteroanterior chest radiographs (CXRs)	ROI+Lung region symmetry Computation	Normal: 498, Abnormal: 478	U.S. National Library of Medicine, and National Institute of Tuberculosis and Respiratory Diseases, India
	2020	[14]	Image	3D lung airway tree	Deep airway segmentation embedded in n Mimics	COPD:190, HC: 90	Central Hospital Afiliated to Shenyang Medical College
	2017	[21]	Image	CT Scan to feature vector	Lung Mask extracted by region-growing algorithm	8 x4 x 10 = 320 features	Danish Lung Cancer Screening Trial, COPD Gene study, National Jewish Center in Denver, Colorado, Frederikshavn hospital in Denmark
	2020	[13]	Text	Demographic information, Electronic medical record information, examination results, health self-scores, and follow-up information	-	Total:1999, COPD - 829, non-COPD - 1021 and undiagnosed -149	Clinical Medical Science Data Center

### C. Generative Adversarial Networks (GAN):

The lung image segmentation is performed by generative adversarial networks on the chest X-Ray image. The generative model is trained by GANs by addressing the problem in a supervised fashion. It works on two-phase, generator model and discriminator model. The generator model(U-net) produces a binary image (generated image) as output for the respective X-Ray image. The discriminator model classifies whether it is a mock or not based on ground truth. To restore the gradient weight of the network and maintain morphology, skip connections of U-Net are used. High computing power and cost are considered as an impediment of GAN [3,19].

### D. Support Vector Machine (SVM):

SVM is a widely used classification algorithm in machine learning (supervised algorithm) due to its effectiveness, reliability, and ability to handle high dimensional feature space, where numbers of features are considered to be very small. It can resolve both classification and regression problems. For SVM, each data object is plotted as a point for n-dimensional space with the value of each variable being a common coordinate value. Then perform classification by locating the hyperplane which very well distinguishes the two groups[4,6,17].

### E. K-Means Clustering(Knn):

K-Means Clustering is a kind of Unsupervised Machine Learning Algorithm. Simply put, it takes datapoint as input and returns k no of cluster i.e., clustering means clumping data based on the rife feature. The concept behind K-Means is that we want to assign k new points to the data received, known as a Centroid – is about to try to concentrate on the middle of one of the clusters. Once that ends, our algorithm stops clustering [2].

### F. Adaboost Algorithm:

A multitude of weak classifiers is combined with the AdaBoost algorithm to carry out the strong classifier. This algorithm is iterative. Training the different weak classifiers with the same training set and combining them pursuant to the weight ratio to cast a strong classifier is the fundamental hypothesis [4].

### G. GDCNN:

Deep convolutional neural network in short DCNN is a classification algorithm. Here experiment is performed on a genetic-based DCNN model which was developed based on gene map. It is necessary to observe DCNN operations(encoding) are based on gene schematic and trained model is close to 100 epochs so GDCNN has high

Zain Ul Abideen et al, predict tuberculosis by implementing B-CNN. The images from Montgomery and Shenzhen are taken and fed to the CNN model [24]. The performance of CNN fails when uncertain cases occur. To conquer this, B-CNN is revived. Results show the accuracy of B-CNN is higher than other deep learning models.

From table I and II we discuss the feature that characterize the model (i) Image type, (ii) Segmentation, (iii) Augmentation, (iv) Technique used, (v) Epoch, (vi) Learning rate, (vii) Performance.

For the 1<sup>st</sup> feature (image type) most of the survey papers have used 2D CXR or CT scan where the majority of CXRs is in the PA view and suited most diseases. R. Du et al., [14] proposed B-CNN would wear a 3D lung airway tree as input. 3D views of 2d images are taken and modeled. This means that top, front and side views are taken into consideration. This helps the model to overcome data insufficiency and gives different views of the nodule. On the other hand, V. Cheplygina et al., [21] used feature vector and shape vector as input for the model. These are the results of segmentation.

For 2<sup>nd</sup> feature(segmentation), instead of giving the entire image as input, we do segmentation as the intermediate step of pre-processing which results in vector-like texture, shape, feature. These results are target attribute given for the model

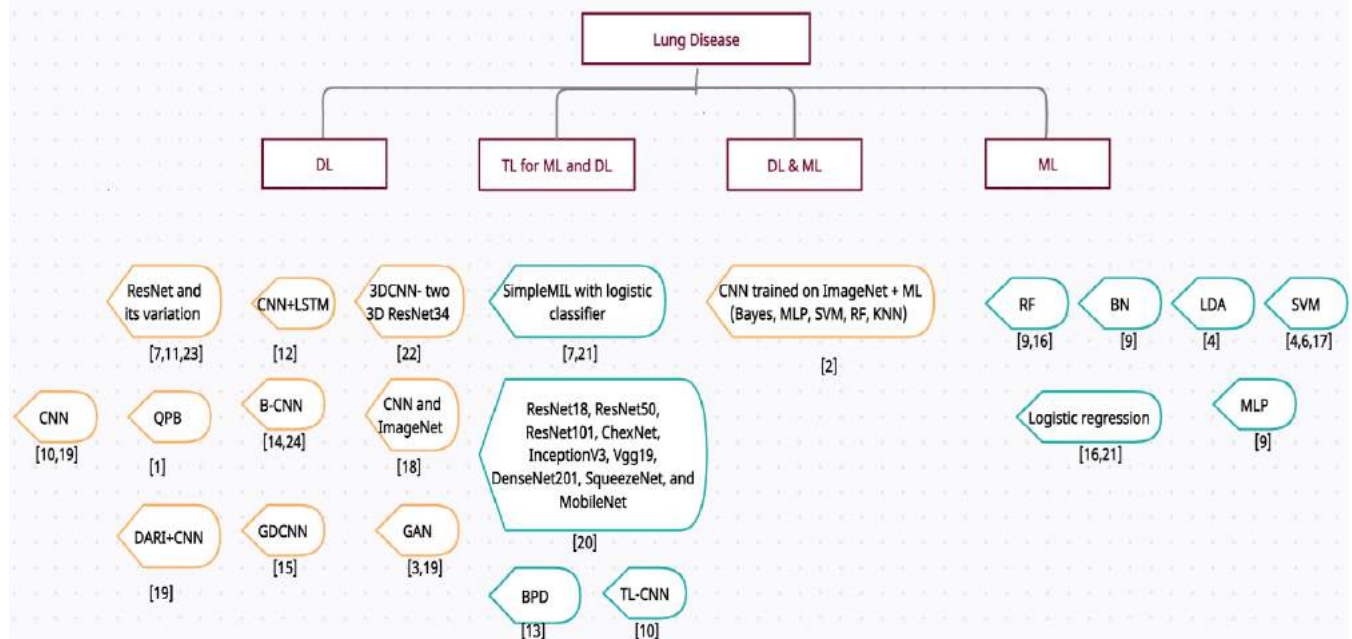


Fig. 1. Literature Survey Diagram.

computational and spatial complexity [15].

## IV. DISCUSSION

Faizan Munawar et al., presents GAN [2]. The lung CXR image segmentation is performed by generative adversarial networks on the chest X-Ray image. The generative model is trained by GANs by addressing the problem in a supervised fashion. It works on two-phase, generator model and discriminator model. The generator model(U-net) produces a binary image (generated image) as output for the respective X-Ray image. The discriminator model classifies whether it is a mock or not based on ground truth.

for classification. In [5-7,20], authors used U-Net for lobe or nodule segmentation. It segments the pixels of similar types and produces a bounded nodule area as input to the target model. Generally, segmentation is meant to increase accuracy, but in the recent past [1,2,4,18,19] authors have given superior accuracy without the need for segmentation.

With the third feature (augmentation), it reduces the need to collect fresh data by increasing the data size. It adds heterogeneity in existing data by doing [5,7,8,15] rotation, flipping, elastic, deformation, and scaling. In each case, the DL model takes these images as a new one for the process. Most of the covid19 and pneumonia classifier needs augmentation where S. Sakib et al., [19] have used GAN [3] (neural network model) for augmentation for the classification of covid19.

TABLE II. ANALYSIS OF TECHNIQUES IN EXISTING WORKS

Prognosticate	Reference	Classes	Augmentation	ML/DL	Name of the Algorithm/Model Description	Image Size	Epoch	Learning Rate	Performance Metrics (Accuracy)
Covid 19	[7]	Non-Pneumonia, ILD (Interstitial Lung Disease), Covid19	Rotation, Flipping	DL	Detection and Classification based separate blocks. 3D ResNet + Prior Attention Strategy	96x96x96	30k	0.0001	Non-Pneumonia: 91.5% ILD:89.4% Covid19:93.3%
	[11]	Normal, Covid19, Pneumonia	Yes	DL	Transfer Learning - VGG16, VGG19, Xception, InceptionResNet, InceptionV3, NASNetLarge, DenseNet121, ResNet50V2	224_224 pixels for VGG16/19 and 299_299 pixels for InceptionV3	100	$10^{-3} - 10^{-5}$	VGG19 86% for X-Ray, 100% for Ultrasound and 84% for CT Scans
	[5]	Covid-19, Pneumonia, Control	horizontal_flip, Gaussian noise with a variance of 0:015, rotation, elastic deformation, and scaling		Modified COVID-NET <sup>2</sup>	224x224	24 Batch size = 32	$2^{-5}$	91.67
	[15]	Normal, Covid19, Pneumonia	rotation, flipping, over-sampling and small distortion	DL	GDCNN	-	100	-	98.84%
	[19]	COVID-19, pneumonia, and normal	Yes - generative adversarial network (GAN)[3] model for data augmentation	DL	DARI algorithm + CNN	-	20	0.001	COVID-19 - 99.45% pneumonia - 84.84%. Normal - 83.04%
	[2]	COVID-19 and normal	Yes	DL + ML	Transfer Learning: CNN trained on ImageNet + ML (Bayes, MLP, SVM, RF, KNN)	224x224	10	-	MobileNet + SVM: 98.5%
Pneumonia	[8]	Viral-Pneumonia and Normal	random cropping, and horizontally flipping	DL	confidence-aware anomaly detection (CAAD)model	448x448x3	20, the batch size to 40	$5 \times 10^{-4}$ , which decays linearly to $10^{-6}$	87.57%
	[10]	i) Normal and COVID-19 pneumonia; ii) Normal, viral and COVID-19	rotation, and translation. After Augmentation=10,027	DL	Transfer Learning - Pre-trained deep Convolutional Neural Networks (CNN)	SqueezeNet:227x227, ResNet18, ResNet101, VGG19 and DenseNet201: 224x224, Inceptionv3: 299x299	mini-batch size: 16,Back Propagation epochs: 20	$10^{-3}$	>97% for all categories
	[12]	Normal, Pneumonia	No	DL + NLP	Multi-channel Multi-modal Recurrent Convolutional Neural Network along with LSTM	25x25x25 voxels converted into vectors for LSTM	3000	Initial Rate $10^{-4}$	95%
	[23]	Healthy, COVID-19, H1N1, CAP	Yes	DL	ResNet-50 based Novel Model	512x512	110/90	0.01/0.001	94.18%
	[22]	Community Acquired Pneumonia (CAP), COVID-19	Yes	DL	Online attention module With a 3D convolutional network (CNN) - two 3D ResNet34 networks with attention map	32x32x3	20	0.0002	87.5%

	[1]	Normal, Pneumonia	No	Neural Model	Quantum BP neural network (QBP) based on quantum particle swarm optimization (QPSO)	64x64	3000	0.01	96.07%
	[4]	Normal, Pneumonia	No	ML	LDA-SVM	6-D feature Vector	-	-	92.75%
TB	[18]	Normal, Viral and bacterial Pneumonia	No	DL	Customized CNN and ImageNet Pretrained model	224x224	100	-	94.1%
	[16]	MDR and XDR tuberculosis - Resistant and susceptible	No	ML	Logistic regression, Random forest and Gradient boosting	-	-	-	90%
	[17]	Normal, Abnormal	No	ML	SVM	2048x2048	-	-	84%
	[6]	Lesion lung portion with abnormality	No	ML	multiple-instance learning (MIL): si-mi SVM +PEDD	1024x1024	-	-	90%
	[24]	TB and non-TB	No	DL	B-CNN	224 x224	-	-	Montgomery -96.42% and Shenzhen and 86.46%
	[20]	TB and non-TB normal	Yes	DL	Transfer Learning: ResNet18, ResNet50, ResNet101, ChexNet, InceptionV3, Vgg19, DenseNet201, SqueezeNet, and MobileNet	InceptionV3: 227x227, ResNet, DenseNet, ChexNet, VGG, MobileNetV2 and SqueezeNet:224 x224	50	10 <sup>-3</sup>	97%
COPD	[9]	Normal, Abnormal that includes TB	No	ML	Voting-based combination of three different classifiers: Bayesian network (BN), Multilayer Perception (MLP) and RF	-	-	-	96%
	[14]	COPD, HC (Healthy Control)	Yes	DL	B-CNN	64x64x64	2200	1e-4 to 1e-1	88.2%
	[21]	COPD and Non-COPD	No	ML	Transfer Learning - SimpleMIL with logistic classifier	41x41x41 voxels	-	-	90.4%
	[13]	COPD and Non-COPD	-	ML	Transfer Learning - balanced probability distribution (BPD) algorithm	26 features	-	-	92.1%

For 4<sup>th</sup> feature is “technique used”. As we step forward into AI, it expounds ML and DL are the pillars. ML algorithm is applied on images for features acquisition, then based on conformity of features, the disease is classified. SVM classifier [17] is used in TB identification and gave an accuracy of 95.2% but it fails in few cases because one should know how the feature is selected by the algorithm for the given input data and this process is based on scenario too. On the other hand, DL peculiarly CNN, which is a multiple-layer model made of neurons that learns features on the process itself. There is no separate technique needed for learning, and one can be free of selecting a feature as it learns automatically. The first layer of CNN takes input, where we give the processed image. Then it is sandwiched with a hidden layer where the learning process is done. At last, the output is given by fully connected (last layer) with

the help activation function. The models of CNN are various in the hidden layers and the output they produce.

For 5<sup>th</sup> and 6<sup>th</sup> feature is said as hyperparameters of the model. They define how the model getting adapted to the given problem. Usually, the Learning rate (Lr) should be small so that the model reaches global minima. J. Wang et al., [7] used 30K epoch with Lr as 0.001 and achieved an accuracy of 93.5%. When Lr is high it takes less epoch i.e., learns quickly. Epoch is just the number of times the model is trained. Training a model with high epoch results in learning more about input data. Sometimes leads to spare data learning that causes the model to overfit. K. C. Santosh et al., [19] have used just 20 epochs with Lr as 0.001 and gave 99.45% as accuracy. The goal of developing DL or ML model is to be generalized i.e., it should have rock-bottom variance and bias for unseen and seen data respectively. K-

fold CV is adopted by a majority of the model to avoid overfitting.

For 7<sup>th</sup> feature(accuracy), the model's precision depends on the quality of data, training method, the framework used, and many factors affecting it. The contemporary need is to develop a generalized model which we have seen in the survey. For the type of lung disease considered, the DL model plays a good role compared to ML. The following table III gives virtues and shortcomings in few papers.

TABLE III. COMPARISON OF EXISTING WORKS

Reference	Advantages	Limitations
[4]	LDA-SVM shows higher accuracy in prediction without Adaboost algorithm.	SVM does not provide probability estimate and fusion of several algorithm leads to many iterations.
[15]	Pattern of disease is identified by genetic algorithm.	High computational and spatial complexity.
[14]	Grayscale, RGB, and binary images are using in training of model.	Just COPD and HC are remembered for the examination. Only 2D snapshots of 3D lung airway tree is used.
[24]	BOA increase the efficiency of CNN model.	No focus on potential validation and BOA time consuming process.
[3]	Accuracy of algorithm is high when compared to another model.	Tedious because it has two phase execution. The model highly dependent on ground truth.
[16]	No augmentation technique adapted and produced an accuracy of 90%	Sputum is used as an input data for the model. Most of latest technologies uses medical images for DL and ML models and produces greater accuracy.

## V. CONCLUSION

In the literature survey, most of the works have been done based on the two major fields, machine learning, and deep learning. Studies resulted that, in machine learning, the combination of LDA-SVM algorithms gives higher accuracy. Likewise, the deep learning GDCNN algorithm gives an accuracy of 98.84%. The accuracy that had been laid out; the attained ones are not sketchy. Additionally, to date, there's no common system that categorizes all possible lung diseases like TB, Pneumonia, COPD, and COVID19, indicating that the proposed systems are dedicated to a sole disorder are some of the gaps identified in our study. On the contrary, the system going to be proposed manifests all the patterns in lung diseases.

## REFERENCES

- [1] D. Yumin, M. Wu and J. Zhang, "Recognition of Pneumonia Image Based on Improved Quantum Neural Network," in *IEEE Access*, vol. 8, pp. 224500-224512, 2020.
- [2] E. F. Ohata et al., "Automatic detection of COVID-19 infection using chest X-ray images through transfer learning," in *IEEE/CAA Journal of Automatica Sinica*, vol. 8, no. 1, pp. 239-248, January 2021.
- [3] F. Munawar, S. Azmat, T. Iqbal, C. Grönlund and H. Ali, "Segmentation of Lungs in Chest X-Ray Image Using Generative Adversarial Networks," in *IEEE Access*, vol. 8, pp. 153535-153545, 2020.
- [4] G. Ling and C. Cao, "Automatic Detection and Diagnosis of Severe Viral Pneumonia CT Images Based on LDA-SVM," in *IEEE Sensors Journal*, vol. 20, no. 20, pp. 11927-11934, 15 Oct.15, 2020.
- [5] J. D. Arias-Londoño, J. A. Gómez-García, L. Moro-Velázquez and J. I. Godino-Llorente, "Artificial Intelligence Applied to Chest X-Ray Images for the Automatic Detection of COVID-19. A Thoughtful Evaluation Approach," in *IEEE Access*, vol. 8, pp. 226811-226827, 2020.
- [6] J. Melendez, B. van Ginneken, P. Maduskar, R. H. H. M. Philipsen, H. Ayles and C. I. Sánchez, "On Combining Multiple-Instance Learning and Active Learning for Computer-Aided Detection of Tuberculosis," in *IEEE Transactions on Medical Imaging*, vol. 35, no. 4, pp. 1013-1024, April 2016.
- [7] J. Wang et al., "Prior-Attention Residual Learning for More Discriminative COVID-19 Screening in CT Images," in *IEEE Transactions on Medical Imaging*, vol. 39, no. 8, pp. 2572-2583, Aug. 2020
- [8] J. Zhang et al., "Viral Pneumonia Screening on Chest X-Rays Using Confidence-Aware Anomaly Detection," in *IEEE Transactions on Medical Imaging*, vol. 40, no. 3, pp. 879-890, March 2021.
- [9] K. C. Santosh and S. Antani, "Automated Chest X-Ray Screening: Can Lung Region Symmetry Help Detect Pulmonary Abnormalities?," in *IEEE Transactions on Medical Imaging*, vol. 37, no. 5, pp. 1168-1177, May 2018.
- [10] M. E. H. Chowdhury et al., "Can AI Help in Screening Viral and COVID-19 Pneumonia?," in *IEEE Access*, vol. 8, pp. 132665-132676, 2020.
- [11] M. J. Horry et al., "COVID-19 Detection Through Transfer Learning Using Multimodal Imaging Data," in *IEEE Access*, vol. 8, pp. 149808-149824, 2020.
- [12] Q. Wang, D. Yang, Z. Li, X. Zhang and C. Liu, "Deep Regression via Multi-Channel Multi-Modal Learning for Pneumonia Screening," in *IEEE Access*, vol. 8, pp. 78530-78541, 2020.
- [13] Q. Wang, H. Wang, L. Wang and F. Yu, "Diagnosis of Chronic Obstructive Pulmonary Disease Based on Transfer Learning," in *IEEE Access*, vol. 8, pp. 47370-47383, 2020.
- [14] R. Du et al., "Identification of COPD From Multi-View Snapshots of 3D Lung Airway Tree via Deep CNN," in *IEEE Access*, vol. 8, pp. 38907-38919, 2020.
- [15] R.G. Babukarthik, V. A. K. Adiga, G. Sambasivam, D. Chandramohan and J. Amudhavel, "Prediction of COVID-19 Using Genetic Deep Learning Convolutional Neural Network (GDCNN)," in *IEEE Access*, vol. 8, pp. 177647-177666, 2020.
- [16] R. S. Sergeev et al., "Genome-Wide Analysis of MDR and XDR Tuberculosis from Belarus: Machine-Learning Approach," in *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 16, no. 4, pp. 1398-1408, 1 July-Aug. 2019.
- [17] S. Jaeger et al., "Automatic Tuberculosis Screening Using Chest Radiographs," in *IEEE Transactions on Medical Imaging*, vol. 33, no. 2, pp. 233-245, Feb. 2014.
- [18] S. Rajaraman and S. K. Antani, "Modality-Specific Deep Learning Model Ensembles Toward Improving TB Detection in Chest Radiographs," in *IEEE Access*, vol. 8, pp. 27318-27326, 2020.
- [19] S. Sakib, T. Tazrin, M. M. Fouda, Z. M. Fadlullah and M. Guizani, "DL-CRC: Deep Learning-Based Chest Radiograph Classification for COVID-19 Detection: A Novel Approach," in *IEEE Access*, vol. 8, pp. 171575-171589, 2020.
- [20] T. Rahman et al., "Reliable Tuberculosis Detection Using Chest X-Ray With Deep Learning, Segmentation and Visualization," in *IEEE Access*, vol. 8, pp. 191586-191601, 2020.
- [21] V. Cheplygina et al., "Transfer Learning for Multicenter Classification of Chronic Obstructive Pulmonary Disease," in *IEEE Journal of Biomedical and Health Informatics*, vol. 22, no. 5, pp. 1486-1496, Sept. 2018.
- [22] X. Ouyang et al., "Dual-Sampling Attention Network for Diagnosis of COVID-19 From Community Acquired Pneumonia," in *IEEE Transactions on Medical Imaging*, vol. 39, no. 8, pp. 2595-2605, Aug. 2020.
- [23] X. Qian et al., "M3Lung-Sys: A Deep Learning System for Multi-Class Lung Pneumonia Screening From CT Imaging," in *IEEE Journal of Biomedical and Health Informatics*, vol. 24, no. 12, pp. 3539-3550, Dec. 2020.
- [24] Z. Ul Abideen et al., "Uncertainty Assisted Robust Tuberculosis Identification with Bayesian Convolutional Neural Networks," in *IEEE Access*, vol. 8, pp. 22812-22825, 2020.

# Extraction of Tumour in Breast MRI using Joint Thresholding and Segmentation – A Study

Seifedine Kadry

Faculty of Applied Computing and  
Technology, Noroff University College,  
Kristiansand, Norway  
email: skadry@gmail.com

Venkatesan Rajinikanth

Department of Electronics and  
Instrumentation Engineering  
St. Joseph's College of Engineering  
Chennai 600119, India  
email: v.rajinikanth@ieee.org

Robertas Damaševičius

Faculty of Applied Mathematics  
Silesian University of Technology  
44-100 Gliwice, Poland  
email: robertas.damasevicius@polsl.pl

Isah A. Lawal

Faculty of Applied Computing and  
Technology, Noroff University College,  
Kristiansand, Norway  
email: Isah.Lawal@noroff.no

David Taniar

Faculty of Information Technology  
Monash University  
Clayton, Victoria 3800, Australia  
email: David.Taniar@monash.edu

**Abstract**— Breast Cancer (BC) is one of the harsh conditions, which largely affects the women group. Due to its significance, a range of procedures are available for premature detection and early treatment to save the patient. The clinical level diagnosis of BC will be done using; (i) Image supported detection and (ii) Core-Needle-Biopsy (CNB) assisted confirmation. The proposed work aim to develop a computerized scheme to detect the Breast-Tumor-Section (BTS) from the breast MRI slices. This work implements a joint thresholding and segmentation methodology to enhance and extract the BTS from the 2D MRI slices. A tri-level thresholding based on Slime-Mould-Algorithm and Shannon's-Entropy (SMA+SE) is implemented to enhance the BTS and Watershed-Segmentation (WS) is implemented to mine the BTS. After extracting the BTS, a study between the BTS and Ground-Truth image is performed and the necessary Image-Performance-Values (IPV) are computed. In this work the axial, coronal and sagittal slices of 2D breast MRI are separately examined and the attained results are presented.

**Keywords**— Breast abnormality, Breast MRI, Slime-Mould-Algorithm, Thresholding, Segmentation..

## I. INTRODUCTION

Cancer is one of the harsh diseases in humans and the untreated cancer will lead to painful casualty. The earlier reports confirm that, every cancer is curable, when it is detected in its pre-mature stage. Breast-Cancer (BC) is one of the major cancer in women group and in the recent years, the occurrence rates of the BC is gradually rising due to various factors including the heredity [1-4].

The early stage recognition of BC will be done only using the recommended clinical protocol. The later stage of BC can be recognised with one/all of the following symptoms; lump in breast, alteration in the shape/size of breast, change in the skin colour in infected section, pain, etc . Further, the BC is classified into various groups including the Ductal-Carcinoma-in-Situ (DCIS) and Lobular-Carcinoma-in-Situ (LCIS) [5,6].

When the a symptom of the breast cancer is identified the patient, the patient will approach the doctor for further examination using the prescribed clinical protocol including the initial examination using the image supported method and the final stage confirmation with Core-Needle-Biopsy (CNB). The implementation of CNB is one of the more painful and invasive technique in which the breast tissue samples are collected, evaluated and preserved in a fully controlled clinical environment. The initial stage (image

based assessment) is a common methodology in which the suspicious breast section is examined using the chosen imaging modalities. The common modalities for BC detection include Mammogram, breast-ultrasound, breast-thermal-imaging and breast MRI [3,4,6-9]. The Mammogram is one of the oldest and widely accepted BC screening techniques [7]. The recent approaches includes the ultrasound, thermal imaging and MRI and in which the MRI based assessment gives the reconstructed three-dimensional (3D) view of the breast; which can be analysed more precisely compared to other approaches. Further, the 3D breast MRI can be converted into 2D slice and the evaluation of 2D slice will help to confirm the dimension and harshness of the cancer and in some cases this kind of assessment also help to take the decision regarding the BC treatment without recommending the CNB.

Image based examination of the breast abnormality is widely reported by the researchers and in every technique, a semi-automated/automated methodology is employed to mine the BTS from the chosen image modality [10]. The assessment of breast MRI is simple and in this modality, the abnormal tumour section will have better pixel visibility compared with others section. The assessment of the breast MRI can be found in [3,11].

In this work, the breast MRI of the benchmark images of the Reference Image Database to Evaluate Therapy Response (RIDER) [12-14] is considered for the assessment. For the assessment 10 patient's images (10 patient x 20 slices = 200 images) are collected and the 3D to 2D conversion is implemented using the ITK-Snap [15,16]. The 2D slices of Axial (AX), Coronal (CO) and Sagittal (SA) slices are individually examined and the results are separately evaluated.

Extraction of the BTS from the considered 2D slices is achieved using the joint thresholding and segmentation approach. Initially, a tri-level thresholding is executed using SMA+SE and the enhanced tumour segment is mined with WS. The proposed technique is an automated segmentation procedure and helps to mine the BTS with better accuracy in all the three views of the breast MRI. After extracting the BTS, a comparison is implemented with the Ground-Truth (GT) and the Image-Performance-Values (IPV) are computed and based on the superiority of these values; the merit of the examination scheme is validated.

The other division of the research are organised as below; Section 2 discuss the context, Section 3 demonstrates the

methodology, Section 4 and 5 give results and the conclusion of the proposed work, respectively.

## II. CONTEXT

Breast abnormality is one of the medical emergencies and a considerable number of research works are proposed and implemented to examine the breast-cancer using image assessment techniques. The earlier works on the breast abnormality can be categorized based on the imaging methods.

The examination of the DCIS using the image processing technique can be found in the earlier procedures [2,6,9]. This work implemented a chosen tri-level thresholding and segmentation process to extract and evaluate the abnormal fragment from the breast images [8-11]

The earlier works on the mammogram supported breast assessment can be found in [7,17]. In this method the existing abnormality in the breast X-ray is examined using a suitable imaging technique and the results are presented and discussed. The breast ultrasound is one of the non-invasive imaging approach and the implementation of joint thresholding and segmentation can be found in [8].

The thermal imaging supported breast disease detection using a semi-automated technique is discussed in the work of Raja et al. [4] and Steven et al. [6]. From these works, it can be noted that the early detection of the breast cancer can be found with the carefully collected breast thermal images. The segmentation of infected breast section based on the thresholding and segmentation methodology is clearly discussed by the work of Raja et al. [18]. This work adopted the benchmark Magnetic-Resonance-Angiogram (MRA) images for the assessment and the attained results confirms that the segmentation technique implemented helped to achieve a better result.

In the proposed research, similar methodology is adopted to pre-process and post-process the considered breast MRI slices using a chosen thresholding and segmentation scheme. In order to attain better result, a much case is paid towards the selection of the thresholding process and based on the previous research work, the SE based thresholding is adopted. Further, to achieve a better thresholding, the threshold selection is optimized using the Slime-Mould-Algorithm (SMA). The BTS from the threshold image is automatically extracted using WS approach. In this work, the essential test images are collected from the RIDER-TCIA database and very image is separately examined with the considered procedure and based on the outcome, the performance is validated.

## III. METHODOLOGY

Evaluation of medical images is very essential to identify the disease in patients. The needs for an essential image processing methodology are essential to extract the infected section from the images with better accuracy. The implemented breast MRI examination process is depicted in Figure 1. The various stages of this scheme are; collection of the 3D clinical grade MRI, 3D to 2D conversion with ITK-Snap, implementation of tri-level thresholding with SMA+SE to enhance the tumour segment, extraction of BTS with WS, comparing the BTS with GT and computing the IPV. Based on the attained average values of the IPV, the

significance of the proposed breast MRI based tumour examination technique is confirmed.

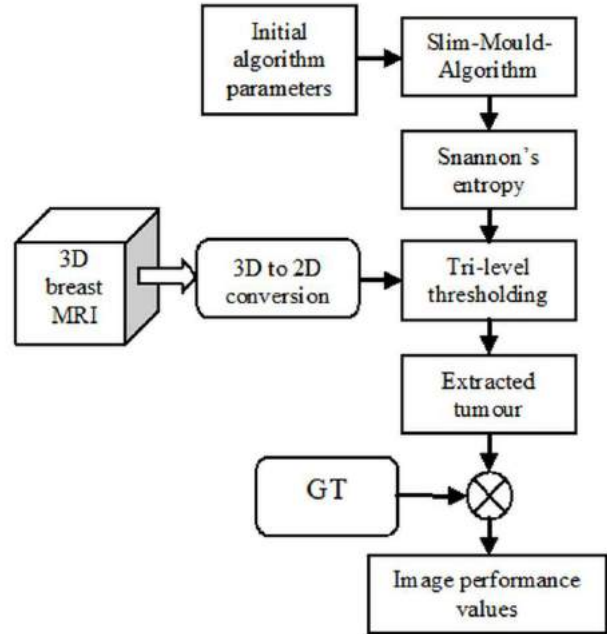


Fig.1. Scheme of proposed breast MRI assessment system

### A. RIDER-TCIA Database

Most of the clinical grade images are protected with the ethical and copyright law and hence the availability of clinical grad images is very limited. In this work, the breast MRI existing in RIDER-TCIA [12-14] is considered for the assessment and this image is balanced like clinical images and in this study, only 10 patient's images are considered for the evaluation. From each patient, 20 slices are collected for the investigation along with its GT. Figure 2 depicts the sample images of 4 patients.

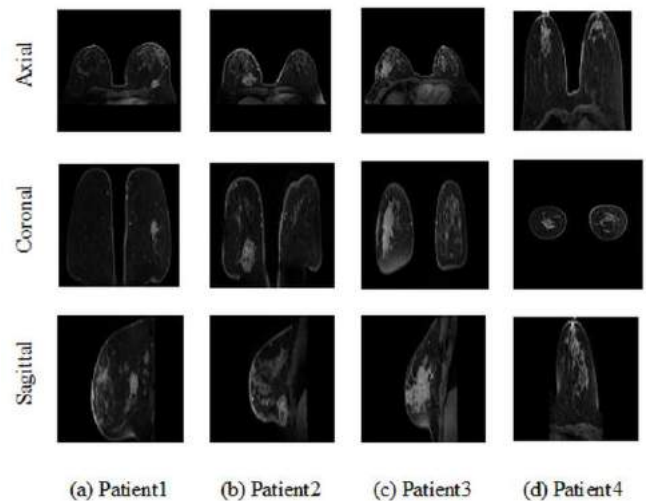


Fig.2. Sample test images of the RIDER-TCIA database

During the image collection process, a 3D to 2D conversion is implemented with the ITK-Snap [15,16] and this helps to split the 3D image into axial, coronal and sagittal slices and based on the need, we can extract the necessary slices, which include the tumour. The initial image which is extracted is of dimension 288x288x3 pixels (RGB scale) and to have better visibility, all these images are



resized to 512x512x3 pixels and the resized images are then examined. The main merit of this database is, all the images are associated with the GT.

Figure 3 depicts the sample trial images of different 2D views along with its GT images. Similar case is applicable for all other images of the considered database.

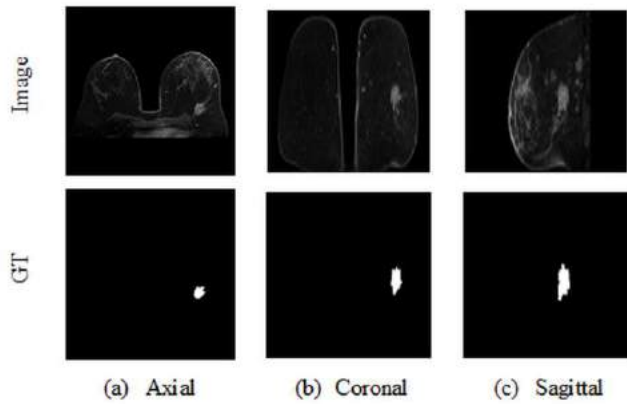


Fig.4. Sample 2D slices of MRI with different views

### B. Tri-level Thresholding

The earlier work in the literature confirms that the thresholding combined with the segmentation helped to achieve a better result during the medical image assessment. Moreover, heuristic algorithm supported thresholding is one of the commonly implemented procedure for the medical image assessment and the result achieved with this technique are superior compared to traditional thresholding [18,19]. Hence, in this work, a tri-level thresholding is executed using SMA+SE.

#### 1) Slime-Mould-Algorithm

SMA is one of the recently developed heuristic approaches by Li et al. [20] to solve complex optimization problem. This algorithm is formulated by imitating the food searching pattern of Slime-Mould. Compared to other methods, the number of algorithm parameters to be assigned is very minimal; number of agents, dimension, upper-bound, lower-bound, search methodology and the stopping criteria. The basic code for the SMA is available in [21] and in this work the SMA is implemented to maximizing the SE by randomly varying the breast MRI's threshold.

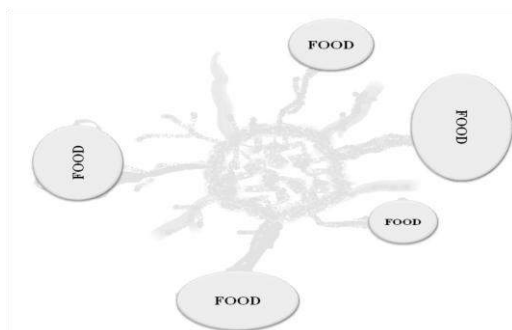


Fig. 4. Structure of SMA searching the food

Figure 4 depicts the basic structure of the SMA based food search. The essential information and the mathematical description can be found in [22]. In this work, the following initial algorithm parameters are assigned; agent size=30,

search dimension = 3, upper-bound=255, lower-bound=0, number of iterations=2000 and stopping criteria = maximal iteration or maximized SE.

#### 2) Shannon's Entropy

SE is one of the established entropy supported thresholding; widely adopted to enhance the medical images with a tri-level thresholding [23].

Let, there is an image of dimension A\*B. The pixel organization in picture (a,b) is defined as P(a,b), for  $a \in \{1,2,\dots,A\}$  &  $b \in \{1,2,\dots,B\}$ . Let L (255) is the total gray-levels in image  $\{0,1,2,\dots,L-1\}$  denoted as O, in such a way that:

$$P(A,B) \in O \quad \forall (a,b) \in \text{Image} \quad (1)$$

Then, the regularized histogram will be;

$$S = (t_0, t_1, \dots, t_{L-1}) \quad (2)$$

For a tri-level thresholding case, Eqn. (2) becomes;

$$S(T) = x_0(T_1) + x_1(T_2) + x_3(T_3) \quad (3)$$

$$E(T) = \max_T \{S(T)\} \quad (4)$$

where,  $T = \{T_1, T_2, \dots, T_L\}$  denotes the thresholds,  $S = \{x_0, x_1, \dots, x_{L-1}\}$  denotes regularized histogram, and  $E(T)$  is the optimal threshold. Other essential information on the SE can be found in [8].

### C. Watershed Segmentation

The WS technique is a famous automated technique workers based on an assigned marker and helps performs a better segmentation. The WS executes operations, such as canny edge-detection, watershed generation, morphological enhancement and segmentation. The earlier works on WS will help to get the essential details [17,23]. In this work, the WS is employed to extract the BTS from the breast MRI.

### D. Image Performance Value Computation

After developing the image examination scheme, its performance is to be confirmed to ensure its clinical significance. In this work, the extracted BTS is evaluated against the GT and the required Image-Performance-Values (IPV) such as True-Positive (TP), False-Positive (FP), False-Negative (FN), and True-Negative (TN) are primarily calculated. All these values denote the pixels of the image background and the BTS and the TP denotes the BTS pixel dimension and TN specifies the background pixel value. From the above computed values; other measures, such as Jaccard, Dice, Accuracy (ACC), Precision (PRE), Sensitivity (SEN), Specificity (SPE), and Negative-Predictive-Value (NPV) computed and the mathematical expression for these IPV can be accessed from [2-4, 24,25].

## IV. RESULT AND DISCUSSION

This section of the paper presents the experimental outcome attained for the sample image depicted in figure 3 and the experimental work is executed in MATLAB®.

The results achieved for the axial view breast MRI is depicted in Figure 5. Fig 5(a) presents the chosen trial picture and Fig 5(b) denotes the result of the SMA+SE thresholding.

From this figure, it can be seen that the implemented tri-level thresholding enhanced the tumor section by reducing the pixel values of other sections. Later, the WS segmentation is implemented with a chosen marker dimension of 10 pixels. The employed WS presents one final image and a number of intermediate images. Fig 5(c) - (f) demonstrate the detected edges, initial watershed, initial and final morphological enhancement, respectively. Fig 5(g) presents the final watershed enhanced pixel groups in which the pixel group depicted in yellow in color is to be extracted. The extracted BTS will be in the form of the binary image as in Fig 5(h). Similar results are attained with all other images and the results attained with the chosen test image are depicted in Figure 6 for coronal (Fig 6(a)) and sagittal (Fig 6(b)) view.

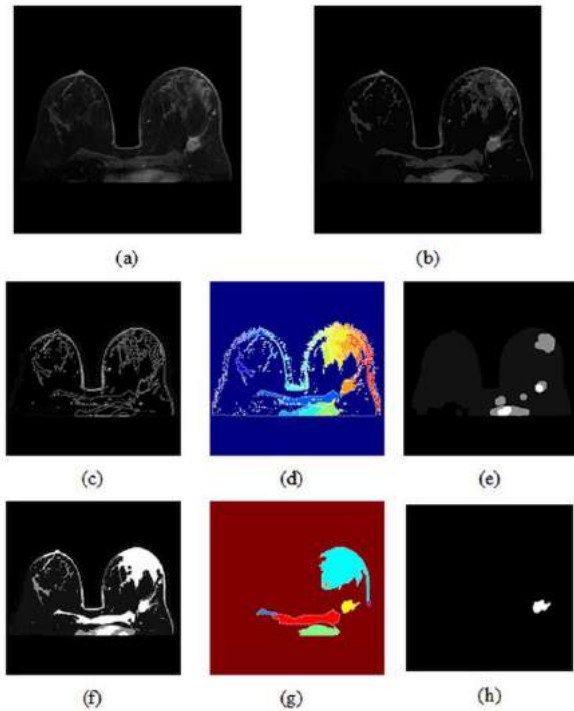


Fig.5. Various results attained with the proposed scheme

(a) Trial picture, (b) Thresholded image, (c) Detected edges, (d) Initialized watershed, (e) and (f) Morphological dilation and enhancement respectively, (g) Final watershed and (h) Extracted tumour

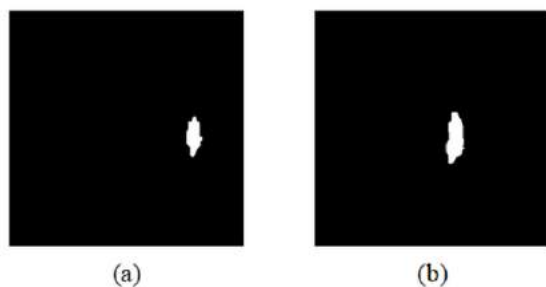


Fig.6. Segmented BTS  
(a) Coronal, (b) Sagittal

After extracting the breast tumour from the MRI slices, comparison is employed with the GT and the IPV's are then computed. The results attained for AX, Co and SA are depicted in Table I and II. This table confirms that the proposed approach is very resourceful in extracting the BTS and offers the better values of IPV's. Similar procedure is repeated for all other test pictures and the average of the

IPV's is considered to validate the tumour assessment performance of the proposed technique. This outcome of this comparison is depicted in Figure 7 and these values confirm that this scheme works well on the slice vies, such as AX, CO and SA.

TABLE I. ESSENTIAL PERFORMANCE MESURE COMPUTED BETWEEN SEGMENTED SECTION AND GT

Image	TP	FN	FP	TN	Jaccard	Dice
AX	811	110	66	261157	82.1682	90.2113
CO	1753	208	53	260130	87.0407	93.0714
SA	2596	462	119	258966	81.7123	89.9359

TABLE II. IMAGE PERFORMANCE MEASURES OF SAMPLE TEST IMAGES

Image	ACC	PRE	SEN	SPE	NPV
AX	99.9329	92.4743	88.0565	99.9747	99.9579
CO	99.9004	97.0653	89.3932	99.9796	99.9201
SA	99.7784	95.6169	84.8921	99.9541	99.8219

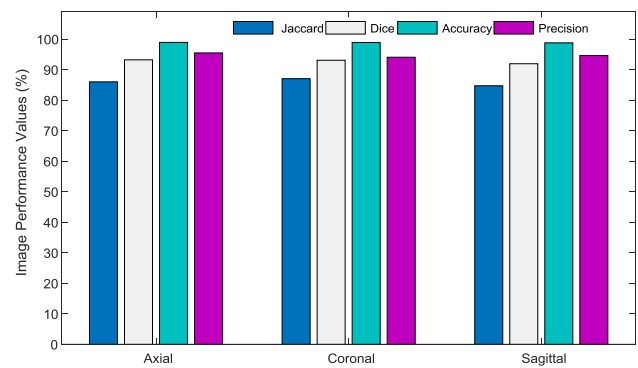


Fig. 7. Evaluation of average IPV's various MRI slices

This research work proposed a methodology to examine the breast MRI using the automated scheme. This work employed the SMA+SE thresholding and WS segmentation to extract and examine the BTS from the chosen picture. In future, the attained results of this work can be improved by replacing the SMA with the Red-Fox-Algorithm invented in 2021 [26]. Further, the performance of SE thresholding can be confirmed with other approaches available in the literature. Also, the segmentation of the WS is to be validated against other segmentation procedures.

## V. CONCLUSION

The early and accurate detection of the breast cancer section will support the appropriate treatment implementation to cure the disease. The proposed work implemented a combined thresholding and segmentation methodology to evaluate the breast MRI slices of RIDER-TCIA database. In this work 200 numbers of images with axial, coronal and sagittal views are investigated using the proposed scheme. This work executed SMA+SE based tri-level thresholding and WS segmentation to extract the BTS. After mining the BTS from the MRI slice, a comparison with GT is performed and the essential values of the IPV's are computed. The overall result attained with the proposed result on the various views of the breast MRI slices confirms that this scheme is clinically significant and in future; it can be employed to evaluate the clinically obtained breast MRI.

## REFERENCES

- [1] <https://www.who.int/cancer/detection/breastcancer/en/>
- [2] N. Dey, V. Rajinikanth, and A.E. Hassanien, "An Examination System to Classify the Breast Thermal Images into Early/Acute DCIS Class," In *Proceedings of International Conference on Data Science and Applications*, pp. 209-220. Springer, Singapore, 2020.
- [3] R. Elanthirayan, K. S. Kubra, V. Rajinikanth, N.S.M. Raja, and S.C. Satapathy, "Extraction of Cancer Section from 2D Breast MRI Slice Using Brain Stom Optimization," In *Intelligent Data Engineering and Analytics*, pp. 731-739. Springer, Singapore, 2021. [https://doi.org/10.1007/978-981-15-5679-1\\_71](https://doi.org/10.1007/978-981-15-5679-1_71).
- [4] N.S.M. Raja, V. Rajinikanth, S.L. Fernandes and S.C. Satapathy, "Segmentation of breast thermal images using Kapur's entropy and hidden Markov random field," *J. Med. Imaging Health Info.*, vol.7, no.8, pp. 1825–1829, 2017. <https://doi.org/10.1166/jmihi.2017.2267>.
- [5] <https://www.mayoclinic.org/diseases-conditions/breast-cancer/symptoms-causes/syc-20352470>
- [6] S.L. Fernandes, V. Rajinikanth, and S. Kadry, "A hybrid framework to evaluate breast abnormality using infrared thermal images," *IEEE Consum. Electron. Mag.*, vol. 8, no.5, pp.31–36, 2019. <https://doi.org/10.1109/MCE.2019.2923926>.
- [7] H.S. Sheshadri, and A. Kandaswamy, "Breast tissue classification using statistical feature extraction of mammograms," *Medical Imaging and Information Sciences*, vol. 23, no. 3, pp. 105-107, 2006.
- [8] R.I.R.Thanaraj, B. Anand, J. A. Rahul, and V. Rajinikanth, "Appraisal of Breast Ultrasound Image Using Shannon's Thresholding and Level-Set Segmentation," In *Progress in Computing, Analytics and Networking*, pp. 621-630. Springer, Singapore, 2020.
- [9] V. Rajinikanth, N.S.M. Raja, S.C. Satapathy, N. Dey, and G.G. Devadhas, "Thermogram assisted detection and analysis of ductal carcinoma in situ (DCIS)," In: *International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICICT)*, IEEE, pp.1641–1646, 2018. <https://doi.org/10.1109/icicict1.2017.8342817>.
- [10] A.Q. Al-Faris, U. K. Ngah, N. A. M. Isa, and I. L. Shuaib, "Breast MRI tumour segmentation using modified automatic seeded region growing based on particle swarm optimization image clustering," In *Soft Computing in Industrial Applications*, pp. 49-60. Springer, Cham, 2014.
- [11] N. Shrivastava, and J. Bharti, "Breast tumor detection and classification based on density," *Multimedia Tools and Applications*, vol.79, no. 35, pp.26467-26487, 2020.
- [12] C.R. Meyer, T.L. Chenevert, C.J. Galbán, T.D. Johnson, D.A. Hamstra, A. Rehemtulla, and B.D. Ross, "Data From RIDER Breast MRI," *The Cancer Imaging Archive*, 2015. <https://doi.org/10.7937/K9/TCIA.2015.H1SXNUXL>.
- [13] K. Lark , B. Vendt, K. Smith et al., "The Cancer Imaging Archive (TCIA): Maintaining and Operating a Public Information Repository," *Journal of Digital Imaging*, vol. 26, no.6, pp.1045-1057, 2013. <https://doi.org/10.1007/s10278-013-9622-7>.
- [14] <https://wiki.cancerimagingarchive.net/display/Public/RIDER+Breast+MRI>
- [15] P.A. Yushkevich, J. Piven, H. C. Hazlett, R. G. Smith, S. Ho, J. C. Gee, and G. Gerig, "User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability," *Neuroimage*, vol.31, no.3, pp. 1116-1128, 2006. <https://doi.org/10.1016/j.neuroimage.2006.01.015>.
- [16] <http://www.itksnap.org/pmwiki/pmwiki.php?n=Main.Publications>
- [17] S.P.S. Raj et al., "Examination of digital mammogram using otsu's function and watershed segmentation", In. *Fourth International Conference on Biosignals Images and Instrumentation (ICBSII)*, pp. 206-212, 2018. DOI: 10.1109/ICBSII.2018.8524794.
- [18] N. S. M. Raja, S. L. Fernandes, N. Dey, S. C. Satapathy, and V. Rajinikanth, "Contrast enhanced medical MRI evaluation using Tsallis entropy and region growing segmentation," *Journal of Ambient Intelligence and Humanized Computing*, pp.1-12, 2018. <https://doi.org/10.1007/s12652-018-0854-8>.
- [19] V. Rajinikanth, S.C. Satapathy, N. Dey, S.L. Fernandes and K.S. Manic, "Skin melanoma assessment using Kapur's entropy and level set—A study with bat algorithm," vol. 104, pp. 193-202, 2019, [online] Available: [https://doi.org/10.1007/978-981-13-1921-1\\_19](https://doi.org/10.1007/978-981-13-1921-1_19).
- [20] S. Li, H. Chen, M. Wang, A. A. Heidari, and S. Mirjalili, "Slime mould algorithm: A new method for stochastic optimization," *Future Generation Computer Systems*, vol. 111, pp.300-323, 2020.
- [21] <https://in.mathworks.com/matlabcentral/fileexchange/76619-slime-mould-algorithm-sma-a-method-for-optimization>
- [22] M. Premkumar et al., "MOSMA: Multi-Objective Slime Mould Algorithm Based on Elitist Non-Dominated Sorting," *IEEE Access*, vol.9, pp. 3229 – 3248, 2020. DOI: 10.1109/ACCESS.2020.3047936.
- [23] V. Rajinikanth, K.P. Thanaraj, S.C. Satapathy, S.L. Fernandes, and N. Dey, "Shannon's entropy and watershed algorithm based technique to inspect ischemic stroke wound," In *Smart intelligent computing and applications*, pp. 23-31. Springer, Singapore, 2019. [https://doi.org/10.1007/978-981-13-1927-3\\_3](https://doi.org/10.1007/978-981-13-1927-3_3).
- [24] M. Nisa, J. H. Shah, S. Kanwal, M. Raza, M. A. Khan, R. Damaševičius, and T. Blažauskas, "Hybrid malware classification method using segmentation-based fractal texture analysis and deep convolution neural network features," *Applied Sciences*, vol. 10, no. 14, pp.4966, 2020.
- [25] S.A. Roseline, S. Geetha, S. Kadry, and Y. Nam, "Intelligent Vision-Based Malware Detection and Classification Using Deep Random Forest Paradigm," *IEEE Access*, vol. 8, pp.206303-206324, 2020.
- [26] D. Połap, and M. Woźniak, "Red fox optimization algorithm," *Expert Systems with Applications*, vol. 166 pp.114107, 2021. <https://doi.org/10.1016/j.eswa.2020.114107>.

# Machine Learning Approach to Study the Impact of Obesity on Autonomic Nervous System using Heart Rate Variability Features

SR Rathod<sup>1</sup>, UM Chaskar<sup>2</sup>, CY Patil<sup>3</sup>

Department of Instrumentation and Control Engineering<sup>1,2,3</sup>

College of Engineering, Pune-India

rathodsr@yahoo.com<sup>1</sup>, u\_chaskar@yahoo.com<sup>2</sup>, cypatil@gmail.com<sup>3</sup>

**Abstract**—Obese people have high chances of cardiovascular disease (CVD), which is supposed to be due to the alteration in autonomic nervous system (ANS) activity. The changes in ANS activity can be identified using heart rate variability (HRV). HRV is a non-invasive tool to measure the ANS activity using linear and non-linear HRV features. The paper presents an aim to understand the effect of obesity on ANS using HRV parameters. Initially, sixteen control and sixteen obese subjects of both the gender between ages 20 to 50 were involved in the study after that synthetic minority oversampling technique (SMOTE) was used to increase the sample size of control and obese subjects from sixteen to forty-eight. The statistically significant difference between two groups was observed using the Independent  $t$  test. The statistical results of the study indicate the sympathovagal imbalance due to reduced parasympathetic activity. The statistical results were validated by incorporating the machine learning technique into the study. Machine Learning (ML) algorithm helps to identify the most important predictor that can clearly differentiate control and obese subjects. The statistical and ML algorithm result shows changes in the sympathovagal balance due to decreased parasympathetic activity.

**keywords**-Obesity, Cardiovascular disease (CVD), Autonomic nervous system (ANS), Heart rate variability (HRV), Synthetic minority oversampling technique (SMOTE), Machine Learning (ML).

## I. INTRODUCTION

As per the World health organization (WHO), obesity is one of the leading disorder that enhance mortality in an obese person. The definition of obesity says an excessive fat accumulation in the body that resulted in chronic diseases like hypertension, CVD, myocardial infarction (MI), and diabetes. Many researchers have found a strong correlation between obesity and CVD [1]. The study has suggested that an imbalance of autonomic activity increases CVD chances

in obesity [2]. The ANS is a control mechanism of the body that generally maintains homeostasis in the body. ANS regulates the glands, blood vessels, and internal organs. The ANS is divided into two branches sympathetic nervous system (SNS) and the parasympathetic nervous system (PNS). The SNS mobilizes the body systems to provide energy for the fight or flight response, whereas PNS conserve the energy by regulating the rest and digest response.

HRV measures the effect of the ANS function on heart as the vagal nerve is the mediator between ANS and heart. Even a small change in ANS resulted in changes in heart rhythm. HRV is a variation in the RR interval of electrocardiogram (ECG). Thus, HRV could be the most important and non-invasive method to investigate the impact of obesity on ANS. The significantly decreased HRV in obesity increase the chances of CVD[2,3]. ANS control vital organs of the body, fluctuation in these organs can be represented using linear and non-linear HRV parameters.

The paper is organized as follows- Section II presents the methodology where subjects, criteria for obesity, statistical test, and machine learning algorithm are discussed, section III discussed the statistical and machine learning results , and finally, the conclusion is given in section IV.

## II. METHODOLOGY

### A. Subjects

This study was performed solely for research purposes at the institute level with the permission of Dean Research and Development of College of Engineering Pune. The researcher and subjects have made a voluntary agreement. The study involves the electrocardiogram(ECG) acquisition of sixteen

normotensive obese individuals and sixteen control subjects between 20 to 50 years of age of both genders who participated in the study. However, sixteen sample size of control and obese are not sufficient to analyse the statistical results. Thus we have synthetically increased the sample size of control and obese subjects using the Synthetic Minority Oversampling Technique(SMOTE)[4]. It is powerful and most widely used technique. It creates random set of samples to balance minority class. New synthetic data samples are generated between randomly chosen minority class sample and its nearest neighbors samples. The details about the implementation of SMOTE technique is given in Algorithm 1.

---

**Algorithm 1** SMOTE Algorithm
 

---

**Input:** Dataset  $D, \{y_i \in T\}$  where  $i = 1, 2, \dots, T$

Number of minority samples( $T$ )

SMOTE Percentage( $P$ )

Number of nearest neighbors( $k$ )

**Output:** Synthetic Samples

**for**  $i = 1, 2, \dots, T$  **do**

1) Find the  $k$ -nearest neighbors of  $y_i$

2)  $\hat{P} = \lceil \frac{P}{100} \rceil$

**while**  $\hat{P} \neq 0$  **do**

1) Select randomly one of the  $k$ -nearest neighbors  $y_{n(i)}$

2) Choose randomly  $\delta \in [0, 1]$

3)  $y_{new} = y_i + \delta (y_{n(i)} - y_i)$

4)  $T \leftarrow T + y_{new}$

5)  $\hat{P} = \hat{P} - 1$

**end while**

**end for**

---

The SMOTE percentage we have selected to increase the sample size was 200%, then the data was increased from sixteen to forty-eight. So forty-eight control and forty-eight obese subjects HRV measures were used in the analysis.

### B. Criteria to decide obesity

As per the World health organization(WHO) guidelines, obesity is determined using BMI. The BMI is calculated as  $BMI = \text{Weight}(kg)/\text{Height}(m^2)$ [10]. Subjects with a BMI value of 18 to 25( $kg/m^2$ ) are considered normal or non-obese, and subjects with a BMI value of more than 30( $kg/m^2$ ) are considered obese[5, 6].

### C. ECG Recording and HRV Analysis

The ECG of sixteen obese and control subjects was recorded using a standard ECG system at sampling frequency of 500Hz in a supine resting position for 15 minutes. The last five minutes segments were used for HRV analysis in the Heart rate variability analyzer of Biomedical Workbench LabVIEW. The parameters of HRV were determined using a linear and non-linear method[7]. The linear process involves the analysis of the time domain and frequency domain. In the time domain, the RR interval signal is used to extract the statistical parameters such as *mean HR*, *mean RR*, SDNN, and RMSSD. The *mean HR* and *mean RR* represents the average values of heart rate and RR interval. The SDNN indicates the standard deviation of the normal to normal(NN) interval, and RMSSD represents the root mean square of the standard deviation of the NN interval. The Fast Fourier Transform (FFT) technique was used to calculate the frequency domain parameters of HRV. The frequency-domain parameters that are extracted from HRV are total power(TP)( $ms^2$ ), low frequency(LF) and high frequency(HF) power, low frequency(LF) and high frequency(HF) in normalized unit, and LF/HF ratio. The two non-linear HRV features are extracted from the HRV signal, i.e., SD1 and SD2. The SD1 and SD2 are two Poincare Plot measures. SD1 represents the short term variability in the NN interval, and SD2 is the long term variability in NN interval. The control and obese group were analyzed based on the linear and non-linear HRV parameters[8,9,10,11,12].

### D. Statistical Analysis

The statistically significant difference between the control and obese group was observed using the Independent  $t$  test. The results are presents as *mean  $\pm$  standard deviation*. An  $\alpha$  level of  $p < 0.05$  was considered statistically significant. The data were analyzed with Epi.Info statistical software tool.

### E. Machine Learning Algorithm

While analyzing the control and obese group using a statistical test, it was observed that most of all, the HRV linear and non-linear parameters show a significant difference. But the important predictor that can differentiate the control and obese group was not identified using a statistical test. Thus we have used a non-linear machine-learning algorithm to find important predictors between control and obesity. In this study, we have used two non-linear machine learning algorithms, i.e., Classification and Regression Tree(CART)[13] and Gradient

Boosting Decision Tree(GBDT)[14]. Both of these algorithms are used to find important predictors. The important predictor was found out using the feature importance score. Feature importance indicates the importance of each features. A feature importance score greater than 90% is considered an important predictor. The important predictor feature will be given as input to the ML algorithm, and its performance was evaluated using performance metrics.

Six different classification quality evaluation measurements such as accuracy, sensitivity, specificity, precision, F1 score, and Area under the receiver operating characteristic curve (AUC) were used. These classification measures calculated using the following confusion matrix.

TABLE I  
CONFUSION MATRIX

		Actual Value	
		Positive	Negative
Predicted Value	Positive	TP	FP
	Negative	FN	TN

TP-True Positive, FP-False Positive, FN- False Negative, TN- True Negative.

The details about the performance measure are as follow-

**Accuracy** is the ratio of the total number of instances of the correct prediction. Accuracy calculated as follows-

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

**Sensitivity** is used to determine the portion of the actual positive instances case classified adequately by the classifier. Sensitivity calculated as follows-

$$Sensitivity = \frac{TP}{TP + FN} \quad (2)$$

**Specificity** is used to know the ability of classifiers to identify incorrectly classified negative cases.

$$Specificity = \frac{TN}{TN + FP} \quad (3)$$

**Precision** It is an indicator that defines the true portion of the instances when predicted to be true. Precision calculated as follows-

$$Precision = \frac{TP}{TP + FP} \quad (4)$$

**F1 Score** is a harmonic mean of recall and precision. It must be one for good performance and zero for the

bad performance of the classification algorithm. F1 score calculated as follows-

$$F1Score = \frac{2 * Precision * Recall}{Precision + Recall} \quad (5)$$

**Area Under Curve** Area under ROC curve is another important metrics to evaluate the performance of machine learning algorithm. AUC near to 1 indicates the perfect performance and near to 0.5 indicates worst performance of the machine learning model[15].

### III. RESULTS AND DISCUSSION

#### A. Statistical Results

1) *Time-domain HRV Parameters:* The time-domain HRV parameters *mean* RR, SDNN, and RMSSD were significantly reduced in the obese group. Reduced *mean* RR and SDNN indicate that the RR interval time series signal variability is reduced, and total variance is also reduced. The lower value of RMSSD represents parasympathetic activity(Table II).

TABLE II  
TIME-DOMAIN HRV PARAMETERS

Time-domain HRV Features	Healthy Group (n=48)	Obese Group (n=48)	P-Value
<i>mean</i> HR	67.33 ± 9.82	72.10 ± 10.51	0.0237
<i>mean</i> RR	855.17 ± 124.75	798.69 ± 116.47	0.0241
SDNN	57.42 ± 8.49	49.90 ± 7.50	0.0001
RMSSD	47.44 ± 7.25	40.23 ± 6.11	0.0001

2) *Frequency-domain HRV Parameters:* In the frequency domain, TP( $ms^2$ ), HF( $ms^2$ ), HF( $nu$ ), and LF: HF ratio were significantly decreased whereas LF( $ms^2$ ) and LF( $nu$ ) were comparable in obese as compared to control.

TABLE III  
FREQUENCY-DOMAIN HRV PARAMETERS

Frequency-domain HRV Features	Healthy Group (n=48)	Obese Group (n=48)	P-Value
TP ( $ms^2$ )	3265.70 ± 537.58	2543.04 ± 451.68	0.0001
LF ( $ms^2$ )	948.12 ± 156.55	926.26 ± 171.12	0.5154
HF( $ms^2$ )	1011.06 ± 171.96	637.48 ± 111.17	0.0001
LF ( $nu$ )	48.07 ± 7.16	57.11 ± 8.36	0.0001
HF ( $nu$ )	49.94 ± 7.42	40.88 ± 6.07	0.0001
LF:HF	1.59 ± 0.26	1.08 ± 0.16	0.0001

Reduced values of TP( $ms^2$ ) indicate the less variance . A lower value of HF( $ms^2$ ) indicates the declined parasympathetic tone. Less value of LF: HF ratio indicates the autonomic imbalance(Table III).

3) *Non-linear HRV Parameters*: The non-linear HRV parameters SD1 and SD2 were analyzed and found less in obese compared to control. The SD1 feature value was significantly less in obese that indicates the reduced short-term variability in the HRV signal(Table IV).

TABLE IV  
NON-LINEAR HRV PARAMETERS

Non-linear HRV Features	Healthy Group (n=48)	Obese Group (n=48)	P-Value
SD1	33.72 ± 5.07	28.11 ± 4.30	0.0001
SD2	66.00 ± 9.62	65.83 ± 9.70	0.9303

### B. Machine Learning Results

The ML algorithms are used to find the most important predictor that separates obese subjects from the control. However, in the statistical analysis, it was observed that most of the time domain, frequency domain, and non-linear HRV parameters are significantly reduced but do not give an important predictor. The important predictor can be found out using the feature importance technique, which provides a feature importance score to each feature. The feature importance score greater than 0.90 or 90%, indicates the most important predictor. In this study, we found that *mean RR*, LF: HF ratio and HF( $ms^2$ ) was the most important predictor. We have obtained *mean RR* and LF: HF important predictor using the CART algorithm, whereas HF( $ms^2$ ) was the important predictor obtained using the GBDT. We have used only these predictors as input to the CART and GBDT ML algorithm. When we have applied the *mean RR* and LF: HF ratio as input to the CART algorithm, we got an accuracy of 96.55%, a sensitivity of 100%, a specificity of 92.86%, precision of 93.75%, F1 score of 0.96 with an AUC of 0.96. When we have used HF( $ms^2$ ) as input to the GBDT algorithm, we observed accuracy of 93.10%, the sensitivity of 93.33%, the specificity of 92.86%, precision of 93.33%, F1 score of 0.93 with an AUC of 0.92. The important predictor indicates that the CART and GBDT ML algorithm can classify the obese and control subjects with an accuracy of 96.55% and 93.10%, respectively(Table V). This suggests that obesity modulates the function of ANS, as *mean RR*, HF( $ms^2$ ), and LF: HF ratio are reduced in obese subjects. Thus changes in ANS modulates cardiac activity.

## IV. CONCLUSION

In the present study, we have used short-term HRV analysis of obese and control subjects to study the impact of obesity

TABLE V  
EVALUATION OF IMPORTANT HRV PREDICTOR

Algorithm	Important HRV Predictor	Feature Importance Score	AC (%)	SE (%)	SP (%)	PR (%)	F1 Score	AUC
CART	<i>mean RR</i>	0.93	96.55	100	92.86	93.75	0.96	0.96
	LF/HF	0.94						
GBDT	HF( $ms^2$ )	0.95	93.10	93.33	92.86	93.09	0.93	0.92

AC-Accuracy, SE-Sensitivity, SP-Specificity, PR-Precision.

on ANS. We have used real and synthetic HRV data for analysis. The results of the study are presented using statistical tests and a machine learning algorithm. The statistical analysis shows the significant reduction in HRV parameters of obese subjects compared to control and machine learning algorithm was used to found important HRV predictor. The statistical results suggests an alteration in sympathovagal balance due to less parasympathetic activity. Further, this was confirmed using the CART and GBDT algorithm, which showed a classification accuracy of 96.55% and 93.10%, respectively.

#### REFERENCES

- [1] Hurt RT, Kulisek C, Buchanan LA, McClave SA. The obesity epidemic: challenges, health initiatives, and implications for gastroenterologists. *Gastroenterol Hepatol (N Y)*. 2010;6(12):780792.
- [2] Thayer JF, Yamamoto SS, Brosschot JF. The relationship of autonomic imbalance, heart rate variability and cardiovascular disease risk factors. *Int J Cardiol*. 2009; 141(2):122131.
- [3] Rossi RC, Vanderlei LC, Gonalves AC, Vanderlei FM, Bernardo AF, Yamada KM, da Silva NT, de Abreu LC. Impact of obesity on autonomic modulation, heart rate and blood pressure in obese young people. *Autonomic neuroscience*. 2015;193:138-41.
- [4] Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: synthetic minority over-sampling technique. *Journal of artificial intelligence research*. 2002 ;16:321-57.
- [5] World Health Organization: obesity and overweight. Fact sheet No. 311, [updated June 2016]. Available from: <http://www.who.int/mediacentre/factsheets/fs311/en/>. Accessed December 13, 2016.
- [6] Romero-Corral A, Somers VK, Sierra-Johnson J, et al. Accuracy of body mass index to diagnose obesity in the US adult population. *Int J Obes (Lond)*. 2008;32(6):959966.
- [7] Heart rate variability: standards of measurement, physiological interpretation and clinical use. task force of the European society of cardiology and the North American Society of Pacing and Electrophysiology. *Circulation*. 1996;93(5):10431065.
- [8] Vanderlei LCM, Pastre CM, Hoshi RA, Carvalho TD, Godoy MF. Basic notions of heart rate variability and its clinical applicability. *Rev Bras Cir Cardiovasc*. 2009;24(2):205217.
- [9] Laederach-Hofmann K, Mussgay L, Ruddle H. Autonomic cardiovascular regulation in obesity. *J Endocrinol*. 2000;164(1):5966.
- [10] Billman GE, Huikuri HV, Sacha J, Trimmel K. An introduction to heart rate variability: methodological considerations and clinical applications. *Front Physiol*. 2015;6:55.
- [11] Billman GE. The effect of heart rate on the heart rate variability response to autonomic interventions. *Frontiers in physiology*. 2013;4:222.
- [12] Shenoy AR, Doreswamy V, Shenoy JP, Prakash VS. Impact of obesity on cardiac autonomic functions in middle aged males. *National Journal of Physiology, Pharmacy and Pharmacology*. 2014;4(3):236-9.
- [13] Breiman L, Friedman J, Stone CJ, Olshen RA. *Classification and regression trees*. CRC press; 1984.
- [14] Natekin A, Knoll A. Gradient boosting machines, a tutorial. *Frontiers in neuroinformatics*. 2013;7:21.
- [15] Hajian-Tilaki K. Receiver operating characteristic (ROC) curve analysis for medical diagnostic test evaluation. *Caspian Journal of internal medicine*. 2013;4(2):627-635.



# Image Character Recognition using Convolutional Neural Networks

Adith Narayan\*, Raja Muthalagu\*\*

\* *Department of Computer Science, Birla Institute of Technology and Science Pilani,*  
Dubai Campus, Dubai, UAE  
f20170044@dubai.bits-pilani.ac.in

\*\* *Department of Computer Science, Birla Institute of Technology and Science Pilani,*  
Dubai Campus, Dubai, UAE  
raja.m@dubai.bits-pilani.ac.in

**Abstract**—This paper aims the study and implementation of Convolutional Neural Network (CNN) for Image character recognition. Handwritten Character recognition involves recognition of texts present in digital images and documents and processing them for various applications such as machine translation, pattern recognition and so on. This paper studies the use of CNN in detecting and recognizing handwritten text images with a higher accuracy. The CNN model is tested on English handwritten characters and validated on its performance. The model performs feature extraction from images through multiple layers. These are later used for training the model and thereby recognizing characters.

**Keywords**- Convolutional neural network, character recognition, deep learning

## I. INTRODUCTION

Optical Character Recognition refers to recognition of text from images, documents and other devices. Handwritten character recognition also involves optical recognition but of handwritten characters from various sources. OCR has a wide range of applications over various domains such as text mining and machine translation. Some of the most common approaches used for recognition are statistical methods and neural networks. In this paper, the neural network approach for character recognition is discussed in detail along with their results.

### A. Neural Network

Deep Learning forms a subset of Machine learning and Neural Network acts as a backbone to the deep learning methods and algorithms. Deep learning signifies the depth (or number of layers) in a neural network. Neural networks, and in specific

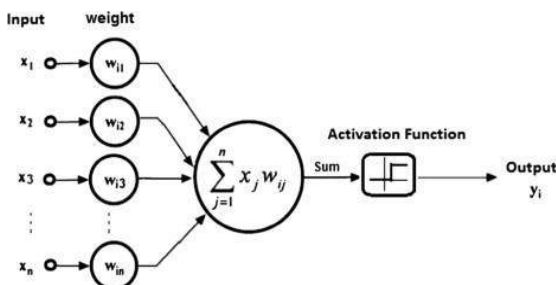


Fig. 1. Basic Neural Network

artificial neural networks, are information processing systems which mimic the human brain. The information processing elements called neurons form the basis of Neural Networks that draws parallels from the biological neural networks. The principal components of a neural network are- inputs, weights associated

with the connection links, bias and output. Each node in a neural network is called as a perceptron.

Convolutional Neural Network (CNN) is a type of deep neural network which is used for visual image analysis. It is a multilayered perceptron network that is fully connected i.e. each neuron in the layer is fully connected to all the neurons in the following layer. It uses a three dimensional structure where a group of neurons analyze a region of the image known as feature. It works primarily in three stages.

The first stage is the convolution layer in which pixels of the image is scanned to learn the features of the input and a feature map is produced as the output. The second stage is Pooling which involves Dimensionality Reduction of the feature space along with retaining essential information. The third stage involves flattening matrix to a vector that are fed into a fully connected layer that combines features into a model and decides the class to which the image belongs to.

## II. LITERATURE SURVEY

CNN model has been guaranteed to improve the performance of character detection. Saidane [1] proposed a CNN model for character detection and proved to outperform the traditional recognition methods using binary characters. The model achieved better results. Jacobs et al. [2] proposed a recognizer system using CNN for detection on grayscale images. The model outperformed the OCR software on documents with low resolution. The CNN model performed better in the detection of Chinese characters which was mainly due to absence of large public datasets for Chinese characters.

In [3], English and Kannada datasets were used for feature extraction such as edge methods, texture representation and shape to evaluate their common parametric values. It was concluded that on blur feature descriptors and shape context, SVM classifier did not perform as well as Nearest neighbor classifier.

Jain et al. [4] implemented a CNN-RNN hybrid network model for Arabic text recognition in natural scene images and videos. A synthetic dataset for natural scenes was created for training the network. The model attained an accuracy of 75.05% on the synthetic dataset and 98.17% on the video text. In [5], HoG (Histogram of Oriented Gradients) feature extraction was performed along with two new

descriptors Co-occurrence and Convolutional Co HoG for character detection in natural images. The Co-occurrence HoG extracts co-occurrence of gradient pairs of the pixels along with greater contextual and spatial information. Convolutional Co HoG captures co-occurrence from all patches of the image for greater spatial knowledge. These were evaluated on three dataset characters apart from Bengali and Chinese.

Kale et al. [6] used a feature extraction approach with the help of feed-forward Artificial Neural Network and Support Vector Machine (SVM). They claimed an accuracy of 90% using a rule based classifier. The increase in the size of the vocabulary contributed to a decrease in the performance of the methodology. In [7], two feed-forward Neural Networks were used along with gradient and density features for the detection of Malayalam characters. LeCun et al. [8] was the first to propose a Convolutional Neural Network, which was one among the most successful models as it considers the spatial arrangement of an image as well.

In [9], three CNN networks were used for separate feature extraction from scene images. They were then combined into a single feature map for classification. ElAdel et al. [10] implemented a Neural Network architecture with FWT (Fast Wavelet Transform) using Adaboost and MRA(Multi Resolution Analysis) algorithms. Arabic handwriting recognition was done using IESK-arDB dataset with an accuracy of 93.92%.

### III. DESIGN AND IMPLEMENTATION

#### A. Dataset

The dataset that we have used is the Chars74K dataset of English language. It contains 26,416 images of English alphabets with 1016 images belonging to each class to be trained i.e. each English letter. A sample from the character dataset is shown in Fig. 2.

#### B. Splitting Dataset for Training, Testing and Validation

The image dataset is split into training, testing and validation sets. An optimal ratio of 80:20 is used for splitting of training and testing datasets. Further, a split up of 80:20 ratio is employed for training and validation datasets. After the split, training dataset contains 16905 images with 5284 and 4227 images for testing and validation.

#### C. Preprocessing

Image Preprocessing refers to techniques that account for distortions and enhance the features of the image, allowing efficient performance on the model.



Fig. 2. A sample character of the dataset

Some of the important preprocessing techniques are conversion to grayscale, even distribution of contrast and intensity, normalization of pixel values to 0-1 range for greater convergence.

Image augmentation is also performed which is responsible for producing transformations of the image. The transformations applied are width and height shifts, and shears performed by a factor of 10%, zooms and random rotations by factors of 20% and 10 degrees respectively.

#### D. CNN Model

A CNN model is created where layers are sequentially

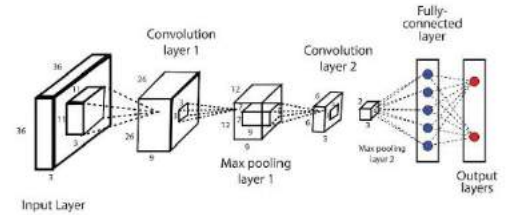


Fig. 3. Structure of CNN

stacked from the input layer, all through the intermediate layers to the output layer. The layers in the model are: Convolutional layer, Pooling layer, Dense or Fully Connected layer. Each of their structure and functions are discussed below.

#### E. Convolutional Layer

Conv2D is the first two dimensional and convolutional layer which transforms input into an output tensor. The convolution is performed by filters that stride over the input image and high level features are extracted from them producing an output feature map. The size of the filters are 5x5 and 3x3.

#### F. Pooling Layer

The next layer is the Pooling layer which is responsible for dimensionality reduction by reducing the spatial size of the features. Relu (Rectified linear unit) is an activation function that outputs the input if its positive, else 0.

The activation function makes the output of the convolutional layer non-linear. Here we have used Max pooling which returns the maximum aspect of the features (most dominant) covered by the image kernel. Max pooling provides much better performance than average pooling which computes average of the values of the portion covered by image kernel. Fig. 4 portrays max and average pooling on the kernel.

#### G. Dropout

Dropout is a regularization method that is applied here which ignores certain set of random input neurons during training phase. This helps to reduce overfitting of the model.

#### H. Flattening

Flattening converts the input into 1-D feature vector which are fed into the final layer of the CNN called the Dense or fully-connected layer.

#### I. Dense Layer

The last layer of the model is the Dense layer which aggregates the feature vectors learnt and predicts the class of the image. Adam Optimization algorithm is used here which is an adaptive learning algorithm that uses specific learning rates for for each of the parameters. The image is classified using Softmax technique.

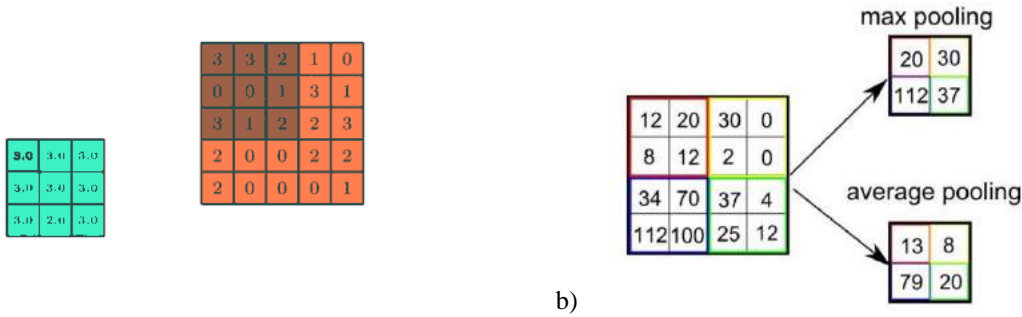


Fig. 4. a) 3x3 and 5x5 pooling and b) Max and average pooling

**J. Training The Model**

The CNN model is trained for 10 epochs with 2000 steps per epoch. The batch size of the training samples is 50. The Test loss and Test accuracy is shown in Fig. 5. The plots for Accuracy and Loss for the model is displayed in Fig. 6.

**K. Real Time Recognition**

The CNN model implemented also performed real-time recognition of English characters, both handwritten and text images. The maximum probability of recognition of a character was found to be 98% for real-time detection of handwritten image.

**L. Contour Detection**

The detection of individual characters from images of words have been implemented in this model using contours. This is done with the help of Tesseract optical character recognition engine.

Contours help in estimating the coordinates of each of the characters and place a bounding box over it. This helps in successful detection of each character in the image. This has successfully detected characters of real-time handwritten images in addition to natural scene images.

**IV. RESULTS AND CONCLUSION**

This paper discusses the implementation of Convolutional Neural Network Model to recognize English handwritten characters along with text images. The model has been successfully able to recognize characters in real-time which further increases its scope. Preprocessing is an important factor in ensuring high performance of the model. Image preprocessing techniques enhance the features of the image, thereby increasing the accuracy of recognition. The recognition is shown in Fig.7. The model attained a high recognition accuracy of 97.59% with a loss of 6.6%. Real-time detection of characters is also implemented using contour based techniques as shown in Fig. 8 . The accuracy rate with average pooling technique has resulted in marginal decrease of accuracy and marginal increase in loss as depicted in Fig.9 . It could be concluded that deep learning CNN model is an efficient character recognition model which yields high accuracies in recognition and has established its performance for handwritten characters as well.

**V. FUTURE SCOPE**

The performance of the model in the recognition of characters can further be enhanced with the help of large datasets for training.

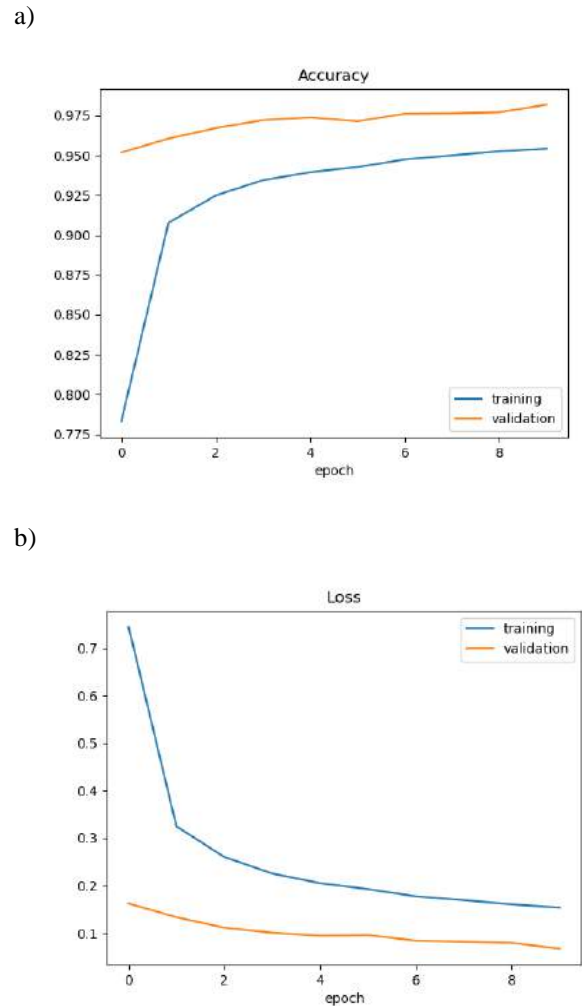


Fig. 6. a) Accuracy plot and b) Loss plot

Further character recognition of natural scene and handwritten images have had limited successes due to the difficulty in contour based techniques, especially when applied on unevenly distributed images. This requires to be explored and addressed with implementation of efficient contour detection methods.

This paper does not compare the results with recognition using state-of-art methods. Efficiency of the same compared to neural networks could be explored in the future.

```

1991/2000 [=====] - ETA: 3s - loss: 0.1663 - accuracy: 0.9515
1992/2000 [=====] - ETA: 3s - loss: 0.1664 - accuracy: 0.9515
1993/2000 [=====] - ETA: 3s - loss: 0.1663 - accuracy: 0.9515
1994/2000 [=====] - ETA: 2s - loss: 0.1663 - accuracy: 0.9515
1995/2000 [=====] - ETA: 2s - loss: 0.1663 - accuracy: 0.9516
1996/2000 [=====] - ETA: 1s - loss: 0.1663 - accuracy: 0.9516
1997/2000 [=====] - ETA: 1s - loss: 0.1663 - accuracy: 0.9515
1998/2000 [=====] - ETA: 0s - loss: 0.1663 - accuracy: 0.9516
1999/2000 [=====] - ETA: 0s - loss: 0.1663 - accuracy: 0.9516
2000/2000 [=====] - 885s 443ms/step - loss: 0.1663 - accuracy: 0.9516
Test Score = 0.86699541571431383
Test Accuracy = 0.9793716669002642
Process finished with exit code 0
    
```

Fig. 5. Test score and accuracy



```

1991/2000 [=====] - ETA: 3s - loss: 0.1698 - accuracy: 0.9502
1992/2000 [=====] - ETA: 3s - loss: 0.1698 - accuracy: 0.9502
1993/2000 [=====] - ETA: 2s - loss: 0.1698 - accuracy: 0.9502
1994/2000 [=====] - ETA: 2s - loss: 0.1697 - accuracy: 0.9502
1995/2000 [=====] - ETA: 2s - loss: 0.1697 - accuracy: 0.9502
1996/2000 [=====] - ETA: 1s - loss: 0.1697 - accuracy: 0.9502
1997/2000 [=====] - ETA: 1s - loss: 0.1697 - accuracy: 0.9502
1998/2000 [=====] - ETA: 0s - loss: 0.1697 - accuracy: 0.9502
1999/2000 [=====] - ETA: 0s - loss: 0.1698 - accuracy: 0.9502
2000/2000 [=====] - 863s 432ms/step - loss: 0.1697 - accuracy: 0.9502
    
```

Fig. 9. Accuracy and loss with average pooling

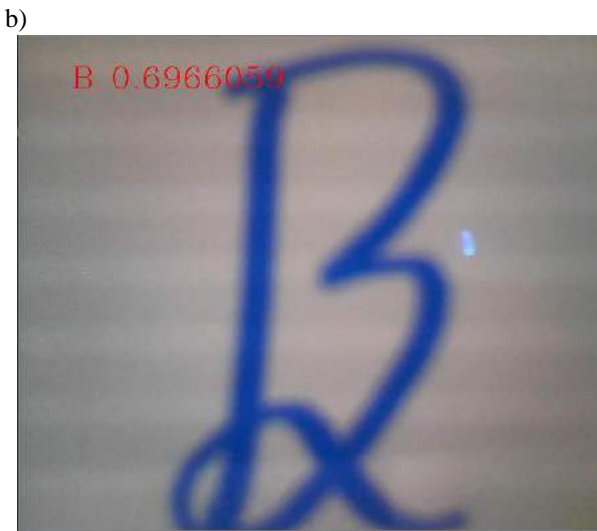


Fig. 7. a) and b) Real-time recognition of handwritten characters

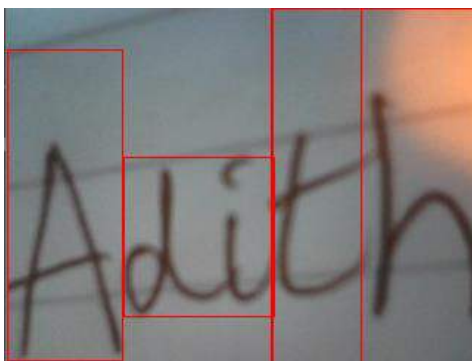


Fig. 8. Real-time Contour detection of handwritten word

## VI. REFERENCES

- [1] Zohra Saidane and Christophe Garcia, "Automatic scene text recognition using convolutional neural network," in Proceedings of the Second International Workshop on Camera-Based Document Analysis and Recognition (CBDAR), 2007.
- [2] Charles Jacobs, Patrice Y Simard, Paul Viola, and James Rinker, "Text recognition of low-resolution document images," in Document Analysis and Recognition, 2005. Proceedings. Eighth International Conference on. IEEE, 2005, pp. 695–699.
- [3] T. D. Campos, B. R. Babu, and M. Varma, "Character recognition in natural images," in Proc. Int. Conf. Comput. Vis. Theory Appl., 2009, p. 273\_280.
- [4] Jain, M. Mathew, and C. V. Jawahar, "Unconstrained scene text and video text recognition for arabic script," in Proc. 1st Int. Workshop Arabic Script Anal. Recognit. (ASAR), Nancy, France, Apr. 2017, pp. 26\_30.
- [5] S. Tian, U. Bhattacharya, S. Lu, B. Su, Q. Wang, X. Wei, Y. Lu, and C. L. Tan, "Multilingual scene character recognition with co-occurrence of histogram of oriented gradients," Pattern Recognit., vol. 51, pp. 125\_134, Mar. 2016.
- [6] Kale, Karbhari V., Prapti D. Deshmukh, Shrinivas V. Chavan, Majharoddin M. Kazi, and Yogesh S. Rode. "Zernike moment feature extraction for handwritten Devanagari compound character recognition." In 2013 Science and Information Conference, pp. 459-466. IEEE, 2013.
- [7] Chacko, Anitha Mary MO, and P. M. Dhanya. "Multiple classifier system for offline malayalam character recognition." Procedia Computer Science 46, pp. 86-92. Elsevier, 2015.
- [8] Y. LeCun, B. Boser, J.S. Denker, D. Henderson, R.E. Howard, W. Hubbard and L.D. Jackel, "Backpropagation applied to handwritten ZIP code recognition," Neural Computations, vol. 1, issue 4, pp. 541-551, 1989.
- [9] W. Xue, X. Dai, and L. Liu, "Remote sensing scene classification based on multi-structure deep features fusion," IEEE Access, vol. 8, pp. 28746\_28755, 2020.

- [10] ElAdel, R. Ejbali, M. Zaied, and C.B. Amar, Dyadic Multi-resolution Analysis-Based Deep Learning for Arabic Handwritten Character Classification, Proc. Tools with Artificial Intelligence (ICTAI), 2015 IEEE 27th International Conference on , 2015, pp. 807-812.
- [11] Yuanping Zhu, Jun Sun, and Satoshi Naoi, "Recognizing natural scene characters by convolutional neural network and bimodal image enhancement," in Camera-Based Document Analysis and Recognition, pp. 69–82. Springer, 2012.
- [12] Chirag I Patel, Ripal Patel, Palak Patel, "Handwritten character recognition using neural network",International Journal of Scientific and Research Volume 2,Issue may 2011".
- [13] Batuhan Balci, Dan Saadati, Dan Shiferaw, "Handwritten Text Recognition using Deep learning".
- [14] Nafiz Arica, and Fatos T. Yarman-Vural, —Optical Character Recognition for Cursive Handwriting, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.24, no.6, pp. 801-113, June 2002.
- [15] Husnain M, Missen MMS, Mumtaz S, Jhanidr MZ, Coustaty M, Luqman MM, et al. Recognition of Urdu Handwritten Characters Using Convolutional Neural Network. Applied Sciences. 2019;9(13):1–21.
- [16] Pragathi, K. Priyadarshini, S. Saveetha, A. S. Banu and K. O. Mohammed Aarif, "Handwritten Tamil Character Recognition Using Deep Learning," 2019 International Conference on Vision Towards Emerging Trends in Communication and Networking (ViTECoN) Vellore, India, 2019, pp.1-5.
- [17] Jemimah K, "Recognition of Handwritten Characters based on Deep Learning with TensorFlow", Research Scholar, School of Computer Science and Engineering, Bharathidasan University, Trichy, India, International Research Journal of Engineering and Technology (IRJET), 2019, pp 1164-1165.
- [18] Megha Agarwal, Shalika, Vinam Tomar, Priyanka Gupta, "Handwritten Character Recognition using Neural Network and Tensor Flow", Computer Science and Engineering, SRM IST Ghaziabad, India, International Journal of Innovative Technology and Exploring Engineering (IJITEE), 2019, pp 1445.
- [19] D. S. Prashanth, R. V. K. Mehta and N. Sharma, "Classification of Handwritten Devanagari Number - An analysis of Pattern Recognition Tool using Neural Network and CNN", *Procedia Computer Science*, vol. 167, pp. 2445-2457, 2020.

**2021 IEEE Seventh  
International Conference on  
Bio signals, Images and  
Instrumentation  
ICBSII 2021**

**SESSION IV  
RESEARCH  
PAPERS**

# Comparison of Watershed and Local Center of Mass Techniques in Segmentation of Breast Mammogram

Mohamed Yacin Sikkandar  
 Department of Medical Equipment Technology  
 College of Applied Medical Sciences, Majmaah University  
 Al Majmaah, 11952, Saudi Arabia  
 m.sikkandar@mu.edu.sa

**Abstract**— Tumor interpretation on cancer image can lead to early detection of cancer with the prevalence of breast cancer being the second leading cause for death in women. Segmentation of mammographic images is a challenging space because of the complexity in extracting data without causing any source of artefacts on the image. Diagnosis becomes difficult when data extraction is challenging. In order to cater effective diagnosis and effective treatment, segmentation is a vital process. This paper discusses about two segmentation method, Watershed and Local Center of Masses. Comparison between these two algorithms based on the amount of data extraction for six different categories of mammographic images, the apt segmentation method for data extraction is found out. Watershed provides an average dice score of 0.53, and occupies 665 KB memory space and avails 5 seconds running time whereas LCM avails on an average of 0.59 dice score, 250 KB memory space and 1008 seconds computing time. This can pave way for selection of effective diagnosis that can result in early diagnosis in turn leading in faster cure.

**Keywords**—LCM, Watershed, Breast Cancer, Segmentation

## I. INTRODUCTION

Cancer refers to the rapid multiplication of cells that grow out of control and spread neighbouring tissues and mutation occurs at a faster pace in the cell DNA. Cancer cells can originate at any part in the human body which are made of more than trillion of cells resulting in the breakdown of an orderly process of normal cell origination. When extra cells originate in an uncontrolled manner it results in formation of tumor cells. The occurrence of cancer cells depend on multiple factors such as age, gender, diet and genetics.

Depending on the region they originate, cancer takes specific name with breast cancer being the commonly occurring cancer in women population with over more than 2 million cases yearly as per the reports given by World Health Organization (WHO). With various screening methods to detect and diagnose cancer tissue and its growth, mammography stands as a dedicated radiographic technique for breast tissue imaging. With rapid development in image processing techniques, analysis of cancer images have had a very big impact on the diagnosis and treatment procedure. Advent of image processing has resulted in many novel techniques for precise monitoring of breast tissue tumor. Extensive research on mammography images based on image processing has been employed to extract multiple features and parameters. Image processing techniques applied on mammographic images from which mathematical values were extracted to access whether the image is benign

or malignant showed potential results [1]. Comparing mammography breast tissue images with thermal texture mapping indicated pronounced variation 63% to 88% when analysing benign and malignant cells. Most Mammographic images employed shearlet transform to differentiate between benign and malignant images [2, 3]. The extracted features were reduced by kernel principal component analysis and were ordered based on T value and a k-nearest neighbour classifier was used to classify the images which showed a specificity of 93.8% [3]. Classifying micro calcifications on digital mammography using morphological descriptors and artificial neural network focused on monitoring tiny calcium micro calcifications that can possibly develop into a cancerous tumor thereby working on early stages of cancer formation showing output results with an accuracy of 80%, where they worked on Breast Imaging-Reporting and Data System (BI-RADS) stage 2,3 and 4 categories, on which they primarily focused on the advancement of cancer tumour cell growth [4]. Automated mass segmentation in mammography whose work performed localization of lesion in mammogram images using deep learning methods, pixel wise segmentation was performed which differentiated between benign and malignant tumor cells helping radiologist[5]. But the major drawback was the need of millions of image data sets for training the system. Detection and analysis of breast masses obtained from MRI images and dual energy contrast enhanced digital mammography images from which texture and shape features were extracted .Obtained results were processed on to improve the detection of breast cancer showed significant characterization of tumor cells [6].

Feature based registration applied on infrared imaging focusses on an alternative technique for breast cancer screening by extracting feature based approach for screening cancerous tumor[7]. Feature extraction in mammographic images derive a range of values depending on the extent of growth of the tumour cell [8]. Gravitational search algorithm were used to detect breast cancer tumour particles which employs KNN classification method to classify various stages of tumour [9]. Classification of mammographic masses based on gradient and texture analysis by employing Mahalanobis distances and discriminating as benign or malignant have given significant results[10]. Method to improve detection of calcifications by employing adaptive variance stabilization focusses on intensity dependent quantum noise on the input digital mammographic images. The performance of the transformed mammogram was much better than the raw input data and the results were superior [11]. Computer aided detection and a diagnostic system for monitoring breast cancer from mammographic images mainly focusing on anomaly detection without any

requirement of image segmentation [12]. Existing literature have applied various image processing techniques to extract information about the cancerous tissue but had drawbacks like reduction in spatial resolution, failing to retain the exact edges present in the mammographic image, effect of noise while processing images, over segmentation, false boundary retention. Taking into consideration all these factors this paper discusses about segmentation and processing methods applied on the mammographic image database where the above mentioned failures are eradicated and the proposed system is trained in a manner that it can access the pathology and classify the tumor accurately over the other proposed systems available in the literature.

The proposed method starts with pre-processing and segmentation is done on six different categories of mammographic image using Watershed algorithm and Local Center of Masses. Each of these method is executed to extract detailed information on the intensity distribution with which the exact location and boundaries of the tumor cells can be extracted to analyze which algorithm is best suitable for mammographic image segmentation.

The organization of the paper is as follows: Section II discusses about the methodology involved in preprocessing on the mammographic image by employing watershed algorithm and local center of mass method. Following this Section III discusses about results and discussions and how both the algorithm differ in the process on the image to extract the tumor regions, spatial resolution based texture analysis is done for detailed extraction of data. Section IV discusses about the results and the future work.

## II. METHODOLOGY

### A. Preprocessing

Preprocessing is one of the lowest levels of abstraction performed on input image. Preprocessing primarily aims at improving the quality of image data by suppressing the factors that cause undesired distortions and enhance the features present on the image for effective processing and segmentation. Every input image are processed using watershed algorithm and local center of mass. Preprocessing plays an important role, where the most relevant data from an image are extracted on the mammographic image. The entire work concept where the input mammographic image are preprocessed and analyzed using Watershed algorithm and Local Center of Mass (LCM) methods and the results are compared to conclude which method is most sufficient in interpreting tumor regions on mammographic images.

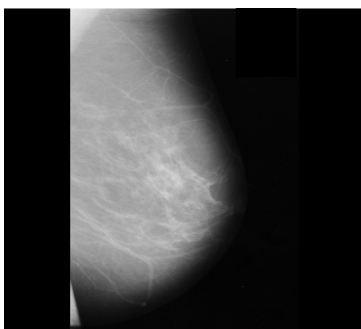


Fig. 1. Preprocessed breast mammographic image

During preprocessing, anisotropic diffusion filter is used to reduce the noise on the input mammographic image without distorting the significant data zones on the image. Edges, lines and significant details that are important for data interpretation are retained while performing this process. Fig.1 indicates the preprocessed image that is used as an input for segmentation using watershed and local center of mass algorithm by employing anisotropic filtering method with the edges on the image preserved and pulmonary bones removed using diagonal edge detection method, which will benefit better interpretation of data in the processes to be done.

### B. Segmentation

Once preprocessing of images are done, images are partitioned into multiple images as sets of pixels to extract in depth data from the mammographic image. Every single pixel on the image are assigned with a specific label and the pixel that carry the same data will be carrying a same label value. Each pixel with a unique label value will have a unique color, intensity and texture when compared with the neighboring pixels.

*Watershed Algorithm:* Image segmentation is a key process for any image recognition system. Watershed algorithm primarily focusses on separating various objects present in the image. With user defined markers every single pixel values are treated as a local topography. The images are read by treating it as a single surface where light pixels are represented by high elevations and dark pixels are represented by low elevations. Each subset pixels are related with respect to their location and intensity and are interpreted in detail for data extraction [13-14].

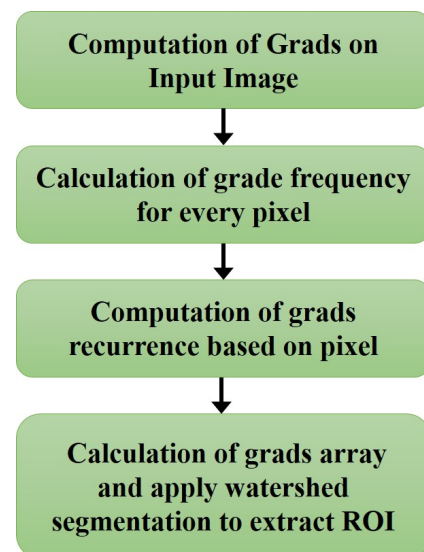


Fig. 2. Watershed algorithm flow

Watershed algorithm primarily focusses on extracting exact boundaries. Watershed algorithm, a region based method for extracting data from images which is processed by selecting seed points with respect to every object in the image, considering the background as a separate entity and the tumor region as a separate entity. This significantly removes the presence of regions other than the tumor cells, making the lymph nodes more visible. This feature has made watershed algorithm a very powerful tool in image segmentation applications making it a robust method. Linear



features are extracted while performing watershed algorithm on the preprocessed image which gives us in depth information about the morphological features which helps in retaining the boundaries of the tissue. Fig.2 highlights the flow of input image undergoes while performing segmentation using watershed algorithm.

*Local center of mass segmentation:* Local center of mass is another segmentation method applied on the image, which is an unsupervised medical image segmentation modality which works on computing one dimensional local center of masses. The entire image is analyzed based on pixel based on center of masses containing the pixel. Local center of masses are computed from multiple orientations and then each pixel is estimated based on iteration done and each pixels are updated.

*Computation of 1D Local Center of Mass:* Let  $f: \Omega \rightarrow R$  be a 1D discrete-domain image-intensity signal of length  $N$ , with  $\Omega := \{1, \dots, N\}$ . Pixels of  $f$  are grouped into disjoint regions based on the calculated center of masses of every single pixel putative region, which is termed as local center of mass,  $C: \Omega \rightarrow R$ .  $C_n$  is the center of mass region containing the  $n^{th}$  pixel. Local center of mass is calculated as:

$$C_n = \frac{\sum_{m=1}^N w_{m,n} m}{\sum_{m=1}^N w_{m,n}} \quad (1)$$

where the  $w_{m,n}$  is the non-negative weighting is computed from  $f$  such that it is large when  $m^{th}$  and  $n^{th}$  are present in the same region and take a lesser value when not present in the same region.

*Computational Complexity reduction:* To have a tractable solution, the value of  $w$  should be in such a way that it solves the purpose of above mentioned pixel grouping purpose and in turn the  $C$  – computational complexity can be reduced

$$W_{m,n} := e^{-|D_m - D_n|} \quad (2)$$

Where  $D_n$  is given as

$$D_n := \alpha \sum_{i=1}^n |f_{i+1} - f_i|^p \quad (3)$$

With values of  $\alpha$  and  $p$  greater than zero, the entire flow carried in LCM segmentation method is shown in Figure 3.

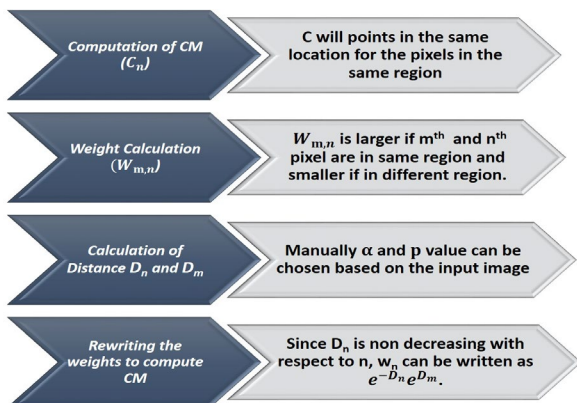


Fig. 3. Algorithm of Local Center of Mass technique

### III. RESULTS AND DISCUSSION

Texture analysis done on the mammographic images where the image is partitioned into various regions of interest

and are classified separately giving us information about the various levels of intensities on the input image. Six different types of cancer images categorised as CALC-Calcification, CIRC-well defined/circumscribed masses, SPIC-Speculated Masses, MISC-III-defined masses, ARCH-Architectural Distortion, ASYM-Asymmetry and Norm-Normal. Statistical methods analyse the spatial distribution of gray values by measuring local characteristics at each point in the picture and deriving statistics from local character distributions and this can be done when ROI is extracted properly. The explanation behind this is the fact that one of the distinguishing characteristics of texture is the spatial distribution of gray values. Texture is considered as quantitative indicator in monitoring the category of tumor cell. The optimised parameter which have been chosen for segmenting breast cancer images shown in Table I give us a proper segmented output only when the values are chosen in proper limits.

TABLE I. OPTIMIZATION PARAMETERS USED

S.No	Parameter	LCM	Watershed Algorithm
1	Dice coefficient	$0.59 \pm 0.02$	$0.53 \pm 0.03$
2	Optimal Parameters	$\alpha^* = 2100 \pm 300$ $t^* = 1900 \pm 500$	$h^* = 0.41 \pm 0.05$
3	Tested Range of Parameters	$\alpha: 1400 - 2400$ $t = 300 - 2000$	$h: 0-1.5$
4	CPU Runtime	$1004 \pm 80$ seconds	$4 \pm 1$ seconds
5	Memory Space (KB)	$250 \pm 30$	$665 \pm 70$

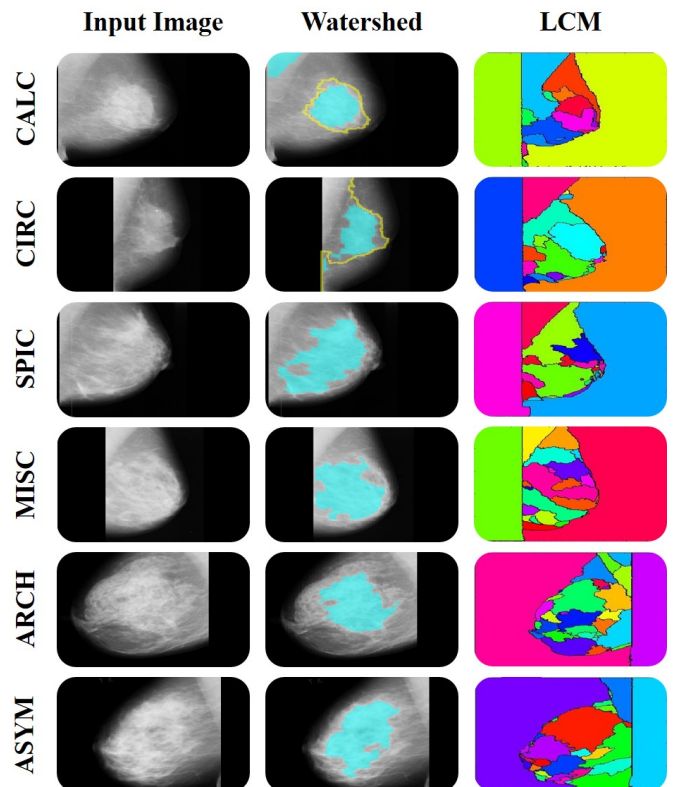


Fig. 4. Chart indicating the outputs obtained through Watershed and LCM Algorithms

The segmentation results done with both LCM and Watershed methods are compared with ground truth image,

which is manually cropped by the clinician and Dice score values are calculated and tabulated in Table I. Segmented image highlighting the tumor region help the clinician to exactly narrow down on the affected part so that the treatment plan can be effective enough to target only the region of interest. The results were compared with the state of art segmentation techniques as discussed by Setarehdan et al., [15]. The results prove that watershed algorithm stands better than LCM method on tracing boundaries of the tumor cells. Fig.4 highlights the output obtained by employing both the algorithm and we can clearly visualize significant differences in the output obtained from both the algorithms, with watershed algorithm being much efficient than LCM algorithm for diagnosing masses and classifying from the six mentioned types of input mammographic image.

**TABLE II. COMPARISON OF WATERSHED AND LCM ALGORITHMS**

S.No	Parameter	Watershed Algorithm	LCM Algorithm
1	Computation Time	Rapid and consumes lesser time	Slow and consumes prolonged time interval for analysis
2	Optimization	Accurate Seed points are chosen depending on GRAD Values	Optimized value of $\alpha$ and $p$ is mandatory for efficient Segmentation
3	Intensity	Output not based on intensity	Gives exact output based on intensity
4	Boundary tracing	Efficient in tracing boundaries of tumor cells	Comparatively less efficient on tracing boundaries
5	Memory	Consumes large space	Consumes less space
6	Execution Time	Fast	Slow

The segmentation results of LCM and Watershed algorithm are compared and is presented Table II, which clearly incites us that Watershed Algorithm is better than LCM method on Mammographic images in order to segment images and classify benign and malignant tumour cell with the normal cells available on the same image.

#### IV. CONCLUSION

The experimental results indicate pronounced differentiation between normal cell and tumor cell without causing any sort of deformation on the original image, retaining the original data on the image intact for any of the six categories of images. The entire flow starting from pre-processing, segmentation and texture analysis have been effective enough to differentiate between benign and malignant tumor cells. At any step in this flow there has not been any degradation in the picture quality nor there any introduction of noise, which is a very unique feature when compared with the existing methods available in literature. With the results obtained from watershed and local center of mass segmentation methods, both the methods are unique in their own manner with deep insights offered by both the methods, with Watershed algorithm having an overhand over Local Center of Mass algorithm. Watershed algorithm was best effective in detecting cancer tissue boundaries and could precisely monitor the intensity distribution while LCM was accurate in extracting the intensity distribution but failed to accurately detect the edges between different regions on the mammographic image. This study further can be extended by

constructing a neural network and the entire process can be automated and trained with large number of mammographic images so that the system becomes robust enough to analyze any sort of mammographic images.

#### REFERENCES

- [1] A. Asaduzzaman, P. Mitra, K. K. Chidella, K. A. Saeed, K. Cluff and M. F. Mridha, "A novel technique to analyze mammography images for breast cancer treatment," 2017 4th International Conference on Advances in Electrical Engineering (ICAEE), Dhaka, 2017, pp. 105-110, doi: 10.1109/ICAEE.2017.8255336.
- [2] Yune Yuan et al., "Analysis of breast diseases examination with thermal texture mapping, mammography and ultrasound," The 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, San Francisco, CA, 2004, pp. 1166-1169, doi: 10.1109/IEMBS.2004.1403373.
- [3] A. M. Ibrahim and B. Baharudin, "Classification of mammogram images using shearlet transform and kernel principal component analysis," 2016 3rd International Conference on Computer and Information Sciences (ICCOINS), Kuala Lumpur, 2016, pp. 340-344, doi: 10.1109/ICCOINS.2016.7783238.
- [4] E. D. Avalos-Rivera and A. De J. Pastrana-Palma, "Classifying microcalcifications on digital mammography using morphological descriptors and artificial neural network," IEEE CACIDI 2016 - IEEE Conference on Computer Sciences, Buenos Aires, Argentina, 2016, pp. 1-4, doi: 10.1109/CACIDI.2016.7785990.
- [5] D. Abdelhafiz, S. Nabavi, R. Ammar, C. Yang and J. Bi, "Convolutional Neural Network for Automated Mass Segmentation in Mammography," 2018 IEEE 8th International Conference on Computational Advances in Bio and Medical Sciences (ICCBS), Las Vegas, NV, 2018, pp. 1-1, doi: 10.1109/ICCBS.2018.8542071.
- [6] S. T. Ben Ameer, L. Wendling and D. Sellami, "Detection and analysis of breast masses from MRIs and dual energy contrast enhanced mammography," 2016 International Image Processing, Applications and Systems (IPAS), Hammamet, 2016, pp. 1-5, doi: 10.1109/IPAS.2016.7880152.
- [7] N. R. Mudigonda, R. M. Rangayyan and J. E. Leo Desautels, "Detection of breast masses in mammograms by density slicing and texture flow-field analysis," in IEEE Transactions on Medical Imaging, vol. 20, no. 12, pp. 1215-1227, Dec. 2001, doi: 10.1109/42.974917.
- [8] M. Lyra, S. Lyra, B. Kostakis, S. Drosos, C. Georgosopoulos and K. Skouroliakou, "Digital mammography texture analysis by computer assisted image processing," 2008 IEEE International Workshop on Imaging Systems and Techniques, Crete, 2008, pp. 73-76, doi: 10.1109/IST.2008.4659944.
- [9] V. Agostini, S. Delsanto, F. Molinari and M. Knaflitz, "Evaluation of feature-based registration in dynamic infrared imaging for breast cancer diagnosis," 2006 International Conference of the IEEE Engineering in Medicine and Biology Society, New York, NY, 2006, pp. 953-956, doi: 10.1109/IEMBS.2006.260808.
- [10] H. Al-Shamlan and A. El-Zaart, "Feature extraction values for breast cancer mammography images," 2010 International Conference on Bioinformatics and Biomedical Technology, Chengdu, 2010, pp. 335-340, doi: 10.1109/ICBBT.2010.5478947.
- [11] F. Shirazi and E. Rashedi, "Feature weighting for cancer tumor detection in mammography images using gravitational search algorithm," 2016 6th International Conference on Computer and Knowledge Engineering (ICCKE), Mashhad, 2016, pp. 310-313, doi: 10.1109/ICCKE.2016.7802158.
- [12] A. Bria et al., "Improving the Automated Detection of Calcifications Using Adaptive Variance Stabilization," in IEEE Transactions on Medical Imaging, vol. 37, no. 8, pp. 1857-1864, Aug. 2018, doi: 10.1109/TMI.2018.2814058.
- [13] Lu, Siyuan, Shuihua Wang, and Yudong Zhang. "A note on the marker-based watershed method for X-ray image segmentation." Computer methods and programs in biomedicine 141 (2017): 1-2.
- [14] Hill, Paul R., Cedric Nishan Canagarajah, and David R. Bull. "Image segmentation using a texture gradient based watershed transform." IEEE Transactions on Image Processing 12.12 (2003): 1618-1633.

- [15] Setarehdan, S. Kamaledin, and Sameer Singh, eds. Advanced algorithmic approaches to medical image segmentation: state-of-the-art applications in cardiology, neurology, mammography and pathology. Springer Science & Business Media, 2001.

# Prediction of cardiac rhythm based on heart rate variability features in sinus tachycardia conditions

B. Dhananjay, J. Sivaraman,\* *Member, IEEE*

Bio-signals and Medical Instrumentation Laboratory

Department of Biotechnology and Medical Engineering

National Institute of Technology Rourkela, Odisha, India

[518bm1004@nitrkl.ac.in](mailto:518bm1004@nitrkl.ac.in); [jsiva@nitrkl.ac.in](mailto:jsiva@nitrkl.ac.in)

**Abstract**— The variability observed in cardiac rhythm measurements makes classification and prediction of Sinus Tachycardia (ST) ECG signals challenging. The study aims to predict ST volunteer's heart rate by designing and developing a Long Short Term Memory (LSTM) based Recurrent Neural Network (RNN) model. The Heart Rate Variability (HRV) features comprising the time and frequency domain were considered as the input feature set for the designed and developed LSTM-RNN model. The proposed LSTM-RNN model works on the principle of multivariate time series forecasting. The LSTM-RNN model's training accuracy is 96.72%, and the validation accuracy is 85.21%. The performance of the LSTM model works well for predicting time series of unknown duration. The time gap in the input time series data does not affect the LSTM model, giving an advantage over other time series prediction models.

**Keywords**— Heart Rate Variability; Long Short Term Memory; Multivariate Time Series Forecasting; Recurrent Neural Network; Sinus Tachycardia

## I. INTRODUCTION

The preliminary change observed during a Sinus Tachycardia (ST) condition is that the heart rate is more than 100 bpm [1]. The listed reasons for ST to occur are physical, emotional, pharmacological, and physiological stresses, and sedentary lifestyle is one of the leading causes of ST [2, 3]. ST is classified into Inappropriate Sinus Tachycardia (IST) and Postural Orthostatic Tachycardia Syndrome (POTS). To date, it is challenging to diagnose, classify and predict ST, IST, and POTS [3, 4]. Wang et al. [5] concluded that predicting POTS using metoprolol is the most efficient as it shows a spontaneous change to a higher heart rate. Kaczmarek et al. [6] studied the prediction response to ivabradine in IST patients and concluded that IST patients spontaneously responded to ivabradine due to pathophysiological changes observed during IST. Jarkovsky et al. [7], in their study, suggested that heart rate is an independent feature to predict acute heart failure among sinus rhythm (SR) volunteers. The studies [5-7] predict ST, IST, and POTS by manually analysing, leading to inter and intra-variability differences. Machine Learning (ML) and Deep Learning (DL) in medicine and research have shown much potential. The models developed by ML and DL automate the tasks performed by humans and have performed better than human capabilities [8].

Yoon et al. [9] stated that tachycardia is an essential marker for cardiorespiratory instability. They trained machine learning classifiers Logistic Regression (LR) and Random Forest (RF) to predict the tachycardia using intensive care unit data. Kim et al. [10], in their work, trained

various machine learning classifiers such as LR, RF, Support Vector Machines (SVM), Extreme Gradient Boosting (XGBoost), Naive Bayes (NB), and Artificial Neural Networks (ANN) to predict post-intubation tachycardia.

Previous studies [8-10] have used machine learning classifiers such as LR, RF, SVM, XG Boost, NB, ANN to predict tachycardia conditions. The LR classifier's major drawback is that it assumes linearity between the dependent and the independent variables. LR classifier overfits when the input features are less than the input observation. The RF classifier's major drawback is that it gives low prediction accuracy for the given input dataset compared with available machine learning algorithms. SVM classifier is not suitable for large datasets. XGBoost is very difficult to interpret, and visualization is challenging. The input parameters to the XGboost classifier need to be appropriately tuned; otherwise will result in overfitting. The hyperparameters present in the XGBoost model are too many, thus making it difficult to tune. XGBoost is suitable for classification rather than prediction. NB is a lousy estimator, as it does not give equal weightage to the entire input feature set.

The proposed work aims to predict ST volunteer's heart rate by designing and developing a Long Short Term Memory (LSTM) based Recurrent Neural Network (RNN) model. The proposed LSTM-RNN model works on the principle of multivariate time series forecasting. The proposed work is different from the existing literature [5-7, 9, 10] as to predict ST, drugs like metoprolol and ivabradine have been injected manually into the patients, and statistical machine learning models (LR, RF, SVM, XGBoost, NB) have been used. LSTM is a deep machine learning model and has advantages over other time series models as the time gap in the input time series data does not affect the model.

## II. METHODOLOGY

### A. Data Collection

The study population considered in the study were from the National Institute of Technology Rourkela and gave informed consent to participate in the study. The ethical clearance committee of the study excluded smokers, alcoholics, hypertension, volunteers at a risk of cardiovascular diseases, and also barred volunteers who were on medication six months prior to the study [11, 22]. The volunteers (n=50) of mean age ( $23 \pm 3.4$ ) years participated. The ECG recording took place in two phases, where the first phase consisted of sinus rhythm (SR) ECG recording, and the second phase consisted of ST ECG recording from the same volunteers. Only in the ST phase of the recording, the volunteers underwent physical stress. In both phases, the ECG signal recording followed standard conventional

\*Corresponding author: Dr. J. Sivaraman, Bio-signals and Medical Instrumentation Laboratory, National Institute of Technology Rourkela

recording. The ECG signal recording was done using EDAN SE 1010 PC ECG instrument. The sampling frequency of the instrument is 1000 samples/second. The study required a six minute ECG signal recording of all the volunteers in both the recording phases.

### B. Feature Set

The input feature set given to the designed and developed LSTM-RNN model consists of Heart Rate Variability (HRV) features. The acquired six minutes ECG signal was subdivided into the first five minutes and last one minute. The first five minutes of the acquired ECG signal from the volunteers in both phases was used to calculate HRV. The last one minute ECG signal obtained from the second phase volunteers was used to predict the heart rate estimated by the designed and developed LSTM-RNN model. Table 1 displays the different HRV parameters considered the input feature set for the designed and developed LSTM model.

### C. Signal Processing

The P and R peak detection was performed on the acquired last one minute ECG signal from the ST phase of recording. To detect peaks of P-wave and R-wave from the ST ECG signal amplitude thresholding technique was considered. Fig. 1 represents the peak detection of the ST ECG signal. The number of samples is represented on the x-axis and microvolts on the y – axis. The sampling frequency is 1000 samples/second. The x-axis shows 1800 to 3000 samples, i.e., 1 sample equals 0.001 seconds (1/fs). The limits of the x-axis are from 1800 ms to 3000 ms (1.8 seconds to 3 seconds).

### D. Statistical Analysis

The HRV features obtained for all the volunteers in both phases were statistically analysed. The student's T-Test was performed to check the statistical significance between the recorded values ( $P < 0.05$ ). All the values are expressed in mean  $\pm$  SD.

### E. LSTM Architecture

The architecture of LSTM consists of three types of gates, namely forget gate, input gate, and output gate. LSTMs are a crucial component in the memory cell as it can sustain a particular state for a long time and has an explicit memory (cell state vector) and different gate units. Gating units help in the exchange of detailed information in the memory. Cell state vectors generally give the memory of the LSTM as it erases the old memory (forget gate) and adds the new input memory (input gate).

Gate is represented by a neural network consisting of sigmoid as the activation function, responsible for the information flow into the memory of the LSTM. The designed LSTM network in this work is stateful as the outputs observed in hidden states and cell states acted as the input to the next iteration of the network. Fig. 2 displays the LSTM architecture. Equation (1) represents the cell state equation of the LSTM cell.

$$\chi_{\tau} = \phi_{\tau} \chi_{\tau-1} + \iota_{\tau} \varpi_{\tau} \quad (1)$$

$\chi_{\tau}$  The current cell state equation  $\phi_{\tau}$  represents the forgotten gate,  $\chi_{\tau-1}$  representing the previous cell state,  $\iota_{\tau}$  representing the input gate,  $\varpi_{\tau}$  representing the input modulation gate.

### F. Recurrent Neural Network design

Recurrent neural networks (RNNs) are a branch of neural networks used to process sequential data. RNNs can also be used to predict the sequence based on the input time series data information. In this work, many to one architecture of RNN was used and training and testing ratio was 70:30 [12].

The designed RNN model has two hidden layers with a dropout of 20%. The first hidden layer consists of sigmoid as the activation function and the second hidden layer consists of the rectilinear unit (ReLU) as the activation function. The optimizer used is the adam, and loss was calculated using mean squared error. Equation (2) represents the sigmoid activation function, where  $S(z)$  represents the sigmoid function,  $e$  represents the Euler's number. Sigmoid is used generally where the model has to predict the probability, as the output of sigmoid ranges between 0 and 1. Equation (3) represents the ReLU activation function. Equation (4) represents the working of adam optimizer,  $\rho$  representing the predicted weight,  $\psi$  representing the optimizer's previous weight,  $\sigma$  representing the step size, and  $\varepsilon$  representing the estimator. The estimator estimates the given input data, calculates the first and second momentum, and adapts to the learning rate.

$$S(z) = (1 + e^{-z})^{-1} \quad (2)$$

$$R(z) = \max(0, z) \quad (3)$$

$$\rho = \psi - \sigma * \varepsilon \quad (4)$$

TABLE 1 HRV PARAMETERS OF STUDY POPULATION

HRV Parameters	units	SR	ST	P-value*
Heart rate	bpm	80 $\pm$ 7.72	104 $\pm$ 9.88	<0.05
Average PP Interval	ms	747 $\pm$ 78.99	577 $\pm$ 51.96	<0.05
SDNN	ms	45 $\pm$ 22.84	29 $\pm$ 18.91	<0.05
RMSSD	ms	39 $\pm$ 16.58	17 $\pm$ 9.88	<0.05
NN50	beats	15 $\pm$ 12.45	4 $\pm$ 4.98	<0.05
pNN50	%	21.2 $\pm$ 17.37	4 $\pm$ 6.18	<0.05
Total Power	ms <sup>2</sup>	6 $\pm$ 5.37	13 $\pm$ 8.52	<0.05
Low Frequency	ms <sup>2</sup>	3 $\pm$ 2.22	4 $\pm$ 2.97	<0.05
High Frequency	ms <sup>2</sup>	1 $\pm$ 1.87	1 $\pm$ 1.29	<0.05
Low Frequency (Normalized)	N.A.	46 $\pm$ 21.57	42 $\pm$ 24.67	<0.05
High Frequency (Normalized)	N.A.	39 $\pm$ 20.75	22 $\pm$ 11.80	<0.05
LF/HF	N.A.	3 $\pm$ 1.85	4 $\pm$ 2.87	<0.05

\*Paired student T-Test

### G. Performance Analysis

Performance analysis metrics consisted of training and validation accuracy. The formulas for the training and validation accuracy are given below:

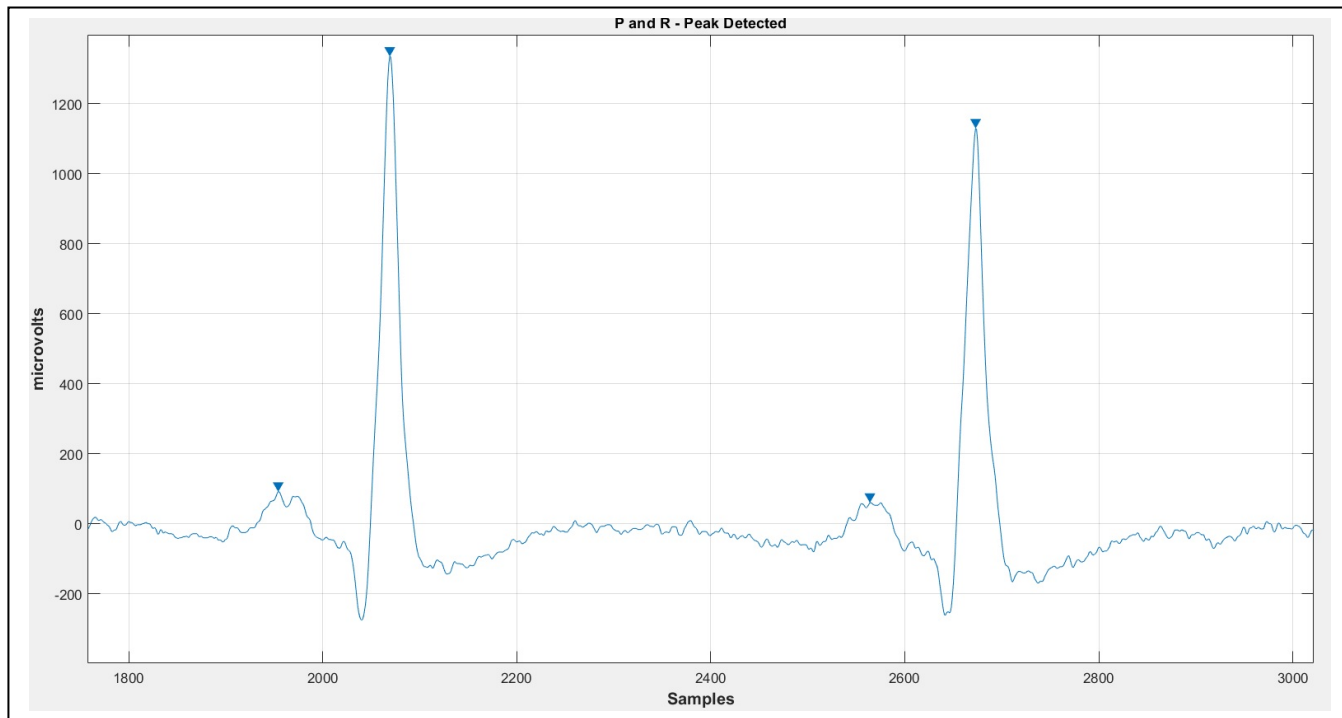


Fig. 1. P and R peak detection performed on ST ECG signal

$$T_A \% = 1 - ((O_S - T_V)/100) \tag{5}$$

$T_A\%$  represents the training accuracy in percentage,  $O_S$  represents the HRV values obtained from the 5 minute ECG recording, and  $T_V$  represents the HRV values in the training set.

$$V_A \% = 1 - ((O_P - V_V)/100) \tag{6}$$

$V_A\%$  represents the validation accuracy in percentage;  $O_P$  represents the RRI values obtained from the 1 minute ECG recording, and  $V_V$  represents the RRI values in the validation set.

(Q1) of the average PPI for SR and ST volunteers is 690 ms and 550 ms, respectively. The median average PPI for SR and ST volunteers is 747 ms and 577 ms, respectively. The third quartile value (Q3) of the average PPI for SR and ST is 790 ms and 610 ms, respectively. The maximum average PPI for SR and ST volunteers is 900 ms and 650ms, respectively.

Fig. 4 shows the standard deviation of the PPI (SDNN) boxplot for SR and ST volunteers. The data classes are represented on the boxplot's x-axis and PPI's standard deviation in ms on the y-axis. The minimum SDNN for SR and ST volunteers is 10 ms and 5 ms. The Q1 SDNN for SR and ST volunteers is 30 ms and 15 ms, respectively. The median of SDNN for SR and ST volunteers is 35 ms and 25 ms, respectively. The Q3 SDNN for SR and ST volunteers is 65 ms and 45 ms, respectively. The maximum SDNN for SR volunteers is 115 ms and 80 ms for ST volunteers

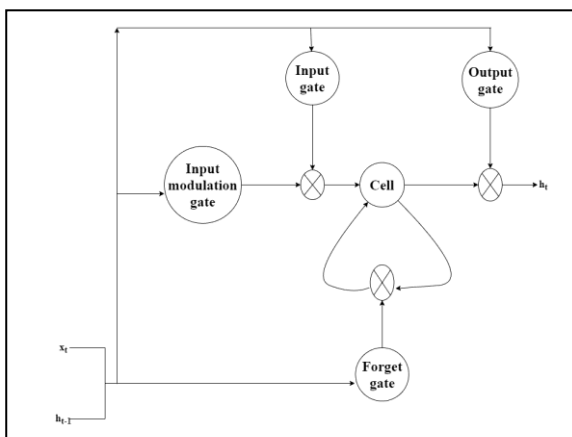


Fig. 2. LSTM architecture

### III. RESULTS

Fig. 3 represents the boxplot of average PP interval (PPI) for SR and ST volunteers. The two data classes are represented on the x-axis and the average PPI (ms) on the y-axis. The minimum average PPI for SR and ST volunteers is 650 ms and 450 ms, respectively. The first quartile value

Fig. 5 displays the root mean square of PPI's successive differences (RMSSD) boxplot for SR and ST volunteers. The data classes are represented on the boxplot's x-axis and RMSSD in ms on the y-axis. The minimum RMSSD for SR and ST volunteers is 10 ms and 5 ms, respectively, and Q1 RMSSD for SR and ST volunteers is 30 ms and 10 ms, respectively, and the median RMSSD value of SR and ST volunteers is 38 ms and 20 ms, respectively. The Q3 RMSSD for SR and ST volunteers is 60 ms and 30 ms, respectively. The maximum RMSSD for SR and ST volunteers is 85 ms and 35 ms, respectively.

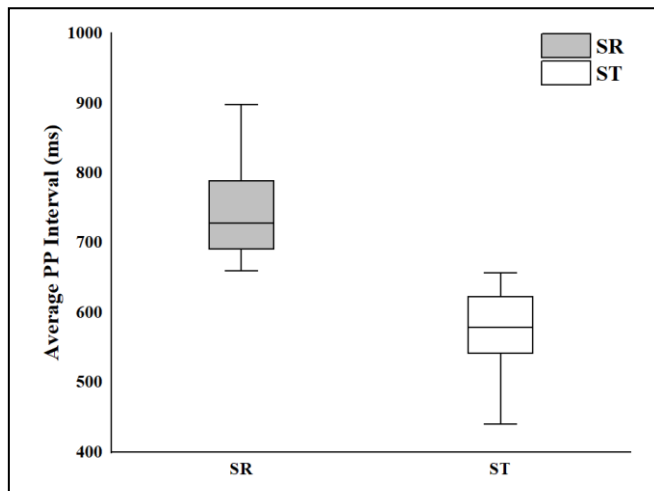


Fig. 3. Boxplot of average PP interval for SR and ST volunteers

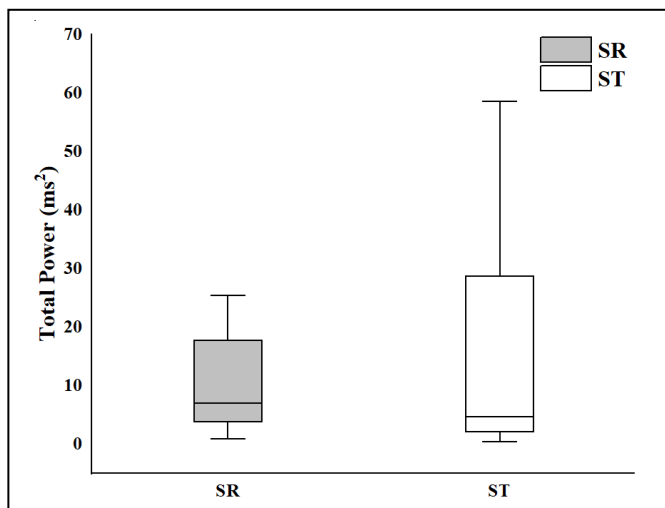


Fig. 6. Boxplot of Total Power for SR and ST volunteers

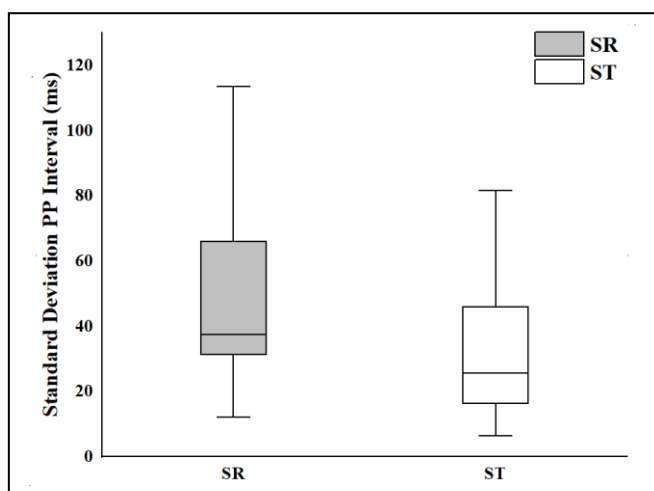


Fig. 4. Boxplot of Standard deviation of PP Interval for SR and ST volunteers

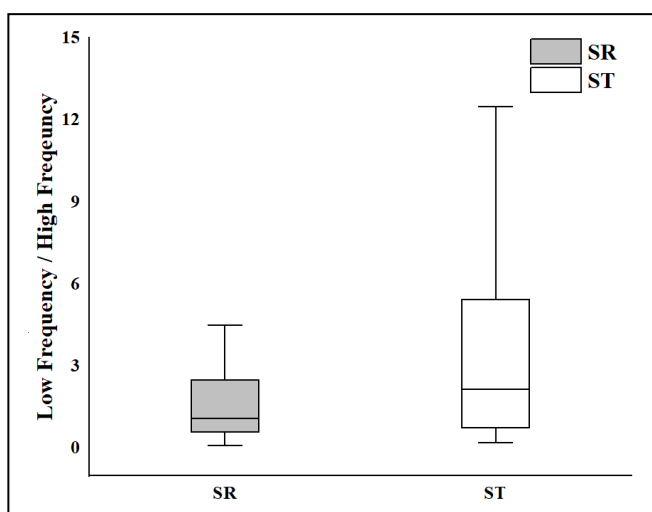


Fig. 7. Boxplot of ratio low frequency to high frequency for SR and ST volunteers

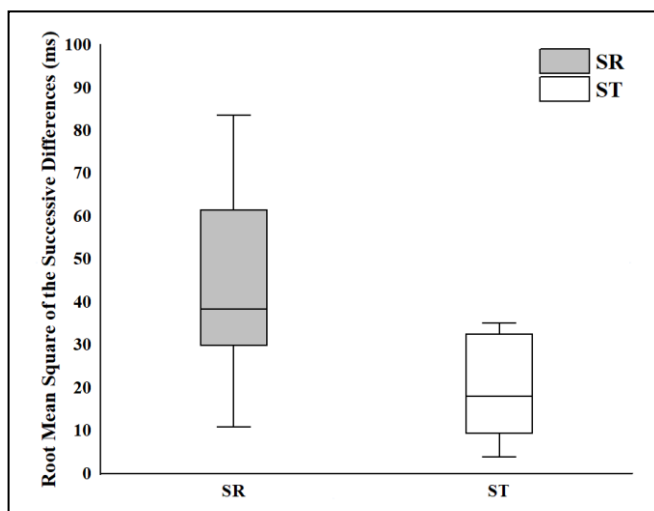


Fig. 5. Boxplot of Root Mean Square of the Successive Differences SR and ST volunteers

Fig. 6 exhibits the total power (TP) boxplot for SR and ST volunteers. The data classes are represented on the boxplot's x-axis and TP in  $ms^2$  on the y-axis. The minimum TP for SR and ST volunteers is  $1 ms^2$  and  $0 ms^2$ , respectively. The Q1 TP for SR and ST volunteers is  $3 ms^2$  and  $1 ms^2$ , respectively. The median TP value for SR and ST volunteers is  $8 ms^2$  and  $5 ms^2$ . The Q3 TP for SR and ST volunteers is  $18 ms^2$  and  $28 ms^2$ , respectively. The maximum TP for SR and ST volunteers is  $25 ms^2$  and  $58 ms^2$ .

Fig. 7 demonstrates the boxplot of the ratio of low frequency to high frequency. The data classes are represented on the boxplot's x-axis, and LF/HF ratio between SR and ST volunteers on the y-axis. The minimum LF/HF ratio between SR and ST volunteers is 0.1 and 0.2 respectively, and the average LF/HF ratio for SR and ST volunteers is 3 and 4, respectively, and the maximum LF/HF ratio for SR and ST volunteers is 5 and 13 respectively.

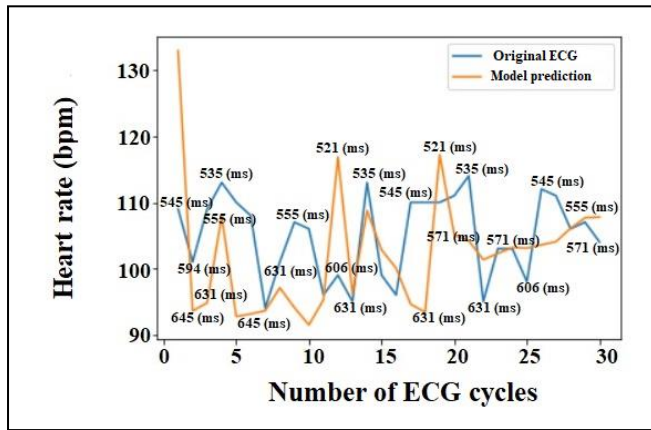


Fig. 8. Actual prediction and Model prediction

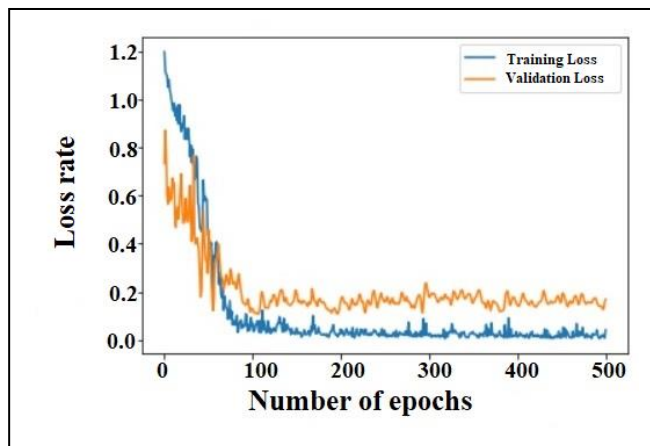


Fig. 9. Training and Validation loss rates of the designed and developed LSTM model

Fig. 8 displays the original ECG and model prediction of the ST volunteer's heart rate. The x-axis shows the ECG cycles. The y-axis represents the heart rate in beats per minute (bpm). The figure RRI (ms) has also been mentioned for every observable change in the heart rate. The first heart rate predicted at the first ECG cycle of the original ECG is 110 (bpm) (545 ms), and the model prediction is 93 (bpm) (645 ms). The heart rate predicted at the 30<sup>th</sup> ECG cycle of the original ECG, and the model prediction is 105 (bpm) (571 ms) and 108 (bpm) (555 ms), respectively.

Fig. 9 illustrates the training and validation loss rates of the designed and developed LSTM-RNN. The epochs count is displayed on the x-axis, and the y-axis represents the model's loss rate after each epoch. The model's training accuracy is 0.9672, and the model's validation accuracy is 0.8521.

Table 1 showcases the HRV features of the study population. The time domain parameters of HRV analysis was majorly influenced by the mean heart rate. The frequency domain parameters of HRV analysis like LF, and LF/HF ratio were comparatively high as the LF increased for healthy volunteers under physical stress. The study on Total Power (TP) as a frequency domain parameter is rare and thus it has been presented in this work.

#### IV. DISCUSSION

To evaluate the functions of the autonomic nervous system (ANS) in different cardiac arrhythmias, analysis of HRV is the optimal method [13]. Analysis of HRV is done in two different methods, namely time and frequency analysis. Time-domain parameters such as SDNN reflect the ANS balance, RMSSD represents the vagal influence during respiration, NN50 and pNN50 divulge the parasympathetic activity's information. Frequency domain parameters such as HF reveal the parasympathetic nervous system (PNS), and LF states the PNS and ANS information. The ratio of LF and HF indicates the sympathovagal balance [14, 15].

In healthy individuals, PNS is more active or prominent when the individual is at rest. Parasympathetic activity tends to decrease as the intensity of the exercise increases in a healthy individual, giving rise to sympathetic activity [16]. Savin et al. [17] concluded that reduced heart rate in the recovery phase of an exercise is seen because of delay in parasympathetic reactivation and not loss of sympathetic withdrawal. Crouse et al. [18], in their work, stated that sympathetic blockade had a very minimal influence on heart rate. Increased LF is observed in healthy volunteers due to physical activity [19]. The statistical analysis shown in this work is in line with [16-19] as the average PPI is more for SR than ST, resulting in significant changes being observed in the time domain parameters of the HRV analysis.

In the current study, HRV analysis comprising of time and frequency domain has been performed. The statistical analysis of PPI shown in Fig. 3 states that the atrial cycle of SR volunteers is more than ST volunteers due to increased heart rate affecting the ECG cycle length, and similar results have been observed in [20, 22-24]. The statistical analysis of SDNN observed in Fig. 4 displays variation for SR compared with ST because, in ST, the heart rate is lower; thus, the variation observed is minimal [25]. The analysis of RMSSD, an HRV parameter shown in this work, is very low for ST volunteers compared with SR volunteers because the HRV parameters generally depend on the mean heart rate [25, 26]. Analysis of total power an HRV parameter is infrequent as in frequency domain analysis mostly focuses on LF and HF [16-19, 26]. Comparing the time domain parameters SDNN and RMSSD decrease during higher heart rate [13-26], but on the contrary total power increased in this work. In [25] concluded that TP reduces as the heart rate increases. The reason for an increase in TP as the age group considered for this work is not varied, as discussed in [13-26]. LF/HF's ratio is more for ST volunteers than SR volunteers depicted in this work as LF increases in healthy volunteers under physical activity [19].

In the past, few studies [20-24] have developed statistical models like ARX and ARMAX to predicted ECG using HRV analysis. Taye et al. [27] developed a Convolutional Neural Network (CNN) to predict the onset of ventricular tachyarrhythmia using HRV as the input features and obtained an accuracy of 84.6%. Parsi et al. [28] ranked the HRV features based on the minimum redundancy maximum relevance (mRMR) method by combining mRMR and statistical machine learning classifiers such as SVM, kNN, and RF and obtained an accuracy of 86%. In the present



work, the accuracy obtained by the designed and developed LSTM-RNN model is 85.21%.

## V. CONCLUSION

In the present work, HRV analysis for SR and ST conditions was performed. The 12 different HRV parameters acted as the feature set for the designed and developed LSTM-RNN model. The logic behind designing and developing the LSTM-RNN model is to use it for multivariate time series forecasting. The LSTM-RNN model was used to predict the heart rate for ST volunteers. The training accuracy obtained by the model is 96.72%, and the validation accuracy is 85.21%. The LSTM-RNN model's advantages are that it performs well to predict time series of unknown duration. The time gap in the input time series data does not affect the LSTM model, giving an advantage over other time series prediction models.

## ACKNOWLEDGMENT

The authors acknowledge the support from the Ministry of Education, Government of India. The present study was supported by financial grants from the Science Engineering Research Board (SERB), Department of Science and Technology (DST), Government of India (EEQ/2019/000148).

## REFERENCES

- [1] B. Olshansky, "What's so inappropriate about sinus tachycardia?" *J. Cardiovasc. Electrophysiol.*, vol. 19, pp. 977-980, September 2008.
- [2] R. S. Sheldon et al., "Heart Rhythm Society expert consensus statement on the diagnosis and treatment of postural orthostatic tachycardia syndrome, inappropriate sinus tachycardia, and vasovagal syncope," *Heart Rhythm*, vol. 12, pp. 41-63, June 2015.
- [3] B. Olshansky, R. M. Sullivan, "Inappropriate sinus tachycardia," *EP Europace*, vol. 21, pp. 194-207, February 2019.
- [4] B. Olshansky, R. M. Sullivan, "Inappropriate sinus tachycardia," *J. Am. Coll. Cardiol.*, vol. 61, pp. 793-801, February 2013.
- [5] S. Wang et al., "Heart rate and heart rate difference predicted the efficacy of metoprolol on postural tachycardia syndrome in children and adolescents," *J. Pediatr.*, vol. 224, pp. 110-114, September 2020.
- [6] K. Kaczmarek et al., "Baseline intrinsic heart rate and response to ivabradine treatment in patients with inappropriate sinus tachycardia," *Ann. Noninvas. Electro.*, vol. 25, pp. e12709, May 2020.
- [7] J. Jarkovsky et al., "Heart rate as an independent predictor of long term mortality of acute heart failure patients in sinus rhythm according to their ejection fraction: data from the AHEAD registry," *Eur. J. Intern. Med.*, vol. 78, pp. 88-94, August 2020.
- [8] A. K. Fenny et al., "Artificial intelligence and machine learning in arrhythmias and cardiac electrophysiology," *Circ. Arrhythm. Electrophysiol.*, vol. 13, pp. e007952, August 2020.
- [9] J. H. Yoon et al., "Predicting tachycardia as a surrogate for instability in the intensive care unit," *J. Clin. Monit. Comput.*, vol. 33, pp. 973-985, February 2019.
- [10] H. Kim et al., "Prediction of post-intubation tachycardia using machine learning models," *Appl. Sci.*, vol. 10, pp. 1151-1171, February 2020.
- [11] M. Samuels et al., "Effectiveness and cost of recruiting healthy volunteers for clinical research studies using an electronic patient portal: A randomized study," *Clin. Transl. Sci.*, vol. 1, pp. 366-372, December 2017.
- [12] H. B. Braiek, F. Khomh, "On testing machine learning programs," *J. Syst. Softw.*, vol. 164, pp. 110542, June 2020.
- [13] P. K. Stein, M. S. Bosner, R. E. Kleiger, B. M. Conger, "Heart rate variability: a measure of cardiac autonomic tone," *Am. Heart J.*, vol. 127, pp. 1376-1381, May 1994.
- [14] C. R. Cole, E. H. Blackstone, F. J. Pashkow, C. E. Snader, M. S. Lauer, "Heart-rate recovery immediately after exercise as a predictor of mortality," *N. Engl. J. Med.*, vol. 341, pp. 1351-1357, October 1999.
- [15] U. Nussinovitch et al., "Reliability of Ultra – Short ECG indices for heart rate variability," *Ann. Noninvas. Electrocardiol.*, vol. 16, pp. 117-122, April 2011.
- [16] R. Shephard, "Exercise Physiology", Philadelphia, PA, B.C. Decker Inc., 1987.
- [17] W. M. Savin, D. M. Davidson, W. L. Haskell, "Autonomic contribution to heart rate recovery from exercise in humans," *J. Appl. Physiol. Respir. Environ. Exer. Physiol.*, vol. 53, pp. 1572-1575, December 1982.
- [18] S. F. Crouse, J. Sterling, H. Tolson, S. Hasson, "The effect of beta-adrenergic blockade on heart rate recovery from exercise," *J. Cardiopulm. Rehabil. Prev.*, vol. 9, pp. 202-206, May 1989.
- [19] A. Malliani, M. Pagani, F. Lombardi, S. Cerutti, "Cardiovascular neural regulation explored in the frequency domain," *Circulation*, vol. 84, pp. 482-492, August 1991.
- [20] S. Karimulla, J. Sivaraman, "The role and significance of Atrial ECG components in standard and modified lead systems," *Electronic Systems and Intelligent Computing (ESIC 2020)*, vol. 686, pp. 347-355, September 2020.
- [21] A. Jyothisana, J. Sivaraman, "A study on stability analysis of QT Interval dynamics of ECG using ARMAX model," *Electronic Systems and Intelligent Computing (ESIC 2020)*, vol. 686, pp. 307-316, September 2020.
- [22] A. Jyothisana, B. Arya, J. Sivaraman, "Stability analysis on the effects of heart rate variability and premature activation of atrial ECG dynamics using ARMAX model," *Phys. Eng. Sci. Med.*, vol. 43, pp. 1361-1370, December 2020.
- [23] J. Sivaraman, G. Uma, P. Langley, M. Umopathy, S. Venkatesan, G. Palanikumar, "A study on stability analysis of atrial repolarization variability using ARX model in sinus rhythm and atrial tachycardia ECGs," *Comput. Meth. Prog. Bio.*, vol. 137, pp. 341-351, December 2016.
- [24] B. Dhananjay, J. Sivaraman, "The role of heart rate variability in atrial ECG components of normal sinus rhythm and sinus tachycardia subjects", *Advances in Intelligent Systems and Computing (AISC 2020)*, vol. 1171, pp. 637-644, August 2020.
- [25] M. Malik et al., "Task force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology. Heart rate variability: standards of measurement, physiological interpretation and clinical use." *Circulation*, vol. 93, pp. 1043-1065, March 1996.
- [26] A. K. Reimers, G. Knapp, C. D. Reimers, "Effects of exercise on the resting heart rate: a systematic review and meta-analysis of interventional studies," *J. Clin. Med.*, vol. 7, pp. 503, December 2018.
- [27] G. T. Taye, H. J. Hwang, K. M. Lim, "Application of a convolutional neural network for predicting the occurrence of ventricular tachyarrhythmia using heart rate variability features," *Sci. Rep.*, vol. 10, pp. 1-7, April 2020.
- [28] A. Parsi, D. Byrne, M. Glavin, E. Jones, "Heart rate variability feature selection method for automated prediction of sudden cardiac death," *Biomed. Signal Proces.*, vol. 65, pp. 1-13, March 2021.

# Modelling of Electrocardiogram Using Autoregressive Moving Average Model and Linear Predictive Coefficient A Comparative Study

Ms. Rashmi Annamma George<sup>1</sup>, Dr.R. Periyasamy<sup>2</sup>

<sup>1</sup> Research Scholar, ICE Department, National Institute of Technology, Tiruchirappalli.

<sup>2</sup> Assistant Professor, ICE Department, National Institute of Technology, Tiruchirappalli.

rashmi.ann21@gmail.com<sup>1</sup>, periyasamy25@gmail.com<sup>2</sup>

**Abstract**—This paper investigate ECG prediction based on the Autoregressive model, Autoregressive Moving Average, and Linear Predictive Coefficient model. It is a complicated and tedious process to predict if the data set is stochastic. The autoregressive model is one of the foremost technique for statistical analysis. Being a linear combination, the model is outlined by the coefficients or the weights within the linear arrangement. By using the best fit models, different models can be evaluated against ECG signal to decide which model portrays the details more reliably. Modelling time series is used for forecasting. The most common methods of modelling time series data for the forecasting of stochastic signals are autoregressive process and autoregressive moving average process. Comparative results have been reported by selecting a model in terms of root mean square error according to the best fits findings. This paper concludes that using a time series model has better forecasting results than the predicted values from existing techniques. The best fit using the model was 97.9%.

**Keywords**—Ecg Modelling, ARMA, LPC

## 1. INTRODUCTION

An ECG is a recording of the heart's electrical activity. These signals change with time and are caused by the expansion and contraction of the heart. The ECG signal is obtained by determining the potential difference between electrodes mounted on the skin's surface. The ECG's various waves reflect the series of depolarization and repolarization of the atria and ventricles, respectively. The ECG signals are marked as P, Q, R, S and T [8]. The ECG signals are also segmented as PQ, QRS complex and ST Segment. The corresponding location and magnitude of these peaks carry valuable information about the heart functions. Modelling is a very useful medium for this electrical activity to be portrayed. Modelling is essentially a

mathematical representation of a method, action, or experimental effect. There are many ways in which ECG modelling is used.

The autoregressive (AR) and autoregressive moving average (ARMA) models are the most widely used linear regression models for time series analysis [1].

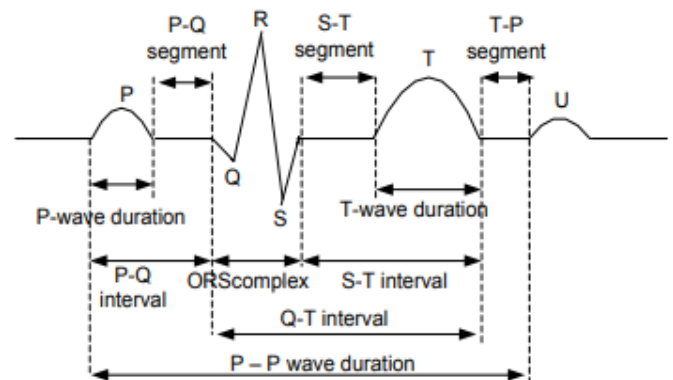


Figure 1: Typical ECG signals and its peaks

The idea of signal modelling is to represent the signal via model parameters. Signal modelling is mainly used for two purposes. Synthesis of signals and educational purpose. Educational purpose includes Research and development, which is used for the categorization of various disease conditions and also to generate signals models. The objective of this endeavour is to identify various approaches to model ECG waveform and to identify their constraints.

## 2. LITERATURE REVIEW

In the recent past, several theoretical investigations have been carried out to model physiological waveforms. Kim et al [1] proposed an

adaptive forecasting algorithm that forcefully changes the autoregressive system (AR) coefficients. The order of the AR process was determined through fractal dimension. Eva et al [2] suggested a forecasting technique based on fuzzy transformation for the time series. Fuzzy rules are automatically generated from the data given and used to predict potential components. Ronaldo et al [5] proposed an autoregressive (AR) process for the respiratory sinus arrhythmia to obtain the tachograms. AR process is used to determine the frequency which controls the R-R intervals. AR system power spectrum has a form that is similar to the experimental effect than the power spectrum derived from the Gaussian distributions. Kansal [6] proposed a univariate autoregressive model for forecasting future values of a time series based on current and previous values. In multidimensional time series, the model is based on multiple variables, including the forecast variable, present, and past values. The predictive accuracy is measured by root mean square error. The key drawback of this design is that many parameters may be selected.

Hamzacebi [7] proposed and tested an artificial neural network (ANN) architecture for seasonal time series forecasting and analysis on four real-world time series. The results of the proposed ANN were compared to those of other ANN architectures and standard statistical models. According to the proposed model, it has a lower prediction loss than other models. Zhong Gao et al [9] proposed a data prediction method that combines back propagation neural network (BPNN) and variational mode decomposition. The performance assessment indexes chosen are RMSE and MAE. Bodisco et al. [11] suggested that statistical models are helpful for the segment-wise synthesis of ECG. The model is made up of a series of piece-wise continuous periodic functions that start at a peak and end at a trough. In most of the models proposed in the literature, either they cannot be applied for clinical practices or real time signals. Therefore, the objective is to develop a model that can be used for real time signals and thereby predicting the future patterns of the signals.

### 3. PROPOSED METHOD

The proposed work is to investigate ECG prediction based on AR, MA, ARMA and LPC model represented as a time series, to predict the ECG signals if given the previous values. In time series analysis, autoregressive models (AR), or autoregressive moving average (ARMA) models are often used to fit the time series data. The purpose of this analysis is to help predict the future trend more accurately. In the AR model, the current state is highly linear based on the previous results.

AR recognition methods are currently being used in signal modelling for a broad class of physical signals [3]. The MA system predicts data sets dynamically, offering a plethora of time series options

for forecasting patterns. The ARMA model generates and predicts linear data. Since each model has different order, determining the correct order is important. The aim of this paper is to compare various parameter estimation methods for nonstationary signal identification. The assumption that the non-stationary signal satisfies the Yule-Walker equations and is a common method for solving Autoregressive processes. This method has the advantage of having calculations are accurate, particularly for a large number of data points [4].

Autoregressive Moving Average model is given by

$$y[n] = \sum_{k=1}^p a_n y(n-k) + G \sum_{l=1}^q a_n x(n-l) \quad (1)$$

where  $x[n]$  is the inputs and  $y[n]$  the output.

Auto-regressive model parameters are related to auto-covariance using Yule Walker equations. The model parameters can be found by yule walker equations. Auto correlation function (ACF) is used to calculate the model order. The yule walker equations are given in matrix form as

$$\begin{bmatrix} r_{xx}[0] & \cdots & r_{xx}[1-N] \\ \vdots & \ddots & \vdots \\ r_{xx}[N-1] & \cdots & r_{xx}[0] \end{bmatrix} \begin{bmatrix} a_1 \\ \vdots \\ a_N \end{bmatrix} = \begin{bmatrix} r_{xx}[1] \\ \vdots \\ r_{xx}[N] \end{bmatrix} \quad (2)$$

The goodness of fit and number of model parameters are identified. The ARMA (autoregressive-moving average) model describes a stationary signal by combining two terms: one term for auto-regression (AR) and the other for moving average (MA). Based on previous samples, the Linear Predictor Coefficient predicts the current value of a real-valued time series. The prediction value can be approximated as a weighted linear combination of the past samples. Two types of linear predictor are available, Forward linear Predictor and backward linear predictor. Here we look into forward linear predictor.

The correct order of the system was chosen after determining the best fit design. This was done by autocorrelation function. The next step is to predict the patterns. In the proposed method, we use three methods for testing. We have taken MIT BIH database and found the model parameters and noise variances by yule walker equations. The predicted patterns are then compared with patterns generated and found the accuracy in terms of root mean square error. The performance of the models is evaluated using root mean square error (RMSE). The root mean square error is defined as the square root difference between actual value and predicted value which is commonly used for measuring the difference between predicted and measured value.

The 28 recordings of ECG signals from MIT-BIH ST Change data base is first sampled at 250Hz and then derived a time series of instant Heart Rate. The next procedure is to find out the location of each R wave peak. For that the ECG signal passes through the various

stages of the PAM- TOMPKINS algorithm to detect the location of each R wave peak. Later, location information of R wave is converted into time series of Instant HR by calculating the time difference between two consecutive R peaks. The next step is to found the model parameters by Yule Walker equations. After finding the model order, the signal is predicted with AR, ARMA and LPC

The proposed process is as shown

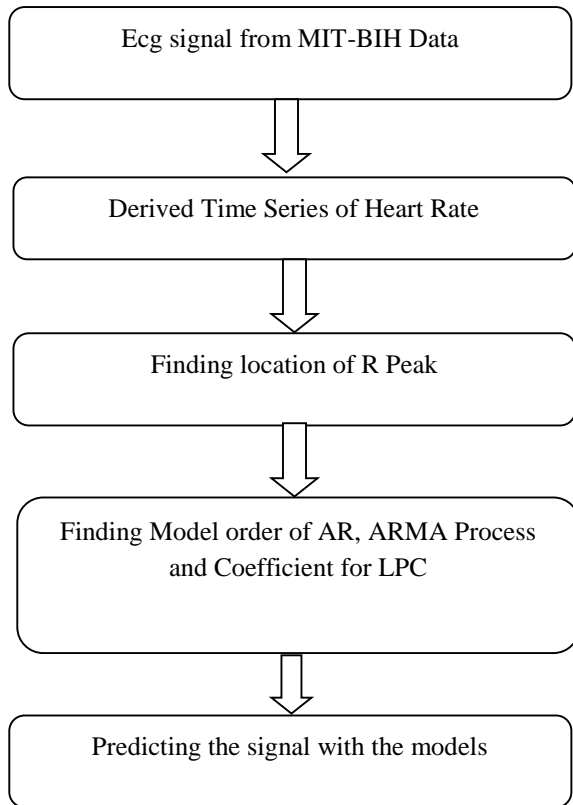


Figure 2: Flowchart of the proposed system

#### 4. RESULTS AND DISCUSSIONS

This section explains the results of the three predicted models and the error comparisons. The three models are applied on the 28 recordings of ECG signals from MIT-BIH ST change data base and performance was evaluated using RMSE. The mean Heart Rate Variability was found by 53.10. The data base is first sampled and then derived a time series of instant Heart Rate. And then found out the location of each R wave peak. After that the ECG signal passes through the various stages of the PAM- TOMPKINS algorithm and the location of each R wave pea was detected. Later, location information of R wave is converted into time series of Instant HR by calculating the time difference between R peaks. The model order was found by yule walker equations and the model parameters were identified. Using the model parameters of AR and ARMA process, the signal was predicted and the RMSE

was found. The results obtained are shown below. Figure 3 is a raw ecg signal. The predicted signals using AR, ARMA and LPC are shown in Figures 4,5 and 6 respectively.

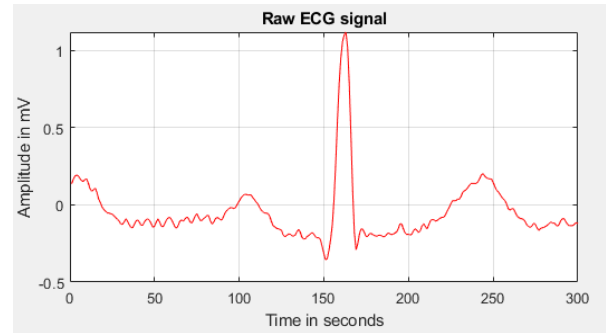


Figure 3: A Raw ECG signal

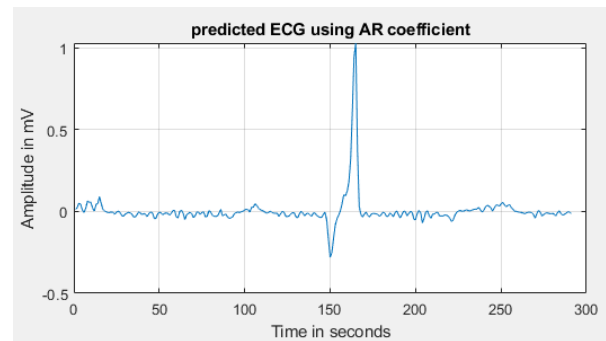


Figure 4: Predicted signal using AR process.

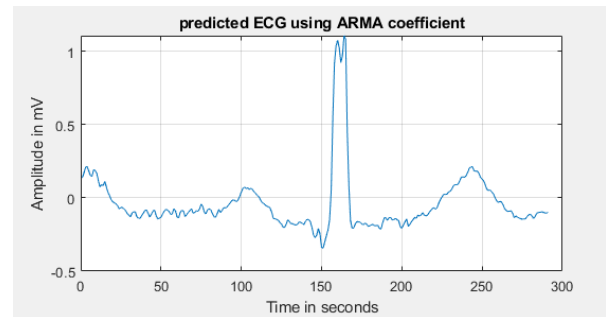


Figure 5: Predicted signal using ARMA process

The predicted data using LPC is shown

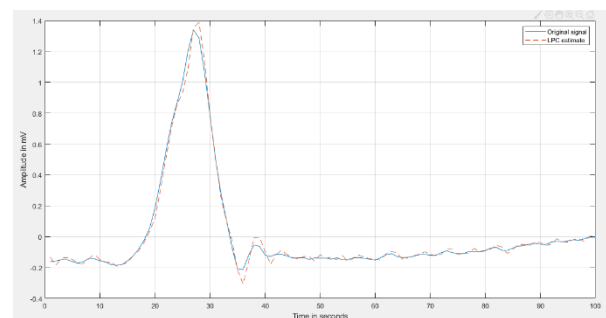


Figure 6: Measured and forecasted signal using LPC.

The order and root mean square error for different methods is shown in Table 1.

TABLE 1: Comparison of RMSE for AR, ARMA and LPC

Methods	Order	RMSE
AR	10	.2066
ARMA	10	.2064
LPC	3	.0208

After comparing the original data and the data we predict, 97.9% our prediction using LPC matches the original data. We would try to apply our model and algorithm to different types of data in the future particularly ECG signal from the exercise testing and forecasting the future patterns thereby avoiding the risk occurring during Tread Mill Test.

## REFERENCES

- [1] Sun-Hee Kim, Christos Faloutsos, and Hyung-Jeong Yang, Coercively Adjusted Auto Regression Model for Forecasting in Epilepsy EEG, Computational and Mathematical Methods in Medicine, The Scientific World Journal, Volume 2013
- [2] Eva Volna, Martin Kotyrba, and Hashim Habiballa, ECG Prediction Based on Classification via Neural Networks and Linguistic Fuzzy Logic Forecaster, Hindawi Publishing Corporation, The Scientific World Journal, Volume 2015
- [3] H.T. Wu, H.K. Wu, C.L. Wang, Y.L. Yang, W.H. Wu, T.H. Tsai and H.H. Chang, "Modeling the Pulse Signal by Wave-Shape Function and Analyzing by Synchro-squeezing Transform", Published: June 15, 2016.
- [4] J.M. Boulakia, S. Cazeau, M.A. Fernández, J.F. Gerbeau and N. Zemzemi, Mathematical Modeling of Electrocardiograms: A Numerical Study, Annals of Biomedical Engineering vol. 38, 2008, pp 1071–1097.
- [5] Ronaldo M Evaristo, Antonio M Batis ta, Ricardo L Viana, Kelly C Iarosz, José D Szezech Jr, Moacir F de Godoy, "Mathematical modelling with autoregressive process for Electrocardiogram signals, Elsevier Science Publishers, Commun Nonlinear Sci Numer Simulat, June 15, 2017
- [6] Sarita Kansal, prasant P Bansod, Multivariate autoregressive model for ECG signal forecasting", International Journal of multivariate data analysis, Volume 1, 2017
- [7] C. Hamzacebi, "Improving artificial neural networks' performance in seasonal time series forecasting", Information Sciences 178 (2008), pages: 4550-4559
- [8] Om Prakash Yadav, Shashwati Ray. "Smoothing and Segmentation of ECG Signals Using Total Variation Denoising Minimization-Majorization and Bottom-Up Approach". International Conference on Computational Modeling and Security (CMS 2016).
- [9] Zhong Gao Sun, Ying Lei, "An ECG signal analysis and prediction method combined with VMD and Neural Network", 2017 7th IEEE International Conference on Electronics Information and Emergency Communication (ICEIEC), 2017.
- [10] Zeeman, E.C., 1972a. Differential Equations for the Heartbeat and Nerve Impulse. Mathematics Institute, University of Warwick, Coventry, UK
- [11] T. Bodisco, J.D. Netto, N. Kelson, J. Banks, R. Hayward and T. Parker, characterizing an ECG signal using statistical modeling: a feasibility study, Australia and New Zealand industrial and applied mathematics (ANZIAM) journal, vol. 55, pp. C32-C46.
- [12] S. Das and K. Maharatna, Fractional dynamical model for the generation of ECG like signals from filtered coupled Van-der-Pol oscillators, Computational methods of Biomed, vol. 112, issue 3, 2013, pp. 490-507.
- [13] P.E. McSharry, G.D. Clifford, L. Trassenko and L.A. Smith, A dynamical model for generating synthetic electrocardiogram signals, IEE Transactions on Biomedical Engineering, vol. 50, issue: 3, March 2003.
- [14] N.J. Dabanloo, D.C. McLernon, H. Zhang, A. Ayatollahi and V. Johari-Majd, A modified Zeeman model for producing HRV signals and its application to ECG signal generation, Journal of Theoretical Biology, vol. 244, issue 2 244, 2017, pp. 180-189.
- [15] J. S. Kim, P. Smyth and S. Luther, Modeling Waveform Shapes with Random Effects Segmental Hidden Markov Models, Computational Engineering, Finance, and Science Jul 11, 2014.

# Extraction of Abnormal Skin Lesion from Dermoscopy Image using VGG-SegNet

Seifedine Kadry  
*Faculty of Applied Computing and  
 Technology, Noroff University College,  
 Kristiansand, Norway*  
 email: skadry@gmail.com

Venkatesan Rajinikanth  
*Department of Electronics and  
 Instrumentation Engineering  
 St. Joseph's College of Engineering  
 Chennai 600119, India*  
 email: v.rajinikanth@ieee.org

David Taniar  
*Faculty of Information Technology  
 Monash University  
 Clayton, Victoria 3800, Australia*  
 email: David.Taniar@monash.edu

Isah A. Lawal  
*Faculty of Applied Computing and  
 Technology, Noroff University College,  
 Kristiansand, Norway*  
 email: Isah.Lawal@noroff.no

Robertas Damaševičius  
*Faculty of Applied Mathematics  
 Silesian University of Technology  
 44-100 Gliwice, Poland*  
 email: robertas.damasevicius@polsl.pl

**Abstract**— Skin is one of the vital and well-known sensory organs in human physiology and due to various reasons, the abnormality in skin arises. Skin Melanoma (SM) is one of the medical crisis in humans and the untreated SM will cause various abnormalities, such as skin irritation, spreading the cancerous cells through the blood stream, etc. Efficient assessment of the SM is essential to identify the severity of the disease and hence the proposed work implemented a Convolutional-Neural-Network (CNN) based approach to support the automated SM examination. This work employed the VGG-SegNet scheme to extract the SM section from the Digital-Dermoscopy-Image (DDI). After the extraction, a relative assessment between the segmented SM and the Ground-Truth (GT) is executed and the essential performance indices are then computed. The proposed scheme is tested and validated using the benchmark ISIC2016 database and the average result attained with the proposed study helped to achieve a better values of Jaccard-Index, Dice, and Accuracy for the DDI with and without the artifacts. These results confirm that, proposed technique is significant in evaluating the clinical grader DDI.

**Keywords**—Skin-Melanoma, Digital dermoscopy, VGG-SegNet, Segmentation, Evaluation.

## I. INTRODUCTION

The skin is one of the vital sensory organ and also responsible to protect the inner organs from the outer environment. The disease in skin will cause various difficulty in humans and hence a considerable precautionary measures are needed to prevent the skin from the diseases [1-3].

In humans, skin cancer is one of the medical emergency and timely recognition and handling is essential to cure the disease [4,5]. The skin cancer is categorized as non-melanoma and melanoma and the report of World-Health-Organisation (WHO) confirms that, globally 3 million active non-melanoma and 132,000 melanoma cases are existing and every year the infection rate is rising in alarming rate [6,7]. The WHO also confirms that the risk factor in Caucasian populations is more than dark-skinned populations. To reduce the skin cancer occurrence rates, the WHO recommended various guidelines and also insisted to conduct the awareness programs to save the human community from skin cancer.

From the earlier study, the main cause for the Skin Melanoma (SM) is predicted to be the intermittent and high exposure of the skin to ultraviolet (UV) radiation. The early study also reported that the untreated sun burn will lead to the

skin melanoma. When the SM is not treated in its early phase, the cancerous cells will spread through the blood stream; which cannot be cured completely. Hence, to support the early detection and treatment of SM, a number of scheduled health examinations is recommended by the doctors to patients.

The various stages involved in the detection of SM includes; (i) Self examination by the patient to identify the doubtful skin sections, (ii) Detailed examination of suspicious skin section by a dermatologist, (iii) Dermoscopy based assessment with prescribed clinical protocol, (iv) Recommending the needle biopsy test to confirm the cancer stage, and (v) Minor/major surgery to completely remove the cancerous skin section.

The clinical level assessment of the SM is commonly performed using the ABCD/ABCDE rule, in which the shape and the structural features of the abnormal skin section and examined by the dermatologist and based on the finding treatment related decision will be taken by the dermatologist. When the number of patients to be examined is more, then the skin clinics will recommend the Digital-Dermoscopy-Image (DDI) based examination procedures in order to support the accurate estimation of the SM. If the recorded DDI is very clear, then evaluation of SM is uncomplicated and the skin infection level can be accurately examined. If the abnormal section in DDI is associated with artefacts, such as, hair, medical gel, markings, etc, then the assessment of the SM seems to be complex and needs a special tool for efficient diagnosis.

Development of an automated scheme for the accurate diagnosis of SM for clinical grade DDI is very essential and hence, the proposed work employed the Convolutional-Neural-Network (CNN) based system. The extraction of the SM section from DDI is one of the common practice and this work implements the VGG-SegNet scheme to extract the SM section from the DDI with and without the artefact. In this work, a pre-trained VGG-SegNet is employed to examine the SM fragment of the benchmark DDI images of International-Skin-Imaging-Collaboration (ISIC2016) database [8-11]. This database consist 1250 DDI images (900 training+350 testing) along with the related Ground-Truth (GT). The employed VGG-SegNet helps to get the binary form of the SM fragment. After the extraction, a relative appraisal of SM and the GT is implemented and based on the computed values of Jaccard-Index, Dice, Accuracy (ACC), Precision (PRE), Sensitivity (SEN), Specificity (SPE) and Negative-Predictive-

Value (NPV) the merit of the proposed segmentation scheme is confirmed.

The other sections are organised as below; Section 2 discuss the Related works, Section 3 demonstrates the methodology, Section 4 and 5 give the experimental outcome and discussions and Section 6 concludes the proposed work.

## II. RELATED WORKS

Due to its significance, a number of DDI assessment methods are proposed and implemented for efficient detection of SM. Most of the earlier research suggests segmentation or a classification approach to detect the SM from the DDI. The employment of semi-automated/automated segmentation and machine/deep learning classification can be found in earlier works [1-3, 12,13].

In the SM segmentation approach, the main aim is to develop a methodology which supports the mining of the SM section from the considered DDI. The implementation of thresholding and segmentation is widely found in the literature and the aim of this technique is to mine the SM with better accuracy. Normally, the segmentation methods helps to get the binary version of the SM fragment, which is then evaluated using the standard techniques. The earlier work implemented a method to evaluate the skin melanoma harshness based on the ABCD/ABCDE rule [1,13]. In this approach, essential information such as area, boundary, and diameter are measured using the binary version of the SM section.

Classification of the DDI into normal/melanoma class is also one of the promising works, in which a suitable machine/deep-learning scheme is implemented to categorize the considered DDI using suitable two/multi class classifiers [12].

The ultimate aim of the SM examination procedure is to develop an appropriate methodology to confirm the melanoma and its harshness in the DDI using the implemented method. To support the efficient diagnosis of the SM, this work implemented an automated CNN scheme to support the efficient SM fragment mining. In this work, a pre-trained VGG-SegNet is implemented to evaluate the DDI with and without artefact and the attained result with ISIC2016 confirms that, proposed work is appropriate to examine the clinical grade DDI.

## III. METHODOLOGY

This section presents the methodology employed in this research work for DDI assessment. After gathering the necessary trial picture from the ISIC2016 database, every image is resized to the dimension of 224x224x3 pixels. Initially, this work considered the pre-trained Vgg-SegNet scheme to extort the SM part from the adopted trial images. The total numbers of trial images considered are 1250 numbers. Further, the image augmentation (horizontal & vertical flip,  $\pm 45^\circ$  rotation) is also employed during the pre-training process of the VGG-SegNet for the considered DDI.

Figure 1 presents the implemented scheme to extract the SM. Initially, the essential images of dimension 2048x1536x3 pixels is collected from the database and all the trial images and GT are then resized to a dimension of 224x224x3 pixels (recommended size for VGG scheme). The resized images are

then considered to train the SegNet scheme and after the essential training, SM part of every image is segmented and compared against the GT. This comparison will help to get the essential values of the Image-Performance-Measures (IPM) and based on the average of IPMs of all 1250 images, the segmentation performance of the VGG-SegNet is validated.

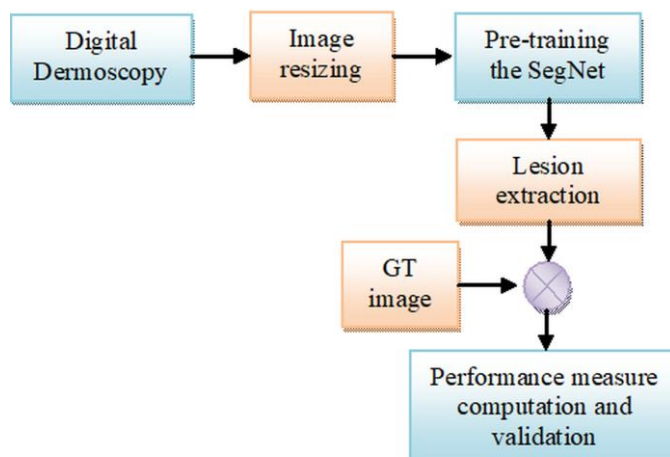


Fig.1. Proposed SM segmentation scheme

### A. Image database

Due to its significance, a number of SM evaluation systems are proposed to examine the cancerous fragment from the DDI. In this work, the proposed CNN scheme is tested and validated using the benchmark ISIC2016 database. It is a commonly adopted DDI dataset consist 1250 numbers of RGB scaled pictures with a dimension of 2048x1536x3 pixels and every image is associated with its related GT.

The recommended image dimension for VGG-SegNet is 224x224x3 pixels; hence every image and GT of ISIC2016 is resized before the assessment. This dataset consist images with various categories, such as clear, with hair section, with medical-gel, with scale markings, etc. Segmentation of SM fragment from clear image is quite uncomplicated compared with the DDI with artefact. Hence, the considered scheme is appropriately trained with the original as well as the augmented test images.

Figure 2 depicts the sample trial pictures collected from ISIC2016 and Fig 2(a) and (b) depicts the image without and with artefact, respectively.

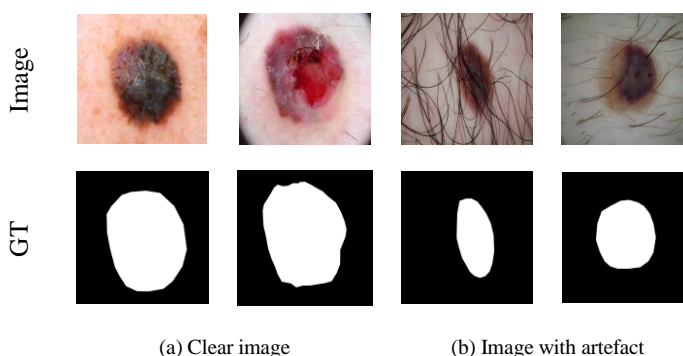


Fig. 2. Sample test images of ISIC2016

### B. VGG-SegNet

This work employs the pre-trained VGG-SegNet scheme to extract the SM fragment from the DDI and the architecture of this scheme is depicted in Figure 3. This method consists of two divisions namely the encoder and decoder parts as depicted in Figure 3. The Encoder consist a down convolution, which converts the given images into possible learned features and the decoder consist the up-convolution, which converts the images from the learned features [14-17]. The final segment of the decoder division consists of the SoftMax layer, which will support a binary classification to separate the SM from background. The proposed scheme helps to get a binary SM fragment, which is then compared against the GT for validation. Other essential information regarding the VGG-SegNet can be found in earlier works [18-22].

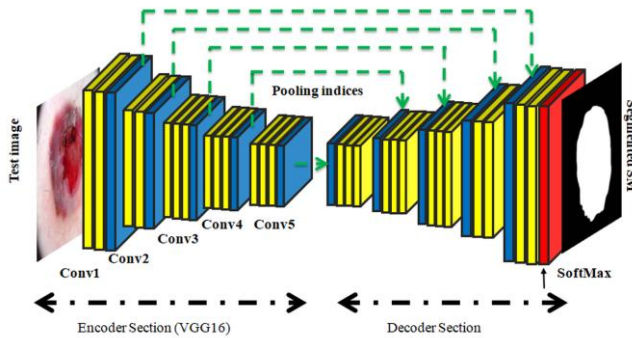


Fig. 3. Scheme of pre-trained VGG-SegNet

Considered VGG-SegNet scheme is initially trained using the considered DDI and later every image (1250 numbers) are tested with the trained system and the extracted binary SM is then considered for further evaluation.

### C. Validation

In medical image assessment, there is a proved practice to validate the constructed disease recognition scheme using the prescribed protocol to confirm its clinical significance.

This protocol suggests the validation of the proposed scheme by computing the important Image-Performance-Measures (IPM) during the assessment between the SM fragment and GT. In this work, the IPMs, such as True-Positive (TP), False-Positive (FP), False-Negative (FN), and True-Negative (TN) are computed. From these measures, other values, such as Jaccard, Dice, Accuracy (ACC), Precision (PRE), Sensitivity (SEN), Specificity (SPE), and Negative-Predictive-Value (NPV) are also accomplished, and based on the computed IPMs; the significance of VGG-SegNet is validated [21-25].

## IV. EXPERIMENTAL RESULT

This division of the research disclose the experimental results attained with the implemented scheme. This work is performed using the workstation; Intel i5 2.5GHz processor with 16GB RAM and 2GB VRAM set with MATLAB®.

Initially, the VGG-SegNet is trained using the resized trial images of dimension 224x224x3 pixels and during this task the original and augmented images are considered. When the scheme is properly trained, then every test image of the ISLC2016 is separately tested and the attained results are tabulated for further assessment. The attained result of this

work confirms that, proposed scheme helps to get better result on the clear as well as the DDI with artefact.

Figure 4 depicts the result attained for a clear DDI. Fig 4(a) depicts the chosen trial picture and Fig. 4(b)-(e) depicts the results obtained from various sections of VGG-SegNet. Finally, the segmented binary SM is depicted in Fig 4(f).

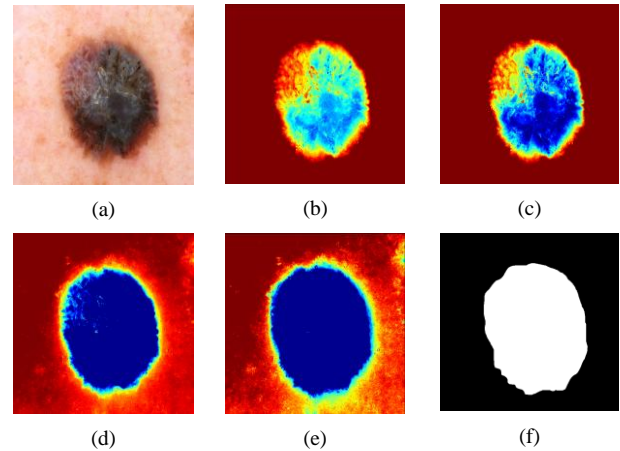


Fig. 4. Segmentation result attained with sample DDI

Test picture, (b),(c) sample result from encoder, (d), (e) Sample result of decoder, (f) Extracted SM

Similarly, Figure 5 depicts the outcome of the DDI with hair section. Extraction of the SM when the DDI is associated with the hair is very complex and from the results of Fig. 5, it is observed that the VGG-SegNet helped to extract SM with better accuracy. The sample image is depicted in Fig 5(a), Fig (b)-(e) depicts the intermediate results of this scheme and the final outcome is depicted in Fig 5(f).

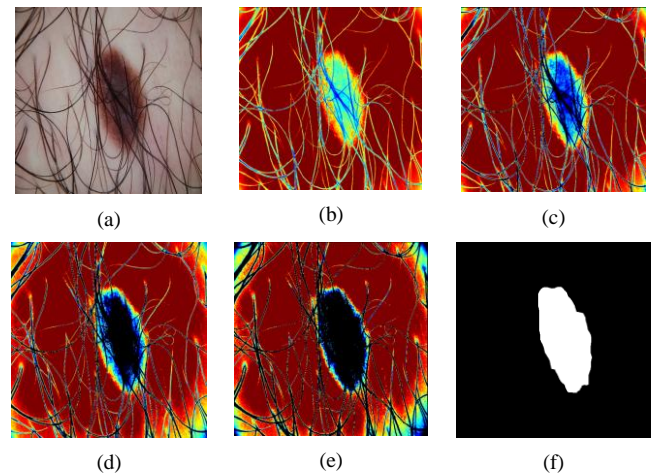


Fig. 5. Segmentation result attained with DDI with hair

(a) Test picture with artefact, (b),(c) sample outcome from encoder, (d), (e) Sample outcome of decoder, (f) Extracted SM

From the results of figure 4 and 5, it can be noted that the proposed scheme is efficient in localizing and segmenting the SM section from the considered DDI. Similar procedure is executed on all the test images and the extracted SM fragments are recorded.

After extracting all the SM sections of the database, an assessment is implemented with the GT and the necessary values of the IPMs are computed.



## V. DISCUSSION

This research work aims to implement the CNN segmentation to extract the SM fragment from the DDI with superior accuracy. The implemented method is tested and evaluated on ISIC2016 database and the results attained for sample DDI (DDI1 and DDI2) are depicted in figure 4 and 5 respectively.

After extracting the essential SM, an assessment is implemented with the GT and the essential IPMs are computed. The IPMs attained for DDI1 and DDI2 is depicted in Table I and II and these values confirm that the proposed VGG-SegNet scheme works well on the clear and the DDI with artefact. The IPM attained with the clear image is better compared to the image with the hair section. Similar procedure is employed to evaluate all the 1250 test images and the average of the computed IPMs are considered to validate the implemented scheme.

TABLE I. VITAL IMAGE PERFORMANCE MEASURE CALCULATED USING COMPARISON OF SM AND GT

Image	TP	FP	TN	FN	Jaccard	Dice
DDI1	16647	1854	31607	68	89.65	94.54
DDI2	5213	1523	43405	35	76.99	86.99

TABLE II. TIMAGE PERFORMANCE MEASURES FOR CHOSEN TEST PICTURESOF SAMPLE TEST IMAGES

Image	ACC	PRE	SEN	SPE	NPV
DDI1	96.17	89.98	99.59	94.46	99.78
DDI2	96.89	77.39	99.33	96.61	99.92

The main advantage of the proposed scheme is, it is an automated technique and does not need any hair removal approaches as discussed in the earlier research work. Table III compares the performance of the proposed technique with other existing method in the literature. The results presented in this table confirms that, proposed approach help to get better values of ACC, PRE and SEN compared to the other existing methods. The SPE attained in earlier works are better compared to VGG-SegNet.

TABLE III ASSESSMENT OF IMPLEMENTED SCHEME WITH EXISTING RESULTS

Method	ACC (%)	PRE (%)	SEN (%)	SPE (%)
EXB [26]	95.30	-	91.00	96.50
CUMED [26]	94.90	-	91.10	95.70
Mahnudar [26]	95.20	-	88.0	96.9
ALL-FCNs [12]	95.20	89.90	92.40	96.00
OSO-FCNs [12]	95.80	91.30	92.50	96.40
Rajinikanth et al. [2]	92.16	89.18	93.16	92.18
Dey et al. [3]	92.84	90.06	94.18	92.31
Rajinikanth et al. [13]	92.29	89.74	92.88	92.25
VGG-SegNet	97.16	92.81	95.04	94.75

The Glyph-plot depicted in figure 6 confirms that the overall performance attained with VGG-SegNet is superior compared to other techniques.

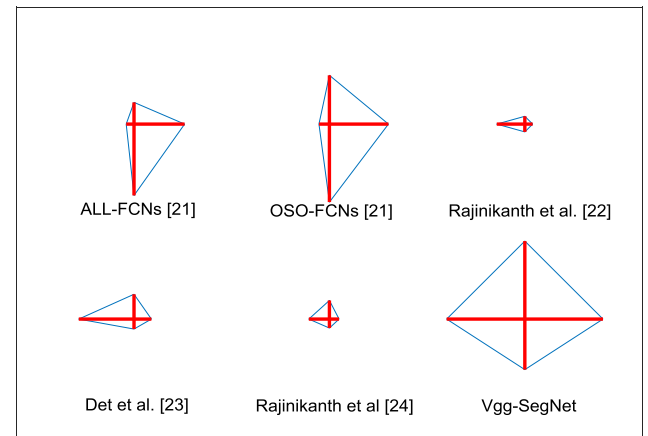


Fig. 6. Glyph-Plot to confirm the merit of VGG-SegNet

In the proposed research, only the segmentation is performed with the considered VGG-SegNet scheme. This scheme consist the combination of Encoder and Decoder section and the encoder section is the traditional VGG16 architecture without the Fully-Connected-Layer (FCL). When this section is associated with the FCL and the SoftMax classifier, then it is possible for us to classify the existing DDI into melanoma/non-melanoma class. The implementation of VGG16 based classification of the ISIC2016 is one of the recommended future scopes of this research. Further, in future, the extracted SM section can also be considered to get the essential texture and the shape features to implement the ABCD/ABCDE rule supported examination.

## VI. CONCLUSION

Examination of the skin cancer using the computerized tool will considerably reduce the diagnostic burden during the mass screening. Hence a number of methods are proposed and implemented to detect the skin abnormalities using the DDI. This work employed a pre-trained VGG-SegNet scheme to extract and evaluate the SM fragment from the DDI with dimension 224x224x3 pixels. The proposed section consist encoder-decoder section and at the end it has a SoftMax classifier to execute a binary classification. The classifier unit separates the SM from the background and the binary form of SM is then compared ageist the GT to find the IPMs. The average values of IPMs of all the 1250 images are then computed and validated against the existing methods in the literature. The result of this work confirms that, proposed work offers better result with the ISIC2016 database.

## REFERENCES

- [1] F. Nachbar, W. Stolz, T. Merckle 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.
- [2] V. Rajinikanth, S.C. Satapathy, N. Dey, S.L. Fernandes and K.S. Manic, "Skin melanoma assessment using Kapur's entropy and level set—A study with bat algorithm," vol. 104, pp. 193-202, 2019, [online] Available: [https://doi.org/10.1007/978-981-13-1921-1\\_19](https://doi.org/10.1007/978-981-13-1921-1_19).
- [3] N. Dey, V. Rajinikanth, A.S. Ashour and J.M.R.S. Tavares, "Social group optimization supported segmentation and evaluation of skin melanoma images," *Symmetry*, vol. 10, no. 2, pp. 51, 2018.
- [4] N. Kowsalya, A. Kalyani, T.D.V. Shree, N.S.M. Raja and V. Rajinikanth, "Skin-Melanoma Evaluation with Tsallis's Thresholding and Chan-Vese Approach," *In. IEEE International Conference on System Computation Automation and Networking (ICSCA)*, pp. 1-5, 2018.

- [5] N. Dey, V. Rajinikanth, H. Lin, and F. Shi, "A Study on the Bat Algorithm Technique to Evaluate the Skin Melanoma Images," In *Applications of Bat Algorithm and its Variants*, pp. 45-60. Springer, Singapore, 2021.
- [6] WHO ([https://www.who.int/news-room/q-a-detail/radiation-ultraviolet-\(uv\)-radiation-and-skin-cancer](https://www.who.int/news-room/q-a-detail/radiation-ultraviolet-(uv)-radiation-and-skin-cancer))
- [7] <https://dermnetz.org/topics/melanoma-pathology/>
- [8] M.A. Marchetti et al., "Results of the 2016 International Skin Imaging Collaboration International Symposium on Biomedical Imaging challenge: Comparison of the accuracy of computer algorithms to dermatologists for the diagnosis of melanoma from dermoscopic images," *Journal of the American Academy of Dermatology*, vol. 78, no. 2, pp. 270-277, 2017.
- [9] N.C.F. Codella et al., "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 *IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*, pp. 168-172, 2018.
- [10] <https://challenge.isic-archive.com/landing/2016>
- [11] D. Gutman, N.C.F. Codella, E. Celebi, B. Helba, M. Marchetti, N. Mishra, and A. Halpern, "Skin lesion analysis toward melanoma detection: A challenge at the international symposium on biomedical imaging (ISBI) 2016, hosted by the international skin imaging collaboration (ISIC)," *arXiv preprint arXiv:1605.01397*, 2016.
- [12] L. Huang, Y.-G. Zhao, and T.-J. Yang, "Skin lesion segmentation using object scale-oriented fully convolutional neural networks," *Signal, Image and Video Processing*, vol. 13, no. 3, pp.431-438, 2019. <https://doi.org/10.1007/s11760-018-01410-3>.
- [13] V. Rajinikanth, N.S.M. Raja, and S. Arunmozhi, "ABCD rule implementation for the skin melanoma assesment—a study," In 2019 IEEE International Conference on System, Computation, Automation and Networking (ICSCAN), pp. 1-4. IEEE, 2019. DOI: 10.1109/ICSCAN.2019.8878860.
- [14] V. Valiūskaitė, V. Raudonis, R. Maskeliūnas, R. Damaševičius, and T. Krilavičius, "Deep learning based evaluation of spermatozoid motility for artificial insemination," *Sensors*, vol. 21, no. 1, pp.72, 2021.
- [15] O. O. Abayomi-Alli, R. Damaševičius, M. Wiczorek, and M. Woźniak, "Data Augmentation Using Principal Component Resampling for Image Recognition by Deep Learning," In *International Conference on Artificial Intelligence and Soft Computing*, pp. 39-48. Springer, Cham, 2020.
- [16] A. T. Sahlol, D. Yousri, A.A. Ewees, A. A. Mohammed, R. Damaševičius, and M.A. Elaziz, "COVID-19 image classification using deep features and fractional-order marine predators algorithm," *Scientific Reports*, vol 10, no. 1, pp.1-15, 2021. <https://doi.org/10.1038/s41598-020-71294-2>.
- [17] M.A. Khan, I. Ashraf, M. Alhaisoni, R. Damaševičius, R. Scherer, A. Rehman, and S.A.C. Bukhari, "Multimodal brain tumor classification using deep learning and robust feature selection: A machine learning application for radiologists," *Diagnostics*, vol. 10, no. 8, pp. 565, 2020.
- [18] M. Nisa, J. H. Shah, S. Kanwal, M. Raza, M. A. Khan, R. Damaševičius, and T. Blažauskas, "Hybrid malware classification method using segmentation-based fractal texture analysis and deep convolution neural network features," *Applied Sciences*, vol. 10, no. 14, pp.4966, 2020.
- [19] M.A. Khan, N. Hussain, A. Majid, M. Alhaisoni, S.A.C. Bukhari, S. Kadry, Y. Nam, and Y.D. Zhang, "Classification of positive COVID-19 CT scans using deep learning," *Computers, Materials and Continua*, vol. 66, no. 3, pp.2923-2938, 2021. 1. doi:10.32604/cmc.2021.013191.
- [20] S.A. Roseline, S. Geetha, S. Kadry, and Y. Nam, "Intelligent Vision-Based Malware Detection and Classification Using Deep Random Forest Paradigm," *IEEE Access*, vol. 8, pp.206303-206324, 2020.
- [21] V. Rajinikanth, A.N.J. Raj, K.P. Thanaraj, and G.R. Nai., "A customized VGG19 network with concatenation of deep and handcrafted features for brain tumor detection," *Applied Sciences* 10, no. 10, pp.3429, 2020. <https://doi.org/10.3390/app10103429>.
- [22] N. Dey, Yu-Dong Zhang, V. Rajinikanth, R. Pugalenth, and N.S.M. Raja, "Customized VGG19 architecture for pneumonia detection in chest X-rays," *Pattern Recognition Letters*, vol.143, pp.67-74, 2021. <https://doi.org/10.1016/j.patrec.2020.12.010>.
- [23] B.A. Muthu, C. B. Sivaparthipan, G. Manogaran, R. Sundarasekar, S. Kadry, A. Shanthini, and A. Dasel, "IOT based wearable sensor for diseases prediction and symptom analysis in healthcare sector," *Peer-to-peer networking and applications*, vol. 13, no. 6, pp.2123-2134, 2020.
- [24] R. Roufayel, and S. Kadry, "Expression of miR-23a by apoptotic regulators in human cancer: A review." *Cancer biology & therapy*, vol.18, no. 5, pp. 269-276, 2017.
- [25] S. Oueida, P. A. Char, S. Kadry, and S. Ionescu, "Simulation models for enhancing the health care systems," *FAIMA Business & Management Journal*, vol. 4, no. 4, pp.5-20, 2016.
- [26] L. Yu, H. Chen, Q. Dou, J. Qin, and P.A. Heng, "Automated melanoma recognition in dermoscopy images via very deep residual networks," *IEEE Trans. Med. Imaging*, vol. 36, no.4, pp. 994–1004, 2017.

# Emotional Classification of EEG Signal using Image Encoding and Deep Learning

Anjana K A, Ganesan M, Lavanya R  
 Department of Electronics and Communication Engineering,  
 Amrita School of Engineering, Coimbatore  
 Amrita Vishwa Vidyapeetham, India

anjanakokkuvayil@gmail.com, m\_ganesan1@cb.amrita.edu, r\_lavanya@cb.amrita.edu

**Abstract**—For humans, emotions are important and play a significant role in human insight. Emotions are commonly identified by speech, facial expression and gesture. Recently Electroencephalograph (EEG) based emotion recognition have accumulated solid interest in the research community and it provides cheap, portable and reliable techniques for emotion recognition. In this work, classification of seed database having three emotions like positive, neutral and negative was performed, which was publicly available. This paper has two parts, time-series to image conversion of EEG signal and classification of emotion. First part, the data is transformed to an image that is used to analyse the EEG signal and in second part, the transformed image passes through deep learning to understand the emotions encountered during the EEG signal generation. Experimental results indicate that the scalogram of image encoding provides the best classification accuracy of 98%, compared to spectrogram and Hilbert Huang Transform (HHT) 78% and 75% classification accuracy respectively.

**Keywords**—EEG, Spectrogram, Scalogram, Hilbert-Huang Transform (HHT).

## I. INTRODUCTION

In human cognition and speech of everyday life activities, emotions play a crucial role. It is connected with decision-making, observation, and interpersonal recognition. Understanding and realizing how to respond to an individual's appearance mostly enhances the interactions. Psychology has an important role in realizing an individual's emotion. The study of non-verbal actions was pioneered by psychologist Ekman and it helped to recognize the six basic emotions such as rage, fear, disgust, excitement, surprise, and sadness [1].

By incorporating information from facial expressions, body movement, gestures [2] and speech [3], different forms of emotions were identified, which are known as external emotion expressions. Sometimes emotional states remain internal and cannot be detected externally. So the accurate rates of the external emotion expressed is high under extreme circumstances and low at normal situations. This work mainly focuses on emotion recognition based on electroencephalograph (EEG). EEG is an electrical activity of neural cells and it is reasonably objective assessment of emotions compared to non-physiological clues.

Spectrogram, Scalogram, Hilbert-Huang Transform (HHT) are techniques of image encoding that are used to translate time series data into images. A spectrogram, which is a visualisation of frequencies over time in a signal. It is the first image encoding technique that is used in this work. Time domain analysis and frequency domain analysis has been proven to be advantages of this technique [4]. Some researchers have recently explored automated extraction from image spectrograms with validated deep learning models. Yuan et al., [5] proposed spectrogram based image classification for patients' EEG data analysis with CNN

algorithm. Ruffin et al., [6] employed spectrogram based image classification for seizure detection from EEG signals with deep learning algorithms, which yielded an accuracy of 77.57%. The second technique employed is the scalogram, which is a wavelet representation, plotted as a time and frequency function. It provides a high time period at low frequencies windowing and low time period at higher frequencies windowing [7]. Li et al., [8] implied Convolutional Recurrent Neural Network (C-RNN) based emotion recognition from multi-channel EEG data with 72.06 percent valence and 74.12 percent arousal accuracy. Kaya et al., [9] suggested a minimum Redundancy Maximum Relevance-Convolutional Neural Network (mRMR-CNN) for classifying EEG signals based on an influential approach to the system for aid decision. And Turk et al., [10] employed scalogram based epilepsy detection using CNN. Energy representations of scalograms have limitations such as signals cannot be reconstructed from the phase information and cross-terms appear as interferences in the T-F plane between patterns [11].

The third technique is HHT, which has two-step methods to analyse the non-stationary and nonlinear signal [12]. In HHT, the initial step is the Empirical Mode Decomposition (EMD) that parts the first sign into a limited number of intrinsic mode functions (IMFs). In the following stage, the Hilbert Transform (HT) that creates an orthogonal pair is the second step of HHT. A few researchers have analyzed the HHT based EEG signal. Li et al., [13] employed the approach for extracting features and pattern recognition of ictal EEG was suggested using analytical mode and Support Vector Machine (SVM). Wang et al., [14] proposed a methodology to define and decompose non-stationary and non-linear objects, for example by combining CNN with HHT. The HHT provides more accurate time spectral representations in nonstationary signals [15] and the findings are presented in time-frequency-energy space [16] and are based on an adaptive basis.

In this paper, emotions are classified from EEG signals by employing the three time series to image encoding techniques in CNN. Furthermore these models are evaluated on the basis of their accuracy and performance.

The accompanying areas of this paper are organized as follows: section II clarifies the methodology. The result and discussion are found in segment III, and the conclusions are given in area IV.

## II. METHODOLOGY

Our proposed methodology has two parts, image encoding and classification of emotions. The time series signal is converted to an image in the first section and the second portion involves emotion classification by the use of deep learning. Fig. 1. shows the block diagram of methodology.

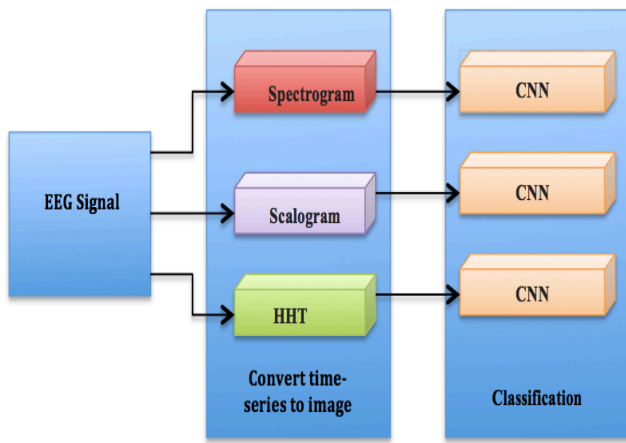


Fig. 1. Block diagram of methodology.

### A. Database

The dataset is taken from a seed database and is openly available [17]. The dataset is a deep belief network to construct an EEG based emotion recognition model for three emotions. It involved fifteen students from Shanghai Jiao Tong University (7 males and 8 females) with self-reported normal or corrected vision and normal hearing. Chinese film clips of positive, neutral, negative emotions are used to evoke their feelings. There are fifteen trials in an experiment and it consists of five for positive, five for neutral and five for negative. EEG data was recorded according to the international 10-20 schemes, from a 62-channel electrode cap. In this methodology, preprocessed data is collected from the database. It comprises three emotional classifications, such as positive, neutral and negative, and data is down sampled to 200 Hz. The EEG information was prepared with a 0.3 to 50Hz band pass channel. Here, Fig. 2. shows the EEG signal of 2s segment in the dataset. EEG signal, x-axis reflects time of 2s and y-axis represents amplitude of the signal.

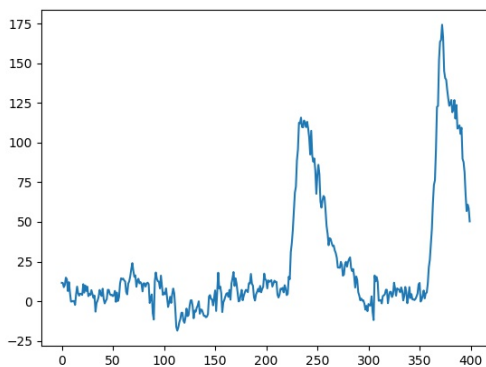


Fig. 2. EEG signal of 2s segment.

### B. Image Encoding

Image encoding is done using three techniques such as Spectrogram, Scalogram and HHT for converting time series signals to images.

a) Spectrogram: The transform is a visual representation of signal strength by time at different frequencies. In the visual representation, the x-axis represents time and y-axis addresses frequency. The third dimension defines the amplitude at a particular time. The colour intensity at each point reflects the signal amplitude. Spectrogram [18] was determined from the time signal by utilizing Short Time Fourier Transform (STFT). The EEG data is split into segments of nonoverlap. The segments are then passed through the Fourier transform to obtain the spectrogram. Every vertical line shows the estimation portions of amplitude versus frequency briefly as expected. The vertical lines are laid together to shape a three-dimensional surface image of the spectrogram.

Equation (1) shows the spectrogram that can be determined by processing the square size of the STFT itself.

$$\text{Spectrogram}(n, k) = |\text{STFT}(n, \omega)|^2 \quad (1)$$

$$\text{STFT}(n, \omega) = \sum_{-\infty}^{\infty} y[x] \omega[n-x] \cdot e^{-j\omega n} \quad (2)$$

where equation (2) represents the STFT of the signal.

In this work, single channel EEG signals is considered from the spectrogram. First, it splits the EEG signal into smaller parts. Here the EEG signal is divided using hanning window technique into 2s segments. To each of the smaller segments, STFT is then applied. As a result, the signal spectrogram can be computed from the STFT. Fig. 3. represents the spectrogram of the EEG signal.

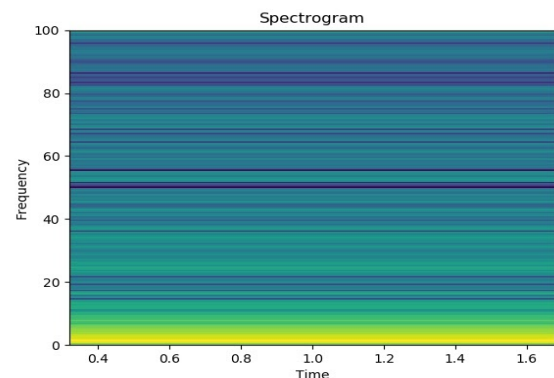


Fig. 3. Spectrogram of EEG signal.

b) Scalogram: The squared amplitude of the Continuous Wavelet Transformation (CWT) is the scalogram [19]. The CWT goes through a function known as the mother wavelet. The relation between scales and frequencies can be used in the conversion process to construct a time-frequency map from the scalogram, which is both scaled and shifted. At high frequencies, Scalogram provides short time interval windowing and long time interval windowing at low frequencies. CWT has the ability to divide windows at different sizes, allowing it to best analyse the high and low frequency information in the time series. Therefore, this technique is used for high frequencies on a small scale and low frequencies on a high scale for better resolution.

The scalogram's mathematical expression follows equation (3).

$$W_x(s, \tau) = \frac{1}{\sqrt{s}} \int_{-\infty}^{\infty} x(t) * \psi\left(\frac{t-\tau}{s}\right) dt \quad (3)$$

where  $x(t)$  represents the time signal,  $\psi(t)$  is wavelet function.  $s$  represented the scale and  $\tau$  represented the position parameter. Fig. 4. show the EEG signal scalogram image.

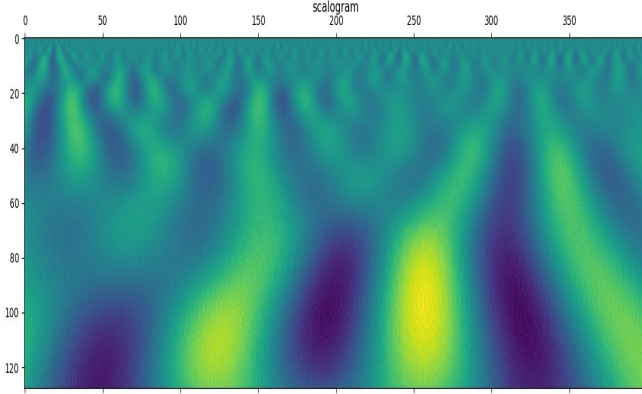


Fig. 4. Scalogram images of EEG signal.

c) Hilbert-Huang Transform: It is a non-stationary and non-linear analysis of signals. It requires two steps, such as the EMD and HT. Firstly, The original EEG signal is split into number finite functions in EMD. IMF for signal division captures the largest frequency portion of each event in the signal. Later process, HT produces a couple for every IMF, which is phase, moved to 90 degrees [20]. These techniques help to measure the instantaneous difference in amplitude and frequency as for time.

Mathematically, the original signal  $X(t)$  can be given by the equation (4).

$$X(t) = \sum_{i=1}^n c_i(t) + r_n(t) \quad (4)$$

where  $i$  represent the number of the corresponding IMF in  $c_i(t)$  and residue portion represents  $r_n(t)$ . To get instantaneous frequency  $\omega_i(t)$  and amplitude  $a_i(t)$  initially the signal is decomposed into IMFs, later the HT can be performed on each IMF.

HT of every IMF component  $c_i(t)$  is given in equation (5).

$$u_i(t) = \frac{1}{\pi} P \int_{-\infty}^{\infty} \frac{c_i(\tau)}{t-\tau} d\tau \quad (5)$$

where, Cauchy Principal Value represents  $P$ .  $c_i(t)$ ,  $u_i(t)$  can be combined to form the analytic signal  $z_i(t)$ , which has been explained in equation (6).

$$z_i(t) = c_i(t) + ju_i(t) = a_i(t) e^{j\theta_i t} \quad (6)$$

where,

$$a_i(t) = \sqrt{c_i^2(t) + u_i^2(t)} \quad \text{and} \quad \theta_i(t) = \arctan \frac{u_i(t)}{c_i(t)}$$

and instantaneous frequency is

$$\omega_i(t) = \frac{d\theta_i(t)}{dt}$$

The EEG data  $X(t)$  can be expressed as a form of an equation by applying the HT to the segments of every IMF.

$$X(t) = \Re \sum_{i=1}^n a_i(t) e^{j\int \omega_i(t) dt} + r_n(t) \quad (7)$$

where equation (7) represents the three-dimensional representation of instantaneous frequency and amplitude as a function of time. The hilbert amplitude spectrum  $H(w, t)$  was designed from the three dimensional distribution on the time-frequency plane. Fig. 5. shows the HHT spectrum of a single channel EEG signal. Time is expressed by the x-axis, frequency is expressed by the y-axis and colour expresses energy.

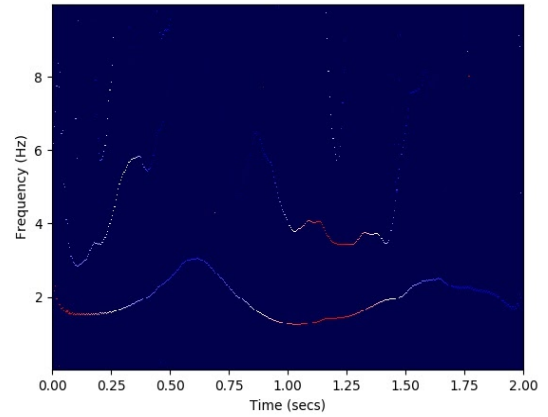


Fig. 5. HHT images of EEG signal.

### C. Deep Learning

a) Convolutional Neural Network: It is a subset of deep learning network, which is commonly used for image recognition [21]. CNN is a very effective and fine-tuned model that is ideal for a wide variety of applications [22]. In this work, RGB image is considered for CNN classification because it gives a better choice of what kind of features can be detected by the CNN and provides better classification accuracy [23]. The CNN architecture shown in Fig. 6. includes several layers, such as the layer of convolution, pooling layer, and completely linked layer. The convolution layer performs input extraction via the convolutional kernel function. Sigmoid activation function is used for enhancing the speed of convergence here and used 50 epochs for training the dataset for better performance. The pooling layer down, samples the function acquired from the convolution layer. Yield from the pooling layer is passed across the flatten and fully connected layer. Softmax layer gives the last yield of the CNN.

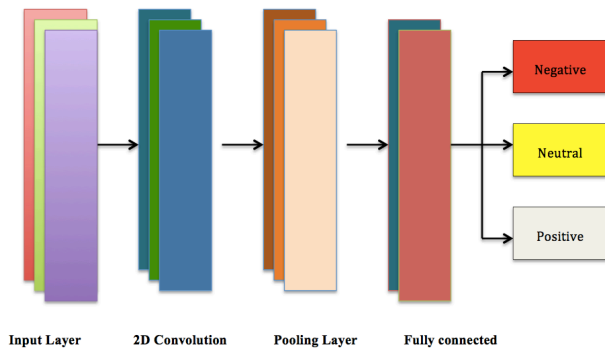


Fig. 6. Architecture of CNN

Table I. contains the specifications of the parameters used to train the Spectrogram, Scalogram, and HHT.

TABLE I. Hyper parameter of Network

Hyper parameters	Parameter
Learning rate	0.0001
Strides	2
Loss Function	Sparse Categorical Cross entropy
Pooling	MaxPooling
Epochs	50

### III. RESULT AND DISCUSSION

In all the three approaches examined in the previous section, 70% of the images are used for training and rest were utilised for testing. From the EEG signal, Spectrogram, Scalogram, and HHT were produced. The overall accuracy obtained for EEG based emotion classification from spectrogram-CNN is 78%. HHT-CNN and scalogram-CNN accuracy obtained is 75% and 98% respectively for emotion classification from CNN. In this work, the scalogram shows the best accuracy for emotion classification compared to other two techniques. Although, at low frequencies, it offers a long time window and a short time window at high frequencies. The training and validation accuracy of the Scalogram are shown in Fig. 7 and losses is shown in Fig. 8.

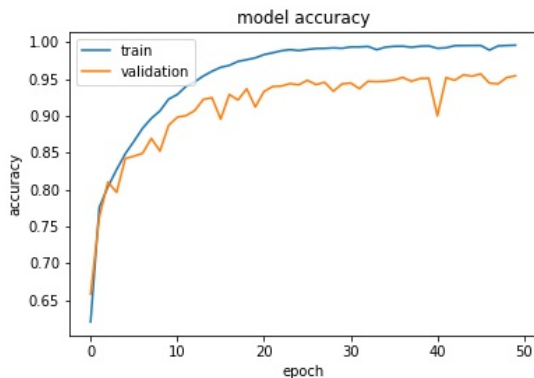


Fig. 7. Training and Validation accuracy of Scalogram

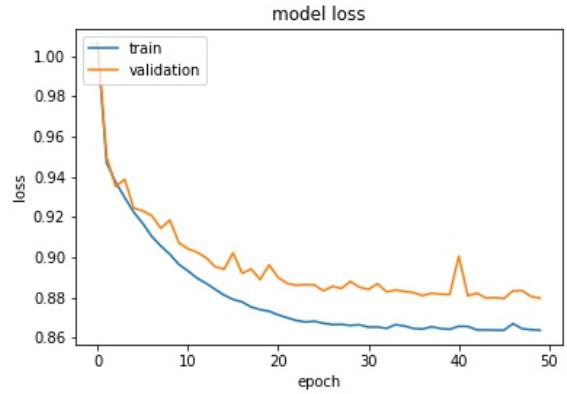


Fig. 8. Training and Validation loss of Scalogram

The training and testing accuracy of three techniques are compared as shown in Table II.

TABLE II. Accuracy and Loss of three techniques

Approach	Training		Testing	
	Accuracy %	Loss	Accuracy %	Loss
Spectrogram	86	0.913	78	0.941
HHT	90	0.904	75	0.961
Scalogram	99	0.863	98	0.867

Recall, Precision, and F1 Score performance of three approach measures are shown in Table III.

TABLE III. Performance of models

Approach	Precision	Recall	F1 Score
Spectrogram	78.6	78	78.3
HHT	74.6	75	74.3
Scalogram	98	98.3	98.3

### IV. CONCLUSION

In this paper, human emotions contained in the SEED dataset have been classified. Various types of image-encoding approaches- Spectrogram, Scalogram, and HHT have been utilized in the proposed system. Human emotions have been classified, and features of emotions have been extracted by using a deep learning model. The experiment has concluded that the proposed system, Scalogram-CNN,

outperformed the other techniques, Spectrogram-CNN and HHT-CNN.

#### REFERENCES

- [1] Kobayashi, Hiroshi, and Fumio Hara. "Recognition of six basic facial expression and their strength by neural network." In [1992] Proceedings IEEE International Workshop on Robot and Human Communication, pp. 381-386. IEEE, 1992.
- [2] Varghese, Ashwini Ann, Jacob P. Cherian, and Jubilant J. Kizhakkethottam. "Overview on emotion recognition system." In 2015 International Conference on Soft-Computing and Networks Security (ICSNS), pp. 1- 5. IEEE, 2015.
- [3] Gopalakrishnan, Athira, K. P. Soman, and B. Premjith. "Part-of-Speech Tagger for Biomedical Domain Using Deep Neural Network Architecture." In 2019 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT), pp. 1-5. IEEE, 2019.
- [4] Parsons, Stuart, Arjan M. Boonman, and Martin K. Obrist. "Advantages and disadvantages of techniques for transforming and analyzing chiropteran echolocation calls." *Journal of Mammalogy* 81, no. 4 (2000): 927-938
- [5] Yuan, Longhao, and Jianting Cao. "Patients' EEG data analysis via spectrogram image with a convolution neural network." In International Conference on Intelligent Decision Technologies, pp. 13-21. Springer, Cham, 2017.
- [6] Ruffini, Giulio, David Ibanez, Marta Castellano, Laura Dubreuil-Vall, Aureli Soria-Frisch, Ron Postuma, Jean-Francois Gagnon, and Jacques Montplaisir. "Deep learning with EEG spectrograms in rapid eye movement behavior disorder." *Frontiers in neurology* 10 (2019).
- [7] Garg, Divya, and Gyanendra K. Verma. "Emotion Recognition in Valence-Arousal Space from Multi-channel EEG data and Wavelet based Deep Learning Framework." *Procedia Computer Science* 171 (2020): 857-867.
- [8] Li, Xiang, Dawei Song, Peng Zhang, Guangliang Yu, Yuexian Hou, and Bin Hu. "Emotion recognition from multi-channel EEG data through convolutional recurrent neural network." In 2016 IEEE international conference on bioinformatics and biomedicine (BIBM), pp. 352-359. IEEE, 2016.
- [9] Kaya, Duygu. "The mRMR-CNN based influential support decision system approach to classify EEG signals." *Measurement* 156 (2020): 107602
- [10] Turk, Omer, and Mehmet Sirac Ozerdem. "Epilepsy detection by using scalogram based convolutional neural network from EEG signals." *Brain sciences* 9, no. 5 (2019): 115.
- [11] Rioul, Olivier, and Martin Vetterli. "Wavelets and signal processing." *IEEE signal processing magazine* 8, no. 4 (1991): 14-38.
- [12] Phan, Sonal K., and Cathy Chen. "Big Data and Monitoring the Grid." In *The Power Grid*, pp. 253-285. Academic Press, 2017.
- [13] Li, Shufang, Weidong Zhou, Qi Yuan, Shujuan Geng, and Dongmei Cai. "Feature extraction and recognition of ictal EEG using EMD and SVM." *Computers in biology and medicine* 43, no. 7 (2013): 807-816.
- [14] Wang, Shuang, Bin Guo, Chenjie Zhang, Xuemei Bai, and Zhijun Wang. "EEG detection and de-noising based on convolution neural network and Hilbert-Huang transform." In 2017 10th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI), pp. 1-6. IEEE, 2017.
- [15] Deng, Siyi, Ramesh Srinivasan, Tom Lappas, and Michael D'Zmura. "EEG classification of imagined syllable rhythm using Hilbert spectrum methods." *Journal of neural engineering* 7, no. 4 (2010): 046006.
- [16] Huang, Norden E. "Introduction to the Hilbert–Huang transform and its related mathematical problems." In *Hilbert–Huang transform and its applications*, pp. 1-26. 2014.
- [17] Zheng, Wei-Long, and Bao-Liang Lu. "Investigating critical frequency bands and channels for EEG-based emotion recognition with deep neural networks." *IEEE Transactions on Autonomous Mental Development* 7, no. 3 (2015): 162-175.
- [18] Li, Penghua, Minglong Chen, Fangchao Hu, and Yang Xu. "A spectrogram-based voiceprint recognition using deep neural network." In *The 27th Chinese Control and Decision Conference (2015 CCDC)*, pp. 2923-2927. IEEE, 2015.
- [19] Türk, Ömer, and Mehmet Siraç Özerdem. "Epilepsy detection by using scalogram based convolutional neural network from EEG signals." *Brain sciences* 9, no. 5 (2019): 115.
- [20] Tang, Shifu, Hong Ma, and Liyun Su. "Filter principle of Hilbert-Huang transform and its application in time series analysis." In 2006 8th international Conference on Signal Processing, vol. 4. IEEE, 2006.
- [21] Rajkumar, A., M. Ganesan, and R. Lavanya. "Arrhythmia classification on ECG using Deep Learning." In 2019 5th International Conference on Advanced Computing & Communication Systems (ICACCS), pp. 365-369. IEEE, 2019.
- [22] Saiharsha, B., B. Diwakar, R. Karthika, and M. Ganesan. "Evaluating Performance of Deep Learning Architectures for Image Classification." In 2020 5th International Conference on Communication and Electronics Systems (ICCES), pp. 917-922. IEEE, 2020.
- [23] Sachin, Rajan, V. Sowmya, D. Govind, and K. P. Soman. "Dependency of various color and intensity planes on CNN based image classification." In *International symposium on signal processing and intelligent recognition systems*, pp. 167-177. Springer, Cham, 2017.

# Novel Method for Analyzing the Relation of miRNA Seed Length and the miRNA-mRNA Interaction Strength

Binthiya Suny Gabriel

Department of Electronics, CUSAT  
Cochin University of Science and  
Technology  
Kochi, India  
binthiya\_gabriel@hotmail.com

Tessamma Thomas

Department of Electronics, CUSAT  
Cochin University of Science and  
Technology  
Kochi, India  
tessamma1@gmail.com

**Abstract**— microRNAs have been recognized as an important regulator of gene expression among a number of species. Although analysis done in genomics, on a comparative basis, indicate that miRNAs have many targets, how these targets respond to the miRNA targeting at the seed region, have not been fully explored. The application of correlation method for identifying the breast cancer specific seed region of miRNA for two breast cancer specific mRNAs BRCA1 and ERBB2 is presented in this paper. The results obtained indicated a peak at the point of maximum coincidence between miRNAs and its target. A nearly linear relationship between the miRNA seed length and the correlation strength between the mRNA and its corresponding miRNAs was also observed.

**Keywords**- binding region, correlation strength, normalised correlation, seed length, seed region

## I. INTRODUCTION

In the past three decades, cancer was considered to be a result of changes in the expression of genes that code for proteins. MicroRNAs have recently been found to associate with the formation of tumour by either acting as a suppressor or an oncogene [1,2].

MicroRNAs regulate the expression of a gene by binding to the complementary sites in the 3'-UTRs of mRNAs. This interaction between miRNA and the target is importantly influenced by the seed region. The most significant prerequisite for proper repression, is a perfect seed base pairing. The strength of base pairing contributes to the stability of the miRNA-target duplex and is also important in miRNA repression [3].

In this study, the strength of the base pairing is expressed in terms of normalized correlation between the 3'UTR of the target gene and its corresponding miRNAs. The study also includes finding a relationship between the length of the seed sequence and the strength of the correlation.

## II. LITERATURE STUDIES

The interaction between a miRNA and an mRNA is affected by the strength of binding between the two at the seed region. Researchers have suggested that longer seed regions have higher efficacy on the repression of mRNA. Tests were done to evaluate the hypothesis by calculating the differential expressions of all the miRNAs that target the mRNA. This study did not find a linear relationship between

the length of the seed and expression profiles of miRNAs [4,5].

SVM based binary classification method has been developed to predict the targets of miRNA. This approach also uses an approach to model both the canonical and non-canonical matches in the seed. A superior performance was achieved by making seed enrichment metric [6].

Unsupervised computational methods have been used to find out the seed regions without considering the miRNA biology. Though the method successfully found several known seeds, its novelty could not be considered as a proper evidence for a sequence to be considered as a seed [7].

A very much simplified method for studying the miRNA-mRNA interaction, chimera PCR, has obtained considerable results in detecting specific miRNA-mRNA interactions. However, this study claims that seed region is not sufficient for target interactions that are functional in nature [8].

The seed region/sequence plays a very significant role in analysing the extent of interaction between miRNA and mRNA towards cancer. Our study on seed region and its relation to the strength of correlation will help further in studying the characteristics of the breast cancer specific miRNA-target gene-TF disease network.

*As mentioned earlier, when some of the miRNAs bind with mRNAs, deadenylation may occur, leading to cancer. Our purpose is to find out such seed binding regions within the binding regions that are specific to breast cancer. The new method; correlation method, is used for finding these seed binding regions to then analyze the relationship between the seed length and the strength of correlation in the interaction between miRNA and mRNA.*

## III. METHOD USED

The importance of methods based on signal processing, is attributed to their application in collecting, processing and interpreting the information present in both genomics and proteomics data. The importance of genomic signal processing is rising as it has been recognized that the characterization of genomic and proteomic regulations require various disciplines related to signal processing [9-11].

A deeper understanding of the systems theory is needed for topics related to systems biology, which also entails



various methods of signal processing. Bioinformatics and Genomic Signal Processing both apply computational methods to tackle various problems related to biology. Several Digital Signal Processing techniques including STFT have found application in the search for genomic repeats and also in finding both the thermodynamic and bending properties of DNA using Fourier analysis [12-14].

Several DSP based algorithms have been applied to studying the characteristics of DNA and RNA sequences. However, the solutions provided by these algorithms include much background noise that is computationally complex and less accurate. Additional techniques need to be used in order to remove the noise [15-18].

When some of the miRNAs bind with mRNAs, degradation may occur, leading to cancer. Our purpose is to find out such binding regions and seed binding regions specific to breast cancer, without the presence of any background noise. Normalised Correlation method is used for this purpose.

#### A. Determination of mRNA binding region and the miRNA seed region

The steps involved in finding the binding region of miRNA to the mRNA, and the seed region are as follows:

1. The miRNA and the mRNA, whose binding region to the miRNA is to be found, are selected.
2. The normalized correlation between the reverse complimented miRNA and the mRNA, is calculated by shifting the miRNA by one base and finding the location of maximum correlation [19].
3. The region of maximum correlation/coincidence between mRNA and the reverse complimented miRNA is noted.
4. The region where at least 6 to 8 bases coincide between the two sequences is considered the seed region of the miRNA.
5. The graph of normalized correlation versus the number of times the reverse complimented miRNA sequence is shifted along the length of the mRNA sequence is plotted.
6. The above steps are repeated for all the miRNAs that target the breast cancer specific mRNA.
7. The relationship between the seed length and the correlation strength values for BRCA1 is also plotted.

## IV. DATABASE

The lists of the miRNAs that target an mRNA are obtained from the exiqon website [<https://www.exiqon.com/miRSearch>] and the miRNA sequences are obtained from the miRBase website [<https://www.mirbase.org/>]. The 3'UTR of the breast cancer specific mRNAs are obtained from the UCSC Genome Browser website [<https://www.genome.ucsc.edu/>]. The breast cancer specific mRNAs BRCA1, ERBB2 and the miRNAs that target both these mRNAs, were used for analysis, in this study.

## V. IMPLEMENTATION OF THE METHOD

The analysis to find a relation between the correlation strength values and the number of shifts was done using the normalized correlation method. The normalized correlation

between the 3'UTR of breast cancer specific mRNA, BRCA1, and the reverse complimented sequence of miRNA, hsa-miR-6720-5p was done to first obtain the region of maximum correlation/binding region and later to obtain the seed binding region. The method was also repeated for another breast cancer specific mRNA, ERBB2, with the same miRNA.

#### A. Illustrating Example 1

Input mRNA: BRCA1 (3'UTR: 1383 Bases)

Input miRNA: hsa-miR-6720-5p (23 Bases)

BRCA1 or Breast Cancer 1 is a gene that usually plays a role in restraining the growth of cells in the breast. However, a mutation in this gene predisposes to breast cancer [20]. In this example, as shown in Table I, the seed sequence obtained is GGCTGGAA.

TABLE I. CORRELATION AND LATERAL SHIFT DETAILS FOR hsa-miR-6720-5p (BRCA1)

Sl. No.	miRNA Name and Length	Actual Seed Binding Region wrt mRNA, Sequence	miRNA Seed Region	Maximum Correlation value at the binding region	Binding Region of mRNA	No. of shifts needed to reach maximum coincidence
1	hsa-miR-6720-5p (23)	339-346, GGCTGGAA	1-8	0.9836	324-346	323

This seed region is an 8-mer/1-8 length seed. The reverse complimented miRNA sequence when shifted along the BRCA1 3'UTR sequence, gives maximum correlation strength value at the 323<sup>rd</sup> shift. Figure 1 show how the correlation strength value gradually increases till the maximum value of 0.9836 is reached and later decreases.

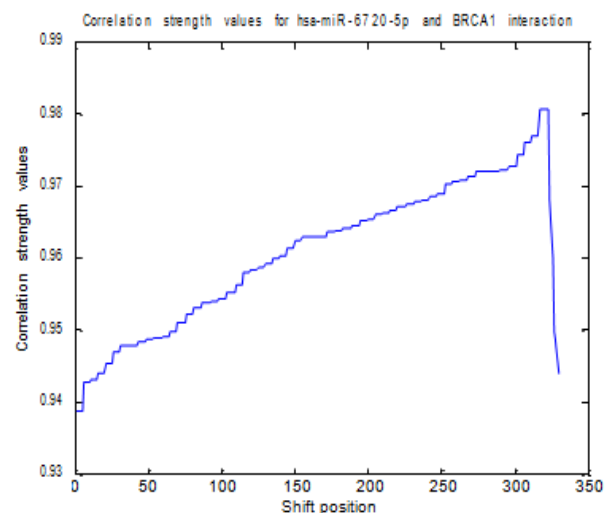


Fig. 1. Correlation strength values versus the no. of shifts of hsa-miR-6720-5p for BRCA1

ERBB2 (erbB-2) is found to be overexpressed in breast cancer up to around 30%. This gene is known to encode a member of the EGF (Epidermal Growth Factor) receptor family.

#### B. Illustrating Example 2

Input mRNA: ERBB2 (3'UTR: 618 Bases)

Input miRNA: hsa-miR-6720-5p (23 Bases)

From the analysis of Table II and figure 2, it was found that the binding position of the mRNA occurs at the maximum correlation position and hence maximum correlation position was chosen as the point of maximum correlation.

TABLE II. CORRELATION AND LATERAL SHIFT DETAILS FOR hsa-miR-6720-5p (ERBB2)

Sl. No.	miRNA Name and Length	Actual Seed Binding Region wrt mRNA, Sequence	miRNA Seed Region	Maximum Correlation value at the binding region	Binding Region of mRNA	No. of shifts needed to reach maximum coincidence
1	hsa-miR-6720-5p (23)	168-175 ,GGCTGGAA	1-8	0.9821	153-175	152

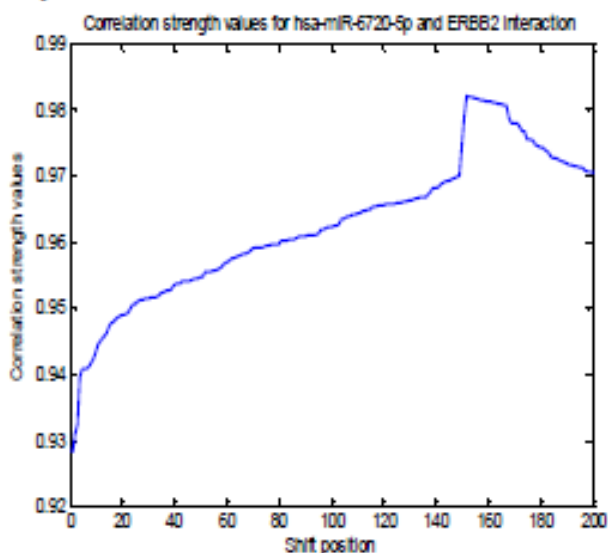


Fig. 2. Correlation strength values versus the no. of shifts of hsa-miR- 6720-5p for ERBB2

For the validation of the results obtained, the analysis was extended to two other breast cancer specific mRNAs, BRCA2 and EGFR. BRCA2 is also a tumor suppressor gene, and is found in all humans. The BRCA2 protein, also known by the name breast cancer type 2 susceptibility protein, is responsible for repairing the DNA. BRCA1 and BRCA2 are expressed in the breast cells where they together help in the repair of damaged DNA or destroy the cells if the DNA cannot be repaired [21].

The figure 3 shows the variation in the correlation strength values with the number of shifts for 3 miRNAs; hsa-miR-146a-5p, hsa-miR-513c-5p, and hsa-miR-514b-5p with maximum correlation values of 0.9051, 0.835 and 0.8 respectively.

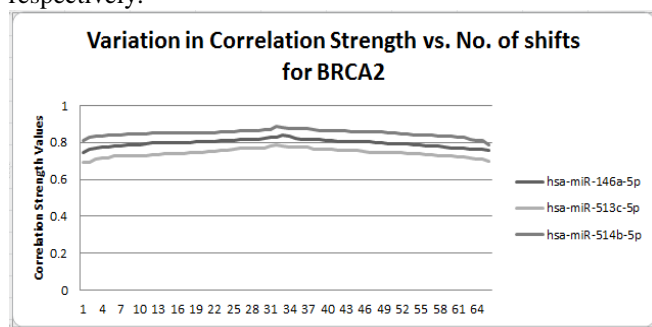


Fig. 3. Variation in Correlation strength values with No. of shifts for BRCA2

The Epidermal Growth Factor Receptor (EGFR) is one among the first identified important targets of the antitumor agents. Approximately half the numbers of cases of triple negative breast cancer and also the inflammatory breast cancer have EGFR in the overexpressed state. Extensive research has indicated that EGFR targeted therapy might have promising roles in different types of breast cancer [22, 23].

Figure 4 shows the variation in the correlation strength values with the number of shifts for 3 miRNAs; hsa-miR-1287-5p, hsa-miR-127-5p, and hsa-miR-885-5p with maximum correlation values of 0.9901, 0.9937 and 0.9956 respectively.

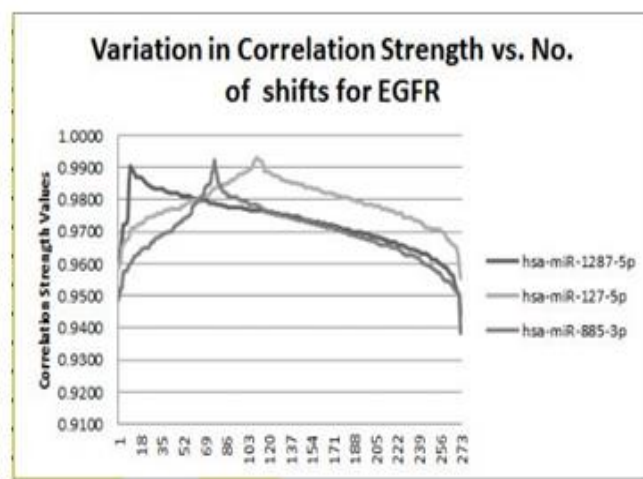


Fig. 4. Variation in Correlation strength values with No. of shifts for EGFR

## VI. RESULTS AND DISCUSSION

1. The observations made about the relationship between correlation strength and the number of shifts of the miRNA, was applied to breast cancer specific mRNAs BRCA1, ERBB2, BRCA2, and EGFR. The validation of the results obtained was done to 25 other breast cancer specific mRNAs to obtain similar results.

2. When it comes to expanding the research on the role of miRNA-miRNA interaction in the progression of breast cancer, the interaction between the types of seed (seed length) with the correlation strength values also paves a way to understanding the extremity of breast cancer. This new relationship was tested between the 3'UTR of BRCA1 and all the miRNAs that target this gene. The figure below shows the result.

Figure 5 shows the 38 miRNAs that target the 3'UTR of BRCA1. The area shaded in orange shows the maximum values of correlation strength between BRCA1 and each of the miRNAs. It can be noted that the highest values of correlation strength is associated with miRNAs having 8-mer seed. The miRNA hsa-miR-4506 (77 bases) is the only miRNA belonging to the 7-mer-A1/1-7 seed group. 7-mer m8 and 6-mer seeds have correlation strength values much lesser than the 8-mer and 7-mer-A1 seeds.

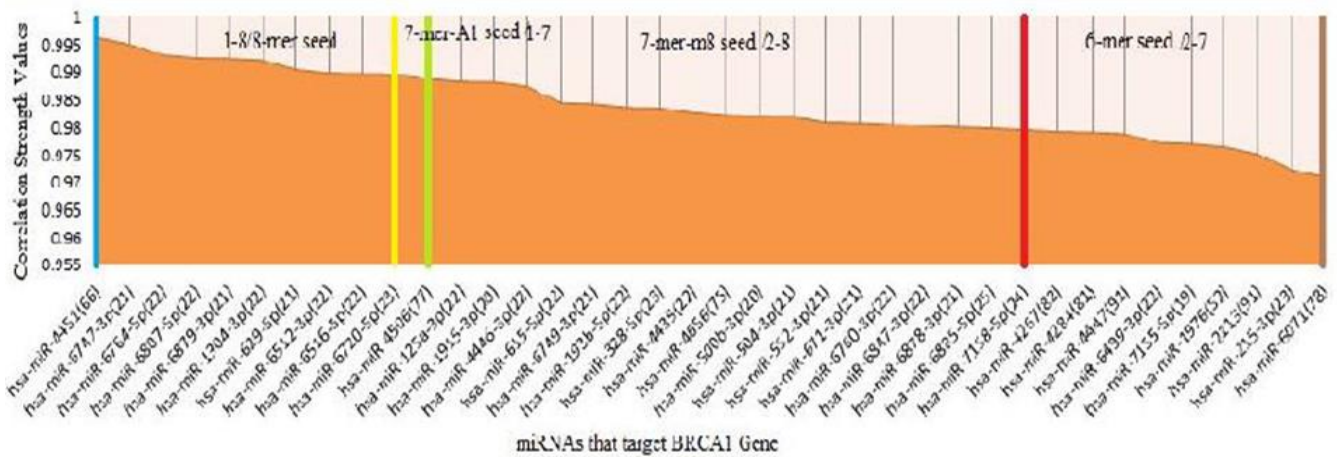


Fig.5. Correlation Strength Values vs. Seed type for BRCA1

VII CONCLUSION

Cancer involves changes in the expression and structure of genes. Mutations in miRNAs lead to abnormalities thus stopping the cells from attaining the fully differentiated form. The seed binding regions of miRNAs specific to the breast cancer specific mRNAs BRCA1, BRCA2, ERBB2 and EGFR were identified using the correlation method and a nearly linear relationship was found between the value of the binding strength (between the mRNA and the miRNAs that target the mRNA) and the seed type.

APPLICATION POTENTIAL OF THE WORK

Breast cancer is a more frequently seen cancer in women. The role of miRNAs in the progression of breast cancer provides a platform for further analysing the repression effects of miRNAs. The impact of miRNA on the expression of mRNA is determined by the seed region of miRNA. The application of normalised correlation in finding the seed region will help in determining the extent of complexity of the cells in cancer. The identification of seed region/sequence using other signal processing methods will be attempted for the purpose of comparing and validating the results obtained.

ACKNOWLEDGMENT

The authors gratefully acknowledge the support from the Department of Electronics, CUSAT.

REFERENCES

[1] Cornwell, A., Palli, R., and Singh, M.V., "Molecular characterization of atherosclerosis in HIV positive persons", *Scientific Reports*, Volume 11, 2021.  
 [2] Sarubala M., Sushmitha S., Antara B. and Surajit P., "Over-Expression of MicroRNA-122 Inhibits Proliferation and Induces Apoptosis in Colon Cancer Cells", *MicroRNA*, Volume 9, Issue: 354, 2020.  
 [3] A. Oulas, M. Reczko and P. Poirazi, "MicroRNAs and Cancer—The Search Begins!", *IEEE Transactions on Information Technology in Biomedicine*, Volume. 13, Issue 1, Pages. 67-77, March 2009.  
 [4] Wang L, "Correlation analyses revealed global microRNA-mRNA expression associations in human peripheral blood mononuclear cells", *MGG Molecular & General Genetics* Volume 293, Issue 1, Pages 95-105, September 2017.

[5] L. E. Mullany, J. S. Herrick, R. K. Wolff, and M. L. Slattery, "MicroRNA Seed Region Length Impact on Target Messenger RNA Expression and Survival in Colorectal Cancer", *PLoS ONE* Volume 11, Issue 4, April 2016.  
 [6] A. Ghoshal, J. Zhang, M. A. Roth, K. M. Xia, A. Y. Grama and S. Chaterji, "A Distributed Classifier for MicroRNA Target Prediction with Validation Through TCGA Expression Data," in *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, Volume 15, Issue 4, Pages. 1037-1051, Aug 2018.  
 [7] Md Izhar Ashraf, and A. Anny Leema, "Identification of Micro-RNA Seed Sequences and Other Possible Conserved Motifs: An Information Theoretic Approach", *Journal of RNA and Genomics*, Volume 13, Issue 1, July 2017.  
 [8] James P.Broughton, and Michael T. Lovci, "Pairing beyond the seed Supports MicroRNA Targeting Specificity", *Molecular Cell*, Volume 64, Issue 2, Pages 320-333, October 2016.  
 [9] Hervé S. "Redefining MicroRNA Targets", *Current Biology*, Volume 19, Issue 10, 2009.  
 [10] Wan JH, and Hsiuying W, "Human microRNA target identification by RRSM", *Journal of Theoretical Biology*, Volume 286, Issue 1, Pages 79-84, 2011.  
 [11] Tarang S, and Weston MD, "Macros in microRNA target identification: A comparative analysis of in silico, in vitro, and in vivo approaches to microRNA target identification", *RNA Biology*, Volume 11, Issue 4, Pages 324-333, 2014.  
 [12] Younger ST, and Corey DR, "Identification and validation of miRNA target sites within nontraditional miRNA targets" *Methods in Molecular Biology*, Volume 1206, Pages 53-67, 2015.  
 [13] Govindarajan R, Duraiyan J, Kaliyappan K, and Palanisamy M, "Microarray and its applications", *Journal of Pharmacy and Bioallied Sciences*, Volume 4, Suppl 2, Pages 310-312, 2012.  
 [14] Li Z, Peng Z, Gu S, Zheng J, Feng D, Qin Q, and He J, "Global Analysis of miRNA-mRNA Interaction Network in Breast Cancer with Brain Metastasis", *Anticancer Research*, Volume 37, Issue 8, Pages 4455-4468, 2017.  
 [15] Plotnikova O, Baranova A, and Skoblov M, "Comprehensive Analysis of Human microRNA-mRNA Interactome", *Frontiers in genetics*, Volume 10, 2019.  
 [16] Plotnikova OM, and Skoblov MY, "Efficiency of the miRNA-mRNA Interaction Prediction Programs", *Molecular Biology*, Volume 52, Issue 3, Pages 543-554, 2018.  
 [17] Binthiya SG, and Tessamma T, "miRNA-mRNA Binding And Seed Binding Region Detection Of Breast Cancer Specific mRNA Using Correlation Method", *Journal of Emerging Technologies and Innovative Research (JETIR)*, Volume 6, Issue 2, February 2019.  
 [18] Savage KI, and Harkin DP, "BRCA1, a 'complex' protein involved in the maintenance of genomic stability", *The FEBS Journal*, Volume 282, Issue 4, Pages 630-646, 2015.  
 [19] Adrienne GW, and Eric PW, "Chemotherapy and HER2-Directed Therapy for Metastatic Breast Cancer", *Oncology (Williston Park)*, Volume 27, Issue 3, Pages 166-175, 2013.  
 [20] Masuda H, and Zhang D, Bartholomeusz C, Doihara H, Hortobagyi GN, Ueno NT, "Role of epidermal growth factor

receptor in breast cancer”, Breast Cancer research and treatment, Volume 136, Issue 2, Pages 331-345, 2012.

- [21] Laura BC, and Amy EP, ”miRNA Targeting:Growing beyond the Seed”, Trends in genetics, Volume 35, Issue 3, Pages 215-222, 2019.
- [22] Wen-Juan N, and Xiao-Min L, ”Dynamic miRNA–mRNA paradigms: New faces of miRNAs”, Biochemistry and biophysics reports, Volume 4, Pages 337-341, 2015.
- [23] Singh R, and Yin-Yuan M, ”Role of microRNAs in breast cancer”, Cancer biology and therapy, Volume 14, Issue 3, Pages 201-212, 2013.

# TEXTURE ANALYSIS OF REACTION DIFFUSION LEVEL SET EVOLUTION OF MULTIPLE SCLEROSIS LESIONS IN BRAIN MR IMAGES

Pandian Ambairam  
Electronics and Communiatiion  
Engineering  
SRM Valliammai Engineering College  
SRM nagar, chengalpattu,Tamil  
Nadu,India  
pandiana.ece@valliammai.co.in

Udhayakumar Ganesan  
Electronics and Communiatiion  
Engineering  
SRM Valliammai Engineering College  
SRM nagar,chengalpattu,Tamil  
Nadu,India  
udhayakumarg.ece@valliammai.co.in

## Abstract

*Multiple sclerosis (MS) is a chronic idiopathic disease that results in multiple areas of inflammatory demyelination in the Human Central Nervous System (CNS) and it causes various trouble with mobility and upper limb function, bladder, vision, speech, swelling, and cognition. Hence, we need early detection and accurate identification of pathological changes of disease progression for MS patients. In this work, automated segmentation method of Reaction-Diffusion Level Set Evolution (RDLSE) is employed for T2-Weighted (T2W) Magnetic Resonance images. Axial view of MR images of Various MS Patients based on Extended Disability Status Scale (EDSS) of the University Medical Centre Ljubljana are used in this analysis. The segmented output image accuracy is validated with Ground Truth images. Texture analysis are employed to Segmented MS region in MR images to improve the accuracy and identification of differences in brain tissue structure. Results show RDLSE methods are able to segment the small MS lesions even in the presence of heterogeneous intensity values and segmented output image feature value show the MS lesion load for various EDSS values in MS patient T2W MR Images.*

**Keywords**—Multiple sclerosis, Magnetic Resonance Imaging, White Matter Lesions, Reaction-Diffusion Level Set Evolution (RDLSE), Texture analysis.

## I. INTRODUCTION

Multiple sclerosis (MS) is a chronic Idiopathic disease that results in multiple areas of inflammatory demyelination in the Human Central Nervous System (CNS) and the brain[1]. The MS disease causes the immune system to attack axons, causing the myelin sheath resulting in conduction to obstruct which leads to permanent loss of brain function. The destruction of the myelin sheath leads to impaired communication between nerve cells and neurological symptoms such as abnormal sensation vision problems, and upper limb function, bladder, vision, speech, swelling, and cognition [2].Extended Disability Status Scale is used to measure the level of disability for MS diseases and also monitor the variations in the level of disability over time [3]. The MS affected brain tissue is detected with noninvasive biomarkers method to assess the importance of disease management and improve treatment effectiveness [4,5]. Magnetic Resonance Imaging (MRI) is the perceptive noninvasive imaging modality for diagnosing of MS diseases and monitoring its progression in various time

periods[6].The MS diseases affected area are shown in the periventricular white matter, internal capsule, corpus callosum of brain region and in the form of demyelination, gliosis, and edema. The MRI of T2 Weighted (T2W) images are used to represent high signal intensity due to deviation in volume of White matter intensity. MS lesions of white matter signal abnormalities (WMSA) are measured in MR brain image using interactive detection and delineation by a radiologist with CAD tools [7].The manual segmentation methods have a lot of demerits like more time consuming, huge intra expert and inter expert variability to detect and validate the myelin loss. Due to overlapping of pixel intensity in MRI brain tissues leads to misclassification. Hence, many researchers are developed many semi-automatic[8] and automatic segmentation [9] methods for MS analysis. The automated segmentation algorithm faces challenges for the Identification of small lesions is more complex, because of noise and heterogeneity present in MR images. Many researchers have used level set methods based on Partial Differential Equations (PDE) to segment the white matter lesions even in the presence of intensity inhomogeneity. Texture-based analysis are used for the extraction of diagnostically meaningful information to discriminate the between Normal and Abnormal-appearing white matter lesions in Brain MR images[10]. In this work, Texture Analysis of Reaction Diffusion Level Set Evolution (RDLSE) is employed to analyze the quantifying disability in Multiple sclerosis in T2W MR Brain images

## II. MATERIALA AND METHOD

### A. Methodology

In this work we used MR Image data set from the University Medical Centre Ljubljana UMCL[11] with Siemens 3 Tesla MRI scanner. The images are selected based on various EDSS value and Input image shown in Fig 1.The axial view of MRI scan and slices range from 98 to 119 are used for this study. The input axial view image are preprocessed with skull stripping method to remove unwanted nonbrain image[12,13]. Gaussian filter is employed to remove additive noise present in the skull stripped image [14,15] and subjected to RDLSE method for segmentation of white matter lesions.

### B. Reaction Diffusion Level Set Method

The Reaction Diffusion level set evolution is given by [16]

$$\begin{cases} \phi_t = \varepsilon \Delta \phi - \frac{1}{\varepsilon} L(\phi), x \in \Omega \subset R^n \\ \text{subject to } \phi(x, t = 0, \varepsilon) = \phi_0(x) \end{cases} \quad (1)$$

Where  $\varepsilon$  is always positive value,  $L(\phi) = -F\delta(\phi)$  for variational LSM,  $\Delta$  -laplacian operator and  $\phi_0(x)$  initial LSF. Diffusion term gradually regularizes the LSF to be piecewise constant every segment domain of given input image  $\Omega_i$ .

Geodesic Active Contour(GAC) of reaction diffusion level set evolution are used to evaluate the image curvature of lesions by Partial Differential Equation is given below

$$F = \text{div}(g(|\nabla I|)) \frac{\nabla \phi}{|\nabla \phi|} + v g(|\nabla I|) \dots (2)$$

where  $\frac{\nabla \phi}{|\nabla \phi|}$  is  $\vec{N}$ ,  $g(|\nabla I|)$  represents the stopping edge

function for contour evaluation on the image boundary,  $v$  is the constant force and assigned as 0.9, and denotes the Gaussian kernel set as 1.5 for image smoothening.

### C. Evaluation measures

The segmented MS lesion image is validated with the ground truth image. In this work, Comparing voxel to voxel in segmented and ground truth image and classified as a true positive (TP), false positive (FP), true negative (TN) or false negative (FN). To validate the segmentation and should get the maximum number of TPs and TNs, and the minimum number of FPs and FNs for perfect segmentation[17,18]. The segmented lesions are utilized to extract features which include Contrast and Dissimilarity. Contrast measures the amount of local variations present in an image.

### D. Results and Discussion

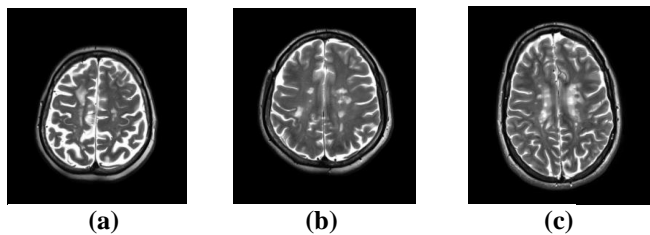


Fig.1 (a,b,c) T2 Weighted MR axial view images.

The axial view of MR images are recorded using T2W sequences from University Medical Center Ljubljana (UMCL) database, considered for this study. T2W images of MS lesions are shown in Fig.1. The MS lesion presented in both T2W sequences has higher pixel values. In T2W images, the MS lesions appear hyper-intensity signals that reflect lesions produced largely due to the loss of axon nerve fibers and demyelination. To remove unwanted brain image information by using Histogram based skull stripping process is adopted to remove non-brain tissue in the axial view of MR images as shown in Fig 2. The enhanced images are subjected to RDLSE methods to segment the MS lesions.

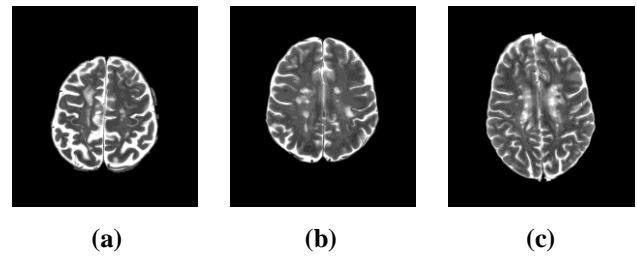


Fig.2 (a,b,c) T2 Weighted skull stripped and enhanced images

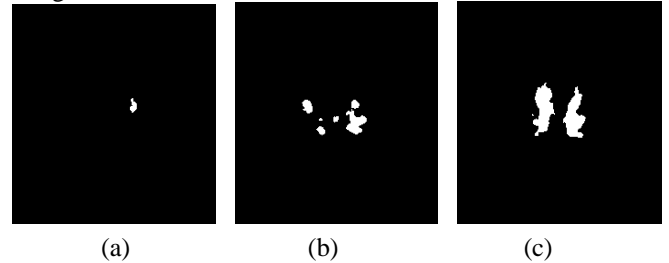


Fig.3 Segmented MS lesions using RDLSE (a,b,c) T2 Weighted Binary MR images.

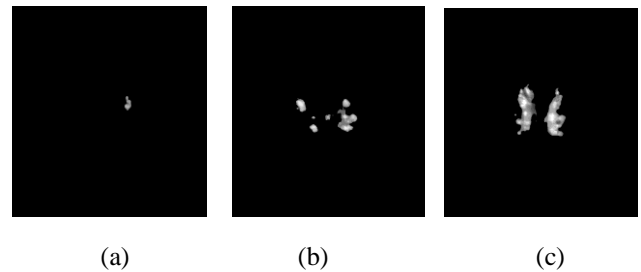


Fig.4 Segmented MS lesions using RDLSE (a,b,c) T2 Weighted Gray scale MR images

Figure 3 shows the segmentation of RDLSE method to T2W binary MR Images. This method used Zero level contour of the final LSF to detect the MS lesion curvature, which represents the boundaries of the WML of MRI sequences. The effect of the RDLSE method provides good performance on the boundary antileakage of MS lesions in MR images. For the quantitative assessment of the RDLSE method, ground truth can be considered as the reference to evaluate the segmentation results. The Dice Similarity (DS) values vary from 0.49 to 0.92 obtained for these segmentation methods. The RDLSE achieve very high DS values and indicates better similarity. This might be due to the edge function for contour evaluation on the image boundary. The values of performance measures are found to be high in RDLSE methods in the segmentation of MS lesion. This may due to the use of the reaction term in the diffusion equation to avoid boundary leakage problems.

The scattergram showing the variation of features are observed from segmented region of T2W MR brain images are shown in Figure 5. Area of Multiple sclerosis region using RDLSE method seems to better correlate with EDSS. The EDSS is a scale value used to quantify the disability in MS variation over time and indicates severity. The values of contrast derived from RDLSE method shown high correlation with EDSS. This could be used to characterize the MS disease progression of various

stages.

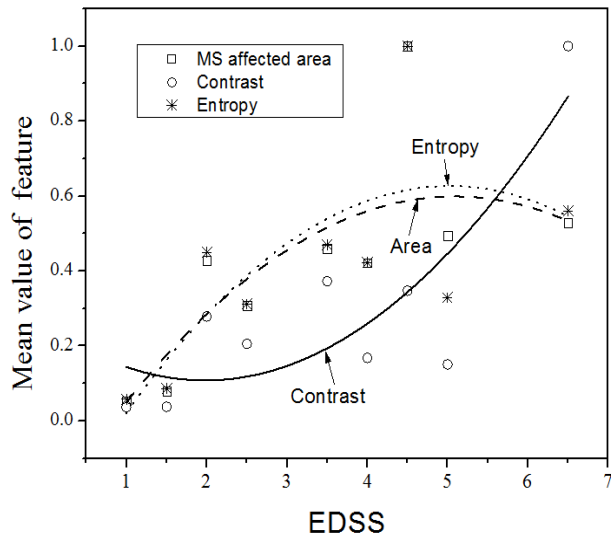


Fig.5 Variations of feature values with EDSS of White Matter Lesions of MR brain images

### E. Conclusion

In this work, we proposed Reaction-Diffusion Level Set Evolution based segmentation is employed to investigate the neurological disorder of MS diseases. RDLSE method used to detect the presence of White Matter Lesions in MR Brain images. Results have shown that the RDLSE method may be useful for the segmentation of MS lesions in T2W images features extracted from segmented region seems to better correlate with Extended Disability Status Scale and could be used to characterize the progression of various stages of MS diseases.

### FUNDING

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

### CONFLICT OF INTEREST

The authors report that they have no conflicts of interest. The submitted article is a reliable work of all mentioned authors.

### REFERENCES

- [1] Young IR, Hall AS, Pallis CA, Legg NJ, Bydder GM, Steiner RE. Nuclear magnetic resonance imaging of the brain in multiple sclerosis. *Lancet* 1981;ii:1063–1066.
- [2] Alan J. Thompson, Jeremy C. Hobart Multiple sclerosis: assessment of disability and disability scales, *J Neurol* (1998) 245:189–196 © Springer-Verlag 1998.
- [3] Patrick DL, Deyo RA (1989) Generic and disease-specific measures in assessing health status and quality of life. *MedCare* 27 [Suppl 3]:S217–232.
- [4] Kurtzke JF. A new scale for evaluating disability in multiple sclerosis. *Neurology* 1955; 5: 580–583.
- [5] Kurtzke JF. Rating neurologic impairment in multiple sclerosis: an Expanded Disability Status Scale (EDSS). *Neurology* 1983; 33: 1444–1452.
- [6] O. Yu, Y Mauss, G. Zollner, I.J. Namer, J. Chambron, “Distinct patterns of active and non-active plaques using texture analysis of brain NMR images in multiple sclerosis patients: Preliminary results”, *Magn. Reson. Imag.*, vol. 17, no. 9, pp. 1261-1267, 1999.
- [7] Filippi M, Horsfield MA, Ader HJ, et al. Guidelines for using quantitative measures of brain magnetic resonance imaging abnormalities in monitoring the treatment of multiple sclerosis. *Ann Neurol* 1998;43:499-506
- [8] A. Achiron, S. Gicquel, S. Miron, and M. Faibel, “Brain MRI lesion load quantification in multiple sclerosis: A comparison between automated multispectral and semi-automated thresholding computer-assisted techniques,” *Magn. Reson. Imag.*, vol. 177, pp. 85–106, 2001.
- [9] Roy, P. K., Bhuiyan, A., & Ramamohanarao, K. (2013). Automated segmentation of multiple sclerosis lesion in intensity enhanced flair MRI using texture features and support vector machine. 2013 IEEE International Conference on Image Processing.
- [10] Harrison LCV, Raunio M, Holli KK, Luukkaala T, Savio S, Elovaara I, et al. MRI texture analysis in multiple sclerosis: toward a clinical analysis protocol. *Acad Radiol* 2010;17:696–707.
- [11] Lesjak, Žiga, Franjo Pernuš, Boštjan Likar, and Žiga Špiclin. "A Novel Public MR Image Dataset of Multiple Sclerosis Patients With Lesion Segmentations Based on Multi-rater Consensus", *Neuroinformatics* (2017).
- [12] Kalavathi P, Prasath VB (2016) Methods on Skull Stripping of MRI Head Scan Images-a Review. *J Digit Imaging* 29(3):365-2-379.
- [13] Bauer S, Nolte L-P, Reyes M: Skull-stripping for tumor-bearing brain images. In Philippe Buchler and Stephen Ferguson, editors, Annual Meeting of the Swiss Society for Biomedical Engineering, page 2, Bern, April 2011. SSBE
- [14] Image Enhancement via Adaptive Unsharp Masking Polesel A Ramponi G Mathews IEEE transactions on image processing 2000 vol: 9 (3) pp: 505.
- [15] Nonlinear unsharp masking methods for image contrast enhancement Strobel N Journal of Electronic Imaging 2006 vol: 5 (3) pp: 353.
- [16] Zhang, K., Zhang, L., Song, H., & Zhang, D. (2013). Reinitialization-Free Level Set Evolution via Reaction Diffusion. *IEEE Transactions on Image Processing*, 22(1), 258–271
- [17] C. Gabriela, L. Diane and P. Florent, “What is a good evaluation measure for semantic segmentation,” in The British Machine Vision Conference, Bristol, 2013.
- [18] Thanh, Dang N. H., et al. “Melanoma Skin Cancer Detection Method Based on Adaptive Principal Curvature, Colour Normalisation and Feature Extraction with the ABCD Rule.” *Journal of Digital Imaging*, Springer Science and Business Media LLC, Dec. 2019.

# Heart Rate Monitoring Using Electrical Impedance

1<sup>st</sup> Yousef Alharbi

Department of Biomedical Technology,  
College of Applied Medical Sciences  
Prince Sattam bin Abdulaziz University  
Al-Kharj, Saudi Arabia  
y.alharbi@psau.edu.sa

2<sup>nd</sup> Anwar Alshrouf

Department of Biomedical Technology,  
College of Applied Medical Sciences  
Prince Sattam bin Abdulaziz University  
Al-Kharj, Saudi Arabia  
a.alshrouf@psau.edu.sa

3<sup>rd</sup> Sofiene Mansouri

Department of Biomedical Technology,  
College of Applied Medical Sciences  
Prince Sattam bin Abdulaziz University  
Al-Kharj, Saudi Arabia  
Laboratory of Biophysics and Medical  
Technologies,  
University of Tunis  
El Manar, Tunisia  
s.mansouri@psau.edu.sa

**Abstract**— Monitoring tissue composition could be persistent for many medical applications. Bioimpedance may be useful for monitoring several cardiovascular parameters such as cardiac output, stroke volume, blood pressure and heart rate. In addition, it could be useful for diagnosing the mitral insufficiency, stenosis, and other valvular heart diseases. The purpose of this study is to use the bioimpedance signal to determine the heart rate. This study involved 51 volunteers (28 male, 23 female) ranging from 20 to 76 years. A current of 1 mA at 30 kHz was injected through two outer electrodes placed on the subject's left arm and then the impedance was collected using two inner electrodes placed between the former outer electrodes. The acquired bioimpedance signal was pre-processed and used to identify the peaks and calculated the average interval. Thus, heart rates of each subject were calculated and then compared with actual measurement from FUZZY model MP3 digital blood pressure monitor. The results showed high correlation, 95.44 %, between the calculated heart rates and measurements. The various clinical applications of bioimpedance will help further promote the technique and several parameters can be determined using the electrical bioimpedance approach.

**Keywords**—Bioimpedance, electrode, cardiovascular, heart rate

## I. INTRODUCTION

The necessity of monitoring tissue composition could be persistent for many medical applications. The measurement of tissue's ability to resist electrical current, provides a non-invasive approach to monitor tissue properties and beyond diseases such as valvular diseases [1] and breast tumours [2]. This approach can be reached using the bioimpedance analysis by injecting a small amount of current on a specific region of human body.

Bioimpedance may be useful for diagnosis and treatment in different clinical situations such as dyspnea [3], hypertension, mechanical ventilation and dialysis [4]. In addition, monitoring of hemodynamic parameters during surgical operations [5] and changes in the cardiovascular system in pregnant women [6]. Bioimpedance can be used for diagnosing some cardiovascular pathologies such as heart failure, myocardial infarction and ischemia [7]. The medical follow-up of pregnant women throughout the course of the pregnancy [8], and the study of various diseases of the autonomic nervous system such as Parkinson's disease [9] also can be monitored.

In the case of cardiovascular applications, cardiac activity causes an increase in blood volume in the aorta and arteries during systole and a decrease in blood volume during diastole. The change in impedance is inversely proportional to the change in blood volume [10]. Due to Ohm's Law, when the current of constant intensity passes through the thorax, changes in voltage are directly proportional to changes in the impedance of the medium. The basic resistance,  $Z_0$ , is the total resistance of the thorax at rest, and represents the sum of the resistance of each of the components of the thorax: adipose tissue, heart and skeletal muscles, lungs, vessels, bone and air. The voltage is collected using two electrodes placed between the injection electrodes [11].

Bioimpedance allows the monitoring, diagnosis and determination of several cardiovascular parameters, in a non-invasive, inexpensive manner, requiring no specialized or highly qualified personnel. The wide clinical applications of bioimpedance will help further promote the technique. The precision and reproducibility of the measurements have been fully demonstrated. Several parameters can be determined by electrical bioimpedance. In this study, we mainly aim to use the bioimpedance signal to determine the heart rate.

## II. METHODS

### A. Study Protocol

This study was approved by the Research Ethics Committee in Health and Science Disciplines of the Prince Sattam bin Abdulaziz University (REC HSD 013 2020). The study involved 51 volunteers (28 male, 23 female) and their age ranging from 20 to 76 years ( $33 \pm 17$ ). Data was read and processed using MATLAB (version R2020a, MathWorks Inc., United States).

### B. Instrumentation

The device used in this experiment is a Siemens plethysmograph, 'Direktheagraph 933'. This device is equipped with an injection module formed by a square pulse generator of frequency 30 kHz at 1 mA. National Instrument data acquisition device (NI USB 6009) was integrated with LabView to collect the electrical bioimpedance output signal. The collected signal is amplitude modulated by the change in impedance of the explored zone. The device also indicates the basic resistance. In order to validate the calculated blood pressure from the aforementioned experiment, a FUZZY model MP3 digital blood pressure monitor was used to obtain the reference heart rate.



Bioimpedance consists in exploring the thorax [12], case of the thoracic bioimpedance or impedance cardiography, or limbs [13] as a case of the peripheral bioimpedance. As shown in Figure 1, in the case of peripheral bioimpedance, a current of 1 mA at 30 kHz was injected through two outer electrodes placed on subject's left arm (electrodes  $I_1$  and  $I_2$ ). Subsequently, two inner electrodes (electrodes  $S_1$  and  $S_2$ ) were used to measure the electrical voltage difference,  $\Delta V$ . During the signal recording, the patient must lie supine, relaxed and in expiratory apnea for a period of 10 seconds. Figure 1 presents the principle of this method. Peripheral bioimpedance exploits the relationship between the change in volume and the change in impedance to determine cardiovascular parameters.

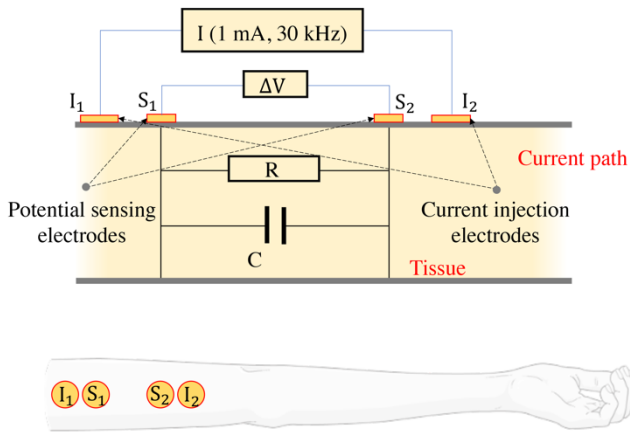


Fig 1. Electrode configuration.

Heart rate (HR) refers to the number of beats per minute (bpm) and it's can vary according to the physical needs of the body. Additionally, it is one of the best indicators of capability of the patient's cardiovascular activities. From the electrical impedance signal we can determine the time interval,  $T$ , by separating two consecutive complexes or peaks (see Figure 2). Thereafter, the HR can be calculated using the following expression:

$$HR = 60/T \quad (1)$$

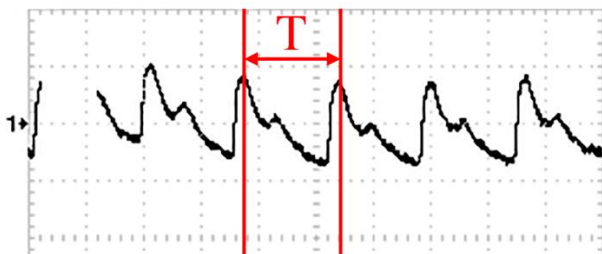


Fig 2. Bioimpedance signal

### C. Signal Processing

The acquired bioimpedance signals require pre-processing in order to identify peaks and subsequently calculating HR based on the average interval,  $T$ , of each subject. The sampling frequency,  $f_s$ , of this tested signal is 500 Hz. The processing workflow applied to the acquired bioimpedance signal is shown in Figure 3. Firstly, the bioimpedance signal is passed

through high pass and low pass filters with a cut-off frequency of 0.05 Hz and 50 Hz, respectively. Subsequently, the signal was pre-processed by a digital 50 Hz notch filter. Additionally, a moving average filter was implemented in series with notch filter using a sliding window of length 60 ms ( $30 \times 2$  ms) along the bioimpedance signal. The former filter was applied in order to produce a smoothed output and then was used for thresholding and peaks detection. In order to detect the peaks of the bioimpedance action potential signal, a non-linear operation was applied (see equation (2)).

$$S_n = \begin{cases} S_m^2 & S_m \geq 0 \\ 0 & S_m < 0 \end{cases} \quad (2)$$

$S_m$  and  $S_n$  represent the signal resulted from the non-linear operation and moving average filter, respectively. Following the above steps, we applied a thresholding on the resulted signal,  $S_n$ , by taking the 40% and above maximum signal. Finally, a periodic loop was applied in order to identify the signal peaks of each period.

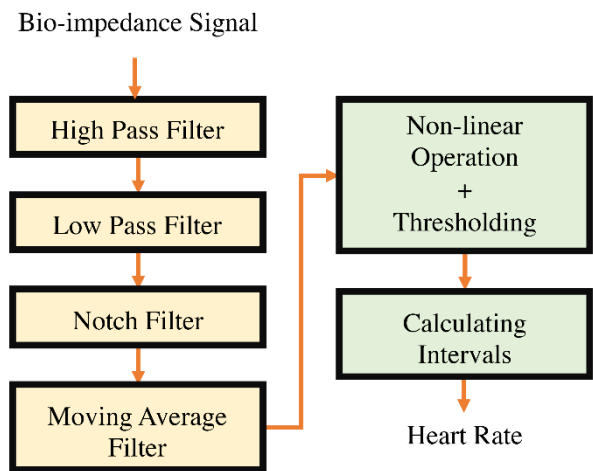


Fig 3. Signal pre-processing and peaks detection workflow.

The resulted outputs of each step at the signal pre-processing workflow are shown in Figure 4 (a – g). Figure 4 (a) shows the original bioimpedance signal before applying any filters. Figure 4 (b, c and d) show highly decrease of noise by applying high pass, low pass and notch filters, respectively. The moving average filter resulted in a smoothed signal and ready for thresholding and peaks detections (see Figure 4 (e)). Figure 4 (f) illustrated the non-linear and thresholding steps applied to the moving averaged signal. Finally, peaks of the signal are shown in the bottom plot of Figure 4. The final HR data from bioimpedance (calculated) and blood pressure monitor (measured) are presented in Table 1.

### III. RESULTS AND DISCUSSION

#### A. Processed Data

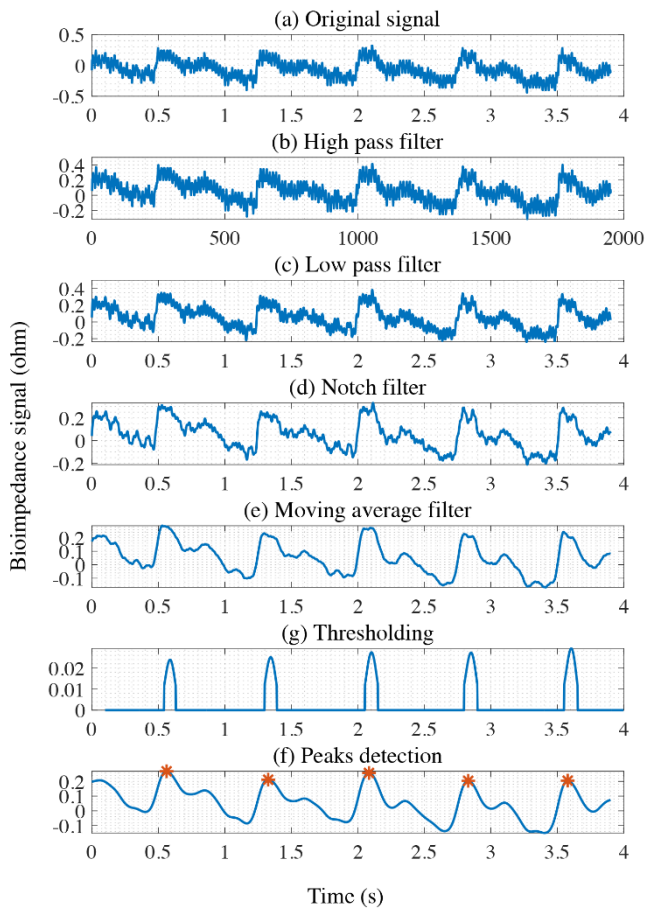


Fig 4. The resulted signal pre-processing operations and final outputs with labelled peaks.

#### B. Signal Interval and HR calculations

As shown in Figure 5, peaks were detected for every subject, and then average interval was calculated. Consequently, HR was calculated using equation (1).

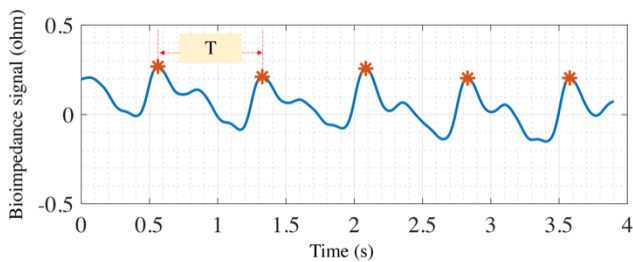


Fig 5. Peaks and periodic interval.

The measured HR using digital blood pressure can be used as reference to validate the calculated HR from the bioimpedance signal. The correlation between these two data was 95.44 %. The correlation coefficients are showing in the correlation matrix in Figure 6.

TABLE 1. Calculated and measured HR from bioimpedance and blood pressure digital monitor.

Record	HR		Record	HR	
	calculated	measured		calculated	measured
1	80	84	27	63	64
2	70	69	28	62	60
3	88	93	29	67	68
4	100	100	30	62	60
5	88	91	31	71	71
6	84	84	32	70	70
7	69	72	33	70	74
8	88	88	34	70	68
9	73	74	35	62	66
10	109	108	36	73	72
11	77	80	37	84	82
12	68	72	38	91	90
13	85	90	39	88	83
14	112	109	40	58	58
15	70	69	41	73	71
16	67	66	42	88	86
17	76	77	43	106	103
18	89	88	44	54	57
19	80	77	45	69	64
20	70	73	46	81	86
21	87	91	47	81	82
22	85	86	48	70	67
23	65	66	49	71	73
24	78	73	50	79	75
25	74	77	51	55	52
26	82	81			

Different studies have been used electrical bioimpedance to monitor several physiological signals and treat some diseases. Li et al. designed a system in order to measure electrical bio-impedance changes for HR detection in motionless condition by using multichannel physiologic recorder MP150 [14]. In their study, four electrodes were placed on the radial artery of the arm near to the wrist. In addition, the same group applied an electrical impedance-based method to analysis respiratory signal in order to identify the difference of the respiratory volume between different sleep posture [15].

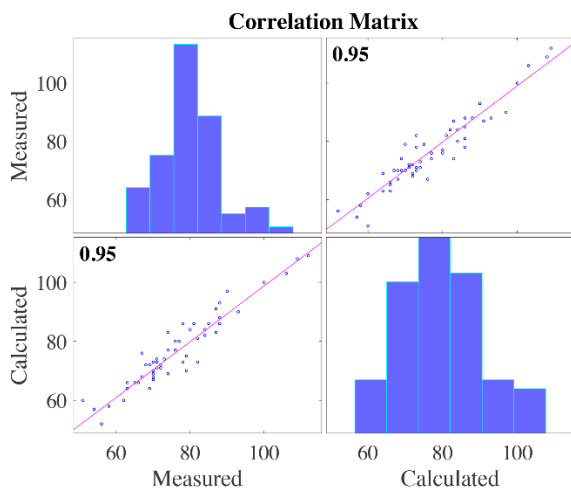


Fig 6. Correlation matrix.

In this study, we calculated the HRs using electrical bioimpedance by placing four electrodes on the volunteer's arm. After calculation and comparison, the results show high correlation between the calculated and measured HRs (> 95%). This revealed that the bioimpedance technique can be accurately reflected the subjects HR. Additionally, bioimpedance can be very useful to monitor the several parameters of cardiovascular systems, such as blood pressure, and diagnose valvular heart diseases. Moreover, a larger number of subjects are needed in order to quantify and recommend the presented method. For future work, we may try to apply electrical bioimpedance approach on wearable device for rapid and real-time measurements.

#### IV. CONCLUSION

The prevailing burden of diseases requires non-invasive, inexpensive and accurate medical devices for monitoring and early diagnosis of diseases. Bioimpedance offers alternative and accessible approach, moreover, and the key for accurate implementation relies on the underlying electrical model of the human body. In this study we were able to determine the HR of 51 volunteers. Their HRs highly agreed with the measured HRs from the digital blood pressure device. Thus, different clinical applications of bioimpedance will help further promote the technique and several parameters can be determined using the electrical bioimpedance approach.

#### ACKNOWLEDGMENT

This project was supported by the Deanship of Scientific Research at Prince Sattam bin Abdulaziz University.

#### REFERENCES

1. Liu, G., et al., *A new approach to detect congestive heart failure using short-term heart rate variability measures*. PLoS One, 2014. **9**(4): p. e93399.

2. Prasad, A. and M. Roy, *Bioimpedance analysis of vascular tissue and fluid flow in human and plant body: A review*. Biosystems Engineering, 2020. **197**: p. 170-187.
3. Strobeck, J.E. and M.A. Silver, *Beyond the four quadrants: the critical and emerging role of impedance cardiography in heart failure*. 2004, Wiley Online Library.
4. Packer, M., et al., *Utility of impedance cardiography for the identification of short-term risk of clinical decompensation in stable patients with chronic heart failure*. J Am Coll Cardiol, 2006. **47**(11): p. 2245-52.
5. Braun, M.U., et al., *Impedance cardiography as a noninvasive technique for atrioventricular interval optimization in cardiac resynchronization therapy*. J Interv Card Electrophysiol, 2005. **13**(3): p. 223-9.
6. Burlingame, J., et al., *Noninvasive cardiac monitoring in pregnancy: impedance cardiography versus echocardiography*. J Perinatol, 2013. **33**(9): p. 675-80.
7. Thangathurai, D., et al., *Continuous intraoperative noninvasive cardiac output monitoring using a new thoracic bioimpedance device*. Journal of cardiothoracic and vascular anesthesia, 1997. **11**(4): p. 440-444.
8. Zielińska, A., H. Dudek, and G. Cybulski. *Application of impedance cardiography for haemodynamic monitoring in patients with ischaemic stroke*. in *XIII Mediterranean Conference on Medical and Biological Engineering and Computing 2013*. 2014. Springer.
9. Nakamura, T., et al., *Role of cardiac sympathetic nerves in preventing orthostatic hypotension in Parkinson's disease*. Parkinsonism Relat Disord, 2014. **20**(4): p. 409-14.
10. Redisch, W., *Electrical impedance plethysmography*. Chest, 1971. **59**(6): p. 36.
11. Grimnes, S. and O.G. Martinsen, *Bioimpedance and bioelectricity basics*. 2011: Academic press.
12. Badeli, V., et al., *Electrode Positioning to Investigate the Changes of the Thoracic Bioimpedance Caused by Aortic Dissection - A Simulation Study*. J Electr Bioimpedance, 2020. **11**(1): p. 38-48.
13. Bera, T.K., *Bioelectrical Impedance Methods for Noninvasive Health Monitoring: A Review*. J Med Eng, 2014. **2014**: p. 381251.
14. Li, K.-Y., et al. *A method to detect heart rate based on electrical bio-impedance*. in *ITM Web of Conferences*. 2016. EDP Sciences.
15. Liu, G., et al., *A principal component analysis based data fusion method for estimation of respiratory volume*. IEEE Sensors Journal, 2015. **15**(8): p. 4355-4364.

**2021 IEEE Seventh  
International Conference on  
Bio signals, Images and  
Instrumentation  
ICBSII 2021**

**SESSION V  
RESEARCH  
PAPERS**

# Relationship between Glaucoma and Complexity Measures of the Electroretinogram

Marc Sarossy  
Dept of Surgery  
University of Melbourne  
Melbourne, Australia  
marc@sarossy.com

Dinesh Kumar  
Dept of Electrical Engineering  
RMIT University  
Melbourne, Australia  
dinesh@rmit.edu.au

Zhichao Wu  
Macular Research Unit  
Centre for Eye Research Australia  
East Melbourne, Australia  
wu.z@umielb.edu.au

**Abstract**—Glaucoma is one of the leading causes of vision loss in the world. Early diagnosis is essential to prevent irreversible vision loss. Such early diagnosis requires the use of tests of visual function rather than tests of anatomical structure. Existing tests of visual function can suffer from the use of bulky expensive equipment and poor subject performance. In this work we describe a novel analysis of the electroretinogram to extract features of complexity and detect glaucoma. Cohorts of participants with and without glaucoma had electroretinogram (ERG) testing. The stimulus used to evoke the signal was that used for the Photopic Negative Response (PhNR), a response known to be reduced in glaucoma. There were 18 control and 21 glaucoma participants. Significant differences were found in PhNR amplitude  $-5.5$  mV vs  $-3.6$  mV ( $p < 0.01$ ), Shannon Entropy  $2.99$  vs  $2.80$  ( $p < 0.01$ ) and Higuch Fractal Dimension  $1.068$  vs  $1.082$  ( $p < 0.02$ ). Differences in box count and Katz Fractal Dimension and Sample Entropy were not significant. These complexity measures may have potential for machine learning classifiers for glaucoma.

**Keywords**—glaucoma, entropy, fractal dimension, electroretinogram

## I. INTRODUCTION

Glaucoma is a disease of the optic nerve of the eye – the structure that electrically connects the eye to the brain. The nerve is comprised of the axons of the ganglion cells of the retina and in a healthy human eye number around one million. Glaucoma is caused by elevated pressure within the eyeball which triggers a cascade of damage within the ganglion cells resulting in death by apoptosis of the ganglion cells [1]. The exact mechanism of the damage remains uncertain however the clinical pattern of evolving glaucoma in an eye is recognizable by a loss of peripheral visual field, stereotypical structural changes at the optic nerve head and changes in the thickness of the ganglion cell and nerve fiber layer on the optical coherence tomography scan. The disease is the second most frequent cause of legal blindness registrations in industrialized countries and the number of people aged 30-80 in the world with glaucoma is estimated to be 64.3 million with 60% of those in Asia [2] Early diagnosis of glaucoma is essential because there are many treatments that reduce the intraocular pressure and consequently reduce the rate and risk of progression.

The electroretinogram is an evoked potential from the eye, measured at the cornea relative to the retina. It is evoked by a brief flash of light. It has been shown [3] that a component of the response – the Photopic Negative Response (PhNR) is reduced by glaucoma, both in experimental monkey models of glaucoma and in human patients with the disease. This response is of low amplitude

and may be contaminated by blink artefact. Despite the publication of a standard [4] the test-retest reliability of the response remains low [5]

The PhNR arises from the spiking potentials of ganglion cells. As the disease progresses, the number of ganglion cells generating the response is reduced. Each ganglion cell produces a stochastic spike train where the spike rate follows a Poisson distribution the parameter influenced by the stimulating light level within that cell's receptive field. A reduction in the number of ganglion cells should be manifested by a reduction in the complexity of the part of the ERG generated by the ganglion cells

Various measures have been described to measure the complexity of signals. Shannon entropy [6] is one measure of complexity that can be applied to a time series. It is a function of the probabilities of the levels within the set of measures – that is a function of the probability density function of the amplitude measures. Sample entropy (SampEn) [7] is another measure which relates to the similarity of sequences of different lengths within the signal.

Fractal dimension is a different method that can be used to measure the complexity of a time series [8]. It describes the rate of change of length as scale increases when plotted on a log-log scale. If an object exhibits fractal properties, it will show a linear relationship on this plot. The measure has found utility in analysis of retinal vasculature.[9]

## II. METHODOLOGY

### A. Participants

This research followed the tenets of the Declaration of Helsinki and was approved by the Human Research Ethics Committee of the Royal Victorian Eye and Ear Hospital. 2 groups of participants were recruited: Glaucoma and non glaucoma controls. Inclusion criterion for the glaucoma cohort was a known diagnosis of glaucoma and a mean deviation on the visual field test of at least  $-3$  dB.

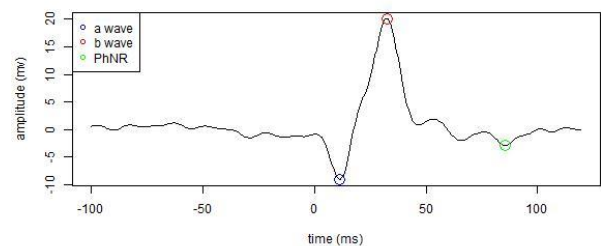


Figure 1 Sample ERG trace

TABLE I. ERG RESULTS

Result			
	Controls	Glaucoma	p
Test subjects	18	21	-
b amplitude (mv)	31.1	30.7	0.66
b latency (ms)	19.3	17.1	0.17
PhNR (mv)	-5.50 (1.95)	-3.60 (2.53)	0.01

TABLE II. COMPLEXITY MEASURES

Result			
	Controls	Glaucoma	p
Test subjects	18	21	-
Box count FD	1.148	1.157	0.57
Higuchi FD	1.068	1.082	0.02
Katz FD	1.005	1.007	0.08
SampEn	0.077	0.072	0.07
Shannon entropy	2.99	2.80	0.01

Normal controls were recruited from family and friends of the glaucoma cohort with no reported history of eye disease. Informed consent was obtained for all participants...

B. ERG Recordings

In this study, only left eyes were considered. ERG responses were recorded from eyes without pupil dilation. ERG signals were recorded with the RETeval™ hand held device (LKC Technologies, Gaithersburg, MD). A red flash of 4ms duration was used as the stimulus on a 10cd/m2 blue background. Flash energy was set to 1.5cd.s/m2. The sample rate of the device was set to 1.9 K Samples per second at 24 bits resolution, an epoch of 220ms and pre-stimulus acquisition starting at 100ms. 200 sweeps were performed for each session. Proprietary 3 pole skin electrodes were used which incorporated active, indifferent and earth electrodes. B-wave amplitude (the first positive going wave measured relative to the a-wave) and implicit time were measured with the provided device software.

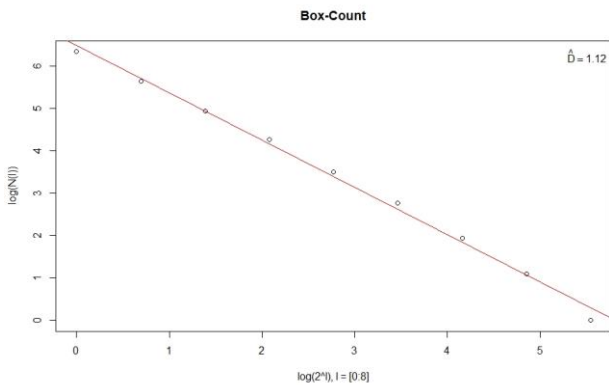


Figure 3 Box Counting Fractal Dimension Example

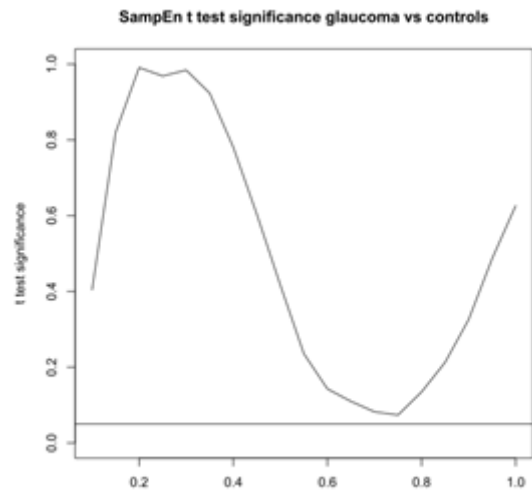


Figure 2 Effect of radius on Sample Entropy

Raw data were exported from the device at the conclusion of each testing session and analysed with R. The a wave, b wave and PhNR amplitude and implicit times were calculated from the processed waveform produced by the device and corresponded to the first negative going wave and the first positively going wave after the stimulus for the a and b wave respectively. The a wave amplitude was measured relative to the detrended baseline and the b wave was measured relative to the a wave. The PhNR was defined as the first negative going wave after the b wave and was referenced to the prestimulus baseline.

C. Mathematical Analysis

All calculations were performed in R. Shannon Entropy was estimated using the Entropy package. Data was discretized into 30 bins for the calculation. Sample entropy was estimated using a custom script in R with a radius of 0.2 times the standard deviation of the signal. Fractal dimensions were estimated using the FractalDim package [10]. Boxcount, Higuchi[11] and Katz [12] entropies were calculated.

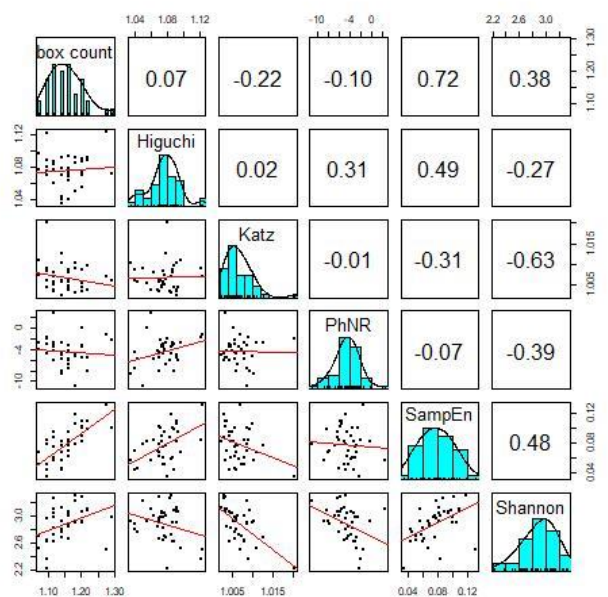


Figure 4 Pairs plot of complexity measures

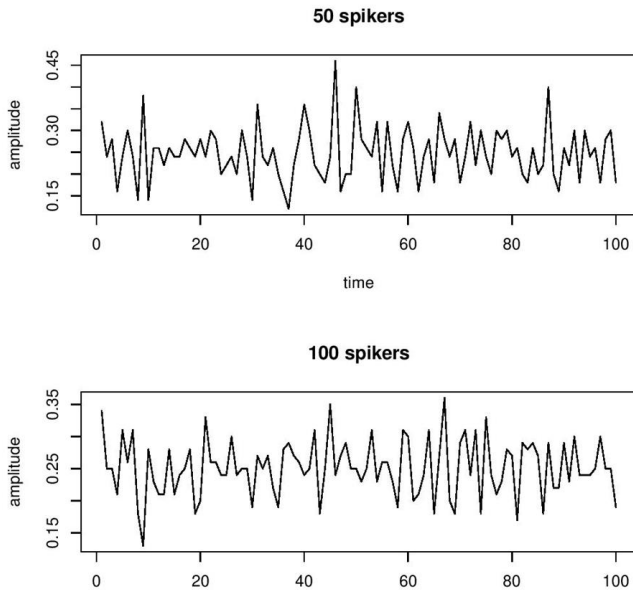


Figure 5 Simulated Electroretinogram signals

### III. RESULTS

#### A. Electroretinogram Waves

Fig 1 shows a typical ERG trace consisting of the average of the 200 sweeps. Markers have been placed on the a and b waves and the PhNR.

Table I shows the results from the glaucoma and control cohort. There were no significant differences in b wave amplitude or latency between the glaucoma and control groups. Glaucoma is known not to affect this wave. There was a significant difference in the PhNR amplitudes between control and glaucoma cohorts.

#### B. Fractal Properties of the Electroretinogram

Fig 2 shows a box count FD calculation for the waveform shown in Fig 1. Good fit was obtained in all signals for the box count log log plot with  $R^2$  measures above 0.95 for all of the ensemble averages

#### C. Complexity Measures and Glaucoma

Table II shows the performance of the various complexity measures in classifying glaucoma and non glaucoma subjects. Box count FD was far from significance whereas Higuchi FD and Shannon entropy had highly significant differences between the cohorts. Katz entropy and sample entropy approached but did not reach significance

Fig 3 shows how the radius affects the performance of SampEn in being able to separate glaucoma from non glaucoma subjects. The radius for SampEn is the tolerance within which subsequences are considered to be the same. Fig 2 shows that for this data set, the optimum radius to discriminate between glaucoma and non glaucoma individuals is 0.7 times the standard deviation

#### D. Relationship between Complexity Measures

Fig 4 shows the relationship between all the various complexity measures and the PhNR for all test subjects (glaucoma and control). There were weak to moderate correlations between the measures. Moderate correlations were found between the PhNR and both Shannon entropy and Higuchi FD – both of which had good classification value when separating glaucoma and non glaucoma. Box count FD was very poorly correlated with PhNR mirroring its inability to differentiate between the cohorts.

### IV. DISCUSSION

Making an early diagnosis of glaucoma is critical to the final outcome in the management of the disease. Early disease is often asymptomatic and significant loss of visual function can occur before diagnosis. Screening of at-risk individuals can help: targeting people with a family history of glaucoma or older people with other risk factors. [13]

Current diagnosis relies upon a mixture of objective and subjective measures. The most important subjective measure is the determination of the visual field. This is a quantitative measurement of the brightness of light required to be detected as a function of spatial location within the visual space

Patients do find the performance of the visual field test (VFT) to be stressful [14]. The test can also suffer from unreliability as it does require patient attention and cooperation over a prolonged period. Tests such as optical coherence tomography (OCT) of the optic nerve or analysis of disc appearance via ophthalmoscopy or photographic imaging of the disc are tests of structure rather than function and as such will not be able to detect the presence of sick nerve fibres within the optic nerve.

The RetEVAL is marketed as a low-cost device suitable for screening glaucoma. The device is hand held, and is used with skin electrodes – compared with other ERG devices which are expensive and require the placement of more invasive contact lens or fiber electrodes. The ERG, like the VFT is a test of function and should be sensitive to changes in function in the ganglion cells and their axons, even if the cells have not yet undergone apoptosis [15]

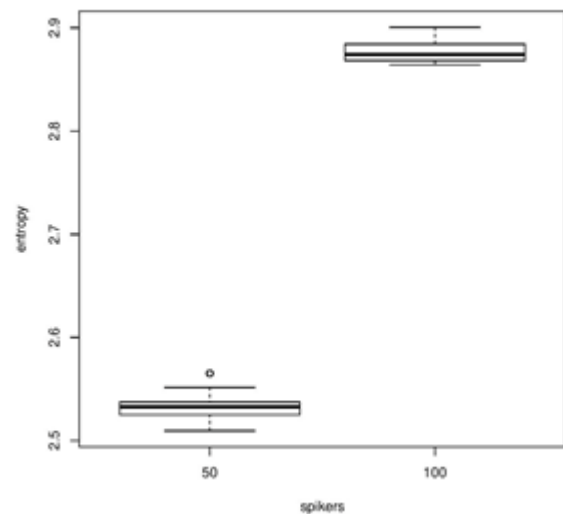


Figure 6 Shannon entropy from simulated electroretinogram signals

The PhNR has been shown to be affected by glaucoma but as yet lacks the reliability to form part of the diagnostic arsenal for the glaucoma specialist and has not been shown to be of value in glaucoma screening.

In this paper, we have shown that the ERG is affected in ways beyond the simple PhNR in glaucoma and it is possible that these measures of complexity either on their own or combined with the PhNR will aid in either the diagnosis or the management of glaucoma

The response is known to be generated by the summing of action potentials from the ganglion cells. Each of those spike trains is a Poisson process.

Fig 5 shows a simulated analog signal – formed from the sum of 50 independently discharging Poisson generators in the top panel and 100 in the lower panel. The signals do not show any obvious difference in the time domain

Fig 6 shows boxplots of Shannon Entropies of similar signals for 1000 such simulations. The graph shows very different entropy levels with the signals that are generated with more independent spiking generators having the higher entropy and therefore complexity. This is a plausible explanation for the differences seen between the glaucoma and non glaucoma cohort.

This study was limited by the small sample size and did not take into account the severity of glaucoma beyond a threshold level. Even given the limitations, we have shown an interesting and plausible fingerprint of glaucoma within the ERG signal that could suggest a feature to consider for future automated diagnosis statistical model.

#### ACKNOWLEDGMENT

The assistance of the staff of the glaucoma clinic at the Royal Victorian Eye and Ear Hospital was very valuable in collecting the data for this manuscript.

#### REFERENCES

[1] McKinnon, S.J.: 'Glaucoma, apoptosis, and neuroprotection', *Current opinion in ophthalmology*, 1997, 8, (2), pp. 28-37

[2] Tham, Y.-C., Li, X., Wong, T.Y., Quigley, H.A., Aung, T., and Cheng, C.-Y.: 'Global prevalence of glaucoma and projections of glaucoma burden through 2040: a systematic review and meta-analysis', *Ophthalmology*, 2014, 121, (11), pp. 2081-2090.

[3] Viswanathan, S., Frishman, L.J., Robson, J.G., Harwerth, R.S., and Smith, E.R.: 'The photopic negative response of the macaque electroretinogram: reduction by experimental glaucoma', *Investigative ophthalmology & visual science*, 1999, 40, (6), pp. 1124-1136.

[4] Frishman, L., Sustar, M., Kremers, J., McAnany, J.J., Sarossy, M., Tzekov, R., and Viswanathan, S.: 'ISCEV extended protocol for the photopic negative response (PhNR) of the full-field electroretinogram', *Documenta Ophthalmologica*, 2018, pp. 1-5.

[5] Tang, J., Edwards, T., Crowston, J.G., and Sarossy, M.: 'The test-retest reliability of the photopic negative response (PhNR)', *Translational vision science & technology*, 2014, 3, (6), pp. 1-1.

[6] Shannon, C.E.: 'A mathematical theory of communication', *The Bell System Technical Journal*, 1948, 27, pp. 379-423.

[7] Richman, J.S., and Moorman, J.R.: 'Physiological time-series analysis using approximate entropy and sample entropy', *American Journal of Physiology-Heart and Circulatory Physiology*, 2000, 278, (6), pp. H2039-H2049.

[8] Naik, G.R., Arjunan, S., and Kumar, D.: 'Applications of ICA and fractal dimension in sEMG signal processing for subtle movement analysis: a review', *Australasian physical & engineering sciences in medicine*, 2011, 34, (2), pp. 179-193

[9] awasaki, R., Azemin, M.C., Kumar, D., Tan, A., Liew, G., Wong, T., Mitchell, P., and Wang, J.: 'Fractal dimension of the retinal vasculature and risk of stroke: a nested case-control study', *Neurology*, 2011, 76, (20), pp. 1766-1767

[10] Gneiting, T., Ševčíková, H., and Percival, D.B.: 'Estimators of fractal dimension: Assessing the roughness of time series and spatial data', *Statistical Science*, 2012, pp. 247-277

[11] Higuchi, T.: 'Relationship between the fractal dimension and the power law index for a time series: a numerical investigation', *Physica D: Nonlinear Phenomena*, 1990, 46, (2), pp. 254-264

[12] Katz, M.J.: 'Fractals and the analysis of waveforms', *Computers in biology and medicine*, 1988, 18, (3), pp. 145-156

[13] Tan, N.Y., Friedman, D.S., Stalmans, I., Ahmed, I.I.K., and Sng, C.C.: 'Glaucoma screening: where are we and where do we need to go?', *Current opinion in ophthalmology*, 2020, 31, (2), pp. 91-100

[14] Glen, F.C., Baker, H., and Crabb, D.P.: 'A qualitative investigation into patients' views on visual field testing for glaucoma monitoring', *BMJ open*, 2014, 4, (1)

[15] Crowston, J.G., Kong, Y.X.G., Trounce, I.A., Dang, T.M., Fahy, E.T., Bui, B.V., Morrison, J.C., and Chrysostomou, V.: 'An acute intraocular pressure challenge to assess retinal ganglion cell injury and recovery in the mouse', *Experimental eye research*, 2015, 141, pp. 3-8



# Identification of Schizophrenia using LSTM Recurrent Neural Network

Abinaya Sundari R

Department of Electronics and Communication Engineering  
Anna University  
Chennai, India

[abinayam.kala@gmail.com](mailto:abinayam.kala@gmail.com)

C M Sujatha

Department of Electronics and Communication Engineering  
Anna University  
Chennai, India

**Abstract—** Schizophrenia is a severe psychiatric illness that greatly affects the quality of life. Early discovery of psychiatric onset and progression are of great significance for bestowing effective treatment to prevent or mitigate further degeneration in the disease. Functional MRI possesses large number of subtle information which helps in early identification of disease. Current fMRI analysis focuses on spatial maps and connectivity patterns rather than temporal information. In this work an attempt has been made to identify schizophrenia with the use of LSTM deep learning model using ROI based time series extracted from fMRI. The classifier accuracy of 81.3% demonstrates that, use of ROI time series can aid as better diagnostic measure in nearby future.

**Keywords—** Schizophrenia, RNN, LSTM, Seed Time Series, fMRI

## I. INTRODUCTION

Schizophrenia is one among the most common neurodegenerative disorder which leads to notable difficulties in everyday functioning and social relations with symptoms like hallucinations, delusion and social withdrawal [1]. World Health Organisation estimates the prevalence of disease about 1% of the population worldwide which is approximately 21 Million and the second common psychiatric disease in India with 8.2 Million schizophrenics.

Schizophrenics often lack consciousness that their difficulties stem from a mental ailment [2], and identification of disease in advanced stages requires lifelong medication compliance. Etiology of Schizophrenia remains unclear and earlier diagnosis is highly essential for starting pharmaceutical treatment to palliate the symptoms [3].

Present non-invasive diagnosis of schizophrenia relies on the qualitative assessment of obvious mental symptoms and patients' self- perceived experience such as Positive and Negative Syndrome Scale, Clinical Assessment Interview for Negative Symptoms, Brief Negative Symptom Scale, Scale for the Assessment of Positive Symptoms, Scale for the Assessment of Negative Symptoms etc. [4]. Behavioural and functionality changes of patients are inferred from these trials and assessments are usually influenced by the subject's educational background and intelligence. However these diagnosis methods help at the latter stage of disease

and are not feasible for earlier identification of disease. Additional diagnosis methods such as neuroimaging and brain signal analysis requires highly experienced professionals to interpret the severity.

Functional Magnetic Resonance Imaging (fMRI) is a non-invasive imaging technique that could capture the brain functions in an effective way with Blood Oxygen Level Dependent (BOLD) signals. It also provides other advantages like higher spatial resolution, lack of ionizing radiation, feasibility, and direct correlation with structural image as well as greater repeatability [5]. Diagnosis with neuro-functional images could be able to detect subtle abnormality at the early stage of schizophrenia [6] while structural images can help merely in progressed phase.

According to the findings of various schizophrenia researches, disrupted functional connectivity among disparate brain regions has been repetitively noticed [7] [8]. Another notable aspect from fMRI data is that the BOLD time courses of independent components with significant temporal information remains unnoticed in functional connectivity. Recurrent Neural Networks (RNN) with feedback connections employs sequential information of each voxel to discriminate schizophrenics and healthy subjects [9]. RNN models such as Gated Recurrent Unit (GRU) [10] and Long Short-Term Memory (LSTM) [11] have been firmly established as state-of-the-art methodologies in sequence investigation. Weizheng Yan, *et al.* [12] have investigated RNN-GRU's using ROI time courses for schizophrenic discrimination with 80.1% of accuracy for leave one site transfer classification. However analysis of ROI time courses or seed time courses with RNN-LSTM remains unexplored.

RNN inputs are dependent on each other and suffers vanishing gradient problem in long sequence data. The LSTM networks are modified RNN networks exclusively designed to abstract patterns in time series to create improved performance. The fact of having long-term and short-term memory cells, LSTM eliminates the vanishing gradient issue in RNN. Thus LSTM is able to memorize characteristic fragments of sequential data in a crucial manner during the learning process [13]. GRU's are similar to LSTM without hidden state and has lesser parameters which make them simple in structure, but it may cost reduced expressibility and accuracy for larger data. GRU's are highly suitable for smaller data [14].

In this work, the ROI time series obtained with resting state fMRI images are utilized for classification module. The ROI time courses obtained from fMRI are of larger size and LSTM is more suitable to achieve higher accuracy. Bidirectional LSTM employed in the network, runs the input in two ways so that the network will have information from past and future, thereby generate better performance.

## II. MATERIALS AND METHODS

### A. Subjects

A set of 20 resting state functional MR images each of Normal(CN) and Schizophrenia(SZ) subjects from open access COBRE (The Center for Biomedical Research Excellence) SZ dataset has been utilized for this work.

Structural MRI of multi-echo MPRAGE sequence with matrix size of 256x256x190, TR=2530ms, TE=29ms, Slice thickness=1mm, were obtained with SIEMENS MAGNETOM TrioTim syngo MR B17.

Resting state BOLD images of the whole brain using an Echo Planar Imaging (EPI) sequence with matrix size of 64x64x33x150 (TR = 2000ms, TE = 29ms, 3.5mm thickness, voxel size(3.8x3.8x3.5mm) with the same machine.

### B. Method

The methodology followed in this work is shown in Fig.1.

#### a. fMRI preprocessing

Pre-processing of resting state fMRI data has been carried out using SPM8[15] and CONN toolbox[16], developed by the Wellcome Department of Cognitive Neurology, London, UK. Functional realignment, unwarping, slice timing, head motion correction, outlier detection, direct functional normalization with MNI coordinates, and functional smoothing (8mm full-width at half-maximum Gaussian kernel) were the steps included.

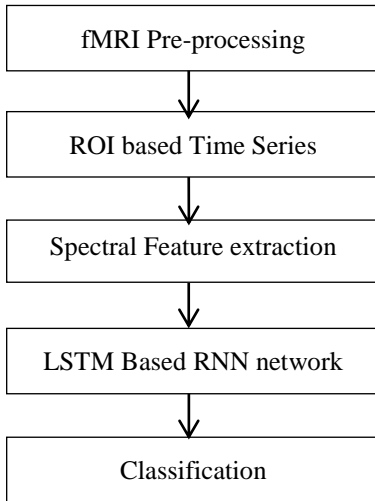


Fig.1 Proposed Method

#### b. Time Course Extraction

The smoothed unwrapped images obtained from preprocessing step have been utilized to obtain the ROI time course. fMRI act as a function of time, meaning each time point in the sequence refers one scan. The Parcellation of the volumes with 3mm AAL atlas into 116 regions has been

performed and their corresponding ROI based time series were obtained with SPM based GREYNA [17] toolbox. The time courses of each region depict the mean BOLD response of the voxels in the region for each scan.

#### c. Feature Extraction

Feature extraction involves in conversion of the sequential inputs to spectrograms (i.e. 2D images) and the 1D features extracted from these spectrograms has been adopted in training the network.

Short Time Fourier Transform (STFT) is employed to map the signal to a 2-dimensional function of time and frequency. STFT is computed by partitioning of signals to short segments and shifting the time window with overlap so that local frequency spectrum for each segment is obtained [18]. Hamming window is utilized to ensure the continuity between the points so that leakage effect in spectrum can be prevented. The general equation for spectrogram computation is given as

$$S(a, k) = \sum_{n=0}^{N-1} s(n + aN') w(n) e^{j2\pi nk/N} \quad (1)$$

where  $k=0,1,\dots,N-1$ ,  $S(a,k)$  is the m-index time-frequency spectrogram.  $N$  is the window segment length and  $N'$  is the shifting step of the time window and  $w(n)$  is the window method of  $N$ -point sequence.  $N'$  should be lesser than  $N$  so as to build overlap in between time windows. It can be carried out in order to comprehensively capture the temporal changes and features of signal [19].

Spectrogram is defined as the magnitude of  $S(a,k)$ , can be represented as  $A(a,k)$

$$A(a, k) = \frac{1}{N} |S(a, k)|^2 \quad (2)$$

Two time moment features were extracted from these spectrograms namely spectral entropy and instantaneous frequency.

Instantaneous frequency  $f_{ins}$  [20] calculates the spectrogram power spectrum  $P(t, f)$  of the data and is given as

$$f_{ins}(t) = \frac{\int_0^{\infty} f P(t, f) df}{\int_0^{\infty} P(t, f) df} \quad (3)$$

where  $t$  is sample time and  $f$  is the frequency of spectrum.

Spectral entropy measures the spectral power distribution of the signal based on Shannon entropy [21] and given as

$$H(t) = - \sum_{m=1}^N P(t, m) \log_2 P(t, m) \quad (4)$$

where,  $H(t)$  is the Spectral entropy,  $P(t, m)$  is the probability distribution at  $t$  and it is given as

$$P(t, m) = \frac{S(t, m)}{\sum_f S(t, f)} \quad (5)$$

$S(t, m)$  is the power spectrum at  $t$ ,  $S(t, f)$  is the time-frequency power spectrogram

#### d. RNN-LSTM

Long Short term memory is one of the deep learning networks used to analyze time-series as an improvement to classical RNN. Traditional RNN suffers from vanishing gradient and exploding gradient in long sequence due to the loss of information carried by gradient and result in no real learning. This drawback can be solved by introducing memory cells in LSTM network. LSTM cells (Fig.2) memorizes the previous information. The network takes three inputs.  $X_t$  is the input of the current time step.  $H_{t-1}$  is the output of the previous LSTM block and  $C_{t-1}$  is the previous

unit memory. As for outputs,  $H_t$  is the output of the current network.  $C_t$  is the memory of the current unit. The output of single unit is generated by considering previous output, current input and previous memory.

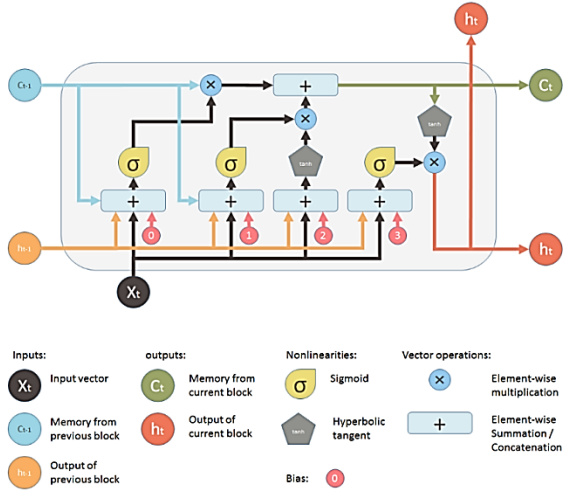


Fig.2 Architecture of LSTM Block[22]

Bidirectional LSTM [23] has the forward and backward hidden layers combined for accessing both the preceding and succeeding data. Bidirectional LSTM can handle the sequential modeling challenge better than conventional LSTM.

The steps involved include identification of the data which is to be removed from the cell state decided by sigmoid activation function between the values of 0 and 1. LSTM has three gates namely Input gate ( $I_t$ ), Forget gate ( $F_t$ ) and Output gate ( $O_t$ ) and the corresponding output are given below

$$I_t = \sigma(w_i[H_{t-1}, x_t] + b_i) \quad (6)$$

$$F_t = \sigma(w_f[H_{t-1}, x_t] + b_f) \quad (7)$$

$$O_t = \sigma(w_o[H_{t-1}, x_t] + b_o) \quad (8)$$

where  $\sigma$  is the sigmoid activation function,  $b$  is biases for respective gates,  $w$  is the weights for respective gate neurons and  $H_{t-1}$  is the output of previous LSTM.

$$C_t = \tanh(w_c[H_{t-1}, x_t] + b_c) \quad (9)$$

$$c_t = F_t * c_{t-1} + I_t * C_t \quad (10)$$

$$H_t = O_t * \tanh(c_t) \quad (11)$$

where  $c_t$  is cell state and  $C_t$  is the candidate for cell state. The output of LSTM is applied to the classifier to perform the final classification task.

#### e. Classification

The classification module has a sequential input layer where the sequential data or feature sequence to be fed. The inputs from the input layer are pipelined to the bidirectional LSTM layer which learns the long term relationship with the sequential data or time steps in both forward and backward direction. The output of the LSTM layers are decided by the input, hidden and output gate and cell state holds the learned information from previous time steps.

In most of the deep networks, fully connected layers (FC) are introduced next to convolution and down sampling layers. The fully connected layer works individually at each

stage for sequential input in LSTM. In case the output of the layer placed prior to the FC layer is an array  $A_1$  of the size  $X \times Y \times Z$ , then the FC output is an array  $A_2$  of the size  $X' \times Y \times Z$ . At time  $t$ , the input of  $A_2$  is  $WA_t + b$ , where  $A_t$ -time step  $t$  of  $A$ . In these studies, Glorot initializer was the initiating algorithm for the weights of this layer [24]. Drop out layers are added next to FC, with the intention of reducing over fitting of network with dropout probability of 0.5.

The penultimate layer is the softmax which is always preceded by a fully connected layer. It uses the softmax activation function as

$$y_r(x) = \frac{e^{a_r(x)}}{\sum_{j=1}^k e^{a_j(x)}} \quad (12)$$

where  $0 \leq y_r \leq 1$  and  $\sum_{j=1}^k y_j = 1$

The final layer computes the cross entropy loss for classification of different classes. The adaptive moment estimation (adam) algorithm [25] has been utilized to train the LSTM network.

### III. EXPERIMENTAL RESULTS

The preprocessed functional MR images are utilized to obtain ROI mean time series for 116 AAL atlas regions. The preprocessing results of sMRI and fMRI were shown in Fig. 3.

These processed fMRI has been co-registered with the sMRI of the same subject and then spatially normalized. Smoothing of data with gaussian kernel of 8mm FWHM has been performed.

The resulting images were used in GRETNA toolbox to perform parcellation and time course extraction as shown in Fig.4.

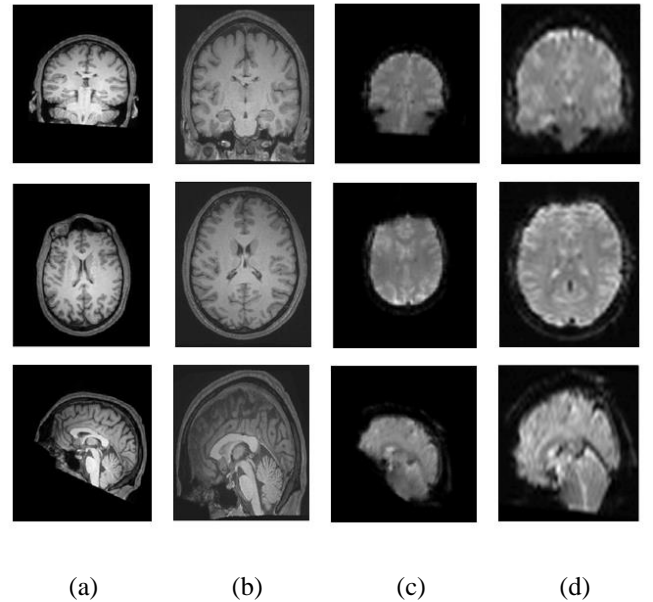


Fig.3 Representative of Preprocessed MR images (a) Original sMRI, (b) Realigned, unwarped, motion corrected, smoothed sMRI, (c) Original fMRI, (d) Realigned, unwarped, motion corrected fMRI

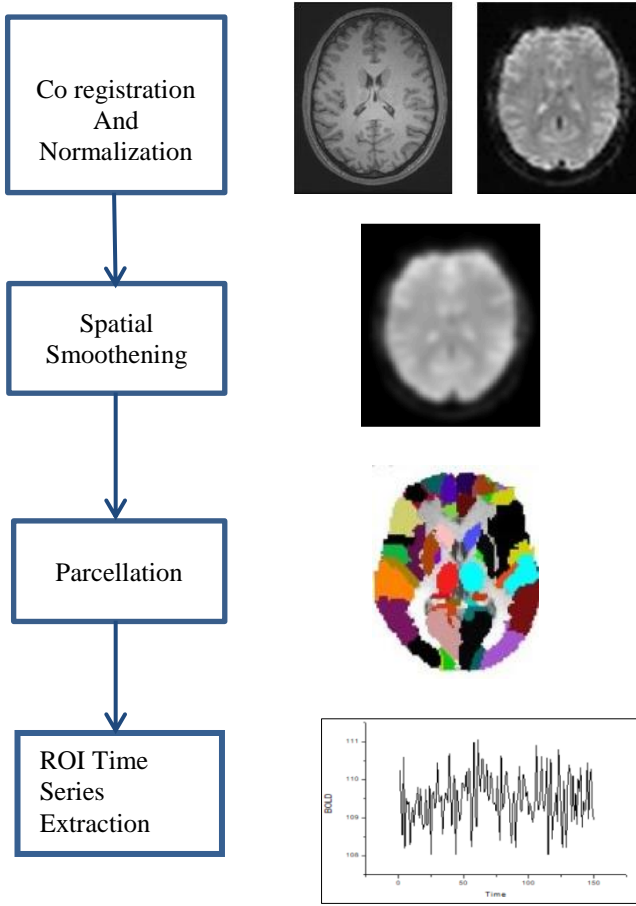


Fig.4. Preprocessing and ROI Timeseries Extraction

These ROI time courses are augmented by duplicating the input and are given to the sequential input layer in the network. The various layers and their activations are given in Table I. The architecture of the network is shown in Fig.5.

TABLE I. Layers of the network

Layer Name	Activations	Learnable
Sequence Input	2	-
BiLSTM	2000	Input Weight- 8000*2 RecurrentWeight- 8000*1000 Bias- 8000*1
Fully Connected	50	Weights 50*2000 Bias 50*1
Dropout	50	-
Fully Connected	2	Weights 2*50 Bias 2*1
Softmax	2	-
Class Output		

The BiLSTM layer has the following parameters: gate activation function- sigmoid, state activation function-tanh, mini batch size = 250, initial learning rate = 0.0001, sequence length = 1000, gradient threshold = 1.

The total number of training data includes 120 samples and 16 samples were used for testing. The training parameters comprises minimum batch size of 50, 1000 hidden units with 200 iterations. The size of training sequence is 1x17400.

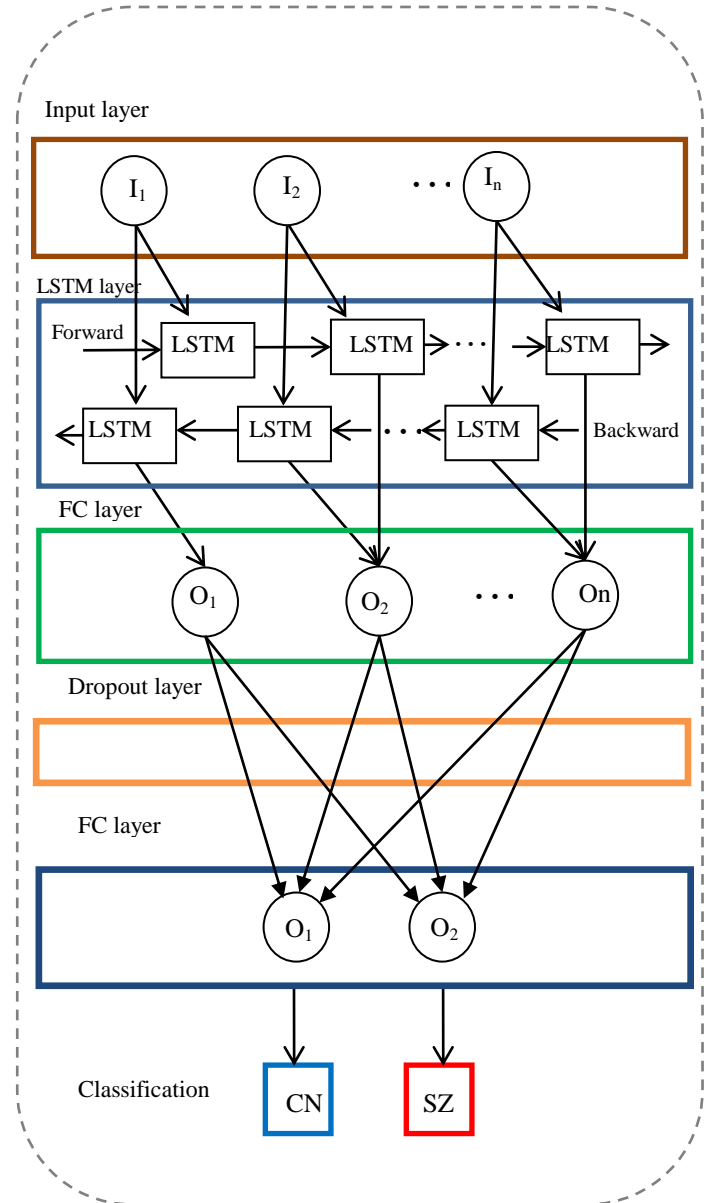


Fig.5 Architecture of Network

In this model the sequence from the input layer is fed to the LSTM blocks which generate output with consideration to previous memory. Due to the small number of training samples, a separate test set is used for validation. The network is trained for 100 epochs where in each epoch the entire dataset is gone through and tested with a validation data for every thirty iterations in order to reduce the computation load on the network and decrease the time taken for the learning process. The accuracy and loss incurred during the training and the test phase are monitored to estimate the effectiveness of the network for classification. This model is able to classify schizophrenics from normal with the accuracy of 77.8%.

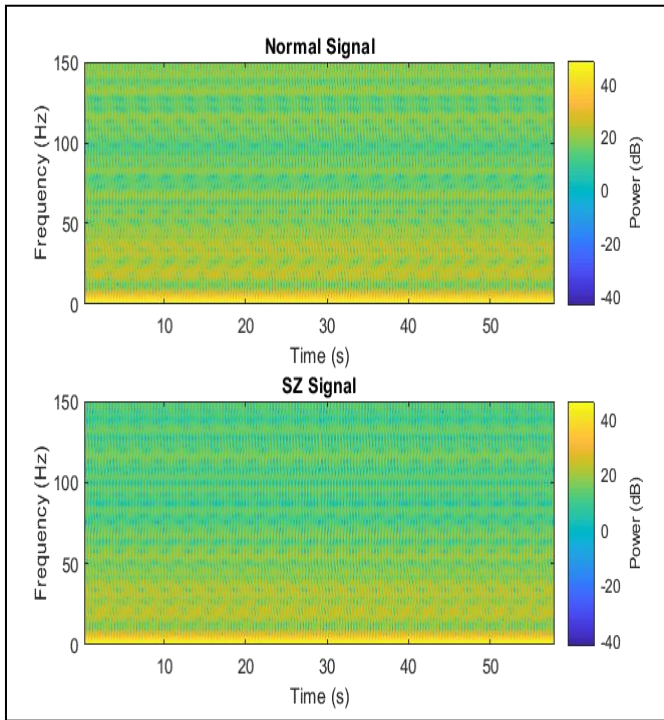


Fig.6 Representative of Spectrogram

In order to enhance the performance of classifier, the input sequence has been transformed to feature sequence.

The ROI time series has been transformed to spectrogram by applying STFT and two time moment features were obtained from these spectrograms. Representative of spectrogram is shown in Fig.6.

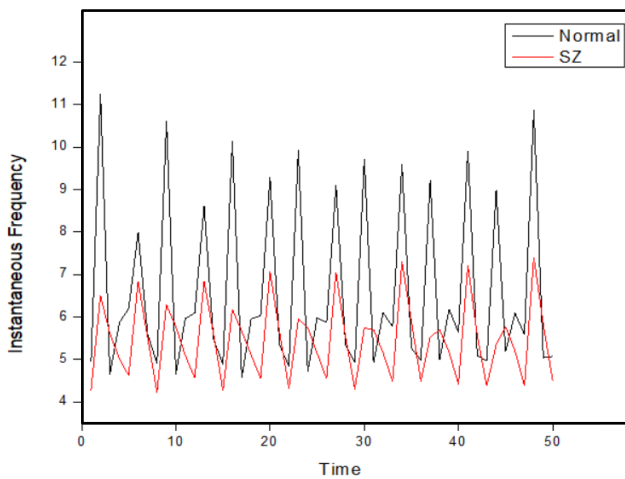


Fig.7 Representative of Instantaneous frequency

These time moment features are concatenated and given as a sequential input to the network. Representative images of the time moment features are shown in Fig.7 and Fig.8.

The obtained accuracy of the model is 81.3% whereas sensitivity and specificity are 60% and 90.9 % respectively. The training and testing accuracy of the model is shown in Fig.9. whereas the loss incurred during the training and the testing cycle is shown in Fig.10.

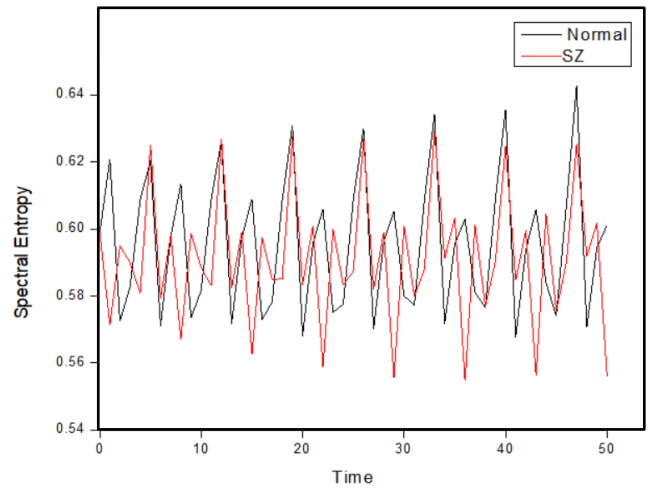


Fig.8 Representative of Spectral frequency

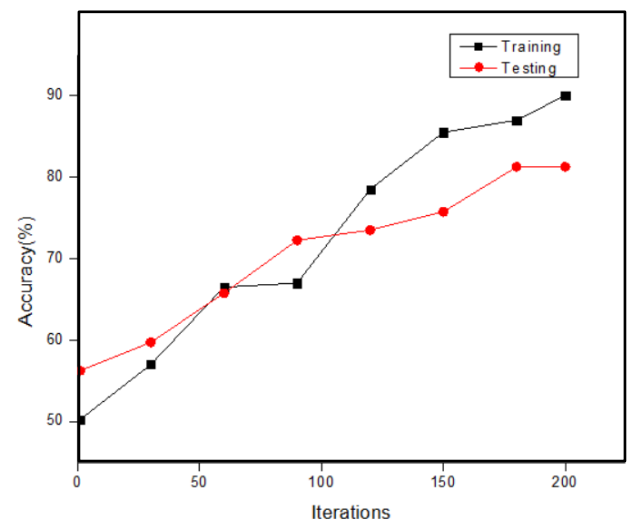


Fig.9 Training and Testing Accuracy of the model

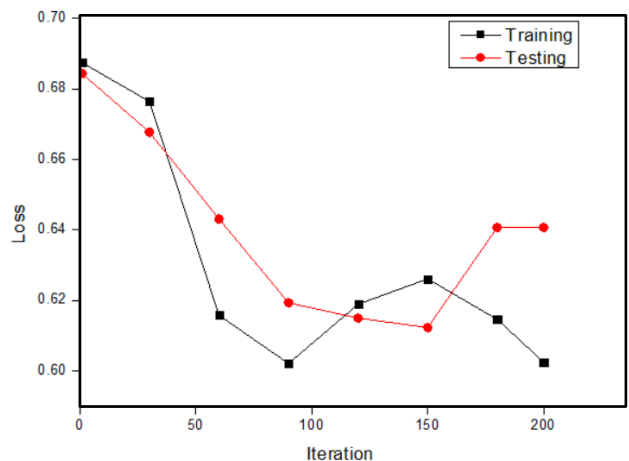


Fig.10 Training and Testing Loss

Table II depicts the comparison of SZ analysis with various classification methods. In the existing literature, schizophrenia discrimination using fMRI ROI time courses has been attempted with RNN-GRU, which yields accuracy

of 80.1% [12]. Analysis of ROI time courses with RNN-LSTM for schizophrenia classification has been addressed in the proposed work with accuracy of 81.3%. Although LSTM utilizes larger memory and more parameters, it performs efficiently with higher accuracy for larger sequences and it can further be improved with inclusion of larger training samples.

TABLE II. Comparison with existing methods

Method	Accuracy(%)
fMRI voxelwise signal classification with RNN[26]	64.9
fMRI voxelwise signal classification with LSTM[26]	66.4
fMRI ROI time series classification with RNN-GRU[12]	80.3
fMRI ROI time series classification with RNN-LSTM (proposed method)	81.3

#### IV. CONCLUSION

In this work an attempt has been made to analyze ROI time series of fMRI using RNN-LSTM network and to discriminate schizophrenics and Normal subjects. fMRI of 20 subjects each from CN and SZ were utilized. The bidirectional LSTM network with direct time sequence input performs with an accuracy of 77.8% whereas the time moment feature input yield 81.3% accuracy. The usage of spectral features also results in feature reduction which enhances the learning rate. Results demonstrate that, LSTM based RNN network using ROI time series is able to characterize fragments of sequential data than GRU and can be a promising aspect in identification of schizophrenia.

In this work a small number of samples are used for training the model. The performance of the model can further be improved by training with more number of samples.

#### REFERENCES

[1] Crow, Tim J. "Molecular pathology of schizophrenia: more than one disease process?." *British medical journal* 280.6207 (1980): 66

[2] Han, Shaoqiang, et al. "Recognition of early-onset schizophrenia using deep-learning method." *Applied Informatics*. Vol. 4. No. 1. SpringerOpen, 2017.

[3] Chong, S. A. "An Overview of Schizophrenia." (2017)..

[4] Kumari, Suneeta, et al. "An assessment of five (PANSS, SAPS, SANS, NSA-16, CGI-SCH) commonly used symptoms rating scales in schizophrenia and comparison to newer scales (CAINS, BNSS)." *Journal of addiction research & therapy* 8.3 (2017).

[5] Yang, Honghui, et al. "A hybrid machine learning method for fusing fMRI and genetic data: combining both improves classification of schizophrenia." *Frontiers in human neuroscience* 4 (2010): 192.

[6] Buckner, Randy L. "Event-related fMRI and the hemodynamic response." *Human brain mapping* 6.5-6 (1998): 373-377.

[7] Woodward, Neil D., and Carissa J. Cascio. "Resting-state functional connectivity in psychiatric disorders." *JAMA psychiatry* 72.8 (2015): 743-744.

[8] Yu, Qingbao, et al. "Brain connectivity networks in schizophrenia underlying resting state functional magnetic resonance imaging." *Current topics in medicinal chemistry* 12.21 (2012): 2415-2425.

[9] Gohel, B., et al. "MEG Based Functional Connectivity: Application of ICA to Alleviate Signal Leakage." *Irbm* 38.3 (2017): 127-137.

[10] Chung, Junyoung, et al. "Empirical evaluation of gated recurrent neural networks on sequence modeling." *arXiv preprint arXiv:1412.3555* (2014).

[11] Hochreiter, Sepp, and Jürgen Schmidhuber. "Long short-term memory." *Neural computation* 9.8 (1997): 1735-1780.

[12] Yan, Weizheng, et al. "Discriminating schizophrenia using recurrent neural network applied on time courses of multi-site fMRI data." *EBioMedicine* 47 (2019): 543-552.

[13] Tan, Jen Hong, et al. "Application of stacked convolutional and long short-term memory network for accurate identification of CAD ECG signals." *Computers in biology and medicine* 94 (2018): 19-26.

[14] Gruber, Nicole, and Alfred Jockisch. "Are GRU cells more specific and LSTM cells more sensitive in motive classification of text?." *Frontiers in Artificial Intelligence* 3.40 (2020): 1-6.

[15] Ashburner, John, et al. "Statistical parametric mapping." (1994).

[16] Whitfield-Gabrieli, Susan, and Alfonso Nieto-Castanon. "Conn: a functional connectivity toolbox for correlated and anticorrelated brain networks." *Brain connectivity* 2.3 (2012): 125-141.

[17] Wang, Jinhui, et al. "GRETNA: a graph theoretical network analysis toolbox for imaging connectomics." *Frontiers in human neuroscience* 9 (2015): 386.

[18] Lynn, Tan Jo, and Ahmad Zuri bin Sha'ameri. "Comparison between the performance of spectrogram and multi-window spectrogram in digital modulated communication signals." *2007 IEEE International Conference on Telecommunications and Malaysia International Conference on Communications*. IEEE, 2007.

[19] Yuan, Ye, et al. "A multi-view deep learning method for epileptic seizure detection using short-time fourier transform." *Proceedings of the 8th ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics*. 2017.

[20] Fulop, Sean A., and Kelly Fitz. "Algorithms for computing the time-corrected instantaneous frequency (reassigned) spectrogram, with applications." *The Journal of the Acoustical Society of America* 119.1 (2006): 360-371.

[21] Lin, Jianhua. "Divergence measures based on the Shannon entropy." *IEEE Transactions on Information theory* 37.1 (1991): 145-151.

[22] Yan, Shi. "Understanding LSTM and its diagrams." *MLReview.com* (2016).

[23] Kaushik, Pallavi, et al. "EEG-based age and gender prediction using deep BLSTM-LSTM network model." *IEEE Sensors Journal* 19.7 (2018): 2634-2641.

[24] Glorot, Xavier, and Yoshua Bengio. "Understanding the difficulty of training deep feedforward neural networks." *Proceedings of the thirteenth international conference on artificial intelligence and statistics. JMLR Workshop and Conference Proceedings*, 2010.

[25] Kingma, Diederik P., and Jimmy Ba. "Adam: A method for stochastic optimization." *arXiv preprint arXiv:1412.6980* (2014).

[26] Dakka, Jumana, et al. "Learning neural markers of schizophrenia disorder using recurrent neural networks." *arXiv preprint arXiv:1712.00512* (2017).

# Segmentation of Spinal Canal using Active Contour Model

Gayathry S Warriar

Department of Computer Science and Engineering  
Noorul Islam Centre for Higher Education  
Kumaracoil, Tamil Nadu, India  
gayathryvarier@gmail.com

Dr K.S.Angel Viji

Department of Computer Science and Engineering  
College of Engineering, Kidangoor  
Kerala, India  
ksangelviji@gmail.com

**Abstract**— Today, the lifestyle of people has paved the way for rise in spinal cord disorders. Severe cases have been reported which could be treated if diagnosed early. Scoliosis is a deformity to spine and ribs which is the primary cause of spinal curvature. The major challenge of scoliosis disease is the unnoticeable change in the orientation of spinal column at its early stage. Moreover, it is visually detectable only in the prodromal stage. The early diagnosis could help cure the disease through exercises and minor surgeries. Depending on manual diagnosis techniques is a tiring task as it can deliver inaccurate results. An automatic segmentation method which helps in the early diagnosis is proposed in this paper. Initially, CT image which is the input is fed into the system. CT images have high contrast between bone and adjacent tissues. Sagittal view datasets have been chosen in order to calculate the Cobb angle for the measurement of scoliosis intensity. Further, distortions are removed from the image and pre processing is performed followed by K-means clustering which detects the spinal canal. In order to segment the required features, the output of clustering is loaded to Active Contour Model. Finally, segmentation of spinal canal is completed. Experimental results prove the accuracy of 95%, 86.86%, 92.22% for Lumbar Vertebrae CT, Lumbar spine CT, Lumbar spine CT with multiple compression fractures respectively for the proposed system which is greater than traditional diagnosis methods. Subsequently, this would be a revolutionary study which assists the doctors for the early diagnosis of the disease.

**Keywords**— *Scoliosis, sagittal view, K-means clustering, Active Contour Model, spinal curvature.*

## I. INTRODUCTION

The progress in technology has altered the lifestyle of people which resulted in spinal disorders. People hardly identify the disease at its early stage. Subsequently, the disease becomes the part of their life as it cannot be cured as whole. Early diagnosis could save the patient affected with the disorders and give a complete cure. The major challenge in disease detection is the unnoticeable change in early stages. Today, various techniques are available for the early diagnosis. Relying on manual detection techniques could provide inaccurate results. Spinal curvature defects refer to the deformity in the column. The vertebrae in human body constitute thoracic, cervical, lumbar and sacro coccygeal vertebrae. The deformities in these vertebrae results in spinal disorders which further affects the posture and body movement. The paper mainly focuses on the patients

affected by scoliosis disease. The experiment is conducted based on the various stages of scoliosis disease. Scoliosis is a disease affecting spine which cannot be cured if not diagnosed early. It is commonly observed in thoracolumbar region.

Automatic segmentation is necessary for the early diagnosis as various studies validate this. Specific diagnostic features are hard to identify through manual techniques or planar imaging techniques. For easy identification of these small diagnostic factors which help in effective treatment, computer aided or automated techniques can be adopted. In Computer-aided diagnosis, the accuracy is high and has less computational time. The paper aims at implementing an automatic segmentation method for the early diagnosis of spinal disorders. The paper is parted as follows; Section II deals with the discussion of various studies performed by researchers. The section II demonstrates the proposed system and the theories regarding the proposed system. The experimental analysis and results have been discussed in section IV followed by a conclusion in section V.

## II. LITERATURE REVIEW

Various studies have been performed in the field of segmentation of spinal canal for the diagnosis of spinal disorders. [1] Proposed a repetitive segmentation which utilizes the whole convolution neural network segment. It analyzes images with the help of data from both image as well as memory in order to identify the adjacent vertebrae. Dice score of vertebrae segmentation was  $94.9 \pm 2.1\%$ . The paper [2] discusses about the automated global level set approach used for the image segmentation of lumbar vertebrae. In this method, hybrid morphological filter and Gaussian mixture model is used for the generation of smooth contour. Quantitative comparisons were performed for validating the proposed result. A level set method based on active contours is illustrated in paper [3]. The method uses minimization of energy functional whose energy parameters are weighted based on the relevance in detecting boundaries. The paper [4] elucidates the various properties within noise and edge points. In order to distinguish noise and edge points, local regional properties were used. In accordance with the local regional properties, a variable regional coefficient and an enhanced ESF is implemented [5]. Computerized Tomography images are analyzed for the detection of spinal disorders and the techniques in [8, 16] discusses about iterative topology refinement and knowledge based automatic segmentation. [7] discusses about a learning based edge detection approach for identifying defects in thoracic vertebrae providing an accurate spinal segmentation.

Numerous segmentation approaches using MRI images are briefed in papers [10-15]. Different approaches are studied regarding the detection of spinal disorders and some are categorized under efficient and robust techniques. Relying on such effective techniques would contribute to the effective early diagnosis of the spinal disorders.

### III. PROPOSED SYSTEM

The study on the early diagnosis of spinal disorders mainly focused on the methods which are reliable, easy to operate and would provide accurate results. It led to the development of an automatic segmentation method. The schematic diagram of the proposed method is shown in Fig 1. In this approach, CT images of the spinal canal are obtained which fed into the system as the input. In computerized tomography, a narrow beam of X-rays are targeted at a patient and is rotated around the body which provide signals processed by computer to deliver cross sectional images of the body. CT images are preferred due to its high contrast within bone and adjacent tissues. Moreover, spine segmentation can be easily performed. Further, preprocessing is done for the removal of distortions in image. Preprocessing is an important step to be performed as it involves the sorting of required data set from larger database available and also to enhance the dataset. Sagittal view datasets have been chosen in order to calculate the Cobb angle for the measurement of scoliosis intensity. K-means clustering is performed for the identification of unknown clusters in the data set which helps in the detection of spinal canal. K-means algorithm determines the k number of clusters and each cluster belongs to the cluster with adjacent mean. In order to segment the required features, the output of clustering is loaded to Active Contour Model. Finally, segmentation of canal is implemented which identifies each area of the spine individually.

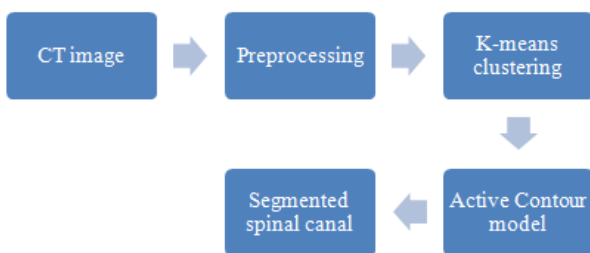


Fig 1: Schematic diagram of proposed system

#### A. Scoliosis

Scoliosis is a deformity to spine and ribs which is the primary cause of spinal curvature. Fig. 2 shows the difference between a normal and scoliosis affected spine. It is mainly seen in the thoracolumbar area. Scoliosis is a three-dimensional deformity, which can be a single C-shaped or S-shaped curve. The major challenge in scoliosis affected patients is the unnoticeable changes at its early

stages. Therefore, a patient hardly gets cured completely. The early diagnosis of scoliosis can cure the disease with some exercises and minor surgeries. The proposed system aims at the early diagnosis of spinal disorders and is tested in scoliosis affected patients.

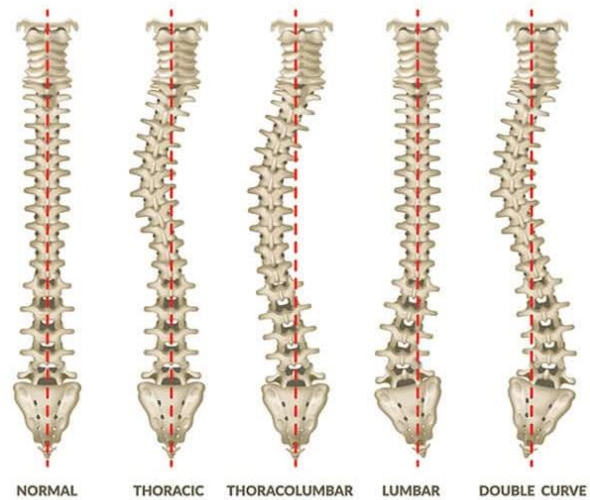


Fig. 2. Difference between a normal and types of scoliosis affected spine [17]

#### B. CT Image

Computerized Tomography is a prevailing imaging modality which combines two dimensional x-ray images taken from various angles around the body and creates three dimensional cross sectional images with the help of computer processing. It possesses least computational time and is less expensive. In the proposed system, CT images are preferred over MRI due to its high contrast within bone and adjacent tissues. Moreover, CT is a conventional method; it is widely used for spine related issues. Therefore, our study will be also based on CT images. The data sets can be of axial and sagittal. In the proposed system, sagittal view is adopted. In the sagittal view, the Cobb angle is calculated, it is the standard angle used to measure the intensity of scoliosis. At first, a parallel line is drawn to the greater vertebral end plate from the most tilted vertebra at the upper part of the curvature of the spine. Then a second parallel line is drawn to inferior vertebral endplate from the most slanting vertebra at bottom of the curvature of the spine. The angle formed by the perpendiculars drawn from the parallel line is called the Cobb angle as shown in Fig. 3. Larger the angle larger will be the intensity of the disease.



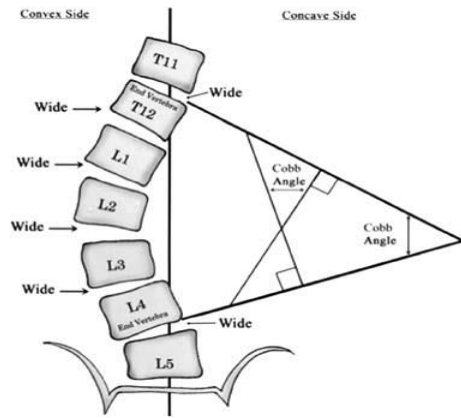


Fig. 3. Sagittal view of the scoliosis affected spinal canal and Cobb angle

### C. Preprocessing

Preprocessing is an important method which should not be neglected. Required dataset is sorted out from the larger database and distortions are removed from the image. In the proposed system, Gaussian filter is adopted for the removal of distortions. Gaussian filter is a linear filter which blurs the image or reduces the distortion as per the requirement.

### D. K- means Clustering

K-means clustering [6] is performed for the identification of K clusters in the data set. It segments input image into relevant feature sets called k clusters required for the investigation. The process of K- means Clustering is as follows:

Step 1: Random selection of K-clusters.

Step 2: Calculation of Euclidean distance (d) from cluster centre to each pixel of the input image. The equation for Euclidean distance is;

$$d = \|P(x, y) - C_k\| \quad (1)$$

Where  $C_k$  is the new cluster centre;  $P(x, y)$  is the pixel value of the cluster in x and y coordinates.

Step 3: Based on the Euclidean distance, each pixel is allocated to a cluster with adjacent center,

Step 4: Recalculation of position of centre using the equation;

$$C_k = \frac{1}{k} \sum_{y \in C_k} \sum_{x \in C_k} P(x, y) \quad (2)$$

Step 5: The process is continued until it converges.

Step 6: The cluster pixels are further joined to an image.

### E. Active Contour Model

Active contour model divides the required feature from a distorted input image [9]. It involves curve of energies. The curve can be moulded to any shape to extract the required image from the input image. There are many methods to achieve the outline of the required area of the input image. Initially, the model needs an initial user approximation of the outline of the required object. Later the

curve is deformed to get the object outline as shown in Figs.4

To get the object outline the weight of the energies is varied and the actual value which gives the required image is found. These weights control the rate of contour minimization to get the object outline. This method is called the final approximation of the contour. Here, the energy value is changed by varying stimulated forces. When these forces reach equilibrium the minimal state is obtained which gives the object outline and finally the segmented spinal canal image is obtained for early diagnosis of spinal defects.

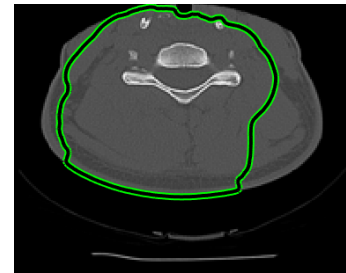


Fig. 4. Active Contour detection result the required area is obtained by removing the noise is represented inner side of the green circle

## IV. EXPERIMENTAL RESULTS

The dataset extracted from typical CT studies which includes multiple, serial and axial computed tomography images derived from thorax and abdomen of Scoliosis patients were used. The images are of size 1280×720 pixels and in jpg format. The results of segmentation are tabulated and the segmentation results of lumbar spine and lumbar vertebrae of various postures of patients are illustrated in Fig 5 and fig 6 respectively. In Fig 5, red color indicates the spinal canal whereas red color in Fig 6 indicates the vertebrae for men.

Fig 5. Segmentation results of lumbar spine of various postures of patients

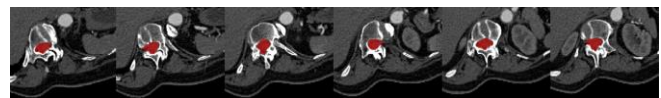


Fig6. Segmentation results of lumbar vertebrae of various postures of patients

Table I demonstrates the performance analysis of the proposed system with various properties. The properties analysed are a true positive, true negative, FP, FN sensitivity,

specificity and accuracy of the proposed system. The system has an overall accuracy of 92.22 with CT of lumbar spine having multiple compression fractures and accuracy of 95 with lumbar vertebrae CT.

TABLE I. PERFORMANCE ANALYSIS OF THE PROPOSED SYSTEM WITH DIFFERENT PROPERTIES

Parameters	Lumbar Vertebrae CT	CT of lumbar spine	CT of Lumbar spine with multiple compression fractures
TP	57	54	56
TN	3	5	4
FP	38	32	27
FN	2	8	3
Sensitivity	0.966102	0.870968	0.949153
Specificity	0.073171	0.135135	0.129032
Accuracy	0.95	0.868687	0.922222

## V. CONCLUSION

Scoliosis is a disease which can be cured when diagnosed early. The proposed method aimed at the early diagnosis of spinal disorders. The scoliosis patients were taken for detailed study about the implementation of the system. It is highly important to identify which minor deformities will eventually result in major deformity the importance of our system is that it helps in early diagnosis of the disease that too in the very starting stage of the disease when there is only a slight variation in the angle of orientation of the spine. For scoliosis, earlier the treatment is started the more effective the treatment will be, this shows how important is the early diagnosis of scoliosis. The automatic spinal canal segmentation system will help in diagnosing scoliosis at the early prodromal stage itself. This not only decreases the need for surgery but also the disease can be easily controlled through exercise and minor treatments. Experimental results validate the accuracy and efficiency of the proposed system.

## REFERENCES

[1] Lessmann, Nikolas, et al. "Iterative fully convolutional neural networks for automatic vertebra segmentation." arXiv preprint arXiv:1804.04383 (2018).

[2] Li, Yang, et al. "Automatic Global Level Set Approach for Lumbar Vertebrae CT Image Segmentation." *BioMed research international* 2018 (2018).

[3] A. Khadidos, V. Sanchez and C. Li, "Weighted Level Set Evolution Based on Local Edge Features for Medical Image Segmentation," in *IEEE Transactions on Image Processing*, vol. 26, no. 4, pp. 1979-1991, April 2017, doi: 10.1109/TIP.2017.2666042.

[4] C. Liu, W. Liu, and W. Xing, "An improved edge-based level set method combining local regional fitting information for noisy image segmentation," *Signal Processing*, vol. 130, pp. 12–21, 2017

[5] Sekuboyina, A., Valentinitich, A., Kirschke, J.S., Menze, B.H., 2017. A localisation-segmentation approach for multi-label annotation of lumbar vertebrae using deep nets. arXiv:1703.04347.

[6] S. Madhukumar and N. Santhiyakumari, "Evaluation of k-Means and fuzzy C-means segmentation on MR images of brain", *The Egyptian Journal of Radiology and Nuclear Medicine*, vol. 46, no. 2, pp. 475-479, 2015.

[7] J. MA, L. LU: Hierarchical Segmentation and Identification of Thoracic Vertebra Using Learning-based Edge Detection and Coarse-to-Fine Deformable Model. *Comput. Vis Image Understand*, **117** (9), 1072 (2013).

[8] N. Archip, P.-J. Erard, M. Egmont-Petersen, J.-M. Haefliger, and J.-F. Germond, "A knowledge-based approach to automatic detection of the spinal cord in ct images," *Medical Imaging, IEEE Transactions on*, vol. 21, no. 12, pp. 1504–1516, 2002.

[9] De Leener, B., Taso, M., Cohen-Adad, J. *et al.* Segmentation of the human spinal cord. *Magn Reson Mater Phy* **29**, 125–153 (2016).

[10] Zhigang Peng, Jia Zhong, William Wee and Jing-huei Lee, "Automated Vertebra Detection and Segmentation from the Whole Spine MR Images," 2005 IEEE Engineering in Medicine and Biology 27th Annual Conference, Shanghai, China, 2005, pp. 2527-2530, doi: 10.1109/IEMBS.2005.1616983.

[11] R. Shi, D. Sun, Z. Qiu and K. L. Weiss, "An Efficient Method for Segmentation of MRI Spine Images," 2007 IEEE/ICME International Conference on Complex Medical Engineering, Beijing, China, 2007, pp. 713-717, doi: 10.1109/ICME.2007.4381830.

[12] El Mendili, M.-M., Chen, R., Turet, B., Péligrini-Issac, M., Cohen-Adad, J., Lehericy, S., Pradat, P.-F. and Benali, H. (2015), Validation of a semiautomated spinal cord segmentation method. *J. Magn. Reson. Imaging*, 41: 454-459.

[13] Zhu, X., He, X., Wang, P. *et al.* A method of localization and segmentation of intervertebral discs in spine MRI based on Gabor filter bank. *BioMed Eng OnLine* **15**, 32 (2016).

[14] C. McIntosh, G. Hamarneh, M. Toom and R. C. Tam, "Spinal Cord Segmentation for Volume Estimation in Healthy and Multiple Sclerosis Subjects Using Crawlers and Minimal Paths," 2011 IEEE First International Conference on Healthcare Informatics, Imaging and Systems Biology, San Jose, CA, USA, 2011, pp. 25-31, doi: 10.1109/HISB.2011.42.

[15] P. D. Barbieri, G. V. Pedrosa, A. J. M. Traina and M. H. Nogueira-Barbosa, "Vertebral Body Segmentation of Spine MR Images Using Superpixels," 2015 IEEE 28th International Symposium on Computer-Based Medical Systems, Sao Carlos, Brazil, 2015, pp. 44-49, doi: 10.1109/CBMS.2015.11.

[16] Q. Wang *et al.*, "Automatic Segmentation of Spinal Canals in CT Images via Iterative Topology Refinement," in *IEEE Transactions on Medical Imaging*, vol. 34, no. 8, pp. 1694-1704, Aug. 2015, doi: 10.1109/TMI.2015.2436693.

[17] <https://www.towsonortho.com/specialties/scoliosis-spine/scoliosis/>

# DCGAN based Pre-trained model for Image Reconstruction using ImageNet

Nandini Kumari

Dept. of Computer Science and Eng.  
Birla Institute of Technology  
Mesra, Ranchi, India  
missnandinikumari@gmail.com

Shamama Anwar

Dept. of Computer Science and Eng.  
Birla Institute of Technology  
Mesra, Ranchi, India  
shamama@bitmesra.ac.in

Vandana Bhattacharjee

Dept. of Computer Science and Eng.  
Birla Institute of Technology  
Mesra, Ranchi, India  
vbhattacharya@bitmesra.ac.in

**Abstract**—Despite recent achievements in generative image modeling, generating better quality image samples from complex datasets such as ImageNet remains an illusory goal. The objective of this paper is to train Deep Convolutional Generative Adversarial Network at the well known CIFAR10 dataset, and study the instabilities specific to such scale and then test the large scale ImageNet dataset for establishment of the proposed DCGAN. We find that applying a pre-trained DCGAN can remove the complexity and also can learn prior details of images and improve the quality of generated image. Our modifications on DCGAN lead to models which set the new cutting edge in class-contingent image reconstruction on pre-trained GAN's. When tested on ImageNet at 128×128 resolution, our model (DCGAN) lowers the loss between the generated and real image samples which shows that the proposed DCGAN model works well with both the datasets.

**Index Terms**—Deep Convolutional Generative Adversarial Network, Generator, Discriminator, ImageNet, Image Reconstruction

## I. Introduction

Algorithms based on machine learning essentially deals with unraveling data from numerous available real-world sources and revealing the different types of patterns, for example, in classification tasks, object detection and segmentation etc. Recently, a variety of dynamic field explore the possibility of how to produce such genuine information through certain generative models. Artificial or synthetic information generation could be utilized for enlarging the training space by creating artificial but natural looking data samples that are excluded from the original data. For this purpose deep learning methods such as convolutional neural network has been embedded to generative modeling by researchers. Generative models use unsupervised learning techniques that include consequently find and learn the patterns and representation of input image samples so that the model can be utilized to create or yield new models that conceivably might have been drawn from the original dataset. Image generation from a latent feature space is what this paper investigates. Image reconstruction is an active research field and with the collateral emergence of deep generative models, two victorious and vitally divergent methods have been allured the attention of the researchers showing optimistic results:

Variational Autoencoders (VAE) [1] and Generative Adversarial Networks (GAN) [2]. The previous classification follow a clear methodology: they utilize a traditional autoencoder (for example an encoder/decoder conspire), however the noise vector is added to the transitional representation to inforce a Gaussian distribution. This way the decoder learns to produce images by following explicit information details conveyances extricated from the input space without compacting the detailed information. Thus, the initial input space contains the finest distribution (distinguished by the encoder), at that point the generator yields a realistic image that contains a visual component complementary to that dissemination.

Generative Adversarial Networks can recreate test samples from complex image space. They comprise of two sub-models: a discriminator which intends to isolate real images from fake (or recreated) images, and a generator which is at the same time streamlined to reproduce sample images that classified as realistic by the discriminator. The two models are trained simultaneously in an adversarial circumstance until the discriminator model is deceived about a fraction of the time, which implies that the conceptual model can be build using generator models. GANs are a stimulating and rapidly developing field, conveying on the assurance of generative models in their ability to make pragmatic models across a scope of real-world issues, most eminently in image to image transition, for instance, making a translation of photos of winter to autumn or noon to night, and in creating photograph sensible photos of articles, scenes and people that even individuals can not distinguish being real or fake. Deep learning has attained magnificent outcomes for classification tasks using discriminative models for which enormous datasets exists; for instance on the ImageNet dataset which comprises of over 1M pictures [3]. Howbeit, for some issues the number of data is not adequate to prepare large number of parameters typically available in these deep based networks. Favorably, deep learning strategies has proved that discovered patterns and information contained on an enormous dataset, (for example, ImageNet) can be easily moved to other computer vision tasks. These networks are either used as feature extractors or for fine

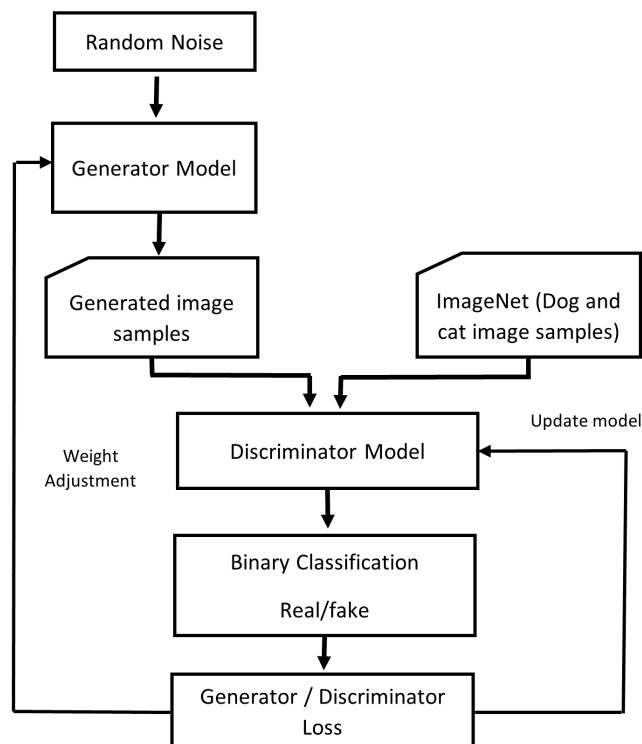


Fig. 1: Proposed DCGAN architecture

tuning by adapting the pre-trained network to adjust the parameters according to the new domain dataset. In deep neural network, the pre-trained network is utilized for initializing the fine tuned weight adjustments for other tasks (viably moving the information gained from the source area). It has also been indicated that less train data were needed to prepare models which were initiated with a pre-trained network. The strategy of utilizing a pre-trained models for reinstatement – that is extremely mainstream for discriminative models is the most awesome aspect of our insight which is not utilized for GANs. Howbeit, as in the instance of discriminative models, the quantity of hyperparameters in a GAN is huge and complex. Particularly in some research areas which need numerous training data, the utilization of pre-trained GANs could essentially enhance the nature of the produced data.

Subsequently, the current work, the utilization of pre-train models for GANs are evaluated. The paper has the following accompanying contributions:

- 1) A few pre-trained parameter's arrangements are evaluated, and described that pre-trained models can viably quicken the learning cycle and give helpful earlier information when training data is limited.
- 2) We trained a Deep Convolutional Generative Adversarial Network (DCGAN) on CIFAR-10 dataset and then utilize that pre-trained DCGAN to evaluate and generate images for two class of ImageNet.

The overall workflow of proposed DCGAN architecture is depicted in Figure 1. The following sections of the paper presents an extensive literature review, followed by the methodology section including dataset description, the proposed architecture and the implementation of the model. At last, the results are documented followed by the conclusion section.

## II. Literature Review

Generative Adversarial Network (GAN) [2] has gain increasingly popularity in research theme in recent years and has achieved modernity performance for a variety of research works in computer vision case, such as video reformulation [4], image transition [5], [6], high-resolution reconstruction [7] and image regeneration [8], [9]. Generative Adversarial Networks have accomplished amazing outcomes in image regeneration [10], [11] via a two-competitor minimax game: a discriminator expects to recognize and try to distinguish the real images from generated images while a generator plans to produce realistic fake images to trick the discriminator. These two modules generally utilize the deep learning mechanization as per the domain area. For images and video data, convolutional neural network and for audio, music and other time series dataset Long Short Term Model has shown good performance. The architecture of GAN utilizes a progression of completely associated layers and in this way is restricted to small image datasets. When moving toward the reconstruction of realistic images of higher intricacy, convolutional models has become more appropriate alternative. Abruptly Deep Convolutional GANs (DCGAN) turned into the definitive GAN design for image reconstruction issues [11]. In DC-GAN, the generator consecutively up-samples the prominent features from raw data by utilizing transpose convolutions, though the discriminator utilizes ordinary convolutions to characterize the input images.

Recently, multi-scale ensemble structures [12], [13] can successfully produce high quality images. It was additionally discovered that these ensemble techniques can be utilized to improve the nature of the created dispersion [14]. A progression of multi-stage generative models has been proposed to produce more reasonable images [15]. [16] proposes Combined Generative Adversarial Network that unravels correlated components of images by utilizing numerous generators to produce various pieces of the picture. The Recursive GANs [17] figures out how to produce picture foundation and closer views independently and recursively. GANs have indicated an extraordinary accomplishment on an assortment of contingent image based applications, e.g., image to image interpretation [18] and text-to-image [12]. Conditional GAN (cGAN) models [19], [20] implements their condition in both sub-structure i.e. generator and discriminator by linking it to the contribution of the deep based layers, for example the random noise vector fed to initial layer or the learned intermediate components for the interior layers. The AC-

GAN system [21] broadens the discriminator with a helper decoder to reproduce class-restrictive data. Likewise, InfoGAN [22] remakes a subset of the inert factors from which the samples were produced.

### III. Methodology

#### A. Dataset

In this paper two datasets has been used. The CIFAR10 dataset is used for training the GAN network and the ImageNet dataset for testing the network on the pre-trained GAN network. The CIFAR10 comprises of  $32 \times 32$  labeled color images originating from 10 unique classes, in which each class contains 60,000,  $32 \times 32$  pixel color images [23]. The class labels and their standard associated integer values are listed here, 0: airplane, 1: automobile, 2: bird, 3: cat, 4: deer, 5: dog, 6: frog, 7: horse, 8: ship, 9: truck. In the current work, 50,000 images are implemented for training, and 10,000 images with image size  $32 \times 32$  are utilized in testing process.

ImageNet is a dataset of over 15 million labeled high-resolution images corresponding to approximately 22,000 categories [3]. The images were collected from the web and labeled by human labelers using Amazon's Mechanical Turk crowd-sourcing tool. ImageNet consists of images with varied image resolutions, while in the current work a constant input dimensionality is needed. Therefore, the images were down-sampled to a fixed image pixels of  $128 \times 128$ . A total of 25,000 images (12,500 each for cat and dog class) has been used to test the pre-trained GAN model.

#### B. Image Reconstruction on Pre-trained DCGAN (Deep Convolutional Generative Adversarial Network)

Since the objective is to regenerate images from the pre-trained GAN model, two CNN modules has been embedded as generative and discriminative model. Specifically, according to the ImageNet Large Scale Visual Recognition Competition (ILSVRC) in recent years, it appears that improved accuracy results regarding image data, generally relies upon the intricacy of CNNs. In this paper, a DCGAN has been implemented and trained on CIFAR10 dataset and then the pre-trained DCGAN model has been tested on two class of ImageNet dataset to evaluate the performance of the pre-trained DCGAN. Here, the significant issues that DCGANs tackles are:

- 1) Discrimination is made so it essentially tackles a directed image discrimination task.
- 2) The feature representation learned by the GAN can be used to reconstruct the image.
- 3) Generator contains complex hyperparameters that can learn complex linguistic rendition of features.

DCGANs use essential standards of CNNs on the generator/ discriminator and have become the most generally utilized design because of their quick combination and furthermore because of the way they can be effortlessly adjusted into more perplexing variations. The Generator network takes an N-dimensional noise vector as input

and transforms it according to the learning parameters and tries to generate some realistic image pixels. The Discriminator then models a binary classification network and yields the possibility that predicts whether the input image is real(1) or fake(0). Considering this, we can characterize the two primary objectives of the generation task:

- 1) Retrain Generator according to the Discriminator's last characterization error rate. (So the reconstructed images look alike realistic).
- 2) Train Discriminator to limit the last classification error. (With the goal that real image data is effectively-recognized from fake image).

To accomplish this, during backpropagation, Generator hyperparameter's weights will be adjusted to minimize the overall loss.

#### C. DCGAN Generator

The proposed DCGAN generator and discriminator architecture is described in Figure 2 and 3 respectively. It tends to be seen from Figure 2, DCGAN generator takes starting information basically a (1, 100) noise vector, which goes through 4 Transpose Convolutional layers (Conv2DTranspose) and a stride rate of 2 to produce a resultant RGB image of size (128, 128, 3). The Transpose Convolutional layer is an inverted convolutional layer that will both upsample input and determine the input feature's details during the training of model. To accomplish this, the noise vector is reshaped into a  $8 \times 8 \times 128$  dimensional yield to coordinate the contribution of the initial Conv2D layer of discriminator which is depicted straightaway. A brief explanation with number of layers, size and different operation along with filter size implemented in DCGAN generator and discriminator model is summarized in Table I and II. The first Conv2DTranspose layer upsamples the input image to  $16 \times 16 \times 128$  with number of filters 128 and kernel size  $3 \times 3$  with stride rate 2. Similarly, other subsequent Conv2DTranspose layer carried the upsampling process with increasing number of filters (256,512,1024) with same stride and kernel size and yields  $128 \times 128 \times 1024$  feature details of the generated image from noise vector with other parameters. Finally, the generated image is fed to a Conv2D layer with filter size 3 to get the output with the image size  $128 \times 128 \times 3$  as real image size of the ImageNet dataset.

#### D. DCGAN Discriminator

The DCGAN Discriminator takes an input image size of (128, 128, 3), same as Generator's output. It passes through 4 standard Conv2DTranspose layers with a stride rate of 2. At the end layer the output image gets flattened to a vector, which is usually fed to a sigmoidal function, which then outputs the Discriminator's prediction for that image.

The first Conv2D layer receives the real images of dog and cat from ImageNet dataset and the generated

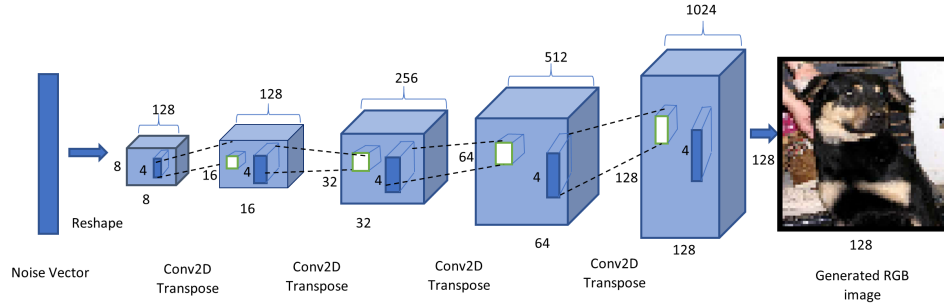


Fig. 2: Proposed DCGAN Generator structure

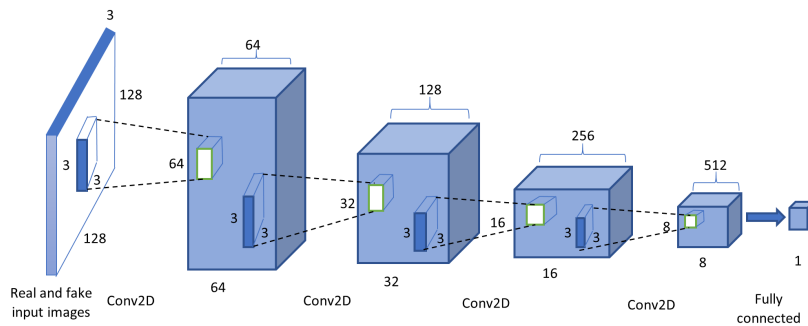


Fig. 3: Proposed DCGAN Discriminator structure

TABLE I: Architecture of DCGAN Generator

Layer	Input	Operation	Filter Size	strides	output
1	$8 \times 8 \times 128$	Dense			
2	$8 \times 8 \times 128$	Conv2DTranspose+leakyRelu	$128 \times 4 \times 4$	2	$16 \times 16 \times 128$
3	$16 \times 16 \times 128$	Conv2DTranspose+leakyRelu	$256 \times 4 \times 4$	2	$32 \times 32 \times 256$
4	$32 \times 32 \times 256$	Conv2DTranspose+leakyRelu	$512 \times 4 \times 4$	2	$64 \times 64 \times 512$
5	$64 \times 64 \times 512$	Conv2DTranspose+leakyRelu	$1024 \times 4 \times 4$	2	$128 \times 128 \times 1024$
6	$128 \times 128 \times 1024$	Conv2D	$3 \times 3$	2	$128 \times 128 \times 3$

fake images from the DCGAN generator, of dimension  $128 \times 128 \times 3$  (3 represents RGB images) as input and outputs a  $64 \times 64 \times 64$  feature map by applying 64 kernels with kernel size  $3 \times 3$  of stride 2, followed by second Conv2D layer with 128 kernel size, which filters out the resultant feature map of layer 1 more precisely and obtain  $32 \times 32 \times 128$  feature map. The third layer comprises of 256 kernels with  $3 \times 3$  kernel size with stride of 2 and yields  $16 \times 16 \times 256$  feature map followed by another Conv2D layer with  $512 \times 3 \times 3$  kernel size which suppresses the output to  $8 \times 8 \times 512$ . Finally the resultant feature map from the previous layer are fully connected with a dense layer of output neurons 1 and derive the probabilities for binary classification between real and fake image classes i.e. using 4% dropout and Sigmoidal function. Additionally, this network was trained with batch size of 32, Binary CrossEntropy loss function and Adam optimizer was used

for 1000 epochs. For the adam optimizer, the learning rate was set to 0.0002 and decay to 0.5. This architecture has been trained on CIFAR10 dataset and tested on ImageNet dataset.

#### IV. Result and Discussion

This paper proposes a Deep learning embedded GAN model namely DCGAN which utilizes the principle of CNN model for working of the generator as well as the discriminator. The objective of this paper is to regenerate images from the pre-trained DCGAN using Imagenet dataset as testing data for the proposed GAN model to enhance the resolution of reconstructed images. In this paper, CIFAR10 dataset is used as the source image space and it uses all the 10 classes for training the DCGAN. The proposed DCGAN trained with 50,000 images and tested on 25,000 images of dog and cat class

TABLE II: Architecture of DCGAN Discriminator

Layer	Input	Operation	Filter Size	strides	output
1	$128 \times 128 \times 3$	Conv2DleakyRelu	$64 \times 3 \times 3$	2	$64 \times 64 \times 64$
2	$64 \times 64 \times 64$	Conv2DleakyRelu	$128 \times 3 \times 3$	2	$32 \times 32 \times 128$
3	$32 \times 32 \times 128$	Conv2D+leakyRelu	$256 \times 3 \times 3$	2	$16 \times 16 \times 256$
4	$16 \times 16 \times 256$	Conv2D+leakyRelu	$512 \times 3 \times 3$	2	$8 \times 8 \times 512$
5	$64 \times 64 \times 512$	Dropout	0.4		
6	$8 \times 8 \times 512$	Dense	1(sigmoid)		$1 \times 65536$

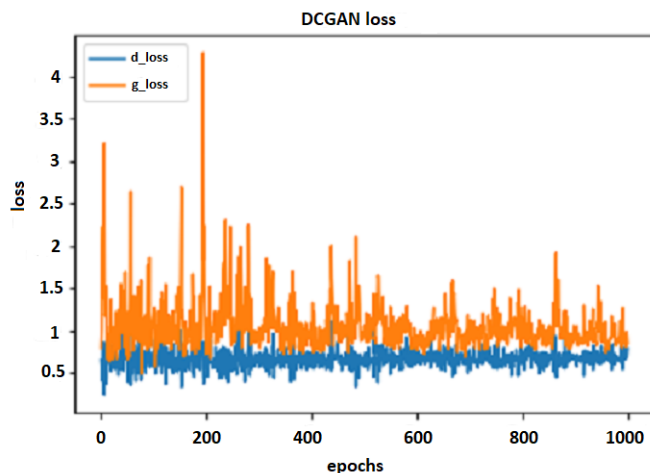


Fig. 4: DCGAN loss graph on ImageNet Dataset

taken from CIFAR10 and ImageNet datasets respectively. The weights are adjusted while training and then the learned parameters are transferred from the pre-trained DCGAN for better performance. The performance for DCGAN model is evaluated with the test set taken from ImageNet dataset. The DCGAN's generator takes input noise and learned parameters and tries to generate realistic images and fed these regenerated images along with the original image to the DCGAN's discriminator for binary classification between real and fake images. The performance of DCGAN is evaluated with the generator's and discriminator's loss depicted in Figure 4. It can be seen from loss graph of DCGAN model that the loss between generated image and the original image almost converges at 1000 epochs which shows that the DCGAN's generator have learned well and tries to generate realistic images. In Figure 5 some of the generated image samples of dog and cat are shown with the increasing epoch size. The quality of generated sample images increases with the number of epochs and is mostly comparable to the original image. The original dog and cat image sample taken from ImageNet dataset are also shown along with the generated image samples. The outcomes show an improvement when utilizing pre-prepared models, with higher resolution images and brings down the loss with all updated weight measurements, proposing that it better catches the genuine information dispersion of the ImageNet dataset contrasted to training from scratch.

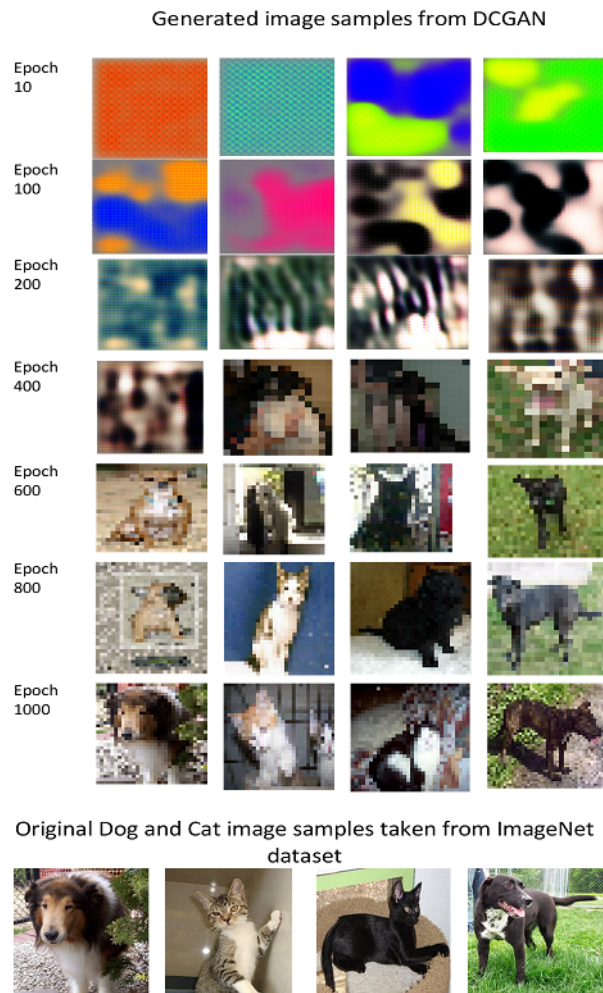


Fig. 5: Generated image from DCGAN on ImageNet dataset (Dog and Cat)

## V. Conclusion

In this paper, a DCGAN based on CNN model has been proposed for image reconstruction using two classes of ImageNet dataset. The modified CNN model is embedded on the generator and discriminator for better performance of the GAN model and named as DCGAN. The contribution of this work is in two essential viewpoints; the first is, a novel strategy DCGAN is utilized to regenerate the image samples for obtaining better results from the large scale datasets. Furthermore, to improve the efficiency and

remove the complexity of overall DCGAN, the model is pre-trained on CIFAR10 dataset and then tested on the ImageNet dataset, hence, generating images with good resolution. The acquired discriminator's and generator's loss and the generated images are also showing that the pre-trained DCGAN has shown significant improvement in images resolution and quality. This approach has also been established as a good practice for testing the datasets with a pre-trained GAN model.

## References

- [1] D. P. Kingma and M. Welling, "Auto-encoding variational bayes," arXiv preprint arXiv:1312.6114, 2013.
- [2] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial networks," arXiv preprint arXiv:1406.2661, 2014.
- [3] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "Imagenet: A large-scale hierarchical image database," in 2009 IEEE conference on computer vision and pattern recognition. Ieee, 2009, pp. 248–255.
- [4] C. Vondrick, H. Pirsivash, and A. Torralba, "Generating videos with scene dynamics," arXiv preprint arXiv:1609.02612, 2016.
- [5] C. Chen, W. Xie, Y. Wen, Y. Huang, and X. Ding, "Multiple-source domain adaptation with generative adversarial nets," Knowledge-Based Systems, vol. 199, p. 105962, 2020.
- [6] Y. Choi, M. Choi, M. Kim, J.-W. Ha, S. Kim, and J. Choo, "Stargan: Unified generative adversarial networks for multi-domain image-to-image translation," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2018, pp. 8789–8797.
- [7] C. Ledig, L. Theis, F. Huszár, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang et al., "Photo-realistic single image super-resolution using a generative adversarial network," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2017, pp. 4681–4690.
- [8] Z. Chen, C. Wang, H. Wu, K. Shang, and J. Wang, "Dmgan: discriminative metric-based generative adversarial networks," Knowledge-Based Systems, vol. 192, p. 105370, 2020.
- [9] H. Li, L. Zhang, B. Huang, and X. Zhou, "Sequential three-way decision and granulation for cost-sensitive face recognition," Knowledge-Based Systems, vol. 91, pp. 241–251, 2016.
- [10] E. Denton, S. Chintala, A. Szlam, and R. Fergus, "Deep generative image models using a laplacian pyramid of adversarial networks," arXiv preprint arXiv:1506.05751, 2015.
- [11] A. Radford, L. Metz, and S. Chintala, "Unsupervised representation learning with deep convolutional generative adversarial networks," arXiv preprint arXiv:1511.06434, 2015.
- [12] H. Zhang, T. Xu, H. Li, S. Zhang, X. Wang, X. Huang, and D. N. Metaxas, "Stackgan: Text to photo-realistic image synthesis with stacked generative adversarial networks," in Proceedings of the IEEE international conference on computer vision, 2017, pp. 5907–5915.
- [13] T. Karras, T. Aila, S. Laine, and J. Lehtinen, "Progressive growing of gans for improved quality, stability, and variation. arxiv e-prints (2017)," arXiv preprint arXiv:1710.10196, 2018.
- [14] Y. Wang, L. Zhang, and J. Van De Weijer, "Ensembles of generative adversarial networks," arXiv preprint arXiv:1612.00991, 2016.
- [15] D. J. Im, C. D. Kim, H. Jiang, and R. Memisevic, "Generating images with recurrent adversarial networks," arXiv preprint arXiv:1602.05110, 2016.
- [16] H. Kwak and B.-T. Zhang, "Generating images part by part with composite generative adversarial networks," arXiv preprint arXiv:1607.05387, 2016.
- [17] J. Yang, A. Kannan, D. Batra, and D. Parikh, "Lr-gan: Layered recursive generative adversarial networks for image generation," arXiv preprint arXiv:1703.01560, 2017.
- [18] C. Wang, C. Xu, C. Wang, and D. Tao, "Perceptual adversarial networks for image-to-image transformation," IEEE Transactions on Image Processing, vol. 27, no. 8, pp. 4066–4079, 2018.
- [19] V. Dumoulin, I. Belghazi, B. Poole, O. Mastropietro, A. Lamb, M. Arjovsky, and A. Courville, "Adversarially learned inference," arXiv preprint arXiv:1606.00704, 2016.
- [20] K. Sricharan, R. Bala, M. Shreve, H. Ding, K. Saketh, and J. Sun, "Semi-supervised conditional gans," arXiv preprint arXiv:1708.05789, 2017.
- [21] A. Odena, C. Olah, and J. Shlens, "Conditional image synthesis with auxiliary classifier gans," in International conference on machine learning. PMLR, 2017, pp. 2642–2651.
- [22] X. Chen, Y. Duan, R. Houthoofd, J. Schulman, I. Sutskever, and P. Abbeel, "Infogan: Interpretable representation learning by information maximizing generative adversarial nets," arXiv preprint arXiv:1606.03657, 2016.
- [23] A. Krizhevsky, G. Hinton et al., "Learning multiple layers of features from tiny images," 2009.



# Classification of Digital Dental X-ray Images Using Machine Learning

Sindu Divakaran

Research Scholar, School of Electrical&Electronics Engineering  
Sathyabama Institute of Science and Technology  
Chennai,India  
\*sindudiva@gmail.com

K. Vasanth

Professor & Head, Department of Electronics & Communication  
Engineering Vidya Jothi Institute of Technology,  
Hyderabad, India  
vasanthecek@gmail.com

Suja D

PG Scholar, Department of Biomedical Engineering  
line 3: Sathyabama Institute of Science and Technology  
Chennai,India

Swedha V

PG Scholar, Department of Biomedical Engineering  
Sathyabama Institute of Science and Technology  
Chennai,India  
swedhavelusamy@gmail.com

**Abstract**—Dental diseases like dental anomalies, periapical and dental caries is increasing day by day in children and adults. Artificial intelligence and neural network with its application in medical imaging is influencing the healthcare industry. X-ray imaging is the commonly employed technique to diagnose diseases of the teeth. Segmentation and classification of differing dental anomalies using neural network is proving to be a boon to the dental field. Application of neural algorithms aids in obtaining images with better detection accuracy. Automated detection reduces the workload of a dentist with classification being accurate. A better penetration of machine learning into these processes highlights its advantages to classify dental X ray images. Different machine learning techniques are deployed to identify and classify the dental abnormalities.

**Keywords**—Classification, Dental X-Ray, Machine Learning, Neural network

## I. INTRODUCTION

Tooth is a dense structure in the human body that decays due to many reasons. Countless conditions like Dental decay, Periodontal disease, Mesioangular impaction, Periapical abscess, Horizontal bone impaction, Vertical bone impaction, Apical periodontitis, Overhanging restoration, Irreversible pulpitis, Cast post restoration, Radiopaque restoration, Proximal caries are detected using an X ray.

Digital images of human body have gained more attention in the field of medical image analysis research. Numerous image processing techniques are emerging to find solutions to diseases found in human body. Human organs are complex and difficult to diagnose. [1] An easiest imaging modality for dense structure of the body is X-ray. These rays penetrate the bones and produce an image which aids in diagnosis. Once anomalies are detected in an X-Ray, the doctor diagnosis the problem and prescribe the therapy.

Digital radiographic image is an advanced X-ray assessment technique which produces images immediately on a computer. X-ray sensitive plates capture the images during investigation and transfer it to the computer immediately. The incident x-ray radiation is converted into an equivalent electric charge and by a detector sensor to a digital image.

Recognition and analysis of dental images have been made easier with the introduction of digital X ray images. The resolution, luminance, noises, contrast are different when different X ray machines are utilised to capture the

images. Analysis of teeth images can be successfully performed by segmentation of the tooth and it forms an important step for treatment planning. Segmentation of the dental images can be done using different techniques like k-means and dual clustering, subtraction of background, methods based on histograms, region growing methodologies, etc. which help to differentiate the normal from the pathologically affected parts of the teeth. [2]

Application of the fundamentals of machine learning in dental imaging is making it easier to segment and classify images. [3] Based on the efficiency and performance of classification algorithms, few of the promising ones like SVM, ANN, KNN are applied on dental data set images.

## II. LITERATURE SURVEY

Anuj Kumar, H. S. Bhadauria Nitin Kumar proposed Fuzzy Clustering with Level Set Segmentation for Detection of Dental Restoration [4] to extract the restoration part from the dental X-ray image by combining the Fuzzy clustering with the iterative level set active contour. Here median filter pre processes the image and segments using Fuzzy clustering.

Jiafa Mao, Kaihui Wang, Yahong Hu, Weiguo Sheng, Qixin Feng, presented Grab Cut algorithm [5] for dental X-ray images based on full threshold segmentation. They obtained the outline image set of Iwholen and Crowns. Morphological open operation and median filtering is applied to the synthetic image of contour and crown, and the resulted image used a Mask for Grab Cut to obtain the target tooth image.

Jie Yang, yuchen Xie, Lin Lui presented the Automated Dental Image Analysis Learning on Small Dataset [6] by Deep datasets, procedures, and results conducted to evaluate dental treatment qualities using periapical dental X-ray images taken before and after the operations. They acted as a support to clinicians to classify diseases as getting better, getting worse and no explicit change.

Said.E.H, Dias .E.M, Nasar .G.F proposed a methodology of Teeth segmentation in digitized dental X-Rays films using mathematical morphology. A grey scale contrast stretching transformation is done to improve the performance of teeth segmentation. They concluded that in addition to its capability of handling bitewing and periapical dental radiographs views, their approach exhibits the lowest failure rate [7][8].

### III. METHODOLOGY

The image dataset made use of in this work is a group of dental X-ray images gathered from dental hospitals, dental clinics and web source dental datasets. Different classes of dental caries and normal images are obtained and stored as database. Figure 1 depicts the stages of processing and analysis of the dental caries from digital X ray images. The unprocessed image is retrieved from the dataset and pre processed for enhancement of images and redundancy removed. This is followed by resizing of the images to reduce distortions and convert to gray scale images as processing of colours images will not give clear data.<sup>[9]</sup>

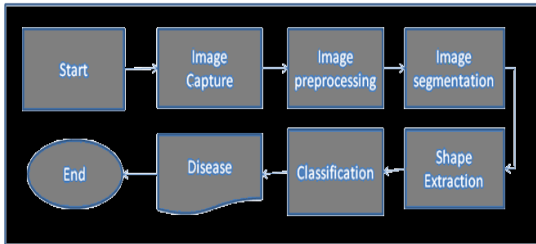


Figure 1: Basic Block Diagram of process

The pre-processed and resized images are segmented into non-intersecting homogeneous regions based on the characteristic pixels and similar attributes of the image. Location and identification of boundaries like lines, curves in images are attained by segmentation. It also determines the accuracy of the computer based output.

Feature extraction produces a reduced group of features which can ease the processing when datasets are large. The features selected usually contain the required information and desired task can be performed using the extracted data. This work uses GLCM for feature extraction. The GLCM function characterises an image's texture by calculating how often pairs of pixels with specific values and in a specified spatial relationship appear in an image, generating a GLCM, and extracting statistical measures from this matrix. Further the images are classified from the dataset fed for training Nd testing.<sup>[10][11]</sup> Classification aids us to differentiate the various classes of dental diseases and strategically diagnose the type of dental disease.

Support Vector Machine (SVM), Artificial Neural Network (ANN) and KNN (Kernal Nearest Neighbour) classification algorithm determines whether there are pathological signs of dental diseases in the analyzed image. SVM generated an optimal classified model by obtaining data from an existing trained set. A promising approach by the application of ANN in the field of dentistry to classify dental caries and impacted teeth created a big influence in image analysis. By considering the best match of new records with an already trained record system, a supervised classification algorithm like KNN is found to reduce complexity.<sup>[12]</sup> The newest neighbour is found by the Euclidean distance and helps to classify accordingly.

### IV. RESULTS AND DISCUSSIONS

This section describes the experimental results of the classification methods using digital dental X ray images for different classes of dental diseases. Dental images, belonging to various dental diseases like vertical impaction, periapical abscess, distal pulp horn caries, missed canal in root canal,

etc are stored in the dental image database. The images were collected from dental clinics and hospitals. Our work consists of a total of 500 images belonging to assorted set of dental caries and normal tooth images. One half of the images were used for training and the other half for the testing.

Digital X ray images of the abnormal teeth are loaded into the system for pre processing.



Figure 2: Input Image of teeth

The goal of pre-processing is to reduce or eliminate unwanted distortions in image data, as well as to improve some image features that are important for subsequent processing. Median filters are employed for preprocessing in this work followed by resizing to a specified dimension and gray scale conversion for a better representation of the images.

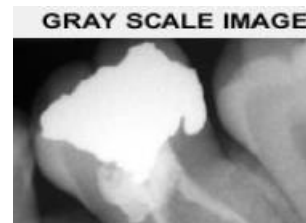


Figure 3: Gray Scale Dental X ray Image

Segmentation<sup>[13]</sup> is performed using Region of Interest (ROI) which separates the uninteresting pixels from the interesting pixels and reassigned with intensity values 0 or 1 for not required and required pixels respectively.



Figure 4: Segmented dental X ray Image

The main aim of feature extraction is to separate favourable characteristics from the image. Gray Level Co-occurrence Matrix (GLCM) is performed to extract the features. It extracts second order texture data from images.<sup>[14][15]</sup> The features extracted from the images stored in the dataset are Skewness, Smoothness, Energy, Entropy, Homogeneity, Performance, Contrast and Correlation along with shape and geometry.

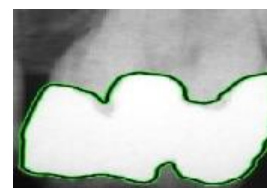


Figure 5: Featured Extracted Image

The Table 1 below gives the extracted feature of some of the dental diseases

Table 1: Extracted feature values of diseased teeth

Disease	Skewness	Smoothness	Energy	Entropy	Homogeneity	Performance	Contrast	Correlation
Dental caries under restoration	2.6047	0.9999	0.81	0.4798	0.9964	0.4044	0.0071	0.9614
Mesioangular impaction	5.3254	0.9995	0.9372	0.2039	0.9992	0.6339	0.0016	0.9743
Proximal caries	6.4644	0.9993	0.9555	0.1544	0.9996	0.4289	0.0085	0.9804
Vertical impaction	2.5354	0.9999	0.8025	0.4921	0.9963	0.2681	0.0074	0.9613
Dental caries involving distal pulp	2.1802	0.9999	0.7618	0.5617	0.995	0.2036	0.01	0.9563

The images are trained according to the dataset that is been stored.



Figure 6: Training of images

The set of dental images are run through the mentioned classification techniques such as SVM, ANN and KNN. The segmented and classified images using the different classifier algorithms are presented in Figure 7.

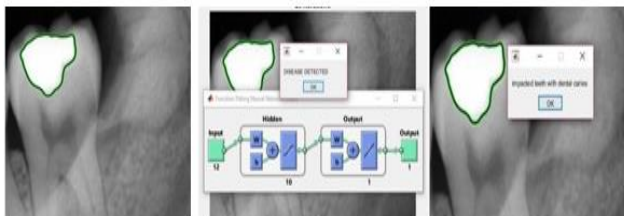


Figure 7: a) SVM Output b) ANN output c) KNN output

### V.CONCLUSION

In this work, it is suggested that by utilising GLCM features and SVM, KNN and ANN classifiers, the teeth affected by dental caries can be set apart from the normal teeth in a more detailed manner. The automated teeth segmentation and classification provide radiologists with a faster and second opinion by reviewing medical images, increasing the sensitivity of disease detection. Hence it is able to produce results in a more precise manner efficiently.

### REFERENCES

[1] Park WJ and Jun-Beom park, "History and application of artificial neural network in dentistry" 2018 dec: 12(4):594-601.  
 [2] Lira P, Giraldi G.A, Neves L, and Feijoo.R, "Dental X-Ray Image Segmentation Using Texture Recognition," Latin America Transactions, IEEE (Revista IEEE America Latina), Vol. 12, No. 4, 2014.

[3] Bethanney Janney J and S.Emalda Roslin, Classification of melanoma from dermoscopic data using machine learning techniques, Multimedia tools and applications, Springer US, Vol. 79, pp. 3713-3728. 2020  
 [4] Abdovahab Ehsani Rad , Mohd Shafry Mohd Rahim, RoselyKumoi and Alireza Norozi, "Dental X-Ray image segmentation and multiple feature extraction" 2012. Page no's 188-197  
 [5] Anuj Kumar; H. S. Bhadauria; Nitin Kumar, Fuzzy Clustering with Level Set Segmentation for Detection of Dental Restoration area, IEEE, June 2018.  
 [6] Jiafa Mao, Kaihui Wang, Yahong Hu, Weiguo Sheng, Qixin Feng, GrabCut algorithm for dental X-ray images based on full threshold segmentation, IET Image Process, Volume 12, Issue 12, December 2018, p. 2330 – 2335  
 [7] T.Sudhakar, Bethanney Janney.J, Haritha.D, Juliet Sahaya.M and Parvathy.V, Automatic Detection and Classification of Brain Tumor using Image Processing Techniques, Research journal of Pharmacy and Technology, 2017, 10(11), 1994-1998.  
 [8] Jie Yang; Yuchen Xie; Lin Liu; Bin Xia; Zhanqiang Cao; Chuanbin Guo, Automated Dental Image Analysis by Deep Learning on Small Dataset, IEEE Xploreer, June 2018  
 [9] E. H. Said, D E M Nassar, G. Fahmy, H. Ammar, Teeth segmentation In digitized dental X-ray films using mathematical morphology, IEEE Transactions on Information Forensics and Security November 2006  
 [10] Bethanney Janney J and Emalda Roslin S, Classification and Detection of Skin Cancer using Hybrid Texture Features, Biomedicine, 2017, 37(2), 214-220.  
 [11] Santhosh B. Review on emerging techniques on detects oral cancer. Int J Electr Sci Eng 2015; 1(1); 41-6.  
 [12] Rad A.E, Rahim .M, Shafry .M, Rehman.A, Altameem .A, and Saba .T, "Evaluation of Current Dental Radiographs Segmentation Approaches in Computer aided Applications", Institution of Electronics and Telecommunication Engineers Technical Review, (Med know Publications & Media Pvt. Ltd.) , Vol. 30, No. 3, PP. 210 – 222, 2013.  
 [13] Bethanney Janney J, S.Emalda Roslin, Analysis of Skin Cancer using K-Means Clustering and Hybrid Classification Model , Indian Journal of Public Health Research & Development July 2019, 10, 1371-1378.  
 [14] Maurya R, Surya K S, Maurya K A and Ajeet. "GLCM and Multi Class Support Vector Machine based Automated Skin Cancer Classification". IEEE 2014 International Conference on Computing for Sustainable Global Development 2014; 12: 444-447.  
 [15] María Prados-Privado , Javier García Villalón,, Carlos Hugo Martínez-Martínez and Carlos Ivorra , "Dental Images Recognition Technology and Applications:" A Literature Review, Applied Sciences, 2020, 10, 2856

# U-Net Supported Segmentation of Ischemic-Stroke-Lesion from Brain MRI Slices

Seifedine Kadry  
*Faculty of Applied Computing and  
 Technology, Noroff University College,  
 Kristiansand, Norway*  
 email: skadry@gmail.com

Robertas Damaševičius  
*Faculty of Applied Mathematics  
 Silesian University of Technology  
 44-100 Gliwice, Poland*  
 email: robertas.damasevicius@polsl.pl

David Taniar  
*Faculty of Information Technology  
 Monash University  
 Clayton, Victoria 3800, Australia*  
 email: David.Taniar@monash.edu

Venkatesan Rajinikanth  
*Department of Electronics and  
 Instrumentation Engineering  
 St. Joseph's College of Engineering  
 Chennai 600119, India*  
 email: v.rajinikanth@ieee.org

Isah A. Lawal  
*Faculty of Applied Computing and  
 Technology, Noroff University College,  
 Kristiansand, Norway*  
 email: Isah.Lawal@noroff.no

**Abstract**— The brain abnormality is one of the major sicknesses in human's health and the untreated brain defect will cause major illness. Ischemic stroke is one of the major medical emergencies and the timely diagnosis and treatment will save the patient from serious sickness. The proposed research employs the U-Net scheme to extort the Ischemic-Stoke-Lesion (ISL) from the brain MRI slices of ISLES2015 database. In this work, a pre-trained U-Net encoder-decoder system is employed to extort the ISL fragment from the chosen test image. After the extraction, a relative assessment is performed with the ground-truth available along with consequent test image. In this work, 20 patients' images (20 patient x 25 slices = 500 images) are adopted for the assessment and the general result achieved with the executed methodology helped to achieve a better value of Jaccard (>90%), Dice (>95%) and Accuracy (>98%) on the considered image dataset.

**Keywords**—Ischemic-stoke, Brain MRI, U-Net, decoder-encoder, assessment.

## I. INTRODUCTION

In human physiology, the brain plays a major role that monitors and controls every other part of the body. The abnormality in the brain will severely affect the sensory signal processing and the decision-making process. Ischemic-stoke is one of the common brain abnormalities in humans, largely occurs due to the lack of blood flow to the brain parts and timely recognition and treatment is very essential to help the patient to recover from the abnormality [1-3]. According to the level, the stroke impact varies from mild to severe and the unrecognised and untreated stroke may lead to temporary/permanent disability. In some cases, the untreated ischemic-stoke will lead to death [4,5].

Due to its importance, in the literature, a considerable amount of works are planned to examine the ischemic-stroke in humans using bio-signals [6] and bio-images [7,8]. The image supported diagnosis will help to get better information regarding the severity of the infection and its location compared to the bio-signal supported methods.

The image assisted detection of ischemic-stoke normally involves the examination of brain CT or MRI slices and due to its better visibility and multi-modality nature, brain MRI slices are widely adopted by the researchers compared to brain CT. The earlier works on Ischemic-Stoke-Lesion (ISL) assessment using the brain MRI slices can be found in [9-15].

The commonly adopted ISL assessment from the brain MRI involves in; (i) Collection of the clinical level brain MRI image, (ii) Pre-processing the image to get the appropriate dimension of MRI slice for the assessment, (iii) Implementing the appropriate imaging approach to extort and assess the ISL fragment and (iv) Validation and confirmation of the proposed scheme to confirm its accuracy.

The previous studies on the ISL assessment has implemented a variety of image processing methodologies on the clinically collected images and benchmark brain MRI slices. The clinical grade images are preserved by the researchers for their own use and most of these images are protected by copyright and not freely available for research purposes. Hence, the brain MRI slices collected from the benchmark database is extensively considered by most researchers to test and validate the developed scheme.

The ISLES2015 [11] is one of the commonly adopted brain MRI databases for the ISL related study and the earlier works implemented using this database can be found in the literature [7-12]. This dataset consists of the carefully recorded clinical-grade brain MRI with diverse modalities, such as Flair, T2, diffusion-weight (DW), and every MRI slice is associated with the essential Ground-Truth (GT) offered by two expert members. The visibility of the ISL is good in the Flair modality MRI slice compared to other approaches and hence, the Flair modality slices are one of the favourite choices of the researchers to test and validate their decrease detection system [8-10].

Recently, the Convolutional-Neural-Network (CNN) based Artificial-Intelligence (AI) practice helped to construct a number of the automated disease detection systems and the results of the earlier approaches also confirmed the merit of the CNN schemes compared to the tradition as well as other disease detection approaches. The merit of the CNN is its improved accuracy compared to other methods and its adaptability on the greyscale as well as RGB scaled images. Further, most of the CNN schemes are available as the pre-trained architecture and the existing structure can be easily implementable on any database with effortless training [16,17].

The proposed research aims to implement a U-Net scheme to extort the ISL area from the trial images with enhanced segmentation accuracy. The earlier work related to the U-Net scheme can be accessed from [18-20]. The pre-trained U-Net is allowed to extract the ISL fragment from the chosen brain MRI slice and after the extraction of the ISL, a

comparative measurement among the ISL section and the related GT is performed and the necessary Image-Quality-Measures (IQM) are computed. Based on the attained values of the IQM, the worth of the planned scheme is validated.

The other sections of this work is organised as below; Section 2 discuss the Context, Section 3 demonstrates the methodology, Section 4 and 5 give the experimental outcome of this research and the conclusion of the proposed work, respectively.

## II. CONTEXT

The earlier works are executed to examine the brain abnormality using Flair modality MRI slices. The earlier works implemented using the ISLES2015 database can be found in [7-12] and in most of the works, segmentation of ISL using a chosen semi-automated/automated image processing approach is implemented.

A number of procedures are proposed and applied to inspect the ISLES2015 datasets and the implemented approaches vary from; traditional methods to the recently developed deep-learning approaches. The ultimate aim of every method is to extort the ISL fragment from the MRI slice of chosen modality; which offers a better result.

In the literature, the MRI modalities, such as Flair, T1, and DW are separately evaluated with; (i) direct semi-automated and automated segmentation techniques, (ii) Combination of the multi-thresholding-based enhancement, and chosen segmentation method, and (iii) Deep-learning assisted segmentation and classification techniques. Every method will have its own merit and the chief aim in each scheme is to extract the ISL with the enhanced image performance measure values.

The traditional assessment and the combined thresholding and segmentation outcome can be found in the earlier literature [7,8]. The earlier works also confirmed that; most of the existing methods are semi-automated procedures and may require partial or full human supporters to execute the evaluation. In this work, an attempt is made to implement the U-Net based segmentation approach to extort the infected segment of the brain MRI with improved accuracy. The main advantage of this approach compared to the existing approach is its availability and its implementation issue. The U-Net considered in this work is a pre-trained system and the implementation and validation of the pre-trained scheme is quite simple compared with the customary methods.

## III. METHODOLOGY

This section shows the methodology adopted in this research work. After collecting the essential test image from the ISLES2015 database, every image is resized to the dimension of 256x256x3 pixels. Initially, this work considered the pre-trained U-Net model to extract the ISL segment from the considered test images. The total number of test images considered is only 500 images and hence image augmentation (horizontal/vertical flip,  $\pm 30^\circ$  rotation and  $\pm 60^\circ$  rotation) is implemented to increase the number of test images to train the U-Net system for the ISLES2015 images. When the U-Net is perfectly trained, then the performance of the U-Net is validated for all the 500 test images. The role of the U-Net is to extract the ISL section

from the test image with better accuracy. After getting the ISL fragment, a relative assessment with the GT is performed and the essential performance measures are then computed. Based on these values, the significance of the U-Net segmentation is validated. Figure 1 depicts the various sections existing in the proposed approach.

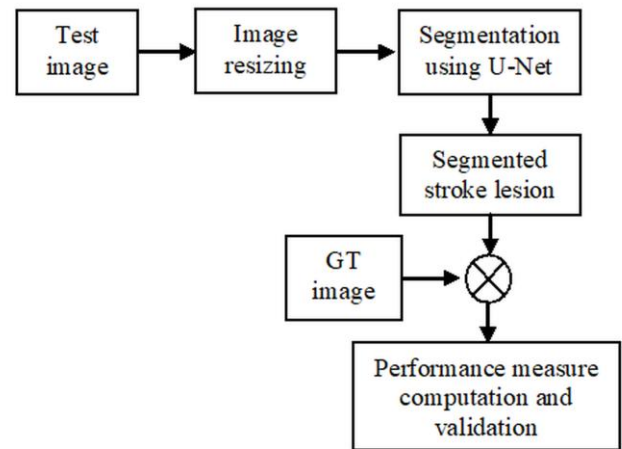


Fig.1. Proposed ISL examination scheme

### A. Image database

The development of an appropriate image examination system is essential for the efficient assessment of the brain abnormality. The proposed work considered the test images from the Ischemic-Stroke-LEsion-Segmentation Challenge 2015 (ISLES2015) [11]. This dataset consists of the clinical-grade, skull-stripped 3D brain MRI of Flair, T2, and DW modality images associated along with the necessary GT. Assessment of 3D brain MRI is quite complex and requires a complex computation; hence a 3D to 2D alteration is realized using the ITK-Snap tool [21,22] and the transformed images are then considered for the evaluation.

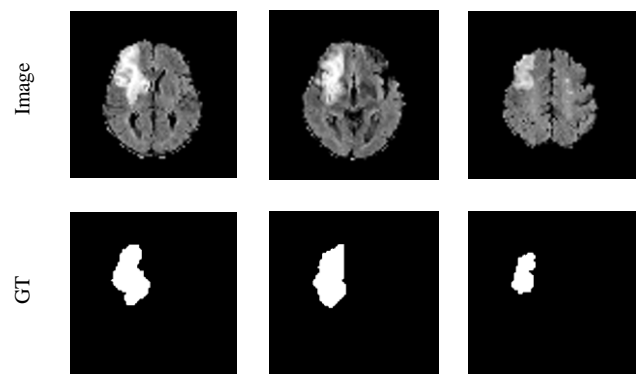


Fig.2. Sample test images of ISLES2015

All the extracted images are available with a dimension of 77x77x3 pixels and every test image is resized to 256x256x3 pixels. In this research, 20 patients' images are considered for the assessment, and from every patient, 25 slices are extracted and included in the test image database. All the 500 images (20 patients' x 25 slices) are separately processed using the U-Net scheme implemented in this study. The sample trial images and the GT is depicted in Figure 2. Other essential information and the earlier works on this dataset can be found in [12-14].

### B. U-Net Scheme

In the proposed work the well-known pre-trained U-Net scheme (VGG-U-Net) shown in Figure 3 is considered to segment the ISL region from the test images.

This scheme consists of two sections namely the encoder and the decoder using as presented in the figure. The encoder section is implemented by considering the VGG11 scheme with the essential convolutional (Conv) and MaxPool layers as presented in the figure and the decoder section is associated with the Up-Conv section which will reconstruct the images from the extracted features of the encoder unit. The final layer of the decoder section consists of the SoftMax layer, which will help to implement a binary classification to separate the ISL section from the background. The proposed scheme helps to get a binary image, which is then compared against the related ground truth for validation. Other essential information regarding the U-Net (VGG-U-Net) can be found in the literature [18-20].

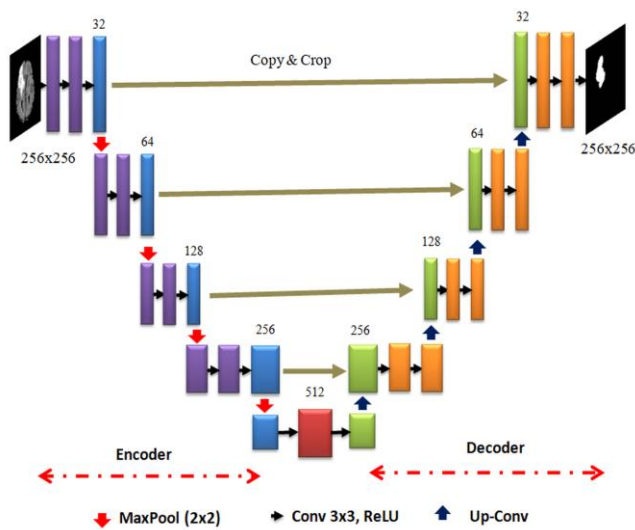


Fig.3. Structure of the VGG-U-Net (U-Net) scheme considered in this research

### C. Validation

The success of every disease detection system depends on its performance on the considered benchmark image database. When better performance is achieved with the system on the benchmark datasets, then it can be assumed that the developed scheme also will provide a better result when the real clinical images are evaluated.

The essential Performance-Measures (PM) such as True-Positive (TP), False-Positive (FP), False-Negative (FN), and True-Negative (TN) are initially computed during the comparison between the segmented ISL and the GT. From these initial measures, other values, such as Jaccard, Dice, Accuracy (ACC), Precision (PRE), Sensitivity (SEN), Specificity (SPE), and Negative-Predictive-Value (NPV) are also attained, and based on these values; the performance of the proposed segmentation approach is validated [7-10].

## IV. EXPERIMENTAL RESULT AND DISCUSSION

This part of research reveals the attained experimental results and its discussions. All the experimental work is executed using a workstation; Intel i5 2.5GHz processor with 16GB RAM and 2GB VRAM set with MATLAB®.

Initially, the considered U-Net scheme is trained using the image database (original and augmented images) and after the complete training process; every test image (500 images) is separately tested and the attained result is compared with its GT image. The comparison among segmented ISL and GT will help to get the necessary PMs and based on its value, the merit of U-Net on the ISLES2015 dataset is confirmed.

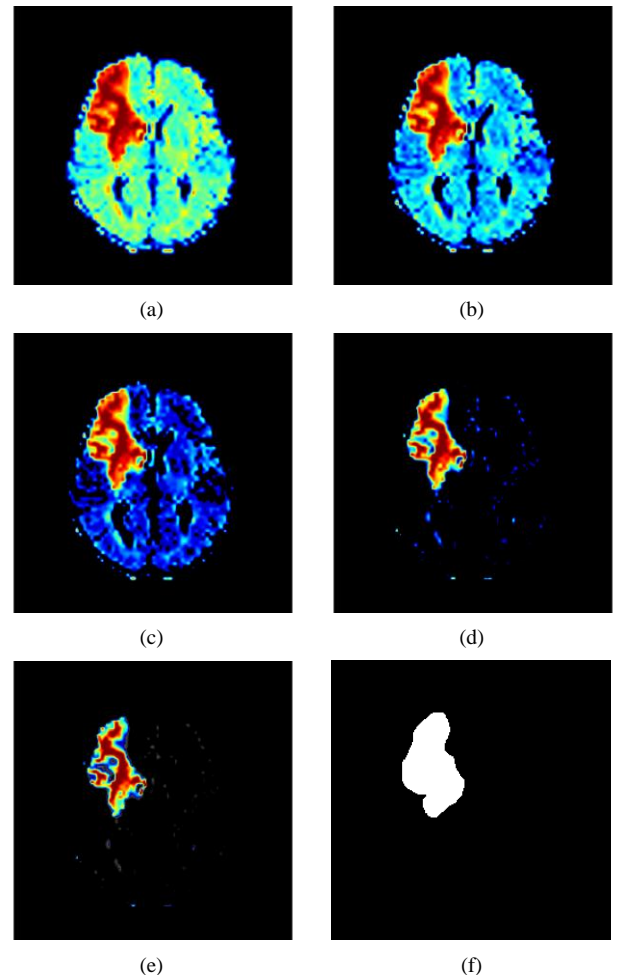


Fig.4. Segmentation outcome attained with U-Net for a sample test image.

(a) – (c) Images extracted from various layers of the encoder section, (d) and (e) reconstructed image by the decoder unit, (f) Segmented ISL by the SoftMax

The pre-trained U-Net scheme is successfully executed to extort the essential ISL fragment from the chosen test images and the result attained for a sample test image can be found in Figure 4. In this figure, both the encoder as well as the decoder outcomes is depicted. Even though the brain MRI slices look like the greyscale images, actually, these images are in the form of RGB and this information is clearly discussed in [7,8]. When the initial image is fed to the encoder part, the essential convolution, as well as the MaxPool operation, is performed to extort the indispensable features from the image. Normally, the encoder section is the pre-trained VGG11 architecture without the fully connected layer. The outcome of the encoder section is in the form of image features. During the encoding operation, every down-convolution layer output is passed towards the up-

convolution layer in which the pixel level segregation image reconstruction is taking place.

The total number of up-convolution operation in the decoder is similar to the encoder unit and this supports the reconstruction of the features into images with varied pixel groups. The final part of this section has a SoftMax classifier unit, which helps to segregate the existing images into the ISL section (binary 1) and the background (binary 0). In this figure, Fig. 4(a) to (c) depicts the results of the encoder section, Fig 4(d), (e) are the results of the decoder sections and Fig 4(f) is the outcome by the SoftMax classifier. Normally, in the medical image segmentation task, Fig 4(f) will be the binary image with the ISL and background and finally, this binary picture is evaluated and confirmed against the GT image.

The comparison of the ISL and the GT will help to get the essential PM values and according to these measures, the merit of U-Net approach is confirmed. In this work, the results attained in the sample test images (depicted in Fig 2) is presented and IM1 to IM3 respectively denotes the Image 1 to Image 3 of brain MRI slices existing in Fig 2. From Table I and Table II values, it can be noted that the U-Net segmentation helps to realize superior values of PMs for every image with healthier values of Jaccard, Dice, and ACC. A similar procedure is repeated for all other images of the considered ISLES2015 database and the average of performance measure is accounted to verify the performance of the proposed CNN supported segmentation.

TABLE I. ESSENTIAL PERFORMANCE MEASURE COMPUTED BETWEEN SEGMENTED SECTION AND GT

Image	TP	FP	TN	FN	Jaccard (%)	Dice (%)
IM1	3254	107	62033	142	92.89	96.31
IM2	3092	163	62196	85	92.57	96.14
IM3	1485	114	63829	108	86.99	93.04

TABLE II. IMAGE PERFORMANCE MEASURES OF SAMPLE TEST IMAGES

Image	ACC (%)	PRE (%)	SEN (%)	SPE (%)	NPV (%)
IM1	99.62	96.82	95.82	99.83	99.77
IM2	99.62	94.99	97.32	99.74	99.86
IM3	99.66	92.87	93.22	99.82	99.83

The overall outcome (average of all 500 images PMs) of the proposed scheme is confirmed alongside the available results in the literature as represented in Table III. This information confirms that the results attained with the proposed scheme are better and it works well on the brain MRI slices of the Flair modality. Further, the other existing methods are requiring a combined traditional thresholding and segmentation methods and most of the existing methods are also are of semi-automated approaches. The U-Net scheme supported method is an automated approach and extracts the ISL fragment with improved PM values compared to the existing works.

This scheme is experienced and authenticated by means of the Flair modality MRI and in the future, the performance of U-Net segmentation can be tested with other brain MRI

modalities. Further, the merit of U-Net can be compared and confirmed against other pre-trained CNN segmentation methods available in the literature.

TABLE III EVALUATION OF PROPOSED APPROACH WITH EXISTING METHODS

Method	Jaccard (%)	Dice (%)	ACC (%)
Rajinikanth and Satapathy [7]	78.60	88.54	91.17
Rajinikanth et al. [8]	90.34	94.92	-
Lin et al. [9]	Otsu	88.48	90.48
	Kapur	90.36	92.21
Rajinikanth et al. [10]	Watershed	90.34	94.92
	Active Contour	88.05	93.63
	Markov random field	86.12	92.46
Proposed U-Net segmentation	<b>90.74</b>	<b>95.18</b>	<b>98.68</b>

## V. CONCLUSION

Assessment of a chosen brain MRI slice is one of the commonly adopted real-time investigation approaches. Evaluation of the ISL section with an appropriate image processing scheme is very essential during the brain abnormality assessment. The extracted ISL will help to locate the brain section where the problem existing and the severity of the problem. This work implemented a pre-trained U-Net scheme to extract the ISL with superior accuracy. The outcome of the extracted section is then compared with the GT image and the disease detection performance of proposed scheme is confirmed based on the computed values of the PMs. In this work, 500 numbers of brain MRI slices of ISLES2015 is examined and the average values of the PMs confirmed that the proposed scheme offers better result compared to the results of other existing methods.

## REFERENCES

- [1] M. A. Khan, I. Ashraf, M. Alhaisoni et al., "Multimodal brain tumor classification using deep learning and robust feature selection: A machine learning application for radiologists," *Diagnostics*, vol. 10, no. 8, pp.565, 2020. <https://doi.org/10.3390/diagnostics10080565>.
- [2] Q. Ke, J. Zhang, W. Wei, R. Damaševičius, and M. Woźniak, "Adaptive independent subspace analysis of brain magnetic resonance imaging data," *IEEE Access*, vol. 7, pp.12252-12261, 2019. DOI: 10.1109/ACCESS.2019.2893496.
- [3] S.L. Fernandes, U. J. Tanik, V. Rajinikanth, and K. A. Karthik, "A reliable framework for accurate brain image examination and treatment planning based on early diagnosis support for clinicians," *Neural Computing and Applications*, vol. 32, no. 20, pp. 15897-15908, 2020. <https://doi.org/10.1007/s00521-019-04369-5>.
- [4] S. Kadry, V. Rajinikanth, N.S.M. Raja, D. J. Hemanth, N. M. S. Hannon, and A. N. J. Raj, "Evaluation of brain tumor using brain MRI with modified-moth-flame algorithm and Kapur's thresholding: a study," *Evolutionary Intelligence*, pp. 1-11, 2021. <https://doi.org/10.1007/s12065-020-00539-w>.
- [5] K. Revanth, N.S.M. Raja, and V. Rajinikanth, "Computational investigation of stroke lesion segmentation from Flair/DW modality MRI," In 2018 Fourth International Conference on Biosignals, Images and Instrumentation (ICBSII), pp. 206-212. IEEE, 2018. DOI: 10.1109/ICBSII.2018.8524617.
- [6] N. Krishna, M. K. Sekaran, A. V. N. Vamsi et al., "An efficient mixture model approach in brain-machine interface systems for extracting the psychological status of mentally impaired persons using EEG signals," *IEEE Access*, vol. 7, pp. 77905-77914, 2019. DOI: 10.1109/ACCESS.2019.2922047.

- [7] V. Rajinikanth, and S. C. Satapathy, "Segmentation of ischemic stroke lesion in brain MRI based on social group optimization and Fuzzy-Tsallis entropy," *Arabian Journal for Science and Engineering*, vol. 43, no. 8, pp. 4365-4378, 2018. <https://doi.org/10.1007/s13369-017-3053-6>.
- [8] V. Rajinikanth, K. P. Thanaraj, S. C. Satapathy, S. L. Fernandes, and N. Dey, "Shannon's entropy and watershed algorithm based technique to inspect ischemic stroke wound," *Smart intelligent computing and applications*, vol.105, pp. 23-31. 2019. [https://doi.org/10.1007/978-981-13-1927-3\\_3](https://doi.org/10.1007/978-981-13-1927-3_3).
- [9] D. Lin, V. Rajinikanth, and H. Lin., "Hybrid Image Processing-Based Examination of 2D Brain MRI Slices to Detect Brain Tumor/Stroke Section: A Study," In *Signal and Image Processing Techniques for the Development of Intelligent Healthcare Systems*, pp. 29-49. Springer, Singapore, 2021. [https://doi.org/10.1007/978-981-15-6141-2\\_2](https://doi.org/10.1007/978-981-15-6141-2_2).
- [10] V. Rajinikanth, S. C. Satapathy, N. Dey, and H. Lin, "Evaluation of ischemic stroke region from CT/MR images using hybrid image processing techniques," In *Intelligent multidimensional data and image processing*, pp. 194-219. IGI Global, 2018. DOI: 10.4018/978-1-5225-5246-8.ch007.
- [11] O. Maier, B.H. Menze, J. V. D. Gablentz et al., "ISLES 2015-A public evaluation benchmark for ischemic stroke lesion segmentation from multispectral MRI," *Medical image analysis*, vol 35, pp.250-269, 2017. <https://doi.org/10.1016/j.media.2016.07.009>.
- [12] O. Maier, M. Wilms, and H. Handels, "Image features for brain lesion segmentation using random forests," In *BrainLes 2015*, pp. 119-130. Springer, Cham, 2015. [https://doi.org/10.1007/978-3-319-30858-6\\_11](https://doi.org/10.1007/978-3-319-30858-6_11).
- [13] A. Carass, S. Roy, A. Gherman et al., "Evaluating white matter lesion segmentations with refined Sørensen-Dice analysis." *Scientific reports*, vol. 10, no. 1, pp.1-19, 2020. <https://doi.org/10.1038/s41598-020-64803-w>.
- [14] S. Winzeck, A. Hakim, R. McKinley et al., "ISLES 2016 and 2017-benchmarking ischemic stroke lesion outcome prediction based on multispectral MRI," *Frontiers in neurology*, vol. 9, pp.679, 2018. DOI: 10.3389/fneur.2018.00679.
- [15] O. Maier and H. Handels, "Predicting stroke lesion and clinical outcome with random forests," In *International Workshop on Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, *Lecture Notes in Computer Science*, vol. 10154, pp. 219-230. 2016. [https://doi.org/10.1007/978-3-319-55524-9\\_21](https://doi.org/10.1007/978-3-319-55524-9_21).
- [16] N. Dey, Yu-Dong Zhang, V. Rajinikanth, R. Pugalenthi, and N.S.M. Raja, "Customized VGG19 architecture for pneumonia detection in chest X-rays," *Pattern Recognition Letters*, vol.143, pp.67-74, 2021. <https://doi.org/10.1016/j.patrec.2020.12.010>.
- [17] V. Rajinikanth, A.N.J. Raj, K.P. Thanaraj, and G.R. Nai., "A customized VGG19 network with concatenation of deep and handcrafted features for brain tumor detection," *Applied Sciences* 10, no. 10, pp.3429, 2020. <https://doi.org/10.3390/app10103429>.
- [18] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," *Lecture Notes in Computer Science*, vol. 9351, pp. 234-241, 2015. [https://doi.org/10.1007/978-3-319-24574-4\\_28](https://doi.org/10.1007/978-3-319-24574-4_28).
- [19] V. Iglovikov and A. Shvets, "Ternausnet: U-net with vgg11 encoder pre-trained on imagenet for image segmentation," *arXiv preprint arXiv:1801.05746*, 2018.
- [20] T. Falk, D. Mai, R. Bensch et al., "U-Net: deep learning for cell counting, detection, and morphometry," *Nature methods*, vol. 16, no. 1, pp.67-70, 2019. <https://doi.org/10.1038/s41592-018-0261-2>.
- [21] P.A. Yushkevich, J. Piven, H. C. Hazlett, R. G. Smith, S. Ho, J. C. Gee, and G. Gerig, "User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability," *Neuroimage*, vol.31, no.3, pp. 1116-1128, 2006. <https://doi.org/10.1016/j.neuroimage.2006.01.015>.
- [22] <http://www.itksnap.org/pmwiki/pmwiki.php?n=Main.Publications>



# DROWSINESS DETECTION SYSTEM USING DEEP LEARNING

Avigyan Sinha<sup>1</sup>, Aneesh R P<sup>2</sup>, Saradha K Gopal<sup>3</sup>

<sup>1</sup> Johns Hopkins University, USA, <sup>2</sup> Regional Centre IHRD, Thiruvananthapuram, India, <sup>3</sup> AJ College of Science and Technology, Thiruvananthapuram,

**ABSTRACT:** Drivers drowsiness is the major problem that causes road accidents. Unlike normal facial expression, drowsiness is defined to be a condition of exhaustion, where the expression of the face is different from usual. The important steps in detecting drowsiness are face detection and expression detection. Many algorithms are being developed to detect face and expressions. But these algorithms give poor performance due to the extrinsic parameters of the environment. Light and position of the camera are the major problems. In this paper, different architectures were used to analyse the performance of face and drowsiness detection. Also we have proposed new detection methods using deep learning technique. To estimate the drivers' state we use facial regions corresponding to entire face. The algorithms employed for face detection are i) Viola Jones ii) DLib iii) Yolo V3. For the Classification, The CNN (Convolutional Neural Network) architectures employed in the drowsiness detection is modified LeNet.

**Keywords:** Convolutional Neural Network (CNN), Drowsiness detection, Face Detection, Viola Jones algorithm, dlib, YoloV3, LeNet, DarkNet-53, Deep Learning, multi-modal analysis

## I. INTRODUCTION

Automobiles have become an essential mode of transportation for people. Number of globally sold vehicles over the last four years was 95 million. After one year 0.3% more vehicles were sold. In 2018 the selling rate was 1 billion and increased by 0.6% in 2019 compared to last year. In 2020 the global selling rate until March is 60.5 million automobiles. With increase in the number of vehicles, the occurrence of traffic and accidents is also gradually. Traffic collisions are a major source of deaths every year. The National Crime Record Bureau (NCRB) reports 496,762 road-related traffic collisions in all states [1]. According to an estimate given by the World Health Organization (WHO), around 1.35 million people worldwide have died due to road traffic injuries [2]. These accidents occur due to fatigue driving approximately 20% - 30%. Drowsy driving is the most dangerous aspect of road accidents.

The methods for drowsiness detection are classified into subjective and objective detection methods. In the objective detection no feedback is given to drivers and detection takes place according to the drivers' physical aspects but subjective detection is all about physical aspects of drivers. The objective type of detection is further grouped into contact and non-contact types. The

proposed system is based on non-contact method, since it is low cost as compared to contact method.

The main goal of this work is to determine whether a driver is drowsy. The algorithmic pipeline analyses each frame image of the video stream and detects the drivers' condition – whether feeling drowsy or not. The proposed system is based on non-contact method, since it is low cost as compared to contact method. After detection, next step is giving an alert alarm to the driver so that he can take necessary action. Here deep learning technique is used with the help of Convolutional Neural Network (CNN).

## II. LITERATURE SURVEY

In various approaches steps are taken for facial landmark detection, object tracking and methods for driver drowsiness detection can be used in two ways either in contact approaches or in non-contact approaches. The nature of the techniques largely depends on the application domain.

Many researchers have attempted various technologies such as monitoring of underlying patterns in steering, monitoring vehicle position in lane, monitoring the eye/face of the driver, physiological measurement etc. It is seen that most approaches use driver eye/face monitoring, physiological measurement. The research done in 2019 uses driver eye/face monitoring which creates a "DriCare" alert for the driver [3]. The system is efficient and its efficiency can be validated using a public driver drowsiness recognition dataset. Some of the research works which made the evolution to proposed work are depicted as follows.

### A. Facial Landmarks Recognition

Facial landmarks are an important, yet challenging phase in drowsiness detection. It has been applied to solve problems like alignment of face, estimation of head pose, swapping of face, blink detection etc. It usually used to pinpoint and characterize significant areas of the face, eyebrows, nose, and mouth.

The research work by the authors Y. Sun, X. Wang and X. Tang propose a new methodology for approximation of the locations of facial key points with convolutional networks having three carefully designed levels [4]. There are two benefits: first, the texture context data over the whole face is applied to find every key point. Second, for the reason that networks are skilled to predict all of the key points simultaneously, the geometric constraints amongst key points are implicitly encoded.

The approach consequently can keep away from local minimum arising due to ambiguity and information corruption in difficult image samples because of occlusions, large pose variations and extreme lightings. Various network architectures are critical for precise and robust facial point detection. Results have been shown in three levels of convolutional networks: initial detection, tuned outcomes with the second and third levels of networks which improve the accuracy. Benefits are high performance convolutional networks. It also improves accuracy of modern-day methods and latest industrial software and additionally incredibly sturdy preliminary estimations. Drawback is that regionally sharing weights of neurons on the same map develops the performance but does not work well on images such as faces when globally sharing weights, and also is not precise under some exceptional poses and expressions.

### B. Tracking of Visual Object

Tracking of visual objects refers to tracing the trajectory of a single object across all frames of a video such that we are given its location only in the first frame. It is one of the major tasks in Computer Vision.

Research work by D. S. Bolme, J. R. Beveridge, B. A. Draper and Y. M. Lui, investigated a simpler tracking strategy [5]. A novel correlation filter, a Minimum Output Sum of Squared Error (MOSSE) filter, which creates robust correlation filters when initialized using a single frame, is introduced. By using modified ASEF, UMACF or MOSSE filters good performance can be obtained under variations in rotation, scale, lighting and partial occlusion. The strength of a correlation peak can be used to detect occlusions or tracking failure, to avoid the online update and to reacquire the trace if the item reappears with a similar form by using "Peak-to-Side" lobe ratio. The subsequent algorithm can be just as precise and is much quicker. The advanced correlation filters are as good as complex trackers and the filter-based tactic is more than 20 times faster and can process 669 frames in a second. The end result is that they are not particularly sturdy to deviations in appearance of the target and fail on perplexing tracking problems.

Research by M. Danelljan, G. Häger, F. S. Khan and M. Felsberg, explore the problem of exact and robust scale approximation for visual tracking in real-time [6]. They have adopted the Discriminative Scale Space Tracker (DSST), which learns distinct correlation filters for unambiguous translation and scale assessment. Strategies are proposed to diminish the computational cost of tracking methods. Thus we get a larger target search space without compromising real time performance. The tracking performance is significantly improved and speed increases two fold.

### C. Detection of Driver's Drowsiness

There are two ways of detecting drowsiness; they are contact methods and non-contact methods. In contact approaches, the driver needs to wear or touch some physical parameter to measure level of drowsiness and on other hand in non-contact approaches, the driver does not need contact physical objects. Non-contact methods have lower cost and are easier.

Research by S.-J.Jung, H.-S.Shin and W.-Y. Chung [7], explores ways to monitor a driver's health state by

using embedded electrocardiogram (ECG) sensors with electrically conductive fabric electrodes on the steering wheel. The ECG signals are sampled at 100Hz from palms and transferred wirelessly to a remote station linked to a server computer in a private area network, to check driver's health condition, fatigue, and drowsiness state from physiological variations in biomedical signals. However, the pipeline was followed by a contact approach which may not be user-friendly and also time consuming.

Research by G. Li, B.-L.Lee and W.-Y. Chung [8], is based on physiological signals among them directly related to drowsiness is brain reflected by Electroencephalographic (EEG) signals. A SVM-based posterior probabilistic model and a smart watch-based wearable EEG device were used. Drowsiness score is measured by probability value of 0 ~ 1 rather than discrete class labels. But the EEG channel setting and the ground truth used are vague and the detection accuracy is not stated properly.

Research by B. Warwick, N. Symons, X. Chen and K. Xiong [9], proposed a wireless wearable sensor used to build Driver Drowsiness systems. To design a drowsiness system there are two phases: collection of physiological data using the biosensor and analysis of measured data to identify the crucial factors related to the drowsiness. In second stage, there is designing of drowsiness detection algorithm and mobile app to alert drowsy people. It is found that a driver's heart rate and breathing rate are good indicators of drowsiness.

Research by M. Omidyeganeh, A. Javadtalab and S. Shirmohammadi [10], captures facial appearance via a camera. The scheme contains four stages: Face extraction from image, eye detection, mouth detection, alert generation in drowsy state. Limitation of the system is that since non-contact approach is used here it depends on factors like light, camera and so on.

The main aims of this paper are to overcome all drawbacks that are mentioned above.

## III. METHODOLOGY

Drowsiness detection system, the entire procedure is carried down according to image processing which is a method to perform some operations on an image. Figure 1 shows the proposed Drowsiness Detection System.

The system flow has five steps which essentially comprise:

- i) Video Acquisition
- ii) Detection of Face
- iii) Detection of Eye and Mouth (alternate approach)
- iv) Assessment of State
- v) Categorization into drowsiness or non-drowsiness state

Here our pipeline has these basic steps: at first video is captured through NIR camera then frames are converted, secondly face is detected, alternatively, eye and mouth detection, next there is state estimation and finally drowsiness detection. The proposed system is based on non-contact method, since it is low cost as compared to

contact scheme. The paper uses deep learning technique for feature expression, keeping away the manual feature extraction.

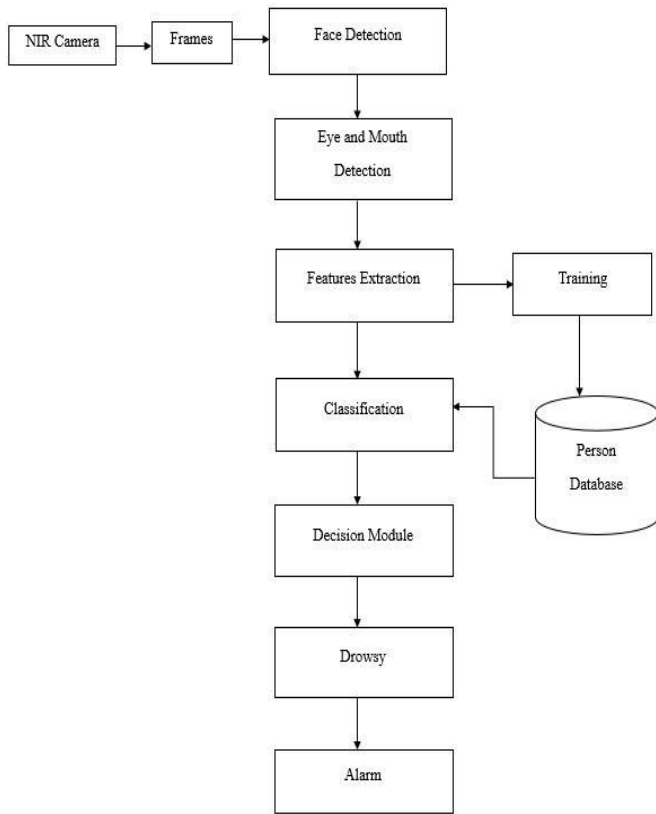


Figure 1 : Proposed Drowsiness Detection System

#### A. Video Capturing

The ULg multimodality drowsiness database (called DROZY) is used here [11]. This is an open dataset intended to support research works. The images are collected using the NIR camera at the university of Liege, Belgium. This dataset contains videos of 14 different persons. Psychomotor vigilance tests (PVTs) were performed to annotate the data. Each image has a resolution of 512 x 424 pixels and gray scale values of 8 bits. Ground truth values of face positions (manual and automatic) were also attached with this dataset. This data contains different subjects, but we use only drowsy and non drowsy conditions. As NIR cameras are used in poor-light conditions, Our proposed work is to classify drowsiness condition from the low light images.

#### B. Face, Eye and Mouth Detection

Viola Jones algorithm is commonly used for face detection, because of its simplicity. Haar features are extracted here to detect landmarks like face, eye, nose, mouth[12]. This method is similar to convolution operation which can detect a feature in a given image. The sum of pixels that fall under white rectangle is subtracted from the sum of pixels within black rectangle to get results of each feature in a single example as mentioned in Eq. (1):

$$\Delta = \frac{1}{n} \sum_{\text{black}}^n 1(x) - \frac{1}{n} \sum_{\text{white}}^n 1(x) \quad (1)$$

Ideal Haar feature pixel intensities are 0 for white and 1 for black. There are 3 types of features that Viola-Jones

identified: features corresponding to edge, line and four-rectangle as mentioned in Fig.3.

Given the Haar features for the face, a method such as integral image as in Eq. (2), adaboost training as in Eq. (3) and finally cascade classifiers can be used. In an integral image the intensity at pixel (x, y) is given by sum of pixels above and left of (x, y).

$$\sum i(x,y) = I(A) - I(B) - I(C) + I(D) \quad (2)$$

In Eq. 2, where  $\sum i(x, y)$  is area to be found and  $I(A)$ ,  $I(B)$ ,  $I(C)$ ,  $I(D)$  are the integral values of each rectangle as shown in Fig. 4.

Adaboost is an algorithm that isolates the best features amongst all these 162,336 features. A weighted combination of all these features determines whether a given window contains a face or not. It builds a strong classifier by using linear combination of weak classifiers.

$$F(x) = \alpha_1 f_1(x) + \alpha_2 f_2(x) + \alpha_3 f_3(x) + \dots \text{(to n terms)} \quad (3)$$



Figure 2 : Frames from the dataset

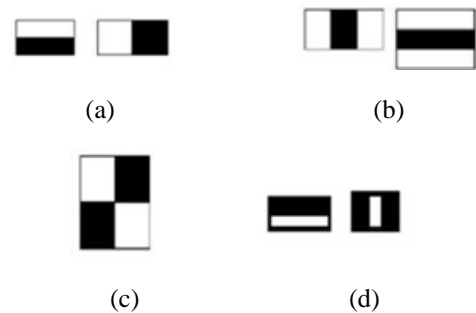


Figure 3 : Haar features (a) Edge features (b) line features (c) four rectangle features (d) face detection features

Cascaded classifiers comprise stages each containing a strong classifier and job of each stage is to determine whether input contains face or no face, But it fails when we have side faces or looking up or down.

Another alternative is to use "Dlib" library for face detection. Dlib is basically an open source platform for machine learning applications. Strength of this library is that this was developed with linear algebra and machine learning tools [13]. This library performs face and landmarks detection based on HoG features and support vector machines. Kernelized correlation filters (KCF) can be incorporated with the face detection for tracking in real time videos. But the initial position of the face should be marked to track the faces in successive images. This KCF algorithm can be modified with Multiple Convolutional

Neural Networks (MC-KCF) to avoid the drawbacks of KCF algorithm. The MC-KCF algorithm use FHOG features from the faces for the initialisation [3].

$$G_x(x, y) = H(x + 1, y) - H(x - 1, y) \quad (4)$$

$$G_y(x, y) = H(x, y + 1) - H(x, y - 1) \quad (5)$$

$$G = \sqrt{G_x^2 + G_y^2} \quad (6)$$

$$\alpha = \arctan\left(\frac{G_x}{G_y}\right) \alpha \in (0^0, 360^0) \quad (7)$$

$H(x, y)$  – Pixel values at  $(x, y)$

$G_x(x, y)$  – Horizontal gradient values at  $(x, y)$

$G_y(x, y)$  – Vertical gradient values at  $(x, y)$

TABLE 1 : DARKNET-53 ARCHITECTURE

Type	Filters	Size	Output
Convolutional	32	3 × 3	256 × 256
Convolutional	64	3 × 3 / 2	128 × 128
1x	Convolutional	32	1 × 1
	Convolutional	64	3 × 3
Residual			128 × 128
Convolutional	128	3 × 3 / 2	64 × 64
2x	Convolutional	64	1 × 1
	Convolutional	128	3 × 3
Residual			64 × 64
Convolutional	256	3 × 3 / 2	32 × 32
8x	Convolutional	128	1 × 1
	Convolutional	256	3 × 3
Residual			32 × 32
8x	Convolutional	512	3 × 3 / 2
	Convolutional	256	1 × 1
8x	Convolutional	512	3 × 3
	Residual		
4x	Convolutional	1024	3 × 3 / 2
	Convolutional	512	1 × 1
4x	Convolutional	1024	3 × 3
	Residual		
Avgpool		Global	
Connected		1000	
Softmax			

Yolo is a deep learning algorithm for pattern matching. As it represents “You Only Look Once” (YOLO), this algorithm efficiently works in object detection as one stage detector [14]. In our work, Yolov3 has been also used for face detection as its speed is more than 1000 times that of R-CNN and 100 times that of Fast R-CNN. In Yolov3, a novel architecture is used for feature extraction. It’s a hybrid between that used in YOLOv2, Darknet-19, and residual architectures. There are successive 3 × 3 and 1 × 1 convolutional kernels along with shortcut connections. It uses Darknet-53 architecture which has 53 convolutional layers. It has more efficiency than ResNet-101 or ResNet-152, making better utilization of GPU as it does highest measured floating point operations in a second [15]. The architecture details is shown in Table 1.

### C. Drowsiness Detection

Drowsiness detection is the classification process in the system. Many machine learning approaches are already developed for classification. But these approaches will not satisfy the goal. As the dataset collected through NIR camera in the vehicle environment, Image has only a low light illumination. Deep learning approaches will give good result. Many Pre trained deep networks are available for drowsiness detection [16]. Two types of classification

methods can be used. i) Analyse eye and mouth region of interests to determine whether the eyes or mouth are opened or closed. ii) Analyse the entire region of interest of the face. Drowsiness can be classified according to the emotion of the person [17]. Here we create a new training algorithm with modified LeNet for two classes. LeNet is the earliest architecture in Convolutional Neural Network[18]. In this work, modified approach of LeNet is used because of its simplicity and performance. The modification is included to improve the efficiency of system by learning the raw face images. CNNs are essential tools for deep learning and are especially suited for analysing image data. Architecture of CNN is mentioned

a). Convolutional Layer :The Convolutional layer is the fundamental building block of a Convolutional Network.

Local Connectivity - It is unrealistic to join neurons of one layer to all neurons in the previous volume, as it unnecessarily increases the number of parameters and computational cost. Rather, every neuron is connected to only a local region of the input volume. The spatial range of this connectivity is a hyper parameter known as receptive field or filter size. The connectivity range along the depth dimension is same as depth of the input. Thus connections are regional in space (along width and height), but to full extent along depth.

Spatial arrangement – The hyper parameters controlling the size of output are depth, stride and zero-padding. Depth is the number of filters to be used. Stride is the amount by which filter is slid. For stride of 1, filters are moved one pixel at a time. For stride of 2, filters move 2 pixels at a time. Thus we get smaller output volumes spatially. Zero padding determines the amount of zeros with which we pad the input around its border. This also gives us control over the size of the output spatially. The size of the output is given by  $(2P+W-F)/S+1$ , where P = zero padding amount, W = input size, F = receptive field size of neurons, S = stride.

Parameter Sharing – While back propagation, each neuron computes the gradient corresponding to its weights. But these gradients get added along each depth slice and only single set of weights are updated per slice.

Back propagation - The backward pass is also a convolution (with spatially-flipped filters)for a convolution operation (for both data and weights).

b) Pooling Layer : This layer gradually diminishes the spatial size of representation along with number of parameters and computational cost in the network. It controls over fitting. The layer operates individually on each depth slice of input and resizes it by MAX operation. Most commonly it has filters of size 2x2, stride of 2 and down samples each depth slice by 2 along the width and height. It adds no parameters since it calculates a permanent function of the input. Besides max pooling, this unit can also compute average pooling or L2-norm pooling.

c) Non-Linearity Layer : This comprises a non-linear activation function which maps the feature map created by previous convolutional layer to an activation map. It is an element-wise operation on the input such that dimensions of input and output are matching. Examples of such

functions are sigmoid (logistic), hyperbolic tangent, rectified linear units (ReLUs) etc.

d) Fully-connected layer :These layers have complete connections to all activations in the previous layer, as in usual Neural Networks like Multi-Layer Perceptron. Their activations can be computed with a matrix multiplication followed by applying a bias offset.

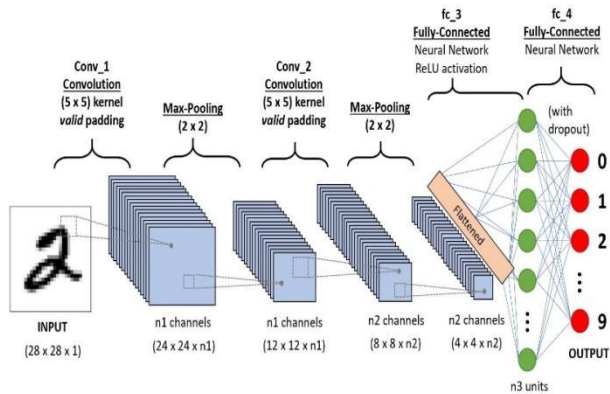


Figure 4: CNN general architecture

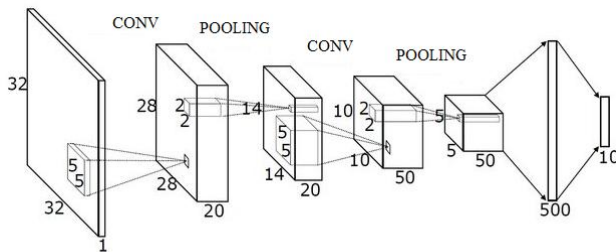


Figure 5 . Modified LeNet architecture

IV. RESULT

Apart from other research works, this works need multiple deep learning approaches. Face detection from the low light images is the first objective of this work. Three different algorithms implemented. Viola Jones algorithm is very fast in execution. But it fails to detect side faces. Dlib based MC KCF algorithm is efficient in landmark detection like eye, nose, eyebrow, mouth etc. from the analysis, Yolo v3, Yolo v3 algorithm is better than viola Jones and Dlib algorithms for face detection. Performance of MC – KCF algorithm is in table 2. The classification step is performed with modified LeNet Architecture. Drowsy data from the university of Liege contains different subjects and are annotated with Psychomotor vigilance tests. These data were trained with our CNN algorithm and created a model.

TABLE 2 . COMPARISON OF PERFORMANCES [3]

Methods	Accuracy
MTCNN	93.2%
DSST	85%
KCF+CNN	93%
MC-KCF	95%

ROC curve is used to measure the performance of our system. It is a graphical plot that displays the identifying ability of a binary classifier system as its discrimination threshold is varied.

The area under the curve in ROC curve for the proposed system in drowsy state is 0.97 and for non-drowsy state is 0.95 as shown in Figure 6.

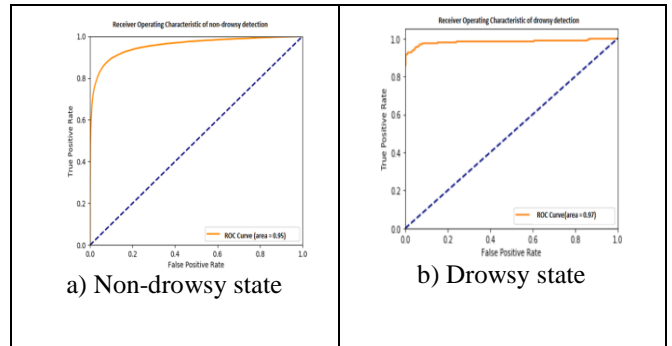


Figure 6: Performance of the proposed system

TABLE 3 . PERFORMANCE ANALYSIS

Algorithm	Accuracy
Zhang[19]	85.9%
Picot[20]	82.1%
Akrout[21]	90.2%
Dricare[3]	93.6%
Proposed Algorithm	97%

V. CONCLUSION

In this paper we demonstrated a way to determine if a driver is in a drowsy state, by using NIR camera. Yolov3 was found to be better architecture for face detection than Viola Jones or dlib method. Modified LeNet is used for classification. Accuracy of the system is 97% for 20 frames per seconds. In some cases where the face is occluded, for example by hand, we might get some erroneous detection.

For future we could extend the analysis by making it more robust to illumination changes in the video frames. A better camera and higher could further improve the system performance. The algorithmic pipeline could be implemented on a single board computer for real time applications like Advanced driver-assistance systems (ADAS). We could also include additional modalities like audio channel etc. along with the video frames and use a multi-modal machine-learning approach to improve performance.

REFERENCES

[1] National Crime Record Bureau(NCRB) (2019), “Cases Reported and Percentage Change in Traffic Accidents during 2017 Over 2016” Table 1A.1, Accidental Deaths & Suicides in India 2017 , page no 129,130. Available: <https://ncrb.gov.in/en/traffic-accidents-2017>

[2] World Health Organization (WHO) , “GLOBAL STATUS REPORT ON ROAD SAFETY 2018 – Summary”, 2019 Available: <https://www.who.int/news-room/fact-sheets/detail/road-traffic-injuries>

- [3] W. Deng and R. Wu, "Real-Time Driver-Drowsiness Detection System Using Facial Features," in *IEEE Access*, vol. 7, pp. 118727-118738, 2019, doi: 10.1109/ACCESS.2019.2936663.
- [4] Y. Sun, X. Wang and X. Tang, "Deep Convolutional Network Cascade for Facial Point Detection," 2013 IEEE Conference on Computer Vision and Pattern Recognition, Portland, OR, USA, 2013, pp. 3476-3483, doi: 10.1109/CVPR.2013.446.
- [5] D. S. Bolme, J. R. Beveridge, B. A. Draper and Y. M. Lui, "Visual object tracking using adaptive correlation filters," 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, San Francisco, CA, USA, 2010, pp. 2544-2550, doi: 10.1109/CVPR.2010.5539960.
- [6] M. Danelljan, G. Häger, F. S. Khan and M. Felsberg, "Discriminative Scale Space Tracking," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 8, pp. 1561-1575, 1 Aug. 2017, doi: 10.1109/TPAMI.2016.2609928.
- [7] S.-J. Jung, H.-S. Shin, and W.-Y. Chung, "Driver fatigue and drowsiness monitoring system with embedded electrocardiogram sensor on steering wheel," *IET Intell. Transp. Syst.*, vol. 8, no. 1, pp. 43–50, 2014.
- [8] G. Li, B. Lee and W. Chung, "Smartwatch-Based Wearable EEG System for Driver Drowsiness Detection," in *IEEE Sensors Journal*, vol. 15, no. 12, pp. 7169-7180, Dec. 2015, doi: 10.1109/JSEN.2015.2473679.
- [9] B. Warwick, N. Symons, X. Chen and K. Xiong, "Detecting Driver Drowsiness Using Wireless Wearables," 2015 IEEE 12th International Conference on Mobile Ad Hoc and Sensor Systems, Dallas, TX, USA, 2015, pp. 585-588, doi: 10.1109/MASS.2015.22.
- [10] M. Omidyeganeh, A. Javadtalab and S. Shirmohammadi, "Intelligent driver drowsiness detection through fusion of yawning and eye closure," 2011 IEEE International Conference on Virtual Environments, Human-Computer Interfaces and Measurement Systems Proceedings, Ottawa, ON, Canada, 2011, pp. 1-6, doi: 10.1109/VECIMS.2011.6053857.
- [11] Q. Massoz, T. Langohr, C. François and J. G. Verly, "The ULg multimodality drowsiness database (called DROZY) and examples of use," 2016 IEEE Winter Conference on Applications of Computer Vision (WACV), Lake Placid, NY, USA, 2016, pp. 1-7, doi: 10.1109/WACV.2016.7477715.
- [12] Anusha AV "Advanced Algorithm For Facial Expression Recognition Using Facial Patches"- *IJIE-International Journal of Innovations & Implementations in Engineering* (ISSN 2454- 3489) 2016 January Edition Volume 1.
- [13] Davis E. King, "Dlib-ml: A Machine Learning Toolkit" *Journal of Machine Learning Research* 10 (2009) 1755-1758, July 2009
- [14] Chen, W., Huang, H., Peng, S. et al. YOLO-face: a real-time face detector. *Vis Comput* (2020). <https://doi.org/10.1007/s00371-020-01831-7>
- [15] Redmon, J.; Farhadi, A., "Yolov3: An incremental improvement", arXiv, 2018; arXiv: 1804.02767
- [16] Zuopeng Zhao, "Driver Fatigue Detection Based on Convolutional Neural Networks Using EM-CNN" *Computational Intelligence and Neuroscience*, Hindawi, November 2020
- [17] Avigyan Sinha, et al. "Real Time Facial Emotion Recognition using Deep Learning", *International Journal of Innovations and Implementations in Engineering* (ISSN 2454- 3489), 2019, vol 1
- [18] Malavika Suresh, et al "Real-Time Hand Gesture Recognition Using Deep Learning", *International Journal of Innovations and Implementations in Engineering* (ISSN 2454- 3489), 2019, vol 1
- [19] Y. Zhang and C. Hua, "Driver fatigue recognition based on facial expression analysis using local binary patterns," *Optik*, vol. 126, no. 23, pp. 4501–4505, Dec. 2015
- [20] A. Picot, S. Charbonnier, A. Caplier, and N.-S. Vu, "Using retina modelling to characterize blinking: Comparison between EOG and video analysis," *Mach. Vis. Appl.*, vol. 23, no. 6, pp. 1195–1208, Nov. 2012.
- [21] B. Akrouf and W. Mahdi, "Spatio-temporal features for the automatic control of driver drowsiness state and lack of concentration," *Mach. Vis. Appl.*, vol. 26, no. 1, pp. 1–13, Jan. 2015

# A Deep Learning Approach to classify the Honeybee Species and health Identification

Karthiga M

Department of CSE

Bannari Amman Institute of Technology

Sathyamangalam, India

[karthigam@bitsathy.ac.in](mailto:karthigam@bitsathy.ac.in)

Suganya E

Research Scholar

Anna University, Chennai

[sugan.ela@gmail.com](mailto:sugan.ela@gmail.com)

Sountharajan S

School of Computing Science and

Engineering

VIT Bhopal University

Sehore, India

[sountharajan@gmail.com](mailto:sountharajan@gmail.com)

Sankarananth S

Department of Electrical and

Electronics Engineering

Excel College of Engineering and

Technology

Tamilnadu, India

[sankarananth99@gmail.com](mailto:sankarananth99@gmail.com)

Nandhini S S

Department of ISE

Bannari Amman Institute of Technology

Sathyamangalam, India

[nandhiniss@bitsathy.ac.in](mailto:nandhiniss@bitsathy.ac.in)

**Abstract**— Honey bee is one of the charming insect that utilizes a collective behavioral nature to achieve the powerful action. Protecting honey bees is one of the important jobs of every human in the world to preserve the ecological balance. Tracking and determining the several species of the bees over their life span electronically is a tedious work. Automated classification of species is important to preserve the various species of honey bees from danger. The diseases that affect the honey bees during their life span have to be detected autonomously and the spread of the diseases to other healthy honey bees has to be preserved. The proposed technique aims in classifying the several species of honey bees and identifying the diseases that are prone to honey bees. Convolution neural network with two dimensional layers are used as a classifier in the proposed model. Data augmentation using Synthetic Minority Over-sampling Technique (SMOTE) is utilized. More than 5000 images of honey bees with lot of features are used for learning purpose. The proposed methodology attained an accuracy of 86% for subspecies classification and 84% for bee health identification.

**Keywords**— Classification, Convolutional Neural Network, Synthetic Minority Over-sampling Technique, Beehive Monitoring System, Rectified Linear Unit, Visual cortex features

## I. INTRODUCTION

Like the essentiality of bacteria in day to day life, honey bees in turn helps in maintaining the ecological balance and it is most important in ecology. If there are no honey bees, the pollinated plants will be exhausted within an ample period of time. The number of honey bees is gradually decreasing due to the increased effect of global warming, modern agriculture and various parasites attack. This in turn leads to the non ripening of fruits and flowers. Recently lots of advancements are handled to increase the sustainability of honey bees. With the evolution of computing technologies and electronic devices, the monitoring of bee hive electronically and data collection regarding bee health is possible. Sensors and other devices are used in the field of ecoacoustics to estimate the unfriendly environments for bees. Forager traffic is one of the useful variables to supervise the availability of food, age of the bee colony and pesticides impact [1]. It also helps in evaluating the health of the honey bees [2]. This forager activity needs real time monitoring of bee hives, pest detection and other hive management issues. Rapid outbreaks in this forager activity may leads to sudden alteration in colony level.

In forager activity, entering and exiting the hives of the honey bees are closely monitored for a certain time span then the related data is gathered without human monitoring. Using human to monitor the status of honey bee is highly complex though accurate. So automation system using Beehive Monitoring System becomes more fabulous. This system collects all needed data without disturbing the usual behavior of bees [3]. The data in the form of audio, video and hive temperature is gathered at constant intervals by EBM.

In the proposed methodology, convolutional neural networks is utilized to classify the species of honey bees and also to correctly identify the diseases that are prone to honey bees. For recognizing the patterns in two dimensionality procedure, one of the standard Machine Learning procedures called ConvNets is utilized. It has special network architecture with layers of sampling and convolution. A two layered Convolutional network model is used in the proposed system. In order to distribute the data in an equal level among all categories, data balancing procedure is handled. The data balancing procedure utilized in the proposed methodology is Synthetic Minority Over-sampling Technique (SMOTE). SMOTE uses synthetic sampling procedure to increase the data samples in case of minor subsets. Visual cortex features are used by convolutional networks for classification. Before inputting to the classifier, image augmentation stage is handled. Rectified Linear Unit (Relu) activation function is used in the augmenting phase. The dataset consists of over 5000 images of worker bees from a bee hive with the attributes such as pollen carrying status, name of the sub species, its health condition along with the time and location. The species of the honey bees in the datasets are Russian bee, Italian bee, Carniolan bee, western honey bee, mixed local stock, VSH Italian bee and some other unknown species. The proposed methodology attained an accuracy of 86% for subspecies classification and 84% for beehealth identification. The rest of the paper is sectioned as follows: Section 2 describes about the various related researches in the proposed area, Section 3 identifies the detailed representation of the proposed method, Section 4 discussed about the training and testing results with accuracy and finally Section 5 concludes about the methodologies used in the proposed work and the accuracy obtained.

## II. RELATED WORK

The health of honey bees and the methodology to safeguard the species of honey bees to maintain the ecological balance is one of the major research fields nowadays and it is carried out by using the forager traffic. Electronic Beehive monitoring system helps in collecting the honey bee related data autonomously without any human intervention and without disturbing the ecological life span of the honey bees. As EBM gains more popularity the ways to improve the design of EBM stands first. For determining the forager traffic levels, by counting the images of omni directional bees, two algorithms are proposed in [4]. The contour detection is the first algorithm where image is augmented and the contour list with connected pixels is computed. Fewer pixel valued contours such as 30 are discarded. The number of final contours reveals the number of available bees in the pad. Binary pixel division of the landing pad is utilized in the second algorithm. Green color in the image pixels is inspected. If the number of green colors in the pixels is more than a certain threshold, then that pixels are marked as pad pixel. Otherwise they are labeled as bee pixel. Then the number of identified bee pixels is divided by 30 to determine the bees count in the landing pad. The number 30 depicts the average count of pixels that are obtained experimentally in a bee.

Omni-directional counting of bees in Langstroth pads is determined by another algorithm illustrated in [5]. The overall performance of the algorithm is tested using situ. This is an accurate algorithm compared to the previous one. Deep Learning is one of the major research areas and shows more advances in artificial intelligence day by day. It provides computation methods with more processing layers to train the system with multiple abstraction levels. This methodology drastically improved the research areas in speech recognition, handwriting recognition, object recognition, video analytics and several other domain areas. One of the challenging tasks is image classification. The various limitations that drain the accuracy level in image classification is the brightness and contrast features, scaling, angle rotation and several other features. Lot of algorithms had been proposed to improve the accuracy in image classification by utilizing lesser number of features. But it is very much difficult to identify the best solution for an application and the constant number of general features used. Deep Neural networks paved a way for improving these accuracies in a more optimized manner. These networks are often complex since high number of features are involved to represent the definite patterns in the images. The features are mainly generalized and robust enough to improve the classification accuracy. With DL models, a sustainable defeat in error rate is achieved nowadays. Yann Lecun et al in 1998 [6] proposed a model using convolutional neural networks called ConvNets to eradicate the needs of manual feature extractors. However, DL algorithms take the way due to the increased processing power, large collection of real time data and emerging new algorithms and techniques. The next outbreak is AlexNet, one of the large deep ConvNet, yielding higher accuracy in huge number of visual recognition methods. AlexNet network composes of 5 layers, max pooling layers, fully connected layers up to 3 and dropout layers. The classification algorithm for determining the prognosis of breast cancer is done by [7]. The classification is made with more than 1000 categories [8].

With the success of AlexNet, more number of models came into view; one of them is ZF Net [8]. This is a novel technique with slight modification in intermediate layers and classifier operation. Mobile based classification to assist the patients is implemented by [9] and the same technique is preferred for automatic honey bee classification. Developing an IoT based smart farm monitoring as proposed in [10] is also studied for developing an IoT based bee monitoring system. Hybrid SVM as proposed in [11] implies a hybrid classification technique for determining the glioblastoma multiforme detection and the technique is surveyed for bee monitoring. Different data sets are used to improve the accuracy of the image by utilizing several neural network models on MNIST dataset as depicted in [12, 13] where the modification in algorithm diminishes the error rate of existing models. Various machine learning models is proposed in [14] for diagnosis of Alzheimer's disease. Big Data is processed by a novel machine learning technique as proposed in [15] wherein the same processing could be used if the image dataset of honey bee monitoring grows. IoT based honey bee activity classification using deep learning techniques is proposed in [16] wherein the overall classification accuracy obtained is more than 94%. Wang et al. [17] developed a new deep learning undirected graph combined model to accurately classify the honey bee poisoning due to pesticides. In [18] computer vision approach is utilized to determine the infection level of the Varroa destructor mite using video recordings from a honey bee hive.

## III. PROPOSED METHODOLOGY

This section describes the 2 layered CNN for bee health identification and sub species classification. The following steps are adopted in the proposed system as shown in Fig 1.

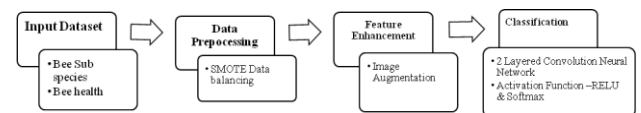


Fig. 1. Flow of the proposed CNN model

### A. Data Balancing

Data balancing is the process of distributing the data in an equal level for all the categories in which it has to be classified. The imbalanced dataset may lead to under fitting or over fitting issues, which in turn reduces the accuracy of the model. The honey bee dataset considered for classification is found to be imbalanced as shown in Fig 2(a) and Fig 2(b).

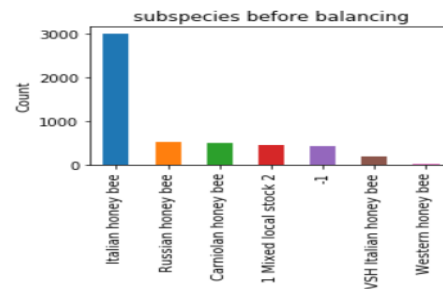


Fig. 2(a). Raw dataset distribution based on subspecies



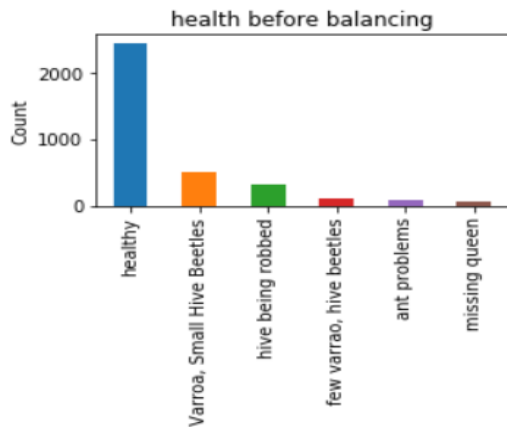


Fig. 2(b). Raw dataset distribution based on bee health

In general, performance of the model and the relation between accuracy and loss is evaluated using the Receiver Operating Characteristics (ROC) curve. The class that has more values has to be under sampled, whereas the class that has fewer values has to be over sampled. In our dataset, the Italian honey bee and healthy bee samples has to be over sampled and other class has to be under sampled. The performance purely depends on the confusion matrix that consists of fundamental values to derive accuracy, recall, precision and f1-score. Synthetic Minority Over-sampling TEchnique (SMOTE) is applied to increase the samples in the minor subsets using over sampling strategy. The SMOTE is chosen for balancing the dataset as most of the classes have less number of samples. The samples are increased by producing synthetic samples. The nearest neighbors are used to produce new samples, where the value of  $k$  is chosen in a random manner. The balanced datasets are obtained once they are processed using SMOTE algorithm available in imbalance package of python library. The balanced datasets are shown in Fig 3(a) and Fig 3(b).

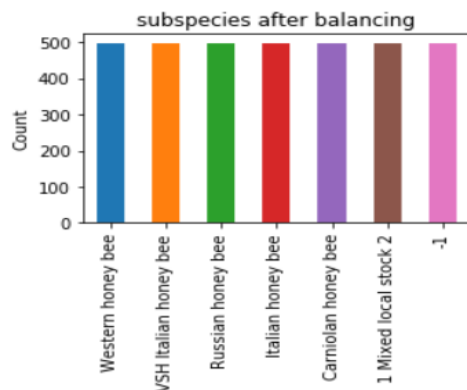


Fig. 3(a) Balanced dataset for subspecies

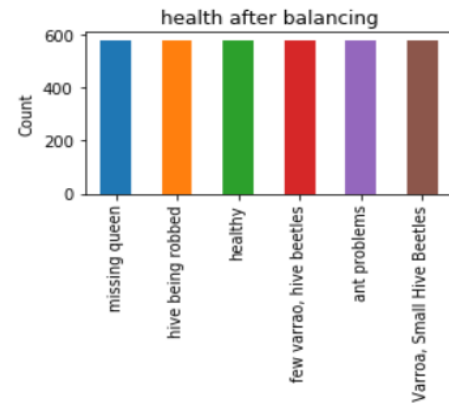


Fig. 3(b) Balanced dataset distribution for bee health

### B. Image augmentation

Transformation of images including zooming, flipping, rotating and shrinking can be done by image augmentation techniques. Augmentation can be either done before training or during the phase of training. Former results in data loss, whereas the latter reduces the loss. Here, the augmentation is done during training where the network is provided with two images as input. The input image produces another image that is masked with a layer, namely, augmented layer, which is given as input for classification model. Hence it results in two types of losses such as one at the augmented phase and other at the classification phase. Both losses are summed up together to compute the total loss. The Rectified Linear Unit (Relu) activation function is used in the augmented network. In addition to this, general flipping, rotating and zooming operations are performed before feeding the image in the augmented network.

### [3]. Convolution Neural Network Based Classification

Convolutional neural networks utilize the visual cortex features for classification which can be termed as a special kind of neural networks. The image gets converted to pixel values represented in the form of a matrix or array via the RGB channels that range between 0 and 255. The pixel intensity is denoted by the values. Initially, the image pixels get through the CNN and it is scanned from left to right and top to bottom. The filter is identified in the first step to identify the convolution. The original pixel values are multiplied and added together to provide a single value in the convolution. A matrix is produced by making the filter to scan through all the value in an image. The output of first layer is the input to the successive layers that acts as a feedback mechanism. The ReLU activation function is applied in the nonlinear layer to dense the network. The sampling operations are performed using Max2D pooling followed by ReLU. Finally, all the layers are connected to the output layer that produces a vector with  $n$  dimensions. During training, the number of iteration is given as epochs. The weights are saved once training is completed. The validation is done using the test data to evaluate the performance of the model. The architecture of the CNN model used for classification is given in Fig 4. The CNN layers for sub species classification is shown in Fig 5(a) and Fig 5(b) respectively. The layer 1 and 2 of the CNN obtained during bee health classification in shown in Fig 6(a) and Fig 6(b) respectively.

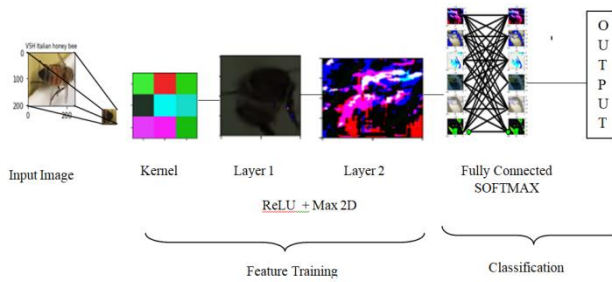


Fig. 4 Architecture of CNN for bee species and bee health identification

D. Rectified Linear Unit (ReLU)

In general, the output of the neurons are modeled using the tanh () function as given below

$$f(y) = \tanh(y) \tag{1}$$

Where,  $\tanh(y) = \frac{1}{(1+e^{-x})}$  (2)

This method of training is termed as saturating nonlinearity which is less efficient than non saturating non linearity. The non saturating nonlinearity can be represented as

$$f(y) = \max(0, y) \tag{3}$$

The ReLU is designed in such a way that it resolves the non linearity issues. Hence, the model takes negligible amount of time for training. The conventional models use the general saturating models that increase the time complexity. Here, ReLU activation function is used to overcome the limitations of sigmoid and tangent function of traditional neural network. The CNN is trained using stochastic gradient descent to backpropagate the errors. The deep learning model makes use of backpropagation that has huge datasets with labels. The linear functions are simple to be handled by neural networks; hence ReLU is applied to rectify the nonlinearities of the model.

[5]. Softmax Activation Function

The deep learning model classifies the health and subspecies of honey bees based on logistic function, hence the softmax function is applied for the multi class classification. The probabilistic sum of softmax activation function represents 1. The estimates of maximum likelihood can be attained using the softmax along with log loss. The frequencies of the classes are considered to provide a better output that has high probability values. The main point to use softmax is that the probability is distributed in all the output nodes. There will be no improvement in the result if softmax is used for binary classification but on coming to multi class classification softmax is the best way to ensure the accuracy of the model. The softmax activation function can be represented in the mathematical form as given below

$$x = \frac{e^y}{e^0 - e^y} \tag{4}$$

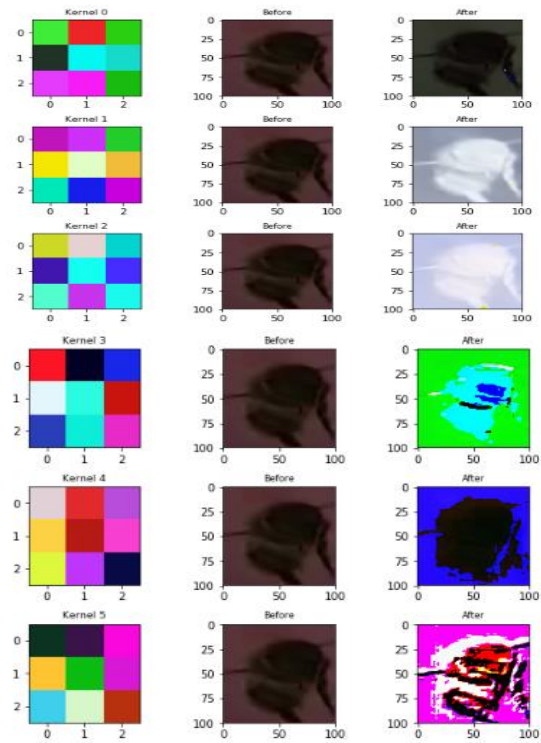


Fig. 5(a). CNN layer 0 Bee Subspecies

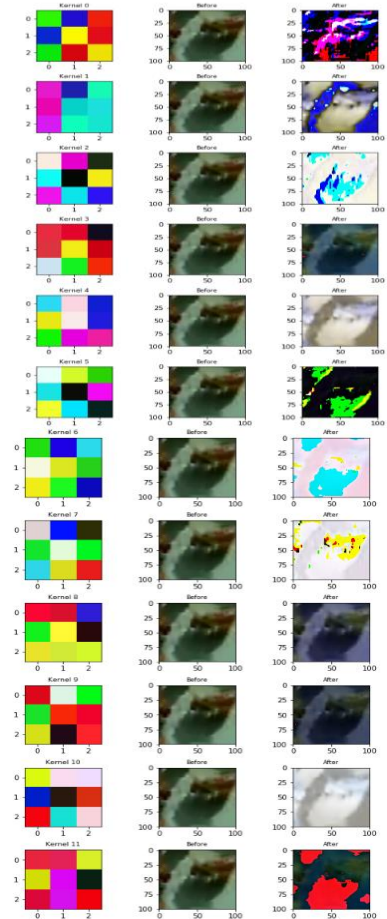


Fig. 5(b). CNN Layer 2 Bee Subspecies

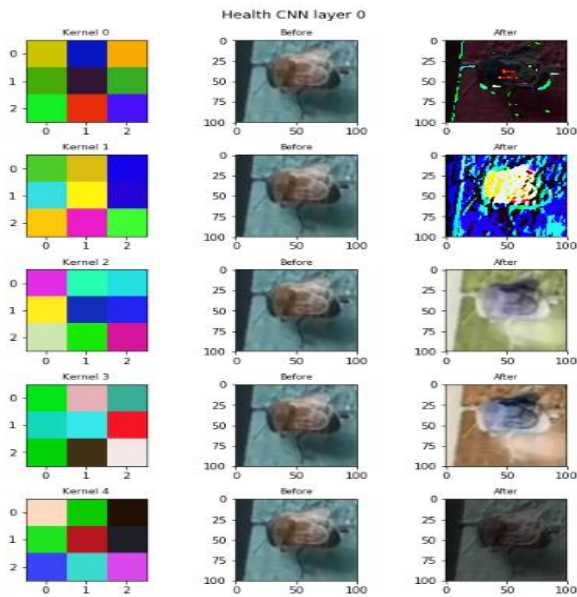


Fig. 6(a). CNN Layer 0 Bee Health

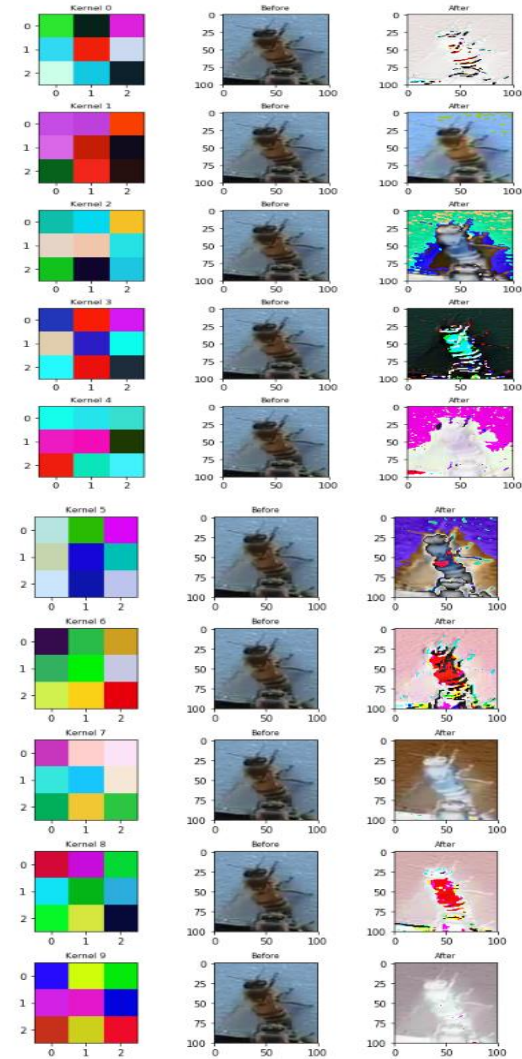


Fig. 6(b). CNN layer 2 Bee health

[6]. *Dataset Description*

The dataset consists of over 5000 images of worker bees from a bee hive with the attributes such as pollen carrying status, name of the sub species, its health condition along with the time and location. The species of the honey bees in the datasets are Russian bee, Italian bee, Carniolan bee, western honey bee, mixed local stock, VSH Italian bee and some other unknown species. The location attribute has different locations from United States of America (USA). The health attribute includes healthy bees, missing queen, bee affected by ant, robbed hives and bees affected by varroa which is a beetle that affects the bee health. Fig7(a) and Fig 7(b) shows the sample images from the dataset based on subspecies and health condition respectively. The dataset is found to be highly unbalanced, hence split-balance mechanism is applied to balance the dataset, which in turn overcome the over fitting problem. The stratified sampling technique is applied for balancing the training and testing dataset. The image dataset is augmented using Image Data Generator class of Keras in python. Then, the dataset is trained using Convolution Neural Networks (CNN) for both bee subspecies and bee health.

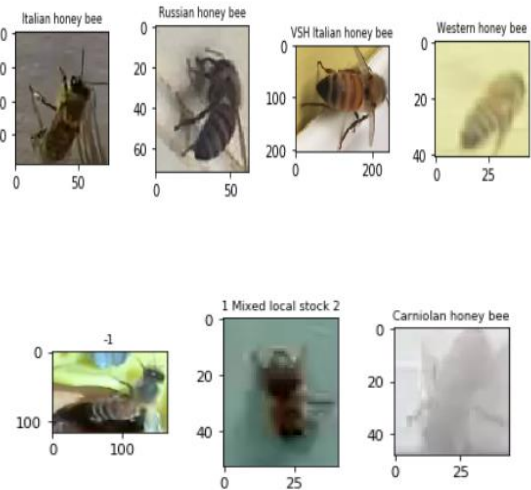


Fig.7 (a). Data samples of different subspecies of bees

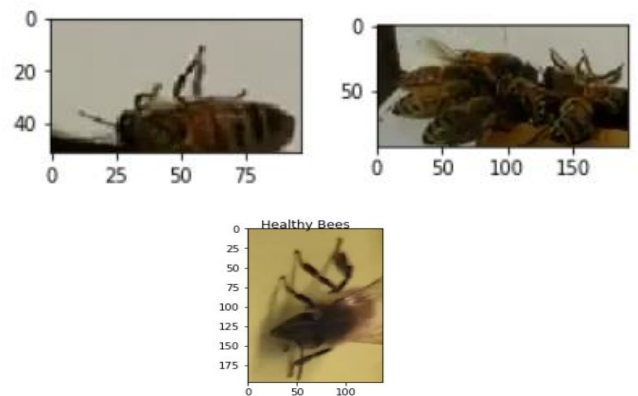


Fig. 7 (b). Data samples of healthy and unhealthy bees

#### IV. RESULT AND DISCUSSION

The CNN with 2 layers is used to learn and validate the image dataset that consists of 5000 images. The data is balanced to reduce the overfitting and the augmentation such as transformation is performed on the image. The data is split as 75% for learning and 25% for validation. A proper distribution of training dataset is ensured as shown in the dataset distribution. The performance of the model is validated using false negative, true positive, false positive and true negative. Using the 25% of testing data, the loss and accuracy of the model is validated for both subspecies of bees and health of bees. The complexity of the CNN is provided with two layers using the keras framework. The size of the kernel is set to 3 and Rectified Linear Unit (Relu) activation function is applied. Hence, there is no possibility of gradient issues and the model has completed training with less computational and less time complexity. The dense layer and Max2D pooling is applied along with softmax activation function, which provides a normalized value. The categorical cross entropy is utilized to compute the loss and the accuracy did not improved after 20 epoches. The accuracy and loss curve for both bee subspecies and bee health is depicted in Fig 8(a) and Fig 8 (b) respectively. The accuracy of classifying the different subspecies and healthy bees is plotted as bar graph in Fig 9(a) and Fig 9(b) respectively for better visualization.

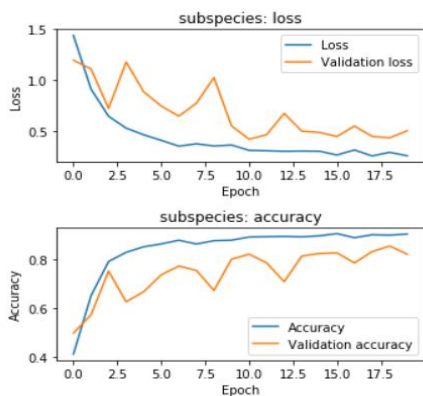


Fig. 8(a). Accuracy and loss for bee subspecies

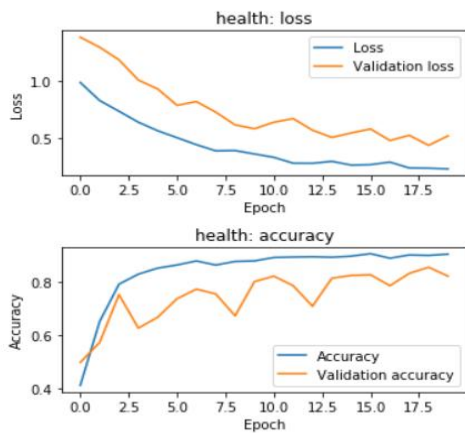


Fig. 8(b). Accuracy and loss for bee health

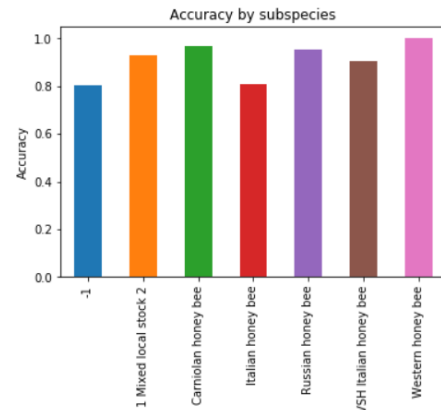


Fig. 9(a). Accuracy for bee subspecies

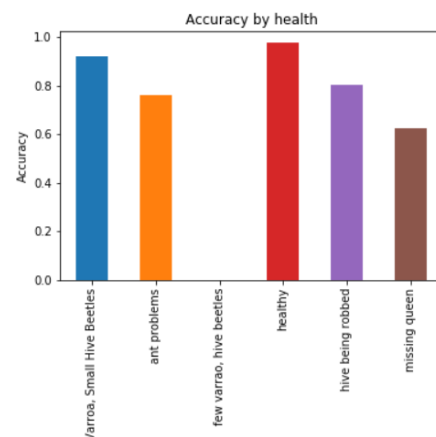


Fig. 9(b). Accuracy for bee health

Table I and Table II shows the detailed performance analysis of the model for subspecies classification and bee health in accordance to precision, recall and f1-score for each category available in the dataset. From the results, it is depicted that the subspecies classification attained an accuracy of 86% and bee health classification 84%. The model classified the western honey bee with 100% precision and recall. The Italian honey bee, Russian honey bee and Carniolan honey bee attained a good precision and recall values, whereas the model deviated for mixed local stock and VSH Italian honey bee, which in turn reduced the overall accuracy of the model. In the bee health classification part, the healthy bees are classified with high accuracy on comparison with other categories. The robbed hives are classified with an average accuracy. The model showed large deviation in the identification of hive beetles, which has to be improved by enhancing the features in the future work.

TABLE 1 PERFORMANCE MEASURES FOR SUBSPECIES CLASSIFICATION

Subspecies	Precision	Recall	f1-Score
Unknown species	0.89	0.89	0.86
Mixed local stock	0.46	0.93	0.62
Carniolan honey bee	0.97	0.97	0.97
Italian honey bee	0.97	0.82	0.89
Russian honey bee	0.99	0.97	0.98
VSH Italian honey bee	0.66	0.93	0.77
Western honey bee	1	1	1
Loss Function	0.3312		
Accuracy	0.8654		

TABLE 2 PERFORMANCE MEASURES FOR BEE HEALTH CLASSIFICATION

Bee Health	Precision	Recall	f1-Score
Varroa, Small hive beetles	0.46	0.94	0.63
Ant problems	1	0.77	0.87
Few Varroa, hive beetles	0.25	0.01	0.01
Robbed hive	0.81	0.82	0.81
Missing queen	1	0.63	0.77
Healthy	0.93	0.99	0.96
Loss Function	0.4052		
Accuracy	0.8491		

## CONCLUSION

The proposed model aims in classifying the various species of honey bees and identify the diseases that affect the honey bees. The classification algorithm that is used is 2 layered convolutional neural networks. Labeled data set with more than 5000 images and features are used for both training and validation and are tuned to increase the performance measures. The performance analysis of the model for subspecies classification and bee health in accordance to precision, recall and f1-score for each category available in the dataset is determined. From the results, it is depicted that the subspecies classification attained an accuracy of 86% and bee health classification is 84%. The model classified the western honey bee with 100% precision and recall. In the bee health classification part, the healthy bees are classified with high accuracy on comparison with other categories. The robbed hives are classified with an average accuracy. In future, the number of features utilized is enhanced to improve the accuracy of the proposed model.

## REFERENCES

- [1]. R. Mclellan, "Honeybee colony weight as an index of honey production and nectar flow: A critical evaluation.SNAP-1," 1977, pp. 401–408.
- [2]. D. A. K. L. D. J. Pham-Delegue, M.-H., "Behavioural methods to assess the effects of pesticides on honey bees," in *Apidologie*, 33, 2002.
- [3]. M. Sanford, "2nd international workshop on hive and bee monitoring," *American Bee Journal*, pp. 1351–1353, 2014.
- [4]. V. A. Kulyukin and S. K. Reka, "Toward sustainable electronic beehive monitoring: Algorithms for omnidirectional bee counting from images and harmonic analysis of buzzing signals." *Engineering Letters*, vol. 24, no. 3, 2016.
- [5]. V. A. Kulyukin, "In situ omnidirectional vision-based bee counting using 1d haar wavelet spikes," in *Proceedings of the International MultiConference of Engineers and Computer Scientists*, vol. 1, 2017.
- [6]. Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, Nov 1998.
- [7]. Sountharajan, S., Karthiga, M., Suganya, E., & Rajan, C. (2017). Automatic classification on bio medical prognosis of invasive breast cancer. *Asian Pacific Journal of Cancer Prevention: APJCP*, 18(9), 2541.
- [8]. A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, 2012, pp. 1097–1105.
- [9]. Suganya, E., Sountharajan, S., Shandilya, S. K., & Karthiga, M. (2019). Mobile cancer prophecy system to assist patients: Big data analysis and design. *Journal of Computational and Theoretical Nanoscience*, 16(8), 3623-3628.
- [10]. Suganya, E., Sountharajan, S., Shandilya, S. K., & Karthiga, M. (2019). IoT in agriculture investigation on plant diseases and nutrient level using image analysis techniques. In *Internet of Things in Biomedical Engineering* (pp. 117-130). Academic Press.
- [11]. Sountharajan, S., Suganya, E., Karthiga, M., & Rajan, C. (2018). Automatic glioblastoma multiforme detection using hybrid-SVM with improved particle swarm optimisation. *International Journal of Biomedical Engineering and Technology*, 26(3-4), 353-364.
- [12]. M. D. Zeiler and R. Fergus, "Visualizing and understanding convolutional networks," in *European conference on computer vision*. Springer, 2014, pp. 818–833.
- [13]. P. Y. Simard, D. Steinkraus, J. C. Platt et al., "Best practices for convolutional neural networks applied to visual document analysis." in *ICDAR*, vol. 3, 2003, pp. 958–962.
- [14]. Karthiga, M., Sountharajan, S., Nandhini, S. S., & Kumar, B. S. (2020, May). Machine Learning Based Diagnosis of Alzheimer's Disease. In *International Conference on Image Processing and Capsule Networks* (pp. 607-619). Springer, Cham.
- [15]. Sountharajan, S., Suganya, E., Karthiga, M., Nandhini, S. S., Vishnupriya, B., & Sathiskumar, B. (2020). On-the-Go Network Establishment of IoT Devices to Meet the Need of Processing Big Data Using Machine Learning Algorithms. In *Business Intelligence for Enterprise Internet of Things* (pp. 151-168). Springer, Cham.
- [16]. M. D. Zeiler and R. Fergus, "Stochastic pooling for regularization of deep convolutional neural networks," *arXiv preprint arXiv:1301.3557*, 2013.
- [17]. J. J. Hopfield, "Artificial neural networks," *IEEE Circuits and Devices Magazine*, vol. 4, no. 5, pp. 3–10, Sept 1988.
- [18]. G. E. Hinton, "To recognize shapes, first learn to generate images," *Progress in brain research*, vol. 165, pp. 535–547, 2007.
- [19]. Szegedy, C., Ioe, S. & Vanhoucke, V. (2016), 'Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning', *Arxiv* pp. 12–12.
- [20]. Zgank, A. (2021). "IoT-based bee swarm activity acoustic classification using deep neural networks". *Sensors*, 21(3), 676.
- [21]. Wang, F., Yang, J. F., Wang, M. Y., Jia, C. Y., Shi, X. X., Hao, G. F., & Yang, G. F. (2020). "Graph attention convolutional neural network model for chemical poisoning of honey bees' prediction". *Science Bulletin*, 65(14), 1184-1191.
- [22]. Bjerger, K., Frigaard, C. E., Mikkelsen, P. H., Nielsen, T. H., Misbiih, M., & Kryger, P. (2019). "A computer vision system to monitor the infestation level of Varroa destructor in a honeybee colony". *Computers and Electronics in Agriculture*, 164, 104898.

# Estimation of The Mean Metabolic Ratio For The Purpose of Grading of Glial Tumours of The Brain\*

\*Note: Sub-titles are not captured in Xplore and should not be used

line 1: 1<sup>st</sup> Given Name Surname  
line 2: dept. name of organization  
(of Affiliation)  
line 3: name of organization  
(of Affiliation)  
line 4: City, Country  
line 5: email address or ORCID

line 1: 4<sup>th</sup> Given Name Surname  
line 2: dept. name of organization  
(of Affiliation)  
line 3: name of organization  
(of Affiliation)  
line 4: City, Country  
line 5: email address or ORCID

line 1: 2<sup>nd</sup> Given Name Surname  
line 2: dept. name of organization  
(of Affiliation)  
line 3: name of organization  
(of Affiliation)  
line 4: City, Country  
line 5: email address or ORCID

line 1: 5<sup>th</sup> Given Name Surname  
line 2: dept. name of organization  
(of Affiliation)  
line 3: name of organization  
(of Affiliation)  
line 4: City, Country  
line 5: email address or ORCID

line 1: 3<sup>rd</sup> Given Name Surname  
line 2: dept. name of organization  
(of Affiliation)  
line 3: name of organization  
(of Affiliation)  
line 4: City, Country  
line 5: email address or ORCID

line 1: 6<sup>th</sup> Given Name Surname  
line 2: dept. name of organization  
(of Affiliation)  
line 3: name of organization  
(of Affiliation)  
line 4: City, Country  
line 5: email address or ORCID

**Abstract**—This electronic document is a “live” template and already defines the components of your paper [title, text, heads, etc.] in its style sheet. **\*CRITICAL: Do Not Use Symbols, Special Characters, Footnotes, or Math in Paper Title or Abstract.** (Abstract)

**Keywords**—component, formatting, style, styling, insert (key words)

## I. INTRODUCTION (HEADING 1)

Glial tumours are becoming increasingly rampant in the present time frame. With this reality borne in mind, there is always a two-sided possibility of the presenting glioma being benign or malignant in nature which emphatically defines the line of treatment and moreover, the fate and future of the patient<sup>1</sup>. There is abundant literature regarding diagnosing gliomas both radiologically as well as histo- pathologically. However, there is a palpable paucity of possible link in between these two modalities availed for diagnosis<sup>1, 2</sup>. There are proper laid down descriptive paragraphs about appearances of different glial tumours on both the diagnostic fronts, but then again, there is sparse literature supporting evidence attempting to correlate the two modalities<sup>3</sup>. Apart from some prominent metabolites such as Choline, Creatine and Lipid lactate, no cut off values are yet defined for most of the other metabolites found within the tumour in deranged state. As such, taking metabolite ratios into account the present study is an attempt to fill in the said research gap to the permissible extent<sup>4</sup>. Cut off values for metabolite ratios as may be observed in the present study vide Magnetic Resonance Spectroscopy in Glial tumours may aid in grading of the presenting brain tumour apart from its diagnosis on the basis of its radiological appearance and correlation thereof with the Histo-pathological appearance. Therefore the aim of this study was to evolve the mean metabolite ratios and evolve mean cut off values for metabolite ratios for the purposes of grading of glial tumors.

## II. MATERIALS AND METHODS

### A. STUDY DESIGN

Cross sectional hospital based observational study.

### B. STUDY SETTING

All patients were examined on GE Brivo MRI machine with 1.5Tesla magnetic field strength in the Department of Radiodiagnosis, Jawaharlal Nehru Medical College, a constituent unit of Datta Meghe Institute of Medical Sciences (Deemed to be University), Sawangi Meghe, Wardha.

### C. STUDY PERIOD

2 years - 12<sup>th</sup> April, 2016 to 11<sup>th</sup> April, 2018.

### D. PARTICIPANTS

All the cases registered with Acharya Vinoba Bhave Rural Hospital attached to Jawaharlal Nehru Medical College, Sawangi, Wardha, diagnosed on histo-pathological findings as Glial tumours, Meningial tumours, and Tumours of the Sellar region were included in this present study. Out of these glial tumour cases were included for this study.

### E. TUMOUR GROUPS STUDIED

Glial tumors.

### F. SAMPLE SIZE

The required sample size for the present study was calculated as under:

$$\text{Sample size} = \frac{z^2 \times p(1-p)}{e^2} \div \left( 1 + \left( \frac{z^2 \times p(1-p)}{e^2 N} \right) \right)$$

Where, z = z-score = 1.96, p = prevalence,

e = margin of error = 5%, N = population size.

Accordingly, sample size availed in the present study was 142. Out of which glial tumour cases were 54.

Inclusion criteria :-

1. All patients aged 4 years and above histopathologically diagnosed as Glial tumors
2. All patients with MRI data of satisfactory imaging quality for the purpose of correlation.
3. Those giving informed consent

Exclusion criteria :-

1. Patients aged less than 4 years of age (MR spectra of metabolites are difficult to distinguish between due to rapid metabolic changes until age of 4 years of age)
2. Patients with poor quality MRI data – due to artefact or contamination by the unwanted metabolites or magnetic inhomogeneities.
3. Patients with pacemakers.
4. Patients with cochlear implants.
5. Patients with claustrophobia.
6. Patients not giving consent.

Metabolite levels were obtained according to the values redeemed upon the application of respective imaging sequences and the individual values were used for calculation of the respective metabolite ratios. Mean metabolite ratios were calculated after getting the respective metabolite values from respective imaging sequences. Multistage logistic regression analysis was applied upon the respective metabolite ratios in order to obtain definite respective cut off point values.

### III. RESULTS

In the present study the observations that are made out are as under:

- Of the total number of 142 patients included in the present study the number of glial tumour cases was 54.

LOW GRADE BRAIN TUMOURS (GR 1 & GR 2):-  
(According to WHO classification)

It was observed that upon MRS, Choline was found to be elevated in all the 54 cases. However, it was seen that the low grade tumours showed a ‘mild’ elevation as against high grade tumours showed ‘severe’ elevation of Choline. Creatine was noted to be decreased in all the cases, to the extent that the low grade tumours showed ‘mild’ decrease whereas high grade tumours showed ‘severe’ decrease to total absence of the said metabolite. N– acetyl aspartate was seen mildly decreased in low grade tumours whereas severely decreased to total absence in high grade tumours. However, Pontine and thalamic gliomas did not show any

notable change with respect to N– acetyl aspartate. Myoinositol showed a decreasing trend as the grade increased in as much as, low grade tumours showed higher levels and vice versa for high grade tumours. Lipid lactate peaks were present in all cases but significantly higher peaks were noted in high grade tumours and mild increase was noted in low grade tumours. Glutathione glutamate metabolite peaks were seen in Meningiomas, though not significant in differentiating between benign and malignant varieties of the same, and were not seen with other tumours, with exactly similar behavior observed with reference to Alanine.

### IV. DISCUSSION

In the present study, all the high grade tumours, inclusive of all Glioblastoma multiformae, Gliomatosis cerebri, Gliosarcoma and Diffuse astrocytoma grade III, there was an evident increase in Choline, maximum in the group of GBM, decrease in Creatine, N – acetyl aspartate, Myoinositol. Lipid lactate peak was also observed to be significantly increased to a ‘severe’ level in high grade tumours. Absence of Alanine and Glutathione Peaks was observed in all high grade tumours. These findings correlate with the study done by Meyerand et al5, Dawoud et al6, Law et al7, Bendszus et al8, Tzika et al9 , Shimizu et al10, Shokry et al11, Balos et al12and Poptani et al13.

In the present study it is revealed that higher Lipid / lactate in tumour was more in favour of Necrosis and Diagnosis of high-grade Astrocytomas. This finding is consistent with the ones reported by Aragao et al14, Shimizu et al10, Darweesh et al14, and Poptani et al13,.

Upon application of logistic regression analysis, ‘cut off values’ for different metabolite ratios were calculated towards facilitation of grading of brain tumours. The observed ‘cut off values’ were as follows :

0.96 for Mi:Cr, 1.03 for Naa:Cr, Cho:LL showed it to be 1.452, 2.15 for Cho:Cr, 2.06 for Cho:Naa, 2.04 for LL:Cr and 2.48 for Cho:Mi respectively.

All of these metabolite ratios were observed to be statistically significant thereby yielding ‘cut off’ values facilitating grading purpose. However, amongst these 8 ratios, there were 4 ratios having higher Sensitivity, Specificity and High accuracy, towards classifying the brain tumour. They were Cho:Cr, Cho:Naa, LL:Cr and Cho:Mi respectively. The present study makes one conclude that that all of these 8 metabolic ratios are useful in grading a case of brain tumour, however, the 4 ratios namely Cho:Cr, Cho:Naa, LL:Cr and Cho:Mi are observed to be of better relevance and utility amongst the 8 utilizable. One of the limitations of this study was that Limited number of metabolites could be studied due to limited magnetic strength vested with 1.5 Tesla MRS Equipment.

TABLE I. STATUS OF INDIVIDUAL METABOLITES WITHIN THE INTRA-TUMORAL TISSUE IN ACCORDANCE WITH THE HISTO-PATHOLOGICAL DIAGNOSES

Table Head	Cho	Cr	NAA	MI	LL	Glx/ Glu	Ala
Diffuse astrocytoma	↑	↓	↓	↓	↑	-	-
Glioblastoma Multiforme	↑	↓	↓	↓	↑	-	-

Table Head	Cho	Cr	NAA	MI	LL	Glx/Glu	Ala
Gliomatosis cerebri	↑	↓	↓	↓	↑	-	-
Gliosarcoma	↑	↓	↓	↓	↑	-	-
Low grade glioma	↑	↓	↓	↓	↑	-	-
Pontine glioma	↑	↓	NC	NC	↑	-	-
Pleomorphic xanthoastrocytoma	↑	↓	↓	↑	↑	-	-
Thalamic glioma	↑	↓	NC	NC	↑	-	-

TABLE II. STATUS OF INDIVIDUAL MEAN METABOLITE RATIOS WITHIN THE INTRA-TUMORAL TISSUE IN ACCORDANCE WITH THE HISTOPATHOLOGICAL DIAGNOSES

TABLE HEAD	CHO:CR	CHO:NAA	CHO:MI	MI:CR	LL:CR	NAA:CHO	CHO:LL	NA:CR
Gliomatosis cerebri	4.18	3.87	5.82	0.54	3.54	0.76	2.51	0.69
Gliosarcoma	5.62	5.11	4.78	0.61	4.31	0.71	2.97	0.71
Glioblastoma Multiforme	4.22	4.88	4.58	0.29	1.68	0.65	2.31	0.58
Diffuse astrocytoma	2.19	2.76	3.22	0.85	2.88	0.98	0.98	0.83
Pleomorphic xanthoastrocytoma	1.37	1.76	1.87	0.89	1.79	1.28	1.38	0.92
Pontine glioma	1.87	1.76	2.01	1.01	1.76	0.87	1.29	1.46
Low grade glioma	1.52	1.87	2.11	1.22	1.77	1.21	1.15	1.28
Thalamic glioma	1.67	1.49	1.81	0.98	1.72	1.28	1.31	1.18

In regard to the Metabolite ratios, the following ratios were calculated and analyzed –

- Choline : Creatine
- Choline :N Acetyl aspartate
- N- Acetyl aspartate :Creatine,
- Myoinositol :Creatine,
- Choline :Myoinositol,
- Lipid Lactate :Myoinositol,
- N-Acetyl Aspartate :Myoinositol,
- Choline :Lipid Lactate
- N-Acetylc Aspartate :Choline,
- Lipid Lactate :Creatine
- Lipid Lactate :N-Acetyl Aspartate.

TABLE III. OVERALL MEAN METABOLITE RATIOS UPON MAGNETIC RESONANCE SPECTROSCOPY IMAGING SEQUENCE (N=54)

PARAMETER (N=54)	MEAN	SD	MIN-MAX
Choline: Creatine cerebri	2.86	1.70	0.97-6.19
Choline : N acetylaspartate	2.58	1.67	0.091-6.57
N-acetyl aspartate :creatine	1.61	8.48	0.36-102

PARAMETER (N=54)	MEAN	SD	MIN-MAX
Myoinositol : creatine	0.89	0.336	0.23-1.54
Choline :myoinositol	2.91	1.72	1.11-6.83
Lipid lactate :myoinositol	3.03	1.71	1.12-6.64
N-acetyl aspartate :myoinositol	1.31	0.409	0.71-2.01
Choline : lipid lactate	1.20	0.45	0.16-1.98
N-acetyl aspartate : choline	1.19	1.29	0.14-1.54
Lipid lactate : Creatine	1.98	0.972	0.92-3.41
Lipid lactate : N acetylaspartate	2.01	0.387	1.12-3.21

The mean metabolite ratio for Cho: Cr was 2.86, for Cho:NAA it was 2.58, for NAA:Cr it was 1.61, for Mi:Cr it was 0.89. For Cho:Mi it came out to be 2.91, it was 3.03 for LL:Mi. It was observed to be 1.31 for NAA:Mi. and for NAA:cho it was 1.19, for LL:Cho it was 1.98 and for LL: NAA, it was 2.01.

TABLE IV. TUMOUR DIFFERENTIATION-(LOW GRADE/ HIGH GRADE) USING LOGISTIC REGRESSION APPLICABLE TO CHARTING OUT CUT OFF VALUES FOR DIAGNOSTIC PURPOSE : MRS

Metabolite ratio	Cutpoint	Sensitivity	Specificity	Correctly classified	LR+	LR-
Mi:Cr	( >= 0.96 )	80.85%	72.63%	75.35%	2.9542	0.2636
NAA:Cr	( >= 1.03 )	80.85%	81.05%	75.35%	2.9542	0.2636
Cho:LL	( >= 1.452 )	74.47%	74.74%	74.65%	2.9477	0.3416
NAA:Cho	( >= 1.277 )	80.00%	83.80%	83.80%	4.5745	0.1064
<b>Cho:Cr</b>	<b>( &gt;= 2.15 )</b>	<b>92.98%</b>	<b>92.96%</b>	<b>92.96%</b>	<b>13.1725</b>	<b>0.0755</b>
<b>Cho:NAA</b>	<b>( &gt;= 2.06 )</b>	<b>94.74%</b>	<b>94.12%</b>	<b>94.37%</b>	<b>16.1053</b>	<b>0.0559</b>
<b>LL:Cr</b>	<b>( &gt;= 2.04 )</b>	<b>92.98%</b>	<b>91.76%</b>	<b>92.25%</b>	<b>11.2907</b>	<b>0.0765</b>
<b>Cho:Mi</b>	<b>( &gt;= 2.48 )</b>	<b>98.25%</b>	<b>98.82%</b>	<b>98.59%</b>	<b>83.5090</b>	<b>0.0178</b>

Fig. 1. Example of a figure caption. (figure caption)

### V. CONCLUSION

In the present study the metabolite ratios were calculated after obtaining individual metabolite counts upon MRS. Cut off values have been identified in respect of significant metabolite ratios in terms of higher sensitivity, specificity



and accuracy (more than 90%) in differentiating low grade from high grade tumors. The research gap analysis towards which research question was framed stands filled up by the generated new knowledge in terms of 'Significant Metabolite Ratios' with identified cut off / demarcating values for gradation of Glial Tumors into Low Grade / High Grade tumors.

Chart 1: Multistage logistic regression analysis of Naa: Cr metabolite ratio: -

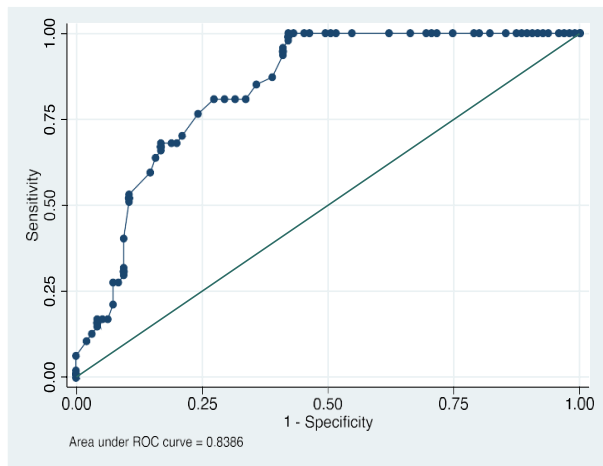


TABLE V. DETAILED REPORT OF CUT POINT, SENSITIVITY, SPECIFICITY AND ACCURACY FOR NAA:CR

CUTPOINT	SENSITIVITY	SPECIFICITY	CORRECTLY CLASSIFIED (ACCURACY)	LR+	LR-
(>=0.96)	80.85%	72.63%	75.35%	2.9542	0.2636

TABLE VI. DETAILED REPORT OF ROC CURVE ANALYSIS FOR NAA:CR

ROC curve		-- Binomial Exact -		
Obs	Area	Std. Err	[95% Conf. Interval]	
142	0.8386	0.0325	0.76694	0.89447

Chart 2: Multistage logistic regression analysis of Mi: Cr metabolite ratio:-

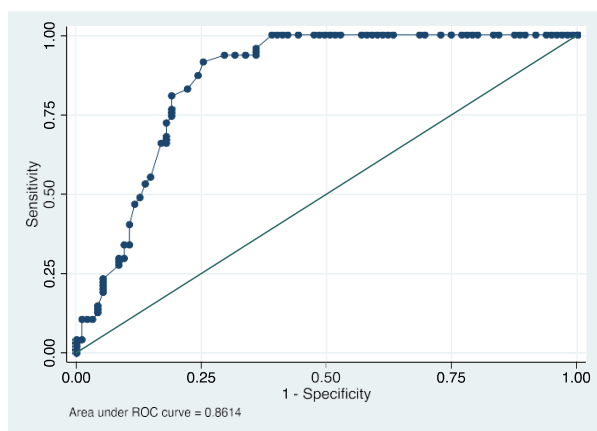


TABLE VII. DETAILED REPORT OF CUT POINT, SENSITIVITY, SPECIFICITY AND ACCURACY FOR MI:CR

Cutpoint	Sensitivity	Specificity	Correctly classified (Accuracy)	LR+	LR-
(>= 1.03)	80.85%	81.05%	80.99%	4.2671	0.2363

TABLE VIII. DETAILED REPORT OF ROC CURVE ANALYSIS FOR MI:CR

ROC curve		-- Binomial Exact --		
Obs	Area	Std. Err	[95% Conf. Interval]	
142	0.8614	0.0302	0.79089	0.91180

Chart 3: Multistage logistic regression analysis of Cho: LL metabolite ratio: -

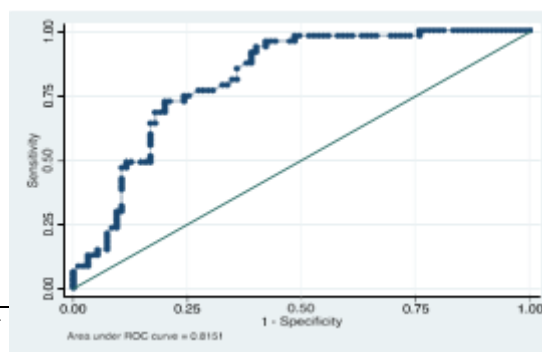


TABLE IX. DETAILED REPORT OF CUT POINT, SENSITIVITY, SPECIFICITY AND ACCURACY FOR CHO:LL

Cut point	Sensitivity	Specificity	Correctly classified (Accuracy)	LR+	LR-
(>= 1.452)	74.47%	74.74%	74.65%	2.9477	0.3416

TABLE X. DETAILED REPORT OF ROC CURVE ANALYSIS FOR CHO: LL

ROC curve		-- Binomial Exact --		
Obs	Area	Std. Err	[95% Conf. Interval]	
142	0.8151	0.0355	0.74334	0.87677

Chart 4 : Multistage logistic regression analysis of NAA:Cho metabolite ratio :-

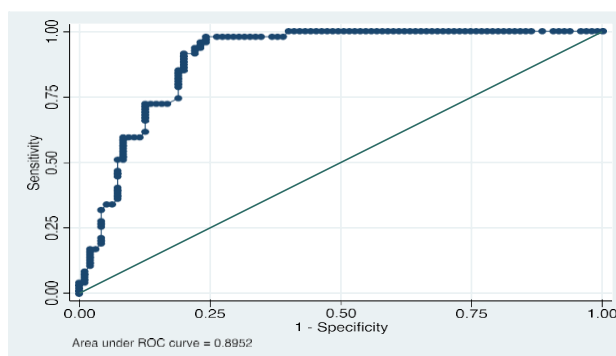


TABLE XI (A) - DETAILED REPORT OF CUTPOINT, SENSITIVITY, SPECIFICITY AND ACCURACY FOR NAA:CHO

Cutpoint	Sensitivity	Specificity	Correctly classified (Accuracy)	LR+	LR-
(>= 1.277)	91.49%	80.00%	83.80%	4.5745	0.1064

TABLE XI (B) - DETAILED REPORT OF ROC CURVE ANALYSIS FOR NAA:CHO

ROC curve		-- Binomial Exact --		
Obs	Area	Std. Err	[95% Conf. Interval]	
142	0.8952	0.0260	0.83176	0.93966

Chart 5 : Multistage logistic regression analysis of Cho: Cr metabolite ratio :-

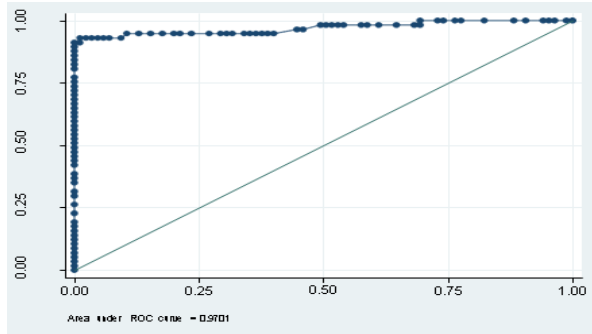


TABLE XII(A) - DETAILED REPORT OF CUTPOINT, SENSITIVITY, SPECIFICITY AND ACCURACY FOR CHO: CR

Cutpoint	Sensitivity	Specificity	Correctly classified (Accuracy)	LR+	LR-
(>= 2.15)	92.98%	92.94%	92.96%	13.1725	0.0755

TABLE XII(B) - DETAILED REPORT OF ROC CURVE ANALYSIS FOR CHO: CR

ROC curve		-- Binomial Exact --		
Obs	Area	Std. Err	[95% Conf. Interval]	
142	0.9701	0.0165	0.92945	0.99227

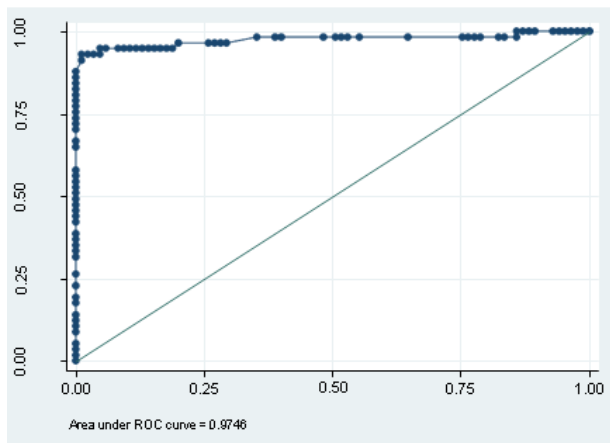


TABLE XIII(A) - DETAILED REPORT OF CUTPOINT, SENSITIVITY, SPECIFICITY AND ACCURACY FOR CHO: NAA

Cutpoint	Sensitivity	Specificity	Correctly classified (Accuracy)	LR+	LR-
(>= 2.06)	94.74%	94.12%	94.37%	16.1053	0.0559

TABLE XIII(B) - DETAILED REPORT OF ROC CURVE ANALYSIS FOR CHO: NAA

ROC curve		-- Binomial Exact --		
Obs	Area	Std. Err	[95% Conf. Interval]	
142	0.9746	0.0164	0.92945	0.99227

CHART 7 : MULTISTAGE LOGISTIC REGRESSION ANALYSIS OF LL:CR METABOLITE RATIO

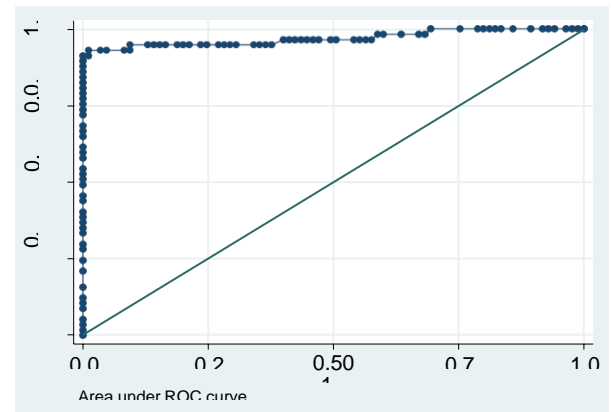


TABLE XIII (A) - DETAILED REPORT OF CUTPOINT, SENSITIVITY, SPECIFICITY AND ACCURACY FOR LL:CR

Cutpoint	Sensitivity	Specificity	Correctly classified (Accuracy)	LR+	LR-
(>= 2.04)	92.98%	91.76%	92.25%	11.2907	0.0765

TABLE XIII(B) - DETAILED REPORT OF ROC CURVE ANALYSIS FOR LL:CR

ROC curve		-- Binomial Exact --		
Obs	Area	Std. Err	[95% Conf. Interval]	
142	0.969	0.0172	0.92945	0.99227

CHART 8 : MULTISTAGE LOGISTIC REGRESSION ANALYSIS OF CHO: MI METABOLITE RATIO :-

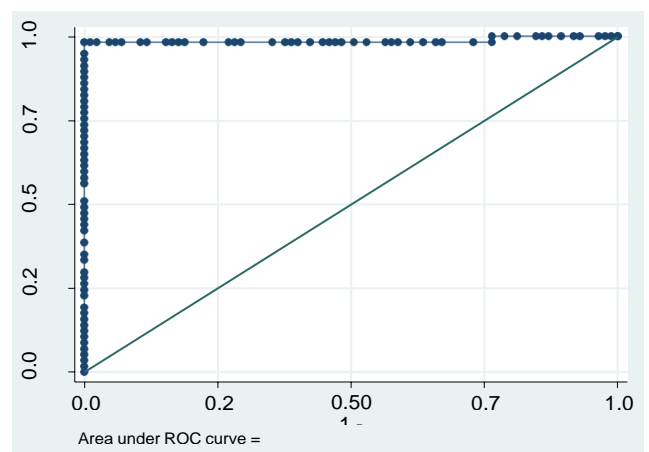


TABLE XIV(A) - DETAILED REPORT OF CUT POINT, SENSITIVITY, SPECIFICITY AND ACCURACY FOR CHO: MI

Cut point	Sensitivity	Specificity	Correctly classified (Accuracy)	LR+	LR-
(>= 2.48 )	98.25%	98.82%	98.59%	83.5090	0.0178

TABLE XIV(B) - DETAILED REPORT OF ROC CURVE ANALYSIS FOR CHO: MI

ROC curve		-- Binomial Exact --		
Obs	Area	Std. Err	[95% Conf. Interval]	
142	0.9866	0.0134	0.95005	0.99829

REFERENCES

[1] Novelize, Robert. Squire's Fundamentals of Radiology. Harvard University Press. 5th edition. 1997. ISBN 0-674-83339-2 p. 1.

[2] Hoeffner EG, Mukherji SK, Srinivasan A, Quint DJ. Neuroradiology back to the future: brain imaging. *AJNR*. 2012;33(1):5-11.

[3] Kieffer SA, Brace JR. Intracranial neoplasm. In: Haaga JR, Dogra VS, Forsting M, Gilkeson RC, Ha HK, Sundram M, editors. CT and MRI of the whole body, 5th ed. Philadelphia: Mosby Elsevier; 2009. Vol 1. pp. 49-144.

[4] Kaddah RO, Khalil ME. Malignant focal brain lesions. Value of MRS tumour biomarkers in preoperative prediction of grades of malignancy. *The Egyptian Journal of Radiology and Nuclear Medicine*. 2014 DEC; 45(4): 1201–1208.

[5] Meyerand ME, Pipas JM, Mamourian A, Tosteson TD, Dunn JF. Classification of biopsy-confirmed brain tumors using single-voxel MR spectroscopy. *AJNR Am J Neuroradiol*. 1999 Jan;20(1):117-23.

[6] Dawoud MAE, Sherif MF, Eltomey MA. Apparent diffusion coefficient and Magnetic resonance spectroscopy in grading of

malignant brain neoplasms. *The Egyptian Journal of Radiology and Nuclear Medicine*. 2014 Dec; 45(4), 1215–1222.

[7] Law M, Yang S, Wang H, Babb JS, Johnson G, Cha S, et al. Glioma grading: sensitivity, specificity, and predictive values of perfusion MR imaging and proton MR spectroscopic imaging compared with conventional MR imaging. *AJNR Am J Neuroradiol*. 2003 Nov-Dec;24(10): 1989-98.

[9] Bendszus M, Warmuth-Metz M, Klein R, Burger R, Schichor C, Tonn JC, et al. MR spectroscopy in gliomatosis cerebri. *AJNR Am J Neuroradiol*. 2000 Feb;21(2):375-80.

[10] Tzika AA, Vajapeyam S, Barnes PD. Multivoxel proton MR spectroscopy and hemodynamic MR imaging of childhood brain tumors: preliminary observations. *AJNR Am J Neuroradiol*. 1997 Feb;18(2):203-18.

[11] Shimizu H, Kumabe T, Tominaga T, Kayama T, Hara K, Ono Y et al. Noninvasive evaluation of malignancy of brain tumors with proton MR spectroscopy. *AJNR Am J Neuroradiol*. 1996 Apr;17(4):737-47

[12] Shokry A. MRS of brain tumors: Diagrammatic representations and diagnostic approach. *The Egyptian Journal of Radiology and Nuclear Medicine*. 2012 Dec;43(4):603-612.

[13] Balos DR, Gavrilović S, Lavmić S, Vasić B, Macvanski M, Damjanović D, et al. Proton magnetic resonance spectroscopy and apparent diffusion coefficient in evaluation of solid brain lesions. *Vojnosanit Pregl*. 2013 Jul;70(7) :637-44.

[14] Poptani H, Gupta RK, Jain VK, Roy R, Pandey R. Cystic intracranial mass lesions: possible role of in vivo MR spectroscopy in its differential diagnosis. *Magn Reson Imaging*. 1995;13(7):1019-29.

[15] De Fatima Vasco Aragao M, Law M, Batista de Almeida D, Fatterpekar G, Delman B, Bader AS, et al. Comparison of Perfusion, Diffusion, and MR Spectroscopy between Low-Grade Enhancing Pilocytic Astrocytomas and High Grade Astrocytomas. *AJNR*. 2014 Aug;35(8):1495-502.

[16] Darweesh AN, Badawy ME, Hamesa M, Saber N. Magnetic resonance spectroscopy and diffusion imaging in the evaluation of neoplastic brain lesions. *The Egyptian Journal of Radiology and Nuclear Medicine* . 2014 Jun;45 (2): 485–493

**2021 IEEE Seventh  
International Conference on  
Bio signals, Images and  
Instrumentation  
ICBSII 2021**

**SESSION VI  
RESEARCH  
PAPERS**

# Retinal Vessel Segmentation with Slime-Mould-Optimization based Multi-Scale-Matched-Filter

Seifedine Kadry  
Faculty of Applied Computing and  
Technology, Noroff University College,  
Kristiansand, Norway  
email: skadry@gmail.com

Venkatesan Rajinikanth  
Department of Electronics and  
Instrumentation Engineering  
St. Joseph's College of Engineering  
Chennai 600119, India  
email: v.rajinikanth@ieee.org

Robertas Damaševičius  
Faculty of Applied Mathematics  
Silesian University of Technology  
44-100 Gliwice, Poland  
email: robertas.damasevicius@polsl.pl

David Taniar  
Faculty of Information Technology  
Monash University  
Clayton, Victoria 3800, Australia  
email: David.Taniar@monash.edu

**Abstract**— Even though a number of sensory organs are existing, eye plays a necessary role in the sensory system; which converts the incident light into meaningful visual information. If any abnormality arises in eye, then the whole sensory system gets stressed. The disease in eye is due to injury, infection and ageing and the untreated eye disease will lead to vision loss. The proposed research aims to propose a Computer-Aided-Procedure (CAP) to extract the blood-vessel section from Digital-Fundus-Image (DFI). In order to accomplish this task, a Multi-Scale-Matched-Filter (MSMF) is designed using the Slime-Mould-Optimization algorithm. In this work, the necessary test images are collected from the benchmark DRIVE and CHASE\_DB1 dataset. After extracting the blood-vessel using the MSMF, an examination among extracted vessel and the Ground-Truth (GT) image is executed and the Image-Performance-Values are separately computed for each database. The attained result with this CAP confirms that the attained Jaccard, Dice and Accuracy of proposed approach is better compared to similar existing approaches in the literature.

**Keywords**—Fundus image, Blood-vessel, Slime-Mould-Optimization, Multi-Scale-Matched-Filter, Validation.

## I. INTRODUCTION

The role of the sensory organs is to collect different information from the body parts and convey the collected information to the brain for further action. During this process, the brain will evaluate the sensory information based on the priority and takes the necessary action to complete the task based on the decision arrived after processing the sensory information [1-3]. Eye is one of the chief units and in this scheme eye disorder will severely influence the sensory part.

In humans, the eye abnormality happens due to; (i) Injury, (ii) Chronic disease and (iii) Aging [4-8]. The eye abnormality due to aging is one of the key causes and hence a scheduled medical check up is recommended for the elderly people. Normally, the eye abnormality will be screen by an experienced ophthalmologist and during this screening process, the complete eye sections are recorded using a special imaging unit called the fundus-camera and the attained RGB scale image from this process is called the fundus-image.

Screening of the eye section to detect the abnormality is one of the commonly recommended medical practice and after the recording the essential image, the retinal section is examined by the ophthalmologist and also using the Computer-Aided-Procedure (CAP). The early detection will

help to cure the disease with appropriate treatment procedures. Usually, the image processing assisted procedures are implemented to examine optic-disc, optic-cup, cup/disc ratio, macula and blood-vessel for various retinal disease assessments [4-10]. The proposed research aims to develop a CAP to extract the blood-vessel from the fundus image with better segmentation performance. Extraction and evaluation of this section is essential to detect the change/infection in blood vessels for the disease assessment and treatment implementation.

Extraction and assessment of the blood-vessel from retinal image is one of the vital research domains and a considerable number of research works are proposed and implemented by the researchers with a variety of techniques. Multi-Scale-Matched-Filter (MSMF) is one of the recent technique helps to extract the blood-vessel section with improved performance. In this work, a MSMF is applied to extract the retinal vessel from the considered test picture. The performance of the MSMF is enhanced my including the Slime-Mould-Optimization (SMO) algorithm. The developed CAP consist different pre-processing stages, such as resizing (to get uniform image dimension of 512x512x3 pixels), green-channel extraction, contrast enhancement, morphological operation, MSMF implementation and comparison of the extracted blood-vessel with GT.

The performance of the MSMF implemented to mine the blood-vessel is enhanced with the SMO algorithm and this technique helps to mine the blood-vessel sector with improved correctness. After the extraction, a comparison is performed along with the Ground-Truth (GT) pictures and the important performance values are computed.

The developed CAP is tested and validated using the benchmark images of DRIVE (40 images) [11,12] and CHASE\_DB1 (28 images) [13,14] and the proposed scheme is separately implemented for each case and the average result is considered to validate the performance. Every test image is initially resized into 512x512x3 pixels and for the assessment only the green channel section of the RGB alone examined and the results are validated. The experimental outcome of this research confirms that the proposed CAP works well on both the datasets and provides an improved outcome.

The remaining sections are allocated as follows; Section 2 presents the earlier research works, Section 3 shows the methodology, Section 4 and 5 presents experimental outcome and conclusion of this research.

## II. RELATED RESEARCH

Medical image examination is one of the key research domains with a grater clinical importance and hence a number of works are proposed to examine a class of images with varied image modalities and disease cases using the benchmark as well as the real clinical images. Further, the assessment of the grayscale images are quite simple compared to the RGB scale images and hence a perfectly developed CAP is essential to examine the RGB scale images recorded using a chosen modality.

Retinal disease assessment using the fundus-picture is one of the proven research domain in which the evaluation of a variety of retinal section is discussed with the aid of the fundus-picture. Blood-vessel extraction from retinal-picture is widely addressed in the literature, since the change in the blood vessel section will help to predict different vital information related to the eye [15-18].

In the earlier work, traditional and heuristic algorithm supported blood-vessel extraction is widely discussed by the researchers; in which the prime task is to employ a semi-automated/automated system to extract the vessel region with recommended accuracy.

In the proposed work, a multi-scale matched filter is implemented to extract the vessel section from the pre-processed fundus-picture. The proposed work is based on the earlier scheme by Sreejini and Govindan [4], who implemented a Particle-Swarm-Optimization based MSMF (PSO-MSMF) to extract the retinal vessels from the DRIVE database. The other work by Keerthana et al. [5] implemented a Firefly-Algorithm technique to achieve the same. Along with the above approaches, few more research are also implemented the customary MSMF to extract the vessel section using a chosen imaging scheme.

Every other technique implemented in the literature implements; a traditional segmentation, image masks based vessel extraction and filter based techniques and in every system; the key task is the extraction of the vessel with an appropriate practice.

The earlier work related to the blood vessel extraction from the fundus image is clearly depicted and in this work, the average result attained with the SMA-MSMF is compared and validated with the result of other procedures ad based on this comparison the performance of the proposed CAP is confirmed.

## III. METHODOLOGY

The aim of this work is to develop an automated blood-vessel mining methodology for the RGB scale retinal pictures. The proposed CAP consist different pre-processing and post-processing schemes and the essential stages existing is depicted in Figure 1.

Initially the collected trial pictures are resized to 512x512x3 pixels and these images are then considered for the experiment. From the RGB scale image, the green-channel section is separated and evaluated. Initially, the contrast of the image is improved and then the morphological enhancement also implemented to improve the visibility of blood-vessel in the chosen trial picture segment. Later, SMA based MSMF is then implemented to trace and extract the blood sections and finally the extracted fragment is compared with GT to validate the performance

of the executed scheme based on the computed performance values.

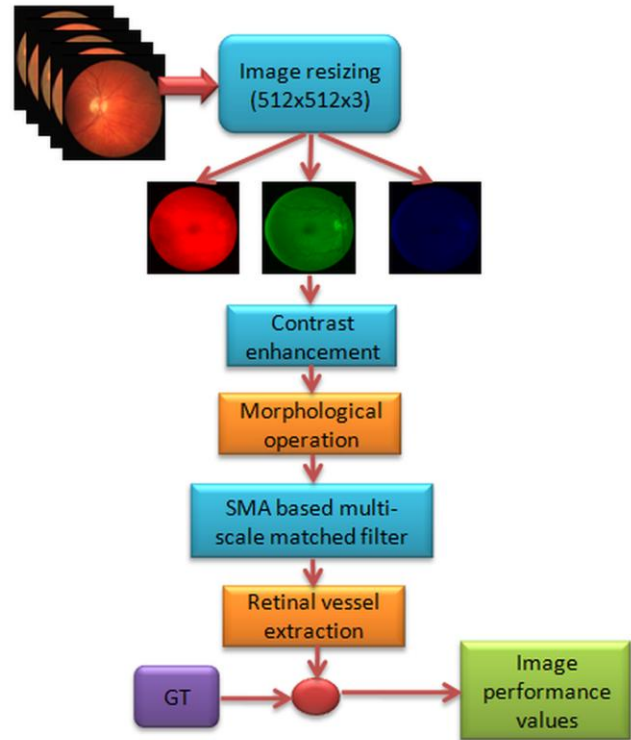


Fig.1. Scheme of the Computer-Aided-Procedure to extract the blood-vessel from fundus image

### A. Image Database

The choice of appropriate image database is essential to confirm the performance of the proposed CAP tool. In this research, the extraction of the blood-vessel is adopted as the prime task and to confirm the merit of the implemented scheme, the benchmark fundus pictures are considered from the DRIVE [11,12] and CHASE\_DB1 [13,14] databases and the sample pictures and the information are presented in Figure 2, Figure 3 and Table 1 respectively.

In both the dataset, every image is associated with the binary form of the Ground-Truth (GT) image in which the vessel section is with a value of binary 1 and the background is with a value of binary 0. The other information on these databases can be found in the earlier works [19-21].

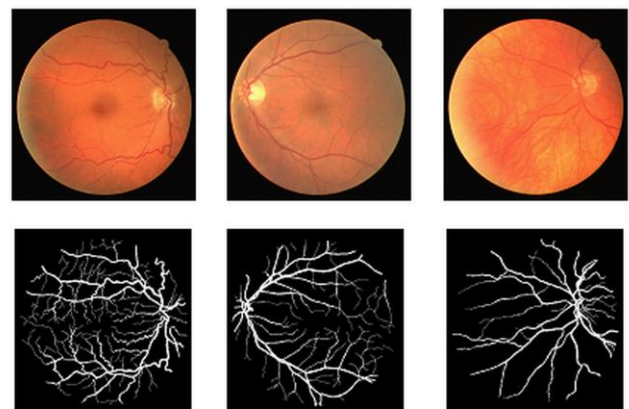


Fig.2. Sample trial picture and GT of DRIVE database

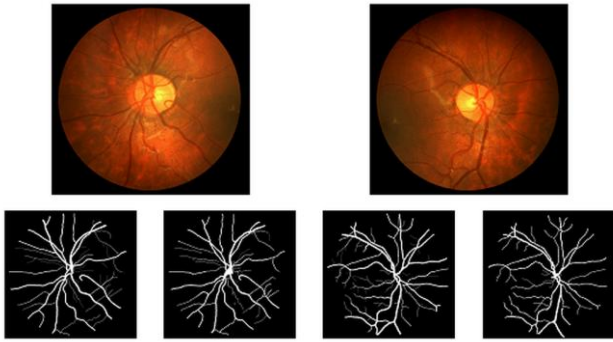


Fig.3. Sample trial picture and GT of CHASE\_DB1 database

TABLE I. RETINAL IMAGE DATABASE INFORMATION

Database	Dimension	Modified size	Total images
DRIVE	565x584x3	512x512x3	40
CHASE_DB1	999x960x3	512x512x3	28

### B. Image Pre-processing

In automated image examination, the essential pre-processing is necessary to enhance/modify the test picture based on the need. In this work, the implemented pre-processing involves; (i) Green-channel slice separation, (ii) Contrast enrichment, and (iii) Morphology supported blood-vessel pixel improvement. All these approaches are implemented to improve the essential section from the retinal picture and after the pre-processing; the proposed MSMF is implemented to extract the vessel section. The complete information regarding the pre-processing can be found in [22].

### C. Multi Scale Matched Filter

The conventional matched-filter (MF) is a general pattern matching practice to perform the blood vessel extraction. The different operations involved in this practice are as follows;

Let a fundus-image is having a dimension of  $H, W$  and the value of the Gaussian MF can be represented as;

$$k(H, W) = -\exp\left(\frac{-H^2}{2S^2}\right); \forall |W| \leq L/2 \quad (1)$$

where  $k$ =kernel,  $S$ =scale value and  $L$ =length.

In Eqn. (1),  $k(H, W)$  denotes the rotation in varied angles to detect the vessel value. The proposed work implemented a rotation of  $15^\circ$  using an appropriate kernel value as below;

$$\text{Kernel}(k) = \frac{180^\circ}{15} = 12 \quad (2)$$

Other essential information regarding the employed MSMF can be found in the earlier works [4,5].

### D. Slime Mould Algorithm

Recently, more numbers of heuristic/evolutionary methods are proposed to solve a variety of optimization works and the SMA is one of the recent approach invented in 2020 [23]. The concept and the implementation of SMA is

simple and in which the mathematical model for the Slime-Mould searching the food is adopted to find the best solution for the chosen task. This method is gaining the popularity in recent days, since the merit of this approach is its exploration capacity and lesser number of parameters to be assigned. In this work, the algorithm values to be assigned is less, such as agent size, search dimension, upper/lower search bound, exploration tactic, and terminating criteria. Figure 4 depicts the exploration of the Slime-Mould in a 2D search space and similar methodology will be implemented for other search space. Essential information regarding the SMO can be accessed from [24,25].

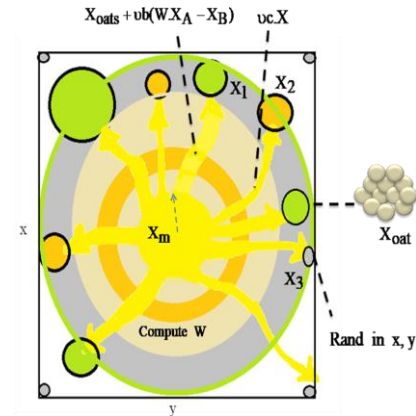


Fig.4. Structure of artificial Slime-Mould searching the food

In this work, the initial algorithm constraints are allocated as; agent dimension=25, search dimension = the size of filter, upper-bound=255, lower-bound=0, number of iterations=5000 and stopping criteria = maximal iteration.

### E. Validation

The performance of the proposed CAP using the SMA-MSMF is to be validated to confirm its clinical significance. In this work, the proposed technique mines the essential blood-vessel section in binary form. After the extraction, the binary vessel image is compared against the GT available in the database and this comparison helps to achieve the performance values. In this work, the values, such as Accuracy, Sensitivity, Specificity, and F1-Score are separately computed for DRIVE and CHASE\_DB1 dataset and the average value is then considered for the validation of the SMA-MSMF with other available methods in the literature.

### F. Implementation

The implementation of the proposed SMA-MSMF is as follows;

Step 1: Collect and resize the retinal images into the prescribed dimension

Step 2: Enhance the contrast and improve the blood vessel section using the morphological approach

Step 3: Assign the optimal values for the SMA and execute the SMA-MSMF technique to extract the blood-vessel

Step 4 Execute the comparison among blood-vessel region and GT and compute essential performance values

Step 5: Compare the attained result with the existing results in the literature and confirm the significance of proposed CAP.

#### IV. RESULT AND DICUSSION

This part of the research presents the experimental results attained with the proposed procedure using the MATLAB software.

Initially, proposed work is executed using the DRIVE database and the pre-processing and attained result for a test image case is clearly depicted in Figure 5 and 6 respectively. Fig. 5(a) to (c) shows the various channel images collected from the RGB scale trial picture. The visibility of the blood-vessel is good in green-channel compared to the alternatives and hence, in the proposed work; only the green-channel is considered for the assessment.

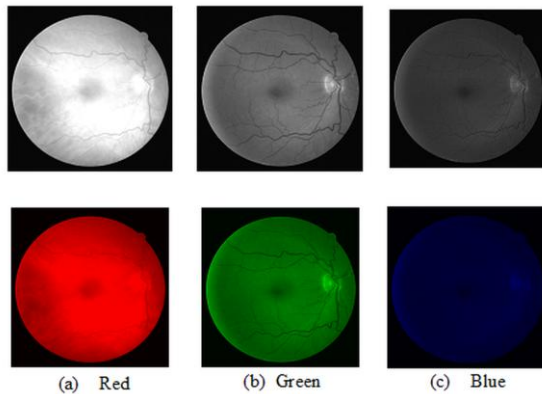


Fig.5. Separated R, G, and B section from RGB image

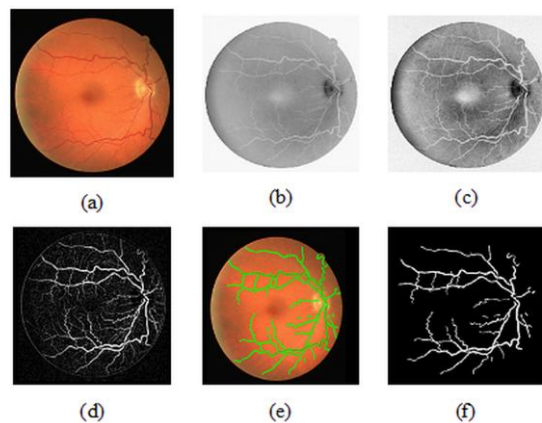


Fig.6. Results attained with proposed CAP scheme

(a) Trial picture, (b) Green-channel picture, (c) Morphological improvement, (d) Enhanced vessel, (e) Traced vessel and (f) Extracted region with SMA-MSMF

Figure 6 depicts the experimental outcome attained with this study. Fig 6(a) shows the trial image, Fig 6(b)-(d) denotes the extracted green-channel picture, morphological enhancement and enhanced image respectively. Fig 6(e) depicts the traces made by SMA-MSMF and Fig 6(f) presents the extracted blood-vessel section. After the extraction, a study with the GT is performed separately for DRIVE and CHASE\_DB1 images and the average values of the performance measures are recorded. The attained values are then compared and confirmed against the similar approaches existing in the earlier research work and these results are presented in Table II and Table III respectively and based on this comparison, the merit of the proposed technique is confirmed. From these table results, it can be noted that the performance values existing in earlier technique are lesser compared to proposed technique.

TABLE II. PERFORMANCE VALUES FOR THE DRIVE DATASET

Method	Accuracy (%)	Sensitivity (%)	Specificity (%)	F1-Score (%)
Sreejini and Govindan [4]	96.33	71.32	98.66	-
Keerthana et al. [5]	87.19	79.16	98.03	-
Tamim et al. [3]	96.07	75.42	98.43	74.75
Khowaja et al. [19]	94.10	74.37	95.96	68.77
Fu et al. [20]	94.12	74.44	96.00	68.84
Proposed	97.15	83.53	98.74	82.28

TABLE III. PERFORMANCE VALUES FOR CHASE\_DB1 DATASET

Method	Accuracy (%)	Sensitivity (%)	Specificity (%)	F1-Score (%)
Tamim et al. [3]	96.32	78.06	98.25	77.17
Khowaja et al. [19]	94.83	71.55	97.46	72.24
Fu et al. [20]	93.36	74.02	94.79	61.87
Sundram et al. [21]	95.01	71.05	96.01	-
Proposed	97.16	82.86	98.61	81.67

These tables confirms that the SMA-MSMF technique helps to achieve a better result on considered image dataset compared to the existing methods, Further, the result available in [26] confirms that, this technique helps to get a better values of the F1-Score for CHASE\_DB1 compared to the Convolutional-Neural-Network (CNN) segmentation results. This confirms that the proposed scheme is very significant in evaluating the funds retinal pictures.

The result of the proposed technique confirms the significance of the proposed technique. In future, the performance of the proposed scheme can be enhanced by employing the Red-Fox-Algorithm proposed in 2021 [27].

#### V. CONCLUSION

Examination of the eye is essential to identify the disease and damage in the eye parts and fundus image supported diagnosis is one of the recommended techniques. In this work, assessment of the blood-vessel is considered for the study and a MSMF is implemented using the SMO algorithm. He proposed technique is an automated methodology and extracts the vessel segment with better accuracy. For the examination, only the green-channel section of the RGB image is considered and the image pre-processing such as, contrast improvement and morphology based vessel enrichment is initially execute. Later, the proposed SMO-MSMF is executed to extract the vessel section of the considered test image. The average result attained for DRIVE and CHASE\_DB1 dataset confirms that, reposed approach helped to extract the vessel with enhanced result compared to the existing related methods. Further, the F1-Score achieved by the proposed technique is superior compared to the CNN segmentation outcome available in the literature. This confirms that, proposed approach is clinically significant.



## REFERENCES

- [1] <https://www.mayoclinic.org/diseases-conditions/retinal-diseases/diagnosis-treatment/drc-20355827>
- [2] D. Kitahar, S.Ananda, and A. Hirabayashi, "Optimization-Based Fundus Image Decomposition for Diagnosis Support of Diabetic Retinopathy," Information Processing Association Annual Summit and Conference (APSIPA ASC) 2019 Asia-Pacific Signal and, pp. 1565-1572, 2019.
- [3] N Tamim, M. Elshrkawey, G.A. Azim, and H. Nassar, "Retinal Blood Vessel Segmentation Using Hybrid Features and Multi-Layer Perceptron Neural Networks," *Symmetry*, vol. 12, no. 6, pp. 894, 2020. <https://doi.org/10.3390/sym12060894>.
- [4] K.S. Sreejini, and V. K. Govindan, "Improved multiscale matched filter for retina vessel segmentation using PSO algorithm," *Egyptian Informatics Journal*, vol. 16, no. 3, pp.253-260, 2015. <https://doi.org/10.1016/j.eij.2015.06.004>
- [5] K. Keerthana, T. J. Jayasuriya, N.S.M. Raja, and V. Rajinikanth. "Retinal vessel extraction based on firefly algorithm guided multi-scale matched filter," *International Journal of Modern Science and Technology*, vol.2, no. 2, pp.74-80, 2017.
- [6] E. Priya, and V. Rajinikanth, *Signal and Image Processing Techniques for the Development of Intelligent Healthcare Systems*. Springer, 2020.
- [7] V. Rajinikanth, E. Priya, H. Lin, and F. Lin, *Hybrid Image Processing Methods for Medical Image Examination*. CRC Press, 2021.
- [8] S. Razeen, F. Ahmed, V. Rajinikanth, P. Tamizharasi, and B.P. Varthini, "Examination of Optic Disc Sections of Fundus Retinal Images—A Study with Rim-One Database," In *Intelligent Data Engineering and Analytics*, pp. 711-719. Springer, Singapore, 2020. [https://doi.org/10.1007/978-981-15-5679-1\\_69](https://doi.org/10.1007/978-981-15-5679-1_69).
- [9] M. F. Armaly, "Genetic determination of cup/disc: ratio of the optic nerve," *Archives of ophthalmology*, vol. 78, no. 1, pp.35-43, 1967. DOI: 10.1001/archoph.1967.00980030037007.
- [10] J.T.A. Rose, S.F. Vinnarasi, Jesline, and V. Rajinikanth, "Assessment of Fundus Images for Retinal Abnormality Screening—A Study," In *Progress in Computing, Analytics and Networking*, pp. 303-312. Springer, Singapore, 2020. [https://doi.org/10.1007/978-981-15-2414-1\\_31](https://doi.org/10.1007/978-981-15-2414-1_31).
- [11] <https://drive.grand-challenge.org/>
- [12] J.J. Staal, M.D. Abramoff, M. Niemeijer, M.A. Viergever, and B. V. Ginneken, "Ridge based vessel segmentation in color images of the retina," *IEEE Transactions on Medical Imaging*, vol. 23, pp. 501-509, 2004. DOI: 10.1109/TMI.2004.825627.
- [13] <https://blogs.kingston.ac.uk/retinal/chasedb1/>
- [14] C.G. Owen et al., "Retinal arteriolar tortuosity and cardiovascular risk factors in a multi-ethnic population study of 10-year-old children; the Child Heart and Health Study in England (CHASE)," *Arteriosclerosis, Thrombosis, and Vascular Biology*, vol. 31, no.8, pp. 1933-1938, 2011. <https://doi.org/10.1161/ATVBAHA.111.225219>.
- [15] N. Kowsalya, A. Kalyani, C. J. Chalcedony, R. Sivakumar, M. Janani, and V. Rajinikanth, "An approach to extract optic-disc from retinal image using K-means clustering," In 2018 Fourth International Conference on Biosignals, Images and Instrumentation (ICBSII), IEEE, pp. 206-212, 2018.
- [16] G. Kavitha and S. Ramakrishnan, "An approach to identify optic disc in human retinal images using ant colony optimization method", *Journal of Medical Systems*, vol. 34, no. 5, pp. 809-813, 2010.
- [17] T.D.V. Shree, K. Revanth, N.S.M. Raja, Raja and V. Rajinikanth, "A hybrid image processing approach to examine abnormality in retinal optic disc", *Procedia Computer Science*, vol. 125, pp. 157-164, 2018.
- [18] J. Amin, M. Sharif, M. Yasmin, H. Ali and S.L. Fernandes, "A method for the detection and classification of diabetic retinopathy using structural predictors of bright lesions", *Journal of Computational Science*, vol. 19, pp. 153-164, 2017.
- [19] S.A. Khowaja, P. Khuwaja, and I.A. Ismaili, "A framework for retinal vessel segmentation from fundus images using hybrid feature set and hierarchical classification," *Signal Image Video Process*, vol.13, pp.379-387, 2019.
- [20] H. Fu, Y. Xu, D.W.K. Wong and J. Liu, "Retinal vessel segmentation via deep learning network and fully-connected conditional random fields," In *Proceedings of the 2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI)*, Prague, Czech Republic, pp. 698-701, 13-16 April 2016. DOI: 10.1109/ISBI.2016.7493362.
- [21] R. Sundaram et al., "Extraction of Blood Vessels in Fundus Images of Retina through Hybrid Segmentation Approach," *Mathematics*, vol. 7, pp.169, 2019. <https://doi.org/10.3390/math7020169>.
- [22] K. R. Remya, and M. N. Giriprasad, "Computer based Technique to Examine Diabetic Retinopathy in Fundus Retinal Images," In 2020 5th International Conference on Computing, Communication and Security (ICCCS), IEEE, pp. 1-8, 2020.
- [23] S. Li, H. Chen, M. Wang, A. A. Heidari, and S. Mirjalili, "Slime mould algorithm: A new method for stochastic optimization," *Future Generation Computer Systems*, vol. 111, pp.300-323, 2020.
- [24] <https://in.mathworks.com/matlabcentral/fileexchange/76619-slime-mould-algorithm-sma-a-method-for-optimization>
- [25] M. Premkumar et al., "MOSMA: Multi-Objective Slime Mould Algorithm Based on Elitist Non-Dominated Sorting," *IEEE Access*, vol.9, pp. 3229 – 3248, 2020. DOI: 10.1109/ACCESS.2020.3047936.
- [26] [https://paperswithcode.com/sota/retinal-vessel-segmentation-on-chase\\_db1](https://paperswithcode.com/sota/retinal-vessel-segmentation-on-chase_db1)
- [27] D. Połap, and M. Woźniak, "Red fox optimization algorithm," *Expert Systems with Applications*, vol. 166 pp.114107, 2021. <https://doi.org/10.1016/j.eswa.2020.114107>.

# SITCOV- A bench mark database for COVID Analysis

\*<sup>1</sup>M.Uma Maheshwari, \*<sup>2</sup>R.Tamilselvi, \*<sup>3</sup>M.Parisa Beham and \*<sup>4</sup>N.Subanandhini

\**Department of ECE, \*Sethu Institute of Technology, Tamilnadu-626 115.*

<sup>1</sup>uma@sethu.ac.in, <sup>2</sup>tamilselvi@sethu.ac.in, <sup>3</sup>parisabeham@sethu.ac.in and subanandhagopal7797@gmail.com

**Abstract--Coronavirus-Covid19 disease causes a respiratory infection in humans. Covid19 was first identified during December 2019 in Wuhan city of china. They are not a new concept for the medical unit around the globe. The virus can spread from person to person among those in close contact. To control the spreading of Covid19 early detection and diagnosis are important factors. It is necessary to have a standard dataset for the researcher for early detection and diagnosis of Covid19. The lack of datasets for Covid19 especially in chest X-rays images and CT images is the main motivation for the research work. In this paper, the main idea is to create a CT Covid19 - images for the research challenges in the identification of Covid19. The proposed dataset is named as "SITCOV" and the number of images collected in this dataset is 73 subjects CT Covid19 images. The features of the proposed datasets are the standard size and equal contrast enhanced images.**

**Keyword: SITCOV, Covid19 dataset, Covid19 CT Images.**

## I. INTRODUCTION

Biomedical imaging is a powerful tool for visualizing the internal organs of the body and its diseases. Image modalities used in biomedical imaging are X-ray, computed tomography (CT), magnetic resonance imaging (MRI) and functional MRI, positron emission tomography (PET) and high-resolution research tomography PET, single photon emission computed tomography (SPECT), and so on [1-5]. Medical imaging has shown significant advancements and many new imaging modalities have been developed in the due course of time. Biomedical imaging techniques have their prominent role both in diagnostic and therapeutic arenas. These techniques had significantly helped to improve the health care of patients. Image guided therapy has drastically reduced the high risk of human errors with improved accuracy in disease detection and surgical procedures.

Coronavirus which is commonly called as Covid-19. The first transmission was observed or the origin of corona virus was in Wuhan, china on December 2019 [6]. They harm the human respiratory system and the initial symptoms of the coronavirus

disease can be developed from 4 to 14 days. There are seven types of coronavirus and which includes: Alpha coronavirus (229E, NL63), Beta corona virus (OC43, HKU1, MERS-COV, SARS-COV) and finally the novel corona virus (SARS-COV2). Severe acute respiratory syndrome coronavirus 2(SARS-COV2) is a latest strain that causes the Covid19 disease. Coronavirus are a large family of viruses that cause illness, ranging from the common cold to more severe disease.

In March 2020, the world health organization (WHO) declared the Covid19 as epidemic [7]. Covid19 is a new course of large number of deaths across the world. As of now, no proper drug or vaccine has been developed for treatment of this disease. However, clinical trials and research is being carried on for the development of vaccine.

Coronavirus 2019(Covid19) was initially transmitted from birds to animals. They belong to the group of viruses that can cause disease in mammals and birds. Now, it spreads from person to person among those in close contact. The virus spreads by respiratory droplets generated from the cough or sneeze of a Covid19 patient. These drops are too heavy to hang in the air, and fall quickly on the floor or surface.

Covid19 symptoms can be very mild to severe. Some people have no symptoms. The most common symptoms are fever, dry cough, and the breathing problem. Apart from this, symptoms like fatigue, sore throat, muscle pain, and loss of taste or smell can also be seen coronavirus patients. Coronavirus is one of the most severe issues that are facing by people around the world. It is necessary that we come out of this situation as early as possible. By 27 Dec 2020, around 79,232,555 confirmed cases and 1,754,493 confirmed deaths were documented by WHO (World Health Organization) [8]. In this paper, a large dataset of CT scan for Covid19 is available. The proposed dataset contain 73 CT Covid19 images of patients.

Many reviews have exposed that CT scans shows the clear radiological findings for COVID-19 patients and promising as a more efficient and easily accessible test method due to the availability of extensive CT devices that can produce results at a fast pace [9,10].

The choice of utilizing CT over other diagnostic modalities is driven by the following characteristics

- CT is highly tissue sensitive, able to capture bones, soft tissues and blood vessels all at the same time.
- Provides more anatomical detail and differentiation.
- Scan data can be manipulated into different views without additional imaging.
- Provide deep and high quality images.
- CT scan focuses on target area better than other imaging modality.

As of now there are two largest sets of public datasets. They are SARS-CoV-2-CT dataset and COVID-CT dataset. Recently, Soares et al. [11] made a set of CT scans database publicly available. This dataset provides 1252 CT scan images which are infected by Covid19 of 60 patients and 1230 CT scan images of 60 patients which are not affected by Covid19 which was taken from hospitals in the city of Sao Paulo, Brazil. Overall it provides 2482 CT scan images of 120 patients. The drawback faced in this database is non-standard size of the images and non-standard contrast of the images.

Another existing database is COVID-CT which is open-sourced dataset. It consists of 349 Covid19 CT images from 216 patients and 463 non-Covid19 CT [12]. The drawback faced in this database is textual information in some images.

Contrast is an important factor in any subjective evaluation of image quality. Contrast enhancement techniques play an important role in image processing to extract information contained in the low dynamic range from the gray level images. To improve the quality of the image, it is necessary to perform operations such as increasing contrast and reducing or eliminating noise [13]. Contrast is determined by the difference in color and brightness of an image with another image. If the images are in low contrast it is difficult to identify the details present in the image. Not all of our images have the right size as needed, so it is important to understand how to resize images properly. Image interpolation

occurs when you resize or distort your image from one pixel grid to another grid. Many algorithms to achieve contrast enhancement and resizing the image have been developed and applied to problems in image processing. There are three parameters to change the size of images. They are types of the image, size of the pixel, and the background resolution. The proposed dataset SITCOV, provide standard size and contrast of images and there will be no textual information in the images and some salient features are included in this dataset. The main contribution of this dataset paper includes,

- A New SITCOV datasets which includes Covis19 images which is used for the detection of COVID and that helps the researcher in analysing and validating their algorithms.
- As contrast Enhancement is a challenge in the existing dataset, a standard contrast enhanced images are created.
- Another challenge in the existing dataset is the textual information that is embedded in the images and Standard size of the Covid19 CT images. So the provided dataset includes images without textual information and standard sized images.

## II. SITCOV DATASET

### A. Construction and Composition Details

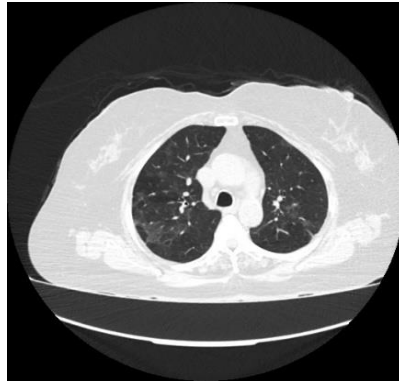
A computed tomography shows more detail than regular X-ray. CT scans use a computer processed combination of many X-ray measurements taken from different angles to produce a tomographic image of a specific area and allow users to see inside objects [14]. It is also referred as Computer axial tomography and computer aided tomography (CAT scan). Patients of all ages can become infected with Covid19 and need to undergo chest imaging. CT imaging has high sensitivity for diagnosis of Covid19. In this paper, 73 Covid19 CT images are used and which is named as SITCOV\_CT. Some sample Covid19 CT images are shown in fig.1.

### B. Labelling the images

In the SITCOV dataset all the images are labelled perfectly for the research work. From the label one can easily identify the subject ID, Age, gender and modality. For example, Covid19 CT image of each subject contain Subject ID, Age, gender and modality are shown in the Figure.2 and Figure.3 respectively.



Covid19\_CT1



Covid19\_CT2



Covid19\_CT3



Covid19\_CT4



Covid19\_CT5



Covid19\_CT6



Covid19\_CT7



Covid19\_CT8



Covid19\_CT9



Covid19\_CT10



Covid19\_CT11



Covid19\_CT12



Fig.1 Sample of Covid19 CT scans images

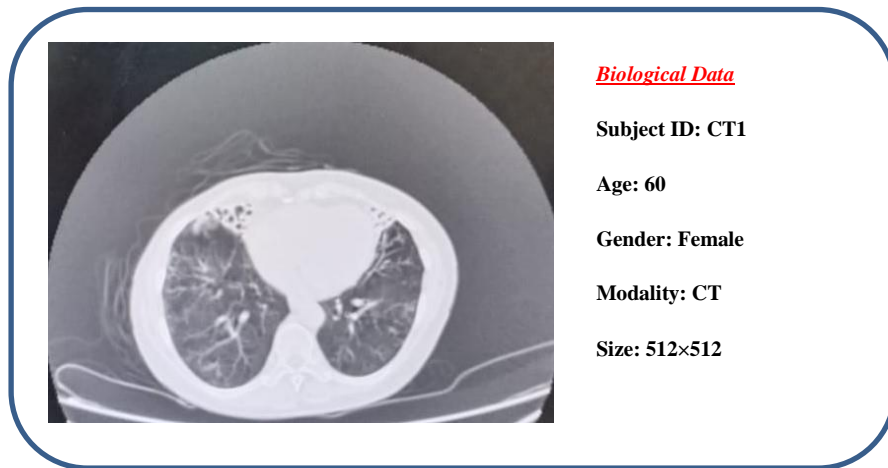


Fig.2 Annotation of Covid19 CT images-sample1

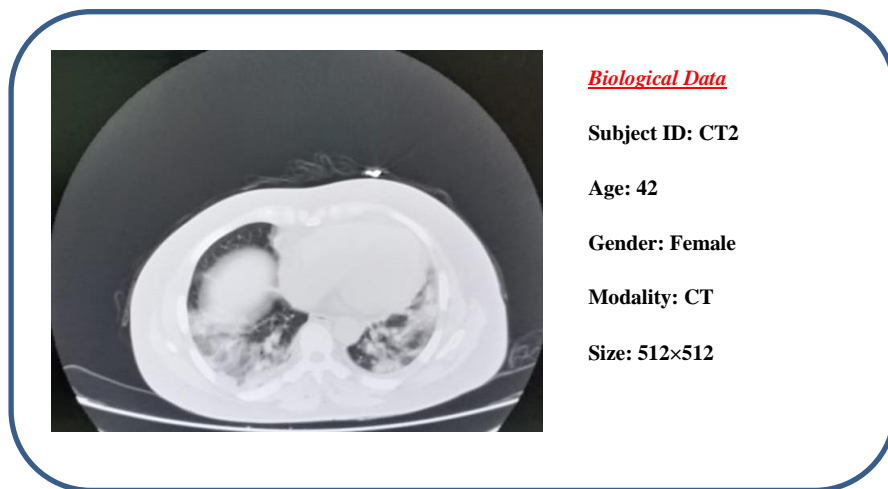


Fig.3 Annotation of Covid19 CT images-sample2

### C. Annotation

SITCOV provides detailed explanations through careful analysis of each CT scan image. From the labelling one can easily identify the detailed description of the subject through following annotation. For example, the label for a Covid19 CT image of the subject is mentioned as Covid19\_CT1.png and Covid19\_CT2.png. The annotated attributes are as follow,

- Subject ID
- Gender
- Age
- Modality
- Size

### III. CONCLUSION

The 2019 novel Coronaviruses (Covid19) are a family of viruses that leads to illnesses ranging from the common cold to more severe diseases and may lead to death according to World Health Organization (WHO). The coronavirus has brought many challenges to academic research. Detection and diagnosis of Covid19 from chest CT images in early stage is a challenging task for the doctor and researcher. The lack of datasets for Covid19 especially in CT images is the main challenges for the scientific study. The dataset used in this research was collected from different radiologist from hospitals and it is available for researchers to download and use it. The insufficiency of benchmark datasets for Covid19 especially in CT images was the main motive of this research. A new medical image dataset called SITCOV which contain a collection of Covid19 CT scan images for research work. The proposed dataset contain 73 CT covid19 images. The dataset can be viewed and downloaded at the following web address: <http://www.sethu.ac.in/SITCOV/>

### ACKNOWLEDGEMENT

The authors thank to Dr. Rajkumar, Radiologist, Government Hospital, Ramnad, India for his help and suggestions in building this database. The authors also immensely thank Dr. Ilayaraja Venkatachalam, Radiologist, Pixel Scans, Trichirappalli, India for providing images with all clinical interpretation to create this database.

### ETHICAL APPROVAL

The "SITCOV" dataset used in this paper is provided by Pixel scans, Trichy. The ethical committee of Pixel scans has reviewed and approved

to conduct research using this SITCOV dataset and publish papers based on the results using those biomedical images.

### REFERENCES

- [1] The Hip Resurfacing Handbook, A Practical Guide to the Use and Management of Modern Hip Resurfacings Woodhead Publishing Series in Biomaterials. 2013; 242-252.
- [2] Brindle K. New approaches for imaging tumour responses to treatment. *Nature Reviews Cancer*. 2008;8(2):94–107.
- [3] Weissleder R, Pittet MJ. Imaging in the era of molecular oncology. *Nature*. 2008; 452(7187):580–9.
- [4] Francisco H. Imai, "Reviewing state-of-art imaging modalities and its potential for biomedical applications", *Journal of Dentistry*, Volume 37, Supplement 1, 2009, ISSN 0300-5712
- [5] Wu, Huaiguang et al. 'Predict Pneumonia with Chest X-ray Images Based on Convolutional Deep Neural Learning Networks'. 1 Jan. 2020 : 2893 – 2907.
- [6] World Health Organization. Pneumonia of unknown cause: China. <https://www.who.int/csr/don/05-january-2020-pneumonia-of-unknown-cause-china/en/>. Published January 5, 2020. Accessed June 6, 2020
- [7] N. Zhu et al., "A Novel Coronavirus from Patients with Pneumonia in China, 2019," *New England Journal of Medicine*, vol. 382, no. 8, pp. 727-733, 2020
- [8] World Health Organization WHO Coronavirus Disease (Covid19) Dashboard. <https://covid19.who.int>
- [9] A. Bernheim, X. Mei, M. Huang, Y. Yang, Z. A. Fayad, N. Zhang, K. Diao, B. Lin, X. Zhu, K. Li, et al., "Chest CT findings in coronavirus disease-19 (covid-19): relationship to duration of infection," *Radiology*, p. 200463, 2020.
- [10] Xuehai He, Xingyi Yang, Shanghang Zhang, Jinyu Zhao, Yichen Zhang, Eric Xing, Pengtao Xie, "Sample-Efficient Deep Learning for COVID-19 Diagnosis Based on CT Scans", medRxiv 2020.04.13.20063941.
- [11] Eduardo Soares, Plamen Angelov, Sarah Biaso, Michele Higa Froes, Daniel Kanda Abe, "SARS-CoV-2 CT-scan dataset: A large dataset of real patients CT scans for SARS-CoV-2 identification", medRxiv, 2020.

- [12] Zhao, Jinyu & Zhang, Yichen & He, Xuehai & Xie, Pengtao. (2020). "COVID-CT-Dataset: A CT Scan Dataset about COVID-19".
- [13] A. Singh, S. Yadav and N. Singh, "Contrast enhancement and brightness preservation using global-local image enhancement techniques," 2016 Fourth International Conference on Parallel, Distributed and Grid Computing (PDGC), Wagnaghat, 2016, pp. 291-294, doi: 10.1109/PDGC.2016.7913162
- [14] Pedro Silva, Eduardo Luz, Guilherme Silva, Gladston Moreira, Rodrigo Silva, Diego Lucio, David Menotti,"COVID-19 detection in CT images with deep learning: A voting-based scheme and cross-datasets analysis",Informatics in Medicine Unlocked,Volume 20,2020,100427,ISSN 2352-9148.

# Positioning the 5-DOF Robotic Arm using Single Stage Deep CNN Model

Sreekar C<sup>1</sup>, Sindhu V S<sup>2</sup>,  
Bhuvaneshwaran S<sup>3</sup>

<sup>1,2,3</sup>UG Student  
Dept. of Electronics Engineering  
Madras Institute of Technology  
Chennai, India

Rubin Bose S\*  
Research Scholar

Dept. of Electronics Engineering  
Madras Institute of Technology  
Chennai, India  
rublins@gmail.com

Sathiesh Kumar V  
Assistant Professor

Dept. of Electronics Engineering  
Madras Institute of Technology  
Chennai, India

**Abstract** — In teleoperation mechanism, the surgical robots are controlled using hand gestures from remote location. The remote location robotic arm control using hand gesture recognition is a challenging computer vision problem. The hand action recognition under complex environment (cluttered background, lighting variation, scale variation etc.) is a difficult and time consuming process. In this paper, a light weight Convolutional Neural Network (CNN) model Single Shot Detector (SSD) Lite MobileNet-V2 is proposed for real-time hand gesture recognition. SSD Lite versions tend to run hand gesture recognition applications on low-power computing devices like Raspberry Pi due to its light weight and timely recognition. The model is deployed using a Camera and two Raspberry Pi Controllers For the hand gesture recognition and data transfer to the cloud server, the Raspberry Pi controller 1 is used. The Raspberry Pi Controller 2 receives the cloud information and controls the Robotic arm operations. The performance of the proposed model is also compared with a SSD Inception-V2 model for the MITI Hand dataset-II (MITI HD-II). The average precision, average recall and F1-score for SSD Lite MobileNet-V2 and SSD Inception-V2 models are analyzed by training and testing the model with the learning rate of 0.0002 using Adam optimizer. SSD MobileNet-V2 model obtained an Average precision of 98.74% and SSD Inception-V2 model as 99.27%, The prediction time for SSD Lite MobileNet-V2 model using Raspberry Pi controller takes only 0.67s whereas, 1.2s for SSD Inception-V2 Model.

**Keywords**— *Single Shot Detector (SSD), SSD Lite, Hand gesture recognition, Amazon Web Services (AWS), Message Queuing Telemetry Transport (MQTT).*

## I. INTRODUCTION

In teleoperation, the doctors position the surgical robots from a remote location by an efficient Human Machine Interaction system. The interaction between the human beings and machines are carried out in more natural forms with the development in the field of Artificial Intelligence. Hand gestures are used to convey a non-verbal type of information's. Gesture-based implementation is associated with a number of fields in many applications, such as Human-Machine Interaction, Virtual Reality, Robot Control, Tele-surgery systems etc. The gesture based control system build a richer bridge between the computers and humans. This method eliminates the keyboard and mouse inputs and to connect directly without any mechanical equipment.

In this paper, a hand gesture recognition system using a single stage deep convolutional neural network (CNN) to control a robotic arm with 5 degrees of freedom (DOF) in a remote location is proposed. The hand gestures are recognized

using a Single Shot Detector (SSD) Lite MobileNet-V2 based CNN model. The MobileNet-V2 CNN model is used as a feature extractor. The information of the recognized gesture is transmitted using a Wi-Fi module and then to the 5 DOF Robotic arm in the remote location. The SSD Lite MobileNet-V2 model predicts the input gestures at a faster rate when compared with the SSD Inception-V2 model.

## II. LITERATURE REVIEW

Saurabh et al. [1] developed a wireless gesture controlled robotic arm. The author captured the hand gestures using a webcam and recognized the gestures using co-relation technique in MATLAB. Robotic arm consists of three DC motors driven by a motor driver and controlled by a PIC 16F877A Microcontroller. The wireless communication is carried out by using a RF module.

Aggarwal et al. [2] proposed a wireless gesture controlled robotic arm with vision. An accelerometer-based device wirelessly (RF Signals) controls a robotic arm. The robotic arm and platform are coordinated with the operator's hand, leg movements and postures. The system is also fitted with an IP camera capable of transmitting video in real time to any Internet-enabled machine.

Sagayama et al. [3] proposed a hand movement recognition model with two deep learning classifiers: the Deep Belief Network (DBN) and CNN. The author analyzed the performance on hand movements of five distinct classes. For each hand gesture, the accuracy rate of DBN, CNN and HOG+SVM are obtained as 98.35%, 94.74% and 89.62%, respectively.

Howard et al. [4] proposed a MobileNet model for embedded vision applications. It comes with a compact architecture in which the network uses depth-wise separable convolutions. The author also introduced two hyper parameters called the width multiplier and resolution multiplier. These parameters help in resizing the model according to the application. The author illustrated the efficiency of MobileNet through a diverse range of applications like object detection, fine grain classification etc.

Othman et al. [5] utilized a SSD-MobileNet model for the detection of objects. The feature extraction is performed using the MobileNet architecture. The authors also reported few enhancements to the model such as default boxes, multi-scale characteristics and depth wise separable convolution layer added to the framework. It resulted in the improvement of efficiency. On the COCO dataset, the authors reported an overall accuracy of 73.00 %.



Wong et al. [6] proposed a deep convolutional SSD architecture (Tiny SSD) specifically developed to recognize objects in an embedded system in real time. A high performance and tailored non-uniform fire subnet module is integrated into the Tiny SSD. It concentrates on the supplementary convolution layers, designed specifically to reduce the effect of model dimension by retaining the detection performance. On the VOC 2007 dataset, the author reported an average accuracy of 61.3 %.

Rubin et al. [7] proposed an Inception-V2 based single stage detector (SSD Inception-V2) model for real-time hand action recognition. The author trained the model using custom-developed hand dataset (MITI-Hand Dataset (MITI-HD)). The model is trained for ten classes. The author reported an Average Precision of 99.00 % and prediction time of 46 ms using a NVIDIA TitanX GPU.

Based on the extensive literature survey it has been identified that the real-time prediction time for the detection of hand gestures could be further reduced. In this paper, a 5 DOF robotic arm is positioned using the vision based real-time hand gestures. The real-time hand gesture recognition is implemented using a single stage CNN model (SSD Lite MobileNet-V2). SSD Lite is a lighter version of the conventional Single Shot Detector (SSD). MobileNet-V2 is a CNN algorithm used for feature extraction. The results of the SSD Lite MobileNet-V2 model is compared with the SSD Inception-V2 model. The computational time of SSD Lite model is comparatively lesser than the existing hand action recognition models.

### III. METHODOLOGY

An overview of vision based robotic arm control using hand gestures recognition is shown in fig.1. The hand gesture recognition is performed using the single stage deep CNN. SSD Lite MobileNet-V2 [8] model and SSD Inception-V2 model [7] are used in this framework for gesture recognition.

MobileNet-V2 and Inception-V2 CNN models are used as feature extractor.

These CNN models are trained and tested using a MITI Hand Dataset-II (MITI HD-II). MITI HD-II is an improved version of MITI HD [7], [9] dataset. It is a custom-created dataset of hand posture obtained from different individuals with differing skin tones, complex profiles, different hand size, conditions of lighting, geometry, fast movements, and different age classes. The dataset has about 10 classes and 970 samples per class. The sample frames of each class from MITI HD-II are shown in fig.2.

The SSD is a CNN based object detector. It uses a single forward pass network and a bounding box regression technique to classify and localize the object. Over Faster R-CNN, SSD offers significant speed gains. SSD is also used in real-time for action recognition. SSD completes the process of region proposal and classification in a single shot process. SSD Lite [6] model is the lighter version of the conventional SSD model. SSD Lite MobileNet-V2 is 20 times more effective and 10 times lighter than YOLO-V2 [10]. It outperforms YOLO-V2 architecture.

MobileNet is a compact architecture that constructs a lightweight deep convolutional neural network using deeply separable convolutions. Depth-separable convolution filters consist of filters for deep convolution and filters for point convolution. On each input channel, the depthwise convolution filter performs a single convolution, and the point convolution filter combines linearly, the 1 to 1 convolutions and the output of depthwise convolution. MobileNets concentrate on latency optimization, but also generate small networks. The depth-wise seperable convolution has separate layers for filtering and combining which drastically contributes in reducing the computation time and model size compared to a standard convolution process. It offers a computation reduction by a factor of  $1/N + 1/D_k^2$ , where N is the number of output channels and  $D_k$  is kernel size. This is 8 to 9 times lesser compared to a standard convolution.

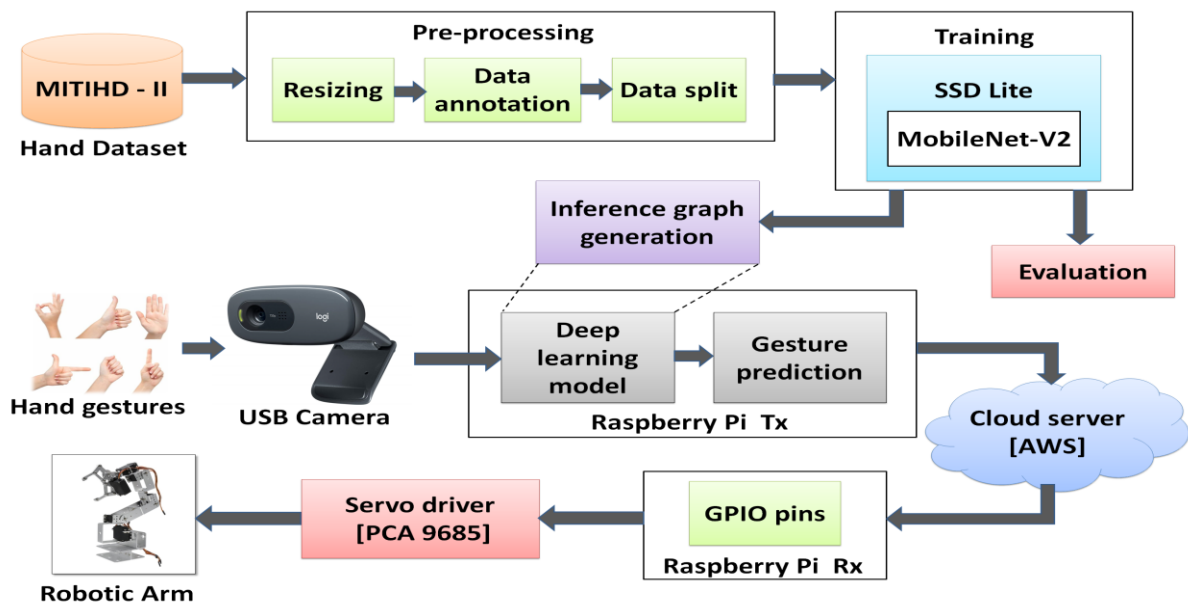


Fig 1. Overall Block Diagram of Vision Based Robotic arm Control using Hand Gestures.



Fig 2. Samples of MITIHD-II dataset.

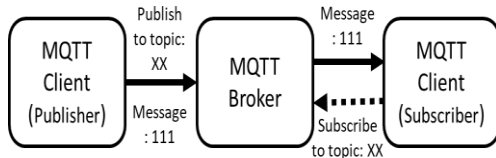


Fig 3. MQTT protocol [10].

Numerous CNN models are tested in object detection problems. Some of them are Region-based convolutional neural network (R-CNN), You Only Look Once (YOLO) [10], and Single Shot Detector (SSD) [7]. In this paper SSD Lite MobileNet-V2 model and SSD Inception-V2 model are used for Hand Gesture Recognition. The performances of the above specified models are evaluated and compared with each other.

Inception-V2 [11] is faster and computationally efficient than other similar architectures like VGG Net and Alex Net. The reduction of dimensionality is carried out effectively in the convolutions of the Inception-V2 model by integrating a smart factorization process. Inception-V2 consists of three layers. The first layer consists of three convolutional layers with dimensionality reduction filter bank. The second layer consists of five layers with deeper filter banks and last layer consists of two wider filter banks. It is then connected to fully connected layers. Out of these two CNN models, SSD Lite MobileNet-V2 model is highly preferred for remotely operated robots because of its light weight. This improves the speed of Hand Gesture Recognition.

The robotic arm has five degrees of freedom and has five MG995 servo motors with a gripper. The robotic arm moves to fixed positions based on the gestures recognized by the SSD model. To control the servo's of the robot, the Adafruit PCA9685 servo driver is utilized. With just 2 pins the Adafruit 16-Channel 12-bit PWM/Servo Driver can drive up to 16 servos over I2C. The robot is interfaced with Raspberry Pi 2 through the PCA9685 driver. The SSD model for Hand gesture recognition is implemented on the Raspberry Pi-1. A USB camera attached to the Raspberry Pi-1 is used to get the input data.

Finally, to integrate with IOT, the Hand Gesture recognition is implemented on Raspberry Pi-1 and the corresponding signals are sent to Raspberry Pi-2, which in turn controls the robotic arm movement. Both the Pi's are connected to different networks. The Raspberry Pi-1 is remotely communicated with Raspberry Pi-2 using Message Queuing Telemetry Transport (MQTT) [12] protocol as shown in fig.3. MQTT is an open source, lightweight and publish- subscribe network that transports messages between devices. The MQTT broker is hosted on an Amazon Web

Services (AWS) server. Table 1 and 2 represents the algorithm used in Raspberry Pi-1 and Raspberry Pi-2 controllers respectively.

TABLE 1 PSEUDOCODE OF GESTURE DETECTION ON RASPBERRY PI-1

Input : Continuous frames of images  $V$  from web camera.  $f$  is the frame at an instant  
Output :  $x1, y1, x2, y2, P_s, P_c$  for each  $f$ , where  $x1, y1, x2, y2$  are bounding box coordinates.  $P_s, P_c$  are predicted score, class.

```

Load CNN Model Into Session ()
WHILE TRUE :
    f = Capture_Frame_From_Webcam ()
    Gesture Recognition using SSD Lite MobileNet-V2
    x1, y1, x2, y2, P_s, P_c = Predict (f)
    IF P_s > threshold:
        Publish P_c to MQTT broker under topic 'Gesture'
    END IF
END WHILE
  
```

TABLE 2 PSEUDOCODE OF ROBOTIC ARM CONTROL ON RASPBERRY PI-2

Input: Predicted class ( $P_c$ )  
Output: Generate corresponding GPIO signals to control Robot.  
Enable to MQTT broker under topic 'Gesture'  
WHILE TRUE :  
IF  $P_c$  is received:  
Perform Desired Robotic arm Action ()  
END IF  
END WHILE

The overall training objective loss function of SSD [7], [13] is calculated based on the weighted sum of the confidence loss  $L_{cnf}$  and the localisation loss  $L_{lcl}$ . It is given by

$$L(a, n, p, t) = \frac{1}{K} (L_{cnf}(a, n) + \rho L_{lcl}(a, p, t)) \dots\dots(1)$$

where  $\rho$  is the weight term and by cross-validation it is set to 1,  $K$  is the number of matched boxes. The loss is said to be 0 if the number of matched boxes default to zero. The term  $a$  is the priors, it is 1 when it matches the ground truth box and 0 otherwise. The term  $n$  is the number of classes,  $p$  is the predicted bounding box parameters and  $t$  is the ground truth bounding box parameter.

#### IV. EXPERIMENTAL SETUP

##### A. Training

The SSD and SSD Lite models are trained utilizing the fine-tuned pretrained weights from COCO dataset. MITI-Hand Gesture dataset is used for training and testing process. The train model is optimized using the gradient decent optimization technique (ADAM optimization [14] [15]). The models are trained for 50000 steps with batchsize-8 and learning rate of 0.0002. The models are trained using the hardware and software setup as described below. Deep Learning toolbox with Python libraries and TensorFlow framework is used. An Intel @ Core TM i7-4790 CPU @ 3.60 GHz, 64-bit processor, 20 GB RAM, Windows 10 PRO operating system and GPU (NVIDIA GeForce GTX TITAN X (PASCAL)) is used to train the architecture. CUDA / CuDNN is used for providing parallel computations in a GPU. The python modules such as Numpy, Matplotlib, Cython, Pandas and Open-CV packages are used in the experimental analysis.

### B. Evaluation and Prediction

The trained models are tested using the test samples and their performances are evaluated by the parameters such as Average Precision, Average Recall, F1 score and Prediction time for various intersection over union (IoU) values. IoU determines the accuracy of hand action detection. IoU is unity when the predicted bounding box exactly overlaps the ground-truth bounding box. The prediction is considered to be true as long as the IoU is 0.5 and for the larger value of IoU's, the prediction is accurate.

### C. Implementation

The robotic arm positioning is implemented by two Raspberry Pi (Model 3B) controllers (Quad Core 1.2GHz Broadcom BCM2837 64bit CPU, 1GB RAM and 40-pin extended GPIO pins) using python framework. The hand gesture recognition is implemented on Raspberry Pi-1 with USB camera. The Raspberry Pi-2 is used to control the robotic arm in remote location. The servo motors of the Robotic arm are controlled by the PWM signals from PCA9685. The Python library adafruit-circuit python-servokit is used to interface Raspberry Pi with PCA 9685. The MQTT broker is hosted on the IOT server of Amazon Web Services (AWS). The Python SDK AWS IoT Python SDK is used to access the AWS server in end devices.

## V. RESULTS AND DISCUSSION

The SSD Lite Mobilenet-V2 and SSD Inception-V2 based hand gesture recognition models are evaluated. The performance of the models are assessed by Average Precision (AP), Average Recall (AR) and F1-Score. The prediction time of the model is calculated during the detection process on a Raspberry Pi processor. Average Precision of different models is calculated for various values of IoU threshold.

The term  $AP_{all}$  is the overall average precision,  $AP_{small}$  is the average precision for the size of the object in an image lesser than  $32 \times 32$  pixels,  $AP_{medium}$  is the average precision for the object size lesser than  $96 \times 96$  pixels and greater than  $32 \times 32$  pixels and  $AP_{large}$  is for the object size greater than  $96 \times 96$  pixels.

Similarly,  $AR_{small}$  average recall for the object is lesser than  $32 \times 32$  pixels,  $AR_{medium}$  and  $AR_{large}$  are the average recall for the object size smaller than  $96 \times 96$  pixels and greater than  $32 \times 32$  pixels, greater than  $96 \times 96$  pixels, respectively.  $AR_1$ ,  $AR_{10}$  and  $AR_{100}$  are the average recall values determined for various number of detections such as 1, 10 and 100.

Both the models are trained for 50,000 epochs with batch size as 8. SSD Lite Mobilenet-V2 and SSD Inception-V2 models are trained with a learning rate of 0.0002. The models are compared on the basis of Average Precision, Average Recall and F1 Score. It may be noted that SSD Inception-V2 performs better on terms of accuracy while SSD Lite MobileNet-V2 has lower prediction time.

Table 3 shows the Average Precision of Single Stage Deep CNN models using MITIHD-II datasets for the various IoU ranges of 0.5, 0.75 and 0.5:0.95. The performance metrics obtained for SSD Inception-V2 is described as  $AP_{0.50}$  is 0.993,  $AP_{0.75}$  is 0.991 and  $AP_{0.5:0.95}$  is 0.884. The metrics of SSD Lite MobileNet-V2 is given as  $AP_{0.50}$  is 0.987,  $AP_{0.75}$  is 0.968 and  $AP_{0.5:0.95}$  is 0.804 respectively. The comparison of these two models shows that the SSD Inception-V2 has higher Average precision values when compared to the SSD Lite MobileNet-V2 model.

The Average Recall of Single Stage Deep CNN models using MITIHD-II datasets for the IoU range of 0.5:0.95 is illustrated in table 4. The Average Recall of SSD Inception-V2 model is 0.911 for  $AR_1$  and 0.914 for both  $AR_{10}$  and  $AR_{100}$ . For SSD Lite MobileNet-V2 model  $AR_1$  is 0.839 and 0.842 for both  $AR_{10}$  and  $AR_{100}$ .

Table 5 demonstrates the comparison of the Performance metrics of Single Stage Deep CNN models using MITIHD-II dataset for the IoU range of 0.5. The  $AP_{0.5}$ ,  $AR_{0.5}$ , and  $F1-Score_{0.5}$  are calculated as 99.27%, 95.58% and 97.39% for SSD Inception-V2 Model and 98.74%, 94.11% and 96.37% for SSD Lite MobileNet-V2 Model. Though the Precision, Recall and F1-score are higher for SSD Inception-V2 model, the prediction time of SSD Lite MobileNet-V2 Model is low (0.67s). This makes a significant contribution in the speed of communication between human and robot.

TABLE 3 AVERAGE PRECISION OF SINGLE STAGE DEEP CNN MODELS USING MITIHD-II FOR VARIOUS IOU RANGES.

CNN Models	Average Precision (AP)					
	0.5	0.75	0.5:0.95	Small <sub>0.5:0.95</sub>	Medium <sub>0.5:0.95</sub>	Large <sub>0.5:0.95</sub>
SSD Inception-V2	<b>0.993</b>	<b>0.991</b>	<b>0.884</b>	<b>0.858</b>	<b>0.877</b>	<b>0.889</b>
SSD Lite MobileNet-V2	0.987	0.968	0.804	0.723	0.742	0.823

TABLE 4 AVERAGE RECALL OF SINGLE STAGE DEEP CNN MODELS USING MITIHD-II FOR IOU RANGES OF 0.5:0.95.

CNN Models	Average Recall (AR)					
	1	10	100	Small <sub>0.5:0.95</sub>	Medium <sub>0.5:0.95</sub>	Large <sub>0.5:0.95</sub>
SSD Inception-V2	<b>0.911</b>	<b>0.914</b>	<b>0.914</b>	<b>0.875</b>	<b>0.898</b>	<b>0.919</b>
SSD Lite MobileNet-V2	0.839	0.842	0.842	0.764	0.791	0.861

TABLE 5 COMPARISON OF PERFORMANCE METRICS SINGLE STAGE DEEP CNN MODELS USING MITIHD-II DATASET.

CNN Models	$AP_{0.50}$	$AR_{0.50}$	$F1-Score_{0.50}$	Prediction time (s)
	(%)	(%)	(%)	Raspberry-Pi 3B
SSD Inception-V2	<b>99.27</b>	<b>95.58</b>	<b>97.39</b>	1.20
SSD Lite MobileNet-V2	98.74	94.11	96.37	<b>0.67</b>

The overall loss curve is plotted with number of training steps in x axis and loss in y axis. The graph is shown in fig.4. The loss is higher at the start and slowly decreases towards the end. It may be noted from the graph that after 50,000 steps SSD Lite MobileNet-V2 has higher loss than SSD Inception-V2 CNN model.

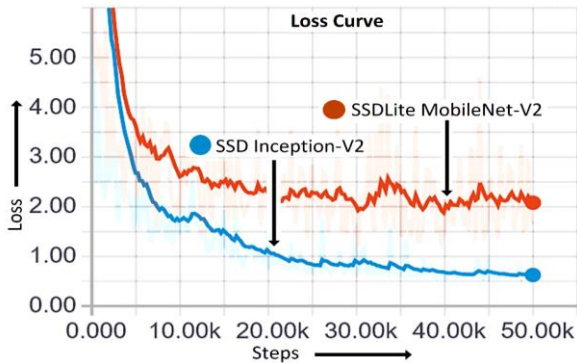


Fig 4. Loss Curve of Single Stage CNN models

The five servo motors that is used in the robotic arm are numbered as 1,2,3,4 and 5. The pulse width range ( $\mu\text{s}$ ) differs for each servo and is determined using trial and error method. For all the DOF's of the robotic arm, the range of pulse width set for the servo motors is shown in table 6.

Based on the gesture detected by the model, the corresponding GPIO signals are generated to move the Robotic arm to a specific position. For different positions, each servo motor in the robotic arm is rotated to a specific angle. The table 7 shows the various angles set for each servo during gesture identification.

TABLE 6 PULSE WIDTH RANGE OF ROBOT SERVO MOTORS.

Servo Motor	Pulse Width ( $\mu\text{s}$ )
1	750-2200
2	650-2150
3	750-2700
4	500-1900
5	800-1670

TABLE 7 ANGLE SETS FOR EACH SERVO FOR VARIOUS GESTURES.

Gesture	Function	Servo 1	Servo 2	Servo 3	Servo 4	Servo 5
One	init	90	90	145	90	180
Two	handshake	180	50	100	180	0
Three	hifi	90	90	90	180	180
Four	akr_gesture	130	180	140	120	180
Five	pick	90	30	110	135	90
Ok	yawn	90	180	130	180	0
Love	handraise	90	90	70	90	0
Tup	lift	180	180	150	80	0
Fold	punch	180	0	70	90	180
Straight	kochi	90	0	100	0	180

The Hardware Implementation setup of Remote Location Robotic Arm Control using Hand Gestures is shown in fig.5. The real-time predicted gestures and the associated robot movements are shown in fig.6. The robotic arm is efficiently positioned using the real-time hand gestures. This setup utilizes the SSD Lite MobileNet-V2 model for hand gesture recognition

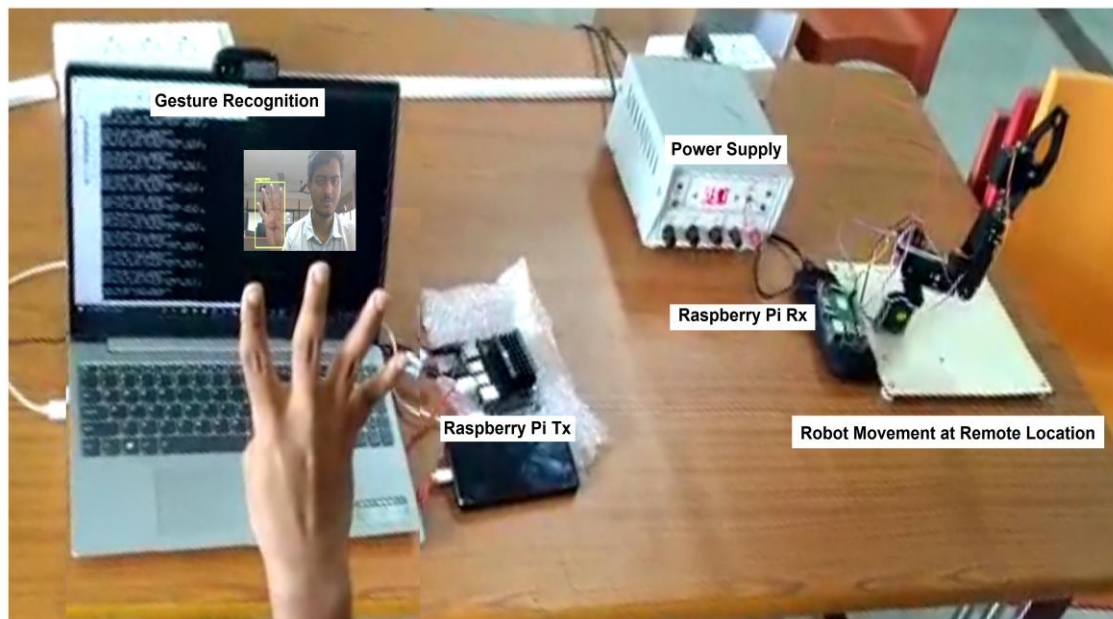


Fig 5. Remote Location Robotic Arm Positioning using Hand Gestures – Hardware Implementation.

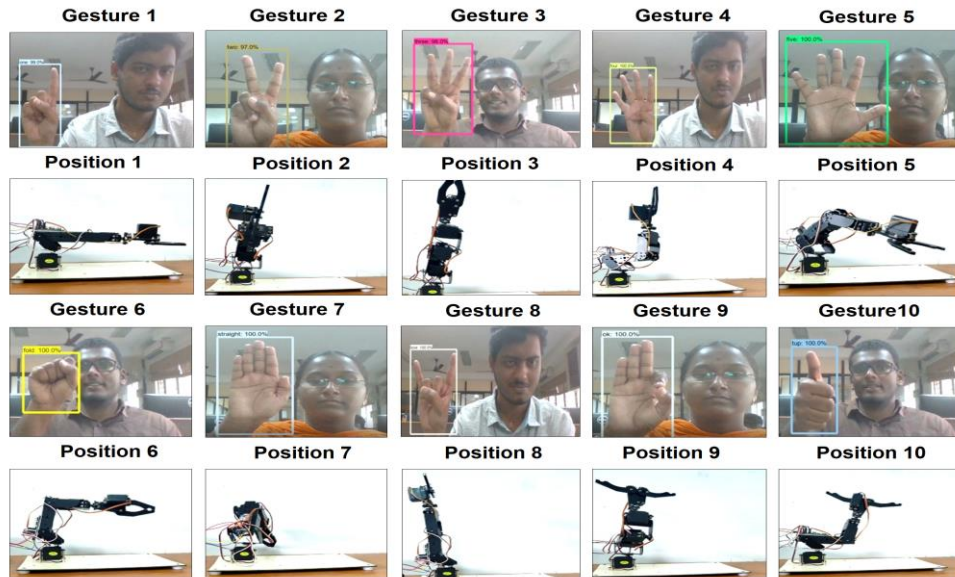


Fig 6. Robotic Arm Position for various gestures recognized.

## VI. CONCLUSION

The remote location robotic arm control is performed using the single stage deep CNN hand gesture recognition model. The SSD lite MobileNet-V2 and SSD Inception V2 models are trained and tested using MITI HD-II dataset. SSD Lite model is preferred to run the hand gesture recognition applications on low power devices such as a Raspberry Pi due to its light weight and speedy recognition. Raspberry Pi controller 1 is used for hand gesture recognition and to transmit the information to cloud. Raspberry Pi controller 2 receives the information from the cloud and control the operations of Robotic arm. Both the Raspberry Pi's are remotely communicated using MQTT Protocol and the MQTT broker is hosted on an AWS server. For IoU threshold of 0.5, the Average precision, Average recall and F1 score of SSD Inception-V2 model is calculated as 99.27%, 95.58%, and 97.39% and the corresponding values for SSD MobileNet-V2 model is 98.74%, 94.11% and 96.37% respectively. The prediction time for SSD Inception-V2 Model using Raspberry Pi controller is 1.2s whereas SSD Lite MobileNet-V2 model consumes 0.67s. The MobileNet model runs faster at the cost of lower precision. Hence to run on low power edge devices, the SSD Lite Mobilenet-V2 model is preferred. Further reduction in prediction time will improve the speed of communication between humans and robots. This prototype can be extended into a Tele-surgery system which is considered as our future work.

## ACKNOWLEDGEMENT

The authors wish to thank NVIDIA for supplying the GPU under University Research Grant Initiative (Model No: NVIDIA TitanX).

## REFERENCES

- [1] Saurabh A. Khajone, S. W. Mohod and V.M Harné, "Implementation of wireless gesture controlled robotic arm", International Journal of Innovative Research in Computer and Communication Engineering, vol. 3, no. 1, 2015.
- [2] Aggarwal, L.; Gaur, V.; Verma, P. Design and implementation of a wireless gesture controlled robotic arm with vision. International Journal of Computer Applications. 2013, 79, 39–43.
- [3] Sagayama KM, Viyasb TV, Hoa CC, Heneseyb LE (2017) Virtual robotic arm control with hand gesture recognition and deep learning strategies. Deep Learn Image Process Appl 31:50.
- [4] Howard, A., G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Marco Andreetto, M., and Adam, H., "MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications", Google Inc., 2017
- [5] Othman N. A. and I. Aydin, "A new deep learning application based on movidius ncs for embedded object detection and recognition", 2018 2nd International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT), pp. 1-5, Oct 2018
- [6] Wong, A., Shafiee, M.J., Li, F., Chwyl, B.: Tiny ssd: A tiny single-shot detection deep convolutional neural network for real-time embedded object detection. arXiv preprint arXiv:1802.06488 (2018)
- [7] Rubin Bose. S, Sathiesh Kumar. V. "Efficient Inception V2 based Deep Convolutional Neural Network for Real-Time Hand Action Recognition". IET Image Processing, Vol. 14, no. 4, March 2020, pp. 688 – 696
- [8] Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., and Chen, L.-C. Mobilenetv2: Inverted residuals and linear bottlenecks. CVPR, 2018, doi: 10.1109/CVPR.2018.00474
- [9] Rubin Bose. S, Sathiesh Kumar. V. "Hand Gesture Recognition Using Faster R-CNN Inception V2 Model". AIR 2019: Proceedings of the Advances in Robotics 2019, ACM digital library, July 2019, Article No.: 19, Pages 1–6.
- [10] Redmon, J., Farhadi, A.: Yolo9000: better, faster, stronger. arXiv preprint, vol. 1612 (2016)
- [11] Ning, C., Zhou, H., Song, Y., Tang, J.: Inception single shot multibox detector for object detection. In: ICME (2017)
- [12] D. Kang et al., "Room Temperature Control and Fire Alarm/Suppression IoT Service Using MQTT on AWS," 2017 International Conference on Platform Technology and Service (PlatCon), Busan, Korea (South), 2017, pp. 1-5, doi: 10.1109/PlatCon.2017.7883724.
- [13] S. Chen, J. Hong, T. Zhang, J. Li and Y. Guan, "Object Detection Using Deep Learning: Single Shot Detector with a Refined Feature-fusion Structure", In 2019 IEEE International Conference on Real-time Computing and Robotics (RCAR), pp. 219-224, 2019, August.
- [14] Kingma, D., & Ba, J. Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980, 2014.
- [15] Ruder, S.: An overview of gradient descent optimization algorithms. arXivpreprint arXiv:1609.04747, <https://arxiv.org/abs/1609.04747> (2017)

# Malnutrition Detection using Convolutional Neural Network

Arun Raj Lakshminarayanan

Department of Computer Science and Engineering  
B.S.Abdur Rahman Crescent Institute of Science and Technology,  
Chennai, India  
[arunraj@crecident.education](mailto:arunraj@crecident.education)

Rajeswari V

Department of Computer Science and Engineering,  
B. S. Abdur Rahman Crescent Institute of Science and Technology,  
Chennai, India.  
[Rajeswari06.vaddi@gmail.com](mailto:Rajeswari06.vaddi@gmail.com)

A.Abdul Azeez Khan

Department of Computer Applications,  
B. S. Abdur Rahman Crescent Institute of Science and Technology,  
Chennai, India.  
[abdulazeekhan@crecident.education](mailto:abdulazeekhan@crecident.education)

Pavani B

Department of Computer Science and Engineering,  
B. S. Abdur Rahman Crescent Institute of Science and Technology,  
Chennai, India.  
[bollinenipavani000@gmail.com](mailto:bollinenipavani000@gmail.com)

Saravanan Parthasarathy

Department of Computer Science and Engineering,  
B. S. Abdur Rahman Crescent Institute of Science and Technology,  
Chennai, India.  
[saravanan\\_cse\\_2019@crecident.education](mailto:saravanan_cse_2019@crecident.education)

K.Javubar Sathick

Department of Computer Applications,  
B. S. Abdur Rahman Crescent Institute of Science and Technology,  
Chennai, India.  
[javubar@crecident.education](mailto:javubar@crecident.education)

**Abstract** — Malnutrition is directly or indirectly responsible for the deaths of children younger than 5 years in many countries. Identification of malnourished children will help to prevent the risk of death and can reduce physical and health issues by taking necessary measures or treatment. The proposed system uses a Convolutional Neural Network (CNN), a Deep Learning algorithm that takes input, analyzes the images, and differentiates one from the other. The architecture we used here is AlexNet for the training process and Transfer Learning. The system takes the image of a child as the input and classifies the image into a malnourished or normal child by comparing the image with the trained model. The objective of the system is to detect malnutrition in children that can help people and healthcare providers to reduce the effects caused by malnutrition by automation implementation instead of a manual process.

**Keywords** — Malnutrition, Convolutional Neural Network, Alexnet.

## I. INTRODUCTION

Malnutrition is a condition that occurs due to fewer intakes or over intake of nutrients. This can lead to health issues such as diabetes, heart disease, eye problems, and stunted growth. Malnutrition is directly or indirectly responsible for the deaths of children younger than 5 years in many countries. According to the World Health Organization's (WHO) 2020 edition, stunting has affected 21.3% or 144 million children under 5 years of age globally. Wasting has affected or threatened the lives of 6.9% or 47 million children under the same age group. About 5.6% or 38.3 million children under the age of 5 years were overweight around the world [1]. Yet, while there has been an improvement, it has been slow and patchy as shown in the Fig 1. Undernutrition leads to physical health issues and growth issues. According to UNICEF's The State of the World's Children 2019 report [2], globally out of 3 children at least 1 is not growing well due to malnutrition, and 1 out of 2 children is suffering from hidden hunger.

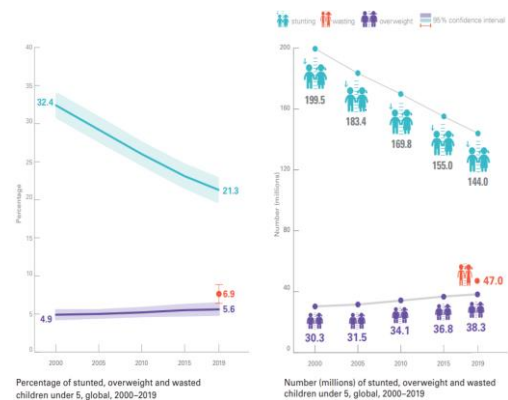


Fig 1: Percentage and Number (millions) of stunted, wasted and overweight under age 5 globally between the years 2000-2019. Source: UNICEF, WHO, World Bank Group joint malnutrition estimates, 2020 edition

Approximately 45% of deaths of children under the age of five years are due to undernutrition [3]. The National Family Health Survey 2005-06 (NFHS-3) reported that in India, 48.0% of children under age 5 years were stunted, 19.8% were wasted, 6.4% were severely wasted and 42.5% were underweight[4]. The National Family Health Survey 2015-16 (NFHS-4) reported that 38.4% were stunted, 21.0% were wasted, 7.5% were severely wasted and 35.8% were underweight [5]. Over the decade there was a slight decline in stunted and underweight children but wasted remains alarming. Children were suffered more with malnutrition and the detection of malnutrition will help to prevent the risk of death and can reduce physical or growth issues by taking necessary measures.

Detection of malnutrition in children can help people and healthcare providers to take preventive measures and can reduce the effect caused by malnutrition on children. To detect the malnutrition in children, a type of artificial neural network

called convolutional neural network (CNN) is used. The architecture we used here is AlexNet, which is a convolutional neural network that is a powerful model capable of achieving high accuracies on challenging datasets.

## II. RELATED WORK

A disease detection concept has detailed by using the human images and analyzing the data from the color of the image. This might not work in every case as some people might look lean and are not malnourished [6]. Tensor Flow algorithm is used to train the dataset which supports various features. The data obtained from semi structured interviews were analyzed using frame analysis and MaxQDA software. The gathering of data required much amount of time and improper data might lead to incorrect results [7]. The nutritional status of the population is characterized using height, weight, and MUAC based anthropometric indicators. The study also revealed that the removal of data based on SMART flag cut-off points improves the data quality of anthropometric surveys. The quality improvement measures are described rather than identifying the children affected with malnutrition [8].

The white paper [9] detailed the role of CNS in the prevention and treatment of malnourished patients who were hospitalized. The accurate documentation of the malnourished patients will help in the appropriate coding, funding reimbursement, and treatment [10]. A decompositional analysis was conducted using two waves (2004-2005 and 2011-2012) data from Indian Human Development. The z-scores of Weight-for-Height (WHZ), Height-for-Age (HAZ), Weight-for-Age (WAZ), and the composite index of Anthropometric Failure were used for undernutrition measures. The analysis provided the up-to-date profile of the nutritional status of India's children using the z-scores. [11].

A system was developed to show the dashboard representation of malnutrition. The final output is the graphical representation of the data in the form of a pie chart or bar graph which contains the data that is maintained in the database and is displayed based on the selected. It represents the data in the form of pie chart or bar graph which might not help to treat the malnourished children but can help us to take preventive measures to decrease the rate of malnourished children [12]. Malnutrition among under-five children is one of the major concerns in India. This review has identified the determinants and strategies that are required to prevent malnutrition under age of five years children in India. Prevention measures are explained and [13]. A system was developed using ID3, Random forest tree algorithm to generate a general result about the nutritional status of the children. Random forest tree algorithm is beneficial to train the data. However, it is comparatively slow to create predictions out of trained data [14].

In 2011 a complete nutrition data about children under five years were collected. Based on this data, a predictive model is developed using PART pruned rule induction which helps program managers and government to identify the children who are at risk. [15]. A Rule-based classification technique with a Multi-Agent System was used to detect malnutrition in

children. The final decision is made based on the number of rules used and shows that there is connectivity between the number of rules and the optimality of final decision. It requires lot of manual intervention like generating rules out of facts and also depletes time while we generate rules for complex data or system [16]. The study says that weight-height based case-detection technique performs worse in identifying malnutrition. MUAC is the best case-detection method and helps to identify severely malnourished children for their admission into community based therapeutic care programs. MUAC is the best method to measure each and every child to know about their nutritional status [17]. The study resulted that 21.7% of children fewer than three years of being affected with undernutrition in rural Western China in 2005 and stated that childhood malnutrition is a large health challenge. This work depicts the overall rate of malnourished children under three years but not techniques to identify malnourished children [18].

## III. SCOPE OF THE STUDY

Identifying malnourished children less than five years age is very important as it causes more problems in children than any other age group as they may lead to growth (both physical and mental) impediment and vulnerability to rehashed diseases. The main aim of the project is to detect the children affected with malnutrition with the help of their images and simple parameters (gender, age, weight and height).

## IV. METHODOLOGY

### A. Dataset

The dataset contains malnutrition and normal children images in different folders labeled as Malnutrition and Normal without any filters or size restrictions.

### B. Graphical User Interface (GUI)

A GUI is created in Matlab for taking inputs from the user as shown in Fig 2, for performing operations and for displaying the results.

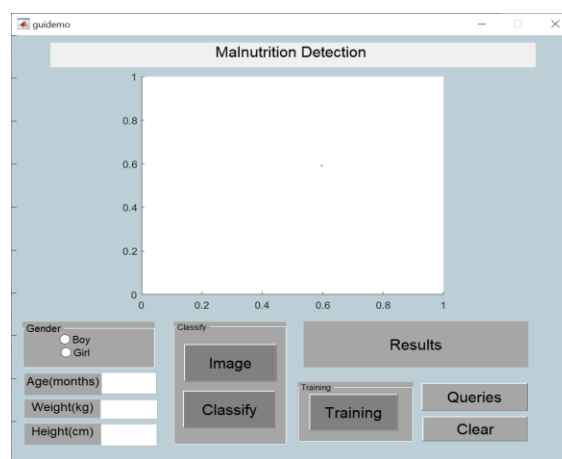


Fig 2: Graphical User Interface for Malnutrition Detection

Input: 1. Child image, 2. Gender, 3. Age (in months), 4. Weight (in kg), 5. Height (in cm).

Processing: Training using Alexnet and transfer learning. Classification using image result and parametric conditions of weight-for-age, weight-for-height, and height-for-age.

Output: Displays the result in string form in Results textbox. Queries button displays frequently asked questions on malnutrition.

### C. Training and transfer learning

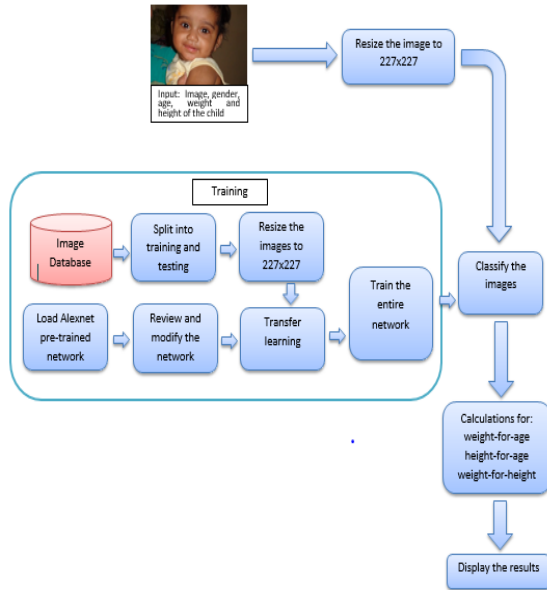


Fig 3: Architecture for Malnutrition Detection

- **Load the dataset:** The dataset consists of malnutrition and normal/healthy children. The very first step in this module is to load the data folders and subfolders.
- **Split the dataset:** The dataset splits into training and testing images randomly and creates two data stores for training and testing data.
- **Load and modify pre-trained network:** AlexNet is a convolutional neural network which is trained on more than one million images from the ImageNet database and can classify images into 1000 object categories, such as a keyboard, mouse, pencil, and many animals. As a result, the model has learned rich feature representations for a wide range of images. AlexNet architecture consists of eight layers. They are five convolutional layers, some of them are followed by max-pooling layers and three fully connected layers. Using a non-saturating Rectified Linear Unit (ReLU) activation function improves training performance. It is a powerful model capable of achieving high accuracies on challenging datasets.

This step loads the pre-trained Alexnet neural network. After loading, an appropriate learning rate, epochs, and mini-batch values are given to the Alexnet architecture and final layers are modified (Fig 4) according to the new dataset categories that are two (malnutrition and normal).

1	'data'	Image Input	227x227x3 images with 'zerocenter' normalization
2	'conv1'	Convolution	96 11x11x3 convolutions with stride [4 4] and padding [0 0 0 0]
3	'relu1'	ReLU	ReLU
4	'norm1'	Cross Channel Normalization	cross channel normalization with 5 channels per element
5	'pool1'	Max Pooling	3x3 max pooling with stride [2 2] and padding [0 0 0 0]
6	'conv2'	Convolution	256 5x5x48 convolutions with stride [1 1] and padding [2 2 2 2]
7	'relu2'	ReLU	ReLU
8	'norm2'	Cross Channel Normalization	cross channel normalization with 5 channels per element
9	'pool2'	Max Pooling	3x3 max pooling with stride [2 2] and padding [0 0 0 0]
10	'conv3'	Convolution	384 3x3x256 convolutions with stride [1 1] and padding [1 1 1 1]
11	'relu3'	ReLU	ReLU
12	'conv4'	Convolution	384 3x3x192 convolutions with stride [1 1] and padding [1 1 1 1]
13	'relu4'	ReLU	ReLU
14	'conv5'	Convolution	256 3x3x192 convolutions with stride [1 1] and padding [1 1 1 1]
15	'relu5'	ReLU	ReLU
16	'pool5'	Max Pooling	3x3 max pooling with stride [2 2] and padding [0 0 0 0]
17	'fc6'	Fully Connected	4096 fully connected layer
18	'relu6'	ReLU	ReLU
19	'drop6'	Dropout	50% dropout
20	'fc7'	Fully Connected	4096 fully connected layer
21	'relu7'	ReLU	ReLU
22	'drop7'	Dropout	50% dropout
23	''	Fully Connected	2 fully connected layer
24	'prob'	Softmax	softmax
25	''	Classification Output	crossentropyex

Fig 4: Modified Alexnet network

- **Transfer learning:** To perform the classification on a new set of images, the pertained network is used as a starting point to learn the task then the final layers are replaced by the new small set of images. Fine-tuning a network with transfer learning is generally a lot quicker and simpler than training a network with randomly initialized weights from scratch. This step performs Transfer learning with the modified pre-trained network and mentioned parameters. Then, it trains the entire network to classify the images, here the images for training requires with the dimension of  $227 \times 227$ , a function is used so all the images get resized to  $227 \times 227$  for training by using resize function.

### D. Classification

This module classifies the input image into either malnutrition or normal. It loads the pre-trained Alexnet model, and then it reads the input image and resizes the image to  $227 \times 227$  as the pre-trained model requires the image size to be the same as the input size of the network and label the image. According to WHO, children who fall under -2 standard deviation (SD) of the WHO Child Growth Standards median for z-scores of weight-for-age, height-for-age, and weight-for-height indicate that they suffer from malnutrition. With the help of these standards and user input for age, weight, and height, the classifier classifies the children into three categories: 1. malnutrition, 2. risk of malnutrition, and 3. Normal.

## V. EXPERIMENTAL RESULTS

The proposed system used 500 images of children under five years, 250 images of malnourished children and 250 images of healthy/normal children. 90% of data is used for training and 10% for testing. The input image for classification and all the images in the dataset resized to the dimensions of  $227 \times 227 \times 3$  for training and testing process.



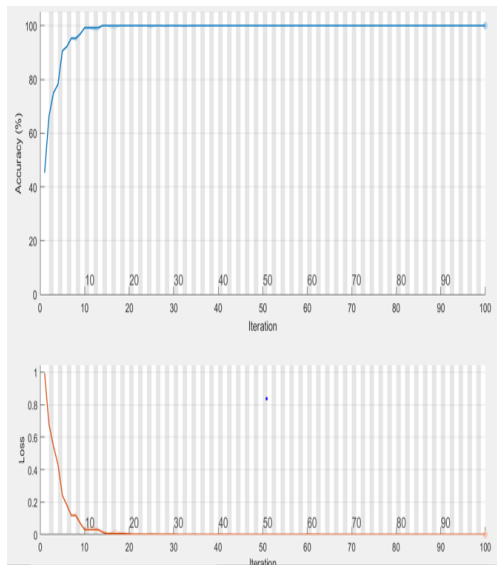


Fig 5: Training progress of the network

The training progress (Fig 5) involves transferred layers and the new layers of the network.

Accuracy is used to find out the performance. It is equal to number of correct predictions. Accuracy changes for every iteration. The accuracy value at each iteration is plotted in graph. The structure of graph shows the increase in accuracy value which means the model learns about the model and loss is reduced. Lower the loss higher the accuracy of the model. If prediction of model is accurate and predicts as expected the loss is lesser or zero. If the prediction of model is not accurate or not as expected, loss value is higher.

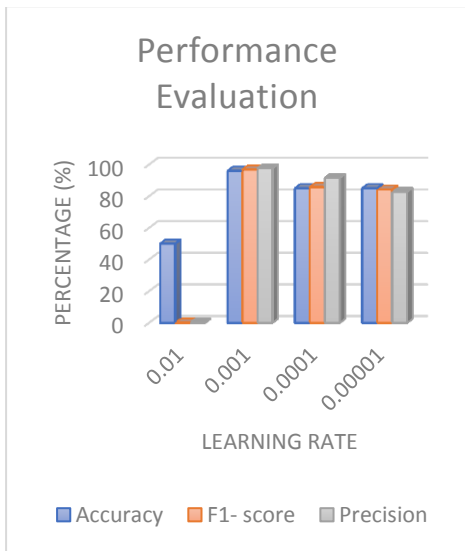


Fig 6: Performance evaluation for different learning rates

After successful training, we tested the proposed model and it is observed that the model achieved 96% of accuracy for the learning rate of 0.001 as shown in Fig 6. The final results were classified into 3 categories based upon image

classification and z-scores of WAZ, WHZ, HAZ: (i) Malnutrition - if the image and z-scores classifications were predicted as malnutrition, the final result shows that the child is malnourished as shown in Fig 7(a), (ii) Risk of nutrition - if either the image or z-scores classification is predicted as malnutrition, the final results shows that the child is at the risk of malnutrition as shown in Fig 7(b), (iii) Normal – if both the image and z-scores classifications were predicted as normal, the final result shows that the child is not malnourished or normal as shown in Fig 7(c).

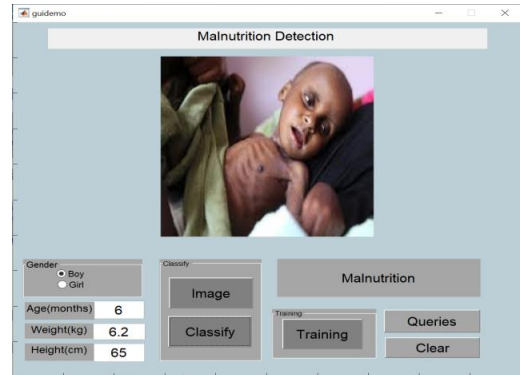


Fig 7 (a)

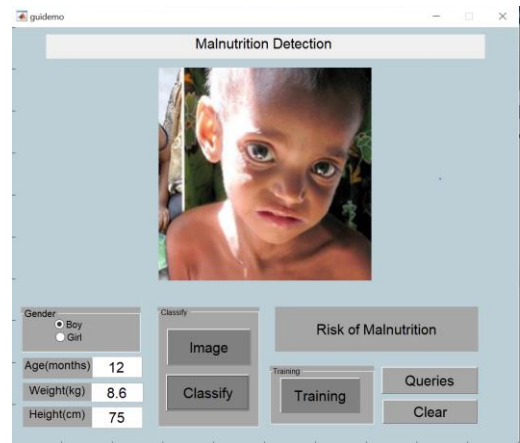


Fig 7 (b)

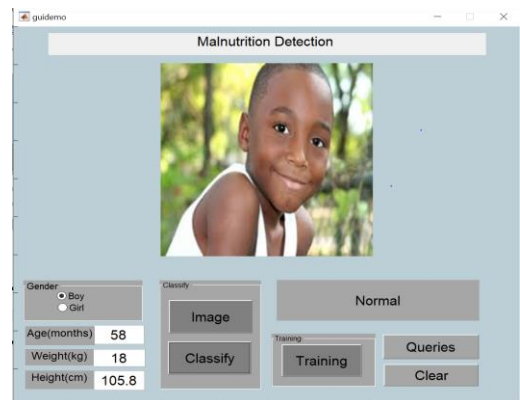


Fig 7 (c)

Fig 7: The child is detected as (a) malnutrition (b) risk of malnutrition and (c) normal using image and simple inputs

## VI. CONCLUSION AND FUTURE WORK

Malnutrition is extravagantly high and has affected many countries in the world by one or more forms. Detecting or predicting malnutrition will help the government or health services to take preventive measures. Conventional Neural Network (CNN or ConvNet) algorithm is used to detect the children affected with malnutrition under age five. Images of children are used as input. Alexnet is a CNN used to find patterns in images to recognize faces and objects and performs classification tasks. By using parametric conditions, Alexnet architecture, and with the help of the extracted features, the system predicts whether children are affected by malnutrition or not. It also categorizes the children who are at the risk of malnutrition. In the next phase of study, we would apply the CNN algorithm to detect the type of malnutrition which affected the children. As the treatment differs for each type, it would be helpful to the parents and healthcare providers.

### REFERENCES

- [1] 'Levels and trends in child malnutrition', UNICEF/WHO/World Bank Group – Joint Child Malnutrition Estimates 2019 edition. <https://www.who.int/nutgrowthdb/jme-2019-key-findings.pdf>
- [2] The State of the World's Children 2019: Children, Food and Nutrition, UNICEF. <https://data.unicef.org/resources/state-of-the-worlds-children-2019/>
- [3] Malnutrition is a world health crisis, [2019], WHO. <https://www.who.int/nutrition/topics/world-food-day-2019-malnutrition-world-health-crisis/en/>
- [4] Nutrition in India, National Family Health Survey (NFHS-3) India 2005-06. <https://dhsprogram.com/pubs/pdf/frind3/frind3-vol1andvol2.pdf>
- [5] India Fact Sheet, The National Family Health Survey (NFHS-4) 2015-16. <https://dhsprogram.com/pubs/pdf/FR339/FR339.pdf>
- [6] Kadam, N., Dabhade, V., Baravkar, R., Saravade, V., & Mankar, C. (2019). Detect Malnutrition in Underage Children by using TensorFlow Algorithm of Artificial Intelligence. (Vol. 06, Issue 12)
- [7] Mohseni, M., Aryankhesal, A., & Kalantari, N. (2019). Prevention of malnutrition among children under 5 years old in Iran: A policy analysis. *PLoS one*, 14(3), e0213136.
- [8] Grellety, E., & Golden, M. H. (2018). Change in quality of malnutrition surveys between 1986 and 2015. *Emerging themes in epidemiology*, 15(1), 1-13.
- [9] Susan Smith, Tracy B. Chamblee, Karie Falder-Saeed, Kelly Haight, Melanie Horbal Shuster, Mary J. Waldo and Hea Lan Yoon, [January 2017], 'MALNUTRITION IN HOSPITALIZED ADULT PATIENTS', National Association of Clinical Nurse Specialists(NACNS). <https://nacns.org/wp-content/uploads/2017/01/Malnutrition-Report.pdf>
- [10] Kellett, J., Kyle, G., Itsiopoulos, C., Naunton, M., & Luff, N. (2016). Malnutrition: the importance of identification, documentation, and coding in the acute care setting. *Journal of nutrition and metabolism*, 2016. (Vol. 2016, Article ID 9026098) <https://doi.org/10.1155/2016/9026098>
- [11] Nie, P., Rammohan, A., Gwozdz, W., & Sousa-Poza, A. (2016). Developments in Undernutrition in Indian Children Under Five: A Decompositional Analysis. <http://ftp.iza.org/dp9893.pdf>
- [12] Madhuri Arya, Pooja Chavhan and Ujjwala Chaudhry, [2015], 'Malnutrition Detection and Management System', *International Journal of Computer Science and Information Technology Research* (Vol. 3, Issue 1, pp: (364-368))
- [13] Sahu, S. K., Kumar, S. G., Bhat, B. V., Premarajan, K. C., Sarkar, S., Roy, G., & Joseph, N. (2015). Malnutrition among under-five children in India and strategies for control. *Journal of natural science, biology, and medicine*, 6(1), 18.
- [14] Markos, Z., Doyore, F., Yifiru, M., & Haidar, J. (2014). Predicting Under nutrition status of under-five children using data mining techniques: The Case of 2011 Ethiopian Demographic and Health Survey. *J Health Med Inform*, 5, 152.
- [15] Markos, Z., Doyore, F., Yifiru, M., & Haidar, J. (2014). Predicting Under nutrition status of under-five children using data mining techniques: The Case of 2011 Ethiopian Demographic and Health Survey. *J Health Med Inform*, 5, 152.
- [16] Dezhi, X., & Ganegoda, G. U. (2011). Rule based classification to detect malnutrition in children. *International Journal on Computer Science and Engineering (IJCSCE)* Vol. 3.

# A Multiclass Skin Lesion classification approach using Transfer learning based convolutional Neural Network

Cauvery K  
Department of Computer Science and  
Engineering,  
Manipal Institute of Technology,  
MAHE, Manipal  
cauk.84@gmail.com

P C Siddalingaswamy \*  
Department of Computer Science and  
Engineering,  
Manipal Institute of Technology,  
MAHE, Manipal  
\*pcs.swamy@manipal.edu

Sameena Pathan  
Department of Computer Science and  
Engineering,  
Manipal Institute of Technology,  
MAHE, Manipal  
sameena.pathan.k@gmail.com

Noel D'souza  
Department of Electrical & Electronics  
Engineering,  
Manipal Institute of Technology,  
MAHE, Manipal  
dsouzanoel2197@gmail.com

**Abstract**— The rapid rise in skin diseases over the past decade has been a growing concern worldwide. Early detection, correct categorization, and accurate identification can result in the successful treatment of melanoma, thereby decreasing the morbidity and mortality rate. Thus, there is a significant need for a system that is capable of identifying skin diseases and precisely classifying them. The proposed work aims to develop a multi class classification system using transfer learning-based convolutional neural networks (CNN). In particular, the proposed solution classifies the dermoscopic images to 8 different categories namely Melanoma (MEL), Basal Cell Carcinoma (BCC), Actinic Keratosis (AK), Benign Keratosis (BKL), Dermatofibroma (DF), Vascular lesions (VASC) and Squamous Cell Carcinoma (SCC). Four state-of-art pre-trained models are used for this task. A functional model-based network is leveraged to embed these sub-models in a larger multi-headed neural network. This will allow the embedded model to be treated as a single large model. An ensemble approach, termed as blending, is employed to combine the predictions efficiently made by the sub-models. Additionally, a robust cropping strategy is implemented to deal with the uncropped images and their impact on the performance of the classifiers is investigated. The impact of applying blending technique to ensemble the pre-trained CNNs are investigated against the performance of the individual classifier. The proposed work is carried out on International Skin Imaging Collaboration (ISIC) 2019 dataset. In this work, the solution for task 1 of the challenge is presented and we obtained balanced multi-class accuracy of 81.2% on the dataset compiled from the original dataset.

**Keywords**— *Balanced Multi-class Accuracy (BMA), Blending, Convolutional Neural Network, Dermoscopic images, Functional model, Skin lesions*

## I. INTRODUCTION

Skin lesions can be referred to as any abnormal growth or appearance on the skin when compared to the skin around it. It can be categorized into many different classes and subclasses [1]. Most of these skin lesions are benign though some, such as certain moles and actinic keratosis, can be pre-cursor to skin cancer. Skin cancers – including melanoma and non-melanoma comprising of Basal Cell Carcinoma (BCC) and Squamous Cell Carcinoma (SCC) –

are one of the most fatal forms of malignancy occurring in humans, with melanoma contributing for the majority of skin-related deaths worldwide. However, these skin cancers are easily curable if detected in the early stages which have led to massive investment in the field of skin lesion analysis to develop automated tools for detection and classification of skin lesions [2]. Previously, conventional computer-aided (CAD) systems relied on the extraction of hand-crafted image features from the lesion area and its border to be fed to a traditional classifier which often involves extensive pre-processing and manual segmentation [3]. Recently, the rise in the application of deep learning in the medical field has led to the development of numerous promising classification methodologies. Specifically, convolutional neural networks (CNNs) [4] have achieved results as precise as certified dermatologists working on the same task and may also be capable of outperforming them [5]. In deep learning-based classification, the feature extraction and classification are both learned and performed as a single unit.

In medical image analysis, transfer learning using pre-trained models has been incorporated extensively by various medical fields. Some of the recent related works using deep learning mentioned below incorporate this strategy. Lee *et al.*, [6], provides a very good example of combining image segmentation using U-Net and DenseNet models and classification using the ensemble of classifiers to provide classification accuracy close to practiced professionals. Their work was done as a part of ISIC-2018 Challenge and claimed to have got a balanced accuracy of 78.9% for task3. Gessert *et al.*, [7] have used ensemble strategy to select the best subset of models that provide good balanced accuracy. Their work was done as a part of ISIC-2019 Challenge and have obtained a balanced accuracy score of 63.6% and 63.4% for task1 and task2 respectively. Frangi *et al.*, [8] presented another novel idea of incorporating dual CNNs (ResNet-50) simultaneously and sending the concatenated output to a synergic network which helps to reduce the intra-class variations and inter-class similarity problem. Their work aimed at identifying melanoma and nevus skin images.

From the literature survey, it is identified that early and accurate diagnosis of skin cancer can dramatically reduce the mortality rate as most of these skin diseases are curable in their nascent stages. It is also identified that diagnostic analysis of dermatologists can be highly subjective because there are high inter-class similarity and intra-class variations as well as low contrast of the skin lesions making it difficult for dermatologists to precisely identify the skin lesions. We can also acknowledge the importance of the role played by deep learning in the classification of skin lesions and how the diagnostic accuracy has improved without being affected by human subjectivity. Therefore, in this work, we will be focusing on addressing the issues of classifying skin lesions with deep learning using the concepts of transfer learning and ensemble by blending technology.

## II. METHODOLOGY

### A. Dataset

The dermoscopic images used in this work are provided by the ISIC organization. The main dataset is a combination of BCN20000 [9], HAM10000 [10], and MSK [11] datasets with a total of 25,331 dermoscopic images acquired at multiple sites and with different pre-processing methods applied beforehand. It contains images of the class's melanoma (MEL), melanocytic nevus (NV), basal cell carcinoma (BCC), actinic keratosis (AK), benign keratosis (BKL), dermatofibroma (DF), vascular lesion (VASC) and squamous cell carcinoma (SCC). The image distribution across lesion classes are depicted in Table I. These images are in 24-bit (8 bits per channel) JPEG format. Dimensions of these images are inconsistent ranging from  $600 \times 450$  to  $1024 \times 1024$ . The HAM10000 dataset contains images that centered and cropped around the lesion. Histogram corrections have been applied to some of these images. The images are of size  $600 \times 450$ . BCN20000 contains images for size  $1024 \times 1024$ .

Table I. Original Train dataset statistics

Diagnostic Category	Count (25,331)
Melanoma (MEL)	4522
Melanocytic nevus (NV)	12875
Basal Cell Carcinoma (BCC)	3323
Actinic keratosis (AK)	867
Benign Keratosis (solar lentigo/seborrheic keratosis/lichen planus-like keratosis) (BKL)	2624
Dermatofibroma (DF)	239
Vascular lesion (VASC)	253
Squamous Cell Carcinoma (SCC)	628

In this work, we have considered only a subset of images from the main dataset due to the issues we encountered during training. Two datasets, having 4999 and 7900 images, are compiled from the original dataset. Datasets are chosen in a way that they maintain the class imbalance present in the original set as much as possible. Note that the original dataset did not have any images of unknown (UKN) class, hence that class is not included in the current dataset. An

additional dataset containing 4242 images are compiled as test dataset for testing purpose. The original test dataset provided by the ISIC Organization is not used in our work. As such, the final prediction is performed for 8 categories.

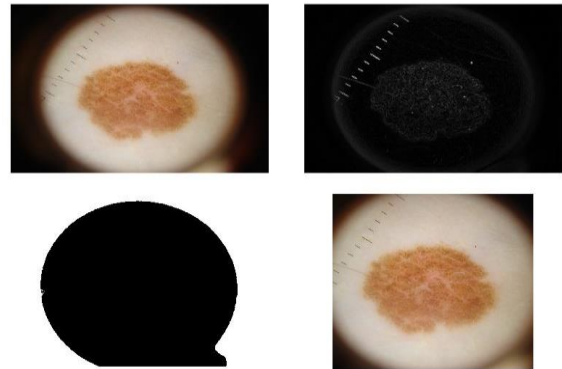


Fig 1. Cropping strategy for uncropped skin images: top left: original image, top right: elevation map, bottom left: watershed segmented image, bottom right: cropped image

As the dataset contains a mixture of cropped and uncropped images, the first step we do is use a cropping strategy to deal with uncropped images as illustrated in Fig. 1. These uncropped images often show large, black mass. Crop algorithm by [12] is partially adapted in this work to identify the lesion region. First, an elevation map is created using the Sobel gradient of the image. Sobel filter allows us to detect the edges in the image. Then the markers are found for the background and the lesion region (foreground) using a very low threshold value (0.3). The threshold value is computed experimentally. Watershed transform is used to then fill regions of the elevation map starting from the markers determined in the previous step. This will give us the segmented lesion. The next step is to find the bounding box for the segmented lesion. To calculate the bounding box, we apply the connected component labeling algorithm. This assigns the same label to all the pixels that are in the same region. In this way, different regions obtained after segmentation are labeled using distinct labels. This also allows us to access the properties associated with each region like area, extent, major and minor axis, and so on. Next, to identify the lesion region among others, we consider the area (the number of pixels of the region) and extent (the ratio of pixels in the region to pixels in the total bounding box) features of each region. The region having the largest area and extent  $> 0.5$  is considered as the target. Otherwise, we consider the three largest regions, and the first one that has an extent  $> 0.5$  is taken as the target lesion region. Once the target region is found, to get the bounding box for that region, we find the centroid (center of mass) and the major and minor axis of that region. Based on these values, the bounding box for cropping the lesion area is calculated. Furthermore, we resize all the images in the dataset to a constant size of 480 pixels (HAM1000 resolution as reference) preserving the aspect ratio.

In this work, we mostly rely on the Inception family of networks that first introduced  $1 \times 1$  convolution for dimensionality reduction in its Inception block which is capable of performing multi-level feature extraction. The

Inception module is based on a pattern recognition network that mimics the animal visual cortex. Inception models are trained on the ImageNet dataset and were selected because of their high-performance accuracies on ImageNet challenge. We have used Inception-v3, Inception-Resnet-v2, and Xception models from the Inception family of models. Another model Densenet was also selected due to its high performance in classification tasks. This also brings variability in our final ensemble. The specification of these models and the number of trainable parameters is given in the Table II.

Table II. CNN architectures specifications for transfer learning

Architecture	Input Size	Trainable Parameters
DenseNet201	224 × 224	19.3 M
Inception-v3	299 × 299	23.63 M
Inception-ResNet-v2	299 × 299	55.87 M
Xception	299 × 299	22.85 M

In this section, we will describe the modifications done on these architectures to perform the classification of skin lesion images into 8 categories. All of these CNN models have been pre-trained on the ImageNet dataset and then on a reduced ISIC-2019 train-split dataset. A few modifications are done to adapt the model to our use cases. The original classifier part of these pre-trained CNNs is replaced by batch-normalization layer followed by dropout layer, then a global average pooling layer, followed by two dense layers of 512 neurons, a second dropout layer, and lastly the classification layer with softmax as activation function to classify the dermoscopic images into 8 categories as shown in Fig 2. The softmax layer gives a prediction score in the range (0.0 to 1.0) distributed over the list of classes. All these models follow the same train pipeline with identical settings. Given a training set  $D_{train} = \{x_i, y_i\}_{i=1}^N$ , where  $x$  denotes the input image of dimension  $d$ ,  $y$  its corresponding label, and  $N$  is the number of input images, first, the images are cropped before passing them for training.

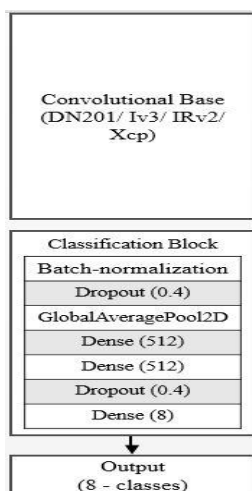


Fig 2. Eight Modified classification block of base-classifiers

The cropped images are then standardized and normalized. Normalizing is done to remove any bias present in the data [13].

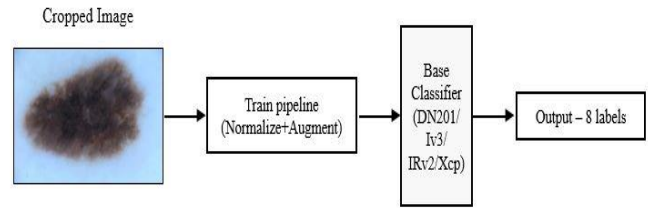


Fig 3. Train pipeline for each base-classifier

The normalized images are then passed to the augmentation pipeline where vertical and horizontal flipping, rotation, zooming, shearing, and width-shift augmentations are applied. In this work, an online-augmentation policy is employed meaning, the augmentations are applied on the fly when the training process takes place. This does not increase the image count rather different versions of augmented images are presented during each pass (epoch) of the training data. These augmented images are then fed to the pre-trained CNN models for training as illustrated in Fig 3.9. The saved models are then loaded and embedded to a meta-classifier and trained using the proposed algorithm. In this work, the blending procedure is implemented to ensemble the predictions done by the base-classifiers. It is a two-level process: level-0 and level-1. The four CNN pre-trained models: DenseNet201, InceptionV3, InceptionResNetV2, and Xception are used as level-0 sub-models to make the initial predictions. These models are trained on the train-split part of the training data and the best versions of each model are saved separately. As neural networks are used as sub-models, we have used a non-linear multi-layer perceptron (MLP) model as a meta-learner (level-1 model) having layers as depicted in Fig. 3, as against the practice of using a linear model like Logistic Regression. Now, this allows sub-networks to be embedded in a larger multi-headed network that then learns how best to combine the predictions from each input sub-model and the ensemble can be treated as a single large model. In our work, each of these sub-models is used as a separate input-head to the meta-learner. All the layers of the loaded models are marked as not trainable so the weights cannot be updated when the blended ensemble model is being trained. The entire blended model is then fitted onto the holdout-split part of the training data. Because the sub-models are not trainable, their weights will not be updated during training and only the weights of the meta-learner will be updated. Thus, the sub-models only predict the holdout-split dataset and the meta-learner will be trained on the predictions made by these sub-models on the holdout-split of the training data. This new ensemble model is used to make predictions on the test dataset. The algorithm I for the blending strategy is given in Table III. The algorithm describes the process to fuse the base classifiers to meta-learner and train the meta-learner using the stacked predictions from the base classifiers. The final predictions are then made on the holdout dataset and test dataset.

Table III: Algorithm for Blending

Algorithm	Blending
1:	Input: training data, $D_{train} = \{x_i, y_i\}_{i=1}^k$ level-0 classifiers, $C_{L0}^1, \dots, C_{L0}^T$ level-1 classifier, $C_{L1}$
2:	Output: ensemble classifier $E$
3:	Split $D_{train}$ into 2 parts, $D_T = \{(x_n, y_n), n = 1, \dots, N\}$ , $D_{HO} = \{(x_m, y_m), m = 1, \dots, M\}$
4:	Step 1: learn base (level-0) classifiers
5:	for $t \leftarrow 1$ to $T$ do
6:	learn base classifiers $C_{L0}^t$ based on $D_T$
7:	end for
8:	Step 2: construct new data set from the predictions of $D_{HO}$ and learn meta-classifier, $E$
9:	embed base classifiers $C_{L0}^1, \dots, C_{L0}^T$ in $C_{L1}$ to get ensemble classifier $E$
10:	for $t \leftarrow 1$ to $T$ do
11:	for $i \leftarrow 1$ to $M$ do
12:	$D_E = \{x'_i, y_i\}$ , where $x'_i = stack\{C_{L0}^t(x_i)\}$
13:	end for
14:	end for
15:	learn $E$ based on $D_E$
16:	return $E$

### III. RESULTS AND DISCUSSION

The dataset is split into two parts, train and holdout (validation) splits. The dataset is split in the ratio of 80:20. The train-split part of the dataset is used to train the level-0 models (DenseNet201, InceptionV3, InceptionResNetV2, and Xception) or sub-models of the blending ensemble model. The validation-split part of the dataset is used by the sub-models to make predictions, which are then concatenated to form the train dataset for the level-1 model or meta-model. The final predictions are then made on the holdout-split dataset and test dataset. Since the data is extremely imbalanced, for training we use cross-entropy loss. Each of the class will be set with different weights and is defined in (1),

$$W_i = \frac{N}{C+n_i} \quad (1)$$

where  $C$  is the number of categories,  $n_i$  is the image count per category and  $N$  is the total number of images in the training data set.

#### A. CNN training Base Models

We leverage on transfer learning to extract features and perform classification. The training of the models is done in 2 steps. In step 1, all the layers in the base models just below the top fully connected (FC) layer are frozen. This allows the base models to behave as initial feature extractors. The training is done for 3 epochs with the learning rate set to 10e-3 and batch size to 64 images. Adam optimizer is used to regulate the learning rate. This step is performed so that the weights of the newly added layers in

the classification block does not get randomly initialized. This will also prevent large fluctuations in the gradients' updates when doing fine-tuning. In step 2 of the training process, all the layers of the base models are made trainable and fine-tuned for 50 epochs. The fine-tuning step starts from the 4<sup>th</sup> epoch onwards and the learning rate is decreased by a factor of 10 and set to 10e-4. Again the other training parameters remain the same with Adam as optimizer and batch size set to 64.

The validation loss is constantly monitored for each epoch during training and the learning rate is set to decrease by a factor of 10 if the validation loss stopped improving after 8 epochs till it reaches 10e-6. To make sure that the model does not over-fit, early stopping is used to monitor the validation loss. Three more checkpoints are used to save the model with the best-balanced accuracy and also the model with minimum validation loss. Finally, the last model after the end of the training was also saved as the latest model. Keras API's with Tensorflow as the backend is used for the above implementation.

Meta-learner training is carried out similarly as the base model training except for the initial 3-epochs training. Before training the meta-learner, all the saved base models with best-balanced accuracy are initially loaded and embedded in a multi-layered perceptron (MLP) model to form a single ensemble model with multiple heads like a hydra. When merging all base-classifiers, the layers of these models are frozen so that their weights do not get updated during the training of the ensemble model.

#### B. Analysis of the performance of base-models and ensemble model for dataset-1 and dataset-2

Using the configuration described in subsection A, the base models and ensemble model was trained for classification using dataset-1 and dataset-2. The balance multi-class accuracy and loss metrics are plotted in the graph at the end of each epoch, as illustrated in Fig 4.

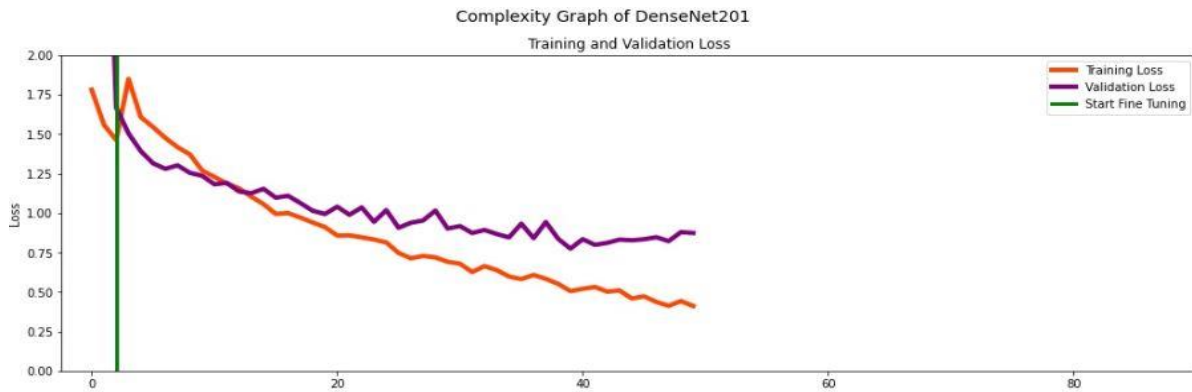


Fig 4 Train and validation loss of DenseNet201 for dataset-1

The prediction performance of the ensemble model for dataset-1 and dataset-2 and also the predictions on the test dataset is given in the Fig 5 and Fig 6.

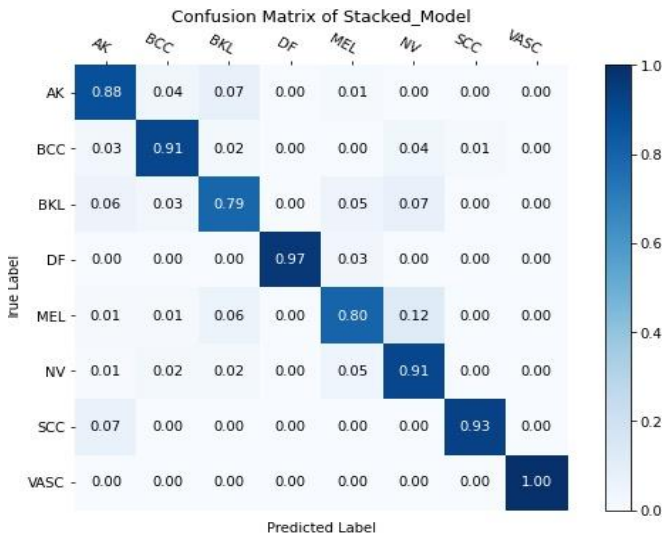


Fig 5. Confusion matrix: Prediction performance of ensemble model on validation-split of dataset-1.

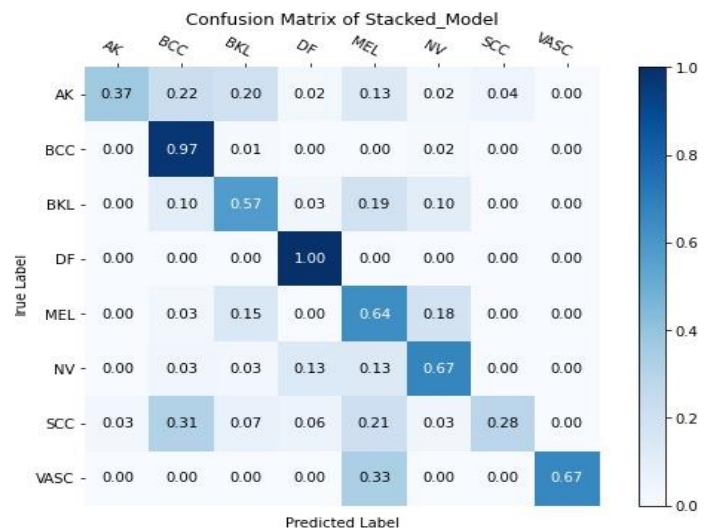


Fig 6. Confusion matrix: Prediction performance of ensemble model on the test dataset-1.

The evaluation results, balance multi-class accuracy, as well as the results of the secondary metrics for the base models and ensemble model on the validation set and test set are given in Table IV and V.

Table IV: Performance metrics of models trained on dataset-1

Models	Performance metrics (%)				
	BMA	P	Re	Sp	F1
DenseNet-V2	<b>69</b>	69	59	79	62
Inception-V3	<b>61</b>	66	54	78	57
Inception-ResNetV2	<b>61</b>	66	54	79	57
Xception	<b>59</b>	66	47	75	52
<b>Ensemble model</b>	<b>64</b>	<b>69</b>	<b>60</b>	<b>78</b>	<b>61</b>

Table V : Performance metrics of models trained on dataset-2

Models	Performance metrics (%)				
	BMA	P	Re	Sp	F1
DenseNet-V2	<b>79</b>	71	62	98	58
Inception-V3	<b>77</b>	68	59	97	55
Inception-ResNetV2	<b>79</b>	68	59	98	55
Xception	<b>72</b>	62	55	98	49
<b>Ensemble model</b>	<b>81</b>	<b>73</b>	<b>62</b>	<b>98</b>	<b>56</b>

From the above table, we can conclude that the ensemble model performs reasonably better on the unseen test images than the base-models. We also observe that the models learn better as the train data-size increases, giving better performance results. Our ensemble model got the *BMA* of 64.7% when trained on dataset-1 (4,999 images), however, its performance increased with the increased dataset (7,900 images). Note that, dataset-2 is more skewed than dataset-1. One more point to note is that DenseNetV2 performance is better than other models used as base-models for our ensemble network. Perhaps, the performance of the ensemble network will improve further on increasing the number of base-classifiers.

#### IV. CONCLUSION

On comparing the prediction results obtained, it can be summarized that the ensemble model trained using the blending technique proposed in this work performs reasonably better than the individual models. Blending based ensemble technique used in this work is one of the ways to train the ensemble model. Another well-known technique Stacking or Stacked generalization can also be used with k-fold validation. The analysis metrics got for both base classifiers and ensemble model during predictions done on test set indicates that the models perform reasonably well. There is still scope for further improvements by varying (increasing/changing) the base classifiers used. The dataset can be expanded to include all the images from the original dataset for better prediction. The metrics have shown that though the ensemble model behaves reasonably, there is still scope for improvement.

#### REFERENCES

- [1] Pathan, S., Prabhu, K. G., & Siddalingaswamy, P. C. (2018). Techniques and algorithms for computer aided diagnosis of pigmented skin lesions—A review. *Biomedical Signal Processing and Control*, 39, 237-262.
- [2] Pathan, S., Prabhu, K. G., & Siddalingaswamy, P. C. (2019). Automated detection of melanocytes related pigmented skin lesions: A clinical framework. *Biomedical Signal Processing and Control*, 51, 59-72.
- [3] Pathan, S., Prabhu, K. G., & Siddalingaswamy, P. C. (2018). Hair detection and lesion segmentation in dermoscopic images using domain knowledge. *Medical & biological engineering & computing*, 56(11), 2051-2065.
- [4] Y. Lecun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436-444, 2015.
- [5] N. C. B, J. Cai, M. Abedini, and R. Garnavi, "Deep Learning , Sparse Coding , and SVM for Melanoma Recognition in Dermoscopy Images Deep Learning , Sparse Coding , and SVM for Melanoma Recognition in Dermoscopy Images," no. October, 2015.
- [6] Y. C. Lee, S.-H. Jung, and H.-H. Won, "WonDerM: Skin Lesion Classification with Fine-tuned Neural Networks," pp. 1-4, 2018.
- [7] N. Gessert, M. Nielsen, M. Shaikh, R. Werner, and A. Schlaefer, "Skin Lesion Classification Using Ensembles of Multi-Resolution EfficientNets with Meta Data," pp. 1-10, 2019.
- [8] A. F. Frangi, J. A. Schnabel, C. D. C. Alberola-lópez, G. F. Eds, and D. Hutchison, and *Computer Assisted Intervention – MICCAI 2018*. 2018.
- [9] M. Combalia *et al.*, "BCN20000: Dermoscopic Lesions in the Wild," pp. 3-5, 2019.
- [10] P. Tschandl, C. Rosendahl, and H. Kittler, "Data descriptor: The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions," *Sci. Data*, vol. 5, pp. 1-9, 2018.
- [11] N. C. F. Codella *et al.*, "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)," *Proc. - Int. Symp. Biomed. Imaging*, vol. 2018-April, pp. 168-172, 2018.
- [12] B. Montaruli, "Skin Lesions Classification using Computer Vision and Convolutional Neural Networks Image Processing and Artificial Vision Master Degree in Computer Science Engineering Polytechnic University of Bari," vol. 10000, no. 2018, pp. 1-33, 2019.
- [13] Convolutional Neural Networks for Visual Recognition, "https://cs231n.github.io/".
- [14] N. Gessert, M. Nielsen, M. Shaikh, R. Werner, and A. Schlaefer, "Skin Lesion Classification Using Ensembles of Multi-Resolution EfficientNets with Meta Data," pp. 1-10, 2019.
- [15] A. F. Frangi, J. A. Schnabel, C. D. C. Alberola-lópez, G. F. Eds, and D. Hutchison, and *Computer Assisted Intervention – MICCAI 2018*. 2018



# EYE TUMOUR DETECTION USING DEEP LEARNING

Avigyan Sinha  
Johns Hopkins University, USA  
avigyan15@gmail.com

Aneesh R P  
College of Engineering,  
Thiruvananthapuram, India  
aneeshprakkulam@gmail.com

Nazneen. N. S  
A J College of Science and  
Technology, Thiruvananthapuram,  
India

**Abstract** -Iris melanocytic tumours, are the most dangerous tumours in the eye, commonly known as eye tumours. This includes freckle, nevus, melanocytoma, Lisch nodule, and melanoma. The detection of eye tumour is very difficult in early stages. Many research works are being carried out to detect eye diseases. But few research works in the eye tumour were published. Most of the system needs specific data acquisition devices to capture the region. This is very expensive. To diagnose eye melanoma, doctors recommend PET - CT, eye ultrasound, angiogram, optical coherence tomography, etc. Here, a new approach is presented to detect the eye tumour from eye images using deep learning technique. The deep network model created with modified LeNet architecture. The model created with the segmented eyeball images. Hough circle transformation could predict the eye ball and iris regions. As the deep learning technique need more data for training, the number of image data has been increased with image augmentation method. Successful testing of this method with an accuracy of 95% shows that this method can implement in real time applications.

**Keywords**- Ocular Melanoma, eye tumour, Cancer, Deep learning algorithm, CNN, LeNet, hough circle, grayscale conversion.

## I. INTRODUCTION

Iris melanocytic tumours are the most dangerous tumours in the eye, commonly known as 'eye tumours'. This includes freckle, nevus, melanocytoma, Lisch nodule, and melanoma [1]. The term ocular is used with tumour to represent that it is accompanied with eye. It can be intraocular, which means inside the eye or extra-ocular which means that it affects the outside part of the eye. The most common types of the intraocular iris tumours are the Cyst, Nevus, and Melanoma [2]. Some other tumours related with eye are Lacrimal Gland Tumour, lid tumour, etc. The exact cause of this disorder is not known, but certain risk factors have been noticed. The disorder is seen more often in people who have light eye colour [3]. Age, certain inherited skin disorders, Exposure to ultraviolet (UV) light, certain genetic mutations etc are also considered as the reasons of eye tumours. The eye tumour can be spread and shall affect the vision.

Routine checkups are the best methods for diagnosing the tumour. Retinoblastoma, the eye tumour, can be identified with naked eye [4]. Ophthalmoscope is commonly used to diagnose. Ultrasound, Fluorescein angiography test OCT, Semiconductor Detectors are the other common methods to diagnose. Needle biopsy is rarely used for diagnosis. CT and MRI are best in diagnosing extra ocular and intracranial extension. Diagnosis of eye tumours is considered according to age, health state, suspected disease, symptoms and past examination records. Cytogenetics and gene expression profiling are used to collect more information about

prognosis. Iris tumours are most common eye tumours and it is classified into different types. For earlier detection of these tumours, generalised procedure is needed to diagnosis the abnormality.

The two most commonly used therapeutic procedures are surgery or radiation therapy. In Radiation therapy damage produced in the tumour cells causes them to die and slowly shrink. The most common radiation therapies are endocrine therapy, brachytherapy, or sealed source radiotherapy [5]. Shrinking of the tumour region can be completely cured with local therapy. The local therapy consists of laser photocoagulation, cryotherapy, thermotherapy, Plaque radiotherapy etc. Laser photocoagulation is the primary therapy with xenon or argon arc. This coagulates all blood supply to the tumour and it can control 70% of the abnormality. Cryotherapy commonly treated for small tumours. The procedures for chemotherapy start with cryotherapy as a preparation step. Thermotherapy is the method of applying heat to the affected area using ultrasound, microwaves, or infrared radiation.

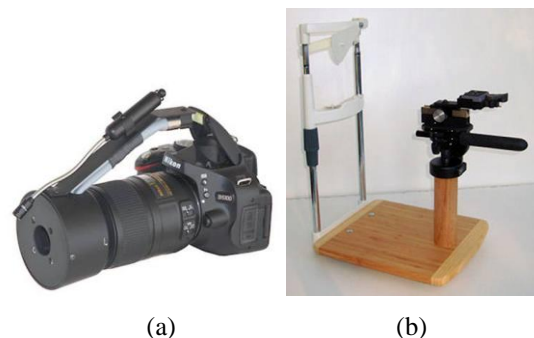


Figure 1: Iris camera and setup (a) Miles Eye Camera MEC-5 (24 megapixel) (b) CRCS-FH4 model Premium Professional Chinrest/Camera [Courtesy: MilesResearch Products]

Ophthalmology is a branch of medical science that investigates the disorders of eye. Bio medical imaging software is the efficient tools for ophthalmologist in diagnosing eye diseases. These imaging technique also help them in surgical treatment. Most of the image diagnosing machines are working with the principle of machine learning. Fundus and OCT image records can reveal the symptoms of diabetes. Many cameras are available now for capturing iris region of eye. These images are used both for biomedical and biometric applications. Figure 1 shows the images of one iris camera and setup for capturing the images. The demonstration of this camera setup is illustrated here to only get an idea for capturing such images. The eye

tumour examinations can be undertaken with these cameras. More research works should come across in the field iris tumour detection

## II. EXPERIMENTAL METHODS

In this paper [6], author proposes an image processing technique to identify the eye tumour. This experiment was conducted with 30 normal and 70 tumour images. The median filtered image segmented into two parts. Canny edge detection and image fusion methods are implemented in locating the affected area.

In this paper [7], the author investigates the symptoms of eye tumour similar to glaucoma with biomedical imaging. The author analysed this case with 9 patients and identified that eye tumours are misdiagnosed as glaucoma. They might be mistaken for glaucoma because of similar symptoms such as pain, severe headache, red eye, etc. In this study, they insisted on the importance of diagnosing tumours correctly by comparing nine cases misdiagnosed as glaucoma initially.

Here [8] the author presented an automated methods to segment iris images for early detection of eye tumour. Vander Lugt correlator based active contour method is used to segment the iris portion. K-means clustering model is used to label the tumourous tissue.

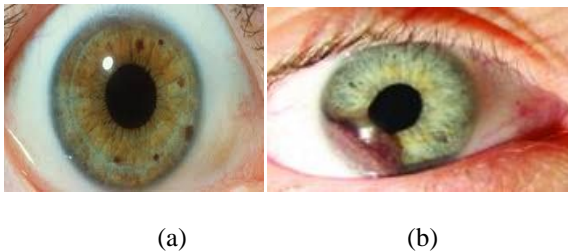


Figure 2: Eye Images (a) Normal (b) tumour

## III. PROPOSED METHOD

Deep learning techniques are mainly employed in the detection of eye tumour. Figure 3 shows the pipeline of the proposed method.

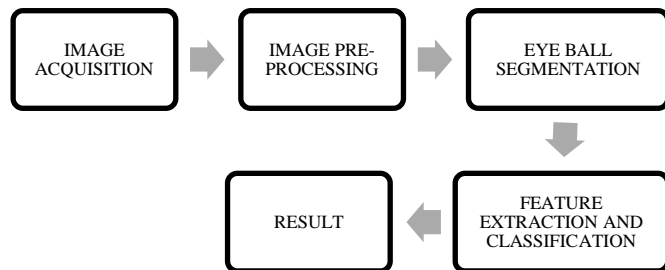


Figure 3 : Pipeline of the system

### A. Image acquisition

Image acquisition is the advent of a digitally encoded illustration of the visible traits of an object, consisting of a bodily scene or the interior shape of an object. The term is frequently assumed to mean or encompass the processing, compression, storage, printing, and display of such

images. As eye tumour has different types, many open dataset are available for research purposes. Only a limited number of image data are available for study. The images were collected from Miles Research, Eye cancer and uveal melanoma image databases. We have prepared 100 abnormal and 100 normal images after analysing the disease conditions. More data is needed for deep learning technique. So image augmentation technique is applied to increase the number. Augmentation step contains a rotation of each image, width shift range of 0.1, height shift range of 0.1, brightness range of (0.3,1.0), horizontal and vertical flip. A new dataset has been created with 2000 images.

### B. Image pre-processing

Image pre-processing step aims to prepare the input data for further analysis. It may use automated algorithms. Filtering is a vital procedure in signal processing, for outlining the capabilities of image, filtering suppresses completely or partially some elements of image. Image pre-processing mainly depends with the source. Two parameters influence in the picture quality. Intrinsic parameters are related with materials made and extrinsic parameters are related with the environment where the image captured [9]. Most of the images are in RGB format. Gray scale conversion is needed for converting the image into two dimensional arrays. Gray scale conversion methods are based on average, lightness and luminosity [10]

Let R, G, B are the colour planes of an image.

$$Average = \frac{(R+G+B)}{3} \quad (1)$$

$$Lightness = \frac{(\max(R,G,B) + \min(R,G,B))}{2} \quad (2)$$

$$luminosity = 0.21 R + 0.72 G + 0.07 B \quad (3)$$

### C. Eye ball segmentation

Segmentation is the process of splitting the entire image into different parts. Here eye ball is the region of interest (ROI) where the features of tumours exist. As the iris region is in circle shape, Hough circle detection is employed to separate the ROI [11]. The Hough transform is a technique of feature extraction which is used in many processes like digital image processing, image analysis, and computer vision. Two circle regions are inside the eyeball. The outer region can be easily separated and the inner region may not clear in diseased image. So outer circle detection is prioritised. Canny based edge detection algorithm prepares the boundary of these regions. Gradient of the images extracted by applying different threshold values [12].

A Hough circle is the transformation of all circular shapes to lines. This helps to identify the number of circles in an image. Circle is commonly represented as

$$(x-a)^2 + (y-b)^2 = R^2 \quad (4)$$

$$x = a + R \cos(\theta) \quad (5)$$

$$y = b + R \sin(\theta) \quad (6)$$

a and b are the centre points of the circle, R is the radius,  $\theta$  is the angle.

The hough transformation finds the probability of a circle within the radius. A search program is associated to find the intensity points within a circular shape and collects probable circle points. The maximum and minimum radii can also be selected. Radius of the eye related with the eye structure of human, distance between camera and eye and camera parameters. For a particular dataset, it should be in a particular range. Figure 4 shows the circle segmentation in iris images.

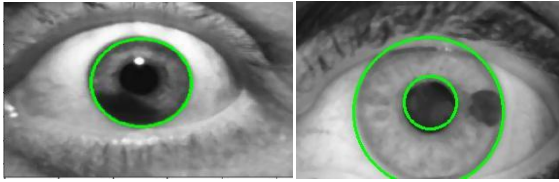


Figure 4: Circle detection using Hough transform

D. Feature Extraction and Classification

In machine learning, feature extraction and classification are two separate processes. The common features of melanoma are Asymmetry, Border irregularity, color, texture, etc [13]. In the conventional ML approach, different features like mean, variance, standard deviation, texture pattern, etc can be selected from the ROI and then proceed with a classifier like SVM, Naive Bayes, ANN etc[14]. In deep learning the entire ROI is put in a deep neural network model where the feature extraction and classification are simultaneously process [15].

Convolutional Neural Networks (ANN) are the efficient deep learning network for computer vision based classifications. It can be easily implemented with highest accuracy. LeNet is a gradient based learning network with multiple layers that is derived from CNN. CNN integrate three architectural design to establish shift, scale and distortion invariance [16]. These algorithms can classify eye tumour pattern with certain modifications. Figure 5 shows the general architecture of CNN. The following four layers do the function of feature extraction and classification [17].

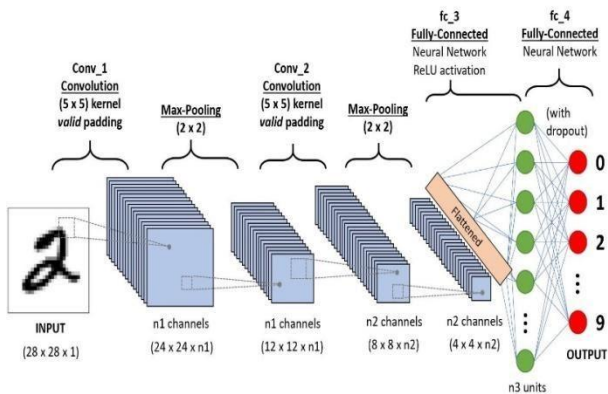


Figure 5: CNN general architecture

i) Convolutonal Layer

The Convolutonal layer is the basic building block of a Convolutional Network. It involves different functions

like Local Connectivity, Spatial arrangement and Parameter Sharing.

Local Connectivity – Connecting neurons from one layer to others is seen as unrealistic, as it increases the parameter size and computational cost. In the CNN approach, every neuron creates a local connection to other neurons. filter size is a hyper parameter which establish the connectivity range. This parameter is similar to the depth of the input.

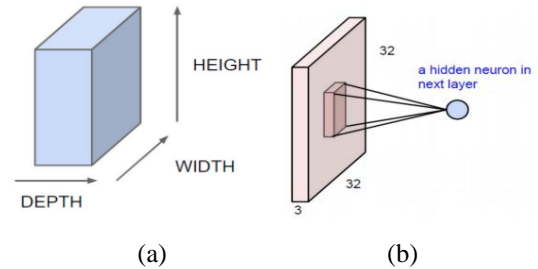


Figure 6: (a) 3D Input representation of CNN (b) Convolution as alternative for fully connected network

Spatial arrangement - The parameters controlling the size of output are depth, stride and zero-padding. The number of filters to be used is the depth. The amount by which the filter is slided is known as stride. For stride of 1, filters are moved one pixel at a time. For stride of 2, filters move 2 pixels at a time. Thus the output volumes we get are smaller spatially. The amount of zeros with which we pad the input around its border is known as zero padding. This also gives us control over the size of the output spatially. The size of the output is given by  $(2P+W-F)/S+1$ , where P = zero padding amount, W = input size, F = receptive field size of neurons, S = stride.

Parameter Sharing - Each neuron computes the gradient corresponding to its weights, but these gradients get added along each depth slice and only a single set of weights are updated per slice, during back propagation.

2. Pooling Layer

This layer gradually diminishes the spatial size of representation along with the number of parameters and computational cost in the network. It controls over fitting. The layer operates individually on each depth slice of input and resizes it by MAX operation. Most commonly it has filters of size 2x2, stride of 2 and down samples each depth slice by 2 along the width and height. It adds no parameters since it calculates a permanent function of the input. Besides max pooling, this unit can also compute average pooling or L2-norm pooling.

3. Non-Linearity Layer

A non-linear activation function which maps the feature map created by the previous Convolutional layer to an activation map is composed of this layer. It is an element-wise operation on the input such that dimensions of input and output are matching. Examples of such functions are sigmoid (logistic), hyperbolic tangent, rectified linear units (ReLU) etc.

4. Fully-connected layer

Complete connections to all activations in the previous layer are present in this layer, as in usual Neural Networks like Multi-Layer Perceptron. Their activations can be computed with a matrix multiplication followed by applying a bias offset.

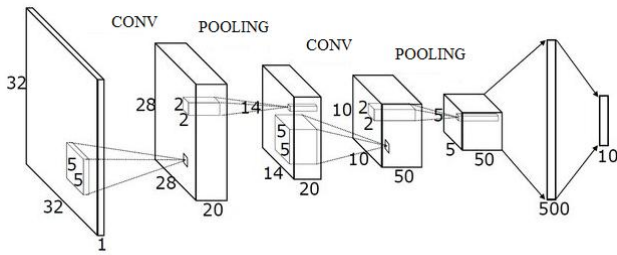


Figure 7: Modified LeNet architecture

IV. RESULTS AND DISCUSSIONS

The aim of this method is to classify the healthy and eye tumour images. Uveal melanoma is a type of eye tumour which is mainly seen in the iris, ciliary body, or choroid. Tumours arise from the pigment cells that are melanocytes which exist within the uvea giving color to the eye. These melanocytes are different from the other retinal epithelial cells originating in the retina that do not form melanomas. With augmentation technique, 2000 images were generated. We used Keras and Tensorflow in Python. Google Colab was our workstation.

For training, 80% of data (1600) was used and for validation and testing, 10 % of data (200) for each was used. Number of epochs = 50, learning rate = 0.001, Batch size = 32, optimizer = “Adam”, loss = “categorical cross entropy”. Hyper parameters were tuned by GridSearchCV method of sklearn library. As the amount of data was adequate, we did not need further data augmentation.

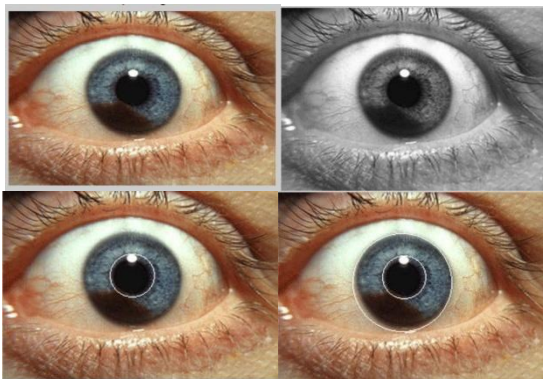


Figure 8: (a) original image (b) Gray scale image (c) pupil detection (d) iris detection

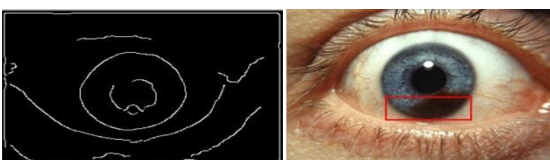


Figure 9: Edge detection and tumour region

Receiver operating characteristic curve is a measure of performance of the system. It plots the true positive rate over false positive rate. The ROC curve in Figure 11, the performance of the model taking Normal condition as class one. Area under the curve is 0.98. Similarly ROC curve shown in Figure 12, treats Tumour condition as class one, Area under the curve is 0.95.

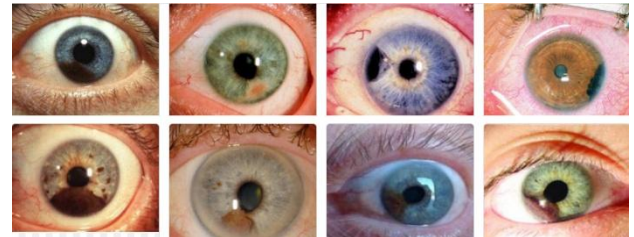


Figure 10: Tumour images in the dataset

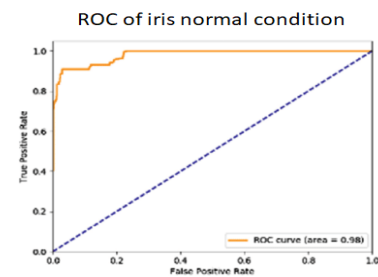


Figure 11: ROC plot of Normal Condition

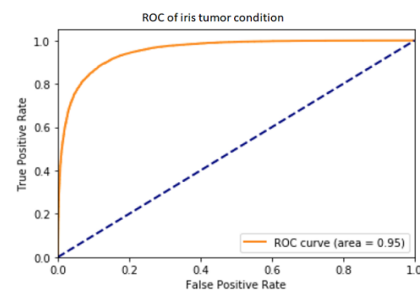


Figure 12: ROC plot of Tumour Condition

$$Accuracy = \frac{(TP+TN)}{(TP+TN+FP+FN)} = 95 \%$$

$$Sensitivity = \frac{TP}{TP+FN} = 97.87 \%$$

$$Specificity = \frac{TN}{TN+FP} = 92.45 \%$$

V. CONCLUSION

Research work in ocular tumour or tumour need more attention in the present world. Many open sources are available for this research work. We selected eye tumour images from different databases for study purposes. We have developed deep learning algorithm with modified LeNet architecture which is a novel work in the field of eye tumour detection. Here a new algorithm is proposed to detect eye tumour. Our method classifies the normal images with 98% accuracy and abnormal images with an

accuracy of 95%. Both of these results show that our model is able to distinguish well between Tumour and Normal condition. Our system can be effectively implemented into real life applications.

More research works are expected to come across in the identification eye tumour. As the iris cameras are used for different purposes, more dataset shall be taken shared for research purposes. Hospitals can Implement eye tumour detection system using iris camera. It can be also implemented in real time mobile applications. More features can be analysed through machine learning algorithms.

#### REFERENCES

- [1] Shields CL, Shields PW, Manalac J, Jumroendarasame C, Shields JA. "Review of cystic and solid tumours of the iris". *Oman J Ophthalmol.* 2013;6(3):159-164. doi:10.4103/0974-620X.122269
- [2] S. Wyder, F. Hennings, S. Pezold, J. Hrbacek and P. C. Cattin, "With Gaze Tracking Toward Noninvasive Eye Cancer Treatment," in *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 9, pp. 1914-1924, Sept. 2016
- [3] R. J. Zawadzki et al., "Progress on Developing Adaptive Optics–Optical Coherence Tomography for In Vivo Retinal Imaging: Monitoring and Correction of Eye Motion Artifacts," in *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 20, no. 2, pp. 322-333, March-April 2014.
- [4] Helen Dimaras Timothy W. Corson , "Retinoblastoma, the visible CNS tumour: A review" *Journal of Neuroscience Research* , January 2018
- [5] K. Svanberg, "Diagnostics and treatment of tumours using laser techniques," 2008 Asia Optical Fiber Communication & Optoelectronic Exposition & Conference, Shanghai, 2008, pp. 1-3.doi: 10.1364/AOE.2008.SuN1
- [6] Helwan, Abdulkader. (2014). ITDS: Iris Tumour Detection System using image processing techniques. *International Journal of Scientific and Engineering Research.* 5. 76-80.
- [7] Min Yang<sup>1</sup>, Wei Wang<sup>2</sup> " Eye Tumours Misdiagnosed as Glaucoma", *Chinese Medical Journal* January 20, 2015 Volume 128 Issue 2
- [8] Moulay Ismail University, Faculty of Sciences, Department of Physics "Automated segmentation of ophthalmological images by an optical based approach for early detection of eye tumour growing", *Phys Med.* 2018 Apr;48:37-46. d. Epub 2018 Mar 27
- [9] Avigyan Sinha, et al. "Real Time Facial Emotion Recognition using Deep Learning", *International Journal of Innovations and Implementations in Engineering(ISSN 2454- 3489)*, 2019, vol 1
- [10] John D Cook , "Three algorithms for converting color to grayscale" , *The Endeavour*, 2009
- [11] Harvey Rhody, "Lecture 10: Hough Circle Transform" , *DIP Lecture 10*, October 11, 2005
- [12] Leshmi Satheesh "Estimation of Diabetic Retinopathy from Retinal Images Using Artificial Neural Network" -*IJIE-International Journal of Innovations & Implementations in Engineering (ISSN2454- 3489)* 2015 June Edition Volume 1.
- [13] Mohammed Thanveersha N et al. "Automatic Brain Hemorrhage Detection Using Artificial Neural Network", *International Journal of Innovations and Implementations in Engineering(ISSN 2454- 3489)*, 2019, vol 1
- [14] Soumya R S "Advanced Earlier Melanoma Detection Algorithm Using Colour Correlogram" ,*IEEE 2016 International Conference on Communication Systems and Networks (ComNet) | 21-23 July 2016 | Trivandrum.*
- [15] Malavika Suresh, et al. "Real-Time Hand Gesture Recognition Using Deep Learning", *International Journal of Innovations and Implementations in Engineering(ISSN 2454- 3489)*, 2019, vol 1
- [16] Yann LeCun, Leon Bottou ,et al. , " GradientBased Learning Applied to Document Recognition" *PROC OF THE IEEE NOVEMBER*
- [17] S. Albawi, T. A. Mohammed and S. Al-Zawi, "Understanding of a convolutional neural network," 2017 *International Conference on Engineering and Technology (ICET)*, Antalya, 2017, pp. 1-6, doi: 10.1109/ICEngTechnol.2017.8308186.

# Classification of Arrhythmia in Time Series ECG Signals Using Image Encoding And Convolutional Neural Networks

Vandith Sreenivas K, Ganesan M, Lavanya R

*Department of Electronics and Communication Engineering,*

*Amrita School of Engineering, Coimbatore*

Amrita Vishwa Vidyapeetham, India

vandithsreenivas@gmail.com, m\_ganesan1@cb.amrita.edu, r\_lavanya@cb.amrita.edu

**Abstract**—Electrocardiograph (ECG) signal analysis has been used extensively to study a patient’s heart and detect problems like arrhythmia for decades. Manual analysis of ECG in real time is laborious and therefore not practical for doctors. Deep learning helps make this job much easier due to quicker learning of signal features and event prediction. Deep Learning classifiers can help doctors differentiate between normal and abnormal ECG signals based on the basic and advanced features of ECG signals. This paper focuses on building a Convolutional Neural Network (CNN) to classify arrhythmia in dual channel ECG signals based on images generated by time series to image encoding techniques. The ECG time series signals were converted into images using Gramian Angular Fields (GAF) and Markov Transition Fields (MTF). These images were fed as input into the deep learning classifier which further classified the signals into various types. Our model achieved an accuracy of 97% for the GAF images and 85% per cent for the MTF images.

**Index Terms**—Gramian Angular Fields, Markov Transition Fields, Convolutional Neural Networks

## I. INTRODUCTION

The classification of ECG signals for cardiac diagnosis is not a new concept. However, the popularity of deep learning and computational biology has encouraged scientists to use data science in analysing and classifying medical data. Moody and Mark depicted the impact of storing ECG records as an online catalogue in the form of the MIT-BIH Arrhythmia Database (mitdb) [1]. This paved the way for a plethora of classification algorithms for various kinds of ECG signals, especially those that represent arrhythmia, a condition involving irregular heartbeat. Krishnan et al in [2] developed an algorithm to classify cardiac arrhythmia using autoregressive modeling. They were able to perform classification for four kinds of arrhythmia. One of the most important features of an ECG signal is its RR interval(interval between the QRS peaks of two ECGs). Sideris et al in [3] were able to perform classification of arrhythmia based only on the RR intervals of the signals, which reduced computational complexity. Some methods also reduced the features of the signals and then performed classification, like Song et al in [4] who used Linear Discriminant Analysis (LDA) for feature reduction and Support Vector Machines (SVM) for classification. Singh et al [5] used Recurrent Neural Networks (RNN) for classifying

multiple beats of ECG. The use of conventional Convolutional Neural Networks (CNNs) for arrhythmia classification was demonstrated by Jun et al. [6], where they used various deep learning methods such as data augmentation in order to prepare the ECG signals. These signals had to be converted to images for input. Most methods for classifying ECG deal with single channel recordings from datasets. In the case of multi channel ECG, the morphology of the signal might change from channel to channel and therefore single channel ECG classification might prove fallacious, as demonstrated by Sanchez et al. [7]. Kim [8] proved that multi-channel ECG classification yields higher accuracy than single channel ECG classification using conventional signal input. Since ECG is a time series signal, it is possible to use time series data analysis in order to encode the signals into images for classification. This is not the direct conversion of signals to images, rather a mapping of a one-dimensional time series ECG signal to a colored image, as depicted by Whang and Oates in [9]. This paper proposes to do the same, for the MIT-BIH Arrhythmia Database. However, for multiple classes, feature learning is important. For example, while analysing Atrial Fibrillation (AF), Christov et al. [10] were able to extract multiple features which could enhance detection of AF in future algorithms.

Analysing features is important to perform classification for various types of arrhythmia simultaneously. Gabbouj et al. in [11] performed patient specific analysis and found that inter patient variations also affect classifier performance. Nevertheless, multi channel ECG data from the mitdb dataset can be converted to images using Gramian Angular Fields (GAF) and Markov Transition Fields (MTF), which have been implemented in this paper for data preparation. In the case of multi channel ECG, Barro et al. [12] showed that the morphology of the signal might change from channel to channel and therefore single channel ECG classification might prove fallacious. So, here we propose to classify dual channel signals. Thiagarajan et al. [13] used an approach of generating the channels for the ECG signals and then classifying them accordingly. Zhang , Xiao and Ji [14] were able to perform classification based on a faster region based Convolutional Neural Networks (CNN), which included image generation and a fast CNN architecture.

Liang et al. [15] used a combination of CNN and a Long Short Term Memory (LSTM) networks, wherein the CNN performed as normal, but the LSTM was able to retain some feedback information which enhanced classification performance. Using wavelet transforms, morphological features of ECG signals were extracted by Coimbra et al. [16]. However, the episodic nature of ECG has contributed to limited applicability in classification due to small datasets and occasional anomalies in signal morphology Clifford et al.[17]. Nevertheless, there are multiple parameters that can determine the presence of AF, like f-waves, absence of P-waves etc. Images capturing the morphology of a heartbeat were implemented by Jiang et al. [18] using one-hot encoding. Sujit et al. in [19] used advanced methods like Synthetic Minority Over-Sampling Technique (SMOTE) to enhance detection ability in Phonocardiogram (PCG) signals.

Single lead signals of arrhythmia like AF can be classified without any pre-processing and still help obtain good classification performance, as exhibited by Sujit et al. [19]. Limitations of ECG arrhythmia classification include requiring to carefully select ECG recordings without cross validation, beat loss during denoising and feature extraction, fewer classes of arrhythmia and lesser classification accuracy for implementation as shown by Song et al [4]. However, these limitations can be overcome as more and more data is collected from patients, and the conversion of one-dimensional signals to two-dimensional images means that initial noise filtering and feature extraction can be avoided. Soman et al. in [20] were able to detect AF in single lead ECG with the help of RNNs with an accuracy of 95%. Since classification of all kinds of arrhythmia in the MIT database would not reap impressive results, major arrhythmia types can be selected like Ganesan et al. [21], who chose 7 classes and proceeded with the conventional CNN. Deep learning enables automated feature learning, which is much faster than manual extraction, which is important especially in the case of a random event like arrhythmia. Comparing models for classification performance is also abundant in the field of ECG analysis, as shown by Saiharsha et al. [22]. In this study, we use the MIT-BIH Arrhythmia Database (mitdb) for performing classification for normal and arrhythmic beats.

## II. METHODOLOGY

The typical conventional classification algorithm for arrhythmia consisted of selecting one of the channels of the signal (V1 or V2), pre-processing the signal to remove artefacts or baseline wandering, splitting each sample into smaller bits, detection of Q peaks and presence of features unique to each arrhythmia, feature extraction and eventually classification. In this paper, however, we chose both the channels (V1 and V2 or MLII and V5) for each record and then converted them into images. The two dimensional nature of images enables us to color code each channel and map both signals onto the plot. Fig.1 shows an ECG signal where arrhythmia is present. This will also reflect in the image when the data is encoded using the GAF and MTF methods. Fig. 2 represents the steps

involved in the entire project. Initially, the signal had to be pre-processed, so we performed median filtering and baseline wandering removal. This made learning of features for the images much easier.

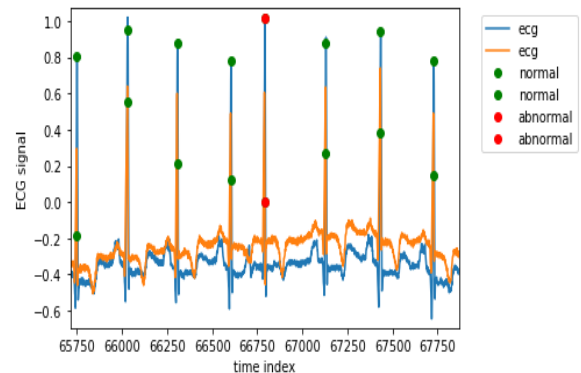


Fig. 1: Presence of an abnormal beat in the dual-channel ECG signal.

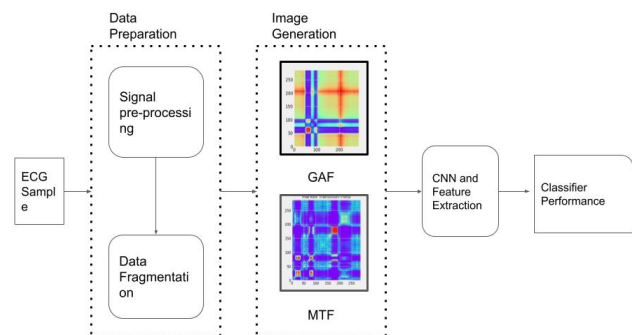


Fig. 2: The project workflow diagram representing the steps involved in the process.

### A. Image Encoding

An ECG is a time series signal indexed in time order with timestamps for each value. There may be recordings from several channels for the same time interval. In this study, we use ECG time series signals consisting of two channels. Analysis of two channels at a single instance is difficult in the one-dimensional signal, but can be made much easier by converting the signal into a two-dimensional color image as demonstrated by Yang et al. [23]. This helps map temporal data of multiple channels (in this case, two) into a single image, which will hopefully help better find features that may not have been found if one-dimensional time series data had been used. This kind of imaging is not a direct conversion, rather an “encoding” of sorts, that is, a change in domain (like with the GAF) or a transition probability (with the MTF). Through this, we assume that the classifier performance will improve for the dual channel ECG data.

### B. Gramian Angular Fields (GAF)

An image obtained from a time series signal, that contains some temporal correlations between each time point. Initially, time series data can be mapped onto the polar coordinate by computing the inverse cosine function (arccos) of the values. This representation based on polar coordinates is a new way of interpreting time series. The corresponding values warp on the spanning circles between different angular points, analogous to ripples in water, as time increases. The angular perspective can be used to define the temporal relation for each distinct time interval after translating the time series into the polar coordinate system by considering the trigonometric sum between each point. The matrix, referred to as the Gram matrix, is represented by  $G$  depicted in equation (1). It represents the polar coordinates of each value in the time series signal.

$$G = \begin{bmatrix} \cos(\phi_1 + \phi_1) & \dots & \cos(\phi_1 + \phi_n) \\ \cos(\phi_1 + \phi_1) & \dots & \cos(\phi_1 + \phi_n) \\ \dots & \dots & \dots \\ \cos(\phi_n + \phi_1) & \dots & \cos(\phi_n + \phi_n) \end{bmatrix} \quad (1)$$

Looking closely at the plot in Fig. 3, it is clear that the intersection of both lines represents the QRS peak. The x-axis represents the time and the y-axis represents amplitude of the signal. The colors also vary as blue represents the first channel and orange represents the second. The GAF has different advantages. Temporal dependency can be preserved as time increases as the position moves from top-left to bottom-right. However, the size of the Gramian matrix could be large if the time series is long.

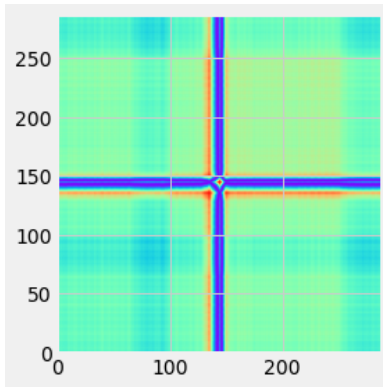


Fig. 3: The Gramian Angular Field of an interval of the ECG time series signal.

### C. Markov Transition Fields (MTF)

In a first order Markov chain along the time axis, the elements of  $W$  constitute the transitions between quantile bins.  $W$  becomes a Markov transformation matrix after normalization. However, this matrix is time insensitive. The Markov Transition Field is therefore defined as shown in equation (2) in order to compensate for the loss due to lack

of temporal dependence.

$$M = \begin{bmatrix} w_{ij}|_{x_1 \in qi, x_1 \in qj} & \dots & w_{ij}|_{x_1 \in qi, x_n \in qj} \\ w_{ij}|_{x_2 \in qi, x_1 \in qj} & \dots & w_{ij}|_{x_2 \in qi, x_n \in qj} \\ \dots & \dots & \dots \\ w_{ij}|_{x_n \in qi, x_1 \in qj} & \dots & w_{ij}|_{x_n \in qi, x_n \in qj} \end{bmatrix} \quad (2)$$

In fact, the MTF  $M$  codes the multi-span transition probabilities of the time series by assigning the probability at each pixel from the quantile at time stage  $t_i$  to the quantile at time phase  $j$ . Each factor denotes the probability of transition with  $k$ -interval time between points. For instance, with a skip step,  $M_{ij}=1$  demonstrates the transition process along the time axis.. The key diagonal  $M_{ii}$ , which is a special case when at time phase  $t_i$   $k=0$  captures the probability of each quantile to itself (the probability of self-transition). Fig. 3 illustrates the MTF over the same signal fragment as was selected for the GAF. A shift in amplitude over time is demonstrated by transition lines and color variations.

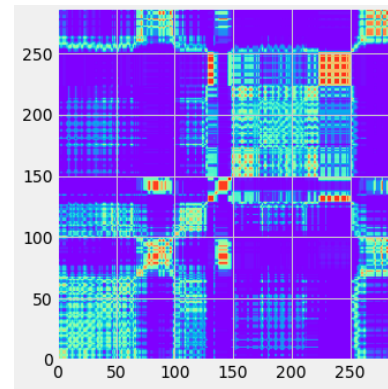


Fig. 4: The Markov Transition Field of an interval of the ECG time series signal.

### D. Dataset

The recordings of ECG arrhythmia used in this paper are taken from the MIT-BIH arrhythmia database [1]. The database for this paper includes 8 ECG recordings (100 to 107), each 30 minutes long, that were chosen from 48 recordings. Each signal is sampled at 360 samples per second. For images, we needed only one waveform per image. Therefore, each sample contained about 288 fragments that contained roughly one cycle each. Therefore our primary goal was to split each sample into approximately 2250 fragments of length 288, which contain at least one PQRST signal containing one peak. This was to ensure that when the images are generated, feature extraction during classification becomes much simpler. Each fragment was then encoded using GAF and MTF. Time series requires that temporal dependency be preserved, which is accomplished by the above two methods. For this study, we selected 2700 images of each GAF and MTF. Beats with 15 distinct forms of arrhythmias are present in the MIT-BIH database, including Standard. For this paper, we combined



all the 15 forms of arrhythmia into one big superclass called 'Abnormal' (A) and all the normal beats into 'Normal' (N). The database contains annotations for all signals. This means that information regarding the occurrence of arrhythmia is predetermined. The annotations give us exact points on the dataset where the arrhythmia has occurred.

### E. Convolutional Neural Network (CNN)

Once the signals are converted to images, they could be labelled. The next step was to set up the classification model. In this study, we implemented a conventional CNN for classifying the dataset, with all the layers and hyper-parameters tuned to our needs. There is a trade-off between classifying for various types and obtaining enhanced classifier performance due to the small data size for some lesser recorded beat types. In this study, we decided to perform binary classification, which would help us achieve better classification performance. We implemented a basic 2D CNN, based on the convolution-pooling architecture. Each image was of size 224 x 224 at the input layer. Three convolution-pooling layers, along with a dense and flatten layer that uses a 'ReLU' (Rectified Linear Unit) activation function. For optimisation, we used the Adam optimiser and trained it at a learning rate of 0.000001. The model was run for 50 epoch for each image type and the classification report was obtained for each case. We implemented the 80:20 split for the training and testing data respectively.

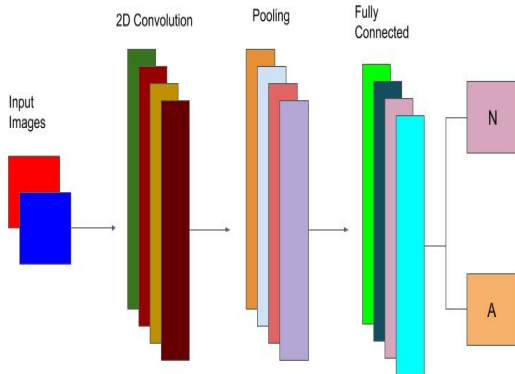


Fig. 5: The model used for this project It is a conventional 2D CNN that uses the ReLU activation function and Adam optimizer.

## III. RESULTS

Comparing the performance for both images, it was found that the model that used GAF images as input had achieved a training accuracy of 97%, as opposed to the model using the MTF images, which obtained 96%. This is most probably due to the lower error rate of the GAF, and the potential over-fitting

that has been observed in previous studies conducted on time series encoding. As for testing, it was found that the model using the GAF images as input achieved an accuracy of 97%, which is higher than the model using MTF images, which could obtain only 84% as per Table I. Although MTF displays information about dynamics, GAF encodes static information. We consider them as two "orthogonal" channels from this point of view, like various colors in the RGB image space.

TABLE I: Evaluation parameters for both image types

Method	Accuracy (%)		Loss	
	Train	Test	Train	Test
GAF	97	97	0.34	0.38
MTF	96	84	0.22	0.51

TABLE II: Accuracy of models for both GAF and MTF

Method	Accuracy (%)		Loss	
	Train	Test	Train	Test
GAF	97	97	0.34	0.38
MTF	96	84	0.22	0.51

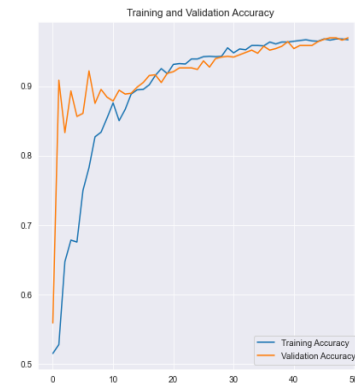


Fig. 6: The training and testing accuracy of the model(GAF).



Fig. 7: The training and testing accuracy of the model(MTF).

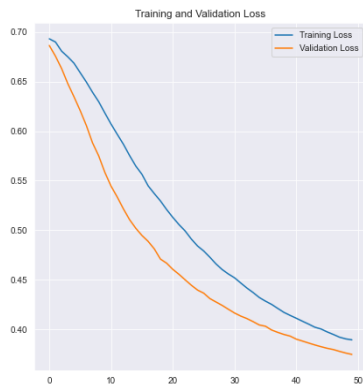


Fig. 8: The training and testing loss of the model(GAF).

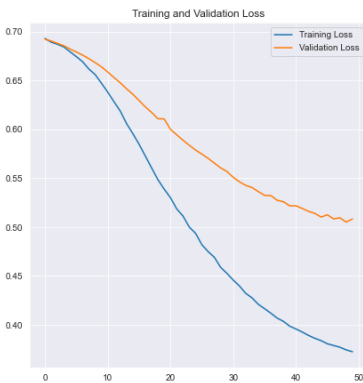


Fig. 9: The training and testing loss of the model(MTF).

Both GAF and MTF are not traditional 'natural' images; they have no image-making characteristics. They are color maps consisting of the ECG signal's different sets of amplitude values, the time representing the x-axis and the amplitude representing the y-axis. The intersection of lines in the GAF represents the QRS peak, and the color variations reflect the change in values. As for the MTF, each transition is recorded and color mapped for. From Fig. 6 it is clear that the accuracy of the GAF model is high for both the training and testing phases. Fig. 7 shows a similar pattern for decrease in loss. The GAF model was able to obtain better precision, recall and F1 score as compared to the MTF model as illustrated by Table II, where N denotes 'Normal' and A represents 'Abnormal'. As observed in Fig. 8, the testing accuracy for the MTF images was found to be much lower than the training accuracy. This could be due to the complexity of the MTF image due to its dynamic nature.

Fig. 9 also says the same story for loss, which does not decrease significantly over the epochs. The accuracy and loss curves for the GAF model are very low at the 50th epoch, which is not the case for the MTF model. These curves represent the performance of the models over time(epochs). The x-axis in each graph represents the epochs and the y-axis represents the value of accuracy or loss accordingly. The

more the model gets trained, the better the graphs turn out to be eventually. This approach is used only for time series data due to the temporal dependency nature of the GAF and MTF methods used in this study.

## CONCLUSION

The classification of dual channel ECG signals using time series encoding was performed successfully. Compared to Coimbra et al. [16], who obtained an accuracy of 96.5% for randomly chosen data from the MIT database(single channel), we were able to achieve 97% for the GAF image model. The potential prospect of multi class models for the MIT database performing well is difficult given the computational complexity involved, lack of adequate data for lesser known classes of arrhythmia, and the hardware constraints of the computer on which the model in this study was executed. However, advanced feature learning algorithms and increase in data recording will help improve classification and detection of various kinds of arrhythmia in the future.

## REFERENCES

- [1] Moody GB, Mark RG. The impact of the MIT-BIH Arrhythmia Database. *IEEE Eng in Med and Biol* 20(3):45-50 (May-June 2001). (PMID: 11446209).
- [2] Ge, Dingfei, Narayanan Srinivasan, and Shankar M. Krishnan. "Cardiac arrhythmia classification using autoregressive modeling." *Biomedical engineering online* 1, no. 1 (2002): 5.
- [3] Tsipouras, Markos G., Dimitrios I. Fotiadis, and D. Sideris. "An arrhythmia classification system based on the RR-interval signal." *Artificial intelligence in medicine* 33, no. 3 (2005): 237-250.
- [4] Song, Mi Hye, Jeon Lee, Sung Pil Cho, Kyoung Joung Lee, and Sun Kook Yoo. "Support vector machine based arrhythmia classification using reduced features." (2005): 571-579.
- [5] Singh, Shraddha, Saroj Kumar Pandey, Urja Pawar, and Rekh Ram Janghel. "Classification of ECG arrhythmia using recurrent neural networks." *Procedia computer science* 132 (2018): 1290-1297.
- [6] Jun, Tae Joon, Hoang Minh Nguyen, Daeyoun Kang, Dohyeun Kim, Daeyoung Kim, and Young-Hak Kim. "ECG arrhythmia classification using a 2-D convolutional neural network." *arXiv preprint arXiv:1804.06812* (2018).
- [7] Barro, S., M. Fernandez-Delgado, J. A. Vila-Sobrino, C. V. Regueiro, and E. Sanchez. "Classifying multichannel ECG patterns with an adaptive neural network." *IEEE Engineering in Medicine and Biology Magazine* 17, no. 1 (1998): 45-55.
- [8] Kim, Kyungna. Arrhythmia classification in multi-channel ECG signals using deep neural networks. Technical Report No. UCB/EECS-2018-80. <http://www2.eecs.berkeley.edu/Pubs/TechRpts/2018/EECS-2018-80.html>, 2018.
- [9] Wang, Zhiguang, and Tim Oates. "Encoding time series as images for visual inspection and classification using tiled convolutional neural networks." In *Workshops at the twenty-ninth AAAI conference on artificial intelligence*, vol. 1. 2015. 000091-000096. IEEE, 2018.
- [10] Christov, Ivaylo, Vessela Krasteva, Iana Simova, Tatiana Neycheva, and Ramun Schmid. "Multi-parametric analysis for atrial fibrillation classification in ECG." In *2017 Computing in Cardiology (CinC)*, pp. 1-4. IEEE, 2017.
- [11] S. Kiranyaz, T. Ince, R. Hamila and M. Gabbouj, "Convolutional Neural Networks for patient-specific ECG classification," 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Milan, 2015, pp. 2608-2611, doi: 10.1109/EMBC.2015.7318926.
- [12] Barro, S., M. Fernandez-Delgado, J. A. Vila-Sobrino, C. V. Regueiro, and E. Sanchez. "Classifying multichannel ECG patterns with an adaptive neural network." *IEEE Engineering in Medicine and Biology Magazine* 17, no. 1 (1998): 45-55.

- [13] Rajan, Deepta, and Jayaraman J. Thiagarajan. "A generative modeling approach to limited channel ECG classification." In 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp. 2571-2574. IEEE, 2018
- [14] Ji, Yinsheng, Sen Zhang, and Wendong Xiao. "Electrocardiogram classification based on faster regions with convolutional neural network." *Sensors* 19, no. 11 (2019): 2558. 10.1109/EMBC.2015.7318926.
- [15] Zheng, Zhenyu, Zhencheng Chen, Fangrong Hu, Jianming Zhu, Qunfeng Tang, and Yongbo Liang. "An automatic diagnosis of arrhythmias using a combination of CNN and LSTM technology." *Electronics* 9, no. 1 (2020): 121.
- [16] Ye, Can, Miguel Tavares Coimbra, and BVK Vijaya Kumar. "Arrhythmia detection and classification using morphological and dynamic features of ECG signals." In 2010 Annual International Conference of the IEEE Engineering in Medicine and Biology, pp. 1918-1921. IEEE, 2010.
- [17] Clifford, Gari D., Chengyu Liu, Benjamin Moody, H. Lehman Li-wei, Ikaro Silva, Qiao Li, A. E. Johnson, and Roger G. Mark. "AF Classification from a short single lead ECG recording: the PhysioNet/Computing in Cardiology Challenge 2017." In 2017 Computing in Cardiology (CinC), pp. 1-4. IEEE, 2017.
- [18] Li, Jia, Yujuan Si, Tao Xu, and Saibiao Jiang. "Deep convolutional neural network based ECG classification system using information fusion and one-hot encoding techniques." *Mathematical Problems in Engineering* 2018 (2018).
- [19] N. R. Sujit, Dr. Santhosh Kumar C., and Rajesh C. B., "Improving the Performance of Cardiac Abnormality Detection from PCG Signal", in AIP Conference Proceedings, 2016, vol. 1715
- [20] V. G. Sujadevi, Dr. Soman K. P., and Vinayakumar, R., "Real-Time Detection of Atrial Fibrillation from Short Time Single Lead ECG Traces Using Recurrent Neural Networks", *Intelligent Systems Technologies and Applications*. Springer International Publishing, Cham, pp. 212-221, 2018.
- [21] Rajkumar, A., M. Ganesan, and R. Lavanya. "Arrhythmia classification on ECG using Deep Learning." In 2019 5th International Conference on Advanced Computing and Communication Systems (ICACCS), pp. 365-369. IEEE, 2019.
- [22] Saiharsha, B., B. Diwakar, R. Karthika, and M. Ganesan. "Evaluating Performance of Deep Learning Architectures for Image Classification." In 2020 5th International Conference on Communication and Electronics Systems (ICCES), pp. 917-922. IEEE, 2020.
- [23] Yang, Chao-Lung, Zhi-Xuan Chen, and Chen-Yi Yang. "Sensor classification using convolutional neural network by encoding multivariate time series as two-dimensional colored images." *Sensors* 20, no. 1 (2020): 168.

# Multimodal Emotion Recognition Using Different Fusion Techniques

Gokul Subramanian, Kotapati Prathyusha, Niranjan Cholendiran, Noviya Balasubramanain, Aravinth J  
*Department of Electronics and Communication Engineering, Amrita School of Engineering, Coimbatore,  
 Amrita Vishwa Vidhyapeetham, India.*

gokulrockstars@gmail.com, prathyusha.kotapati31@gmail.com, niranjancholendiran@gmail.com,  
 noviya2000@gmail.com, j\_aravinth@cb.amrita.edu

**Abstract**— Human beings have the ability to understand and visualize various emotions on a daily basis. This could be done by noticing various features such as facial muscle movements, speech, hand gestures, etc. The automated emotion recognition is an important issue and has also been a lively research topic for the modern time. At the moment, several research workers have taken part in inheriting two or more unimodals for better understanding. This paper shows an approach for emotion recognition that uses three modalities: facial images, audio signals, and electroencephalogram (EEG) signals from FER and Ck+, RAVDESS and SEED-IV datasets respectively. Finally, various fusion techniques were approached and each of these fusion methods gave different results. The maximum accuracy of 71.24% was obtained with help of an autoencoder approach when combined with SVM classifier.

**Keywords**—*Emotion recognition, Multimodal, Feature fusion, Autoencoder, Convolutional neural network (CNN), Deep learning, Electroencephalogram (EEG)*

## I. INTRODUCTION

Emotion is closely related to our daily life as it involves the act of selection and decision making process, which is mainly involved by several external influences such as type of personality, social role, etc. So, emotion analysis is a process by which the emotional state of a person can be easily determined. Each of these emotions has unique features and specifications which can be observed for classification and the key step for machines to identify emotions on their own has been handled by pattern recognition communities for developing a suitable human-machine interaction [1]. Several modalities can be used for this, but we have used EEG, as it directly involves brain signals and gives a sense of authenticity, face and audio, as the emotion can be identified effortlessly by merely looking at the person and listening to their voice. In recent times many researches have manifested that combination of modalities, has shown better performance compared to unimodals in emotion recognition. Hence, in our work, we have approached multimodal emotion recognition instead of unimodal because multimodal involves different datasets and shows the diversity in the features extracted, thereby giving us the best possible result. So, the combined effort of these three modality will deliver a genuine outcome as the comparative study [2] shows that fused multimodals are much better than unimodals. So, four different feature fusion approaches are tried in this work. But, there is a possibility that the features might be of uneven sizes. One main reason is that these datasets are gathered from multiple sources and

they do not share the same origin. After performing feature fusion in four different ways, results were compared with each other.

The layout of the paper is as follows: In the Section.II the available research similar to our objective has been discussed followed by the Section.III describing our proposed methodology in order to obtain the target. Then comes the experimental arrangement and the results for the same in Section.IV, finally performance enhancement techniques has been discussed in Section.V.

## II. RELATED WORKS

Facial expression is a prominent way of exhibiting ourselves as humans. When it comes for an Human-Computer Interaction, facial features have to be extracted before processing. Silva et al. [3] extracted Active Shape Model (ASM) features from MUG dataset and then tried different classifiers: Linear Discriminant Analysis (LDA), K-Nearest Neighbours (KNN), Artificial Neural Network (ANN) and compared the performance of these classifiers. It is proven that LDA showed better performance. Gogić et al.[4] used local binary features (LBF ) for feature extraction and a simple neural network which consists of one layer for classification. They have tried the proposed method on different datasets, among them CK+ dataset showed the best result of 96.48%. Similarly, Mehendale et al. [5] used CNN for both feature extraction and classification. A neural network of 4 hidden layers four filters have been used. This method was performed on different datasets and NIST showed the better result of 96%. Speech is also an evident way of representing ourselves. So, Kadiri et al. [6] used datasets from two different languages and extracted almost 4-7 different emotions with an overall accuracy of about 91.67% and 86.16% for IIIT-H (Telugu) and EMO-DB (German) respectively by extracting 3 main different features and plotting them in a 3D space. This paper proves that language boundaries are not necessary to derive the emotion of a person. Meanwhile, Nantasri et al. [7] used the RAVDESS dataset and extracts mel-frequency cepstral coefficient (MFCC) values to determine the emotion of an individual. 20 MFCC values have been extracted to improve the efficiency. The result shows an accuracy up to 82.3%, which seems satisfying for minimally extracted features.

Meanwhile, Y Li et al. [8] proposed a novel neural network, bi-hemispheric discrepancy model (BiHDM) to learn the bi-hemispheric discrepancy information in EEG emotion signals using three datasets viz. SEED, SEEDIV, MPED. The accuracy of Bi-HD Model classification for both subject-dependent and subject-independent experiments are 74.35% and 69.03% respectively. Apparently, Zheng et al. [9] used SEED Dataset and trained the DBN models with the

DE features from multichannel EEG data. This DBN-based method selects meaningful critical channels and frequency bands through the weight distributions of the trained DBNs and has designed different profiles of electrode sets. They also compared the performance of deep models with shallow models and the results were provided. Ozdemir et al. [10] introduced CNN into the process of emotional EEG recognition. The innovation of this method is to adjust the convolution kernel of the CNN to adapt to the input of EEG signals and the accuracy of 0.8579 was achieved.

From all these surveys, it is visible that successful emotion classification has been done using some verbal, non-verbal and physiological signals. Even though they performed well as unimodal inputs, many researchers have used more than one of them for emotion recognition. For the multimodal analysis different fusion techniques were implemented. The work by Zheng et al. [11] proved that unimodal results can be combined using fusion techniques to improve the performance of the classifier. They were able to achieve the unimodal accuracy for EEG signals and eye gaze as 73.94% and 58.07% respectively. But by combining those unimodal scores using 2 different fusion techniques, they were able to achieve an accuracy of 75.6% using feature level fusion technique and 74.9% using Decision level fusion technique. There are different types of feature fusion techniques that have been widely used. Liu et al. [12] used pre-trained models to extract linear-frequency cepstral coefficients (LFCC) features from raw EEG signals. The raw EEG signals were preprocessed at first. Then LFCC and ResNet-50 features were extracted from preprocessed EEG signals. Both the extracted features from Resnet-50 and LFCC, were concatenated in a traditional method for feature fusion. Eventually, all the fused features are fed into several different classifiers to recognize emotions. Similarly Nemati et al. [13] studied, a hybrid fusion method for multimodal emotion recognition in which the audio and visual modalities are fused using a latent space linear map and then, their projected features were fused with the textual modality. Besides, the implementation result shows that textual and audiovisual contents have improved the performance.

By summing up the literature survey, our aim in this work is to use multiple modalities: EEG signals, Facial expressions and Speech audio signals to classify among four different emotions. The methodology we have adopted for this work is to extract features of all three modalities using CNN and to combine those resultant features by trying out different feature fusion techniques aiming to improve the performance score after classification. The different approaches for feature fusion that have been widely used are discussed below.

### III. METHODOLOGY

To commence the work, various datasets were gathered for various modalities. Then these datasets were pre-processed and annotated separately by following a standard annotation across all the datasets. Later, these pre-processed raw data were sent to separate CNNs for feature extraction. These extracted features should be fused together in order to obtain the precise emotion. So, feature fusion was done and four different fusion methodologies were executed namely Normal concatenation, CNN Fusion, Single Autoencoder (AE) and Multiple autoencoder fusion approach as shown in Fig.1.

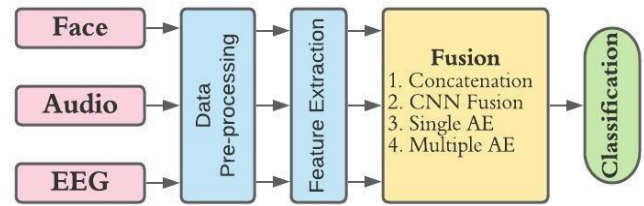


Fig.1. Block Diagram of the experimentation process.

#### A. Data preprocessing

The extracted raw data might have unwanted interference or might not be in a directly usable form. So, such raw data could deteriorate the model performance. In order to overcome this, data pre-processing is widely performed before extracting the features. In our work we have performed diverse data pre-processing techniques for all three modalities. The data pre-processing steps which we have performed includes: Data transformation, Data sampling, Data reduction, Data augmentation, Data annotation, Data normalization. For a particular emotion in each modality, we have restricted it with 180 files, in order to reduce the analysis time. The annotation was done as per the TABLE I.

TABLE.I. DATA ANNOTATION

Emotion	Label
Neutral	0
Sad	1
Fear	2
Happy	3

For face, the data was transformed from (.jpg) images to pixels. The pixel values were normalized between [0, 1] by dividing the pixel values by 255. For audio, each signal was sampled at the rate of 1500 kHz with an offset of 3.5 seconds. The number of audio files in 'Neutral' was comparatively lesser than others, so there was a need for augmentation. In this process, the number of files was replicated by adding some noise, thereby increasing the files to 180. Further, these values were normalized using Min-Max Normalization technique as in (1). Likewise for EEG, 12 among 62 electrodes were chosen, as it displays a stable pattern across all emotions [14]. Further, the signals from the electrodes were sampled at the rate of 200 Hz and normalized with the help of Min-Max Normalization [15].

$$A' = \frac{A - \min(A)}{\max(A) - \min(A)} \times (D - C) + C \quad (1)$$

Where,  $A'$  contains Min-Max Normalized data, if predefined boundary is  $[C, D]$ , the range of original data is  $A$ .

#### B. Feature Extraction

Feature extraction involves extracting the right information from the raw data before classification. In order to increase the accuracy of the training model, feature extraction plays a pivotal role as it is necessary. It gets processed by reducing the vector dimension [16] and

simultaneously removes the redundant data values. In our work, for feature extraction we have adopted a deep learning method called Convolution Neural Network (CNN). CNN consists of different layers in which each layer is formed by performing different operations with the input. CNN is specialized in selecting the right features from the input data. Many research works using CNNs [17, 18, 19] were performed and have achieved better results[20]. Similarly In our work we have designed a two dimensional CNN for facial and EEG feature extraction, and an one dimensional CNN for audio feature extraction as shown in Fig.2.

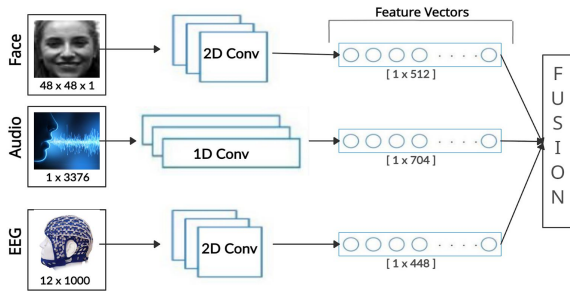


Fig.2. Block diagram of feature extraction using CNNs.

Every CNN model was designed in such a way that the dimension of the input array was scaled down gradually by not missing any valuable information. For facial images, a two dimensional CNN model was designed to extract features from a dataset of grayscale images. This CNN takes an input array of size (48,48,1) and the resultant features were scaled down to an array of size (1,512) after flattening. For audio signals, an one dimensional CNN model was designed to extract features from this sampled raw audio signal. The designed CNN model takes the input array of size (1,3376) and the resultant features were scaled down to an array of size (1,704) after flattening. For EEG, after removing the stable patterns of electrodes, an input array of size (12,1000) is given to the CNN. The resultant features were scaled down to an array of size (1,448) after flattening.

C. Fusion

Data fusion is the way toward coordinating different information sources to deliver more reliable, precise, and useful information than that given by any individual information source. Data fusion processes are often categorized as pixel-level, feature-level and decision-level, based on the processing stage at which the fusion takes place [21, 22]. Description of them has been shown in Table II.

TABLE.II. TYPES OF DATA FUSION

S.No	Data Fusion	Description
1	Pixel-level fusion	Raw data are fused
2	Feature-level fusion	Extracted features from raw data are fused
3	Decision-level Fusion	Final score of each modalities are fused

In our work we have adopted feature-level fusion because of its essential advantage of detecting the correlated feature values generated by different sources, therefore identifying

the best set of features can result in a better performance. For our feature fusion we have tried out four different feature fusion techniques:

1. **Simple Concatenation:** In our first technique the extracted features from different modalities were simply concatenated one after another to form a single fused vector as shown in fig.3. The face features with dimension (1,512), audio features with dimension (1,704) and EEG features with dimension (1,448) were concatenated in the order of face, audio and EEG respectively. The resultant fused features of dimension (1,1664) were passed into a classifier and the results were noted.

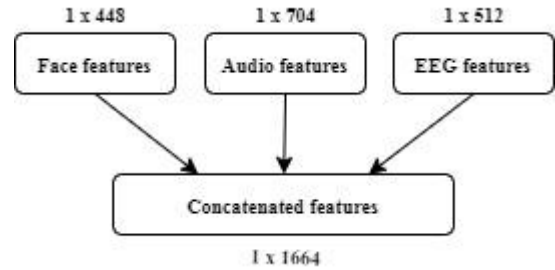


Fig.3. Block diagram for simple concatenation.

2. **CNN fusion:** For our second technique instead of directly passing the concatenated vector to the classifier, In our first technique the extracted features from different modalities were simply concatenated one after another to form a single fused vector as shown in Fig 4. The face features with dimension (1,512), audio features with dimension (1,704) and EEG features with dimension (1,448) were concatenated in the order of face, audio and EEG respectively. The resultant fused features of dimension (1,1664) were fused the individual features using CNN [23]. The resultant fused vectors from the CNN were of dimension (1,320). These vectors were passed into the classifiers and the results were noted.

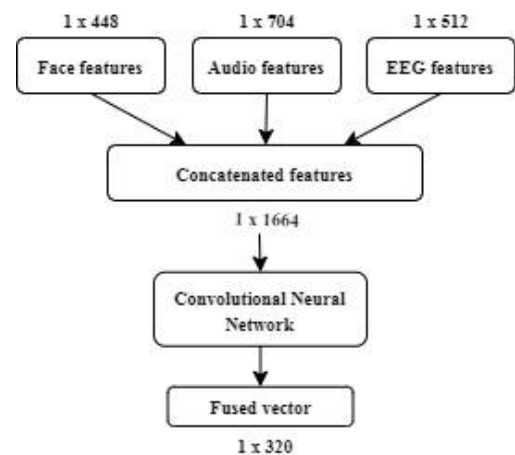


Fig.4. Block diagram for CNN fusion.

3. **Single autoencoder fusion:** Similarly, for our third technique, we have replaced an autoencoder in place of a CNN for fusion [24]. Autoencoder is a type of unsupervised neural network which is widely used in

dimensionality reduction. The basic working of an autoencoder is, it first encodes the input and forms a latent feature called bottleneck. Later the bottleneck will be decoded to construct back the input. In our third fusion technique an autoencoder was trained with the simple concatenated feature as shown in. Later the bottleneck of that autoencoder was extracted as shown in Fig 5. and that was the fused vector used for classification. The resultant bottleneck of the autoencoder was of dimension (1,64) which was passed into classifiers and the results were noted.

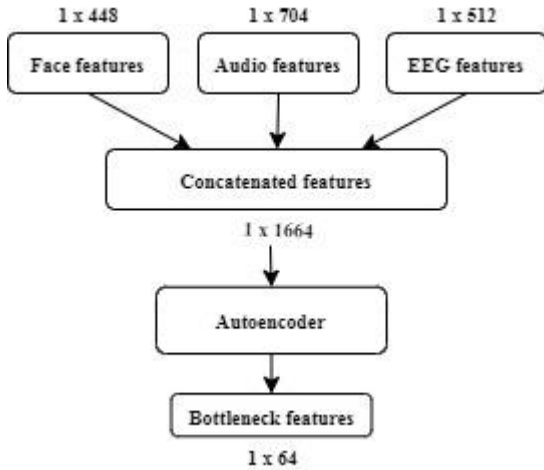


Fig.5. Block diagram for Single autoencoder fusion.

4. **Multiple autoencoder fusion:** Similarly, For the final technique, we used multiple autoencoders for fusion as shown in Fig 6. Initially, three autoencoders were used to extract the encoded version of all three individual modality features which were of a standard dimension (1,64). Later they were concatenated and sent into a separate autoencoder. The bottleneck of that autoencoder was of dimension (1,128) which is the final fused vector used for classification. noted.

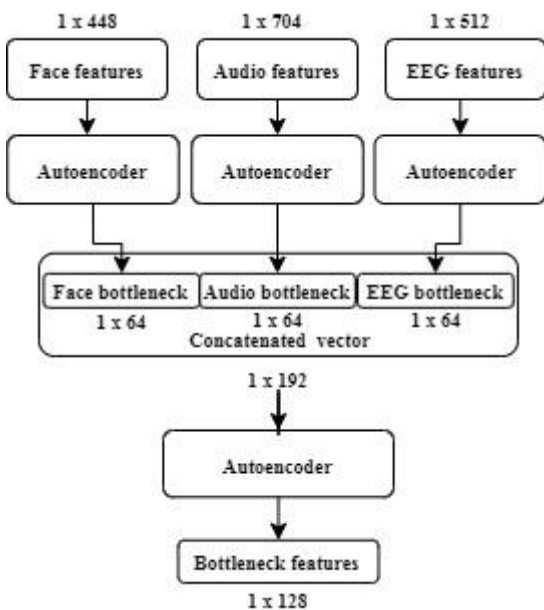


Fig.6. Block diagram for Multiple autoencoder fusion.

#### D. Classification

After fusing the features using different techniques, they were split into 70% for testing and 30% for training. Later the training features were used to train three machine learning classifiers:

1. Support Vector Machine (SVM)
2. K Nearest neighbors(KNN)
3. Random Forest.

Further, the respective testing features were used to analyse the performance of each model and the classification accuracy of them were noted. The results of all four fusion techniques were tabulated for comparison.

## IV. RESULTS AND DISCUSSION

#### A. Dataset description

Four datasets have been used for emotional recognition. Ck+ and FER'13 for facial emotion recognition, RAVEDSS for audio and SEED IV for EEG based emotion recognition. The description about each dataset for individual modality has been given below:

- **Facial:** we have availed images from two different datasets: CK+ [25] and FER'13[26]. All the images are  $48 \times 48$  pixel, gray-scale images of faces. Total of 123 subjects have participated in CK+ experimental setup. Emotions include neutral, sadness, happiness, fear, Anger, contempt and disgust. The FER'13 dataset consists of 35k images categorized into eight different emotions namely, happiness, neutral, sadness, anger, surprise, disgust, fear.
- **Audio:** we have taken a publicly available dataset RAVEDSS [27]. In this dataset 24 voice actors (12 male, 12 female) performed 60 trials each trying out 7 different emotions. Each raw audio was around 2-3 seconds long and sampled at the rate of 1500 Hz.
- **EEG:** we have taken the SEED IV dataset [28]. It contains EEG signals extracted from 15 participants. For each participant, 3 sessions were performed on different days, and each session contains 24 trials. Signals were recorded with the help of 62-channel ESI Neuro scan system.

#### B. Experimental setups and results

The experiment was done using an open source Google product Collaboratory, which is based on Jupyter Notebooks. The interpreter language Python version 3.8 and its inbuilt libraries were used in this experiment. Initially we have extracted the feature vectors from three individual modalities and fused them using four different fusion techniques. Finally the fused vector has been passed into three classifiers namely, SVM, KNN and Random forest. Performance of the four fusion techniques has been evaluated based on accuracy. Comparative results of the four fusion techniques we performed are presented in TABLE III.

From the results it is visible that the SVM model has shown better results comparatively, and the fusion technique using multiple autoencoders has shown a superior performance compared to other tested techniques. The combination of SVM and multiple autoencoder fusion has outperformed other possible combinations. It has achieved a higher accuracy of 71.24%. Followed by simple

concatenation and SVM has performed well with an accuracy of 67.59%. CNN fusion has the lowest accuracy compared to all fusion techniques.

TABLE III. CLASSIFICATION RESULTS FOR FUSION TECHNIQUES

Fusion technique	Accuracy (%)		
	SVM	KNN	Random Forest
Simple concatenation	67.59	55.55	64.81
CNN fusion	52.77	42.12	52.77
Single autoencoder fusion	64.35	54.62	64.351
Multiple autoencoder fusion	71.24	61.11	61.57

## V. CONCLUSION

This paper mainly demonstrates the way in which emotion can be discovered for putting together various features from audio, face and EEG. Later, four feature fusion methods were done and compared, out of which one autoencoder fusion outperformed the others with an accuracy of 71.24%. But, the performance of the work can be enhanced in various ways. One way of doing this is by converting the time series data of audio and EEG to an image format before extracting the features [29]. Additionally, due the diversity in data used, our unimodals might have missed the correlation with others. Hence, performance would improve if the multimodal data were extracted at the same instant, i.e. Dataset containing facial, audio and EEG features extracted at the same instance using common stimuli, from a familiar subject [30].

## REFERENCES

- [1] Veni, S., & Thushara, S. (2017). Multimodal approach to emotion recognition for enhancing human machine interaction-a survey. *International Journal on Advanced Science, Engineering and Information Technology*, 7(4), 1428-1433.
- [2] Shrivastava, V., Richhariya, V., & Richhariya, V. (2018, December). Puzzling Out Emotions: A Deep-Learning Approach to Multimodal Sentiment Analysis. In 2018 International Conference on Advanced Computation and Telecommunication (ICACAT) (pp. 1-6). IEEE.
- [3] Silva, C., Sobral, A., & Vieira, R. T. (2014, October). An automatic facial expression recognition system evaluated by different classifiers. In X Workshop de Visão Computacional, at Uberlândia, Minas Gerais, Brazil (pp. 208-212).
- [4] Gogić, I., Manhart, M., Pandžić, I. S., & Ahlberg, J. (2020). Fast facial expression recognition using local binary features and shallow neural networks. *The Visual Computer*, 36(1), 97-112.
- [5] Mehendale, N. (2020). Facial emotion recognition using convolutional neural networks (FERC). *SN Applied Sciences*, 2(3), 1-8.
- [6] Kadir, S. R., & Alku, P. (2020). Excitation Features of Speech for Speaker-Specific Emotion Detection. *IEEE Access*, 8, 60382-60391.
- [7] Nantasri, Panuwit, Ekachai Phaisangittisagul, Jessada Karnjana, Surasak Boonkla, Suthum Keerativittayanun, Anocha Rugchatjaroen, Sasiporn Usanavasin, and Takahiro Shinozaki. "A Light-Weight Artificial Neural Network for Speech Emotion Recognition using Average Values of MFCCs and Their Derivatives." In 2020 17th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON), pp. 41-44. IEEE, 2020.
- [8] Li, Yang, Lei Wang, Wenming Zheng, Yuan Zong, Lei Qi, Zhen Cui, Tong Zhang, and Tengfei Song. "A novel bi-hemispheric discrepancy model for eeg emotion recognition." *IEEE Transactions on Cognitive and Developmental Systems* (2020).
- [9] Zheng, W. L., Zhu, J. Y., & Lu, B. L. (2017). Identifying stable patterns over time for emotion recognition from EEG. *IEEE Transactions on Affective Computing*, 10(3), 417-429.
- [10] Ozdemir, M. A., Degirmenci, M., Izci, E., & Akan, A. (2020). EEG-based emotion recognition with deep convolutional neural networks. *Biomedical Engineering/Biomedizinische Technik*, 1(ahead-of-print).
- [11] Zheng, W. L., Zhu, J. Y., & Lu, B. L. (2014, July). Multimodal emotion analysis in response to multimedia. In 2014 IEEE International Conference on Multimedia and Expo Workshops (ICMEW) (pp. 1-2). IEEE..
- [12] Liu, N., Fang, Y., Li, L., Hou, L., Yang, F., & Guo, Y. (2018, April). Multiple feature fusion for automatic emotion recognition using EEG signals. In 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 896-900). IEEE..
- [13] Nemati, S., Rohani, R., Basiri, M. E., Abdar, M., Yen, N. Y., & Makarenkov, V. (2019). A hybrid latent space data fusion method for multimodal emotion recognition. *IEEE Access*, 7, 172948-172964.
- [14] Zheng, Wei-Long, Jia-Yi Zhu, and Bao-Liang Lu. "Identifying stable patterns over time for emotion recognition from EEG." *IEEE Transactions on Affective Computing* 10, no. 3 (2017): 417-429.
- [15] L. Al Shalabi and Z. Shaaban, "Normalization as a Preprocessing Engine for Data Mining and the Approach of Preference Matrix," 2006 International Conference on Dependability of Computer Systems, Szklarska Poreba, 2006, pp. 207-214, doi: 10.1109/DEPCOS-RELCOMEX.2006.38.
- [16] Prasad, Sandra, and J. Aravindh. "Serial multimodal framework for enhancing user convenience using dimensionality reduction technique." In 2016 International Conference on Circuit, Power and Computing Technologies (ICCPCT), pp. 1-6. IEEE, 2016.
- [17] Gao, Zhongke, Xinmin Wang, Yuxuan Yang, Yanli Li, Kai Ma, and Guanrong Chen. "A channel-fused dense convolutional network for EEG-based emotion recognition." *IEEE Transactions on Cognitive and Developmental Systems* (2020).
- [18] Ahmed, T. U., Hossain, S., Hossain, M. S., ul Islam, R., & Andersson, K. (2019, May). Facial expression recognition using convolutional neural network with data augmentation. In 2019 Joint 8th International Conference on Informatics, Electronics & Vision (ICIEV) and 2019 3rd International Conference on Imaging, Vision & Pattern Recognition (icIVPR) (pp. 336-341). IEEE.
- [19] Kiranyaz, Serkan, Onur Avci, Osama Abdeljaber, Turker Ince, Moncef Gabbouj, and Daniel J. Inman. "1D convolutional neural networks and applications: A survey." *Mechanical Systems and Signal Processing* 151 (2019): 107398.
- [20] Saiharsha, B., Diwakar, B., Karthika, R., & Ganesan, M. (2020, June). Evaluating Performance of Deep Learning Architectures for Image Classification. In 2020 5th International Conference on Communication and Electronics Systems (ICES) (pp. 917-922). IEEE.
- [21] Pohl, C., & Van Genderen, J. L. (1998). Review article multisensor image fusion in remote sensing: concepts, methods and applications. *International journal of remote sensing*, 19(5), 823-854.
- [22] J. Aravindh and S. Valarmathy, "Multi classifier-based score level fusion of multi-modal biometric recognition and its application to remote biometrics authentication," *Imaging Sci. Journal* vol. 64, no. 1, 2016
- [23] Williams, Jennifer, Steven Kleinogesse, Ramona Comanescu, and Oana Radu. "Recognizing emotions in video using multimodal dnn feature fusion." In *Proceedings of Grand Challenge and Workshop on Human Multimodal Language (Challenge-HML)*, pp. 11-19. 2018.
- [24] Nguyen, Dung, Duc Thanh Nguyen, Rui Zeng, Thanh Thi Nguyen, Son N. Tran, Thin Nguyen, Sridha Sridharan, and Clinton Fookes. "Deep Auto-Encoders with Sequential Learning for Multimodal



- Dimensional Emotion Recognition." arXiv preprint arXiv:2004.13236 (2020).
- [25] Lucey, P., Cohn, J. F., Kanade, T., Saragih, J., Ambadar, Z., & Matthews, I. (2010, June). The extended cohn-kanade dataset (ck+): A complete dataset for action unit and emotion-specified expression. In 2010 IEEE computer society conference on computer vision and pattern recognition-workshops (pp. 94-101). IEEE.
- [26] Giannopoulos, P., Perikos, I., & Hatzilygeroudis, I. (2018). Deep learning approaches for facial emotion recognition: A case study on FER-2013. In Advances in hybridization of intelligent methods (pp. 1-16). Springer, Cham.
- [27] Livingstone, S. R., & Russo, F. A. (2018). The Ryerson Audio-Visual Database of Emotional Speech and Song (RAVDESS): A dynamic, multimodal set of facial and vocal expressions in North American English. PloS one, 13(5), e0196391.
- [28] Zheng, Wei-Long, Wei Liu, Yifei Lu, Bao-Liang Lu, and Andrzej Cichocki. "Emotionmeter: A multimodal framework for recognizing human emotions." IEEE transactions on cybernetics 49, no. 3 (2018): 1110-1122.
- [29] Lyon, R. F., Rehn, M., Walters, T., Bengio, S., & Chechik, G. (2013). U.S. Patent No. 8,463,719. Washington, DC: U.S. Patent and Trademark Office.
- [30] Soleymani, M., Lichtenauer, J., Pun, T., & Pantic, M. (2011). A multimodal database for affect recognition and implicit tagging. IEEE transactions on affective computing, 3(1), 42-55.

# Modelling And Simulation Of High Flux Hemodialyzer Membranes Of Different Porosities To Identify The Optimal Membrane Design

Ahana Fatima Alex  
*Department of Electrical and  
 Electronics Engineering*  
*Manipal Academy of Higher Education  
 (MAHE)*  
 Dubai, UAE  
 Ahana.Fatima@dxb.manipal.edu

R Vinoth  
*Department of Electronics and  
 communication Engineering*  
*Manipal Institute of Technology (MIT)*  
*MAHE*  
 Manipal, India  
 vinoth.nair@manipal.edu

Ravishankar Dudhe  
*Department of Electrical and  
 Electronics Engineering*  
*Manipal Academy of Higher Education  
 (MAHE)*  
 Dubai, UAE  
 ravishankar.dudhe@manipaldubai.com

**Abstract**—High flux hemodialyzer membranes of different average porosities were modelled. Diffusion and convection property of a toxin molecule through the membrane was observed through simulation using Finite element method. In this study, a porous membrane having a porosity of 0.15 was compared to that of 0.3 membrane porosity. Different synthetic polymers having such porosities were thus considered for this purpose. Effective diffusion and convection of a toxic molecule through the porous membrane over a period of time was modelled and then analysed to draw conclusion on the optimal porosity value. PMMA with a porosity of 10-20% shows better performance in terms of toxin clearance as well as endotoxin removal.

**Keywords**— High flux membrane, Porosity, Diffusion, Convection, Fick's Law, Diffusion coefficient, Dynamic viscosity, Finite Element Analysis.

## I. INTRODUCTION

According to the latest statistics of World Health Organization (WHO) and International society of kidney federation, more than 900 million people around the globe are affected by some or the other kind of kidney diseases which is roughly double the number of diabetes patients. Kidney diseases, unlike any other ailment are asymptomatic and could be diagnosed only when its functionality falls below 25%. Hence hemodialysis becomes the only effective treatment until a healthy kidney is available for transplant [1].

Because of this fact, Hemodialyzer and the dialyzer membranes are of great research interest. The membrane depending on their pore sizes are called as low flux membrane or high flux membrane. Superior membranes such as medium cut-off membranes and high cut-off membranes are also currently gaining popularity. However, an optimal dialyzer design that incorporates all the parameters involved in the process- both functional and geometric is yet to be modelled. Many studies and researches are carried out in this optimization processes and many theories have been proposed for the same [2, 3].

The material used for the membranes and its porosity are also of great importance as it helps in improving the dialyzer performance and thus in optimization. Different synthetic polymers which are biocompatible are in use for the exiting dialyzer membranes but membrane fouling has not been eliminated completely. Therefore more

membrane materials such as Activated Carbon(AC)/ Polyethersulfone/ Polyvinylpyrrolidone Mixed Matrix Membrane (MMM) which helps in attaining endotoxin free dialysate [2] and Polymethylmetacrylate (PMMA) are widely being studied. The porosity of such membranes also plays a great role as in highly removing the uremic toxin- which is the primary need of any dialyzer membrane.

Researches and studies have proposed the use of an AC MMM for achieving the first time, endotoxin free dialysate solution and high elimination of toxins using a single membrane. With the use of such a membrane, endotoxin removal from the dialysate solution could be improved 10 times than that of commercial fibers. Protein bound toxins could be removed easily using MMM [4].

While various other studies focused mainly on factors leading to protein fouling. Such studies took measures that completely avoid protein fouling such as careful observation of rate of clearance of uremic toxins in order to see whether the parameters introduced to avoid membrane fouling has affected the clearance rate [5]. Many papers threw insights on various materials that could be effectively used for dialysis treatment and various techniques to alter the membrane composition for better results. Membrane manufacturing and major steps involved are also accounted [6].

Research papers also discussed about the membrane fouling and how it blocks the pores. Membrane modification and ways to cleanse and distill these membranes were considered too [7]. A novel anti-fouling membrane was proposed. The paper claims to be an effective way to avoid fouling without increasing the mass transfer resistance [8].

This paper focuses on the modelling and simulation of high flux membranes, having different porosities and material properties. From the literature, PMMA is assumed to show greater results compared to existing membrane materials. Hence, the two membranes in the study will follow close approximation of PMMA and that of commonly used polyamix. Uremic toxin molecule namely urea is being passed through a very thin membrane into a dialysate solution and the diffusion and convection property of the molecule is determined and thus the effectiveness of membrane is identified. The

results thus obtained could be incorporated in the optimal dialyzer design.

COMSOL Multiphysics- the finite element analysis software is used here for the design and analysis of the diffusion and convection properties of different high flux membranes. COMSOL Multiphysics is a cross media used for simulating designs, processes and infra-structures in all areas of physics and engineering. It helps in replicating the original model which could be fully done at user's freedom. Transport of diluted species in the chemical reaction module of COMSOL is used here. A time dependent study is considered so as to determine the concentration profile of the toxin through the porous membrane at different time instances. The exact process was repeated again for different membrane porosities. The major design equations for the diffusion and convection include Fick's law and partial differential equations [9, 10, 11].

The paper explains modelling and simulation of high flux membranes with different porosities and material properties so as to determine the better membrane material of the two. Section I show the importance of membrane material selection and how it helps in better toxin clearance and why membrane fouling should be avoided. Section II focuses on the methodology used in designing the porous membranes with different porosities for the diffusion and convection of a toxin molecule through them. It could be considered as two cases- case I simulation a porous membrane of material PMMA. PMMA has porosity of 10-20% [12, 13]. While case II involves the simulation of a porous membrane of material polyamix having around 30-45% porosity [10, 13, 14]. Urea having a diffusion co-efficient of  $1690.346 \times 10^{-12} \text{ m}^2/\text{s}$  is considered as the toxin molecule that falls on the porous membrane in both the cases. . Section III provides the results of the two simulations. Comparative study of the same provides the final result on which membrane is more effective. Conclusions are thus drawn from the obtained results.

## II. METHODOLOGY

The study is divided into two cases. Case one is design and analysis of a porous high flux hemodialyzer membrane having porosity 0.15 and has membrane properties of that of a PMMA membrane. Diffusion and convection of uremic toxin, urea is analyzed. Concentration of this toxin molecule is determined over a time period. Case two is design and analysis of a porous membrane having porosity 0.3 and membrane properties of polyamix. In case II the design process remains the same as in case I. In both the cases transport of dilutes species is considered for porous media design. Appropriate porosity values were given for the membrane material. Other input parameters of the membrane considered were density and dynamic viscosity.

In either case, chemical reaction module of COMSOL Multiphysics was used. Also, Transport of diluted species was the interface used and a time dependent study was

considered. Laminar flow was taken and concentration profiles were determined [10, 15].

### A. Case I

#### a) Design and simulation of a high flux hemodialyzer membrane having 15% porosity

Here, diffusion and convection of urea through the membrane was studied. For this purpose, a droplet of impurity was considered to fall on the membrane surface which separates the toxin from the dialysate compartment. As shown in fig 1 the diffusion of the toxin molecule through the very thin membrane of thickness 0.007mm and into the dialysate compartment is observed.

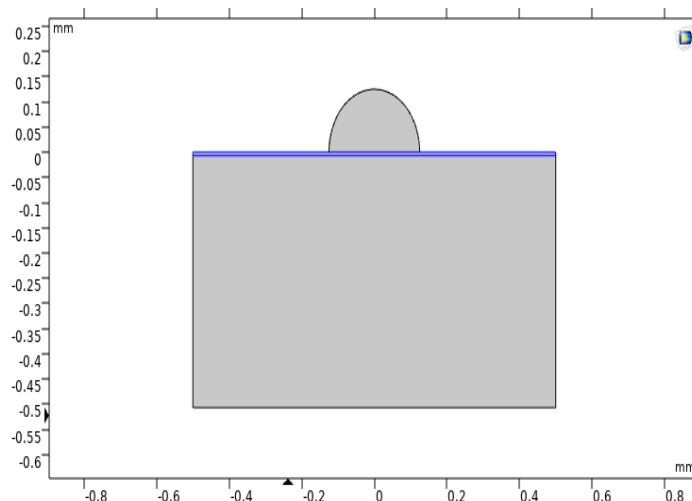


Fig. 1. Geometry of membrane (highlighted in blue), dialysate compartment and toxin molecule in COMSOL multiphysics

#### b) COMSOL Simulation

In this model only the urea clearance was considered across the membrane. Chemical engineering module of COMSOL Multiphysics is used here considering the toxin molecule to be a solvent mix of blood and urea. Therefore, Fick's law of this module can be made use of. The below mentioned simplified partial differential equation (PDE) describes the process of diffusion and convection in the blood and dialysate compartments [2, 10]

$$\nabla \cdot (-D_i \nabla c_i + c_i u) = 0 \quad \dots\dots 1$$

where  $c_i$  denotes the concentration of the toxin ( $\text{mol}/\text{m}^3$ ) in the respective phase.  $D_i$  denotes the diffusion coefficient ( $\text{m}^2/\text{s}$ ) in the liquid phase and  $u$  denotes the velocity ( $\text{m}/\text{s}$ ) in the respective liquid phase.

For different toxins, the diffusion coefficient is different and is calculated from [10, 13].

$$D = 1.62 \times 10^{-4} (MW)^{-0.552} \quad \dots\dots 2$$

Here, MW is the molecular weight of respective toxins. Toxin diffusion at the surface of the membrane, within the membrane and in the dialysate compartment is determined. Very fine meshing was done at the interfaces of toxin molecule and membrane, and membrane and dialysate compartment. Figure 2 depicts the overall meshing. Extremely fine meshing at the interfaces could also be seen. The concentration of the toxin is not much effected at time = 0 minute as no reaction takes place at that instant. Same can be seen in figure 3. Figure 4 and figure 5 shows the concentration profiles at time 2 and 4 minutes respectively.

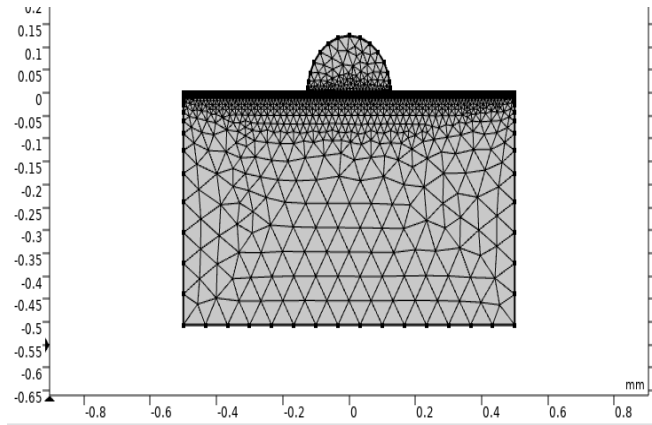


Fig. 2. Mesh created. Extremely fine meshing at the interfaces

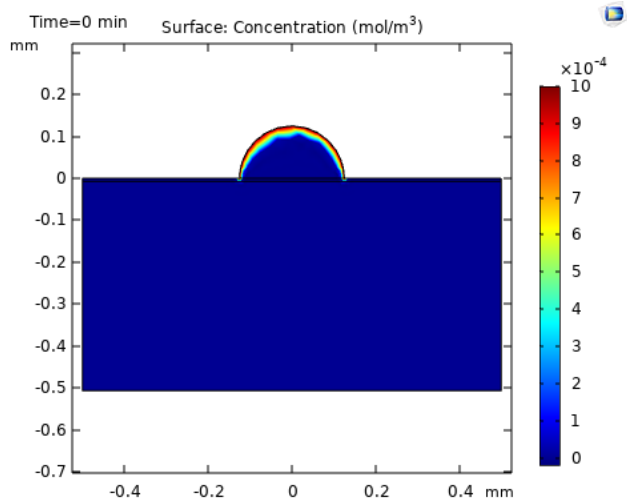


Fig. 3. Diffusion of toxin molecule at time = 0 min, Case I

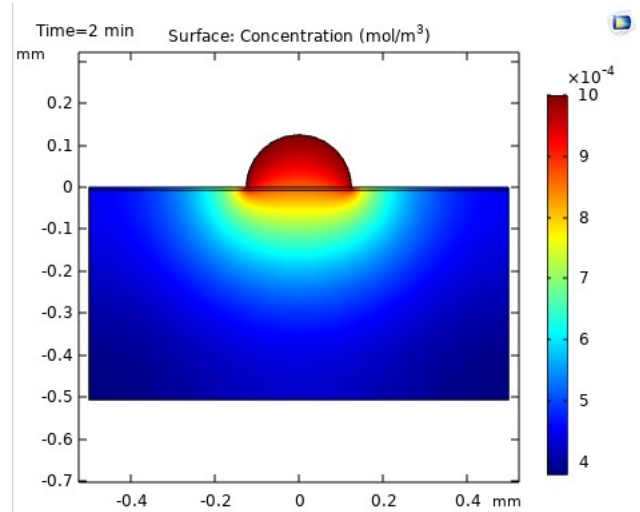


Fig. 4. Diffusion of toxin molecule at time = 2 min, Case I

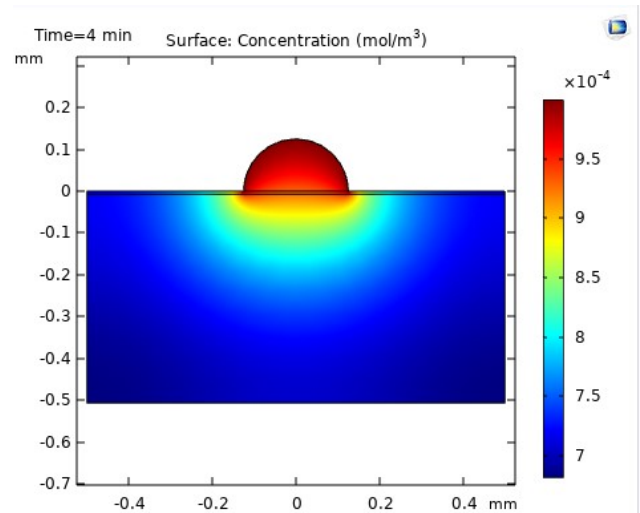


Fig. 5. Diffusion of toxin molecule at time = 4 min, Case I

**B. Case 2**

*a) Design and simulation of a high flux hemodialyzer membrane having 30% porosity*

Here too the exact process followed in case one was done but the membrane properties were altered. The dynamic viscosity, density and porosity of membrane was changed as per that of polyamix. 0.3 was the porosity given.

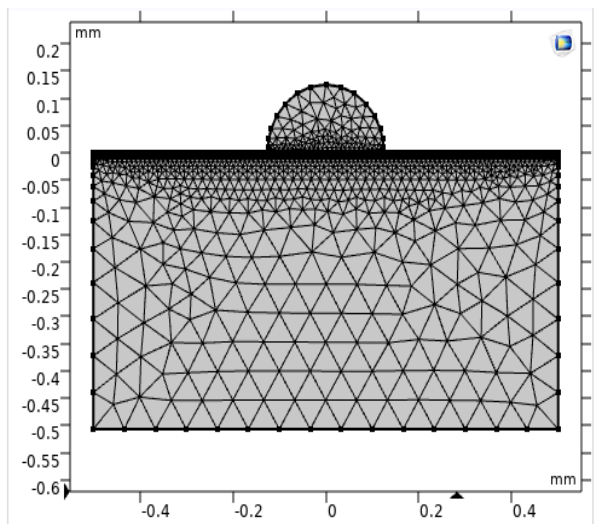


Fig. 6. Mesh created for case II

*b) COMSOL Simulation*

All the processes taken place in COMSOL simulation and parameter values other than the mentioned were similar as in case one. Here too very fine meshing was done for the interfaces. Figure [6] shows the meshing

Figure 7 shows the concentration of urea molecule at time = 0 minute and as in case I, here too at 2 and 4 minutes, the plots were observed which is as shown in figure 8 and figure 9 respectively.

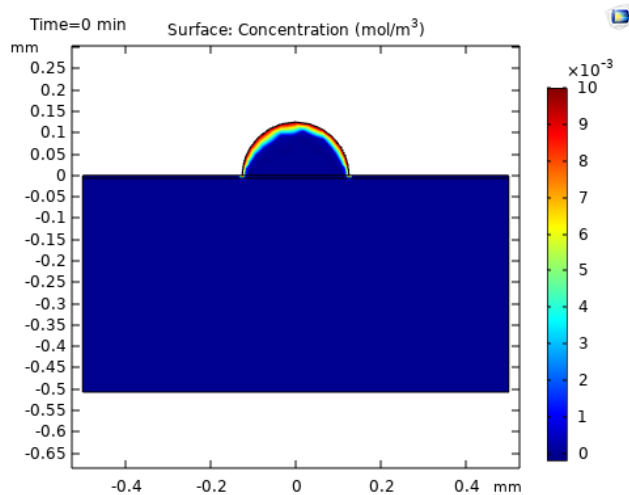


Fig. 7. Diffusion of toxin molecule at time = 0 min , Case II

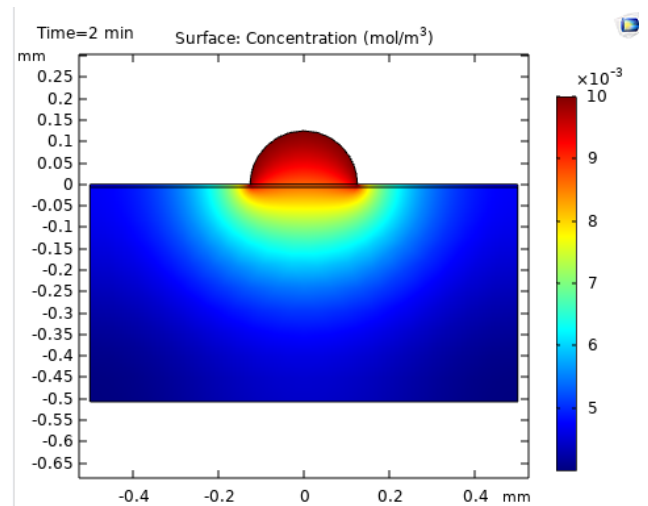


Fig. 8. Diffusion of toxin molecule at time = 2 min , Case II

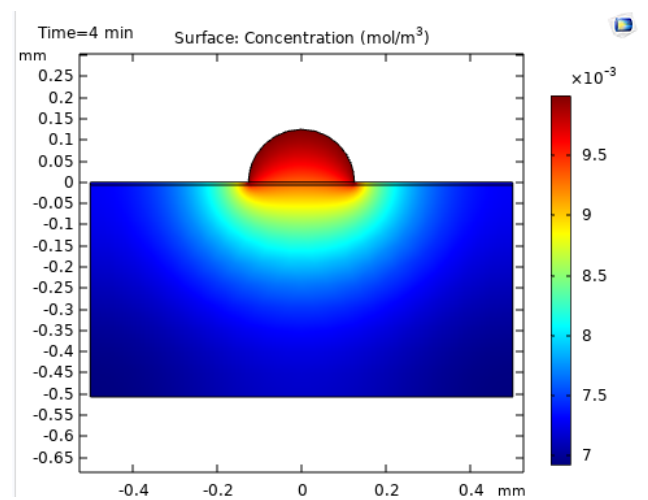


Fig. 9. Diffusion of toxin molecule at time = 4 min, Case II

III. RESULTS

The model was simulated using COMSOL multiphysics 5.4 under a time dependent study. Concentration of toxin molecule, urea was set to be same at initial condition and the diffusion effects were determined. From the results of the two cases, it is seen that the PMMA membrane with porosity 0.15 has faster diffusion properties as the initial concentration changes to  $0.00075 \text{ mol/m}^3$  at 1 minute time and diffuses completely around 5 minutes. While polyamix with 0.3 porosity, takes around 10 minute for the complete diffusion of the toxin molecule as shown in fig. 10-13. The fig 10 and 12 shows how the urea molecule diffuses along the dialysate compartment. At one minute time, only a portion of molecule is being diffused into the solution hence the initial concentration is slightly lower than for higher time ranges where the toxin further moves into the dialysate compartment through the porous membrane. Fig 11 and 13 shows the diffusion process that happen at the surface, the initiation of reaction, at the middle of dialysate compartment and finally at the bottom of the

compartment. These figures help us to have a clear idea on rate of diffusion as well as performance of the porous structure in toxin removal. The better performance of PMMA membrane thus obtained through simulation is in strong agreement with the data available from literature [10, 12-14]. Porosity and its effect on diffusion was thus correl

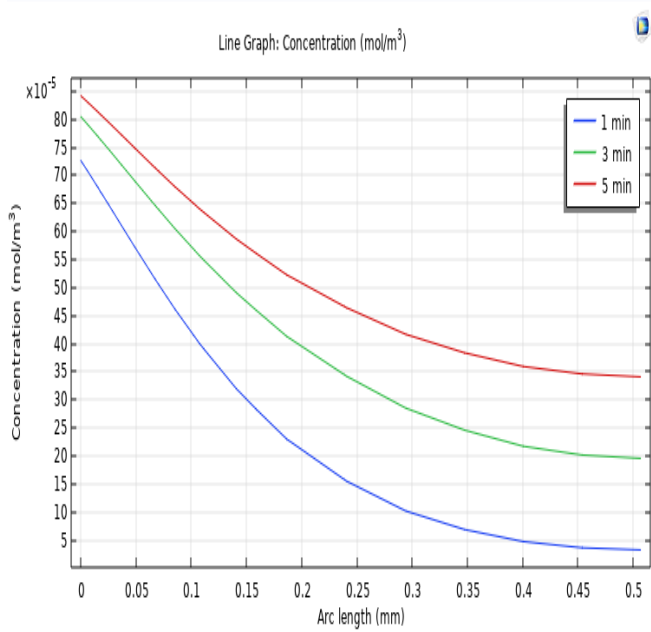


Fig. 10. Diffusion of toxin molecule, Concentration plot across the membrane, Case I

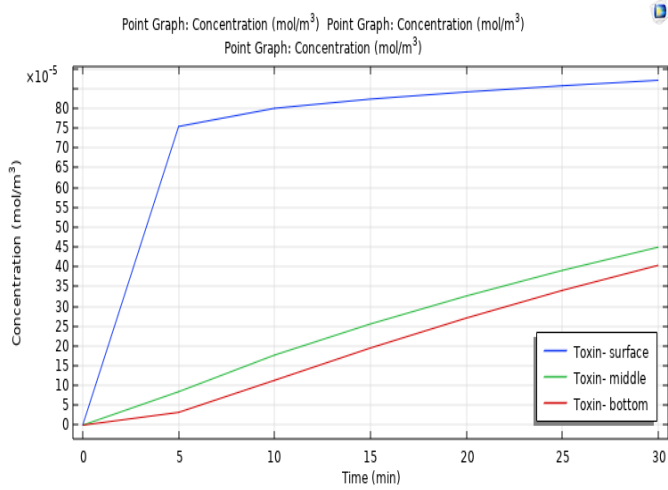


Fig. 11. Concentration Vs time, Case I

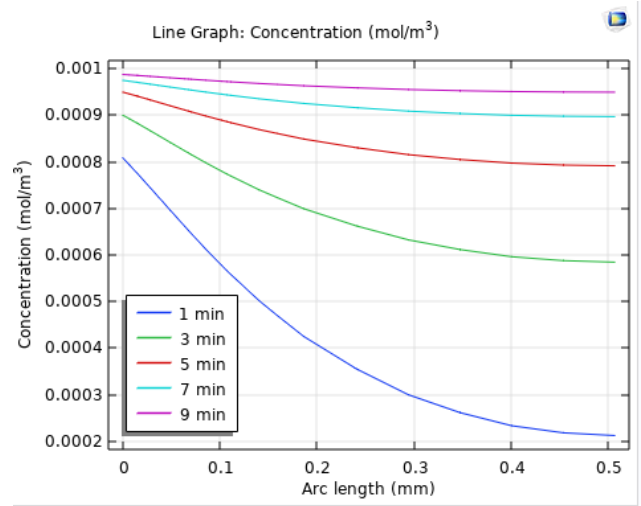


Fig. 12. Diffusion of toxin molecule, Concentration plot across the membrane, Case II

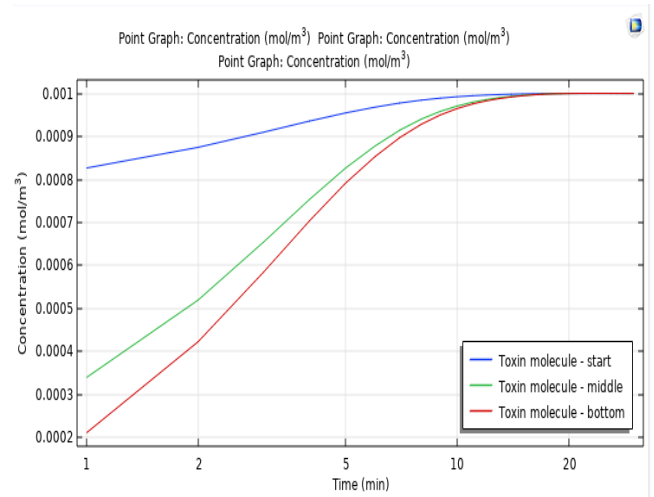


Fig. 13. Concentration Vs time, Case II

#### IV. CONCLUSION

The study implicates that, diffusion process- the basic working principle of hemodialyzer varies with different membranes of different material properties and porosities. The design of high flux hemodialyzer membrane using PMMA shows better toxin clearance at a faster rate with minimal fouling. This process of better material selection could in turn helps in increasing the overall dialyzer performance further helping one step ahead in the design optimization procedure. Therefore, further studies and researches on modelling high flux and medium cut off membranes using PMMA polymer is much recommended.

## REFERENCES

- [1] Namratha Krishnan .Hemodialysis Kinetics 101. Yale University School of Medicine Veterans Affairs Medical Center, 2017.
- [2] Donato D, Boschetti-de-Fierro A, Zweigart C, Kolb M, Elout S, Storr M, Krause B, Leyppoldt K, Segers P. "Optimization of dialyzer design to maximize solute removal with a two-dimensional transport model". *Journal of Membrane Science*. 2017 Nov 1; 541:519-28.
- [3] Donato D, Storr M, Krause B. "Design optimization of hollow fiber dialyzers to enhance internal filtration based on a mathematical model". *Journal of Membrane Science*. 2020 Mar 15; 598: 117690.
- [4] Waniewski J, Poleszczuk J, Pietribiasi M, Debowska M, Wojcik-Zaluska A, Zaluska W. "Impact of solute exchange between erythrocytes and plasma on hemodialyzer clearance". *Biocybernetics and Biomedical Engineering*. 2020 Jan 1; 40(1): 265-76.
- [5] Geremia I, Bansal R, Stamatialis D. "In vitro assessment of mixed matrix hemodialysis membrane for achieving endotoxin-free dialysate combined with high removal of uremic toxins from human plasma". *Acta Biomaterialia*. 2019 May 1; 90:100-11.
- [6] Boschetti-de-Fierro A, Beck W, Hildwein H, Krause B, Storr M, Zweigart C. "Membrane innovation in dialysis". In *Expanded Hemodialysis*, Karger Publishers, 2017 (Vol. 191, pp. 100-114).
- [7] Etinger A, Kumar WA, Soiefer L, Chun J, Singh P, Grossman E, Matalon A, Holzman RS, Meijers B, Lowenstein J. "The effect of isohydric hemodialysis on the binding and removal of uremic retention solutes". *PloS one*. 2018; 13(2).
- [8] Qiu H, Peng Y, Ge L, Hernandez BV, Zhu Z. "Pore channel surface modification for enhancing anti-fouling membrane distillation". *Applied Surface Science*. 2018 Jun 15;443:217-26.
- [9] Ronco C, et al. "Hemodialyzer: from macro-design to membrane nanostructure; the case of the FX-class of hemodialyzers". *Kidney International*. 2002 May 1; 61: S126-42.
- [10] Islam MS and Szpunar J. "Study of dialyzer membrane (Polyflux 210H) and effects of different parameters on dialysis performance". *Open Journal of Nephrology*. 2013 Sep 6; 3(03):161.
- [11] Haas K. "Modeling Blood Cell Concentration in a Dialysis Cartridge" (Doctoral dissertation, Worcester Polytechnic Institute).
- [12] Ronco C. "Polymethylmethacrylate: a flexible membrane for a tailored dialysis". *Contributions to nephrology*. 1998.
- [13] Yamashita AC, Sakurai K. "Dialysis membranes—physico-chemical structures and features". *Updates in Hemodialysis*. 2015 Sep 9:159-87.
- [14] Hedayat A, Szpunar J, Kumar NA, Peace R, Elmoselhi H, Shoker A. "Morphological characterization of the polyflux 210h hemodialysis filter pores". *International journal of nephrology*. 2012 Nov 6.
- [15] Ronco C, Brendolan A, Lupi A, Metry G, Levin NW. "Effects of a reduced inner diameter of hollow fibers in hemodialyzers". *Kidney international*. 2000 Aug 1; 58(2):809-17

**2021 IEEE Seventh  
International Conference on  
Bio signals, Images and  
Instrumentation  
ICBSII 2021  
SESSION VII  
RESEARCH  
PAPERS**



# microErgo: A Concept for an Ergonomic Self-Assessment Tool

João Rodrigues\*, Phillip Probst\*, Catia Cepeda\*, Federico Guede-Fernández\*, Sara Silva\*, Patricia Gamboa\*, Carlos Fujão<sup>‡</sup> Cláudia R. Quaresma\*<sup>†</sup> and Hugo Gamboa\*

\*LIBPhys (Laboratory for Instrumentation, Biomedical Engineering and Radiation Physics),  
Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, Caparica, Portugal

<sup>†</sup>Departamento de Física, Faculdade de Ciências e Tecnologias,

Universidade Nova de Lisboa, 2829-516 Monte da Caparica, Portugal

<sup>‡</sup>Volkswagen Autoeuropa, Quinta da Marquesa, 2954-024 Q.ta do Anjo, Portugal

**Index Terms**—micro, ergonomoy, risk, awareness, recommendation, self-assessment, office

**Abstract**—Work-related musculoskeletal disorders (WRD) have a high prevalence among office workers, which is even worsen with the current shift to home working office. Tools to assess their work setup in order to create awareness and reduce ergonomic risks are in need. There are tools available that allow ergonomic risk assessment to be carried out in a quantitative and qualitative way. However, quantitative tools require specialists' collaboration, while qualitative tools do not provide a quantification of risk and do not consider other variables, such as social and psychosocial factors. In this sense, we introduce the concept of a self-assessment tool, named microErgo, that calculates the risk of developing pain or WRD based on ergonomic exposure for its prevention and promoting self-awareness. This tool's construction requires an understanding and determination of the risk factors present in the office scenario. The risk associated with each factor must be considered and modeled by additional variables, such as individual characteristics or psychosocial. The final ergonomic risk is calculated by combining all sources of risk and transcribing it into a microErgo scale. microErgo is a concept to build an easy-to-use tool that combines existing ergonomic models with statistical data to calculate the risk of developing pain or WRD based on completing a self-assessment questionnaire and provide recommendations based on the risks identified.

## I. INTRODUCTION

Work-related disorders (WRD) have a major impact on the well being of workers. WRDs negatively influence productivity at work and increase the amount of absenteeism among the work force, thus also directly affecting companies economically and financially. In total, 7.4% of European workers suffered from work-related health problems, being mostly reported as musculoskeletal, stress, depression and anxiety [1]. In office jobs, mostly computer work, 10 to 62% of workers have reported musculoskeletal disorders [2]. Moreover, with the increasing need for remote work at home offices, this tendency is set to rise, which demonstrates the utmost importance to develop improved strategies for WRD prevention.

Several tools are currently available to quantify the exposure to ergonomic risk factors, namely RULA, EAWS and ROSA, among others [3], [4], [5]. However, these typically require (1) the supervision of experts on the field, (2) extensive

observation of variables and (3) identification of all possible risk factors. Additionally, these tools usually are validated based on pain reports thus only viewing the problem of WRDs from a physical perspective. However, research has shown that also psychosocial and environmental factors contribute to WRDs [5]. Other tools are also online available to provide a qualitative assessment of ergonomic risk. These are interesting as self-assessment tools, providing postural awareness to the user. Moreover, the improvements in postural awareness have been associated with reduced pain reports [6]. However, these tools lack (1) the quantification of the associated risk and (2) how the cumulative effect of this risk increases the chance of a disorder.

In this work, we propose a concept for a tool that is able to provide a quantifiable risk measure of exposure to ergonomic risk factors using evidence based on ergonomic guidelines. The primary focus of the tool lies in an user-friendly self-assessment that generates an ergonomic risk. The user is informed about this risk and which strategies could use to reduce it, thus increasing the user's awareness. The proposed tool is called *microErgo*.

It is relevant to note that this tool will not diagnose or provide medical expertise to deduce occupational injuries. It does present an estimation based on the available evidence provided by the user's inputs and the bibliography available. The inherent purpose is to enable ordinary people to do back-of-the-envelope calculations and, based on that, make personal decisions. With informative content, people can be more aware of their occupational habits and environment, as well as the risk factors to which they are exposed to, and take better decisions to prevent occupational disorders from happening.

This manuscript is structured in the following manner: (1) description of current available tools for ergonomic assessment, (2) tool concept description, (3) work plan to generate the concept, (4) demonstration of an example of what could be this tool and (5) conclusion and future work.

## II. EXISTING TOOLS

In the literature some available tools for ergonomics risks assessment can be found. They can be divided in two main groups: quantitative and qualitative assessment tools. Most of

these methods are based on the observation of the physical workload at work and monitoring the effects of the ergonomic changes.

#### A. Quantitative Tools

Quantitative tools are able to give a detailed and quantifiable assessment of a work process or a work station setup. They produce scores that allow for a categorization of the risk. However, these tools usually need to be administered by ergonomic specialists in order to be correctly used and, thus, there is chance for misuse among non-professionals. In the following section we present some of these tools.

Rapid Upper Limb Assessment (RULA) [3] is a tool designed to assess the exposure of workers to risk factors associated with work-related upper limb disorders. It has demonstrated low-moderate validity and moderate-good inter-rater reliability. RULA considers posture, force and movement. The final score is between 1 and 7 and a higher score is associated with greater risk level. This tool focuses on the postures adopted by the worker but it is not able to assess psychosocial factors or work style through the observational methods used.

Another available tool to assess biomechanical and ergonomic factors is the Rapid Office Strain Assessment (ROSA) [5]. It was designed not only to quantify risks associated with each component of a typical office workstation but also to establish an action level for change based on reports of worker discomfort. ROSA proved to be an effective and reliable method for identifying computer use risk factors related to discomfort. It was created using postures that were described in the *CSA Z412* guidelines for office ergonomics which were developed by a panel of ergonomic experts [7]. This method is organized in several sections such as: chair, monitor, telephone, keyboard and mouse. The scores verified in each subsection are combined to achieve a ROSA final score, which indicates the overall risk of musculoskeletal discomfort. The individual posture and equipment scores for each risk factor were modelled after deviations from a neutral posture given by the *CSA*. These deviations are also supported as risk factor for the onset of musculoskeletal disorders. ROSA should be used by an observer which selects the appropriate scores based on the posture of the worker at the computer workstation.

Occupational Repetitive Action (OCRA) [8] is one of the most widespread observational risk tools for assessing biomechanical risks related to repetitive tasks performed by upper limbs. The OCRA index is based on the ratio between the daily number of actions actually performed by the upper limbs in repetitive tasks, and the corresponding number of recommended actions. The main drawback of this method is the time consumption required for the computation of all considered factors.

Moreover, there is a self-report instrument which assess mixed ergonomic and psychosocial factors called Maastricht Upper Extremity Questionnaire (MUEQ) [9]. It assesses the

occurrence and nature of Complaints of Arm Neck and Shoulder (CANS) in computer workers and it associates physical and psychosocial risk factors. This questionnaire covers six domains: work station, posture, quality of break time, job demands, job control and social support.

#### B. Qualitative Tools

Qualitative tools are easy to use and mostly online based. These tools can be used in order to quickly acquire knowledge on ergonomic risks and get information on how to set up work stations and reduce possible risks. These tools, however, lack the quantification of risks and, thus, lack the possibility to assess potential risks. In the following section, we will present qualitative tools.

One of these qualitative tool is Ergotron [10]. It is a tool that helps to build an ergonomic workstation using simple web applications to create a more ergonomic work space. These enable the user to check the ergonomic fitness of the workstation through questionnaires and get recommendations from the provided answers. Additionally, Ergotron provides guides for the promotion of the ergonomics of standing. Their approach to improve workstation ergonomics is based on three key elements (1) neutral posture, (2) voluntary motion, and (3) rest time.

The European Agency for Safety and Health at Work has also released an extensive online risk assessment tool, called Online interactive Risk Assessment (OiRA) [11]. OiRA has a set of risk assessment tools for a variety of different working environments, ranging from agriculture to working with metal. The website also provides an assessment tool for office work. This tool assesses risks associated with office workstations, office layout, office environment, workers, and work organisation. The tool is set up in a simple questionnaire format with additional information presented with each question. Thus the users can easily inform themselves on the risks before answering a question. After the questionnaire is completed, an action plan with recommendations on how to reduce risk together with a timeline when measures against possible risks should be implemented, is generated.

There also exist other tools with a similar purpose as Ergotron and OiRA, such as [12], [13].

### III. CONCEPT DESCRIPTION

The proposed tool will be based on a model that considers available ergonomic tools of both quantitative and qualitative nature, but additionally takes into account statistical information to modulate the risk based on individual and psychosocial variables. The information gathered from these sources can be categorised as (1) information for *risk factor assessment* and (2) information for *risk factor modulation*. The concept is visualised in Figure 1.

The quantitative and qualitative tools, such as ROSA [5], RULA [3], MUEQ [9] or OiRA [11], will be leveraged to produce a set of user-friendly and easy understandable questions used for risk factor assessment. These questions will thus ask for information to identify potential ergonomic risks,

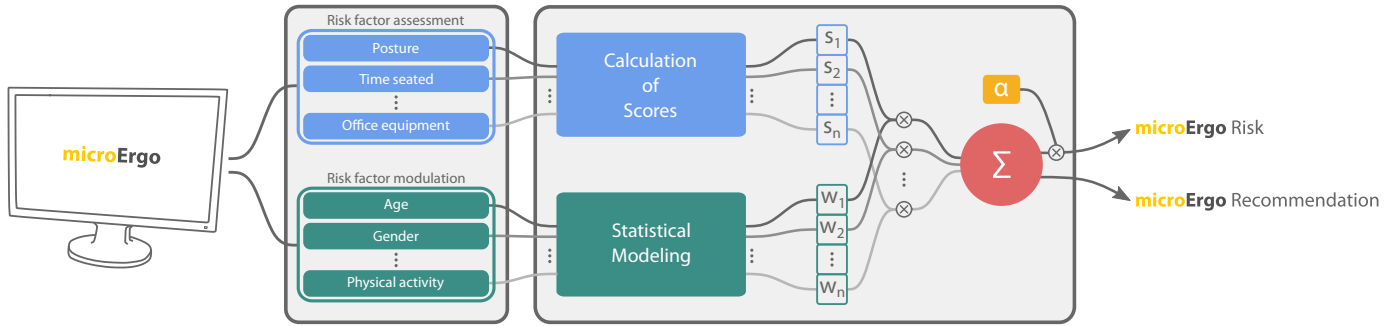


Fig. 1. microErgo Concept. A mathematical model combines a set of risk factors for assessment with a set of risk factors for modulation to generate a *microErgo* Risk and respective recommendations.

which can be attributed to, for example, having a bad posture during work, the time spent seated, the setup and the type of office equipment used, etc. Each question, or section of questions, will produce  $N$  risk scores that are calculated based on the existing quantitative and qualitative tools.

Risk factor modulation is based on statistical information that could be derived from age, gender, the tendency to make physical exercise, etc.. As Korhonen et al. [14] have shown, these examples of risk factor modulation are strong predictors for WRDs. The hereby gathered parameters are then used in a statistical model that determines a set of  $N$  weights. The exact type of statistical model still has to be determined but each weight can generally be expressed as

$$w_i = \sum_{j=1}^N p(d_i|m_j), \quad (1)$$

where  $w_i$  is the  $i$ -th weight for a WRD and  $d_i$  and  $m_j$  are a WRD (e.g. neck pain), and a  $j$ -th modulation factor (e.g. age), respectively. In case there are multiple modulation factors associated with the same WRD, the final  $w_i$  is the sum of the modulation factors weights.

The model has two outputs: a final *microErgo* risk score and a set of *microErgo* recommendations. The *microErgo* risk score is calculated by

$$\text{microErgo Risk} = \alpha \sum_{i=1}^M w_i s_i, \quad (2)$$

where  $s_i$  is a score that was calculated from the risk factor assessment and  $M$  is the total number of risk factors for assessment in use.  $\alpha$  is a variable that changes according to  $M$  and  $N$ , and corresponds to the conversion factor to *microErgo*.  $\alpha$  is calculated based on the maximum of each risk factor assessment and the maximum of each risk modulation factor, and in that scenario the *microErgo* Risk is equal to 1.000.000 *microErgo*. This is based on the definition of a *microErgo* to be a one-in-a-million chance of contracting an injury from a WRD. The *microErgo* score that is calculated takes inspiration from the micromort concept [15].

The recommendation will contain an informative part and a suggestive part. The informative part gives context and explains from what actions certain risks stem from. This could

be, for example, that the user is informed about ergonomic risks that are attributed to postural misalignment during work. By adding this informative part to the recommendation, we aim to give the user a substantive overview of the associated risks and thus increase the user's awareness for these. In the suggestive part of the recommendation, the user will be presented with a possible suggestion on how to reduce the risk. Suggestions will be given as a combination of text and images in order to make these as understandable as possible.

#### IV. HOW DO WE GET THERE

The concept and idea of a *microErgo* scale to measure the *microErgo* risk that contributes to musculoskeletal pain and disorders, demonstrates a complexity inherent to the open set of existing scenarios and risk factors. As mentioned in Section I, this tool will not give a medical assessment, but rather put in perspective the user regarding sources of micro risk and how do these contribute to pain or injury, triggering a sense of awareness in the user. Such functionality will have to answer to several questions, namely (1) *what represents a musculoskeletal risk factor*, (2) *how much risk does it imply*, (3) *are there variables that aggravate risk factors*, (4) *how to validate the microErgo scale* and (5) *how to setup a set of validated recommendation based on the risks identified*.

This section will provide guidelines in answering the mentioned questions, being divided in (A) Exploratory Information Gathering, (B) Building the formulation, (C) Test and validation.

##### A. Exploratory Information Gathering

The first stage in designing *microErgo* requires an exhaustive research of scientifically validated information. From the gathered information, we should be able to define the list of risk factors that are relevant for office scenarios, sort their relevance and attribute a risk score. This should be based on the existing screening tools mentioned in Section II. For instance, for office scenarios, *ROSA* and other qualitative tools are available to give guidelines on the risk factors that exist and quantify them, such as the workstation setup and certain postural angles [5].

Besides the presence of risk factors, the literature indicates that several other variables affect and aggravate pain, posture

and therefore risk of injury, such as age, gender, smoking habits, psychosocial variables, among others [5], [14]. Therefore, an exhaustive search on risk modulation variables should be made. The increase impact of these in pain reports should be accounted and used in the formulation step.

Finally, a thorough research over which recommendations are indicated by health organizations for specific risk scenarios has to be made. As recommendations have to be reliable and lead the user in improving current conditions, this work section is of utmost importance.

### B. Building the Formulation

After gathering the necessary information, the *microErgo* formulation will be made for all possible scenarios. This process requires two layers of information: (1) the front-end section, which corresponds to questions that are used on the self-assessment survey and (2) the back-end section, which receives the answers, combines the inputs in a mathematical model and calculates the final *microErgo* score.

The set of questions of the survey has to be user-friendly, while at the same time, be associated with a quantifiable risk factor from the available screening tools. Given that the aim is to provide a self-assessment survey and the existing quantification of risk usually relies on the collaboration of a specialist, a translation of the existing tools to an easy to use questionnaire is needed.

Additionally, personal and psychosocial variables have to be included, when found on the literature, to be associated with a change in pain reports. These variables will be used as modulation factors of the risk scores, attributing a higher weight to the risk when the variable aggravates pain reports. Further, in Section V is presented an example that demonstrates how these variables are used to modulate the risk.

As presented in Figure 1, the process regards the usage of two sources of information: variables that give a risk value, and variables that modulate the risk values used. All the modulators that contribute to a certain risk variable will be combined with it, as presented in Section III.

### C. Test and Validation of the Formulation

The developed formulation will have to be tested and validated. The process will require the expertise of a multi-disciplinary team of specialists in engineering, ergonomists, psychologists and occupational therapists. This process will follow the validation strategies found in existing screening tools [3], [5], which will involve (1) correlating the final scores with current methodologies and (2) correlating the final scores with pain reports. Based on the finding results, the type of questions included in the tool and their respective formulation will be adjusted.

Additionally, the recommendations that will be provided to the user will have to be tested in terms of comprehension, effectiveness and value. Based on these metrics, we will adjust the way recommendations are presented to the user.

## V. DEMONSTRATION

In this section, we present a simple example that demonstrates how *microErgo* could combine assessment tools and statistical data to estimate an ergonomic risk.

To provide this example, we decided to use an effective tool that assesses ergonomic factors in office workstations (ROSA) [5] and two risk modulation factors that were already associated with an increase of ergonomic risks: age and stress [16]. In a future approach, it is important to translate ROSA to a self-assessment survey, to be answered directly by the final user and not by a specialist.

In Figure 2 it is illustrated how the *microErgo* risk is estimated for the same subject and compares the result of two different postures. This illustration is aligned with all the further explanation.

The stages that the proposed concept follows to reach its outputs are described below.

### A. Interaction

An interface will be provided to the user, in which is presented the ergonomic assessment tool: ROSA assessment based on Chair, Keyboard and Mouse scores, scaled from 1 to 10. To assess the risk modulation factors, the user fills a questionnaire with the age in years (*Options*: <35, 35-44, 45-54, 55-64) and mental stress (*Options*: Not at all, Little, To some extent, Much).

The reported values for the two scenarios of the example are presented in Table I.

### B. Score based on ROSA

The ROSA assessment is made based on Chair score, Keyboard+Mouse score and duration. Given the reported data presented on Table I, the two scenarios have different scores, with a result of 6 for the top scenario and 4 for the bottom scenario.

### C. Statistical Modelling

As was already mentioned, this example includes age and mental stress as modulation factors. To define the impact of these factors in posture, which corresponds to the weight mentioned in Equation 1, we used the Odds Ratio (OR), which measures the strength of the association between two events [16]. Viikari-Juntura *et al.* [16] reported that the OR between workers with age of 55-64 years old and pain reports is 1.8 and the OR for a mental stress of "to some extent" is 2.2. Regarding that we have just one score from ROSA, and these two modulation have an impact on the reported WRD, the final weight from the statistical modeling will be the sum of the two OR: 4. This weight is the same for both scenarios.

In case there is a scenario with some modulation factor that does not has an impact on posture, this would not be regarded in the weight calculation associated with this WRD.

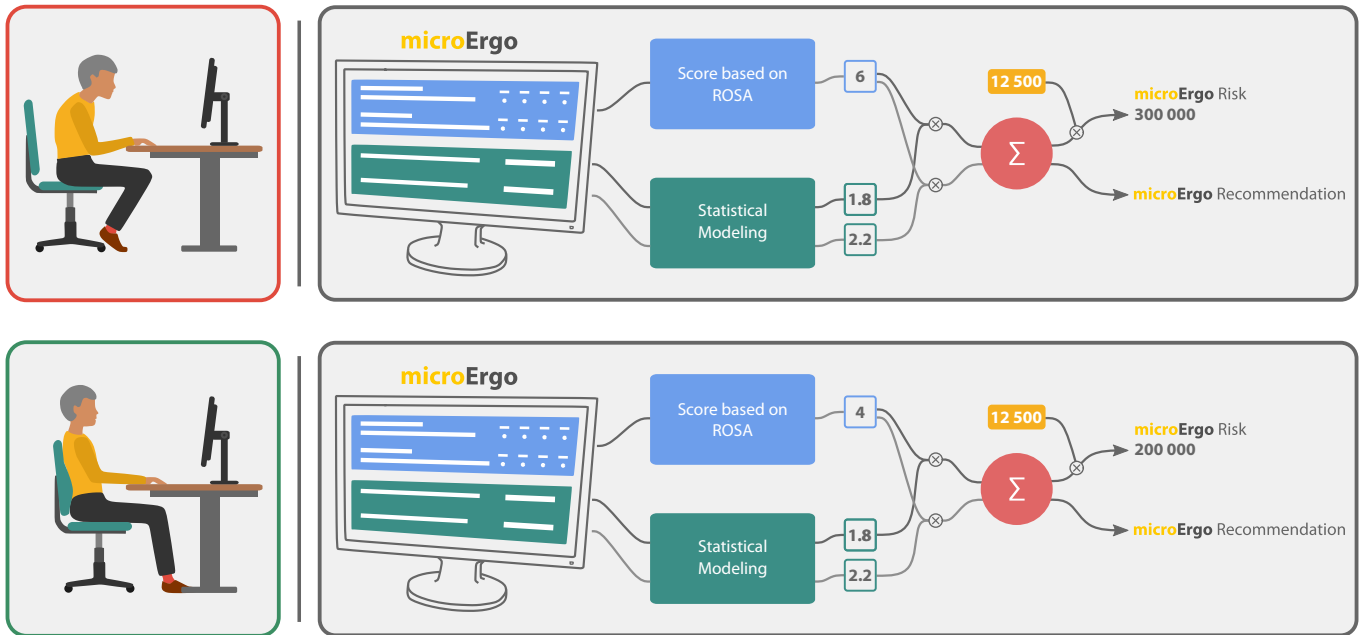


Fig. 2. Example of the microErgo tool application. Two examples with different inputs are given, on the top is presented an user with a incorrect posture and on the bottom the same user has a more correct posture. The respective outputs of the model are presented, as well as the microErgo Risk.

#### D. Estimation of Risk

As mentioned in the Section III, a microErgo is a one-in-a-million change of contracting an injury, so this risk ranges from 0 to 1.000.000.

Based on Equation 2, and given that we have the score of ROSA and the weights of age and mental stress, the first step is to calculate the  $\alpha$  factor to microErgo. For that, we use the maximum risk value and the maximum values of the tools and modulation factors, which are:

- Maximum risk: 1.000.000
- Maximum ROSA score: 10
- Maximum age factor: 1.8 (from [16])
- Maximum mental stress factor: 6.2 (from [16])

Applying these values in Equation 2:

$$\alpha = \frac{1000000}{10 \times (1.8 + 6.2)} = 12500 \quad (3)$$

With the  $\alpha$  value for this example, the microErgo risk for this example is estimated by Equation 4.

$$microErgo\ risk = \alpha \times (w_{age} + w_{mental\ stress}) \times s_{ROSA} \quad (4)$$

Applying Equation 4 for the two scenarios of the example, for the top scenario the risk is of 300.000 microErgo, while for the bottom scenario the risk is lower, being 200.000 microErgo.

#### E. Output

From this tool there are two outputs that will be provided to the user through the interface: the informative part and the suggestive part. The informative part includes the microErgo Risk that was calculated, and the suggestive part includes

specific recommendations according to what was assessed and respective modulation factors. Comparing to the existing tools presented in Section II, the first part is similar to what the quantitative tools propose, and the last approach is similar to what the qualitative tools provide.

An example of recommendation provided in the bottom scenario could be to place the monitor at a correct distance, since it is too far, and explain the risks associated with this type of postures.

#### VI. CONCLUSION AND FUTURE WORK

In this work, we have presented a concept for **microErgo**, an ergonomic tool that combines quantitative and qualitative assessment tools, as well as statistical data in order to give a more comprehensive and easy to interpret risk self-assessment for WRDs.

microErgo is a flexible tool and can be adapted to the user needs. In the example provided in V, posture, age and mental stress were included in the tool, but the final score could be adapted when different tools or modulation factors are added.

In relation to previous work, we aim to provide a unique risk measure, instead of different scores that are obtained from the existing quantitative tools. In addition, we pretend to have self-assessment surveys built from the existing tools to fulfill the major aim of **microErgo**: increase the awareness for WRDs and the associated ergonomic risks that lead to these, which have been already associated with reduced pain reports [6].

Comparing with the suggestions that the qualitative tools provide, we aim to have a mix of text and images in order to make these as understandable as possible. In a future perspective, we can create a database with the outputs from all users and present to each user a statistical study about the

Assessment	Top Scenario	Bottom Scenario
Duration	>4hours of daily sitting	>4hours of daily sitting
Chair	Too low height with adjustable Too long pan depth without adjustments No arm rest support and non adjustable No back support and non adjustable	Correct height with adjustable Correct pan depth without adjustments No arm rest support and non adjustable Back support and adjustable
Monitor	Eye level, but too far	Eye level, correct distance
Mouse+Keyboard	Mouse in line with shoulder Wrist in positive angle on the keyboard	Mouse in line with shoulder Wrist in correct angle on the keyboard
Age	55-64 years	55-64 years
Mental Stress	To some extent	To some extent

TABLE I  
REPORTED DATA IN ROSA AND QUESTIONNAIRE FOR THE TWO SCENARIOS PRESENTED IN FIGURE 2.

WRD based on the available data and a comparison between the user and the population assessed.

An interesting input to the modeling processing, that should be regarded in the future, is to have the microErgo risk as a risk accumulated over time. This would answer the question: how the cumulative effect of this risk increases the chance of a disorder and therefore provide a higher level of awareness over the cumulative effect that risk has if maintaining the same habits.

The presented concept still finds itself in the early stages of the design and development process. Thus, in the next steps, the development plan presented in section IV will be followed. Findings and executed studies during the development process will be published in future works.

#### ACKNOWLEDGMENT

This work was partly supported by Fundação para a Ciência e Tecnologia, under projects OPERATOR (ref. 04/SI/2019) and PREVOCUPAI (DSAIPA/AI/0105/2019), and Ph.D. grants PD/BDE/142816/2018 and PD/BDE/150672/2020.

#### REFERENCES

- [1] S. Agilis, "Final statistical report on the quality assessment and statistical analysis of the 2013 ad hoc module on accidents at work and other work-related health problems." 2015.
- [2] J. Wahlström, "Ergonomics, musculoskeletal disorders and computer work," *Occupational Medicine*, vol. 55, no. 3, pp. 168–176, 05 2005. [Online]. Available: <https://doi.org/10.1093/occmed/kqi083>
- [3] L. McAtamney and E. Nigel Corlett, "Rula: a survey method for the investigation of work-related upper limb disorders," *Applied Ergonomics*, vol. 24, no. 2, pp. 91 – 99, 1993. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/000368709390080S>
- [4] K. Schaub, G. Caragnano, B. Britzke, and R. Bruder, "The european assembly worksheet," *Theoretical Issues in Ergonomics Science*, vol. 14, no. 6, pp. 616–639, 2013. [Online]. Available: <https://doi.org/10.1080/1463922X.2012.678283>
- [5] M. Sonne, D. L. Villalta, and D. M. Andrews, "Development and evaluation of an office ergonomic risk checklist: Rosa – rapid office strain assessment," *Applied Ergonomics*, vol. 43, no. 1, pp. 98 – 108, 2012. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0003687011000433>
- [6] H. Cramer, W. Mehling, F. Saha, G. Dobos, and R. Lauche, "Postural awareness and its relation to pain: Validation of an innovative instrument measuring awareness of body posture in patients with chronic pain," *BMC Musculoskeletal Disorders*, vol. 19, 04 2018.
- [7] C. S. Association *et al.*, "Csa-z412 guideline on office ergonomics," *CSA International, Toronto, Ontario*, pp. 148–217, 2000.
- [8] E. Occhipinti, "Oera: a concise index for the assessment of exposure to repetitive movements of the upper limbs," *Ergonomics*, vol. 41, no. 9, pp. 1290–1311, 1998.
- [9] S. Eltayeb, J. B. Staal, J. Kennes, P. H. Lamberts, and R. A. de Bie, "Prevalence of complaints of arm, neck and shoulder among computer office workers and psychometric evaluation of a risk factor questionnaire," *BMC musculoskeletal disorders*, vol. 8, no. 1, pp. 1–11, 2007.
- [10] Ergotron Inc., "Ergonomics & Wellness," <https://www.ergotron.com/en-us/ergonomics>, 2021, (accessed: 2021-01-29).
- [11] European Agency for Safety and Health at Work, "Online interactive risk assessment," <https://oiraproject.eu/en>, 2021, (accessed: 2021-01-29).
- [12] Health and Safety Authority Ireland, "Business electronic safety management and risk assessment tool," <https://www.besmart.ie/>, 2021, (accessed: 2021-01-29).
- [13] Mark Middlesworth, "Recommended ergonomic assessment tools," <https://ergo-plus.com/ergonomic-assessment-tools/>, 2020, (accessed: 2021-01-30).
- [14] T. Korhonen, R. Ketola, R. Toivonen, R. Luukkonen, M. Häkkinen, and E. Viikari-Juntura, "Work related and individual predictors for incident neck pain among office employees working with video display units," *Occupational and environmental medicine*, vol. 60, no. 7, pp. 475–482, 2003.
- [15] R. A. Howard, "Microrisks for medical decision analysis," *International Journal of Technology Assessment in Health Care*, vol. 5, no. 3, pp. 357–370, 1989.
- [16] E. Viikari-Juntura, R. Martikainen, R. Luukkonen, P. Mutanen, E.-P. Takala, and H. Riihimäki, "Longitudinal study on work related and individual risk factors affecting radiating neck pain," *Occupational and environmental medicine*, vol. 58, pp. 345–52, 05 2001.

# Real Time Heart Beat Monitoring Using Computer Vision

Arppana A R<sup>\*</sup>, N K Reshma<sup>\*</sup>, Gopika Raghu<sup>\*</sup>, Nita Mathew<sup>\*</sup>, Harsha R Nair<sup>\*</sup>, Aneesh R P<sup>\*\*</sup>

<sup>\*</sup>Department of Biomedical Engineering, TKMIT, Kollam, Kerala, India

<sup>\*\*</sup>Regional Centre IHRD, Thiruvananthapuram, Kerala, India [aneeshprakkulam@gmail.com](mailto:aneeshprakkulam@gmail.com)

**Abstract** -Among the different biological parameters, heart rate (HR) plays a very significant part in the disease diagnosis. It confirms the state of living of a person. Heart rate gives an idea about the person's health and hence it should be observed. Nowadays, a non-contact based method is more comfortable, low cost, simple and easy to use. In this era, with an increase in automobiles, people who are engaged in driving are also on the surge. Such people are also victims of cardiac-related diseases. The proposed method demonstrates real-time non-contact heart rate monitoring systems that enable the detection of cardiac activity cycles without any cutaneous contact using machine learning. The methodology involves the extraction of heart beats signals from face images using independent component analysis. Neural network is employed to generate the BPM from the samples. As part of preparing the dataset, a heartbeat sensor is developed with a microcontroller. The algorithm is successfully tested with real time face images and the non-contact based detection method more conveniently helps to measure HR with more acceptable accuracy. It can apply to the drivers for the prevention of accidents or manual labourers who work in.

**Keywords:** Heart Rate (HR), Machine learning, High-resolution camera, Independent Component Analysis (ICA), Principal Component Analysis (PCA), Neural network.

## I. INTRODUCTION

Heartbeat is one of the diagnostic tools for heart failure[1].Monitoring the rate of heart and surge in pressure of blood is an important thing for human health and plays an important role for the various dangerous situations. Contactless systems are convenient, cheap and user friendly [2].The color differences occur in the facial region because of the heart's nature of sending blood to different parts of our body. Thus, we obtain our heart rate from the color sequences produced. The heart is a muscular organ which pumps blood through blood vessels in most animals. Heartbeat refers to the cycle of heart muscle contraction. It occurs with an impulse from the Sino atrial node. A heart rate below 60 bpm is considered to be very low. The heartbeat of a person can be checked manually in two ways. One of them is by checking the pulse on the wrist, known as radial pulse and another method is by checking the pulse from the neck known as the carotidpulse [3].

The procedure involves placing 2 index fingers on either the wrist or neck and then taking count of the number of

pulses for 30 seconds. It's then multiplied by 2 to get the actual heartbeat. Another method to measure pulse is by the process of heart rate sensor. It works on the principle of photoplethysmography, ie, any change in volume of blood causes a change in the light intensity through that organ[4].Heart rate refers to the number of times the heart beat occurs in a minute. The normal rate is 60-70 beats per minute. Some of the cardiovascular diseases are: 1) Atherosclerosis 2) Heart attack 3) Stroke 4) Heart failure 5) Arrhythmia 6) Heart valve problems[5].

The Heart Beat is mainly represented with the help of QRS waveform. The P wave (0.25mV, <80ms) represents atrial depolarization (activation). The QRS complex (2.5-3.0mV, 80-100ms) represents the repolarization of auricles. The time passed by two successive R waves is given by the R-R interval. The T wave (<5mV, 160ms) represents the repolarization of ventricles [5].

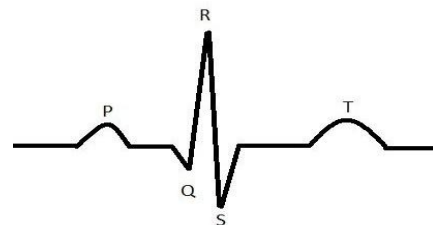


Figure 1 : Normal ECG

## II. LITERATUREREVIEW

In 1995, Costa et al. investigated the first non-contact health monitoring system in which camera images were used to obtain different health indicators from varied colour sequences occurring in the human body [6]. This reported only a graph of heartbeat but did not report quantitative results and any relation with ECG signals used for comparison [7]. Takano et al. simultaneously extracted some parameters. MATLAB custom functions were used for the extraction of HR and HRV. RR and BVP are other parameters, extracted from the images [8]. In this, heart rate monitoring occurs for a particular period, but the efficiency

cannot be defined. Kenneth et al, proposed a method for obtaining two PPG signals at the same time using contactless systems but with varied wavelengths. The system extracted SPO2 successfully [9].

In 2008, Verkruyse et al used a method of obtaining HR and RR from pre-recorded frames using a proper light source from 30 seconds to a few minutes using simple and inexpensive digital cameras[10]. In 2009, Banitsas et al. presented a method involving usage of a smartphone camera to obtain heart rate. Finger was used as the main source of obtaining information[11]. In 2011, Poh et al. introduced an algorithm with computer webcam. It was used to directly obtain the information from the primary colours. Sensors were used to record the BVP with the respiration. 256 Hz sampling rate was used. They were extracted from a 1 minute color facial video using fast Fourier transform[12]. In 2014, Zhang et al. introduced a webcam based approach for observing the life indicators of drivers contactlessly. HR and RR is processed using fast fourier transform. Once, the camera obtains the face images, extraction into primary color channels is done [13].

Guo et al. introduced a method to track the heart rate of a driver, by creating an actual scenario of driving challenges. Face regions were divided into 7 sub regions and average HRV is calculated from BVP obtained by the Independent Component Analysis method[14]. Xiaobai Li et al. introduced another method. Heart rate was obtained by creating an actual man-machine interactivity scenario [15]. In 2015, Rahman et al. introduced a contactless method for life indicator value abstraction. Apart from the heart rate and RR, another parameter named inter bit interval was also measured. Even this method included the usage of a laptop with a webcam [2].

### III. PROPOSED METHOD

“Real Time Heart Rate Monitoring using Computer Vision” uses a camera for imaging and signal processing to measure heart rate. Figure 2 shows the block diagram. This proposed method demonstrates real-time non-contact heart rate monitoring systems that enable the detection of cardiac activity cycles without any cutaneous contact using machine learning and a high-resolution camera.

### IV. METHODOLOGY

#### a) Image Acquisition

Image is taken using a webcam. The successive images of the person are taken for a dataset collection for further analysis. High resolution camera with proper lighting is used to capture the video. This video is converted to frames. LogiTech HD camera with a frame rate of 30 per sec is used. 50 person's face videos were collected in a standard environment. Each video contains 3000 frames[16]. A sensor based heart beat detection is also developed to record the HR

while capturing the frames. This hardware contains the sensor and a microcontroller.

#### b) Image Preprocessing

The source is an important part in collecting the data. The small change in the camera parameters, may mislead the final calculations. For the portability of the algorithm, camera parameters are analyzed. The image full camera projection matrix is obtained only when external and internal camera criteria are known. To preprocess images before the detection of the face, The algorithms, used, are Intrinsic image decomposition algorithm and Retinex algorithm. Camera Parameters include intrinsic and extrinsic parameters [17]. They are as follows:

- (i) Intrinsic - It refers to camera parameters which are internal and constant. It's camera specific.
- (ii) Extrinsic - It refers to camera parameters which are external and vary according to world frame.

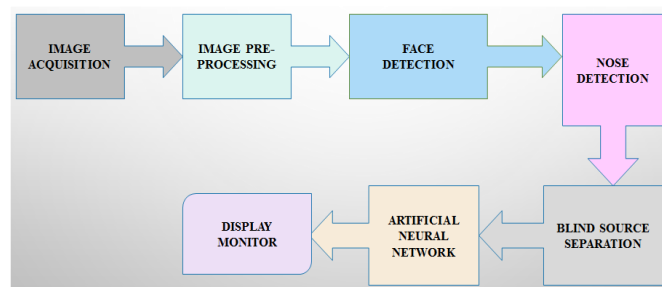


Figure 2: Block Diagram

#### i) Intrinsic image decomposition

Intrinsic image decomposition resolves the problem faced during decomposition of an image, which is inherently ill-posed[18]. A single image is typically solved by optimization based on Retinex theory. A common approach to apply deep learning to image processing, is to extract high level features by down sampling and then up sampling the features to the original resolution, is effective for obtaining overall structural information, but it is not appropriate due to possible loss of details. An image  $I$  is decomposed into reflectance  $R$  and shading  $S$ , satisfying:

$$I = R \times S \quad (1)$$

In the logarithmic domain, reflectance  $R$  can be represented as a residual of  $I$  when we have computed shading  $S$ .

$$\log(R) = \log(I) - \log(S) \quad (2)$$

#### ii) Retinex

The Retinex theory, as originally developed by Land and McCann can be seen as fundamental theory for several state-of-the-art intrinsic image algorithms[19]. This theory is perceived to be related to color receptivity of human



vision in the real world. The entire image is responsible for the appearances in our real-life scenes. The L,M,S can occur in varied colors. The three lightness seen in long, middle and short waves determines the color sensations. Algorithms that copy spatial interaction of vision to find out lightness is being named after this.

#### c) Face Detection and Nose Detection

The facial skin indicates the color changes during the variation of the heart beats. The region around the nose covers a larger area of the face. the color variation in those areas will be capable of showing the color variation of circulation. Viola Jones. Viola Jones Algorithm is used to detect the face and nose[20]. Adaboost training algorithm with Haar features detects the regions . The detected nose regions can be expanded , if the area of the detected nose is smaller. Then eye and mouth are also detected to fix the region of interest[21]

#### d) Blind Source Separation

The disjoining of source signals from a group of various signals, regardless of the signal information and process is called Blind Source Separation(BSS).This method commonly used to separate source devices without any references [22]. Application of this method is being implemented in biomedical signal processing [23].

##### i) Principal Component Analysis(PCA)

PCA reduces the complexity of the equation while retaining certain information [24]. The covariance matrix can be represented as

$$R_s = \frac{1}{N} \sum_{t=1}^N S(t)S^T(t) \quad (3)$$

The covariance matrix factors can be equated as

$$R_s = DQD^T \quad (4)$$

Here Q represents a M X M matrix of diagonal form with eigen value of  $R_s$  and D is the eigenvector matrix.

##### ii) Independent Component Analysis (ICA)

The method for splitting multivariable signals into its extra sub constituents by a computational approach is called ICA. Considering two instruments are played in a room simultaneously and two microphones are also used[ 25]. The outputs from the instruments are taken as  $y_1(t)$  and  $y_2(t)$  and the weighted sum as  $s_1(t)$  and  $s_2(t)$  respectively.

$$y_1 = b_{11}s_1(t) + b_{12}s_2(t) \quad (5)$$

$$y_2 = b_{21}s_1(t) + b_{22}s_2(t) \quad (6)$$

where  $b_{11}$ ,  $b_{12}$ ,  $b_{21}$ , and  $b_{22}$  are the parameters which depend on the distance of microphones. For n linear mixtures observed signals  $y_1, y_2, \dots, y_n$  are mutually independent.

$$y_j = b_{j1}s_1 + b_{j2}s_2 + \dots \dots b_{jn}s_n \quad \text{for all } j. \quad (7)$$

Using this representation of vector matrix, the above mixing model can be specified as:

$$X = AS \quad (8)$$

It can be written in the form:

$$x = \sum_{i=1}^n a_i s_i \quad (9)$$

There are two main pre-processing stages in ICA-Centering and whitening. The main reasons include algorithms simplification, deduction of dimensionality and deduction of indicator numbers. In centering, the mean is subtracted from each signal which is similar to conversion of the coordinate centres to the origin which means it can be always re-added to the final product obtained. Whitening transformation converts the data if it has an identity covariance matrix . The input data matrix x must be centred and whitened before applying the algorithm of ICA. Contrary to the other methods, ICA is very efficient for determining different factors [26].

##### (i) EFICA

Efficient ICA is an ICA algorithm designed to distinguish signals from non-Gaussian independent identically distributed (i.i.d)[27]. The basic assumption is that each source signal is,  $s_k, k=1,2,\dots,d$  consists of an independent N realization of a random variable q with a non-Gaussian distribution function:

$$F_k(x) = P(\xi_k \leq x) \quad (10)$$

##### (ii) FAST ICA

Fast ICA is the implementation of a rapid static point algorithm for the analysis of separate components [28]. The algorithm used in Fast ICA is an easy to use graphical user interface. Since the algorithm of Fast ICA is computationally efficient, it is used for the estimation of ICA. Initially, by considering Fast ICA for One Independent Component. Let the collected sample be random vector x (i.e. a linear mixture of independent source signals based on the BSS). A column of the B matrix of demixing form corresponds to the final vector w(k). If blind source separation approach is used, w(k) differentiates one of the non-Gaussian source signals so w(k)Tx(t), t = 1, 2, 3..... is equal to one of the source signals. For the estimation of n separate components, the program must be run n number of times. By adding an orthogonalizing projection inside the loop, different independent components can be estimated.

##### d) Artificial Neural Network

The neural network that mimics an animal brain is known as an artificial neural network. The artificial neural network consists of input, output and weights. A supervised

learning method is adapted to train the heart beat ground truth. Input features are the mean , variance ,standard deviation of fast fourier transform of ICA signal. The difference in successive images shows the variation of blood in the skin region. There are 15000 frames with different HB measurements . These values are trained with the extracted feature . This creates a Generative model for testing. This model is used to predict the heart beat from the new frames [29].

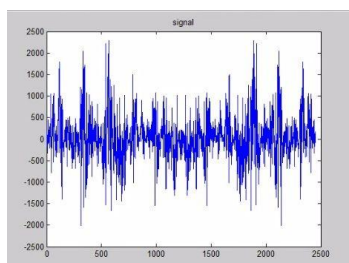


Figure 3: Signal generated through blind source separation

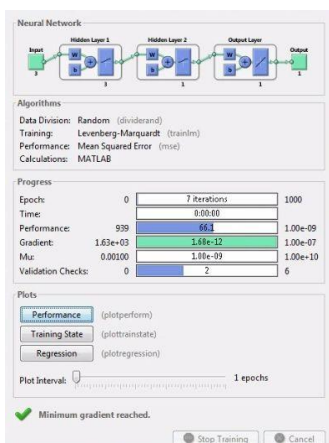


Figure 4 : Neural Network Training

## V. RESULTS AND DISCUSSIONS

An experimental setup was created indoors by placing a laptop with built-in webcam on a table with ambient sunlight. The participants were informed beforehand and during the experimental time , they were asked to focus into the camera for capturing of different image frames. The images for real time heart rate extraction were recorded sequentially with pixel resolution and were stored in the laptop using PNG format. During the experiment, an ECG sensor was also used to capture the readings.

The captured frame indicates the beginning and ending of a video sequence. The count of image frames during a fixed time is used to obtain heart rate in real time. But, it is essential to maintain the pixel at a constant rate for the purpose of calculation. Using the face detection function ‘CascadeObjectDetector’ of MATLAB along with VJ (Viola and Jones ) method, the face is tracked in real time. From the facial region, Region Of Interest(ROI) was

detected. The region had 60% width and 40% height and had only a useful region.

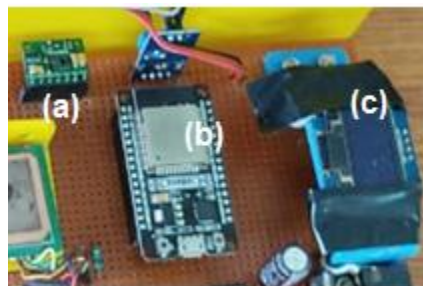


Figure 5: Hardware based heart rate monitoring system a) Heart beat sensor, b) ESP node32 microcontroller, c) OLED

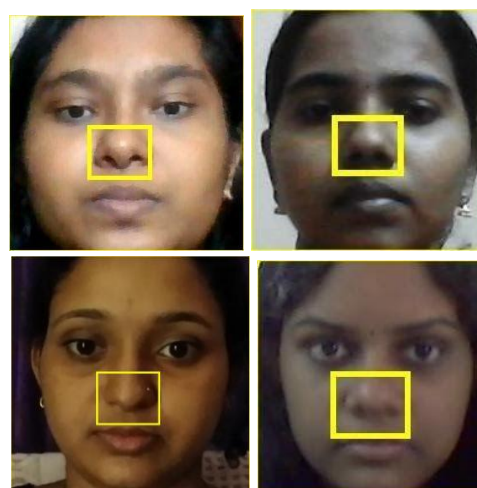


Figure 6: Face and Nose Detection of different persons

The RGB signals were obtained from the cropped ROI, the three desired signals. The RGB signal production involves two phases. In the first phase, the RGB color values with respect to each frame is averaged and in the second phase, the RGB signals are obtained by summing the mean RGB color values. Signal de-trending is used for the removal of trends occurring due to noise and temperature. The collected RGB is usually drifted and noisy. The RGB signal has been detrended using the method based on smoothness priors approach.

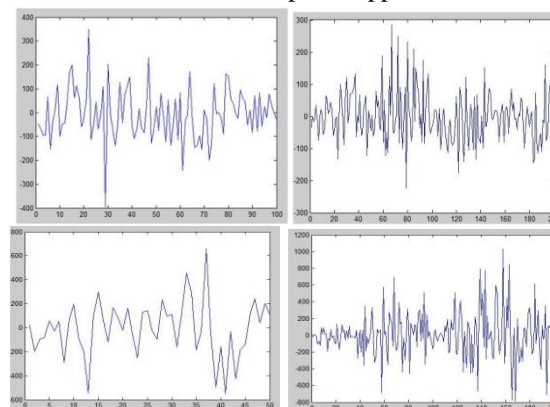


Figure 7: Different output signals in different time scale

Mean absolute error (MAE) is used to measure the performance of the system. 15000 frames with known heart rate value is trained and 4000 frames used for testing.

$$MAE = \frac{1}{N} \sum_{k=1}^N |Y_k - X_k| \quad (11)$$

$Y_k$  = Predicted Heart rate through the algorithm

$X_k$  = Heart rate measured from hardware sensor

$N$  = 4000 frames ( 1000 frames of 4 persons)

Mean absolute error = 0.05

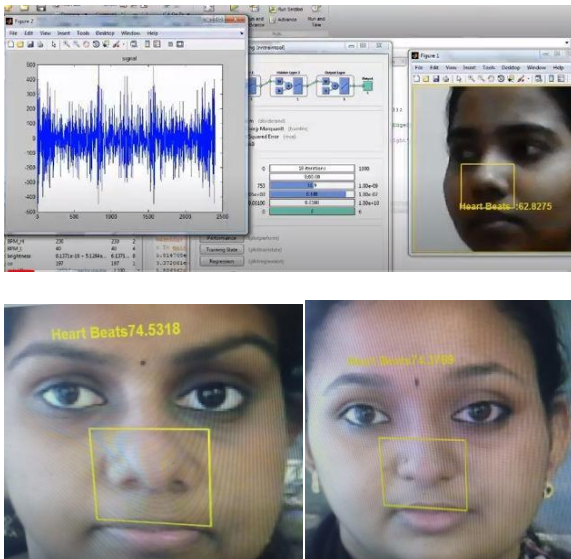


Figure 8 : Simulation Results

## VI. CONCLUSION AND FUTURE ENHANCEMENT

Heart beats detection is an inevitable part of disease diagnosis. Contact based pulse monitoring is being carried out. Non-invasive based heartbeats detection paves ways to disease detection. This will help in drowsiness detection also. Here computer vision based algorithm is developed for heart beats detection. Independent component analysis method is used to convert the image region to signal. Nose region is selected as the ROI . An embedded hardware is developed with contact based HB detection. This is used to prepare the ground truth while capturing the face images. With these annotations The ICA signal is trained using Neural Network. This algorithm is successfully tested with real time video. The obtained mean absolute error is 0.05.

## REFERENCES

- [1] S Havlin, L.A.N Amaral, Y Ashkenazy, A.L Goldberger, P.Ch Ivanov, C-K Peng, H.E Stanley,"Application of statistical physics to heartbeat diagnosis", Physica A: Statistical Mechanics and its Applications,Volume 274, Issues 1–2,pigno: 99-110, 1999
- [2] H. Rahman, M.U. Ahmed, S. Begum and P. Funk," Real Heart Rate Monitoring From Facial RGB Color Video Using Webcam", The 29th Annual Workshop of the Swedish Artificial Intelligence Society (SAIS), June 2016.
- [3] Sungjun Kwon, Hyunseok Kim and Kwang Suk Park, "Validation of heart rate extraction using video imaging on a built-in camera system of a smartphone," 34th Annual International Conference of the IEEE EMBS San Diego, pp. 2174-2177, August 2012.
- [4] L. Iozzi, L. Cerina and L.T. Mainardi, "Assessment of beat-to-beat heart rate detection method using a camera as contactless sensor", pp. 521-524, 2016.
- [5] Toshihiro Kitajima, SangOn Choi and Edwardo Arata Y. Murakami, "Heart Rate Estimation based on Camera Image", International Conference on Intelligent Systems Design and Applications (ISDA), pp. 50-55, 2014.
- [6] G. D. Costa, "Optical remote sensing of heartbeats," Opt. Commun., vol. 117, pp. 395–398, 1995
- [7] H. Rahman, M.U. Ahmed, S. Begum, P. Funk, "Non-Contact Physiological Parameters Extraction Using Facial Video Considering Illumination, Motion, Movement and Vibration," IEEE Transactions on Biomedical Engineering, vol. 67, no. 1, pp. 88-98, January 2020.
- [8] Chihiro Takano, Yuji Ohta, "Heart rate measurement based on a time-lapse image" ,Medical Engineering & Physics, Volume 29, Issue 8, 2007, Pages 853-857,
- [9] Kenneth HumphreysTomas E WardTomas E Ward Charles Markham, "Noncontact simultaneous dual wavelength photoplethysmography: A further step toward noncontact pulse oximetry" Review of Scientific Instruments , American Institute of Physics ,May 2007
- [10] W.Verkruyse, L.O. Svaasand, and J. S. Nelson, "Remote plethysmo-graphic imaging using ambient light. " Opt. Express, vol. 16, no. 26, pp.21434-21445, 2008.
- [11] K. BanitsasP. et al. "A Simple Algorithm to Monitor HR for Real Time Treatment Applications", Information Technology and Applications in Biomedicine, 2009. ITAB 2009. 9th International Conference on December 2009
- [12] M.-Z. Poh, D. J. McDuff, and R. W. Picard, "Advancements in noncontact, multiparameter physiological measurements using a webcam," IEEE Trans. Biomed. Eng., vol. 58, no. 1, pp. 7–11, Jan. 2011.
- [13] Zhang Q, et al "A machine learning empowered system for long-term motion-tolerant wearable monitoring of blood pressure and heart rate with ear-ECG/PPG," IEEE Access, vol. 5, pp. 10547–10561, 2017.
- [14] Weiwei Guo 1, a, Xiaoting Tian 1, b , Jiyuan Tan 1, c and Li WANG , "Change in Heart Rate Variability Indexes due to high driving workload in
- [15] X. Li, J. Chen, G. Zhao and M. Pietikäinen, "Remote Heart Rate Measurement from Face Videos under Realistic Situations," 2014 IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 2014, pp. 4264-4271, doi: 10.1109/CVPR.2014.543.
- [16] Malavika Suresh et al., " Real- Time Hand Gesture Recognition Using Deep Learning ", International Journal of Innovations and Implementations in Engineering(ISSN 2454- 3489) , December 2019 vol 1
- [17] Avigyan Sinha, et al. " Real Time Facial Emotion Recognition using Deep Learning " International Journal of Innovations and Implementations in Engineering(ISSN 2454- 3489) , December 2019 vol 1
- [18] P. Laffont, A. Bousseau and G. Drettakis, "Rich Intrinsic Image Decomposition of Outdoor Scenes from Multiple Views," in IEEE Transactions on Visualization and Computer Graphics, vol. 19, no. 2, pp. 210-224, Feb.

2013, doi: 10.1109/TVCG.2012.112.

- [19] Edwin H. Land and John J. McCann, "Lightness and Retinex Theory," *J. Opt. Soc. Am.* 61, 1-11 (1971)
- [20] Anusha A V ,Jayasree J K , Anusree Bhaskar, "Facial Expression Recognitin And Gender Classification Using Facial Patches"2016 International Conference on Communication Systems andNetworks (ComNet) pp.21-23,July 2016.
- [21] Anusree Bhaskar "Advanced Algorithm ForGender Prediction With Image Quality Assessment" International Conference on Advances in Computing, Communications and Informatics (ICACCI) pp. 1848-1855, 2015.
- [22] Sneha K A et al. " Real Time Source Camera Identification using Advanced Blind Source Separation Method "International Journal of Innovations and Implementations in Engineering(ISSN 2454- 3489) , April 2018 vol 1
- [23] M. G. Devika, et al. "Myocardial infarction detection using hybrid BSS method," 2016 International Conference on Communication Systems and Networks (ComNet), Thiruvananthapuram, 2016, pp. 167-172, doi: 10.1109/CSN.2016.7824008.
- [24] R. V. Savitha, S. R. Breesha and X. F. Joseph, "Pre Processing the Abdominal ECG Signal Using Combination of FIR Filter And Principal Component Analysis," 2015 International Conference on Circuits, Power and Computing Technologies [ICCPCT], pp. 1-4, July2015.
- [25] V. A.V, C. Markose and A. R.P, "Non-contact Heart Rate Monitoring Using Machine Learning," 2019 2nd International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICT), Kannur, Kerala, India, 2019, pp. 1400-1404, doi: 10.1109/ICICT46008.2019.8993251.
- [26] Alaa Tharwat,"Independent component analysis: An introduction," Applied Computing and Informatics, 2018.
- [27] Z. Koldovsky, J. Malek, P. Tichavsky, Y. Deville and S. Hosseini, "Extension of EFICA Algorithm for Blind Separation of Piecewise Stationary Non Gaussian Sources," 2008 IEEE International Conference on Acoustics, Speech and Signal Processing, Las Vegas, NV, USA, 2008, pp. 1913-1916, doi: 10.1109/ICASSP.2008.4518009.
- [28] E. Oja and Z. Yuan, "The FastICA Algorithm Revisited: Convergence Analysis," in *IEEE Transactions on Neural Networks*, vol. 17, no. 6, pp. 1370-1381, Nov. 2006, doi: 10.1109/TNN.2006.880980.
- [29] J. S. Smith, B. Wu and B. M. Wilamowski, "Neural Network Training With Levenberg–Marquardt and Adaptable Weight Compression," in *IEEE Transactions on Neural Networks and Learning Systems*, vol. 30, no. 2, pp. 580-587, Feb. 2019, doi: 10.1109/TNNLS.2018.2846

# Development of Reconfigurable Control Schemes for Epileptic Seizures

1<sup>st</sup> Nambi Narayanan S  
Research Scholar

Department of Instrumentation Engg  
MIT Campus, Anna University  
Chennai-44, Tamil nadu, India  
vinoth2852@gmail.com

2<sup>nd</sup> Sutha Subbian  
Assistant Professor

Department of Instrumentation Engg  
MIT Campus, Anna University  
Chennai-44, Tamil nadu, India  
sutha\_man@hotmail.com

**Abstract**—Advanced research in neuroscience recognized electrical stimulation as a well-accepted treatment to suppress neurological disorders like Epilepsy, Parkinson's, mental depression and schizophrenia. Among these disorders, epilepsy is one of the most common neurological disorders. When epileptic seizures occur, the nervous system generates abnormal action potentials with repetitive firing spikes which cause permanent brain damage and unexpected death. Hence, it must be controlled by either medication or surgical treatment. If it is severe, it cannot be controlled by medication. One of the alternate choices to control epilepsy is electrical nerve stimulation. However, development of a control scheme for the nervous systems to control neurological disorder is a challenging task due to complex biochemical reactions associated with electrical action potential. In this paper, Reconfigurable Control schemes are proposed to control epileptic seizure using Proportional Integral (PI) controller and Model Predictive Controller. Based on Hodgkin-Huxley (HH) model, the transfer function and state-space models are developed for a neuron under normal and Epileptic/abnormal conditions. It is assumed that the epileptic seizure is bifurcations in the neural system due to a change in sodium transconductance. When an epileptic seizure is detected, a reconfigurable controller will take control action to recover the nervous system by suppressing the repetitive firing. In addition, the stability and performances of reconfigurable control schemes are evaluated and compared. The results show the feasibility of proposed reconfigurable control schemes.

**Keywords**— *Epileptic seizure, Hodgkin-Huxley model, PI Controller, Model Predictive Controller, Bifurcation, Reconfigurable controller.*

## I. INTRODUCTION

Epilepsy is a common neurological disorder that is caused due to abnormal electrical discharge in the nervous system. Epilepsy may be controlled by Antiepileptic Drugs (AEDs), diet therapy, implantable devices and surgery. Among these, implantable devices with electrical stimulation have drawn more attention because it does not produce side effects as compared to medications. Electrical stimulation relies on dynamics of the central nervous system. However, understanding of its dynamics is very difficult during seizures. The mathematical modelling is an essential tool for analysing the complexity of epileptic seizures. Roxana et al (2012) presented various computation models and its applications for epileptic therapy. They reported the need of Deep Brain Stimulation (DBS) and vagal nerve electrical stimulation for controlling seizures. There are two different types of

models used for DBS, namely microscopic and mesoscopic models. In the microscopic model, the neural activity is described at the level of each neuron, whereas in the mesoscopic model, it is described at the level of small neural populations. Wim van Drongelen (2013) described the behaviour of epileptic seizures using different types of cellular models and network models. Among the cellular models, Hodgkin-Huxley (HH) model is one of the most popular microscopic conductance based neuron models to describe the generation and propagation of action potential. The HH model can be used for modelling sinus atrial node cell in human heart. Nambi Narayanan et al (2017) developed a nonlinear model for human heart using HH model by considering action potential as input and Electrocardiogram (ECG) signal as output.

Bifurcation phenomenon is one of the roots for damaging the nervous system which leads to neurological disorders such as epilepsy, Parkinson's etc. Hence, bifurcation phenomenon analysis and development of various control strategies have become emerging research areas in neuroscience. Ozgur et al (2018) developed repetitive firing stopper mechanism exerted on nerve fibre using electric field. They have designed Linear Quadratic Regulator (LQR) to change the bifurcation characteristics to suppress repetitive firing pulse. Yexin Lin et al (2019) applied projective control theory to suppress Hopf bifurcation phenomenon in HH model. The authors designed the controller by augmenting washout filter dynamics with HH neuron dynamics. Shaolong Chen (2019) proposed a novel control method with fractional order washout filter for bifurcation control of fractional order neuron using HH model. They analysed Hopf bifurcation by varying the external current. Shuang Liu et al (2020) analysed the bifurcation points in nonlinear electromechanical coupling main drive system of a rolling mill and designed LQR using feedback linearization method. Harleen K Brar et al (2018) reported that Proportional Integral (PI) controller provides stable neural activity under healthy condition of brain. In epileptic seizure conditions, PI controller leads to destabilization, hence they proposed model predictive control with differential dynamic programming to control epileptic seizure.

Reconfigurable control scheme plays a major role in healthcare monitoring and control to increase the life span of the patients and reduce the complexity of the system. For the epileptic seizure patients, the proposed reconfigurable control scheme can suppress the repetitive

firing due to bifurcations. However, the reconfiguration can be done if the bifurcation model is known. Rieke Fruengel et al (2020) developed local and global reconfigurable network during the pre-seizure based on different node centrality indices. Chen Liu et al (2020) proposed Neural Network based PID control scheme to control different Parkinson disease states by readjusting its controller parameters. Andrew Haddock et al (2017) designed a model predictive controller for Parkinson Disease (PD) control and optimized the controller to achieve minimum patient's symptoms and device power consumption. Sutha et al (2012) developed multi-objective reconfigurable output feedback controller for Multi-Input Multi Output three tank systems under partial actuator fault conditions. The same scheme is extended for biomedical application to control the abnormality in neuron action potential. Yue Zhang et al (2013) proposed the bifurcation analysis in electrophysiological system. Indeed, there are many diseases which are close relations with bifurcations such as Parkinson, epilepsy and pathological heart rhythms etc. John E. Fleming et al (2020) developed a control scheme for deep brain stimulation to control Parkinson's disease by suppressing beta band activity. They designed PI controllers for amplitude modulation and frequency modulation to control PD. Ning Li (2017) proposed adaptive periodic intermittent control to control epilepsy with minimum cost.

The major contribution of the paper includes I) Development of transfer function models and state space models for neuron with healthy (normal) condition and epileptic seizures conditions using HH model. II) Development of PI control scheme and MPC scheme based on HH model III) Development of reconfigurable control scheme to readjust its controller parameters based on abnormal model parameters to control epileptic seizures. IV) Analyse the performance and stability of the proposed control schemes.

The organization of the paper is as follows: Section 2 describes the modelling of a brain neuron. Section 3 presents the proposed reconfigurable control scheme with PI controller and MPC. The closed loop performance and stability analysis of the proposed control schemes are discussed in section 4. A short conclusion is found in section 5.

## II. MODELING OF BRAIN NEURON

The HH model (Hodgkin et al 1958) is a conductance-based cellular model used to describe the electro- physical activities of a brain neuron is considered in this study for simulation. The HH model represents each component of an excitable brain neuron as electrical quantities like sodium current, potassium current and leakage current. The HH model is comprised of four ordinary differential equations (1)-(4) as follows:

Total current through the brain neuron cell membrane ( $I_{ext}$ ) is

$$I_{ext}(t) = C_M \frac{dV}{dt} + g_{Na} m^3 h (V - V_{Na}) + g_K n^4 (V - V_K) + g_L (V - V_L) \quad (1)$$

$$\frac{dm}{dt} = \alpha_m (1 - m) - m \beta_m \quad (2)$$

$$\frac{dh}{dt} = \alpha_h (1 - h) - h \beta_h \quad (3)$$

$$\frac{dn}{dt} = \alpha_n (1 - n) - n \beta_n \quad (4)$$

$$\alpha_n = \frac{0.1 - 0.01V}{\exp(1 - 0.1V) - 1} \quad \beta_n = 0.125 \exp\left(\frac{-V}{80}\right)$$

$$\alpha_m = \frac{2.5 - 0.1V}{\exp(2.5 - 0.1V) - 1} \quad \beta_m = 4 \exp\left(\frac{-V}{18}\right)$$

$$\alpha_h = 0.07 \exp\left(\frac{-V}{20}\right) \quad \beta_h = \frac{1}{\exp(3 - 0.1V) - 1}$$

One of the causes of epileptic seizure is bifurcation in the neurological system with the influence of either sodium channel conductance or potassium channel conductance. In this paper, bifurcation occurs due to a change in sodium channel conductance is considered. The parameters of Neuron using HH model under normal operating conditions and epileptic seizure conditions are given in Table. I

TABLE I. PARAMETERS OF NEURON USING HODGKIN AND HUXLEY MODEL UNDER NORMAL AND EPILEPTIC CONDITIONS

Parameter	Normal values	Epileptic (Abnormal) values
External Current injection $I_{ext} \left(\frac{\mu A}{cm^2}\right)$	0	0
Sodium Channel Conductance $g_{Na} \left(\frac{ms}{cm^2}\right)$	120	550
Potassium Channel Conductance $g_K \left(\frac{ms}{cm^2}\right)$	36	36
Leakage Channel Conductance $g_L \left(\frac{ms}{cm^2}\right)$	0.3	0.3
Sodium Channel resting potential $V_{Na} (mV)$	115	115
Potassium Channel resting potential $V_K (mV)$	-12	-12
Leakage channel resting potential $V_L (mV)$	10.613	10.613
Membrane Capacitance $C_M \left(\frac{\mu F}{cm^2}\right)$	0.91	0.91
Activating molecule of the potassium channel (n)	0.3177	0.3177
Activating molecule of the Sodium Channel (m)	0.05301	0.05301
Inactivating molecule of the Sodium Channel (h)	0.5960	0.5960

To develop control scheme, the nonlinear HH model described by equations (1)–(4) is linearized around the steady state operating points given in Table.I

The transfer function model under normal and epileptic seizure conditions are given below:

Normal:

$$\frac{V(s)}{I_{ext}(s)} = \frac{s^3 + 4.5251s^2 + 1.2914s + 0.0909}{s^4 + 4.8251s^3 + 2.6490s^2 + 0.4783s + 0.0273}$$

Epileptic seizure (Abnormal):

$$\frac{V(s)}{I_{ext}(s)} = \frac{s^3 + 3.6561s^2 + 1.1095s + 0.0848}{s^4 + 3.9561s^3 + 2.2063s^2 + 0.4176s + 0.0254}$$

State space model of the system is

$$\dot{x} = A x + B u$$

$$y = C x$$

Under normal/healthy condition of neuron ,the state space model is given by

$$\begin{bmatrix} \dot{V} \\ \dot{m} \\ \dot{n} \\ \dot{h} \end{bmatrix} = \begin{bmatrix} -0.30 & 0 & 0 & 0 \\ 0.0034 & -0.1832 & 0 & 0 \\ 0.0154 & 0 & -4.2245 & 0 \\ -0.0035 & 0 & 0 & -0.1174 \end{bmatrix} \begin{bmatrix} V \\ m \\ n \\ h \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} I_{ext}$$

$$Y = [1 \quad 0 \quad 0 \quad 0] \begin{bmatrix} V \\ m \\ n \\ h \end{bmatrix}$$

Under Epileptic seizure (Abnormal) condition, the state space model is represented by

$$\begin{bmatrix} \dot{V} \\ \dot{m} \\ \dot{n} \\ \dot{h} \end{bmatrix} = \begin{bmatrix} -0.3000 & 0 & 0 & 0 \\ 0.0042 & -0.1947 & 0 & 0 \\ 0.0206 & 0 & -3.3307 & 0 \\ -0.0027 & 0 & 0 & -0.1307 \end{bmatrix} \begin{bmatrix} V \\ m \\ n \\ h \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} I_{ext}$$

$$Y = [1 \quad 0 \quad 0 \quad 0] \begin{bmatrix} V \\ m \\ n \\ h \end{bmatrix}$$

The open loop response of the neuron under normal and epileptic seizure conditions are shown in Fig. 1. For the normal person, a single neuron generates a single pulse of action potential for external membrane current zero. Under epileptic seizure condition due to bifurcation, a single neuron cell produces a series of spikes.

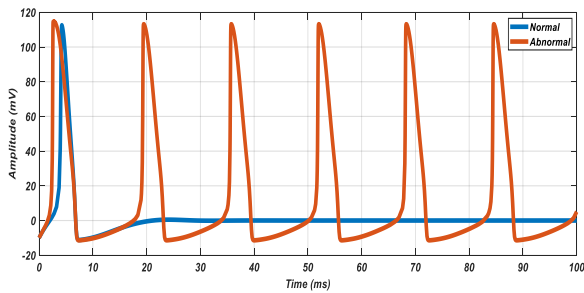


Fig.1 Action potential under normal and epileptic conditions

### III. DEVELOPMENT OF RECONFIGURABLE PI CONTROL AND MPC SCHEMES:

In this section, PI controller and model predictive controller are designed using transfer function model and state space model for a normal healthy person and epileptic seizure person. In addition, reconfigurable control scheme is designed to retune the controller parameters of PI and MPC based on the parameters of abnormal condition.

#### A.PI Control Scheme

The block diagram in Fig.2 shows the PI control scheme for a brain neuron. The PI control parameters are tuned for a normal person and an epileptic seizure person to control action potentials by manipulating total ionic current through neuron cell membrane. The PI controller parameters are given in Table. II.

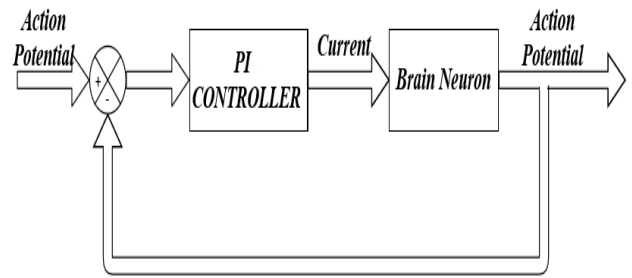


Fig.2. Closed loop control scheme for brain neuron with PI Control

TABLE II. PI CONTROLLER PARAMETERS FOR NORMAL AND EPILEPTIC PERSONS

Parameters	Normal	Epileptic
Kp	0.31169	2.5288
Ki	0.280521	4.722

Most of the automation industries use practical PI controller for controlling various process variables. However, in health care application, optimal control theory can provide valuable insight to develop effective control strategies to avoid mortality and morbidity. Hence, the epileptic seizure suppression is performed with an optimal control strategy. It is discussed in the subsequent subsection.

#### B. MPC scheme

The block diagram of closed loop system with MPC is shown in Fig 3. The controller parameters are tabulated in Table III.

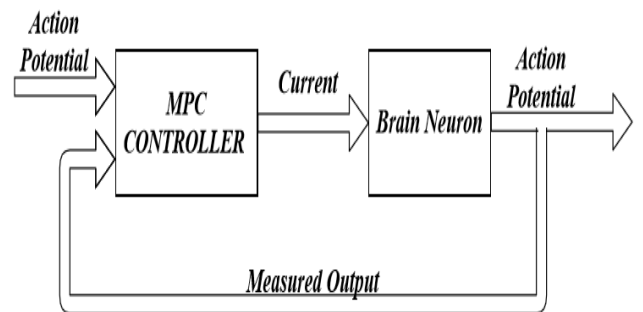


Fig.3. Closed loop Control Scheme for brain neuron with MPC

TABLE III. MPC CONTROLLER PARAMETERS FOR NORMAL AND EPILEPTIC PERSONS

Parameter	Normal	Epileptic
Control Horizon	3	5
Predicted Horizon	5	20

#### C. Reconfigurable control scheme

The reconfigurable control loop is working with normal PI and MPC controller parameters when there is no abnormality detected. When the abnormality is identified in brain neuron, then the control loop will reconfigure its parameter to suppress the spike due to abnormality. The block diagram of the proposed reconfigurable control scheme is shown in Fig.4

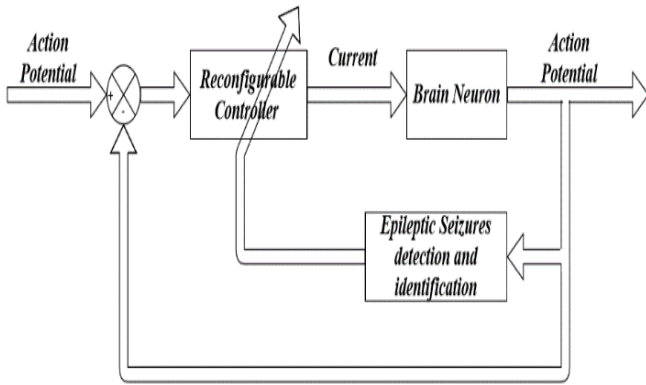


Fig.4. Reconfigurable control scheme

IV. CLOSED LOOP PERFORMANCE OF PROPOSED CONTROL SCHEMES

A. Reconfigurable control scheme with PI controller

The closed loop performances of the system with reconfigurable PI controller for normal healthy condition, epileptic condition and recovered conditions are shown in Fig.5.

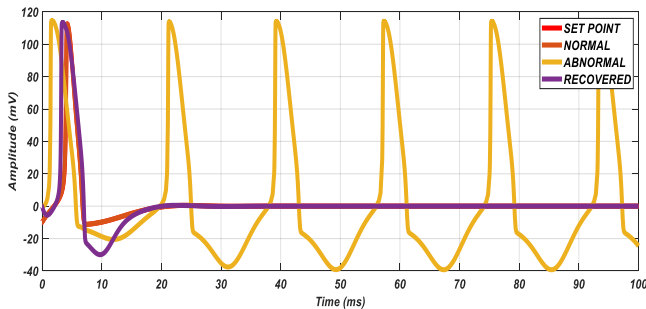


Fig. 5. Closed response of reconfigurable control scheme with PI Controller

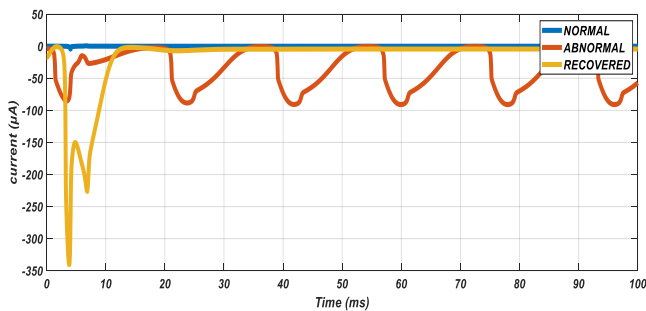


Fig.6. Associated manipulated variables of reconfigurable PI Control scheme

The normal healthy condition of the brain is simulated with sodium channel conductance of 120ms/cm<sup>2</sup>. From the Figure 5, it is found that under a healthy condition, brain neuron produces a single pulse with 110mv of maximum peak and width around 6ms. The epileptic seizure condition is simulated with sodium channel conductance of 550ms/cm<sup>2</sup>. Under epileptic/abnormal condition, bifurcation takes place and the neuron generates a series of spikes for the healthy condition neuron controller parameters. To suppress the spike potentials, the controller parameters are retuned based on the model developed under epileptic seizure condition. The recovered brain

neuron generates action potential with a single pulse and shown in Fig.5. The corresponding manipulated variables are shown in Fig.6. The performance measures of control Effort (CE), Integral Absolute Error (IAE) and Integral Square Error (ISE) of the proposed reconfigurable control scheme with PI controller are tabulated in Table IV. It is observed that ISE, IAE and CE are very small under normal/healthy condition of neuron. It is also noted that these values are increased abruptly from their normal values under abnormal/Epileptic condition. After tuning the controller parameter properly by reconfigurable control mechanism based on abnormal model, the performance measures are reduced from their abnormal values.

TABLE IV. PERFORMANCE MEASURES OF RECONFIGURABLE PI CONTROLLER

PI-Controller	NORMAL	ABNORMAL	RECOVERED
CE	0.3582	260900	221100
ISE	65.37	216300	8476
IAE	9.019	3568	212.3

B. Reconfigurable control scheme with MPC

The closed loop performance of the system with reconfigurable Model Predictive Controller for normal healthy condition, epileptic condition and recovered conditions are shown in Fig.7. The manipulated variables associated with different conditions are shown in Fig.8. When it is compared with a reconfigurable control scheme with PI controller, the number of spikes generated by MPC control is less than the number of spikes produced by PI controller. Simulation results show that Model Predictive Controller outperforms PI Controller in epileptic seizures condition. The performance measures of MPC controllers are tabulated in Table V. As compared to reconfigurable PI controller, MPC provides better performances.

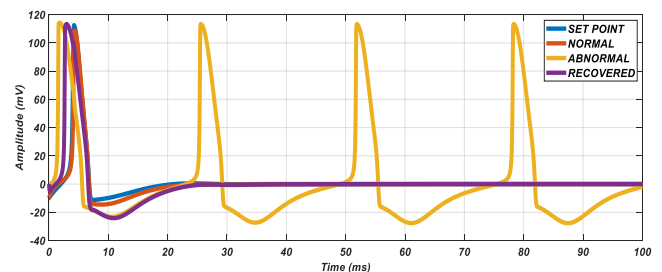


Fig. 7. Closed response of reconfigurable control scheme with Model Predictive Controller

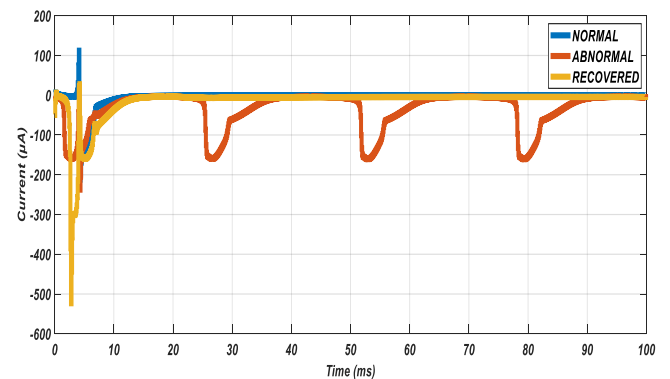


Fig. 8. Associated manipulated variables of reconfigurable controller with Model Predictive Controller



Table V. PERFORMANCE MEASURE OF RECONFIGURABLE CONTROLLER WITH MODEL PREDICTIVE CONTROLLER

MPC-Controller	NORMAL	ABNORMAL	RECOVERED
CE	51.81	221100	132582
ISE	2.266	123100	5373
IAE	2.383	1938	172.9

### C. Stability Analysis

The stability and bifurcation analysis are carried out for open loop model and closed loop control systems. The effect of sodium conductance ( $g_{Na}$ ) on action potential is studied. The actual values of  $g_{Na}$  for normal healthy person and epileptic persons are 120 mS/cm<sup>2</sup> and 550 mS/cm<sup>2</sup> respectively. The eigenvalues of neuron model under normal and epileptic seizure conditions are tabulated in Table VI. From the eigenvalue analysis it is observed that  $g_{Na}$  increases eigenvalues of the system move towards imaginary axis from left side of s-plane. Hence the stability of the system decreases as  $g_{Na}$  increases.

TABLE VI. STABILITY ANALYSIS OF OPEN LOOP SYSTEM

Parameter	Normal ( $G_{NA} = 120$ )	Abnormal ( $G_{NA} = 550$ )
Eigen Values	-4.2245 -0.3000 -0.1832 -0.1174	-3.3307 -0.3000 -0.1947 -0.1307

The eigenvalues of the neuron under closed loop condition with normal, epileptic and recovered conditions are evaluated in TABLE VII. The table shows the overdamped open loop system becomes underdamped system under closed loop condition and provides complex conjugate poles. From the normal and abnormal closed loop system analysis, it is found that complex eigenvalues are unaffected by epileptic seizure in closed loop condition. In recovered condition, to suppress the spikes controller take more control effort and shift the complex eigenvalues to the left side from their original positions.

TABLE VII. STABILITY ANALYSIS OF CLOSED LOOP SYSTEM

Parameter	Normal	Abnormal	Recovered
Eigen Values	-4.22 + 0.00i -0.31 + 0.43i -0.31 - 0.43i -0.18 + 0.00i -0.12 + 0.00i	-3.33 + 0.00i -0.31 + 0.43i -0.31 - 0.43i -0.20 + 0.00i -0.131 + 0.00i	-3.33 + 0.00i -1.42 + 1.65i -1.42 - 1.65i -0.20 + 0.00i -0.131 + 0.00i

## V. CONCLUSION

In this paper, linear models are developed for brain neuron using HH model under a healthy condition and epileptic seizure condition. Based on models, reconfigurable control schemes with PI controller and Model predictive controller are proposed to control epileptic seizure. Then the performances of the proposed control schemes are compared qualitatively as well as quantitatively. From the comparison, it is found that MPC

based reconfigurable controller provides better performance than PI based reconfigurable control scheme in terms of frequency of occurrence and control effort. Finally, stability of the proposed reconfigurable control scheme is investigated.

## REFERENCES

- [1] Roxana A. Stefanescu, R.G. Shivakeshavan and Sachin S. Talathi, "Computational models of epilepsy" Seizure ,Vol,21, pp.748-759, 2012.
- [2] Wim van Drongelen, " Modeling Neural activity", Biomathematics, Vol. pp.1-37,2013.
- [3] Nambi Narayanan. S, Sutha. S and Divya G " Development of nonlinear model for human heart " International Conference Trends in Measurement and Automation (TIMA), 2017
- [4] R. Ozgur Doruk, "Control of the repetitive firing in the squid giant axon using electrical fields", Middle East Technical University,pp.no.1-14, 2018.
- [5] Yexin Lin, Haiyuan Liu\* and Li Hu, "LQR Suppression of Hopf Bifurcation in HodgkinHuxley Neurons" NeuroQuantology, Vol.17, pp.22-32, 2019.
- [6] Shalong Chen, Yuan Zou, and Xudong Zhang, " An Efficient Method for Hopf Bifurcation Control in Fractional-Order Neuron Model" IEEE Access,Vol.7,pp.77490-77498, 2019.
- [7] Shuang Liu, Hongling Ai, Zhenfang Pang, Zhenjun Lin & Dingxuan Zhao, " Hopf bifurcation control of nonlinear electromechanical coupling main drive system of rolling mill"European Physical Journal plus,Vol.135,pp.1-14, 2020.
- [8] Harleen K. Brar, Ioannis Exarchos, Yungpeng Pan, Evangelos Theodorou, and Babak Mahmoudi, "Seizure Reduction using Model Predictive Control",Vol.18, 2018.
- [9] Rieke Fruengel, Timo Bröhl, Thorsten Rings & Klaus Lehnertz "Reconfiguration of human evolving large-scale epileptic brain networks prior to seizures: an evaluation with node centralities", scientific report natural research,Vol.NO. 21921, pp.no.1-10,2020
- [10] Chen Liu, Ge Zhao, Jiang Wang, Hao Wu, Huiyan Li, Chris Fietkiewicz, And Kenneth A. Loparo," Neural Network-Based Closed-Loop Deep Brain Stimulation for Modulation of Pathological Oscillation in Parkinson's Disease", Vol. 8, 2020.
- [11] Andrew Haddock , Anca Velisar , Jeffrey Herron , Helen Bronte-Stewart and Howard J. Chizeck, "Model Predictive Control of Deep Brain Stimulation for Parkinsonian Tremor" International conference on Neural Engineering Shanghai, China, May 25 - 28 , 2017.
- [12] Yue Zhang, Kuanquan Wang, Yongfeng Yuan, Dong Sui, Henggui Zhang,"Stability and Bifurcation Analysis of Hodgkin-Huxley Model", International conference on bioinformatics and biomedicine Vol. 1, pp: 49-54, Shanghai, China, 2013.
- [13] Sutha.S, Lavanya.S and Thyagarajan.T, "Multi-objective reconfigurable output feedback controller for MIMO system using MOEAs", International Journal of Chemical Engineering Communications, Vol.199, No.9,pp.1125-1143, 2012.
- [14] John E. Fleming, Eleanor Dunn and Madeleine M. Lowery, "Simulation of Closed-Loop Deep Brain Stimulation Control Schemes for Suppression of Pathological Beta Oscillations in Parkinson's Disease" Journal of Frontiers in Neuroscience, Vol.14,pp.1-22, 2020.
- [15] Ning Li, Haiyi Sun,and Qingling Zhang, "The Controller Design of the Epilepsy Therapy Apparatus" Journal of Mathematical Problems in Engineering, pp.1.8, 2017.

## NON-INVASIVE ESTIMATION OF SALIVARY BIOMARKERS FOR THE DIAGNOSIS OF DIABETIC NEPHROPATHY

Purnima. S<sup>1</sup>, Atchaya. V<sup>2</sup>, Avesha. R<sup>3</sup>, Devi Lakshmi. A<sup>4</sup>, Divya Bharathi. S<sup>5</sup>, Dr.  
Prasanna. S

<sup>1</sup>Asst.Professor, <sup>2,3,4,5</sup>UG Scholar, Department of Biomedical Engineering  
Jerusalem College of Engineering, Pallikaranai, Chennai-600 100. <sup>6</sup>Senior Lecturer, Department of Oral  
and Maxillofacial Pathology, Tagore Dental College and Hospital, Chennai.  
Correspondence E-mail: [purni\\_41@rediffmail.com](mailto:purni_41@rediffmail.com) and [devilakshmi186@gmail.com](mailto:devilakshmi186@gmail.com)

**ABSTRACT:** Diabetic Nephropathy (DN), also known as diabetic kidney disease, is the chronic loss of kidney function occurring in those with diabetes mellitus. The development of Diabetic Nephropathy is easier to treat if diagnosed in early stage (appearance of abnormal amount of salivary biomarkers). The conventional methods are mostly invasive and have many limitations like irregular monitoring of the glucose, creatinine, urea etc. Saliva is a multi-constituent biological fluid. It is one of the most abundant secretions and its collection is easy and non-invasive. So, we proposed this novel method for the estimation of salivary biomarkers for the diagnosis of Diabetic Nephropathy. The objective of the project is to reduce the dependence of hemodialysis and kidney transplantation of person suffering from ESRD by estimation of glucose by GOD-POD method, creatinine by Jaffe's method and urea by Berthelot-Urease method.

**KEYWORDS:** Diabetic Nephropathy, Saliva, Glucose, Creatinine, Urea.

### I. INTRODUCTION

Diabetic Mellitus (DM) is a clinical syndrome characterized by abnormalities in carbohydrate, lipid and protein metabolism that leads to the looks of varied complications like retinopathy,

nephropathy, neuropathy, cardiovascular symptoms etc.

Presently, the diagnosis of DM is completed by measuring serum blood sugar levels by a uniform method which are invasive and traumatic to the patient both physically and psychologically. Hence, the present method discourages the individuals from undergoing investigations leading to lack of sufficient diagnosis thus attributing DM as a serious explanation for death worldwide. So, in recent years, efforts are made to exchange blood investigation with another biological material sample that would be collected by non-invasive procedures. Saliva has been put forth as a possible diagnostic tool for surveillance of diseases thanks to its several advantages. It clearly offers a cheap, simple and straightforward approach to use as a screening method.

Saliva offers some distinctive advantages. The entire saliva is often collected during a non-invasive way by individuals with limited training. Diagnosis of disease through the analysis of saliva may be a potential value for youngsters and older adults as a set of the fluid is related to fewer compliance problems as compared with the gathering of blood [1].

Several systemic diseases are

changes in salivary secretion. Chronic renal disorder is one among the systemic diseases which will affect the competition of salivary secretions. More importantly, saliva can indicate, creatinine and urea levels in patients with CKD which are the parameters usually assist in blood samples [2]. Biochemical markers play a crucial role within the accurate diagnosis and in assessing risk and adopting therapy to enhance clinical outcome [3].

The main aim of this study is to estimate the salivary biomarkers like glucose, creatinine and urea non-invasively for the diagnosis of diabetic nephropathy.

## II. MATERIALS AND METHOD

In this study, a total of 30 subjects was selected with a median age group of about 18-60 years and investigated. Out of 30 subjects, 10 subjects were normal patient, 10 subjects were Diabetic Mellitus patients and 10 subjects were Diabetic nephropathy patients. All the 30 subjects are subjected to three tests. They are the estimation of glucose, creatinine and urea.

### (i) Collection of Samples:

Subjects were advised to report 1½ hours after breakfast for serum and saliva collection. About 2 ml blood was collected by veni-puncturing method in an EDTA tube and it was centrifuged at 4000 rpm for 5-10 minutes. The serum alone is used for the estimation. For the collection of saliva, patient is advised to avoid food intake 2hrs prior to collection. Saliva was collected by making the patient rinse his/her mouth with

reported to supply marked and identifiable

water and were insisted to open mouth for five min without swallowing, and about 2ml of saliva was collected in a sterile container by the spitting method. The whole saliva is centrifuged at 2000 rpm for 2-3 minutes. The supernatant alone is used for the estimation.

### (ii) Estimation of Glucose:

Both serum and salivary glucose level were estimated using the Glucose Oxidase-Peroxidase (GOD - POD) method. Four test tubes were taken. The test tubes are labeled as 'Blank', 'Standard', 'Test Serum' and 'Test Saliva' respectively. About 1 ml of Glucose enzyme reagent is added in each test tube. Then, 10µl of the standard was added to the tube marked as 'Standard', followed by 10µl of the test serum sample to the 'Test Serum' test tube and 10µl of the test saliva sample to the 'Test Saliva' test tube. After mixing well, the test tubes were kept incubated at 37°C for 15 minutes. The reagent "Blank" was aspirated in a colorimeter first, followed by standard solution and therefore the absorbance values were noted at 546nm ( $\pm 20$  nm). Then, the test serum sample followed by test saliva sample was aspirated and the absorbance values were noted at 546nm ( $\pm 20$  nm). Finally, the glucose was estimated by using the formula.

### (iii) Estimation of Creatinine:

Both serum and salivary creatinine level were estimated using the Jaffe's method (Alkaline Picrate Method). Four test tubes were taken. The test tubes are labeled as 'Blank', 'Standard', 'Test Serum' and 'Test Saliva' respectively. About 1 ml of

working reagent is added in each test tube. Then, 50 $\mu$ l of the standard was added to the tube marked as 'Standard', followed by 50 $\mu$ l of the test serum sample to the 'Test Serum' test tube and 50 $\mu$ l of the test saliva sample to the 'Test Saliva' test tube. Mix the content well. The reagent "Blank" was aspirated in a colorimeter first, followed by standard solution and therefore the initial absorbance values (A1) were noted exactly after 30 seconds at 490-520 nm. Then, the test serum sample followed by test saliva sample was aspirated and the initial absorbance values (A1) were noted exactly after 30 seconds.

The other absorbance values (A2) for the standard, test serum and test saliva were noted exactly after 120 seconds. Then, the change in absorbance ( $\Delta A$ ) for the standard, test serum and test saliva were calculated. Finally, the creatinine was estimated by using the formula.

**(iv) Estimation of Urea:**

Both serum and salivary urea level were estimated using the Berthelot-Urease method. Four test tubes were taken. The test tubes are labeled as 'Blank', 'Standard', 'Test Serum' and 'Test Saliva' respectively. About 1 ml of working reagent is added in each test tube. Then, 10 $\mu$ l of the standard was added to the tube marked as 'Standard', followed by 10 $\mu$ l of the test serum sample to the 'Test Serum' test tube and 10 $\mu$ l of the test saliva sample to the 'Test Saliva' test tube. After mixing well, the test tubes were kept incubated at 37°C for 5 minutes. Then about 1 ml of chromogen reagent is added in all the test tubes and after mixing well, it is incubated at 37° C for 5 minutes. The reagent "Blank" was aspirated in a colorimeter first, followed by standard

solution and therefore the absorbance values (Abs.S) were noted at 600nm. Then, the test serum sample followed by test saliva sample was aspirated and the absorbance values (Abs.T) were noted at 600nm. Finally, the urea was estimated by using the formula.

### III. RESULTS

In this study total of 30 patients were selected. Of which 10 were control group, 10 were diabetic and 10 were diabetic nephropathy patient. In the control group out of 10 patients, 5 patients had serum glucose level below 100 mg/dl and 5 of them with the level of 100–140 mg/dl. In diabetic patient, out of 10 patients, 7 of them had the serum glucose level of 155–200 mg/dl, 3 had a level of 200–250 mg/dl. Data obtained from 10 control group patients were shown in Table No 1. The salivary glucose level for control was between 0.2-1.0 mg/dl and for a diabetic the level was between 1.0-9.6 mg/dl. However, the salivary glucose levels were highly inconsistent with respect to the serum blood glucose levels, i.e., the salivary glucose levels were not consistent with the severity of diabetes. The serum and salivary creatinine level for control was between 0.6-1.1 mg/dl and between 0.05-0.27 mg/dl respectively. The serum and salivary urea level for control was between 10-40 mg/dl and between 12-30 mg/dl.

Out of 10 diabetic patients, 5 of them had serum urea level of 15-25 mg/dl and the remaining 5 had 15-50 mg/dl. The salivary urea level of 10 patients was between 30-45 mg/dl. The value of creatinine in the serum of 10 diabetic patients was between 1.8-2.5 mg/dl and in saliva 0.3-0.9 mg/dl. The comparison of serum urea values with salivary urea values showed that serum urea

values were significantly lower than salivary urea values. Similarly, the comparison of serum creatinine values with salivary creatinine showed that in all the group

serum creatinine values was significantly similar to salivary creatinine. Data obtained from 10 diabetic patients were shown in Table No 2.

PATIENT	AGE	SEX	BLOOD GLUCOSE (mg/dl)	SALIVA GLUCOSE (mg/dl)	BLOOD CREATININE (mg/dl)	SALIVA CREATININE (mg/dl)	BLOOD UREA (mg/dl)	SALIVA UREA (mg/dl)
1	21	FEMALE	83	0.35	0.89	0.27	13.61	25.71
2	25	FEMALE	107	0.5	0.83	0.20	12.59	18.15
3	27	MALE	89	0.4	0.55	0.11	7.93	13.67
4	30	MALE	102	0.47	0.75	0.18	12.36	18.11
5	33	FEMALE	97	0.2	0.63	0.08	5.29	12.53
6	35	FEMALE	104	0.75	0.81	0.20	12.55	18.15
7	45	FEMALE	92	0.87	0.72	0.24	8.01	13.79
8	46	MALE	96	0.23	0.61	0.08	5.27	12.51
9	50	MALE	110	0.3	0.87	0.23	12.87	18.56
10	55	FEMALE	99	0.61	0.70	0.12	8.20	15.67

Table No 1. Results for control group (Normal Patients)

PATIENT	AGE	SEX	BLOOD GLUCOSE (mg/dl)	SALIVA GLUCOSE (mg/dl)	BLOOD CREATININE (mg/dl)	SALIVA CREATININE (mg/dl)	BLOOD UREA (mg/dl)	SALIVA UREA (mg/dl)
1	44	FEMALE	135	1.12	1.7	0.3	28	30
2	55	MALE	198	2.9	1.9	0.6	30	35
3	45	FEMALE	168	4.5	1.8	0.2	29	38
4	46	MALE	175	1.7	1.9	0.4	34	32
5	49	FEMALE	247	8.7	2.4	0.3	29	44
6	50	FEMALE	243	4.3	2.3	0.9	22	43
7	53	MALE	242	2.6	2.2	0.9	24	42
8	60	FEMALE	230	1.5	2.1	0.8	40	41
9	62	MALE	245	7.9	2.3	0.9	24	42
10	65	MALE	220	3.6	2.0	0.7	39	40

Table No 2. Results for Diabetic Patients

PATIENT	AGE	SEX	BLOOD GLUCOSE (mg/dl)	SALIVA GLUCOSE (mg/dl)	BLOOD CREATININE (mg/dl)	SALIVA CREATININE (mg/dl)	BLOOD UREA (mg/dl)	SALIVA UREA (mg/dl)
1	55	FEMALE	280	14.1	2.2	1.15	110	45.11
2	55	MALE	310	18.5	3.1	1.40	121	50.33
3	59	FEMALE	278	14.7	2.1	1.35	102	49.01
4	60	FEMALE	410	19.6	5.1	1.65	121	64.32
5	63	MALE	325	18.7	3.3	1.49	124	54.21
6	65	FEMALE	290	15.2	2.5	1.21	117	51.89
7	69	FEMALE	320	16.8	3.3	1.45	125	51.01
8	70	MALE	350	17.1	3.8	1.52	125	59.33
9	73	FEMALE	300	15.6	3.1	1.56	120	53.98
10	75	FEMALE	299	14.9	2.7	1.32	111	49.53

Table No 3. Results for Diabetic Nephropathy Patients

As expected, the serum creatinine and urea were significantly higher in patients with CKD. Among all group's serum and salivary glucose, urea and creatinine levels were higher in CKD patients followed by diabetic patients and least in control group patients. The increase in the serum urea, glucose, and creatinine level may be an implication of general disease like renal insufficiency which can cause increased creatinine and uric acid level as a result of reduction of blood flow to the kidneys. It was also observed that the increase of glucose, creatinine and urea level in the diabetic patient might lead to a chronic kidney disease called diabetic nephropathy. At this stage, the patient will completely fail the kidney function. The patient has to depend on frequent dialysis. In serum, it was noted that the value of glucose lies between 220-250 mg/dl, creatinine was between 2-6 mg/dl and urea was between 100-126 mg/dl in serum. In saliva, the value of glucose was between 14-20 mg/dl, creatinine was between 1.0-1.75 mg/dl and urea was between 45-65 mg/dl. Data obtained from 10 diabetic nephropathy patients were shown in Table No 3.

#### IV. DISCUSSION

Diabetes mellitus may be a globally widespread disease that various diagnostic devices are now available. During the past three decades, the incidence and prevalence of End Stage Renal Disease (ESRD) has risen progressively. most significant reasons for rapid increase in CKD patients are rapidly increasing worldwide incidence of diabetes and hypertension.

Saliva may be a wonderful biomarker of early disease detection that results in more practical diagnosis and treatment or monitor systemic illness, like diabetes, diabetic

nephropathy etc., and this study suggests that the saliva contributes a straightforward, quick, and an affordable method for screening of diabetic autonomic nephropathy and diabetic mellitus. Unstimulated whole saliva has been employed in the bulk of diagnostic studies because generally there's a dilution of saliva, modulation in pH in stimulated saliva.

In this study, glucose concentration in unstimulated whole saliva was analyzed. There was a rise in salivary glucose level when there was increase in serum glucose level. Hence, saliva may be used as a screening method in diagnosing DM and diabetic nephropathy.

Creatinine because of its physical properties in an exceedingly healthy state under normal conditions is unable to diffuse easily across the cells and tight intercellular junction of the exocrine gland. But in diseased state, its value increases in saliva possibly thanks to an alteration within the permeability of exocrine gland cells and therefore the increased serum creatinine levels in CKD patients create a amount gradient that facilitates diffusion of creatinine from serum into saliva. Thus, serum creatinine is employed for monitoring disease progression and also salivary creatinine may be used as a monitoring and diagnosing tool. Hence, saliva is used as a screening method in diagnosing diabetes and diabetes nephropathy.

Whenever there's a rise within the blood urea there's concomitant increase in salivary urea also because the kidneys are unable to excrete urea in renal disorder and its concentration in blood increases with increased concentration in saliva thanks to increased serum urea which creates an

increased concentration gradient successively increasing the diffusion of urea from serum to saliva. The serum urea values were significantly below salivary urea values. The serum creatinine values were significantly the same as salivary creatinine. Therefore, salivary creatinine and urea levels correlate well with the serum creatinine and urea respectively so saliva may be used as a non-invasive diagnostic tool for diabetic nephropathy.

## V. CONCLUSION

Frequent monitoring of glucose level is required to cut back its complication. Further studies on large sample size should be done to figure out the diagnostic evaluability of salivary glucose level within the primary diagnosis of diabetic nephropathy. Findings from our study demonstrated elevated levels of salivary glucose, creatinine and urea with correlational statistics to the amount in blood in patients with late-stage. Supported the observations from this study it's often inferred that serum and salivary glucose, urea and creatinine levels were significantly higher in CKD subjects followed by diabetic and normal patients. An increased amount of salivary glucose, urea and creatinine levels were seen only in CKD subjects, diabetic and diabetic nephropathy subjects and no difference was seen in controls. Thus, it often recommends that salivary glucose, urea and creatinine values are often used for screening of renal status in CKD, diabetic and normal subjects. Thus, saliva could even be a stepping stone and has the potential to revolutionize the diagnostic protocol for patients with renal diseases.

## VI. FUTURE WORK

The salivary biomarkers like glucose, creatinine and urea are used to diagnose the diabetic nephropathy. In addition to these, estimation of some other parameters like total protein, albumin, calcium, phosphorous etc. can also be included to diagnose several other diseases like dental caries, oral problems, diabetic Mellitus types etc.

## VII. ACKNOWLEDGEMENT

The authors would like to thanks all the staffs and faculty members of the Department of Oral Pathology at Tagore Dental College and Hospital for their valuable guidance throughout the project.

## VIII. REFERENCES

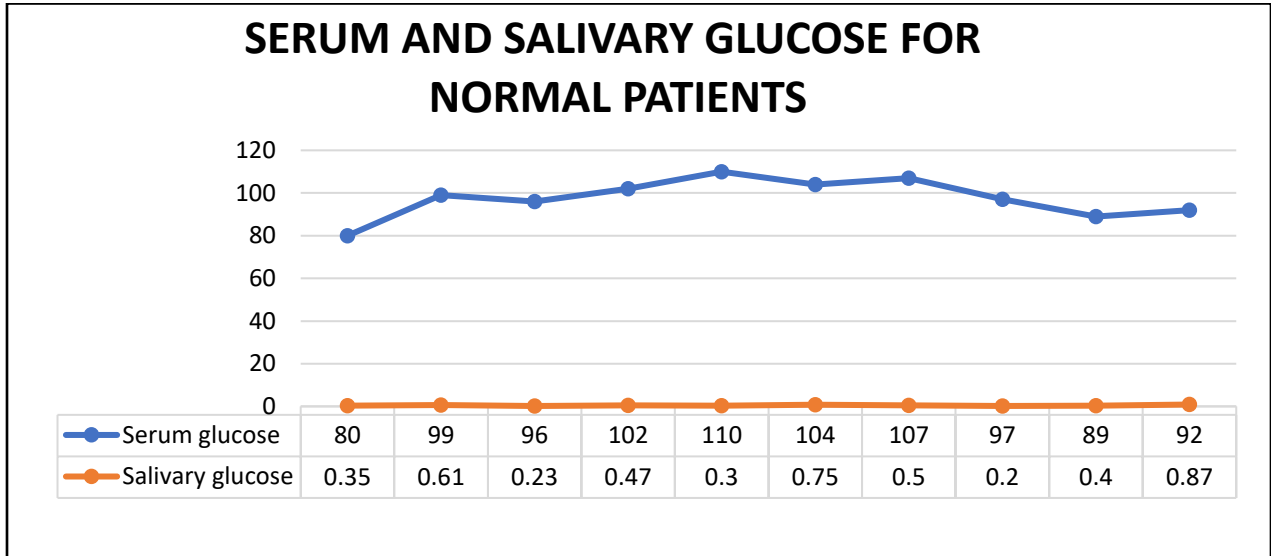
- [1] Harshada Ragunathan, Nalini Aswath, T. Sarumathi, Salivary Glucose Estimation: A non-invasive method, Indian Journal of Dental Sciences, January 29,2020.
- [2] Taye Jemilat Lasisi, Yemi Raheem Raji, Babatunde Lawal Salako, Salivary creatinine and urea analysis in patients with chronic kidney disease: A case control study, BioMed Central, January 16, 2016.
- [3] Amit Ladgotra, Pradhuman Verma, Seetharamaiah Sunder Raj, Estimation of Salivary and Serum biomarkers in Diabetic and Non-Diabetic patients - A comparative study, Journal of Clinical and diagnostic research, June 01, 2016.
- [4] Divya Pandya, Anil Kumar Nagrajappa, K.S. Ravi, Assessment and Correlation of Urea and Creatinine levels in Saliva and Serum of Patients with Chronic Kidney Disease, Diabetes and Hypertension - A research study,

- Journal of Clinical and diagnostic research, October 01,2016. doi:10.1056/nejme1708949. PMID 28854097.
- [5] Ramesh Venkatapathy, Vasupradha Govindarajan, Nirima Oza, Sreejith Parameshwaran, Balamurali Pennagaram, Dhanasekaran, Karthikshree V. Prashad, Salivary Creatinine Estimation as an alternative to Serum Creatinine in Chronic Kidney Disease Patients, Hindawi Publishing Corporation, International Journal of Nephrology, April 10,2014.
- [6] Dr. R. Christeffi Mabel, M.D.S and Dr. Abhinaya. L. M, B.D.S., Noninvasive estimation of Salivary Glucose, Salivary Amylase, Salivary Protein and Salivary pH in Diabetic and Non-Diabetic patients – A case control study, European Journal of Biomedical and Pharmaceutical Sciences, April 09,2018.
- [7] Sugam Shrestha, Prajwal Gyawali, Roojet Shrestha, Bibek Poudel, Manoj Sigdel, Prashant Regmi, Manoranjan Shrestha, Binod Kumar Yadav, Serum Urea and Creatinine in Diabetic and Non-Diabetic Subjects, JNAMLs, Vol.9, No.1, December 2008.
- [8] "Eating Right for Chronic Kidney Disease | NIDDK". National Institute of Diabetes and Digestive and Kidney Diseases. Retrieved 5 September 2019.
- [9] "Global Report on Diabetes" (PDF). World Health Organization. 2016. Retrieved 20 September 2018.
- [10] "Chronic Kidney Disease Tests & Diagnosis". National Institute of Diabetes and Digestive and Kidney Diseases. October 2016. Retrieved 19 December 2017.
- [11] de Boer IH (August 2017). "A New Chapter for Diabetic Kidney Disease". *The New England Journal of Medicine*. 377 (9): 885–887.
- [12] SA. Bamanikar, AA. Bamanikar, A. Arora, Study of Serum Urea and Creatinine in Diabetic and Non-Diabetic Patients in a tertiary teaching hospital, *The Journal of Medical Research* 2016, January-February 2016.
- [13] Glucose in the saliva of the non-diabetic and the diabetic patient, M.J.A.Campbell, *Archives of Oral Biology*, Volume 10, Issue 2, March–April 1965, Pages 197-205.
- [14] S. Malik, S. Gupta, R. Khadgawat and S. Anand, "A novel non-invasive blood glucose monitoring approach using saliva," 2015 IEEE International Conference on Signal Processing, Informatics, Communication and Energy Systems (SPICES), Kozhikode, India, 2015, pp. 1-5, doi: 10.1109/SPICES.2015.7091562.
- [15] A. Soni and S. K. Jha, "Saliva based noninvasive optical urea biosensor," 2017 IEEE SENSORS, Glasgow, UK, 2017, pp. 1-3, doi: 10.1109/ICSENS.2017.8234302.
- [16] Determination of creatinine-related molecules in saliva by reversed-phase liquid chromatography with tandem mass spectrometry and the evaluation of hemodialysis in chronic kidney disease patients, *Analytica Chimica Acta*, Volume 911, 10 March 2016, Pages 92-99.
- [17] Estimation of salivary glucose, amylase, calcium, and phosphorus among non-diabetics and diabetics: Potential identification of non-invasive diagnostic markers, Raphael Enrique G.TiongcoEngracia S.ArceoNicole S.RiveraChastene Christopher D.FlakeArchie R.Policarpio, Elsevier *Diabetes & Metabolic Syndrome*:

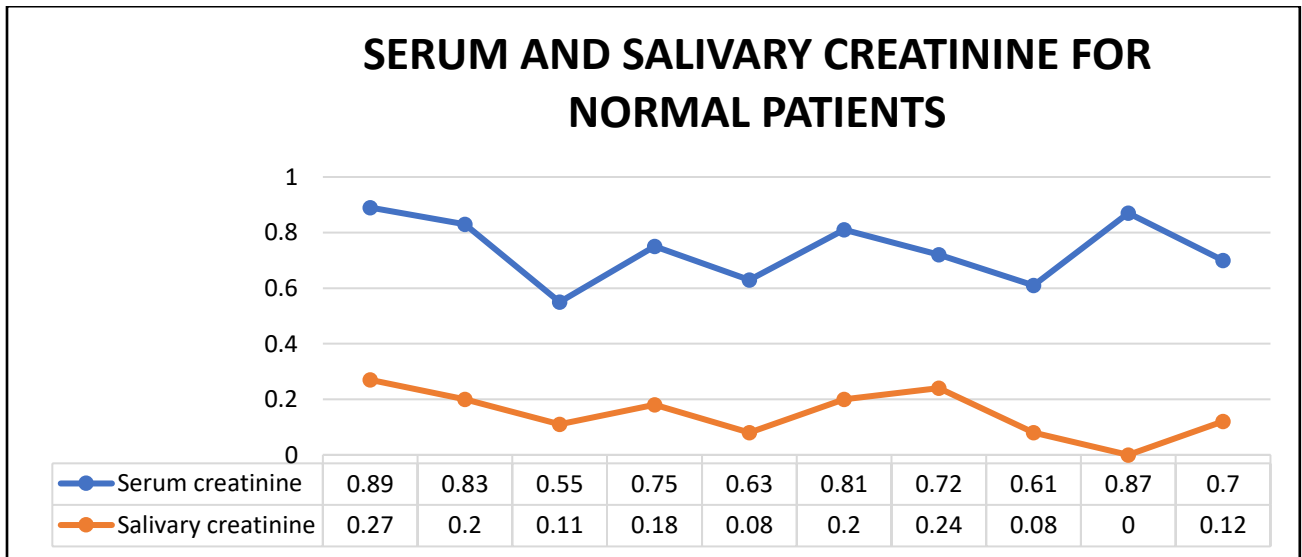


Clinical Research & Reviews, Volume  
 13, Issue 4, July–August 2019, Pages  
 2601-2605.

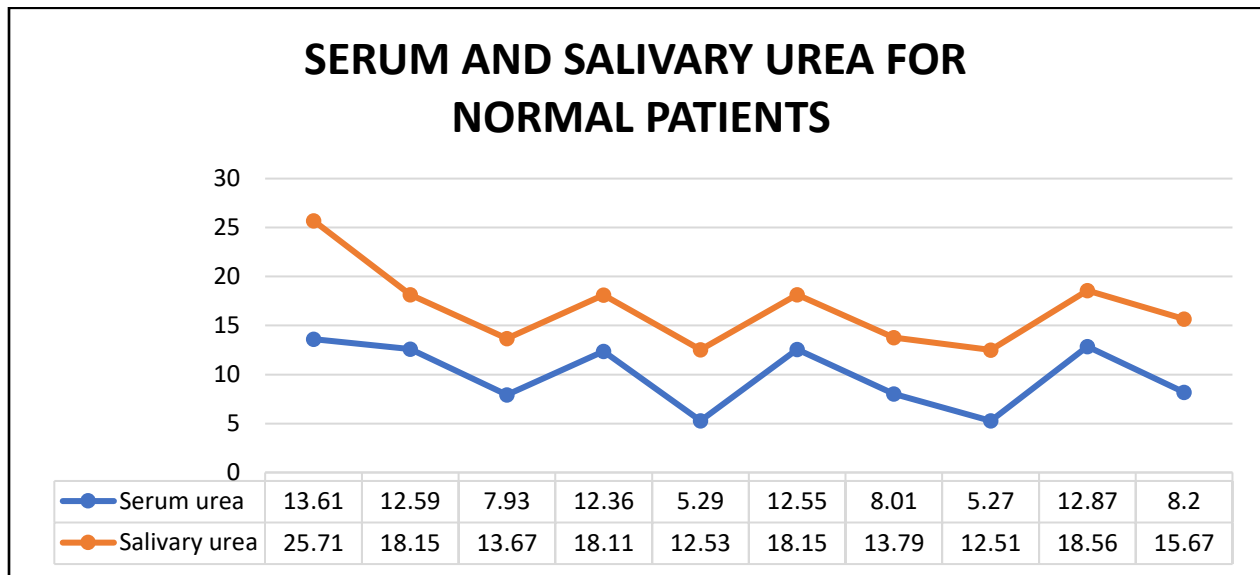
### APPENDIX



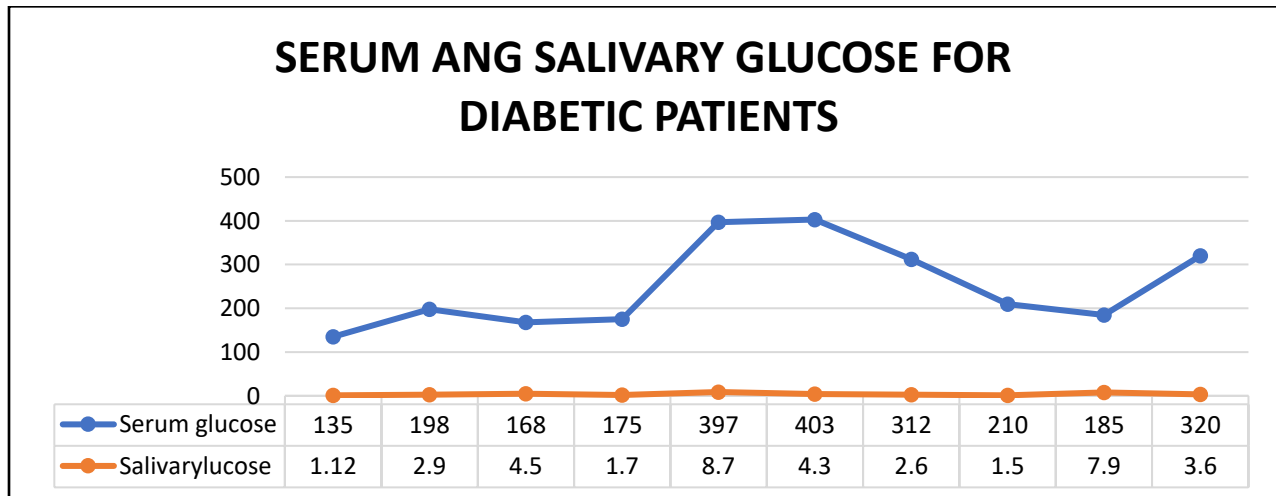
*Figure No 1: Comparison between Serum and Salivary Glucose for Normal Patients*



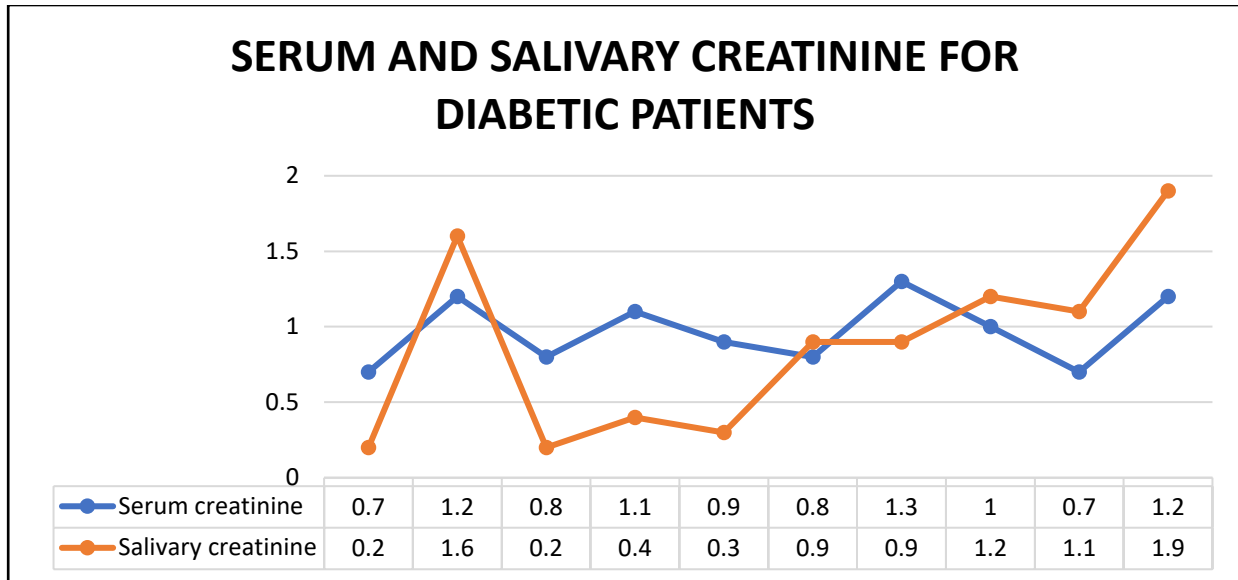
**Figure No 2: Comparison between Serum and Salivary Creatinine for Normal Patients**



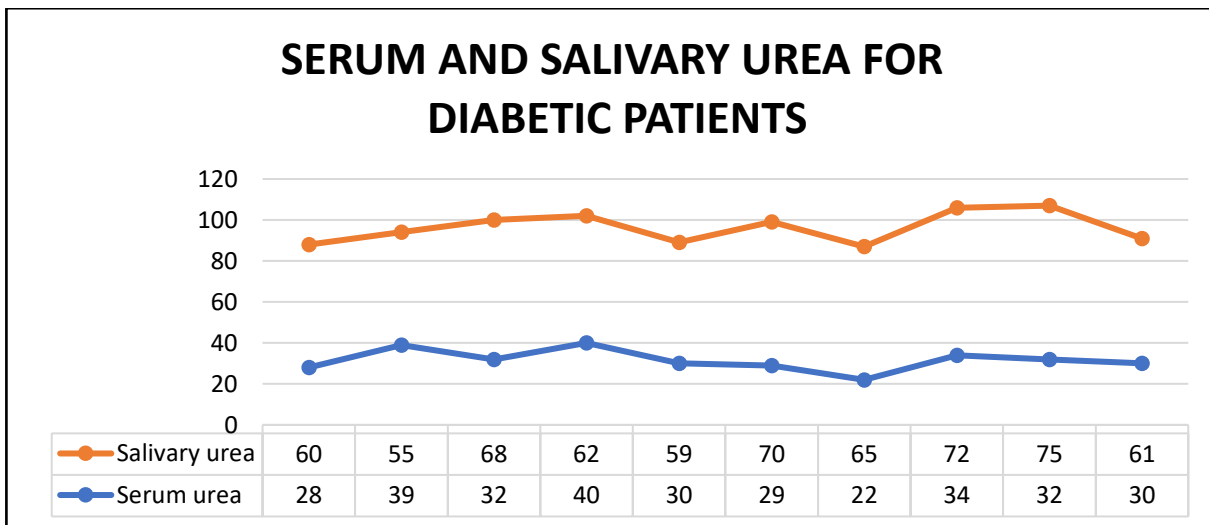
**Figure No 3: Comparison between Serum and Salivary Urea for Normal Patients**



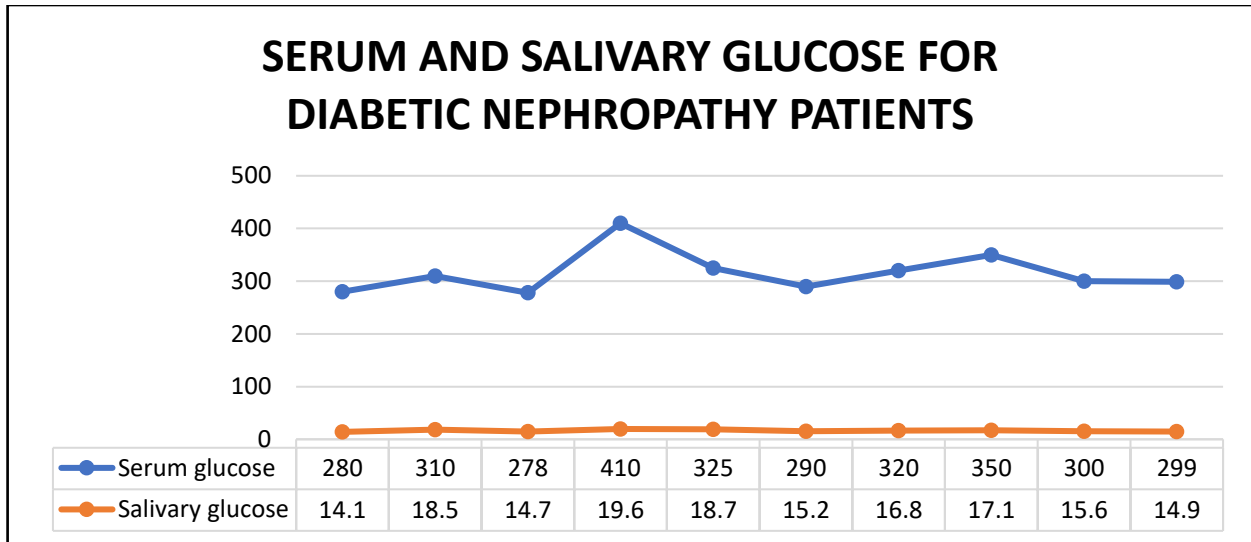
**Figure No 4: Comparison between Serum and Salivary Glucose for Diabetic Patients**



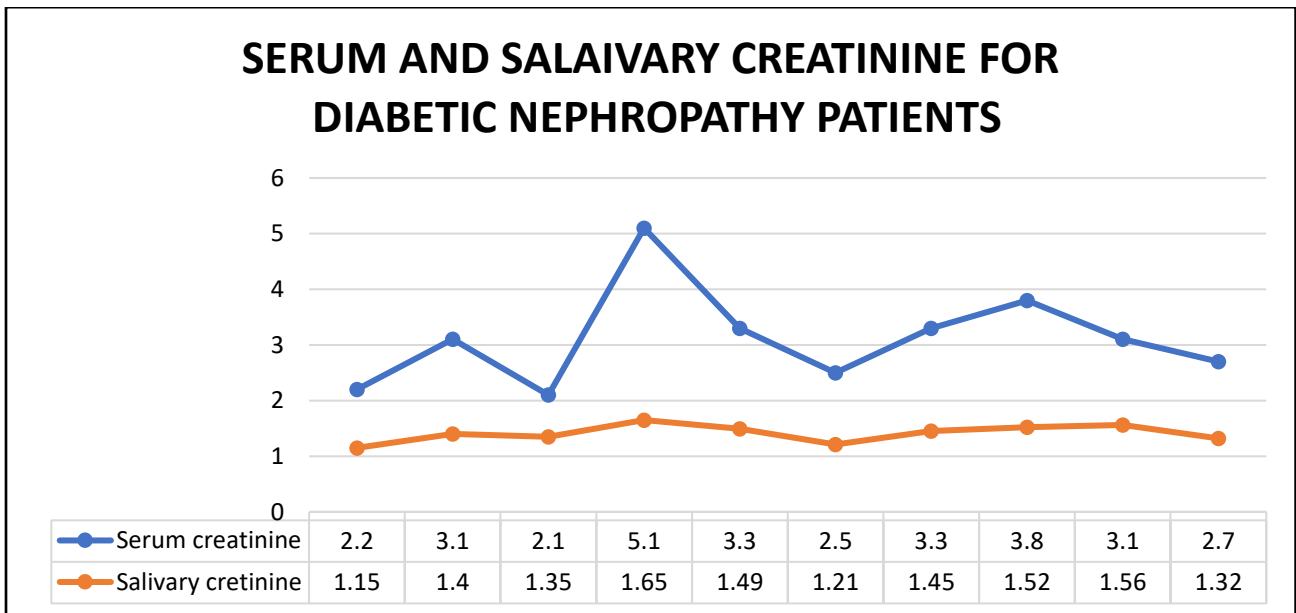
*Figure No 5: Comparison between Serum and Salivary Creatinine for Diabetic Patients*



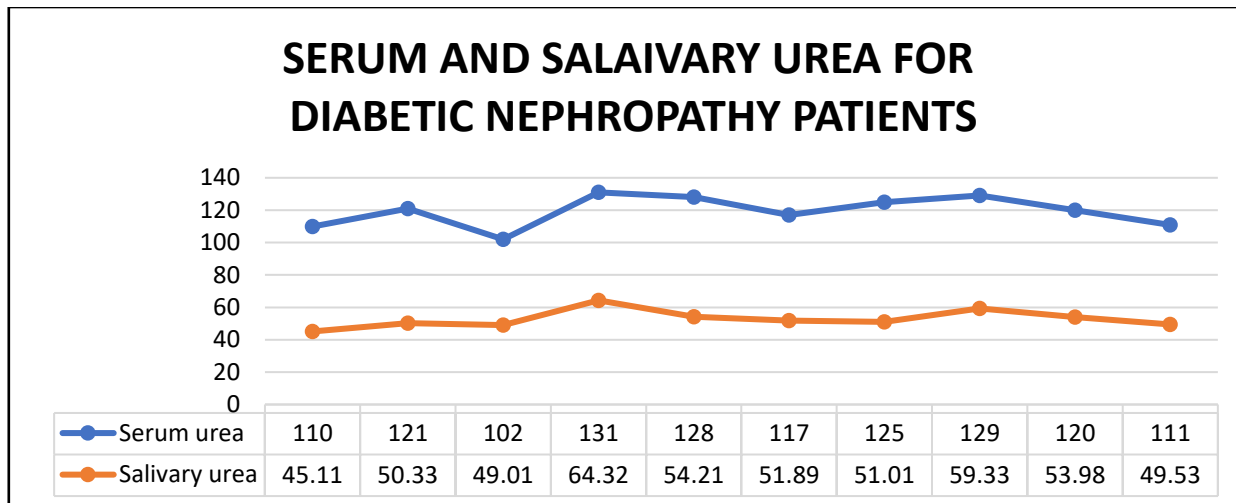
*Figure No 6: Comparison between Serum and Salivary Urea for Diabetic Patients*



*Figure No 7: Comparison between Serum and Salivary Glucose for Diabetic Nephropathy Patients*



*Figure No 8: Comparison between Serum and Salivary Creatinine for Diabetic Nephropathy Patients*



*Figure No 9: Comparison between Serum and Salivary Urea for Diabetic Nephropathy Patients*

# DETECTION OF ESTRUS IN BOVINE USING MACHINE LEARNING

Hemalatha.R.J

Department of Biomedical Engineering  
Vels Institute of Science, Technology  
and Advanced Studies  
Chennai, India  
[hodbiomedical@velsuniv.ac.in](mailto:hodbiomedical@velsuniv.ac.in)

SonaShree.S.P

Department of Biomedical Engineering  
Vels Institute of Science, Technology  
and Advanced Studies  
Chennai, India  
[17615123@velsuniv.ac.in](mailto:17615123@velsuniv.ac.in)

Thamizhvani T.R.

Department of Biomedical Engineering  
Vels Institute of Science, Technology  
and Advanced Studies  
Chennai, India  
[thamizhvani.se@velsuniv.ac.in](mailto:thamizhvani.se@velsuniv.ac.in)

Vijayabaskar.V

Department of Electronics and  
Telecommunication  
Sathyabama Institute of Science,  
Technology  
Chennai, India  
[v\\_vijayabaskar@yahoo.co.in](mailto:v_vijayabaskar@yahoo.co.in)

**Abstract**—Estrus refers to the reproductive stage of cattle during which ovulation occurs. The day of estrus plays an important role in the reproductive life of a cow. Inseminating cow at perfect time of ovulation was a big deal for farmers. Bovines that undergo Silent Estrus will not show any mounting behaviours or physical signs. So detecting estrus in cows undergoing silent estrus is also a challenging task. The main aim of this study is to propose a simple method to detect the estrus in cow by classifying cows into estrus and non-estrus groups, using machine learning algorithm based on correlating the milk parameters such as fat, pH, SNF, specific gravity, density and the age, amount of milking, frequency of milking and breed of the cow. For estrus cows, milk parameters like pH and fat show high significant change whereas SNF and density shows moderate significant change. The selected parameters are fed into machine learning algorithms like Decision Tree Classifiers. The classification based on the selected parameters specifies higher performance for Decision Tree with accuracy 98%. Thus, Decision Tree classifier is defined as an effective classifier and a simple method to detect estrus in cow by utilizing milk parameters.

**Keywords**—Estrus, Bovine, fat, pH, Specific Gravity, SNF, Density, Machine learning algorithm.

## I. INTRODUCTION

Estrus is said to be reproductive stage of cattle during which ovulation occurs which is also known as heat. The estrus plays an important role in the fertilization of the cow's egg which is possible only at this stage. The length of Bovine's estrus cycle is 21 days. The estrus cycle is divided into 4 stages. They are Proestrus (Day: 0-4), Estrus (Day: 4-5), metestrus (Day: 5-8 day) and diestrus (9-21 days). At the state of estrus, the cow ovulates and is ready for insemination. The fertile period of the ovulated egg is very short, only around 18 to 24 hours. The egg will be at its most fertile state immediately following ovulation. So it is important for the semen to be at that site for fertilization. So, it is very much necessary for an accurate heat detection.

If the heat detector does not show accurate result, then insemination will not occur at optimal time and failure of AI

can occur. It takes more than 3 months to confirm the pregnancy of the inseminated cow. This is why it is so important to be able to identify estrus with maximum potential. The estrus for a normal healthy cow can be identified visually. The signs of estrus are standing to be mounted, mounting other cows, increased nervousness, swollen vulva and clear mucus discharge [26]. Farmers may miss out these signs if they do not notice these signs consciously.

Effective heat detection is possible at farms where herd of cow is present. Since the number of herd mates can highly influence heat detection. While small number of cows can be very challenging. Even in herd some will show no interest in riding or being ridden and some cows will undergo silent estrus. Silent estrus is a state where estrus occurs without any physical signs [27]. Farmers use Pressure Sensitive mount detector, tail paint and markers on the tail to detect whether any mounting is occurred or not. But this will not work for cows that undergo silent estrus in small group and cows that do not show interest on riding. Some techniques used to overcome this problem that are available today include ultrasound monitoring of ovary, detection of sex pheromone in excreta [23], P4 estimation in milk [15], measurement of vaginal conductivity using probes [16] and endometrial biopsy [24].

Ayodeji Folorunsho Ajayi et al., (2020), [24] defines the staging of the estrus cycle and induction of estrus in experimental rodents. Methods like visual assessment of vaginal smear cytology, histological examination of reproductive organs which is invasive, vaginal wall impedance method where a probe is inserted into the vagina for 30 seconds and impedance is measured for which animals in estrus shows high impedance. Another method called urine biochemistry where the urine contains high concentration of fatty acids and high levels of protein and lipids in proestrus and estrus phase.

Mozūraitis et al., (2017), [23] analysed the fecal samples using solid phase micro extraction gas chromatography and

found the presence of acetic acid and propionic acid in the headspaces of fecal samples of estrus cows. With low level of acid until a day before ovulation and peaking around 0.5 before ovulation.

Miura et al., (2017), [25] introduced a new method of detecting estrus by monitoring ventral tail base surface temperature using a wearable wireless sensor. The sensor was attached to the ventral tail base at day 11 and temperature is measured for every 2 minutes throughout the day and highest value was analysed. A substantial change is seen around the time of behavioural estrus.

Nayan, V et al., (2020), [21] designed a dipstick using gold nanoparticles conjugated with a novel antipeptide antibodies against LH. This method is a lateral flow assay that detects the ovulation period by assaying a high level of LH in the urine of buffalo at the time of estrus.

Since these methods will be costlier, time consuming, less accurate and invasive, in this study a simple, non-invasive method for estrus detection using milk parameters utilizing machine learning algorithm is proposed based on correlating the milk parameters [2] such as fat [3], pH [1], SNF [3], specific gravity, density and the age, amount of milking, frequency of milking and breed of the cow. 15 estrus cows and 15 non-estrus cows were included in this study. Among the cows (n=50), 37 (74%) cows were HxF breed and 13 (26%) cows were Jersey breed. The mean age of cows were  $9.4 \pm 5.225$  (mean  $\pm$  SD) and majority of cows were in the age group of 3 to 5 years. The milk parameters like fat and SNF were measured using digital milk analyser. pH and temperature of the milk were measured using commercial pH meter (ACETEQ Instruments India Inc.). Specific Gravity of the milk is measured using lactometer. Data like age, amount of milking, frequency of milking and breed of cow were known from the owners of the respective cow. The estrus state of cow is confirmed by a professional veterinary doctor. Machine learning algorithms like decision tree classifier and Naive Bayes classifier are applied for classifying estrus from non-estrus cow based on selected parameters. With the classifier results, Decision tree classifier is defined to be an efficient classifier for describing the estrus stage of cow with higher accuracy.

## II. SAMPLES AND METHODS

### A. Study design and population

Milk samples collected to determine the relationship of the estrus of cow to the milk properties was assessed on Holstein Friesian (n=37) breed and Jersey breed (n=13) cows reared in Dharmapuri, India. The samples were collected (approx. 50 ml) in a clean and dry container. Estrus cows (n=25) and non-estrus cows (n=25) were assessed. Milking was done in the morning between 5.00 AM to 7.00 AM. Samples were collected for 2 days (25samples/day). All the cows belong to different fields. The state of estrus or non-estrus was analyzed using veterinarian guide, information from cow owner, cow records, heat detector and rectal palpation. Milk parameters were measured within 2-3 hours of milking.

### pH

The instrument used of measuring pH is a commercial pH meter that also indicates the temperature. The pH meter used was manufactured by ACETEQ Instruments India Inc. It is a pen type/ pocket model pH & temperature meter with accuracy  $\pm 0.1$  and resolution 0.01pH.

### Specific Gravity and Density

The instrument used to measure the specific gravity (SG) of the sample is commercial lactometer. The Specific Gravity was calculated from the lactometer reading. Density is calculated from specific gravity value.

### Fat and Solid Non Fat

Fat and Solid Non fat (SNF) was measured in a Milk Society. The equipment used is a Digital milk analyzer. Before analysing the milk, it was homogenized using an Ultrasonic Processor, model FCO TU50. The Digital Milk Analyser used is manufactured by EKOMILK ULTRA with DPS milk analyser. It shows the value of fat as well as SNF.

### Other parameters

Variables like age, amount of milking, frequency of milking and breed of cow were known from the information given by the owners. The means values of the parameters measured is given in table 1.

The study samples were divided into the following groups:

Group-I – Estrus cows (n=75)

Group-II – Non-Estrus Cows (n=75)

### B. Feature Selection

An analysis was done between the mean and Standard Deviation values of all the listed parameters. From that analysis 5 features were observed showing significant variation during the time of estrus. They are fat [6], SNF, SG, Density and pH. For estrus cows, milk parameters like pH and fat show high significance whereas SNF and density shows moderate significance. pH is negatively correlated to fat ( $r = -0.4402$ ), SNF ( $r = -0.2108$ ) and Specific Gravity ( $r = -0.1238$ ). pH value of estrus cow is significantly lesser than pH of non-estrus cow. SNF and fat value of estrus cow is significantly higher than the values of non-estrus cow. Temperature doesn't show any significant change towards estrus.

TABLE 1. Parameters and their mean and SD values

SI.NO	PARAMETER	MEAN $\pm$ SD
1.	Age (years)	$9.4 \pm 5.225$
2.	Fat (%)	$3.76 \pm 0.604$
3.	SNF (%)	$9.62 \pm 2.814$
4.	Specific Gravity	$1.027 \pm 0.001$
5.	pH	$6.49 \pm 0.234$
6.	Amount of milking (Litres/day)	$12.82 \pm 2.154$
7.	Frequency of milking (per day)	$2 \pm 0$
8.	Temperature ( $^{\circ}$ C)	$25.95 \pm 2.468$
9.	Density (kg/m <sup>3</sup> )	$1027.66 \pm 1.721$

### C. Classification

Classification of the estrus and non-estrus cow is performed using ensemble tree methods and classifiers.

Decision Tree Classifier is a simple and widely used classification technique. It applies a straight forward idea to solve the classification problem. Decision Tree Classifier poses a series of carefully crafted questions about the attributes of the test record. Each time it receives an answer, a follow-up question is asked until a conclusion about the class label of the record is reached. Naive Bayes classifier is based on the simplifying assumption that the attribute values are conditionally independent given target value. Among the various methods of supervised statistical pattern recognition, the Nearest Neighbour rule achieves consistently high performance, without a priori assumptions about the distributions from which the training examples are drawn. It involves a training set of both positive and negative cases. A new sample is classified by calculating the distance to the nearest training case; the sign of that point then determines the classification of the sample [29, 30]. The k-NN classifier extends this idea by taking the k nearest points and assigning the sign of the majority. Gradient boosting falls under the category of boosting methods, which iteratively learn from each of the weak learners to build a strong model. It can optimize classification, regression and ranking. Gradient boosting falls under the category of boosting methods, which iteratively learn from each of the weak learners to build a strong model. The features that are selected are subjected to the classification process for analysis of estrus state [31]. Classifier results are analyzed using the confusion matrix based on the performance parameters. Accuracy, specificity, sensitivity are determined with the true positive, true negative, false positive and false negative values of the confusion matrix. Decision tree classifier is defined to be more efficient for classification with accuracy of 98%.

### III. RESULTS

There is a statistically significant correlation between the milk parameters like fat, SNF, SG, pH and Density in estrus and non-estrus groups. Fat shows significant positive correlation for SNF ( $r = 0.924$ ). Fat is also positively correlated to SG ( $r = 0.219$ ). pH shows significant negative correlation towards fat ( $r = -0.44$ ), SNF ( $r = -0.21$ ) and SG ( $r = -0.123$ ). Different machine learning models were applied on this value to choose the effective learning model that detects estrus based on these variables. The level of fat in the milk increases during the estrus period. In the same way, SNF and SG also increase at the time of estrus. But the acidity of the milk increases during estrus and so the pH decreases and shows a negative correlation for estrus. The correlation between these parameters towards estrus is given in table 2.

TABLE II. CORRELATION BETWEEN MILK PARAMETERS AND ESTRUS

Sl.NO	Parameter	Correlation
1.	Fat	$r = 0.46102$
2.	SNF	$r = 0.20028$
3.	SG	$r = 0.03520$
4.	pH	$r = -0.80999$
5.	Density	$r = 0.03520$

The features are trained with decision tree classifier, Logistic Regression, Naïve Bayes classifier, Random forest classifier and Linear SVM. The models are trained, tested and validated using MATLAB software (R2019B). The confusion matrix and ROC curve were obtained. The confusion matrix of Decision tree classifier shown in the figure 1 and flowchart is given in figure 2.

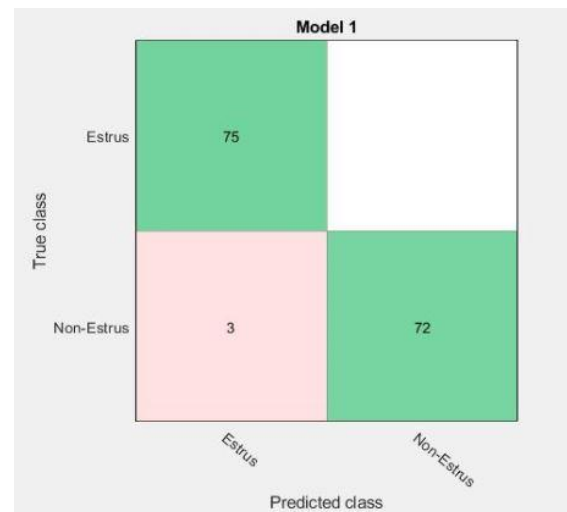


FIGURE 1: Confusion Matrix of Decision tree classifier

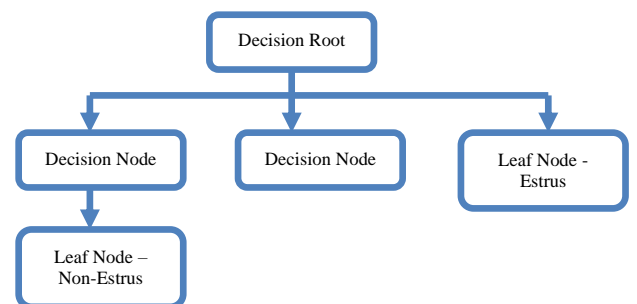


FIGURE 2: Decision Tree Classifier Flowchart

The accuracy, specificity, F1 score, Precision and Recall were calculated from the confusion matrix. The comparison of evaluation metrics for all the 5 models is shown in table 3.



TABLE III. COMPARISON OF EVALUATION PARAMETRIC AMONG MODELS

Si. no	Model	Accuracy	Precision	Recall	F1 Score	Specificity
1.	Decision Tree Classifier	0.98	1	0.96	0.98	1
2.	Logistic Regression	0.9	0.91	0.88	0.89	0.92
3.	Naïve Bayes Classifier	0.90	0.97	0.84	0.90	0.96
4.	Linear SVM	0.94	0.97	0.92	0.94	0.97
5.	Random Forest Classifier	0.5	1	0.5	0.66	0

The accuracy curve is plotted and show in figure 3.

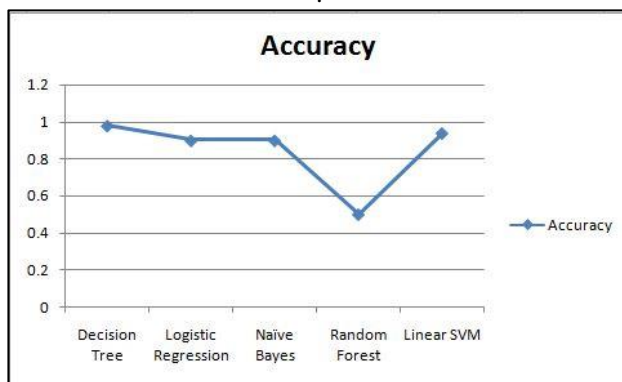


Figure 3: Accuracy of models

#### IV. DISCUSSION

In this study, we tried to correlate between the milk parameters like fat, SNF, SG, Density and pH and to detect the estrus by predicting using machine learning algorithms. Toledo-Alvarado et al.(2018) [2] described changes in milk characteristics during estrous cycle. Milk composition, particularly fat, protein, and lactose showed clear differences among the estrous cycle phases. Fat increased by 0.14% from diestrus high-progesterone to estrous phase, whereas protein concomitantly decreased by 0.03%. Lactose appeared to remain relatively constant over diestrus high-progesterone, rising 1 d before the day of estrus followed by a gradual reduction over the subsequent phases. VarijNayanab et al,(2020) [21] discovered a lateral flow method for detection of LH in urine. A qualitative ELISA for sensing LH was developed based on competitive binding of gold nanoparticles conjugated with epitope peptide and LH towards anti-peptide antibodies against LH. They also further explored the detection of LH in buffalo urine using the gold nanoparticles LHP conjugate (AuNP-LHP) in dipstick format. These experiments provided a proof-of-concept towards applicability of LH based sensor for ovulation prediction in the buffaloes. SerapGöncü et al,(2019) [16] described about the sensor technology that monitors and increase the reproduction system of cattle. Sensor technologies are newly used for cattle production management. These innovative applications are leading to a more efficient cow management in terms of both physiology

and sustainability. Efficiency of ultrasonography is 85-95%. Detection of estrus using sensor based fuzzy logic system gives 84.2% sensitivity. Use of nanotechnology for motion sensing in order to detect the restlessness of cow during heat gives accuracy of more than 82% [28].

In this study the proposed method have a simple method which is even non-invasive to detect estrus. The milk parameters that show significant difference in estrus and non-estrus period are regularly monitored and classified using different machine learning techniques in which decision tree classifier is defined to be accurate and time consuming. This can also pave way for successful artificial insemination. It is also a cost efficient one. In current days, there is already an analyser available which can measure the fat and SNF values. Incorporating a pH and specific gravity measurer in that analyser itself will be quiet more easier. If all these 4 milk parameters are monitored and trained daily, there is a possible easier way of detecting estrus in that analyser itself.

#### V. CONCLUSION

Detecting estrus and successful impregnating of a cow is the single most important factor on the dairy farm. The Conception Rate of cow should be proportionate to that of being inseminated. This ratio shows the effectiveness of estrus detection in the farm. Till now there are multiple ways to detect estrus like Visual observation, Computerized systems like Pedometer or Heat Watch which is a type of real time automated estrus detection system, Chin ball markers, Kamar Estrus Mount Detectors, Teaser animals, Progesterone tests, Videotaping [26]. These techniques will be useful while experiencing with herd of cows in dairy farms. The estrus detection in silent estrus and minimal cow dairy farms are quite difficult which may lead to decrease in livestock and farmers income. In order to introduce a effective, automated, no attention by owners, time consuming, non-invasive and general simple method, we are proposing the Decision Tree classifier as the efficient method for detecting estrus by utilizing milk parameters which gives 98% accuracy

#### REFERENCES

- [1] Paul Canson Berger, "the ph of bull semen and the vagina of cows as related to fertility"(1951). Oregon State College.
- [2] Toledo-Alvarado, H., Vazquez, A. I., de los Campos, G., Tempelman, R. J., Gabai, G., Cecchinato, A., & Bittante, G. (2018). Changes in milk characteristics and fatty acid profile during the estrous cycle in dairy cows. *Journal of Dairy Science*. doi:10.3168/jds.2018-14480
- [3] Lainé, A., Bastin, C., Grelet, C., Hammami, H., Colinet, F. G., Dale, L. M., ... Gengler, N. (2017). Assessing the effect of pregnancy stage on milk composition of dairy cows using mid-infrared spectra. *Journal of Dairy Science*, 100(4), 2863–2876. doi:10.3168/jds.2016-11736
- [4] Penasa, M., De Marchi, M., & Cassandro, M. (2016). Short communication: Effects of pregnancy on milk yield, composition traits, and coagulation properties of Holstein cows. *Journal of Dairy Science*, 99(6), 4864–4869. doi:10.3168/jds.2015-10168
- [5] Loker, S., Miglior, F., Bohmanova, J., Jamrozik, J., & Schaeffer, L. R. (2009). Phenotypic analysis of pregnancy effect on milk, fat, and protein yields of Canadian Ayrshire, Jersey, Brown Swiss, and Guernsey breeds. *Journal of Dairy Science*, 92(3), 1300–1312. doi:10.3168/jds.2008-1425
- [6] Rutten, M. J. M., Bovenhuis, H., Hettinga, K. A., van Valenberg, H. J. F., & van Arendonk, J. A. M. (2009). Predicting bovine milk fat composition using infrared spectroscopy based on milk samples

- collected in winter and summer. *Journal of Dairy Science*.doi:10.3168/jds.2009-2456
- [7] Soyeyurt, H., Bruwier, D., Romnee, J.-M., Gengler, N., Bertozzi, C., Veselko, D., & Dardenne, P. (2009). Potential estimation of major mineral contents in cow milk using mid-infrared spectrometry. *Journal of Dairy Science*, 92(6), 2444–2454. doi:10.3168/jds.2008-1734
- [8] Bohmanova, J., Jamrozik, J., & Miglior, F. (2009). Effect of pregnancy on production traits of Canadian Holstein cows. *Journal of Dairy Science*, 92(6), 2947–2959. doi:10.3168/jds.2008-1782
- [9] Abeni, F., Degano, L., Calza, F., Giangiaco, R., & Pirlo, G. (2005). Milk Quality and Automatic Milking: Fat Globule Size, Natural Creaming, and Lipolysis. *Journal of Dairy Science*, 88(10), 3519–3529. doi:10.3168/jds.s0022-0302(05)73037-x
- [10] Bachman, K. C., Hayen, M. J., Morse, D., & Wilcox, C. J. (1988). Effect of Pregnancy, Milk Yield, and Somatic Cell Count on Bovine Milk Fat Hydrolysis. *Journal of Dairy Science*, 71(4), 925–931. doi:10.3168/jds.s0022-0302(88)79638-1
- [11] Caja, G., Castro-Costa, A., & Knight, C. H. (2016). Engineering to support wellbeing of dairy animals. *Journal of Dairy Research*, 83(02), 136–147. doi:10.1017/s0022029916000261
- [12] Toledo-Alvarado, H., Vazquez, A. I., de los Campos, G., Tempelman, R. J., Bittante, G., & Cecchinato, A. (2018). Diagnosing pregnancy status using infrared spectra and milk composition in dairy cows. *Journal of Dairy Science*, 101(3), 2496–2504. doi:10.3168/jds.2017-13647.
- [13] Johnson, Scott Kay, "Compounds in Bovine Urine Which Change in Concentration During the Estrous Cycle: Possible Pheromones" (1987). All Graduate Theses and Dissertations. 4090. <https://digitalcommons.usu.edu/etd/4090>
- [14] Alexander Forbes Howie, Simon Christopher Riley, Desalew Tadesse Tegegne, "Serum, milk, saliva and urine progesterone and estradiol profiles in crossbred (Zebu x Holstein Friesian) dairy cattle"(2017). <https://www.researchgate.net/publication/320523546>
- [15] B. J. McLeod, J. A. Foulkes, M. E. Williams and R. F. Weller (1991). Predicting the time of ovulation in dairy cows using on-farm progesterone kits. *Animal Production*, 52, pp 1-9 doi:10.1017/S0003356100005638
- [16] Göncü S, Koluman N. The sensor technologies for more efficient cow reproduction systems. *MOJ Eco Environ Sci*. 2019;4(3):128–131. DOI: 10.15406/moes.2019.04.00144
- [17] Selvam RM, Singh D, Akbarsha MA, Archunan G. Urine levels of luteinizing hormone as predictor of the period of ovulation for advantage of timed-artificial insemination in murrha buffalo (*Bubalus bubalis*). *Reprod Dom Anim*. 2017;00:1–5. doi:10.1111/rda.12933.
- [18] SP Mohapatra, SD Ingole, SV Bharucha, AS Nagvekar, PM Kekan and SD Kharde. Changes in urinary dipstick parameters in cyclic and early pregnant Murrah buffaloes. *The Pharma Innovation Journal* 2017; 6(7): 08-10 .
- [19] M.G. Colazo et al., Plasma luteinizing hormone concentrations in cows given repeated treatments or three different doses of gonadotropin releasing hormone. *Theriogenology* 71 (2009) 984–992.
- [20] J.F. Smith , R.J. Fairclough , E. Payne & A.J. Peterson (1975) Plasma hormone levels in the cow, *New Zealand Journal of Agricultural Research*, 18:2, 123-129, DOI: 10.1080/00288233.1975.10421012
- [21] NAYAN, V., Sinha, E. S., Onteru, S. K., & Singh, D. (2020). A proof-of-concept of lateral flow-based luteinizing hormone detection in urine for ovulation prediction in the buffaloes. *Analytical Methods*. doi:10.1039/d0ay00787k
- [22] Forcato, D. O., Carmine, M. P., Echeverría, G. E., Pécora, R. P., & Kivatinitz, S. C. (2005). Milk Fat Content Measurement by a Simple UV Spectrophotometric Method: An Alternative Screening Method. *Journal of Dairy Science*, 88(2), 478–481. doi:10.3168/jds.s0022-0302(05)72709-0
- [23] Mozūraitis, R., Kutra, J., Borg-Karlsen, A.-K., & Būda, V. (2017). Dynamics of putative sex pheromone components during heat periods in estrus-induced cows. *Journal of Dairy Science*, 100(9), 7686–7695. doi:10.3168/jds.2016-12376
- [24] Ayodeji Folorunsho Ajayi 1, Roland Eghoghoso Akhigbe 1(2020), Staging of the estrous cycle and induction of estrus in experimental rodents: an update, doi: 10.1186/s40738-020-00074-3
- [25] Ryotaro Miura 1, Koji Yoshioka 2, Toru Miyamoto 3, Hirofumi Nogami 4, Hironao Okada 5, Toshihiro Itoh 6 , Estrous detection by monitoring ventral tail base surface temperature using a wearable wireless sensor in cattle , DOI: 10.1016/j.anireprosci.2017.03.002
- [26] <https://www.thebeefsite.com/articles/2362/estrus-detection-in-cattle>
- [27] <https://www.milkingcloud.com/how-to-catch-hidden-oestrus>
- [28] Rao TKS, Kumar N, Kumar P, Chaurasia S and Patel NB (2013) Heat detection techniques in cattle and buffalo, 6(6):363-369, doi:10.5455/vetworld.2013.363-369.
- [29] RJ Hemalatha, V Vijaybaskar, TR Thamizhvani (2018) Performance evaluation of contour based segmentation methods for ultrasound images, *Advances in Multimedia*, 2018 : 4976372.
- [30] TR Thamizhvani, KF Tanveer Ahmed, RJ Hemalatha, A Josephin Arockia Dhivya, R Chandrasekaran (2021) Enhancement of MRI images of hamstring avulsion injury using histogram based techniques, *Multimedia Tools and Applications*, 1-18.
- [31] TR Thamizhvani, Syed Uzma Farheen, RJ Hemalatha, A Josephin Arockia Dhivya (2020), Classification of progressive stages of alzheimer's disease in MRI hippocampal region, *Biomedical Engineering: Applications, Basis and Communications*, 32(6): 2050050.

**IEEE conference templates contain guidance text for composing and formatting conference papers. Please ensure that all template text is removed from your conference paper prior to submission to the conference. Failure to remove template text from your paper may result in your paper not being published.**

# Length of Stay Prediction in Acute Intensive Care Unit in Cardiothoracic Surgery Patients

Nafiseh Mollaei

*LIBPhys-UNL*

*Departamento de Física, FCT*

*Nova University of Lisbon*

Portugal

N.Mollaei@campus.fct.unl.pt

Ana Rita Londral

*Value for Health CoLAB*

*Comprehensive Health Research Center*

*Nova Medical School, Nova University of Lisbon, Portugal*

Ana.Londral@vohcolab.org

Catia Cepeda

*LIBPhys-UNL*

*Departamento de Física, FCT*

*Nova University of Lisbon, Portugal*

c.cepeda@campus.fct.unl.pt

Salomé Azevedo

*Value for Health CoLAB, Nova University of Lisbon*

*Dep. Engineering and Management, IST*

*University of Lisbon*

Lisbon, Portugal

Salome.azevedo@vohcolab.org

Jorge Pinheiro Santos

*Nova Medical School*

*Nova University of Lisbon*

Lisbon, Portugal

Jorge.p.santos@nms.unl.pt

Pedro Coelho

*Nova Medical School*

*Nova University of Lisbon*

Lisbon, Portugal

Pedro.coelho@nms.unl.pt

José Fragata

*Nova Medical School*

*Nova University of Lisbon*

Lisbon, Portugal

Jfragata@unl.pt

Hugo Gamboa

*LIBPhys-UNL*

*Departamento de Física, FCT*

*Nova University of Lisbon, Portugal*

Hgamboa@fct.unl.pt

**Abstract**—The goal of this study was to apply machine learning (ML) methods to predict the Length of Stay in an Intensive Care Unit (LOS-ICU) based on preoperative factors. To optimize the capacity of the ICU in surgery department, the prediction of a long stay (>2 days) can support the clinical decision making on accepting or delaying a patient intervention, considering the ICU occupancy. A database with records from 7364 patients that were operated in the Cardiothoracic surgery department of a public Portuguese hospital was used as the base of ML algorithms training. Regarding the risk of the patients to be in the group of long LOS-ICU, we compared five machine learning algorithms including Gradient Boosting, Random Forest, Support Vector Machine (SVM), Adaboost and Logistic Regression. We studied the classifier performance to adjust the sensitivity of a long stay classification, in order to reduce the potential of long LOS-ICU classification being miss classified as a short LOS-ICU.

**Index Terms**—Cardiac surgery, Machine learning, Preoperative risk factor, Classifier tuning

## I. INTRODUCTION

AI applications to Personalized Medicine are growing with increasing interest for medical practice. Personalized treatments and decisions have just started to be used in clinical practice. For example, EHR data can be used for the identification of disease phenotypes and its best treatments [1]. Also, ML can support discussion on treatment options,

time to cure or stability, patient preferences, side effects, treatment burden, or cost, based on patients' characteristics. The collection of outcomes data, including data reported by the patients, will allow to add to EHR data other potentially predictive variables for personalized medicine, that are not yet considered. The regular improvement in outcomes collection and use of healthcare data analysis, including ML algorithms, is going to progress for new treatment pathways. Also, it will progress for better and personalized use of resources in healthcare [2].

ML has the potential of promoting high-value care, by supporting the development of better quality and less cost in healthcare across unique populations or to find better procedures [3]–[6]. Aligned with the aim of the study in this paper, ML application to clinical practice will empower physicians to make high-value decisions [7]–[9].

Postoperative critical care of cardiac surgery patients is where the pressure on service capacity management and cost effort is more critical. As a major bottleneck in the surgery service capacity, more research is needed to predict in advance the patients' needs related to critical care after surgery, towards better intensive care unit (ICU) resource planning [10]. Longer length of stay in intensive care unit (LOS-ICU) is a known indicator of higher risk to mortality, morbidity and lower health-related quality of life [11]. Furthermore, as patients are expected to be admitted to ICU after cardiac surgery,

Supported by science and technology foundation (FCT), Ph.D. grant PD/BDE/142973/2018

the estimation of LOS-ICU is determinant when planning resources to optimize patient throughput. For these facts, the prediction of LOS-ICU based on preoperative factors would be of much relevance for cardiothoracic surgery departments. Other studies related to prolonged LOS-ICU for patients in cardiac surgery reported that increased preoperative creatinine levels, history of atrial fibrillation, and high EuroSCORE values are risk factors for LOS-ICU superior to 2 days [12]. Also, low preoperative hemoglobin, prolonged aortic clamping time, and low PaO<sub>2</sub>/FiO<sub>2</sub> ratio and blood glucose measured, were reported to be related to prolonged LOS in ICU [13].

Automatic prediction of the risk of a patient to be in the group of long LOS-ICU could support quicker and more efficient decisions regarding resources management and, possibly, early clinical support to potential risk factors. This study aimed to explore machine learning (ML) techniques, such as logistic regression, random forest, support vector machine (SVM), Ada boost, and Gradient boosting [14], [15], to predict LOS-ICU, from preoperative data registered on a database of patients that were submitted to cardiac surgery from 2008 to 2020, in a public Portuguese hospital. This predictive model would support the decision-making of clinical teams when ordering patients for surgery planning while optimizing hospital capacity and ensuring patients' safety.

In this paper, we present the results of an observational study to develop a ML-based prediction model to predict the LOS-ICU based on preoperative variables that characterize patients. We include a detailed description of all the important steps that were taken, regarding the pre-processing procedures used to analyze a clinical outcomes database and the classification algorithms. We discuss the validity of the results based on their potential impact on clinical practice and decision-making. The remainder of the paper is structured as follows. In Section 2, we explain the materials and methodology used for data preprocessing. Section 3 gives result of how we implemented training and testing, subsequently used for comparing the different classifiers models. Section 4 presents discussion about the performance for the different classifiers and how these can support the clinical decision-making. In section 5, a future work is presented.

## II. PROPOSED METHOD

### A. Data

Data was collected from 2011 to 2020, in a public Portuguese hospital in Lisbon, Portugal. The dataset [16] contains longitudinal information on outcomes of patients that were operated in the cardiothoracic surgery service. This dataset includes preoperative, operative, and postoperative data. The preoperative data of the population recorded on this database is summarized in Table 1.

### B. Input and output features for the prediction model

Preoperative risk factor data are the input variable. For missing values related to categorical values, we used the mean of numerical values to find the missing values and mode was the approach to substitute the categorical values missing data.

Input information was preoperative data that included diabetes treatment, hypertension, chronic lung disease, cerebrovascular disease, smoking history, hypercholesterolemia, renal disease, last preoperative creatinine, chronic lung disease, extracardiac arteriopathy, neurological dysfunction, preoperative heart rhythm, described in Table 1. The output data is LOS-ICU.

### C. Two groups of LOS: classification

LOS-ICU values are between 0 and 195 days by a mean of 2.84, in the studied dataset. The output of this study in total LOS is from 0 to 282 with a mean of 3.50. The importance of the ML algorithms, in this study, is centered on supporting the doctors on faster and better decisions when grouping the patients by the need of a short LOS-ICU [0,2] (days) or a long LOS-ICU [2,200] (days). In other words, short term, is the group of patients that stay less than 2 days in ICU, and long term, corresponds to patients that stay longer than 2 days. Besides, the data set is imbalanced between these two categories: 2269 belong to long term stay and 5094 cases related to short term stay. In most of the isolated features, the patients from the two groups are mixed (the features overlap) and the differentiation between groups is not an easy problem for the classification methods. The aim of classification is to identify patients in the two groups, short-term or long-term, and further predicting LOS-ICU, in order to optimize the capacity of the surgery department. Based on preoperative risk factors, five ML methods, adaboost [17], Logistic regression [18], random forest [19], support vector machine [20] and gradient boosting [21] have been used for classification.

For each of the classifiers [17]–[21], the experiments were performed by dividing the subjects into two data sets: training (80%), test (20%).

## III. RESULTS

### A. Classification

In this study, the benefit of work with multiple classification models over a single model supports the search for the best classifier [19]. Figure 1 shows the models of classifications. The two types of LOS-ICU categories are called predicted short-term and long-term. The classification answers to the question will the patient be a Long LOS-ICU case. Long LOS is considered the positive class. The classification can be adjusted by defining a threshold on selecting patients to be classified as Long LOS-ICU or Short LOS-ICU, increasing or decreasing false positives vs false negatives.

### B. Receiver operating characteristic curve: ROC curve

The ROC curve is a graph showing the performance of a classification model for all classification thresholds. This curve plots two parameters: true positive rate (TPR) against the false positive rate (FPR), as traced in Figure 2.

There is an efficient method that can provide the different classification thresholds, based on the Area under the Curve (AUC), that provides an overall merit of the classifier. AUC is the integral of the ROC curve and is 1 in case of perfect

TABLE I  
DESCRIPTION OF VARIABLES

	n	(%)
<b>Smoking history</b>		
Never smoked	4996	(67.9)
Ex-smoker (1 month or more)	1516	(20.6)
Current smoker	851	(11.6)
<b>Diabetes treatment</b>		
None	5212	(70.8)
Diet	130	(1.8)
Oral alone	1737	(23.69)
Insulin (with or without oral)	284	(3.9)
<b>Hypertension</b>		
No hypertension	1578	(21.4)
Hypertension	5781	(78.5)
Not known	4	(0.1)
<b>Hypercholesterolaemia</b>		
No hypercholesterolaemia	2585	(35.1)
Hypercholesterolaemia	4778	(64.9)
<b>Renal disease</b>		
None	7138	(96.9)
Functioning transplant	14	(0.21)
Creatinine >200 umol l-1	193	(2.61)
Dialysis - acute renal failure	5	(0.15)
Dialysis - chronic renal failure	13	(0.21)
<b>Chronic lung disease</b>		
No	6693	(90.9)
COPD / Emphysema	629	(8.5)
Asthma	41	(0.6)
<b>Extra cardiac arteriopathy</b>		
No	6979	(94.8)
Peripheral vascular disease	255	(3.5)
Cerebro-vascular disease	163	(2.2)
<b>Cerbovascular disease</b>		
None	6811	(92.5)
CVA	374	(5.13)
TIA	30	(0.43)
Non-invasive >50 % stenosis on Doppler	171	(2.31)
Previous carotid surgery	14	(0.21)
<b>Neurological dysfunction</b>		
No	7194	(97.77)
Yes	169	(2.31)
<b>Pre-operative heart rhythm</b>		
Sinus	5831	(79.25)
Ventricular tachycardia/ Ventricular fibrillation	4	(0.14)
Atrial fibrillation/ Atrial flutter	1221	(16.61)
Complete heart block	4	(0.1)
Other abnormal rhythm	15	(0.21)
Paroxistic atrial fibrillation	134	(1.81)
Rhythm pace	154	(2.11)

classifier or 0.5 in case of a random classifier. Predictions ranked in ascending order in AUC is classification-threshold-invariant. In fact, it measures the quality of the model's predictions irrespective of what classification threshold is chosen.

By tuning the classifier threshold value, the AUC was calculated by using a cutoff where the sum of sensitivity and specificity was maximal. The error rates of the models were calculated by using a cutoff where the sum of sensitivity and specificity was maximal. For the logistic regression, gradient boosting, random forest, support vector machine (svm) and adaboost model used, the accuracy was 0.71, 0.72, 0.76, 0.70 and 0.75, respectively, in defined optimal threshold locations.

The curve shows that the best choice for the classifier is the random forest classifier, on the rates of true and false positives

and the implicit trade-off between the two.

The classifier should then be tuned to operate by adjusting the classification threshold and adjusting the False Negative Rate. Figure 2 presents the classifier adjustment to clinical practice (high specificity to avoid missing long LOS-ICU). Hence, detecting the patients at higher risk of long-stay is the core importance of this ML method, focused on supporting the clinical decision makers when ordering the patients that will be operated. In the confusion matrix it is shown the error rates on False Negatives (True Long LOS classified as Short LOS, that should be reduced) and the False Positives ( True Short LOS classified as Long, that do not pose a risk for the patient). The confusion matrices are computed for point1, point2, and point3 presenting different levels of error

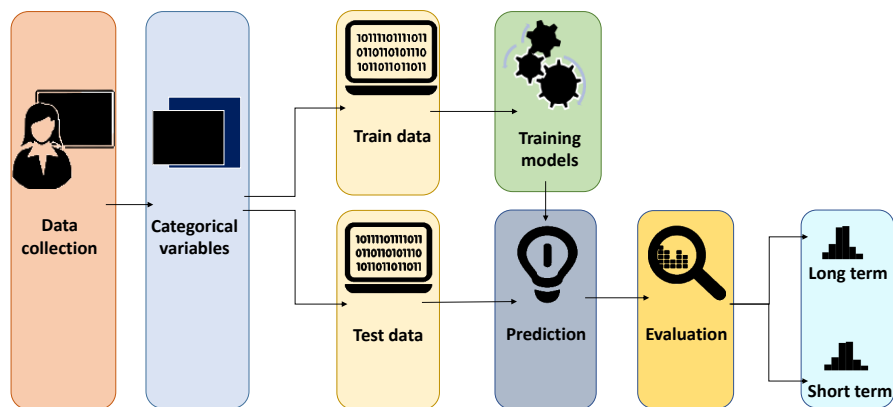


Fig. 1. Flowchart of classification models for LOS-ICU.

for the False Negatives, respectively 5, 35, and 64 in total of 423 true long stay patients. Table 2 shows a summary of statistical measures for the three points marked in the ROC curve. These measures are those relevant for the aim of using ML support to help clinicians in decision-making, preserving patients' safety while optimizing hospital capacity [22]. The points chosen in the ROC curve are positions where sensitivity to Long LOS (true-positive rate) is high ( $>0.860$ ). This is important as it guarantees that most of the patients that will need a long stay will be predicted as so. This tuning will prevent the error Type II, which is the mostly avoided in this problem: to miss classify a patient in the long LOS as a short LOS. This error would put patient safety at risk, as it could lead to high risk of exceeding the ICU capacity. In other words, failure to identify a significant proportion of long-stay patients could result in an unexpected demand for bed space, nursing requirements, and higher unanticipated costs. Another relevant statistical measure for this study is the false omission rate (FOR). This measure represents the ratio between false negatives and predicted negatives, i.e. those patients that are in the group of short LOS but are miss classified in the group of long LOS. Low values of FOR allow the clinical decision-making to be confident when predicting that a patient will have a short stay in ICU. This would allow rapid decisions related to patients that will not put at risk the ICU capacity.

A compromise between the sensitivity for Long LOS and the FOR represent the best outcome for this ML problem, at the expense of a low overall accuracy. For resource management, it may be preferable to accept a decrease in general accuracy of the model but tune it to correctly identify more long-stay patients and increase the confidence in the predicted short-stay patients.

#### IV. DISCUSSION

ICU is an expensive and scarce resource in a hospital. When faced with the decision of a new cardiac surgical intervention,

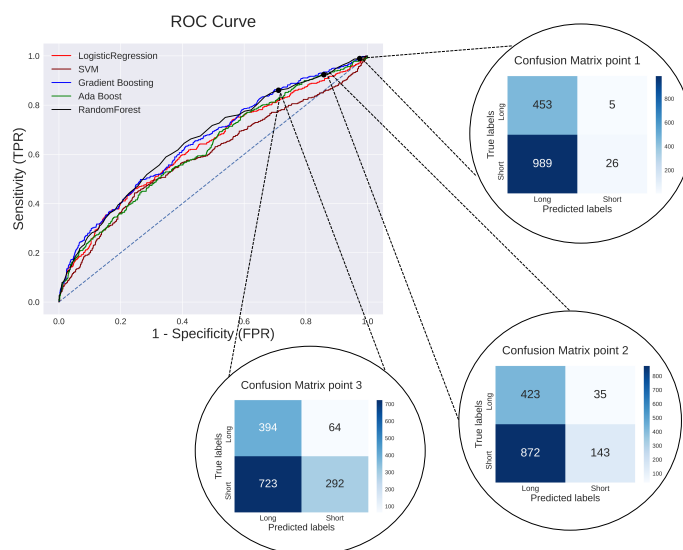


Fig. 2. ROC curve: Adjusting the model in three different points to clinical practice (high specificity to avoid missing long LOS) using five classifiers adaboost, gradient boosting, logistic regression, support vector machine (SVM) and random forest

TABLE II  
SPECIFICITY, FALSE-POSITIVE RATE (FPR), AND SENSITIVITY / TRUE-POSITIVE RATE (TPR) AND FALSE OMISSION RATE (FOR) FOR THREE POSITIONS IN THE ROC CURVE

	Specificity	FPR	Sensitivity	
			TPR	FOR
Point 1	0.025	0.974	0.891	0.161
Point 2	0.140	0.858	0.923	0.196
Point 3	0.287	0.711	0.860	0.179

the clinical decision can be supported by a risk predictor on the number of days that a specific patient will need to stay in the ICU, based on his current clinical status. An accurate prediction would allow a better planing of hospital resources,

while guaranteeing patients' safety.

In this paper, we present a ML method to classify patients in cardiac surgery according to their postoperative needs for ICU stay. By minimizing the risk of a patient being wrongly predicted as short stay, the current algorithm was tuned to force a high precision of risk evaluation of long LOS-ICU. While this model minimizes the most critical error of accepting a new patient that was considered a short stay and reveals as a long stay, it produces an increased number of wrong classifications of patients that need short LOS-ICU as long LOS-ICU. We demonstrate that it is possible to provide to the clinical team, the possibility to adjust the classifier to different sensitivity values, taking into account the ICU resources (e.g. the number of free beds). Like in [23] that provide a validation mechanisms, we refer to the extent to which clinicians can relate to Artificial Intelligence (AI), attach a clinical reasoning to its output and integrate its use in their daily workflows and routines. Our risk prediction tool can support faster clinical decision-making and better resource planning, by predicting the length of stay of each patient in the most critical acute postoperative resource.

#### V. FUTURE WORK

A future study can investigate more machine learning classifiers. In this case, the accuracy could increase in the means of precise prediction in the two labels of LOS-ICU as a short term and long term. To explore this further, regressors are methods that predict LOS-ICU by using categorizing the classifiers in terms of valuating mean absolute error (MAE) in LOS-ICU in long term.

A special interest for resource planning exist to predict the number of days of long LOS-ICU. Such a prediction would allow to differentiate different groups of patients in this class, and improve surgical planning.

This study will also develop an software for risk estimation which could be used for real-time estimation of patients' LOS-ICU after surgery, to be calculated in the preoperative period.

#### ACKNOWLEDGMENT

This work has been partially supported by the project PD/BDE/142973/2018 financed by the Program of FCT. The authors have no conflicts of interest to report

#### REFERENCES

- [1] C. W. Seymour, J. N. Kennedy, S. Wang, C.-C. H. Chang, C. F. Elliott, Z. Xu, S. Berry, G. Clermont, G. Cooper, H. Gomez *et al.*, "Derivation, validation, and potential treatment implications of novel clinical phenotypes for sepsis," *Jama*, vol. 321, no. 20, pp. 2003–2017, 2019.
- [2] D. W. Bates, S. Saria, L. Ohno-Machado, A. Shah, and G. Escobar, "Big data in health care: using analytics to identify and manage high-risk and high-cost patients," *Health Affairs*, vol. 33, no. 7, pp. 1123–1131, 2014.
- [3] I. Ganguli and T. G. Ferris, "Accountable care at the frontlines of a health system: bridging aspiration and reality," *Jama*, vol. 319, no. 7, pp. 655–656, 2018.
- [4] A. Chapman, H. Yang, S. A. Thomas, K. Searle, and C. Browning, "Barriers and enablers to the delivery of psychological care in the management of patients with type 2 diabetes mellitus in china: a qualitative study using the theoretical domains framework," *BMC health services research*, vol. 16, no. 1, pp. 1–10, 2016.
- [5] K. E. Joynt, J. F. Figueroa, N. Beaulieu, R. C. Wild, E. J. Orav, and A. K. Jha, "Segmenting high-cost medicare patients into potentially actionable cohorts," in *Healthcare*, vol. 5, no. 1-2. Elsevier, 2017, pp. 62–67.
- [6] J. Colbert and I. Ganguli, "To identify patients for care management interventions, look beyond big data," *Health Affairs Blog*, April, vol. 19, 2016.
- [7] A. Rajkomar, J. W. L. Yim, K. Grumbach, and A. Parekh, "Weighting primary care patient panel size: a novel electronic health record-derived measure using machine learning," *JMIR medical informatics*, vol. 4, no. 4, p. e29, 2016.
- [8] A. L. Schwartz, A. B. Jena, A. M. Zaslavsky, and J. M. McWilliams, "Analysis of physician variation in provision of low-value services," *JAMA internal medicine*, vol. 179, no. 1, pp. 16–25, 2019.
- [9] L. I. Horwitz, M. Kuznetsova, S. A. Jones *et al.*, "Creating a learning health system through rapid-cycle, randomized testing," *N Engl J Med*, vol. 381, no. 12, pp. 1175–1179, 2019.
- [10] A. Almashrafi, M. Elmontsri, and P. Aylin, "Systematic review of factors influencing length of stay in icu after adult cardiac surgery," *BMC health services research*, vol. 16, no. 1, p. 318, 2016.
- [11] M. R. Williams, R. B. Wellner, E. A. Hartnett, B. Thornton, M. N. Kavarana, R. Mahapatra, M. C. Oz, and R. Sladen, "Long-term survival and quality of life in cardiac surgical patients with prolonged intensive care unit length of stay," *The Annals of thoracic surgery*, vol. 73, no. 5, pp. 1472–1478, 2002.
- [12] R. Eltheni, K. Giakoumidakis, H. Brokalaki, P. Galanis, I. Nenekidis, and G. Fildissis, "Predictors of prolonged stay in the intensive care unit following cardiac surgery," *ISRN nursing*, vol. 2012, 2012.
- [13] T. Kapadodhos, E. Angelopoulos, I. Vasileiadis, S. Nanas, A. Kotanidou, A. Karabinis, K. Marathias, and C. Routsis, "Determinants of prolonged intensive care unit stay in patients after cardiac surgery: a prospective observational study," *Journal of thoracic disease*, vol. 9, no. 1, p. 70, 2017.
- [14] V. Kotu and B. Deshpande, *Predictive analytics and data mining: concepts and practice with rapidminer*. Morgan Kaufmann, 2014.
- [15] R. K. Sevakula, W.-T. M. Au-Yeung, J. P. Singh, E. K. Heist, E. M. Isselbacher, and A. A. Armoundas, "State-of-the-art machine learning techniques aiming to improve patient outcomes pertaining to the cardiovascular system," *Journal of the American Heart Association*, vol. 9, no. 4, p. e013924, 2020.
- [16] P. N. Coelho, L. M. Miranda, P. M. Barros, and J. I. Fragata, "Quality of life after elective cardiac surgery in elderly patients," *Interactive cardiovascular and thoracic surgery*, vol. 28, no. 2, pp. 199–205, 2019.
- [17] J. Sun, H. Li, H. Fujita, B. Fu, and W. Ai, "Class-imbalanced dynamic financial distress prediction based on adaboost-svm ensemble combined with smote and time weighting," *Information Fusion*, vol. 54, pp. 128–144, 2020.
- [18] M. C. Dias, C. Cepeda, D. Rindlisbacher, E. Battegay, M. Cheetham, and H. Gamboa, "Predicting response uncertainty in online surveys: A proof of concept," in *BIOSIGNALS*, 2019, pp. 155–162.
- [19] L. Breiman, "Random forests," *Machine learning*, vol. 45, no. 1, pp. 5–32, 2001.
- [20] V. Vapnik, *Estimation of dependences based on empirical data*. Springer Science & Business Media, 2006.
- [21] J. Cai, K. Xu, Y. Zhu, F. Hu, and L. Li, "Prediction and analysis of net ecosystem carbon exchange based on gradient boosting regression and random forest," *Applied Energy*, vol. 262, p. 114566, 2020.
- [22] I. Ganguli, W. J. Gordon, C. Lupo, M. Sands-Lincoln, J. George, G. Jackson, K. Rhee, and D. W. Bates, "Machine learning and the pursuit of high-value health care," *NEJM Catalyst Innovations in Care Delivery*, vol. 1, no. 6, 2020.
- [23] F. Cabitza and J.-D. Zeitoun, "The proof of the pudding: in praise of a culture of real-world validation for medical artificial intelligence," *Annals of translational medicine*, vol. 7, no. 8, 2019.

# Breast-Cancer Detection using Thermal Images with Marine-Predators-Algorithm Selected Features

Venkatesan Rajinikanth  
*Department of Electronics and  
 Instrumentation Engineering  
 St. Joseph's College of Engineering  
 Chennai 600119, India  
 email: v.rajinikanth@ieee.org*

Robertas Damaševičius  
*Faculty of Applied Mathematics  
 Silesian University of Technology  
 44-100 Gliwice, Poland  
 email: robertas.damasevicius@polsl.pl*

Seifedine Kadry  
*Faculty of Applied Computing and  
 Technology, Noroff University College,  
 Kristiansand, Norway  
 email: skadry@gmail.com*

Hafiz Tayyab Rauf  
*Centre for Smart Systems, AI and  
 Cybersecurity, Staffordshire University,  
 Stoke-on-Trent, United Kingdom  
 email: hafiztayyabrauf093@gmail.com*

David Taniar  
*Faculty of Information Technology  
 Monash University  
 Clayton, Victoria 3800, Australia  
 email: David.Taniar@monash.edu*

**Abstract**— Breast-cancer (BC) is one of the major diseases in women group and the early diagnosis and treatment is necessary to cure the disease. Early detection of BC is very essential to implement appropriate treatment and the proposed research aims to develop an automated BC detection system using Breast-Thermal-Images (BTI). The executed approach is as follows; (i) Recording the image for various breast orientation, (ii) Extracting the healthy/DCIS image patches, (iii) Treating the patches with image processing scheme, (iv) Feature extraction, (v) Feature optimization with Marine-Predators-Algorithm (MPA), and (vi) Two-class classification and validation. In this work, the essential image features, such as GLCM and LBP with varied weights are considered to classify the clinically collected BTI into healthy/DCIS class using a chosen two-class classifier. The result of this study confirms that the Decision-Tree (DT) classifier helps to achieve enhanced accuracy (>92%) compared to other methods adopted in this research.

**Keywords**—Breast cancer, Thermal image, Marine-Predators-Algorithm, GLCM, LBP, Classification.

## I. INTRODUCTION

In the current era, even though a number of modern detection and precautionary methods are available for the use, the diseases in humans are gradually rising because of their personal and environmental reasons [1,2]. Among the existing diseases, the cancer is one of the major illnesses leads to very painful death. The former research on cancer authenticates that; pre-mature detection will help to cure the disease with prescribed clinical treatment methods, such as chemotherapy, radiotherapy and mild/major surgery [3,4].

In women community, the cancer in breast region (Breast-Cancer) is rising due to a variety of reasons and the untreated Breast-Cancer (BC) will lead to various health problems []. The premature phase identification of BC will be completed by means of suggested medical practice. The final phase of BC can be predictable with the following indications; swelling in breast, variation in the silhouette/dimension of breast, transform of breast skin shade, uncontrollable pain, etc. Based on the location of the BC, it can be categorized into (i) Ductal-Carcinoma-in-Situ (DCIS) and (ii) Lobular-Carcinoma-in-Situ [5].

When the symptom of the BC is feat; then the patient should approach the doctor for through examination and confirmation of the disease with a variety of clinical procedures. The clinical level diagnosis of the BC involves in; (i) Visual examination by the disease expert, (ii)

Examination of the suspicious section using the bio-imaging technique, and (iii) Extracting the sufficient breast tissues using the Core-Needle-Biopsy (CNB) in order to confirm the cancer and its stage using recommended clinical tests. The execution of the CNB is one of the hurting invasive practice in which a needle is inserted to collect the tissue for further evaluation. In most of the cases, the CNB will not be recommended and the BC will be diagnosed using the modern imaging techniques [6-8].

The earlier works on bio-imaging based breast abnormality detection considered various imaging modalities, such as X-ray (mammogram), ultrasound, thermal-imaging, Magnetic-Resonance-Angiogram (MRA), and MRI [4,7,8]. Recently, the thermal-imaging approaches are widely adopted in various clinics to detect the disease in organs using the infer-red (IR) radiation coming out of the skin section and due to its non-invasive nature and low cost, these approaches are widely adopted in BC diagnosis [9-12].

The thermal imaging of the barest section is recorded using a specialised thermal imaging device (camera) which converts the captured IR radiation in to RGB scaled image of chosen dimension. In this image, every colour demotes a thermal profile and by simply evaluating the thermal profile, the abnormality in the image can be detected. Thermal-Imaging (TI) supported BC diagnosis can be found in [13] and these works confirms that; this technique helps to detect the disease with better diagnostic accuracy.

In this work, the clinical grade BTI images collected from the DITI-India [14] is considered for the evaluation and the earlier works on this database can be found in [4,8,13]. The recording of the image is carried out in different orientations and after the recording, the suspicions sections having abnormal/normal thermal patterns are cropped and resized to 256x256x3 pixels. The cropped images are then evaluated using; (i) Saliency detection, morphological segmentation and GLCM feature extraction, and (ii) Local-Binary-Pattern (LBP) generation with varied weights ( $W=1$  to 4) and LBP feature extraction. All these extracted features are serially integrated to form a one-dimensional (1D) feature vector. Then the dominant features from this feature vector is selected using Marine-Predators-Algorithm (MPA) and these values are then considered to train and validate the binary classifier. During the classification task, a 10-fold cross validation is implemented and the best value attained is then considered as the classifier output. The task of this section is to categorize the TI into healthy/DCIS group. In this work, the performance evaluation of different classifiers



are executed and based on the classification accuracy, the merit of the chosen binary classifier is confirmed.

Remaining part of this research are prearranged as below; Section 2 demonstrates the methodology, Section 3 presents the experimental outcome, Section 4 and 5 demonstrates the discussion and conclusion of this work.

## II. METHODOLOGY

The detection of the BC using TI is one of the proven approach, in which a chosen image processing schemes are employed to extract and evaluate the abnormal section in the breast. The implementation of various segmentation techniques to extract the abnormal section from the breast TI can be found in [13]. Further, the image executed using the DITI-India dataset can be found in [14]. All the above said techniques are implemented the segmentation techniques.

In the proposed work, a methodology is suggested to extract and evaluate the abnormal breast segment from TI using the MLS. The various phases available in this system can be found in Figure 1. The patient with the breast cancer is completely tested by the doctor and suggests the image supported examination. The BT for the patient are recorded for the patients in a controlled environment using a special camera, which converts the IR radiation into RGB scale images with varied thermal patterns (colours ranging from blue to white). After collecting the TI; cropping and resizing is performed to get the recommended dimension image patches. The image patches are then evaluated using two different feature extraction pipelines, such as (i) saliency enhancement, morphological segmentation and GLCM feature mining, and (ii) LBP enhancement and feature extraction. After collecting these features, serial feature integration is then implemented. Feature optimization is then implemented using the MPA in order to choose the dominant features and the optimized features are then considered to train and validate the binary classifier using a 10-fold cross validation.

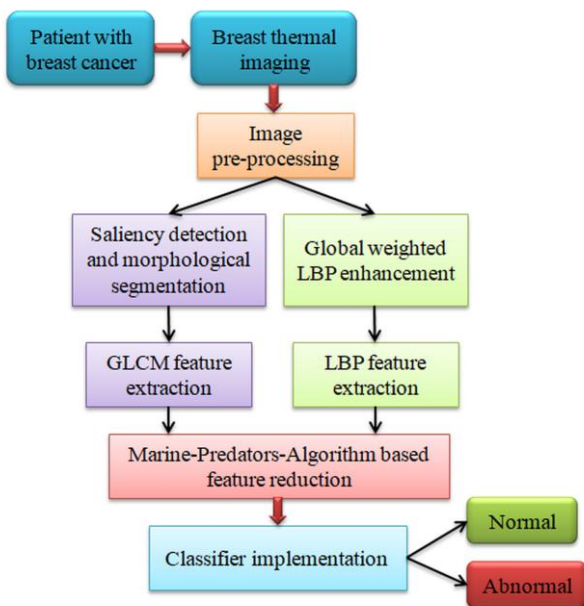


Fig.1. Structure of the proposed thermal image examination scheme

### A. Therma Image Database

Recently, TI assisted disease detection method is widely adopted due to its non-invasive nature. In this work, the real-time images of TI discussed in [4,8,13] is adopted for the examination. This dataset consist both healthy/DCIS class images collected from the volunteers using various angle positions, such as  $\theta = 0^\circ$ ,  $\theta = \pm 45^\circ$  and  $\theta = \pm 90^\circ$  as depicted in Figure 2. In this work, only the thermal patches from the normal and abnormal image section is then extracted and resized to  $256 \times 256 \times 3$  for the assessment. The number of image slices considered in this work for both the class is depicted in Table I. The converted images are then considered for the feature extraction and classifier validation task.

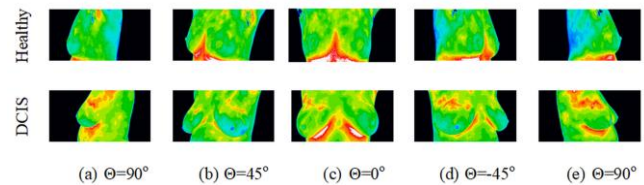


Fig.2. Sample test images collected from volunteers

TABLE I. THERMAL IMAGE PATCHES CONSIDERED

Database	Dimension	Images		
		Total	Training	Validation
Normal (clinical)	256x256x3	300	200	100
Abnormal (clinical)	256x256x3	300	200	100

### B. Feature Extraction

Automated disease detection is important to lessen the diagnostic burden during the mass screening task. In most of the cases, features based disease categorization is employed due to its proven potential and every MLS and deep-learning methods works based on the features. In most of the MLS; the handcrafted-features such as GLCM and LBP are widely adopted to categorize the test pictures.

#### 1. GLCM features mining

GLCM features helps to get the shape and texture information of the images under study. The texture information can be easily attained using the raw/processed images (gray level) and the shape features are collected from the binary image. In this work, the GLCM features are extracted from the pre-processed image; which is achieved using the saliency enhancement and morphological segmentation.

The essential information regarding the saliency supported image enhancement can be found in [15] and the morphology based segmentation scheme is presented in [16].

The GLCM features extracted in the proposed for healthy/DCIS picture is depicted in Eqn. (1),

$$F_{GLCM}(1 \times 25) = GLCM_{(1,1)}, \dots, GLCM_{(1,25)} \quad (1)$$

The GLCM features extracted using this work can be found in the earlier works [17-19].

#### 2. LBP features mining

The LBP is also a handcrafted-feature widely adopted in MLS and in this work, the global weighted LBP invented by

Gudigar et al. [20] is adopted to enhance the image with varied weights ranging from  $W=1$  to 4 and every technique helped to get a one-dimensional features of value  $1 \times 59$  and the total features extracted with the proposed LBP will be  $4 \times 59 = 1 \times 236$ . Eqn. (2) depicts the LBP features of this research;

$$F_{LBP}(1 \times 236) = F_{LBP1}(1 \times 59) + F_{LBP2}(1 \times 59) + F_{LBP3}(1 \times 59) + F_{LBP4}(1 \times 59) \quad (2)$$

The essential work on LBP and its features can be found in [1].

C. Feature Selection

The dominant feature selection is one of the essential tasks in MLS to reduce the problem of over fitting. The traditional feature reduction procedures existing in the literature can be accessed from [18]. In this work, heuristic algorithm based feature selection is employed.

The raw features existing for the assessment is depicted in Eqn. (3) and the feature reduction process will help to get a reduced feature vector value;

$$F_{Total(1,261)} = F_{GLCM(1,25)} + F_{LBP(1,236)} \quad (3)$$

The MPA is a nature-inspired meta-heuristic method invented by Faramarzi et al [21] and this approach combines the Lévy/Brownian-search to find the optimal solutions for a chosen task. As depicted in Figure 3, the MPA consist three search phases in which it uses a Lévy-search when the available prey is less and utilizes the Brownian-search when prey concentration is more. It continuously adopts both the search tactics till it discovers optimal solution. The sample 2D and 3D search strategy in MPA is depicted in Figure 4 and this strategy is continuously adopted in all the phases of MPA till optimal solution is achieved.

Initially, all the agents are arbitrarily dispersed in the investigate space based on Eqn. (4);

$$X_0 = X_L + \mathfrak{R}(X_U - X_L) \quad (4)$$

where  $\mathfrak{R}$  is random value [0,1],  $X_U$  and  $X_L$  are upper and lower bounds.

The MPA uses the survival-of-fittest (SOF) theory to choose the top predator, which survives over a long duration. In the MPA, every search iteration ( $Iter_{max}$ ) is divided into

three phases, such as  $PhaseI = \text{upto } \frac{1}{3} Iter_{max}$ ,  $PhaseII =$

$$\frac{1}{3} Iter_{max} < Iter < \frac{2}{3} Iter_{max} \text{ and } PhaseIII = \text{greater than } \frac{2}{3} Iter_{max}$$

PhaseI: It is the initial section called the high-velocity-phase, in which the predator is assumed faster than the prey.

PhaseII: It is the mid section called the unit-velocity-phase, in which the predator and prey are moving at same velocity.

PhaseIII: It is the final section called the low-velocity-phase, in which prey moves faster than the predator.

When the MPA search is initiated, it will explores the entire search space till the objective-value of the chosen problem is maximized. The complete information of the MPA can be found in [22,23].

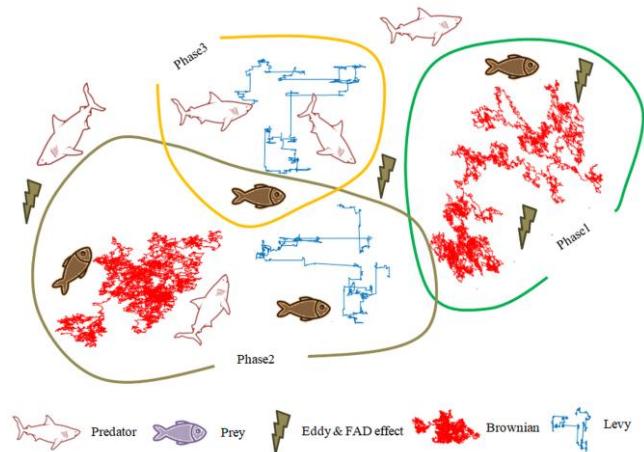


Fig.3. Various phases in the MPA optimization search

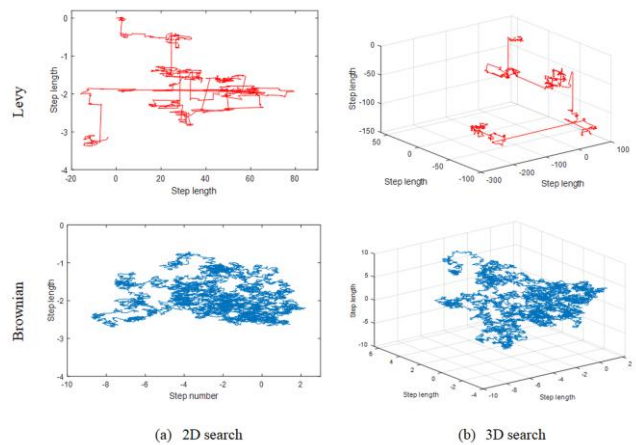


Fig.4. Sample 2D and 3D search pattern

In this work, the following values are considered for the MPA; number of agents = 25, search dimension = features to be selected, objective-value = maximal Hamming-distance, maximum iterations ( $Iter_{max}$ ) = 3000 and stopping criteria =  $Iter_{max}$ .

D. Implementation

The implementation of the feature optimization is depicted in Figure 5 and in this task, every feature (healthy/DCIS case) is individually compared against each other and the feature offers the maximal Hamming-distance is selected as the best feature. Similar procedure is employed for all the 261 features of GLCM and LBP and the features which have the clear difference between the healthy/DCIS cases are selected as the optimized features. In this work, the MPA helps to identify (1x74) optimized values and these features are then considered to evaluate the classifier performance.

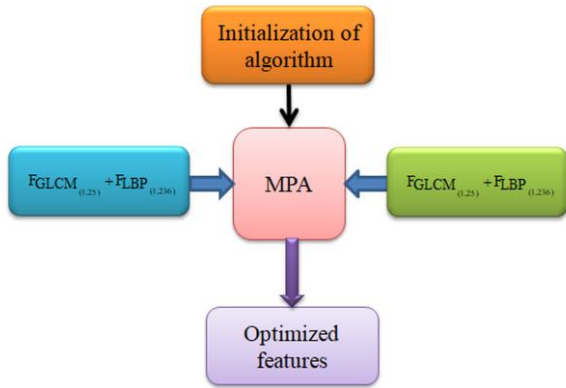


Fig.5. Feature optimization with MPA

### E. Classifier Implementation and Validation

A binary classification with a 10-fold cross validation is then implemented in this work to validate the performance of the BC detection scheme. The binary classification is implemented using SVM variants, such as linear (SVM-L), quadratic (SVM-Q), cubic (SVM-C), Fine-Gaussian (SVM-FG), Medium-Gaussian (SVM-MG), Coarse-Gaussian (SVM-CG) are considered and the results are computed [24].

The performance of the proposed BC detection approach is verified by computing the performance-measures, like True-Positive (TP), True-Negative (TN), False-Positive (FP) and False-Negative (FN), From these PV, additional measures, such as accuracy (AC), precision (PR) sensitivity (SE), specificity (SP) and F1-Score (FS) [17-19].

## III. EXPERIMENTAL RESULT

This part of the research presents the experimental outcome attained using the MATLAB software. Initially, the essential sections from the TI are extracted and resized into 256x256x3 pixels and the trial picture of healthy/DCIS case is depicted in Figure 6. These images are then considered for the evaluation. Fig 6(a) depicts the thermal profiles of healthy class and Fig 6(b) shows the DCIS category.

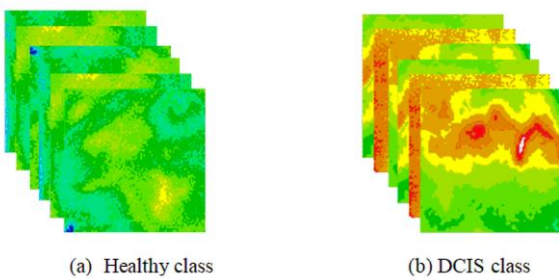


Fig.6. Sample trial images of healthy/DCIS class

Initially, the saliency enhancement and morphological extraction based technique is implemented to mine the abnormal section from the chosen breast TI section and then the GLCM features are extracted. Figure 7 shows the outcome attained with the implemented technique. Fig 7(a) depicts chosen TI segment, Fig 7(b) and (c) shows saliency and feature maps of the enhanced image, Fig 7(d) depicts the saliency feature plot and Fig 7(e) illustrates the extracted DCIS section. The GLCM features are then extracted and then integrated with the LBP features to form the essential handcrafted-features vector.

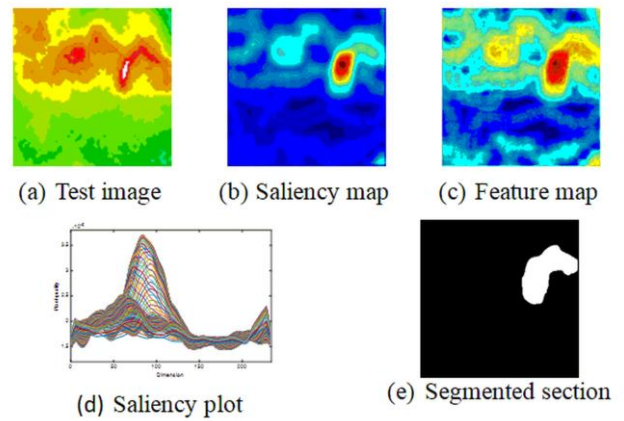


Fig.7. Saliency and morphological segmentation based evaluation of thermal pictures

The image slices are then considered for the LBP supported enhancement with varied weights ( $W=1$  to 4) as in Figure 8 and every image helps to extract 59 LBP features and the total LBP features extracted from this image group is  $4 \times 59 = 236$ . The corresponding LBP histogram is depicted in Figure 9 for healthy and DCIS class.

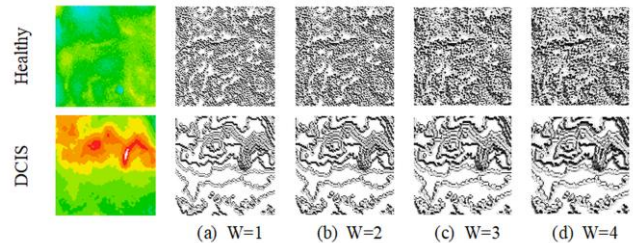
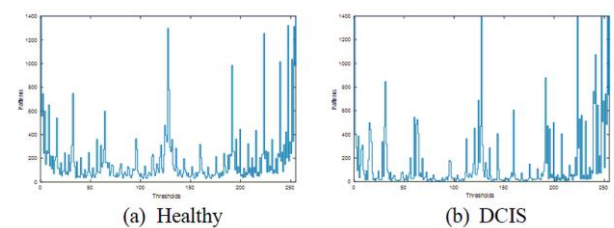
Fig.8. LBP enhanced thermal images for  $W=1$  to 4

Fig.9. LBP histogram of healthy and DCIS class

## IV. DISCUSSION

The integration of the GLCM and LBP features helps to get a feature vector with a dimension  $1 \times 261$  and these features are then reduce to a lower value ( $1 \times 74$ ) using the MPA. These features are then considered to train and validate the binary classifiers considered in this research.

In this work, 200 TI slices are considered to train the SVM classifiers and 100 TI are considered to validate the merit of proposed scheme. The experimental results attained in this work are depicted in Table II. The results of this table confirm that the accuracy of SVM-C and SVM-CG is similar and superior to other methods. The overall performance presented in the Glyph-plot (Fig. 10) confirms that the overall performance of CSVM-C and SVM-CG are approximately similar. If the proposed work is implemented with any one of the above mentioned classifiers, the result will be superior.

TABLE II. PERFORMANCE VALUES ATTAINED WITH TWO-CLASS CLASSIFIERS

Method	TP	FN	TN	FP	AC (%)	PR (%)	SE (%)	SP (%)	FS (%)
SVM-L	92	8	90	10	91	90.19	92	90	91.09
SVM-Q	92	8	94	6	93	93.88	92	94	92.93
SVM-C	93	7	94	6	93.50	93.94	93	94	93.47
SVM-FG	91	9	92	8	91.50	91.92	91	92	91.46
SVM-MG	92	8	93	7	92.50	92.93	92	93	92.46
SVM-CG	94	6	93	7	93.50	93.07	94	93	93.53

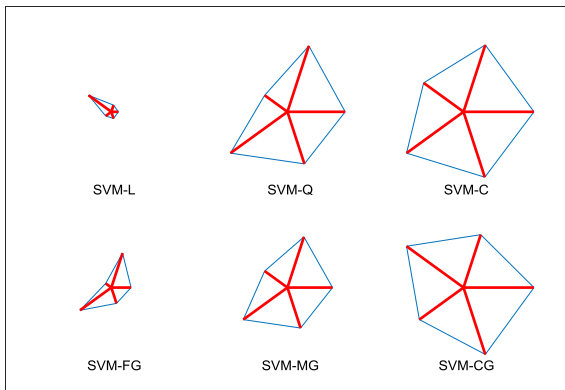


Fig.10. Glyph-plot of the computed performance values

In this research, the clinical grade TI is considered for the assessment and the attained result confirms that this work is clinically noteworthy. In future, the MPA can be replaced with the recent approach called the Red-Fox-Optimization [25] method existing in the literature.

## V. CONCLUSION

Assessment of breast TI is necessary to identify the BC in women community and in this research; a methodology is suggested to examine the TI using chosen handcrafted-features. The proposed work implements saliency based enhancement and morphology segmentation to extract the GLCM features. Later, the LBP features are extracted from the pictures enhanced with  $W=1$  to 4 and these features are then combined to form the hand-crafted feature vector. The dominant features are then selected using the MPA and these features are then considered to test the performance of the considered SVM classifiers. In this work, the result attained with SVM-C and SVM-CG are approximately similar. The proposed scheme works well on the clinical grade breast TI and provides better disease detection accuracy.

## REFERENCES

- [1] <https://www.who.int/cancer/detection/breastcancer/en/>
- [2] N. Dey, V. Rajinikanth, and A.E. Hassanien, "An Examination System to Classify the Breast Thermal Images into Early/Acute DCIS Class," in *Proceedings of International Conference on Data Science and Applications*, pp. 209-220. Springer, Singapore, 2020.
- [3] R. Elanthirayan, K. S. Kubra, V. Rajinikanth, N.S.M. Raja, and S.C. Satapathy, "Extraction of Cancer Section from 2D Breast MRI Slice Using Brain Stom Optimization," in *Intelligent Data Engineering and Analytics*, pp. 731-739. Springer, Singapore, 2021. [https://doi.org/10.1007/978-981-15-5679-1\\_71](https://doi.org/10.1007/978-981-15-5679-1_71).
- [4] N.S.M. Raja, V. Rajinikanth, S.L. Fernandes and S.C. Satapathy, "Segmentation of breast thermal images using Kapur's entropy and hidden Markov random field," *J. Med. Imaging Health Info.*, vol.7, no.8, pp. 1825–1829, 2017. <https://doi.org/10.1166/jmihi.2017.2267>.
- [5] <https://www.mayoclinic.org/diseases-conditions/breast-cancer/symptoms-causes/syc-20352470>
- [6] H.S. Sheshadri, and A. Kandaswamy, "Breast tissue classification using statistical feature extraction of mammograms," *Medical Imaging and Information Sciences*, vol. 23, no. 3, pp. 105-107, 2006.
- [7] R.I.R.Thanaraj, B. Anand, J. A. Rahul, and V. Rajinikanth, "Appraisal of Breast Ultrasound Image Using Shannon's Thresholding and Level-Set Segmentation," in *Progress in Computing, Analytics and Networking*, pp. 621-630. Springer, Singapore, 2020.
- [8] V. Rajinikanth, N.S.M. Raja, S.C. Satapathy, N. Dey, and G.G. Devadhas, "Thermogram assisted detection and analysis of ductal carcinoma in situ (DCIS)," in: *International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICICT)*, IEEE, pp.1641–1646, 2018. <https://doi.org/10.1109/icicict1.2017.8342817>.
- [9] A.Q. Al-Faris, U. K. Ngah, N. A. M. Isa, and I. L. Shuaib, "Breast MRI tumour segmentation using modified automatic seeded region growing based on particle swarm optimization image clustering," in *Soft Computing in Industrial Applications*, pp. 49-60. Springer, Cham, 2014.
- [10] N. Shrivastava, and J. Bharti, "Breast tumor detection and classification based on density," *Multimedia Tools and Applications*, vol.79, no. 35, pp.26467-26487, 2020.
- [11] S.P.S. Raj et al., "Examination of digital mammogram using otsu's function and watershed segmentation", in: *Fourth International Conference on Biosignals Images and Instrumentation (ICBSII)*, pp. 206-212, 2018. DOI: 10.1109/ICBSII.2018.8524794.
- [12] N. S. M. Raja, S. L. Fernandes, N. Dey, S. C. Satapathy, and V. Rajinikanth, "Contrast enhanced medical MRI evaluation using Tsallis entropy and region growing segmentation," *Journal of Ambient Intelligence and Humanized Computing*, pp.1-12, 2018. <https://doi.org/10.1007/s12652-018-0854-8>.
- [13] S.L. Fernandes, V. Rajinikanth, and S. Kadry, "A hybrid framework to evaluate breast abnormality using infrared thermal images," *IEEE Consum. Electron. Mag.*, vol. 8, no.5, pp.31–36, 2019. <https://doi.org/10.1109/MCE.2019.2923926>.
- [14] <http://www.thermography.co/in/>
- [15] X. Hou, and L.Zhang, "Saliency detection: A spectral residual approach," in *2007 IEEE Conference on computer vision and pattern recognition*, pp. 1-8. IEEE, 2007.
- [16] F. Meyer, and S. Beucher, "Morphological segmentation," *Journal of visual communication and image representation*, vol. 1, no. 1, pp.21-46, 1990.
- [17] N. Dey et al., "Social-Group-Optimization based tumor evaluation tool for clinical brain MRI of Flair/diffusion-weighted modality," *Biocybernetics and Biomedical Engineering*, vol. 39, no. 3, pp. 843-856, 2019. <https://doi.org/10.1016/j.bbe.2019.07.005>.
- [18] A. Bakiya, K. Kamalanand, V. Rajinikanth, R.S. Nayak, and S. Kadry, "Deep neural network assisted diagnosis of time-frequency transformed electromyograms," *Multimedia Tools and Applications*, vol. 79, no. 15, pp.11051-11067, 2020.
- [19] N. Dey, Yu-Dong Zhang, V. Rajinikanth, R. Pugalenth, and N.S.M. Raja, "Customized VGG19 architecture for pneumonia detection in chest X-rays," *Pattern Recognition Letters*, vol.143, pp.67-74, 2021. <https://doi.org/10.1016/j.patrec.2020.12.010>.
- [20] A. Gudigar et al., "Global weighted LBP based entropy features for the assessment of pulmonary hypertension," *Pattern Recognition Letters*, vol.125, pp.35-41, 2019. <https://doi.org/10.1016/j.patrec.2019.03.027>.
- [21] A. Faramarzi, M. Heidarnejad, S. Mirjalili, and A. H. Gandomi, "Marine Predators Algorithm: A nature-inspired metaheuristic," *Expert Systems with Applications*, vol. 152, pp.113377, 2020.
- [22] M.A.A. Al-Qaness, A. A. Ewees, H. Fan, L. Abualgah, and M. A. Elaziz, "Marine predators algorithm for forecasting confirmed cases of COVID-19 in Italy, USA, Iran and Korea," *International journal of environmental research and public health*, vol. 17, no. 10, pp.3520, 2020.
- [23] M. Abdel-Basset, R. Mohamed, M. Elhoseny, A.K. Bashir, A. Jolfaei, and N. Kumar, "Energy-aware marine predators algorithm for task scheduling in IoT-based fog computing applications," *IEEE Transactions on Industrial Informatics*, 2020. DOI: 10.1109/TII.2020.3001067.

- [24] M.A. Khan et al., "Computer-aided gastrointestinal diseases analysis from wireless capsule endoscopy: A framework of best features selection," *IEEE Access*, vol. 8, pp.132850-132859, 2020. Doi: 10.1109/ACCESS.2020.3010448.
- [25] D. Połap, and M. Woźniak, "Red fox optimization algorithm," *Expert Systems with Applications*, vol. 166 pp.114107, 2021. <https://doi.org/10.1016/j.eswa.2020.114107>.

**2021 IEEE Seventh  
International Conference on  
Bio signals, Images and  
Instrumentation  
ICBSII 2021  
SESSION VIII  
RESEARCH  
PAPERS**

# Development of a Smart Cervical Collar System for Feedback Control of Intracranial Pressure after Traumatic Head Injury

Frank M. Washko, Ph.D.  
College of Science and  
Engineering  
Flinders University  
Tonsley, South Australia  
[frank.washko@flinders.edu.au](mailto:frank.washko@flinders.edu.au)

William S. Edwards, Ph.D.  
Department of Industrial and  
Manufacturing Engineering  
Kettering University  
Flint, Michigan, USA  
[wedwards@kettering.edu](mailto:wedwards@kettering.edu)

Leslie A. Washko, M.Ed, MBA  
Department of Engineering  
Management  
Saint Martin's University  
Lacey, Washington, USA  
[lwashko@stmartin.edu](mailto:lwashko@stmartin.edu)

**Abstract**— While mild traumatic brain injury may affect brain cells temporarily, more-serious traumatic brain injury can result in bruising, torn tissues, bleeding, and other physical damage to the brain. These injuries can result in long-term complications or death. Studies have shown that care for brain injuries should begin as soon as possible after an injury occurs, and that by reducing intracranial pressure common in traumatic brain injuries some of those complications may be mitigated. The objective of this research is to assess how applying a negative inspiratory pressure instantaneously lowers intracranial pressure and how this reduction persists over time with continued enhanced negative inspiration pressure. There is research that suggests that ICP is reduced approximately 0.75 mmHg for every 1 mmHg reduction in ETP, in animal models, and this study shows promise for positive outcomes on humans using a cervical collar with a vacuum chamber that regulates pressure on the ventral areas of the neck.

**Keywords**—traumatic brain injury, TBI, intracranial pressure, ICP, pressure, control.

## I. INTRODUCTION

The objective of this system is to reduce the various secondary outcomes that can occur as a result of increased intracranial pressure (ICP) in a traumatic head injury [1]. Secondary outcomes may include sensory impairment, moderate cognitive impairment, severe brain injury, coma, or even death. While an ICP of about 0 to 10 mmHg is normal, pressures between 20 and 40 mmHg can lead to substantially reduced positive outcomes for patients. To minimize these secondary effects, care should begin as soon as possible after injury. [2] Ideally, ICP will be reduced by first responders within the “Golden Hour.” [3]

Negative inspiration pressure imposed on the thorax during respiration results in instantaneous decrease in ICP, including increased ICP due to head trauma. [4] In the Yannopoulos

study, it was shown that a pressurized pneumatic band placed around the abdomen resulted in negative inspiration pressure, which resulted in instantaneous decrease in ICP. The resulting data is shown below, showing that, as Endotracheal Pressure (ETP) is reduced, average ICP is correspondingly, and instantaneously, reduced. The reduced ETP, or increased ETP vacuum, is the result of the negative thoracic pressure.

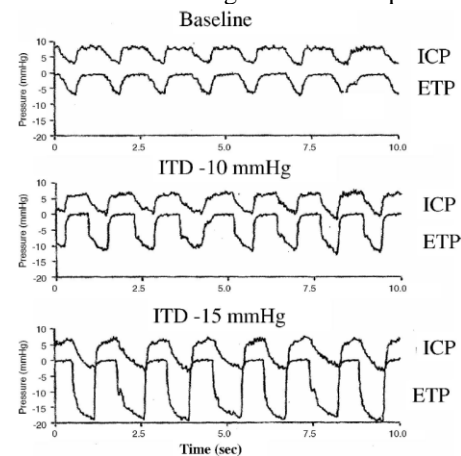


Figure 1. ICP Corresponding to Reduced ETP

The purpose of this research was to replicate this concept in a device that could actually be deployed in the field. Research has shown that ETP can be similarly reduced by applying a negative pressure to the neck, and the same benefit can be realized. [5] That is, negative inspiratory pressure can be created by applying a vacuum to the neck, resulting in decreased ETP, and instantaneously decreased ICP via the vertebral venous system. [6] Therefore, the lightest, smallest, and most efficient way to decrease ICP is by applying vacuum at carotid sinuses of a patient's neck, via a pneumatic chamber in a cervical collar. When coupled with a feedback control system of pressure, this would form the basis of a “Smart Cervical Collar.”

## II. BACKGROUND AND PROPOSED DESIGN

The design of the Smart Cervical Collar with a vacuum diaphragm chamber that applies between 0 and 100 mmHg negative pressure on the ventral neck, is shown in the figure below.

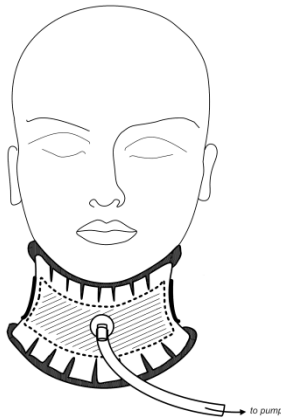


Figure 2. Smart Cervical Collar Basic Design

The system integrates a small pneumatic pump, battery, and controller. The system also includes provisions for various sensors, including a respiration sensor, heart rate sensor, and non-invasive ICP sensor, to provide complete, closed loop feedback control of ICP. Vacuum pressure can be correspondingly increased, or reduced, in response to ICP to keep ICP in a safe range. And, vacuum pressure could further be synced with respiration, or pulse, to provide the most effective and efficient reduction in ICP. A further benefit of the controller is that it can easily record sensor data in memory to analyze in subsequent treatment off the field, in a medical facility, or in Prolonged Field Care (PFC). Using feedback control, the system does not require the medic to set vacuum pressure, but instead lets the system adjust vacuum pressure as needed, while the medic tends to other conditions.

Applying a negative inspiratory pressure instantaneously lowers intracranial pressure and this reduction persists over time with continued enhanced negative inspiration pressure. Indeed, the current research suggests that ICP is reduced approximately 0.75 mmHg for every 1 mmHg reduction in ETP, in animal models. [4] The design concept involves applying negative inspiratory pressure using flexible silicone vacuum chambers sealed to the patient's neck. A polymer shell is shaped and dimensioned to fit a ventral aspect of a patient's neck, and is held in place with a strap – with sufficient stiffness to avoid collapse during vacuum. The research has shown that applying the negative pressure in two areas – at the right and left ventral neck – provides the most effective and efficient negative inspiratory pressure. This side design, which applies vacuum to the carotid sinuses, is shown below.

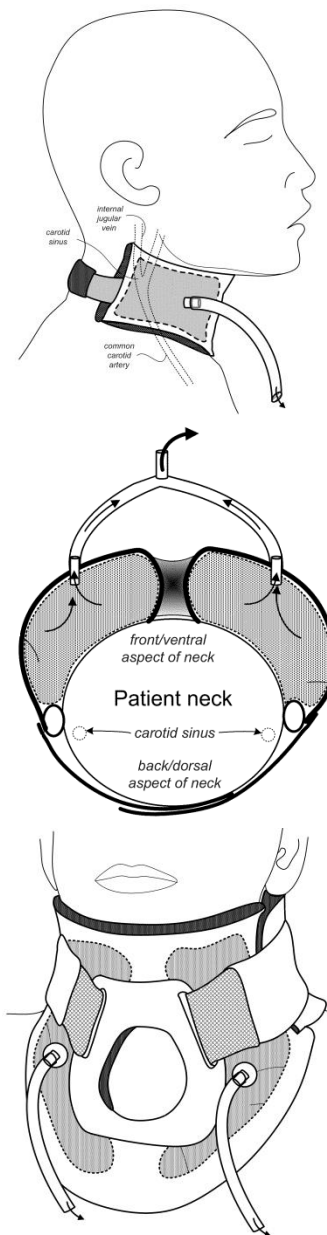


Figure 3. Smart Cervical Collar Side Design

With this design, spacers are placed on the sides of the neck in the shell to avoid placing any positive pressure on the carotid sinuses. The carotid sinus receptors are baroreceptors located in the carotid sinuses of the left and right internal carotid arteries. These baroreceptors monitor the pressure of the blood being delivered to the brain. At normal resting blood pressures, baroreceptors discharge with each heartbeat. When blood pressure falls, baroreceptor firing rate decreases and baroreceptor reflexes help restore blood pressure by increasing heart rate. Impulses from the sinus travel up the carotid sinus nerve to the nucleus of the tractus solitarius (NTS) in the medulla. Stimulation of the vagal nuclei in the medulla results in reduced heart rate (bradycardia). Therefore, avoiding direct pressure on the carotid sinuses, and the dual ventral vacuum



design prevents the corresponding risks and complications of direct stimulation. [4]

Instead, the two-ventral vacuum chamber design places negative pressure, generated through cervical tissues, and is transmitted to the vertebral venous system, thereby lowering intracranial pressure. Thus, intracranial pressure is lowered easily without increasing the work of breathing, without needing to be intubated, and with non-invasive methods. The research is currently incomplete as to the best way to apply the vacuum. The major area requiring development in this study is the extent of vacuum that is required, as the system was designed for a pressure between 0 and 100 mmHg.

A secondary benefit of this system is that it provides field medics with a cervical collar to use to stabilize neck injuries, irrespective of the TBI system. Therefore, the system does not simply take up valuable medic bag space for only one treatment; it also provides a fold-flat cervical collar to use even when increased ICP is not present. For first responder applications, where size and volume are paramount, the design integrates the vacuum chamber system into a single collapsible cervical collar design as shown below. The device will easily fit within a medic or corpsman bag, is ruggedized, and functions in severe environments. For the civilian application, a standard, fixed collar can be designed.

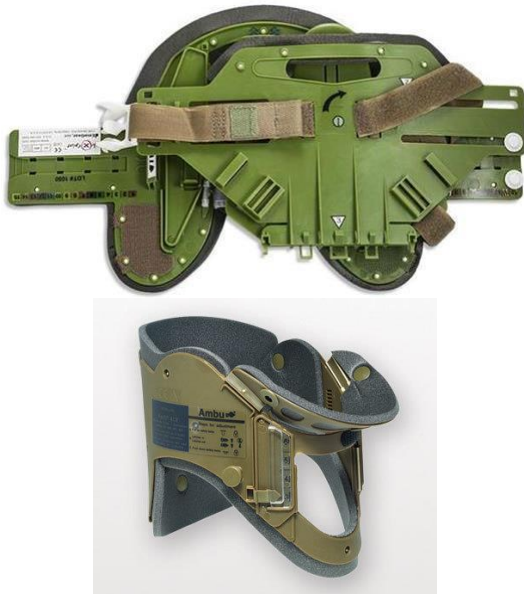


Figure 4. Foldable Cervical Collar for First Responders

The result of the commercialization study shows that the market for such a device is significant. The Smart Cervical Collar is equally applicable to civilian applications where paramedics and first responders can deploy this system to reduce the secondary effects caused by increased ICP. The obvious application is anywhere there is treatment of trauma patients, and anywhere first responder equipment is deployed.

### III. METHODS

For this preliminary study, the system consisted of two air pumps providing a precise vacuum to a flexible chamber on the cervical collar at the carotid sinus. A pressure sensor was inserted in the vacuum chamber and an intracranial pressure sensor was wired into the system. The system also included pulse and respiration sensors, that integrate with a single controller. The majority of components in this system are simple air pumps and vacuum sensors in the appropriate range. The most complicated sensor in this system is the sensor to measure intracranial pressure as the variable to be controlled with the PID algorithm. Several non-invasive sensors using sonography through the ears have become available to measure ICP recently. [7] One successful simple system is the HS-1000 light headphones from HeadSense, which are easily wired into the system. [8] The first round of prototypes is built around an Arduino Uno controller, which provides PID control and integrates all of the sensors required for this project. The Uno easily accommodates the pulse sensor, respiration sensor, multiple pressure sensors, and controls for multiple pumps. It is also a very flexible platform to develop an algorithm. Once this system is fully developed, it is easy to migrate the algorithm to a slightly smaller and more power efficient platform for the next evolution.

#### A. Simulation

Once the designs were completed, the control system was simulated to establish and optimize the baseline algorithm. The PID control system, algorithm, and gains have been simulated and improved in Simulink. The results were used to finalize the one-pump design with low error.

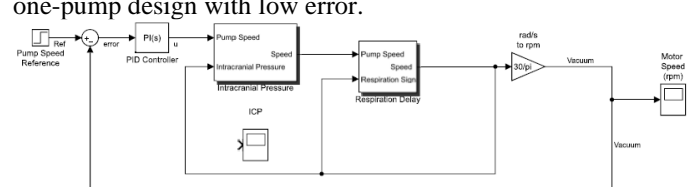


Figure 5. Matlab Control System Model

Some results of the simulation are shown below, showing decreased intracranial pressure as a result of vacuum pressure. Pressures are given in mmHg. Input to the system is a simplified pulse train representing intracranial pressure.

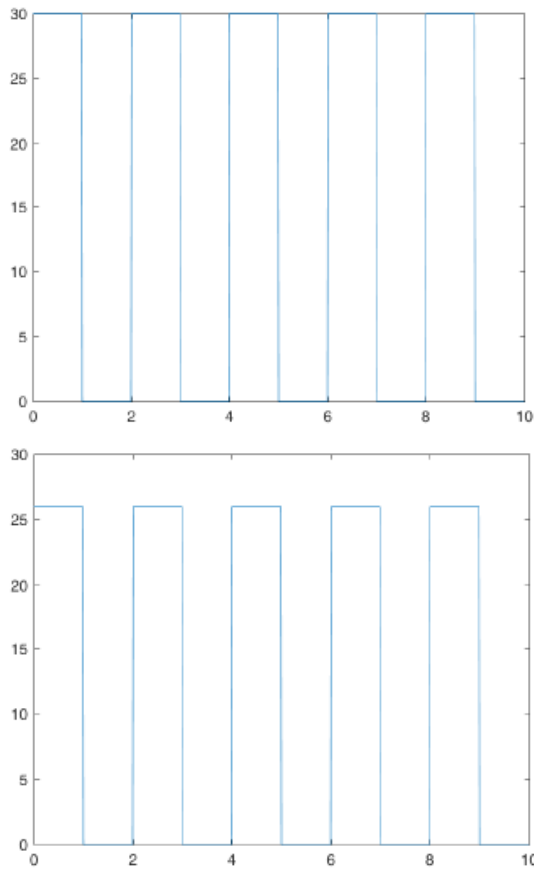


Figure 6. Simulation Input Pressure Pulse Trains

The first figure shows the beginning input ICP signal (at 30 mmHg), and the second figure shows the reduced pulse figure at approximately 27 mmHg after the PID control algorithm.

Additionally, the function of the pneumatic system was being simulated with a 1D fluid simulation model based in Matlab – Simscape (Simscape Fluids). Results of the simulation were fed back into the control system design and the pneumatic system design prior to Phase I hardware testing design. The simulation further bolstered the basis of the simplified, but effective, one-pump design with low error.

The block diagram of the Simscape simulation is shown in the figure below.

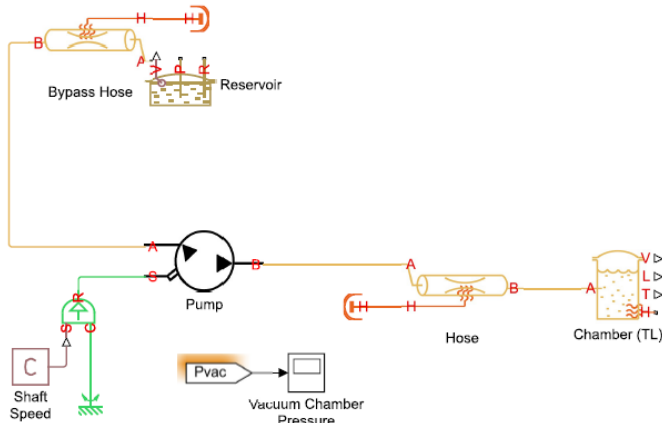


Figure 7. Matlab Simscape – Fluid Simulation Model

**B. Testing**

After simulation and final design improvement, the Phase I prototype test setup was fabricated. Alternative pump concepts were tested, including multiple pumps (one positive and one negative), pumps with a dithering pressure relief valve, and so on. The different configurations were researched to determine the best combination for pressure control and accuracy, and the algorithm was refined with each revision for better performance. Early functional testing of the system indicated some error in the vacuum system, so several backup systems were designed and fabricated, including systems with single pumps, multiple pumps, and with solenoid controlled relief valves.

Ultimately, the system using one pump was refined to the point of optimal performance. One pump with no additional dithering or relief valves is best for cost, weight, power efficiency, and battery life, so considerable effort was expended to make the control algorithm effective with the single pump.

The pneumatic system, including the pumps, tubing, valves, and vacuum chamber have been designed and analyzed. The flexible cervical collar vacuum chamber profile was designed, with a size and shape for a large range of the male and female population range, as shown in the figure below.



Figure 8. Flexible Vacuum Chamber

The silicone vacuum chamber was subjected to further stress and deformation analysis and durability simulation (finite element analysis) to determine that it will not collapse or fail in operation, when subjected to the maximum system pressure. Results of this simulation are shown below.

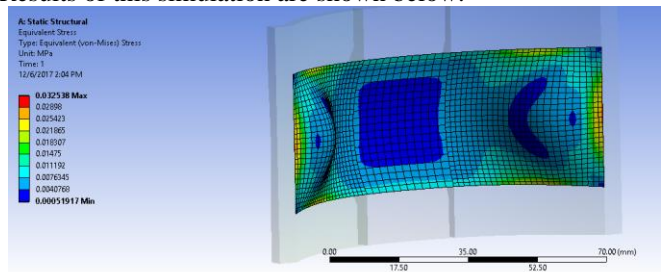


Figure 9. FEA Analysis of Vacuum Chamber

In the bench test, an 8 kilogram plastisol human neck model was fabricated to attach to the Smart Cervical Collar. In the

model, chambers were fabricated into the neck and, and the neck was pressurized with a fluid around fabricated carotid sinuses. In the head, a pressurized chamber was fabricated in fluid communication with the neck chamber. Pressurized pulsatile fluid was pumped through the ICP chamber and through neck and carotid sinus area, and pumped out to a reservoir, using a Cole Parmer pulsatile pump with the same frequency and flow as human blood flow. The carotid sinus area was covered with thin silicone sheet to simulate the neck skin, so that the applied vacuum from the Smart Cervical Collar could be applied through the fluid in the neck, and onto the carotid sinuses.

This fabrication is shown in the figures below.

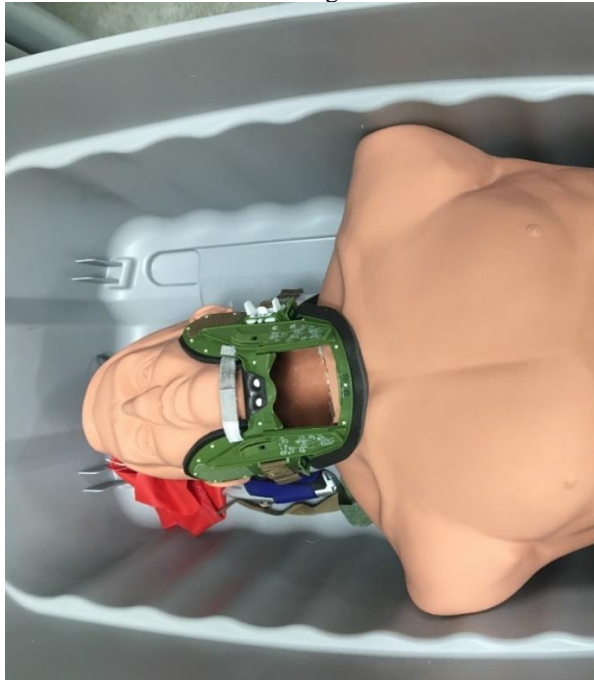


Figure 10. Experimental Test Setup

The intracranial pressure chamber fabricated into the head was instrumented with a variety of different pressure measurements, using the Cole Parmer EW-68075 High Accuracy Pressure Transducer. These readings measure the dynamic ICP pressure changes as a function of pump pressure and the applied pressure in the Smart Cervical Collar chamber, as shown in the figure below.



Figure 11. Intracranial Pressure Sensor

The entire system was designed to be contained in a portable box, approximately 75mm x 150mm, as shown in the figure below.

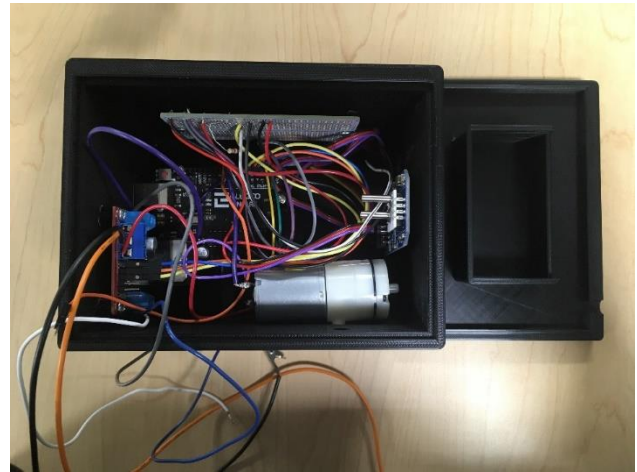


Figure 11. Control System

It is important to note that while a variety of high precision pressure transducers have been used in this system and during testing, it was impossible to use an actual ICP sensor, because of the lack of a human subject and a biological signal to measure with an ICP sensor. Developing testing measurement parameters based on bench test data. The system was tested at two different applied vacuum pressures.

IV. ANALYSIS

The results of the simulation and testing protocols described above are given in the sections below.

A. Simulation

The result of the Simscape pneumatic simulation, showing reduced intracranial pressure, as a simplified pulse wave, is decreased from a higher value, to a lower value, in the system simulation. Units are in mmHg.

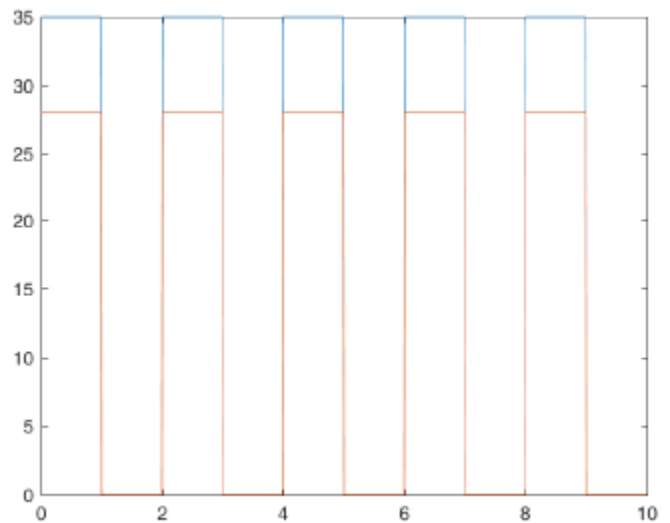


Figure 12. Simulation Pressure Results

In this simulation, the system reduces a 35mmHg pulse train to about 28.5mmHg after the system input. Although the results, in absolute terms, are not meaningful without calibration with the physical model, the simulation indicates that the system should provide pressure reduction. The results of the system simulations were fed back into the control system design to refine and improve the control system design prior to fabrication and testing.

### B. Testing

Initial results show that on the bench, the system functions as intended. On the bench model, the ICP pressure has shown a reduction in pressure of approximately 0.19 mmHg for every 45 mmHg application of vacuum to the carotid sinuses, as shown in the data table below.

Applied Pressure to Carotid Sinus	ICP Chamber Pressure Reduction
45 mmHg	0.19 mmHg
65 mmHg	0.31 mmHg

Table 1. Initial Bench Testing Results

This is significantly less reduction and dynamic response than that observed in the Yannopoulos study discussed above. [4] However, it was well-understood entering this initial study that the rough bench test setup would not have results equivalent to that of the human body. The intent of the preliminary testing was to prove functionality and efficacy of the system – not biomimetic performance.

The system shows that the closed loop control functionality of the system operates correctly. And it showed that application of negative pressure to the neck would likely result in a corresponding reduction in intracranial pressure.

## V. CONCLUSIONS

The main objective of the Smart Cervical Collar research was to design and bench test the system for initial proof of concept. While it has been successful as a proof of concept, the next step is clear: the system must be tested on a human. Because the time and budget of this initial study did not allow for human trials, it was not possible to test the system with real-time feedback from the intracranial pressure (ICP) sensor. This prototype used a simulated ICP signal in simulation, and used pressure generated from a fluid pump for bench testing. The proof-of-concept system was successful in reducing ICP on a simulated person, so there is evidence that continued research may yield positive results. The main objective of the next phase will be to integrate an ICP sensor into this system, and test the complete system on several people, under the FDA First in Human (FIH) early feasibility study program.

## REFERENCES

- [1] Abdelhak, T, Abrego GC., Traumatic brain injury. In: Wartenberg KE, Shukri K, Abdelhak T, eds. *Neurointensive Care: A Clinical Guide to Patient Safety*. New York, NY: Springer; 2015:219-48.
- [2] Joint Theatre Trauma System: Management of Patients with Severe Head Trauma Clin. Prac. Guide, 16 Jun 2014.
- [3] Rasmussen, T., *et al.*, In the “Golden Hour”, Army AL&T, January-March 2015, 80-85.
- [4] Yannopoulos, D. *et al.* “Intrathoracic pressure regulation for intracranial pressure management in normovolemic and hypovolemic pigs” *Crit. Care. Med.* 34 [Suppl.] (2006) 495-500.
- [5] S. Ogoh, *et al.*, “Does pulsatile and sustained neck pressure or neck suction produce differential cardiovascular and sympathetic responses in humans?” *Experimental Physiology* 88.5 (2003) 595-601.
- [6] J.T. Potts and P.B. Raven, “Effect of dynamic exercise on human carotid-cardiac baroreflex latency” *Am. J. Physiol.* 268 (Heart Circ. Physiol. 37): (1995) 1208-1214.
- [7] Avan, P., *et al.*, “Noninvasive In-Ear Monitoring of Intracranial Pressure During Microgravity in Parabolic Flights,” *J. Appl. Physiol.* (Aug. 1 2018).
- [8] Herklots, M.W., *et al.*, “Prospective Evaluation of Noninvasive HeadSense Intracranial Pressure Monitor in Traumatic Brain Injury Patients Undergoing Invasive Intracranial Pressure Monitoring,” *World Neurosurgery* (Oct. 2017).

# Data Acquisition System for Web-based Multi-modal Data Repository

1<sup>st</sup> Munyaradzi Charles Rushambwa  
*Electronic Engineering*  
*Harare Institute of Technology*  
 Harare, Zimbabwe  
 mrushambwa@hit.ac.zw

2<sup>nd</sup> Dr. Anirban Mukherjee  
*Electrical Engineering*  
*Indian Institute of Technology Kharagpur*  
 Kharagpur, India  
 anirban@ee.iitkgp.ac.in

3<sup>rd</sup> Maitreya Maity  
*Electrical Engineering*  
*Indian Institute of Technology Kharagpur*  
 Kharagpur, India  
 maitreya.maity@gmail.com

**Abstract**—The multimodal medical data with annotation helps to build different automated algorithms. Each reported work has used a specific disease database and developed a CAD based on the considered database. Therefore, the availability and quality of medical database play the most crucial role in developing any CAD. However, it has been observed that most of the reported studies used public database (created by foreign universities/centres) or private database. Unfortunately, the availability of a national medical database in India is negligible. However, development of such medical database is possible. Such medical database can encourage new research activity and help a large research community. The proposed study focuses on developing an online public medical multimodal database platform. Here, the data acquisition software is build for collecting various information.

**Index Terms**—database, data acquisition, medical, repository

## I. INTRODUCTION

Computer-aided diagnosis (CAD) has become one of the important research subjects in medical engineering domain. From early '70s researchers have focused on developing algorithms and systems for disease diagnosis and clinical decision making using Artificial intelligence [1-4]. The outcome of these researches brought many useful and reliable applications which help medical professionals to make examination and clinical decision effortlessly. The medical diagnosis devices are upgraded with time and also deliver more detailed information about the disease. In recent time the multimodal medical data, i.e. Biosignal (Eeg, Ecg, Emg etc.), Image (MRI/CT, Ultrasound, X-ray, Micro/Macroscopy etc.), Clinical (Patient history, questionnaire etc.), Genetics etc. are used in disease diagnosis. Due to the availability of a large amount of multimodal medical data, the research on CAD becomes a top priority in the research community. Researches have utilized the knowledge of medical science and trained a computer to make a medical decision. Many medical repositories are found online each with its own specific advantages over the other. However, due to the complexity of some research work, where some of the data provided does not match the localised research of specific areas, regions or countries, there is need to develop custom databases for a certain geography. The World Health Organisation (WHO) Global Health Observatory data repository only allows one to get statistical information

about certain type of medical information but not necessarily giving raw outputs [5]. Most repositories require authors to submit their information or datasets to public data repositories in contrast to a community- specific repository approach without generalising information. The quality and usefulness of information given in some datasets may vary distinctively such that it may have a negative impact on the research outputs of some researchers. One of the widely used database is the Physio-net database which gives a description of each dataset yet some of the data is restricted for access to registered users only. The physionet has a vast collection of datasets that are randomly uploaded onto it but the way one has to search for specific information about a particular dataset makes it very difficult to bring accurate results [6]. The data is uploaded by various researchers and its accuracy and appropriateness is questionable and would require one to compare with real-time data from a clinical setup. This then gives an advantage on this developed dataset that gives verified data from a clinical practitioner since the data acquisition setup is from a medical facility. To make sure that the information is correct and authentic, the data acquisition system will be integrated with the local hospital's electronic health records system (EHR).

## II. METHODOLOGY

The proposed development is distributed into three phases (Fig 1). Initially, a frontend application will be developed for uploading multi-modal medical data into the database. The protocol and format of data coming from various medical devices are different from each other.

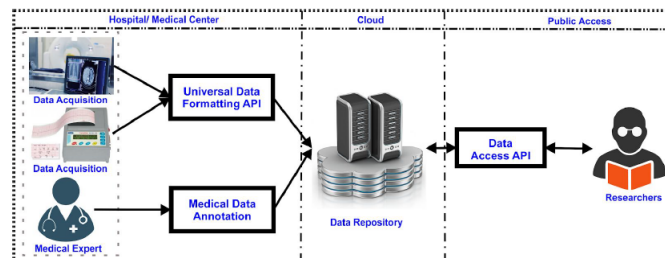


Fig. 1. System Abstraction.

Funded By DST under RTF-DCS India

It is needed to make universal protocol to adapt these data and transferred it into a symmetric format. The Universal Data Formatting API will perform the data transformation task. Besides uploading medical data into the system, it is also required to generate corresponding annotated data. The Medical Data Annotation module will help medical professionals to view and mark abnormalities present in the data. In the second phase, The cloud-based system is responsible for building the data repository. The medical data with annotation are stored into a central database system through web protocol. Therefore, a database schema and corresponding data communication platform will be developed. In the end, the Public Access module is developed for accessing the database remotely by individual/organizations. The Data Access API will help to provide a secure platform for the user to access the database. Here user can search for the required data and view/download the medical data on demand. The overall system will be developed in IIT Kharagpur and deployed in AIIMS, Bhubaneswar. The medical data will be uploaded into the system from AIIMS, Bhubaneswar.

**A. Back End**

1) *Database Setup:* XAMP is a free and open source database configuration software that can be used to locally host Apache, distributions containing MariaDB, PHP among other applications. In Fig. 2 shows the module name and the corresponding process IDs that are used by the applications as they run on the windows-bases services.

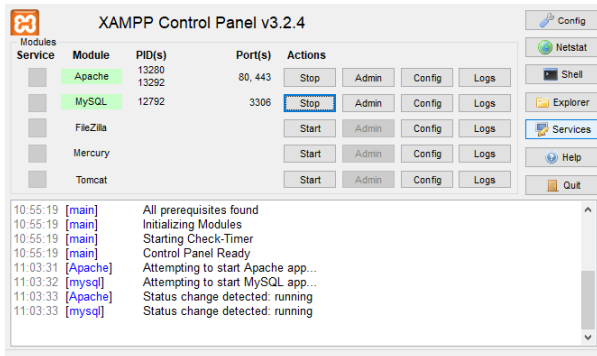


Fig. 2. XAMPP Configurations.

In configuring the Apache module, it is important to note that there is need to change the listen port that binds Apache to the specific IP address to match the computer on which it is hosted. When Apache is running PhpMyAdmin is installed from the web-interface by accessing the installation files on the local server (172.0.0.1) from the web-interface.

2) *MySQL:* In configuring MYSQL it was noted that there is need to specify the port which will be used for all MYSQL clients to establish a connection and in this project, port 3306 was used, In a situation where we need to launch the application on a different server, there is need to change the host address to the match the one that is going to be used which is called the bind address as referred to in the configuration file.

3) *Database Interface:* The database interface is on PhpMyAdmin webpage through which SQL query commands for the database can be run from. The PhpMyAdmin is run on a local host address and there was need to make sure that the binding with the launch address has already been done in the configuration file. Limited access to the databases in the PhpMyAdmin dashboard is given and user groups are created, each with different privileges as shown in Fig 3.

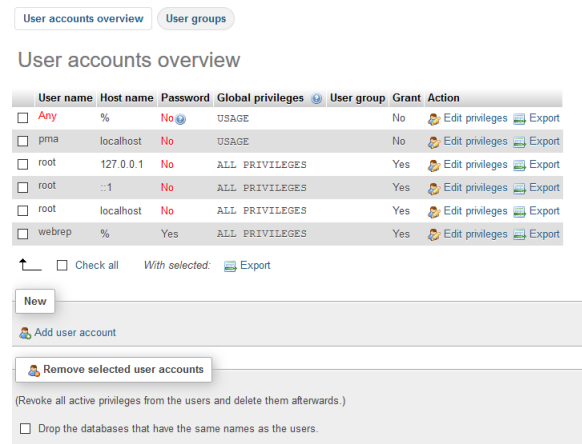


Fig. 3. Privileges Configuration.

There are also options of editing user privileges, deleting or adding new user. for extra security as seen in Fig. 4 the privileges can be limited to data, structure, administration resource limit or secure shell logins (SSL). This then gives an extra security on the access to the data in the system.

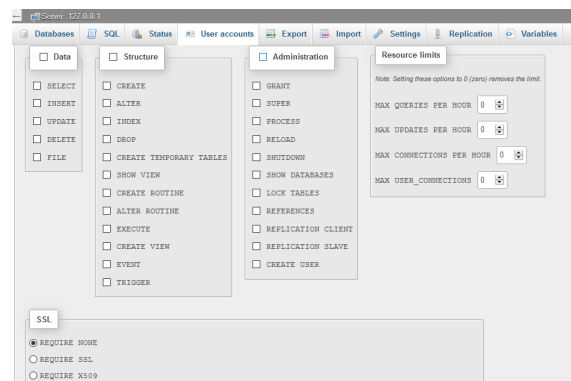


Fig. 4. Privileges Setup.

4) *Schema Design:* The schema design for this project has 13 tables that are used to store information. The first and the most important table is the users table, where there is a column for username, user type, password, administration name and administration password in case one is an administrator as shown in Fig. 5.

The access table has email address, last name, first name user name, station, designation and textual request columns as seen in Fig. 6. The access table is used to capture information about new users who wish to access the system. The biosignals

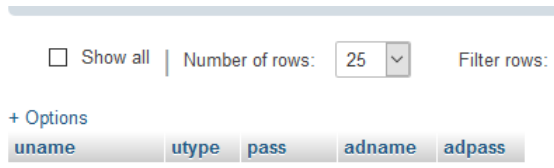


Fig. 5. Users Table.

table is used to store information about EEG, ECG, EMG, speech, MMG, EMG, EOG, GSR and MMG which is all stored in specific standard forms.

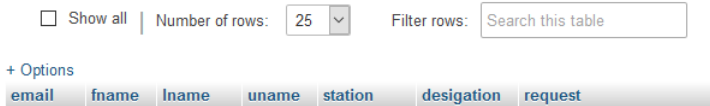


Fig. 6. Request Access Table.

To each entry, a unique ID that is referenced to central columns in all data tables is given for identification of specific individual data. File paths for ZIP, DAT or any other formats can also be saved in this table for downloading on the web-based interface. The cinfo table stores clinical information such as allergies, family history, medical history, current and previous medications, immunizations, diagnoses, previous recorded symptoms and additional information as may be required. The cmeasure table stores clinical measurement data such as height, weight, temperature, blood pressure, heart rate, oxygen concentration (SPO2), blood glucose and body-mass index. The database is formatted such that it only allows numeric recordings saved as decimal values along with corresponding units of each measurement. Diseases from the WHO ICD-10 nomenclature are uploaded onto the database and shown as having an ICD code, class and the corresponding description. Imaging table contains information about ultrasound, X-ray, MRI, fluoroscopy, CT and PET scans. These are stored as images in either raw format, as a file path or as a zipped file. Labtests table contains test type ICD10 modifier, description, type name, values and the corresponding standard units. Pinfo is a table for patient specific information such as race, gender, age, blood group, demography and diet.

**B. Front End**

The front-end application was designed visual studio 2019 version using structured C-sharp language. Different windows forms were built on the platform and the forms are used to collect data from a clinical setup as a bundled application. Information can be collected in different forms which are:

- real time capture using serial ports
- manual single entry
- manual file upload (formatted csv)
- manual file upload (zipped file)

**III. OUTCOME**

The initial form which appear upon running the system is the login form which shows the usertype, username, password, login button, admin button and the cancel button as seen in Fig. 7.

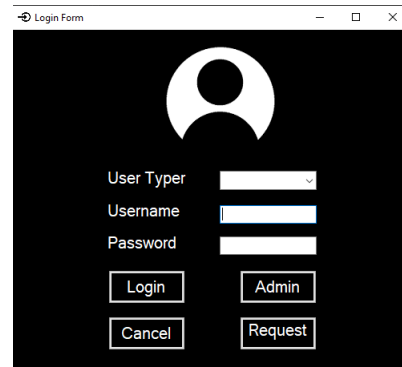


Fig. 7. Login Form.

The admin button is used in case one wants to login as admin and the request button is for when one wants to register/request access to the system. The usertype is a dropdown box where one chooses the user type which he has been given already by the administrator. Upon selecting the admin login, the user is directed to an admin page where they can create users from the requests that would have been submitted, with the latest request being highlighted on top as seen in Fig. 8.

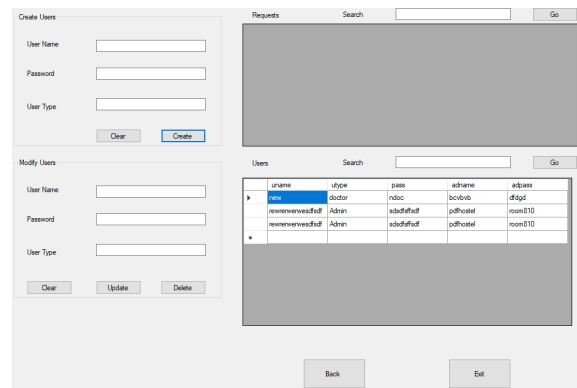


Fig. 8. Admin Panel.

In case of a double creation of an account, the system gives an error. Apart from account creation on this page, the admin has also rights to view, edit or delete users. If one clicks the request access button, they are directed to the page in Fig. 9 where new users can enter their information which is sent to the administrator for verification and account creation.

The administrator will receive a notification for the requisition and act accordingly. Upon successfully logging in as authorized personnel with full rights, one is then taken to the main page which shows links to the other pages as seen in Fig. 10.

Fig. 9. Request Access Form.

select to save to the database.

Index	Modifier	Description	Unit Name
1	BB	DPODF	DPODF
2	noheader	description	ipename
3	BB	DPODF	DPODF
4	SAD	DADAS	DASD
5	noheader	description	ipename
6	BB	DPODF	DPODF
7	SAD	DADAS	DASD
8	BA	Context, serum	BIOCHEMISTRY
9	BA	Context, serum	BIOCHEMISTRY
10	BA	Context, serum	BIOCHEMISTRY
11	BA	Context, serum	BIOCHEMISTRY
12	BA	Context, serum	BIOCHEMISTRY
13	BA	Context, serum	BIOCHEMISTRY
14	BA	Context, serum	BIOCHEMISTRY
15	BA	Context, serum	BIOCHEMISTRY
16	BE	Sodium, serum	BIOCHEMISTRY
17	BE	Proteasum, serum	BIOCHEMISTRY
18	BE	Sodium and data	BIOCHEMISTRY
19	BE	Protein, serum	BIOCHEMISTRY

Fig. 12. Lab Tests Page.

Fig. 10. Main Page.

The patient info page shows the information which needs to be saved in the database. To save time, one can upload a formatted csv file containing information about a number of patients as seen in Fig. 11.

To save time, make it easier and ensure correct information is uploaded, one can search for a test using the search box at the top in figure 13 and select the test from the gridview, this automatically fills the textboxes as seen in Fig. 13. One only need to enter the values and the corresponding units then save to the database. The biosignals page shows where one can save biosignal information and the corresponding formatting required. The file name to be uploaded is shown before uploading as seen in figure 14.

Fig. 11. Patient Info Page.

Fig. 13. Biosignals Page.

To avoid wrong entries, the patient ID is autogenerated from the database and all other fields are drop down menus except for the additional comment box. The clinical information page shows textual information to be saved in the database as seen in Fig. 12 the labtests page shows the type of labtest one can

The imaging page shows the type of imaging format to be uploaded to the database. There is an option of bulk upload of the specific imaging file type as seen in Fig. 15.

The clinical measurements page shows the clinical measurements that can be saved to the database as seen in figure 16, and the corresponding value and units, the page is designed such that the user will only input values in one box and select units from the other box.

IV. CONCLUSION AND DISCUSSION

The data acquisition system was successfully designed for both the front end and the back end. The file types for





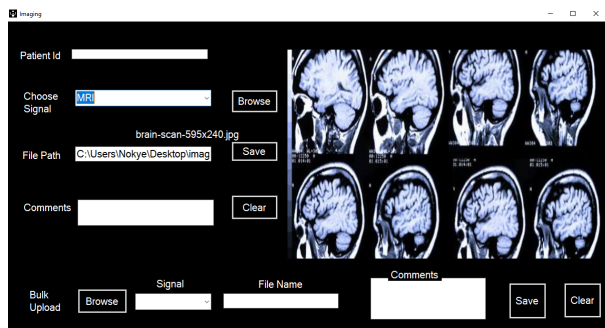


Fig. 14. Imaging Page.

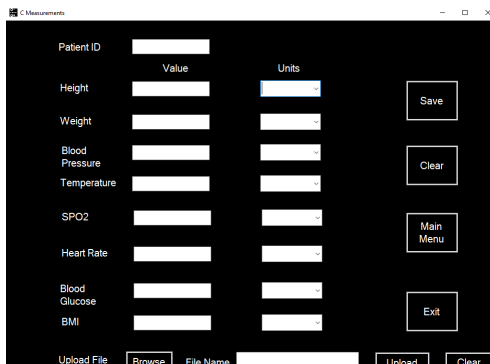


Fig. 15. Clinical Measurements.

each information entry can be formatted to suit the required standard depending on the station which the system will be installed at. The system can automatically format the file paths given upon installation using the publishing wizard in Microsoft visual studio. However, this is an ongoing project and as future work, there is need to work on the web-interface that gives access of the datasets to the general public for use in research or any other authorized work. To note, in this design is that there was need to make sure that the patient identification (name or any other leading information) remain confidential in accordance to the Patient Privacy Act and the Hospitals Act. File formats are to be saved in internationally accepted formats and using the standards SI units in case of measurements. Stability of the system is ensured through a possible hosting on a local server or intranet for every installation, thus reducing risks of attacks from external intruders. The system is designed dynamically to suit all types of clinical environments with special features added to make the interface user-friendly.

## V. REFERENCES

- 1) Rogers, W., B. Ryack, and George Moeller. "Computer-aided Medical Diagnosis: Literature Review", *International Journal of Bio-medical Computing* 10.4 (1979): 267-289.
- 2) Doi, Kunio. "Computer-aided Diagnosis in Medical Imaging: Historical Review, Current Status and Future Potential", *Computerized Medical Imaging and Graphics* 31.4-5 (2007): 198-211.
- 3) Emmanuel, Babatunde S. "A Review of Signal Processing Techniques for Heart Sound Analysis in Clinical Diagnosis", *Journal of Medical Engineering and Technology* 36.6 (2012): 303-307.
- 4) Chua, Kuang Chua, et al. "Application Of Higher Order Statistics/Spectra In Biomedical Signal - A Review", *Medical Engineering and Physics* 32.7 (2010): 679-689.
- 5) Y. Ding and H. Sato, "Derepo: A Distributed Privacy-Preserving Data Repository with Decentralized Access Control for Smart Health," 2020 7th IEEE International Conference on Cyber Security and Cloud Computing (CSCloud)/2020 6th IEEE International Conference on Edge Computing and Scalable Cloud (EdgeCom), New York, NY, USA, 2020, pp. 29-35
- 6) J. Rodrigues, H. Gamboa, V. Kublanov and A. Dolganov, "Storage of Biomedical Signals: Comparative Review of Formats and Databases," 2019 International Multi-Conference on Engineering, Computer and Information Sciences (SIBIRCON), Novosibirsk, Russia, 2019.
- 7) Klann JG, McCoy AB, Wright A, Wattanasin N, Sittig DF, Murphy SN Health Care Transformation Through Collaboration on Open-Source Informatics Projects: Integrating a Medical Applications Platform, Research Data Repository, and Patient Summarization *Interact J Med Res* 2013;2(1):e11
- 8) Muhammad F Walji, Elsbeth Kalenderian, Paul C Stark, Joel M White, Krishna K Kookal, Dat Phan, Duong Tran, Elmer V Bernstam, Rachel Ramoni, BigMouth: a multi-institutional dental data repository, *Journal of the American Medical Informatics Association*, Volume 21, Issue 6, November 2014, Pages 1136–1140
- 9) Frank, E., Maier, D., Pajula, J., Suvitaival, T., Borgan, F., Butz-Ostendorf, M., . . . Orešič, M. (2018). Platform for systems medicine research and diagnostic applications in psychotic disorders—The METSY project. *European Psychiatry*, 50, 40-46.
- 10) Nguyen, T.D., Raniga, P., Barnes, D.G. et al. Design, implementation and operation of a multimodality research imaging informatics repository. *Health Inf Sci Syst* 3, S6 (2015).
- 11) Erin K. Thayer, Daniel Rathkey, Marissa Fuqua Miller, Ryan Palmer, George C. Mejicano, Martin Pusic, Adina Kalet, Colleen Gillespie and Patricia A. Carney (2016) Applying the institutional review board data repository approach to manage ethical considerations in evaluating and studying medical education, *Medical Education Online*, 21:1
- 12) Pennington, Jeffrey W.1; Ruth, Byron1; Miller, Jeffrey M.1; Peterson, Joy2; Xu, Baichen2; Masino, Aaron J.1; Krantz, Ian3,4; Manganella, Juliana5; Gomes, Tamar5; Stiles, Derek5; Kenna, Margaret5; Hood, Linda J.6; Germiller, John7,8; Crenshaw, E. Bryan III2,8 Perspective on the Development of a Large-Scale Clinical Data Repository for Pediatric Hearing Research, *Ear and Hearing: March/April 2020 - Volume 41 - Issue 2 - p 231-238*

# Design and Development of Non-invasive Automated mucus Removal Device By Acoustic Assisted Therapy

K.Anusha Bhavani  
Biomedical Engineer  
GEM Hospital and Research Institute  
Chennai,India  
anusha.smuna97@gmail.com

L.Sujitha  
Biomedical Engineer  
zifo RnD Solutions, Chennai, India  
sujithalakshmanaperumal@gmail.com

P.Uvaraj  
Senior Biomedical Engineering  
St.Thomas Hospital  
Chennai, India  
uvibme@gmail.com

**Abstract**— Mucus is a complex biological material that lubricate and protects the human lung and serves as a physical barrier against foreign particles. If there is a high mucus secretion, it blocks the airway and causes breathing difficulties that leads to many respiratory problems like cystic fibrosis, asthma and Chronic obstructive pulmonary disease (COPD). At present, an invasive device namely, suction catheter is used for removing excess secretion in bedside patients not for outpatients, and it causes more discomfort to the patients. There are no available devices to diagnose and remove the excess mucus secretion from the lungs at an early stage. Physiotherapists perform manual Chest physiotherapy with bronchial drainage method for treating respiratory problems, which is a standard treatment for mobilization and removal of airway secretion. But they are facing difficulties in removing secretion through this airway clearance technique and have found no device to analyze the presence of mucus. In order to overcome these problems, we are proposing a non-invasive auto mucus removal device by acoustic assisted therapy which will help doctors for diagnosing and removing excess lung secretions in both adult and pediatricians.

**Keywords**— cystic fibrosis, mucus removal, bronchial drainage, non-invasive acoustic assisted therapy

**Abbreviations** : Chronic Obstructive Pulmonary Disease (COPD), Liquid Crystal Display (LCD), Conventional Chest Physiotherapy (CCPT),

## I. INTRODUCTION

Mucus is a slippery aqueous secretion which is a normal secretory product of the lining of the epithelium[13]. It is produced from the cells found in the mucus gland. It contains secreted water, sugar, protein, lipids, minerals, and mucins[13]. In the human respiratory system, it is also known as airway surface liquid that protects the lungs by trapping the foreign agents which enter the respiratory pathways through the nose during normal breathing. Under normal condition, mucus protects the airway by moistening the air passing through the respiratory pathways. Overproduction of mucus occurs when the person undergoes different types of stress conditions including smoking, infection etc[16]. If there is a high mucus secretion in the airway, it obstructs the respiratory tract lumen, limits the airflow, and also accelerates decline in lung function[8]. High mucus secretion is a symptom for

common illness such as cold and influenza. All major respiratory problems such as COPD, Asthma and bronchitis also include hyper airway mucus secretion[16]. It is very important to evaluate and manage airway mucus hyper secretion for the patients with severe respiratory problems. In Shen Y, Huang S and Kang J, Jan 2018, for airway inflammatory problems, drug and non-drug therapies were used to manage airway mucus hyper secretion [1][15]. In Adam Rao, Jorge Ruiz, Chen Bao and Shuvo Rov, 2017 controlled audio signal was emitted into the chest using an acoustic device and the changes in the sound transmission for air accumulation in the chest was determined [2][6].

Common reason for abnormal breath sound are asthma, chronic bronchitis, chronic obstructive pulmonary disease are classified by wheeze, rhonchi, crackle sounds[12][20]. Wheeze and rhonchus are continuous sound, wheezes are high pitched and ronchi are low pitched with long duration 80-250ms[12].

There are some devices that are available in the market to remove excess mucus from the respiratory patients [3][11].

**Flutter:** The flutter device is a portable device that combines positive Expiratory pressure therapy with high-frequency oscillations within the airway [4](Fig. 1).



**Fig. 1. Flutter**

**Aerobika:** The Aerobika device is a hand-held device, which is easy to use and drug free (Fig. 2). When the patient exhales through the device, positive pressure and oscillations are created simultaneously due to intermittent resistance[3].



**Fig. 2. Aerobika**

**Acapella:** The Acapella is a hand-held airway clearance device (Fig.3) that works on the same principle as the Flutter. The Acapella device available in three models, a low flow (<15 L/min), high flow (>15 L/min) and the Acapella Choice. In the Acapella[3].



**Fig. 3. Acapella**

All these devices follow same procedures. Patients are asked to cough continuously for about 20 minutes to force the excess mucus out[3][4]. Hence patients suffer a lot and face more discomforts while doing this procedure.

As currently available methods to remove excess mucus have disadvantages such as consumes more time, causes discomforts and manual method, physicians requiring an easy way of approach to overcome those problems. Hence, we found an advantageous method by studying all available methods which will help physicians to provide treatment easily without any discomfort to the patient.

## II. METHODOLOGY

Our work is about to design a non-invasive auto-mucus removal device by acoustic assisted therapy that can be used to diagnose the excessive fluid secretions in the airway and to provide vibration therapy in order to remove those secretions automatically.

By detecting the sound level and intensity of sound in airways by placing microphone acoustic sensor over the chest wall/trachea or on back of the chest wall[7][9]. In Sandra Reichert, Raymond Gass this abnormal sound level is differentiated into continuous sound signals classified into wheezes, rhonchus, crackles which is recorded using auscultatory method. Amplitude of sound determines the lung parenchyma and airways [7][21]. Sound classification and their frequency level.

In Sovijarvi AR, paper it is explained that

- 1) upper anterior chest wall used to detect the normal breath sound.

- 2) Abnormal breath sound detected at the posterior chest wall, containing higher-frequency components and a higher intensity than that of normal breath sounds at the same location [22]. Hence the presence of excessive fluid is diagnosed.

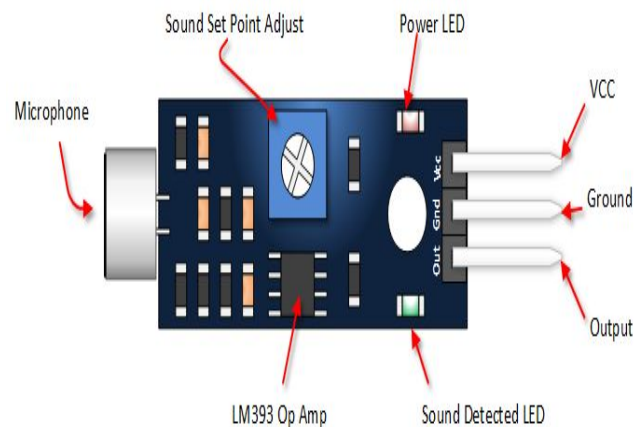
Corresponding to the diagnostic results, vibration therapy is given as alternative method for chest physiotherapy (CPT) which is given by physiotherapist for patients with high mucus.

**Conventional Chest Physiotherapy (CCPT) :** Is done by percussion method by placing the hand in cup shape and clapping the chest wall for about 20-40 minutes manually with same equal force to vibrate the chest muscles that loosen the thick mucus present in airways for lungs, so that it can be easily coughed out by patient. Force given to patient varies by direct contact or by a layer of cloth [19]. This percussion should be done on upper lobes, upper back lobe, left side front lobe and right side lobes, lower back lobes, it should not be done on breast part that may not be effective.

**Vibration Unit :** So we can design the vibration unit in such a way automatically by placing the paddles in specified position to cause vibration to muscles, which is also called as percussion technique to loosen up the mucus present in airway. If sound level and intensity of sound in airways is high, vibration therapy will be done at particular speed for a particular period (according to the setup) which helps to remove the excess mucus secretion. Speed range will be dependent of mucus range and time period of vibration[17], determination of speed and its respective time will be done as in future studies to come up with the product after complete research.

### A. Sensor Description

Microphone Acoustic Sensor also called as electret sensor is used to sense the airway sounds in order to detect the presence of excess mucus in the airway that causes respiratory problems. This sensor works on auscultatory principle used to record the sound waves in frequency range, to observe heart and lung sounds. Sensor indications and pin configuration has been described below ( Fig.4).



**Fig.4. Microphone Acoustic Sensor**

The sensor is connected to the Arduino as mentioned in the sensor pin configuration table ( Table 1 ).

Parameter	Value
VCC	5 VDC from Arduino
GROUND	GND from Arduino
OUT	Connect to Analog input pin
Power LED	Illuminates when power is applied
Sound Detection LED	Illuminates when sound is detected
Sound Set Point Adjust	CW= More Sensitive CCW= Less Sensitive

**Table 1. Sensor Pin Configuration**

It detects the sound level and intensity of sound in airway. After processing the sensor data, Arduino displays the sensor value in the LCD as

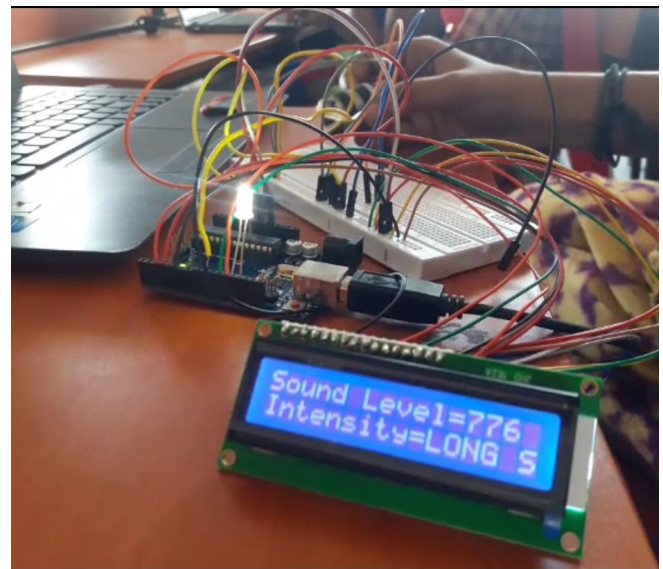
Sound level = 375  
Intensity = medium

#### B. Design

To design a non-invasive auto-mucus removal device by acoustic assisted therapy, we consider two main units: Diagnostic unit and Therapeutic unit.

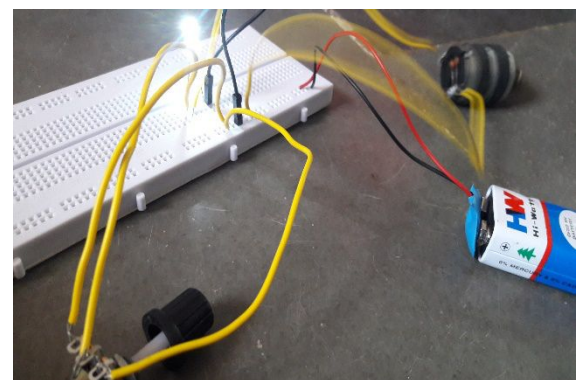
**Diagnostic unit:** This unit controls the diagnostic system. It diagnoses the presence of airway mucus secretion in patients by the use of microphone acoustic sensor which is placed over the chest wall/trachea or on back of the chest so we can clearly hear the abnormal sounds [21]. While breathing, the sensor senses the airway sound and produces corresponding sound value and intensity of sound in the airway. The values that are acquired from the sensor are processed by the Arduino and displays the processed value in the Liquid crystal display LCD that is connected to the Arduino. We have fixed sample values as wheezing and rhoncus sound has dominant frequency above 400hz as referred from [21]. So we have coded 300 to 800hz as with long time duration will be indicated as high and long signal also also LED glows as a alarm indicator. From these diagnosing values, mucus level that is present in the airway can be detected . Observed value is displaced in ( Fig. 5).

Sound level = 776  
Intensity = long signal



**Fig.5. Diagnostic Unit**

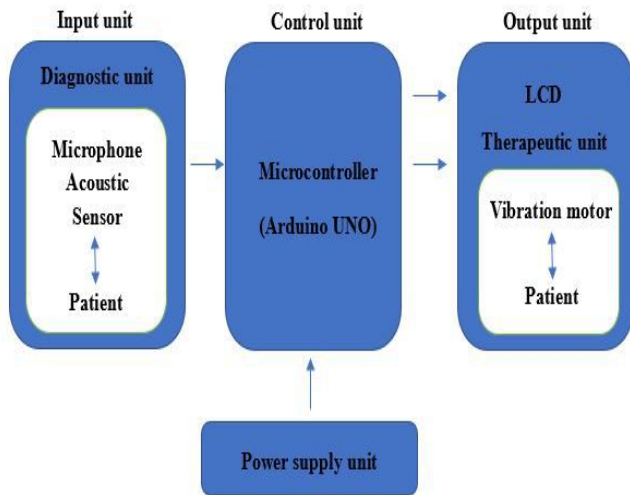
**Therapeutic unit:** When the diagnosing results indicates the excess mucus secretion in the airway, vibration treatment will be provided to the patient automatically using vibration motor to remove excess mucus that narrow the airway. The vibration treatment is provided at particular speed for particular period corresponding to the predefined setup values that are already programmed in the Arduino. Connection of DC motor and its regulator connection is displayed in (Fig .6).



**Fig.6. Therapeutic Unit**

This is designed in such a way to know the condition of a patient (presence of mucus level) accurately and then to provide treatment accordingly. Since the device is automated, there will be no extra force given to the patient in order to remove excess fluids. And it causes no discomfort to the patient during the process. And there is Mean vibration frequency applied to patient fro 10-16hz as in “Investigation of the Frequency and Force of Chest Vibration Performed by Physiotherapists”[19].

### C. Block Diagram



### D. Description

This device consists of input unit, control unit, output unit and power supply unit.

**Sensor Unit :** The microphone acoustic sensor is placed over the chest wall/trachea. Or on back of the chest where lung sound can be heard clearly. Placement of sensor is very important where all the sound signals are recorder to diagnose the abnormalities lung wall. Signal acquired is differentiated as normal and abnormal sound using its frequency range and amplitude[21]. This sound determines the wheeze, rhonchus, crackles on particular milliseconds 80-250ms.

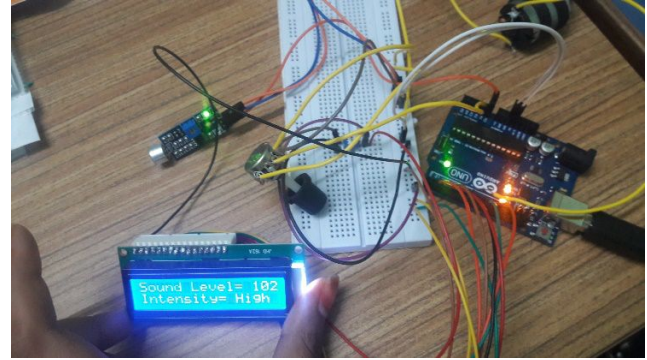
**Control Unit :** The Arduino UNO (control unit) receives the value of microphone acoustic sensor continuously from the input unit. The received data is the input to the program that is inbuilt in the Arduino. so according to the value displayed on the display therapeutic unit is followed. It compares the input value with the predetermined value and sends the condition (sound value and intensity of sound) to the output unit.

**Output unit :** Includes Liquid Crystal Display (LCD) displays the values obtained from the sensor. Normal lungs sound is coded as low signal and frequency range which is defined as abnormal value is displayed as high and there is led indication and displays long signal. Output unit also has led indication when the mucus range is high.

And if the diagnostic result indicates excess mucus secretion, then vibration therapy is provided to the patient corresponding to the predefined values that are already programmed in the Arduino.

**Therapeutic Unit :** vibration unit consist of 3 part as per our current work. Dc motor which creates vibration , regulator for adjusting speed of the motor vibration being delivered to paddle, speed adjustment may vary according to time period the vibration given also by patient

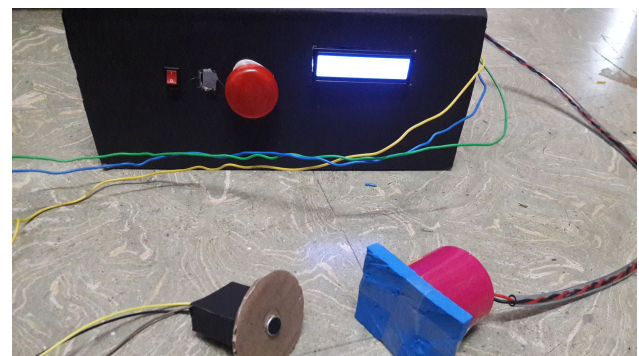
condition[17][19]. And the paddles fixed on the patient chest, which produces vibration to muscles to loosen up the mucus present in airways, mucus in airway causes narrow pathway where inhaling air passe it creates the sound in critical patients,so when these mucus are loosen then will be effective and easy for then to cough out, since this vibration method will be felt as massage to chest area and similar way to ease out the mucus.In Fig.7 the hardware connection of diagnostic sensor and therapeutic unit is shown.



**Fig.7. Hardware Design of Auto mucus removal device**

### III. RESULTS AND CONCLUSION

In this paper, we discussed our work entitled “Design and development of non-invasive automated mucus removal device by acoustic assisted therapy”. We have designed and developed a prototype from our idea using acoustic principle to acquire the minute sound waves from chest wall and displays it, the values differentiated as low, medium,high are completely based on research paper [21]. We have planned to observe and collect data from patient side to determine the exact value for diagnostic unit. And then the Switch can be switched on after viewing patients condition and start artificial chest physiotherapy (CPT )treatment[17]. Still there is lot more future studies to be done to develop this device which combines both diagnostic and therapeutic unit. This will diagnose the mucus level and provides treatment according to the diagnostic results and helps to remove excessive fluid from the airways . Prototype unit designed and developed from our idea in (Fig.8)



**Fig.8 : Idea implemented as prototype unit**

As this device performs both diagnostic and therapeutic procedures, it will be more helpful for physiotherapists as well as physicians to get rid of excessive secretions from the patient easily without any discomforts.

#### IV. FUTURE WORK

The next stage of our work is to develop this prototype including other lung parameters which determine the respiratory diseases, parameters like respiratory rate, pulse oximeter values to be brought into the single diagnostic unit in addition with acoustic signals to bring into pulmonology department monitor. Our future research includes force and frequency that need to be applied for paddles depending upon patient condition, since for percussion method frequency range is from 10hz to 16hz [19].

The sound signals and its data need to be accurately monitor by eliminating the vocal noises using filters to design as a device for commercial use which would be more useful for physicians/ pulmonologists to detect excess secretions and to provide corresponding treatment (Vibration therapy) vibration unit with selected nature of paddles for user comfortable and which will help them to remove those secretions automatically by reducing discomfort to the patients.

#### REFERENCES

- [1] Shen, Yongchun, et al. "Management of airway mucus hypersecretion in chronic airway inflammatory disease: Chinese expert consensus (English edition)." *International journal of chronic obstructive pulmonary disease* 13 (2018): 399.
- [2] Rao, Adam, et al. "Tabla: An acoustic device designed for low cost pneumonia detection." *2017 IEEE Healthcare Innovations and Point of Care Technologies (HI-POCT)*. IEEE, 2017.
- [3] Hristara-Papadopoulou, A., et al. "Current devices of respiratory physiotherapy." *Hippokratia* 12.4 (2008): 211.
- [4] Gastaldi, A. C. "Flutter Device Review: Effects on Secretion and Pulmonary Function." *J Nov Physiother* 6.292 (2016): 2.
- [5] Santos, Mary D., et al. "Pressures and oscillation frequencies generated by bubble-positive expiratory pressure devices." *Respiratory care* 62.4 (2017): 444-450.
- [6] Oletic, Dinko, Bruno Arsenali, and Vedran Bilas. "Low-power wearable respiratory sound sensing." *Sensors* 14.4 (2014): 6535-6566.
- [7] Li, Shih-Hong, et al. "Design of wearable breathing sound monitoring system for real-time wheeze detection." *Sensors* 17.1 (2017): 171.
- [8] Breuer, Thomas, et al. "Acceleration sensors in abdominal wall position as a non-invasive approach to detect early breathing alterations induced by intolerance of increased airway resistance." *Journal of cardiothoracic surgery* 12.1 (2017): 96.
- [9] Wodicka, George R., et al. "Measurement of respiratory acoustic signals: effect of microphone air cavity depth." *Chest* 106.4 (1994): 1140-1144.
- [10] Fu, Xingming, et al. "A novel sound sensor and its package used in lung sound diagnosis." *2014 IEEE 64th Electronic Components and Technology Conference (ECTC)*. IEEE, 2014.
- [11] Marks, John H. "Airway clearance devices in cystic fibrosis." *Paediatric respiratory reviews* 8.1 (2007): 17-23.
- [12] Sarkar M, Madabhavi I, Niranjana N, Dogra M, "Auscultation of the respiratory system", *Ann Thoracic Med* Jul- Sep 2015; 10(3):158-68. doi:10.4103/1817-1737.16083
- [13] Girod, S., et al. "Role of the physicochemical properties of mucus in the protection of the respiratory epithelium." *European Respiratory Journal* 5.4 (1992): 477-487.
- [14] Randell, Scott H., and Richard C. Boucher. "Effective mucus clearance is essential for respiratory health." *American journal of respiratory cell and molecular biology* 35.1 (2006): 20-28.
- [15] Rogers, Duncan F. "Mucoactive agents for airway mucus hypersecretory diseases." *Respiratory Care* 52.9 (2007): 1176-1197.
- [16] Rubin, Bruce K. "Secretion properties, clearance, and therapy in airway disease." *Translational respiratory medicine* 2.1 (2014): 6.
- [17] Sutton pp, Lopez-Vidriero, pavia D "Assessment of percussion, vibration-shaking and breathing exercises in chest physiotherapy, *Eur J Resp Dis*. 1985; 66: 147-152
- [18] Hugh A. *Physiotherapy in respiratory care: an evidence based approach to respiratory and cardiac management*. 3rd. Cheltenham, UK: Nelson Thornes Ltd.; 2001
- [19] S.K. Liand Y.R. Silva "Investigation of the Frequency and Force of Chest Vibration Performed by Physiotherapists" *Physiother Can*. 2008 Fall; 60(4): 341-348.
- [20] Meslier N, Charbonneau G, Racineux JL. "Wheezes" *Eur Respir J*. 1995 Nov;8(11):1942-8. doi: 10.1183/09031936.95.08111942.PMID: 8620967
- [21] Sandra Reichert, Raymond Gass "Analysis of Respiratory Sounds: State of the Art" 16 May 2008, *Clin Med Circ Respirat Pulm Med*.
- [22] Sovijarvi AR, Dalmasso F, Vanderschoot J, Malmberg LP, Righini G, Stoneman SA. Definition of terms for applications of respiratory sounds. *Eur. Respir. Rev.* 2000;10:597-610. [[Google Scholar](#)]

# NOVEL DESIGN APPROACH FOR TREATING SINUS PAIN AT ACUPRESSURE POINTS

Dr. Priya L\*

Associate Professor, Department of  
Electronics and Communication  
Engineering -Biomedical Engineering,  
Vignan's Foundation for  
Science,Technology and Research,  
Tenali,Andhra Pradesh,India,  
priya.god81@gmail.com

Aarthi S

Student, Department of Biomedical  
Engineering  
PSG College of Technology  
Coimbatore,India  
aarthi1296@gmail.com

Preethi M

Student, Department of Biomedical  
Engineering  
PSG College of Technology  
Coimbatore,India  
preethisivaswaamy@gmail.com

Lakshmi Prasanna K

Research scholar , Department of  
Electronics and Communication  
Engineering -Biomedical Engineering,  
Vignan's Foundation for  
Science,Technology and Research,  
Tenali,Andhra Pradesh,India,  
prasannalathakothala@gmail.com

Monisha T

Student, Department of Biomedical  
Engineering  
PSG College of Technology  
Coimbatore,India  
cool.monidude1996@gmail.com

**Abstract--Sinusitis is an inflammation caused in the nasal passage due to cold, allergies or by nasal polyps. Sinusitis is a major Problem among one in eight people around the world. People are interested more in ailment of natural treatments than antibiotic and other medicines. The objective of this work is to provide relief from sinusitis pain through acupressure points. The proposed methodology of sinusitis treatment can be done by applying heating cooling effect and massaging action simultaneously. This paper deals with both the methods of treating the sinusitis. This system integrates an android mobile application with a hardware device that serves this purpose. A face mask which consists of motors and heating and cooling system embedded which can be controlled using user friendly mobile application. The similar methodology can be used for other applications like back pain, muscular pain and knee pain relief.**

**Keywords— Sinusitis, Acupressure, Natural treatment, massaging, pain relief and android mobile application.**

## I. INTRODUCTION

Sinusitis is an inflammation of the nasal passage tissue lining. Healthy sinuses are filled with air. But when they become blocked and filled with fluid, germs can grow and cause an infection. About 35 million Americans are affected by sinusitis at least once each year [1]. It's more likely if the person have swelling inside the nose like from a common cold, blocked drainage ducts and nasal polyps. The sinuses resemble packets or valleys. The passages that begin with nostril lead to a network of nasal passages and cavities in behind the face. These are known as sinuses. Technically as the paranasal sinuses are those in our face, most well-known. The sinuses are pockets of air, which are found in between the bones in face and skull. The paranasal sinuses

are located on either side of the face. The four pair of sinuses shown in Fig. 1(a) Maxillary sinuses are found directly in behind the neck. Ethmoid sinuses are smaller pockets of air, often described as having the appearance of honeycomb. Frontal Sinuses are the uppermost sinuses and are directly behind the fore head. Sphenoid sinuses are further back, close to the brain.

When either the upper or lower sinuses are congested pressure can build causing referred pain. There are several ways to treat sinusitis like Nasal Decongestants that usually come in spray or pill form that have chemicals that are usually quite effective at drying the sinus, one can also breath in hot air or use heat pads on your nose and face to loosen mucus up or Antihistamines that are lifesavers if the sinuses are triggered by allergies. But treating the sinus at pressure points will give more relief.

Based on Ancient Chinese Medicine theory, acupoints treated by special stimuli. Each acupressure point has its own specific therapeutic functions [2]. There are a number of sinus pressure points and pressure points relating to sinus conditions. These pressure points are found all over the body, including, the face, head, arms, hands and neck [3].

Sinus pressure points can be used to treat a number of sinus problems, from simple conditions, such as a runny nose or slight discomfort, to more serious sinus conditions, such as rhinitis, sinus pressure, sinus headaches, and blockages. Sinus pressure points are more reliable for relieving sinus pain than other medications [3]. A hot water bottle, wet compresses, or an electric heating pad applied over the inflamed area also can be comforting for treating sinus. The combination of these two procedures isn't brought together due to their complications in designing. There are six acupressure points for sinus in face. Massaging at those acupressure points will reduce pain from that area. Fig. 1(b) shows the acupressure points for sinus. To relieve from upper or frontal sinuses the pressure should be given at the point A which is located at the bridge of the nose where it meets the ridge of the eyebrows. To open the maxillary sinuses give pressure to the points B and C which are just below to the cheek bones [4-7]. The Pressure applied at these points should be gentle.

---

*Corresponding author: Dr.L.Priya\*, Associate Professor, Department of Electronics and Communication Engineering - Biomedical Engineering, Vignan's Foundation for Science, Technology and Research, Tenali, Andhra Pradesh, India.*

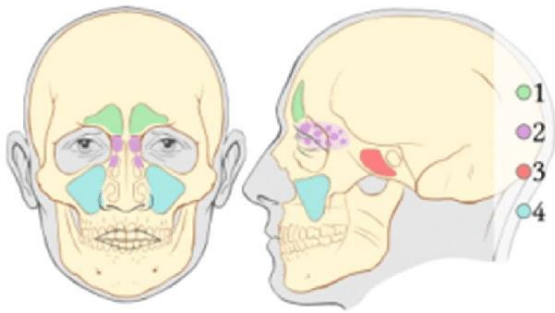


Fig. 1(a). Four pair of sinuses

(Source:researchgate.net/figure/Four-pairs-of-sinuses-1-Frontal-Sinuses-A-small-percentage-dont-have-frontal-sinuses\_fig14\_338777670)

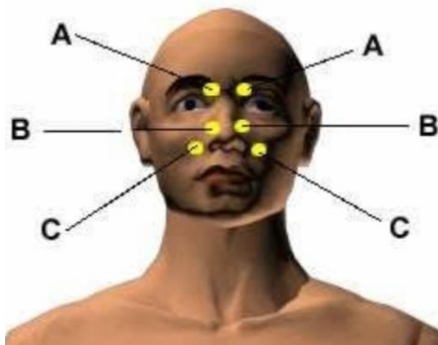


Fig. 1(b). Acupressure points for sinus

(Source: [https://www.researchgate.net/figure/Acupressure-points-for-sinus\\_fig15\\_338777670](https://www.researchgate.net/figure/Acupressure-points-for-sinus_fig15_338777670))

II. METHODOLOGY

The main objective of this work is to develop a device by which the patient can provide self-treatment for sinus pain. Point of care is at home. The patient themselves can relief from pain with the help of this device along with alternate heating and cooling at home.

A face mask has to be designed that will contain the vibrational motor setup peltier chip for pain relief at six acupressure points of sinus treatment. The operation of both the systems can be controlled with help of an android mobile application. The mobile application is designed through Massachusetts Institute of Technology (MIT) app inventor, an easy module for creating android applications.

Existing technology involves the non-invasive method of relieving the sinus pain on the acupressure points using vibration effect. Vibrational motor is used to provide the vibration and the arduino (Atmega 328) is used to control the frequency of the vibration and it can be monitored by the pc[8].

This methodology is to provide a massaging action and heating- cooling effect at the acupressure points of sinus with the help of vibration motor and peltier module respectively shown in Fig. 2 The combination of two technologies was used to provide relief from pain of sinusitis. To provide the massaging action, vibration motors were used at six acupressure points of sinuses and to provide the heating-cooling effect peltier module was used [9-11]. Both motors

as well as peltier module were controlled with the help of a microcontroller ATmega338P. The patient can switch on the Bluetooth via the mobile application in the smart phone. The patient selects the required pressure level and heating and cooling level and that is sent to the microcontroller via bluetooth.

The microcontroller controls the motors and the peltier chip as per the inputs received from the patient’s smart phone. The effects are applied to the patient’s face by the mask. Fig. 3 shows the simple workflow of the device. This device is user- friendly that the user can themselves specify the speed of rotation of vibration motor and also the heating or cooling. This can be controlled with the help of the mobile application which sends the data either “high” or “low” or “medium based upon the massaging level which controls the speed of the vibration motor via Bluetooth. Similarly, the user can specify the heating or cooling via Bluetooth. The voltage level of the vibration motor can be used to control the speed of the motor. When it is “Low” from Bluetooth, the voltage is 0.5 V. When it is “Medium”, the motors runs at 1.75 V and when “High” the voltage is 3V.

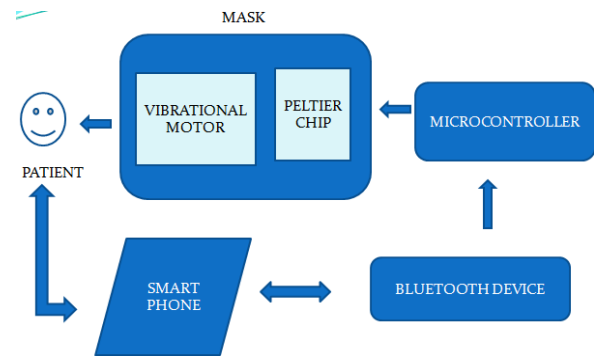


Fig. 2. Block Diagram

III. COMPONENTS AND SPECIFICATION

Linear resonant actuator (LRA) contains a small internal magnetic mass attached to a spring [8]. This mass will create a force when the motor is driven which is then converted into a form of vibrations. The LRA motor is shown in Fig. 3(a). In this work, LRA was used to provide the massaging action at the acupressure points. The specifications of LRA are a Coin Flat Type, 10x3 mm size, voltage supply of 3V, current of 70mA and 9000+/- 3000 rpm (revolution per minute) speed.



Fig. 3(a). Linear Resonant Actuator



Peltier module is a thermoelectric module which is used to produce alternate heating and cooling effect. This module uses the basic principle of Peltier effect. The Peltier effect is a temperature difference created by applying a voltage between two electrodes connected to a sample of semiconductor material. This is useful for transforming heat from one medium to another medium. This effect is named after the Physicist Jean- Charles Athanase Peltier who discovered it. The heat will flow from one side of the module through another when low voltage DC Power is applied to the Thermoelectric (TE) Module. The one face of the module will be heated and the other face will be simultaneously cooled. This can be reversed by changing the polarity of the applied DC voltage[12].

The basic thermoelectric module consists of n and p-type doped semiconductor material that is connected electrically in series and thermally in parallel. The arrangement of the semiconductor materials in the TE module causes the heat to flow through the TE materials. The electrons in the N- material and the holes in the P-material are acting like carriers to move the heat energy. For this work, TEC1- 13706 Peltier module was used to provide alternate heating and cooling effect. TEC1-13706 is of size 40 mm x 40 mm x 4 mm and gives a voltage of 0-13 V at the current rating of 0-6 A. The operating temperature of the Peltier module is 243 to 343 K (-30 to 70°C). Fig. 3(b) shows the TEC- 14706 peltier chip [13].

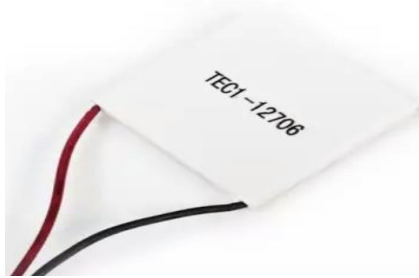


Fig. 3(b). Peltier Chip

Arduino Uno R3 Atmega 328p is used to control the vibration motor and peltier module. The high- performance Atmel pico power 8- bit AVR microcontroller was used. This microcontroller is basically RISC based architecture combines with 32 KB Flash Memory. Even though the power was switched off, the code will be there inside the memory. It contains 1024B EEPROM, 2KB SRAM, 33 GPIO Pins, 32 Registers, along with Timers, Counters, ADC's etc. This microcontroller also contains PWM Pins. The device operates between 1.8- 5.5 volts.

The ULN2003 is a driver IC used to drive both the vibration motor and the peltier module. This Driver IC consists of an array of Transistors which will act like a switch to on and off the motors and TE modules at an input voltage of 30 V and peak collector current of 500mA.

Bluetooth HC- 05 module is a Serial Port Protocol module used for a transparent wireless serial communication

application. It uses 3.5 GHz frequency band. The connection can be point- to- point or multipoint where the maximum range of acceptability is 10 meters. The data transfer rate is 1Mbps. The operating temperature of this module is around 303 K (30°C). It needs 3.3 V DC power supply for operation.

#### IV. RESULTS AND DISCUSSIONS

The user interface of the Android application is shown in the Figure.4. The mobile application was user friendly. The user can himself select the vibration levels with the help of the android application. The application contains buttons for selecting the vibration level as low or medium or high. This command will be sent via the mobile Bluetooth to the Bluetooth module attached to the prototype. This command will alter the level of voltage supplied to the motor with the help of microcontroller. The voltage level for the respective commands was shown in the Table. I.

Table I. Voltage level of vibration motor

Bluetooth data	Voltage level
Low	0.5V
Medium	1.75V
High	3V

The prototype model is shown in the Figure.6. The Prototype model consists of the six vibration motors and the peltier module which were fixed inside a face mask at the respective points. The motors as well as the module were controlled by the microcontroller with user inputs from mobile app. The vibration motor worked effectively with required pressure at regular intervals until the patient himself feels relieved. Before the application of massaging action, heating and cooling given by the peltier module was helpful in aggregating the pressure points for better results. It also helped the patients feel better.

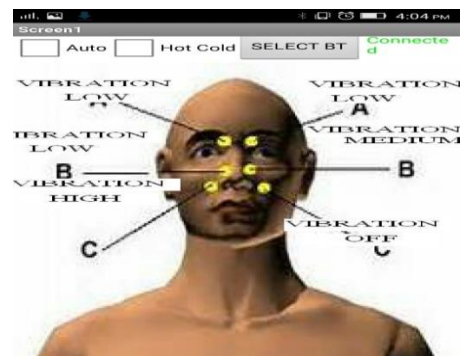


Fig. 4. Android Application User Interface

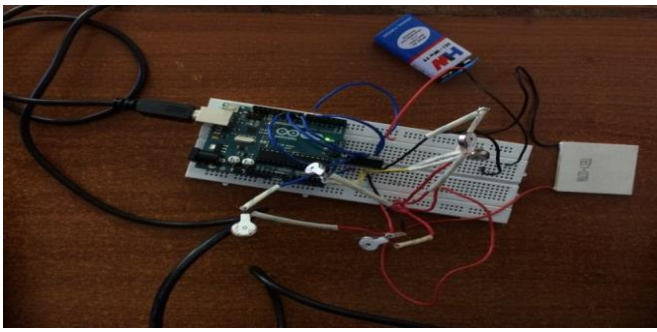


Fig. 5. Prototype model

The voltage levels and its corresponding rpm values were shown in Table. II. These rpm values can be fixed based upon the doctor's advice. The temperature of the peltier module was fixed and is denoted in the Table. III. The temperature was fixed at + 278 K (+5°C) from normal body temperature for heating and subtract 283 K (10°C) from normal body temperature for cooling.

TABLE II. Voltage level and speed of vibration motor

Bluetooth data	Voltage level	Speed (rpm)
Low	0.5V	1500
Medium	1.75V	5250
High	3V	9000

TABLE III. Temperature level of the Peltier module

Heating/ Cooling	Temperature in K (°C)	Time Period in seconds
Heating	Body temperature+ 278 (5)	10
Cooling	Body temperature-283 (10)	20

#### A. Readings and findings

PSG Institute of Medical science and Research (PSGIMSR) Human Ethics Committee has reviewed and approved the research study entitled "Sinusitis pain relief system". The clinical trial was conducted at PSG Hospitals, Coimbatore for a period of one year.

1. **Objectives:** To determine sinusitis pain relief system can improve outcomes in patients with sinusitis pain.
2. **Setting:** one general practice in PSG Institute of Medical Sciences and Research.
3. **Design:** Factorial randomized controlled trial of two year duration.

Patients suffering from Sinusitis of ages between 6 and 60 were chosen. They were grouped into 2 categories based upon the cause of sinusitis. Patient suffering from sinusitis due to heat were grouped into 'A' and patient suffering from sinusitis due to cold were grouped into 'B'. The Patients were given a pain scale and asked to choose the face which is indicated in the scale depends upon the severity of the pain.

A pain scale measures patient's pain intensity. The Wong-Baker faces pain scale shown in the Figure.6 was used as a self-report of pain to assess a patient's experience of pain. It can be used in children aged between 3 and 18 years of age and also for other patients above this age. The patient has to understand how much pain they have, and how this makes them feel[14-15].

Based upon the pain severity, they have to choose the faces. Face 0 is very happy because he doesn't hurt at all (i.e. has no pain). Face 2 hurts just a little bit. Face 4 hurts a little more. Face 6 hurts even more. Face 8 hurts a whole lot. Face 10 hurts as much and the patient may cry at this level. Then the patient was asked to worn the mask and to set the vibration level by himself. The patient can change the level as well as choose the motor which has to be run at that time. At the same time heating and cooling was given in the specified intervals as in the Table.3.

After this treatment, both the group of patients were given again the pain scale and asked to choose the face depends upon how they feel. The patient who chose Face 8 in the first trial now chose Face 2. Thus, the pain level had been reduced. The alternate application of heating and cooling also provides us the best results in both the groups because for 'A' group of patients cooling will help to project the pressure points outside and for 'B' group of patients heating will help to project the pressure points outside.

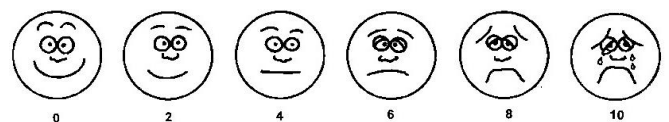


Fig. 6. Wong-Baker faces pain scale

The pain level observations for the Group A and Group B patients as shown in Table. IV and Table. V

TABLE IV. Pain level before and after treatment with the device for Group A Patients

Group A Patients	Pain level before treatment	Pain level after treatment
1	6	2
2	8	2
3	8	4
4	4	0
5	10	4
6	6	0
7	10	2
8	6	4
9	6	0
10	8	2

TABLE V. Pain level before and after treatment with the device for Group B Patients

Group B Patients	Pain level before treatment	Pain level after treatment
1	8	0
2	6	2
3	10	4
4	10	6
5	8	0
6	8	2
7	8	0
8	4	0
9	2	0
10	6	4

## V. CONCLUSION

Thus, the Sinusitis Pain Relief System was developed along with the android application. The device provides better pain relief to patients those who are suffering from sinusitis. This device is also suitable for other body pains such as knee pain, muscle pain and back pain. This device provides point of care treatment and hence, it is very reliable and efficient.

## REFERENCES

- [1] R. Pothman, "The effects of treatment with antibiotics, laser and acupressure upon chronic maxillary sinusitis in children," *Int. J. Sci. Res.*, vol. 22(2), pp. 241 -264, 1982.
- [2] Ying Cheng, BS, Bifeng Gao, MD, "Acupuncture for common cold," *Medicine.* vol. 97(10), March 2018.
- [3] Gelfand and L. Jonathan, "Help for Sinus Pain and Pressure," in *WebMD*, pp. 120-35. 2011.
- [4] M.Piyush, D. Vishwas, K. Shivajirao and D. Vividha "Contemporary acupressure therapy: Adroit cure for painless recovery of therapeutic ailments," *eJTCM*, vol. 7 (2), pp. 251-263 , April 2017.
- [5] Leung RS and Katial R, "The diagnosis and management of acute and chronic sinusitis," *Prim Care, Elsevier*, vol. 35(1), pp. 11-24, 2008.
- [6] Gwaltney JM, Hendley JO, Phillips CD, Bass CR, Mygind N and B. Winther , "Nose blowing propels nasal fluid into the paranasal sinuses", *Clin Infect Dis*, vol. 30(2), pp. 387-91, 2000.
- [7] E. Rössberg, "Comparison of traditional chinese acupressure, minimal acupressure at non-acupoints and conventional treatment for chronic sinusitis", *Addison-Wesley*, pp. 285-315, 1985.
- [8] C. Jim Elliot, S. Anitha and T. S. Nithya "Non Invasive Treatment to Relieve the Pain due to Headaches & Sinus using Vibration on Acupressure Points," *IJEAT*, vol.9(154), December 2019.
- [9] Pletcher, "Use of Acupressure in the Treatment of Sinus and Nasal Symptoms", in *Results of A Practitioner Survey*, vol. 5(2), pp. 786 - 89, 2006.
- [10] Yuval avni, "Vibrating Device for Treating Nasal Congestion and Sinusitis Symptoms and Method Thereof," *Patent WO2006129305A3*, 2008.
- [11] Harry A and Romano, "Self contained heating or cooling suit," in *Pearson Education International 2010*, pp. 112-39.
- [12] *Micro DC Motors: Need For Power Motor*, 2010, Available at: [www.nfpmotor.com](http://www.nfpmotor.com). Accessed Aug 5,2015.
- [13] Peltier, 2014, Available at: [www.hebeiltd.com.cn/peltier.datasheet](http://www.hebeiltd.com.cn/peltier.datasheet). Accessed Sep 16, 2016.
- [14] Ya-Wen Chen and Hsiu-Hung Wang "the effectiveness of acupressure on relieving pain: a systematic review," *Pain Manag Nurs*, vol. 15(2), pp. 539-50, June 2014.
- [15] Sedighe Khalili and Shomia, Mostafa Nasiri, "Comparison of the effect of two methods of acupressure and ice massage on the Hugo point on the severity of pain after appendectomy surgery," *IAMJ*, vol. 10 ( 2), April-June 2019 .

# Intelligent Walking Stick with Static and Dynamic Obstacle Detection Technology for Visually Challenged People

Krishnakumar. S\*

Department of Biomedical Engineering,  
School of Bio and Chemical  
Engineering line 3: Sathyabama  
Institute of Science and Technology,  
Chennai, India  
[drkrishnakumar\\_phd@yahoo.com](mailto:drkrishnakumar_phd@yahoo.com)

Bethanney Janney J

Department of Biomedical Engineering,  
School of Bio and Chemical  
Engineering line 3: Sathyabama  
Institute of Science and Technology,  
Chennai, India  
[jannydoll@gmail.com](mailto:jannydoll@gmail.com)

Umashankar. G

Department of Biomedical Engineering,  
School of Bio and Chemical  
Engineering line 3: Sathyabama  
Institute of Science and Technology,  
Chennai, India.  
[umashankar.bme@gmail.com](mailto:umashankar.bme@gmail.com)

Lumen Christy. V

Department of Biomedical Engineering,  
School of Bio and Chemical  
Engineering line 3: Sathyabama  
Institute of Science and Technology,  
Chennai, India  
[lumenchristyv@gmail.com](mailto:lumenchristyv@gmail.com)

Banani Mridha

Department of Biomedical Engineering,  
School of Bio and Chemical  
Engineering line 3: Sathyabama  
Institute of Science and Technology,  
Chennai, India  
[bananimridha1216@gmail.com](mailto:bananimridha1216@gmail.com)

Prakash Williams. G

Department of Botany and  
Biotechnology,  
Bishop Moore College,  
Mavelikara, India.  
[prakash.gw@gmail.com](mailto:prakash.gw@gmail.com)

**Abstract**— Visually challenged people are facing lots of problems in their day-to-day life. The main aim of the research is to design an intelligent walking stick to guide visually impaired people thereby improving their mobility. In this examination the walking stick is planned with all features of precision and productivity is likewise made high, relatively. The device comprises an array of both ultrasonic sensor and Passive Infrared sensor to detect the static as well as the dynamic obstacle in 180 range of 3m distance in three different directions (left, right, and front). Water sensor is also being attached to the system to detect the presence of both static and dynamic water. When an obstacle is sensed by the sensor, the emergency alert system warns the visually impaired person by a speech warning message. GPS-GSM system is used to track the position of the visually impaired person. On the off chance that the stick is lost, it can be discovered utilizing the remote that actuates the signal to find out the location of the stick. The remote consists of an Rf transmitter and the stick consist of an Rf receiver. Wireless transmission is being carried out between remote and walking stick.

**Keywords**— Walking stick; Ultrasonic sensor; Passive Infrared sensor; GPS-GSM; Rf transmitter; Rf receiver

## I. INTRODUCTION

Visually challenged people face many challenges when they move from one place to another that involves finding the path, detecting and avoiding obstacles in the path, etc. They are unable to do the normal work an average common man does during his daily schedule. They always need tangential assistance to do their daily petty works. Hence, they depend on some assistive device provided to them for the betterment of their life. They require assistive devices to increase navigation and orientation. The regularly utilized assistive gadget by the visually impaired individual is the white stick across the globe [1]. Globally 39 million individuals are tested, out of which 15 million individuals are in India. According to World Health Organization in 2011 evaluations there are 285 billion people on the planet with visual debilitation, 39 billion people are sand-blind and 246 billion are with low vision, and around 15 million people are sand-blind in India [2].

The assistive device technology is a way of supporting them to lead their normal daily activities. The use of an assistive device helps visually challenged people in independent living. A mechanical cane provides information when an obstacle is sensed by the sensors in front of it. There are different sensors used for the detection of the obstacle. Electronic Travel Aids (ETAs) gadgets have been launched to be an assistive technology for visually debilitated individuals. Electronic Travel Aids (ETAs) are gadgets that accumulate the data in and around the ambiance and communicate to the client through sensor cameras, sonar as well as laser scanners [3 & 4]. The fundamental objective of the design is to improve the movement and exactness of visually debilitated individuals. In the present research, an array of ultrasonic (US) sensor and passive infrared (PIR) sensor is being interfaced with Mega Arduino (ATmega2560) for the development of the intelligent guide. This proposed intelligent walking stick can detect static as well as dynamic obstacles. Further to enhance the accuracy of the visually impaired person, the direction of the obstacle is also taken into account i.e., this stick can sense the exact direction in which the obstacle lies. Visually impaired individuals experience tremendous difficulty when they stroll in the city or streets utilizing a regular white cane, however, they have sharp reactivity and great memory power [5]. The present electronic intelligent walking stick helps by providing more convenience to the visually impaired person to lead their normal life.

## II MATERIAL AND METHODS

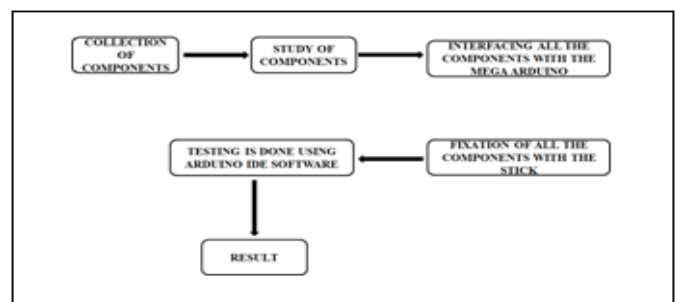


Fig: 1 workflow of the proposed system

### III BLOCK DIAGRAM

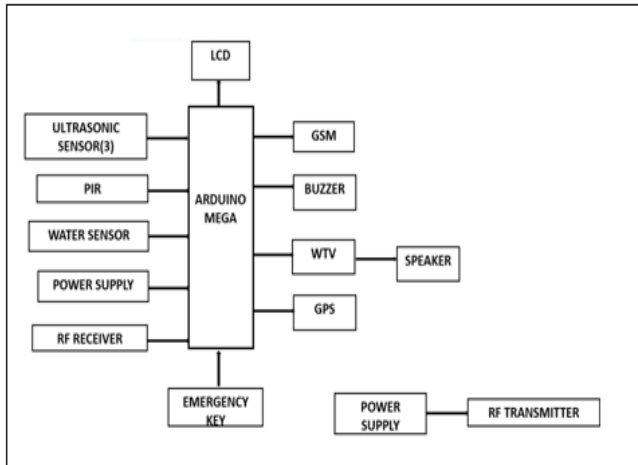


Fig: 2 Block diagram of the proposed walking device

### IV HARDWARE REQUIREMENTS

#### A. Ultrasonic sensor

HC-SR04 ultrasonic sensors are used for obstacle detection and distance measurement. This sensor module contains an ultrasonic transmitter circuit, receiver, and control circuit. This sensor has 0.3 resolutions with a ranging distance of 2 cm to 500cm. It drives from a 5V DC supply and the current is less than 2mA. The sensor is triggered by the 10 $\mu$ s high-level signal. The working mechanism of the ultrasonic sensor is, the point at which the sensor is set off, it sends eight 40 kHz pulse and stays for the reverberation beat. If there is no obstacle present in front of the visually impaired person no echo pulse is received. On the other hand, if there is an obstacle detected by the sensor, then the echo pulse is received in the receiver circuit.

#### B. Passive Infrared (PIR) sensor

A Passive Infrared Sensor (HC-SR501 Pyroelectric Infrared Module KG001) is an electronic sensor that quantifies the infrared light exuding from objects in its field of view. They are often insinuated as "Pyroelectric" or "IR Motion" sensors. The PIR sensor recognizes the impediment by perceiving infrared radiation from objects before it and close to the client.

#### C. Mega Arduino (ATmega2560)

The Mega Arduino 2560 is a microcontroller board considering the ATmega 2560. It comprises 54 computerized input and additionally yield pins, 16 analog load, UARTs, crystal oscillator, USB affiliation, and reset catch. All the parts assist the microcontroller; basically, connected with a PC utilizing a USB connection or power with AC to DC connector or by a battery to start. The Mega Arduino is acceptable with most shields planned for the Arduino Diecimila.

#### D. Water Sensor

Water Sensor is utilized to distinguish any static and dynamic water bodies. The operating mechanism of the water sensor is easy to utilize, smaller, lightweight, cost-effective water droplets recognition, and location sensors. The working standard of the water sensor is to gauge the

size of the trace amount of water beads through the line with a progression of equal wires left uncovered.

#### E. GPS module

Global Positioning System (GPS) is an organization of satellites that impart the signals which help us to discover specific areas. GPS can be utilized to discover explicit areas anyplace over the world. Global Positioning System innovation was initially created for military use; however, it is currently accessible to all people. The proposed gadget is equipped for accepting data from GPS satellites and precisely ascertain its topographical location. It is a 24-satellite route framework that utilizes different satellite signs to discover a receiver's location on the earth's surface. GPS receiver computes its location by a procedure called satellite ranging, which includes estimating the separation between the GPS receiver and the GPS satellites tracking.

#### F. GSM module

Global System for Mobile communications (GSM) is a system broadly utilized for cell phones on the planet. The GSM Association advances its utilization and cases that 80 percent of every single cell phone is utilizing the GSM standard. This implies each of these telephones scans for a cell in the range in which they are being utilized. GSM arrange required to deal with all GSM-based telephone calls.

#### G. RF transmitter (HT12E) and RF receiver (HT12D)

Radiofrequency module (RF module) is a compact electronic gadget that is used to send and receive radio signals between two gadgets. In an embedded framework it is preferable to be in communication with another gadget wirelessly. This may be refined through optical communication or radio frequency (RF) communication framework. The wireless communication system operated at a frequency of 433MHz.

RF transmitter will be with the remote and receiver with the stick. At the point when a switch 'A' in the remote is squeezed, LED 'A' in the stick gleams and the buzzer is turned ON. RF transmitter comprises an encoder, which converts parallel data to sequential data. RF receiver comprises a decoder that changes sequential data over parallel data.

#### H. WTV (APR33A3)

- Operating Voltage Range from 3V ~ 6.5V
- Single-Chip, High-Quality Audio/Voice Recording & Playback
- Solution Voice Recording Length is APR33A3:- 680 sec
- Audio Processor:- Powerful 16-Bits Digital
- Memory:- Non-volatile Flash
- Built-in Audio-Recording Microphone Amplifier
- Resolution :- 16-bits
- Averagely 1, 2, 4, or 8 voice messages record & playback

#### I. PVC pipe

PVC pipe is used to inbuilt all the components are being attached. The diameter of the pipe is 2cm and the height is 3.18 feet.

### V INTERFACING OF THE COMPONENTS

Each component is being interfaced with the mega Arduino for the development of an intelligent walker.

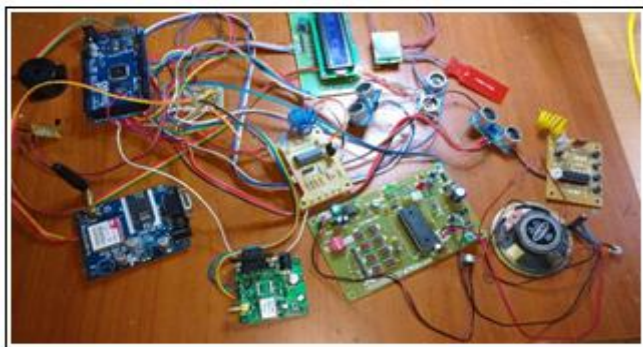


Fig: 3 Interfacing of all components with ATmega2560

### VI ARDUINO IDE SOFTWARE

Arduino IDE software was used. The programme code was written and uploaded to the Arduino board for testing. The software is easy to use. The code is being RUN to perform the necessary working of the proposed system. The principle part of the framework is the Mega Arduino that controls different segments in the proposed framework. At the point when the ultrasonic sensors, PIR sensor, and water sensor recognize any articles or hurdle in 180° way at a reach within 3m, a voice cautioning message is given to the client concerning the direction in which the item lies and the client move forward appropriately.

### VII RESULTS AND DISCUSSION

Nowadays an assistive device is very essential for visually challenged people to lead their normal life without expecting any help from others. The work process and block diagram of the presented walking stick can be seen in fig. 1 and fig. 2 correspondingly. The Intelligent Walking Stick consists of an array of ultrasonic sensor, PIR sensor is interfaced with the Mega Arduino (fig.3). A water sensor is utilized to recognize the presence of water in the way of the visually impeded individual. The water sensor functions fully and can detect water level above 0.5 cm. The voice warning message of the water sensor will stop once the stick is taken out of water. The complete intelligent walking stick is portrayed in fig. 4. The visually impaired person gets information about the presence of human, static as well as dynamic obstacles in front of them with the help of voice commands. Buzzers are used as feedback to the users. An emergency alert system is also being incorporated in case of any emergency. When a moving or static obstacle is detected by the ultrasonic sensor or PIR sensor, a speech warning message is given to a visually impaired person as an alert (Table 1). Three different voice warning messages such as object right (fig. 6) object left (fig. 7) and object front (fig. 8) are given to the user regarding the direction in which the object lies. The moving obstacles also could be detected by using the same sensor (fig. 9). In case of any emergency, if the person presses the emergency button, a message regarding his location is being sent to his/her relative and the location can be tracked using google map. On the off chance that the stick is lost, it can be discovered utilizing a remote that is accessible to the visually debilitated individual.

TABLE I SPEECH WARNING MESSAGE OF INTELLIGENT WALKING STICK

S.No.	Speech warning message	Object detected
1	“object right”	Yes
2	“object left”	Yes
3	“object Infront”	Yes
4	“water detected”	Yes



Fig. 4 complete intelligent walking stick



Fig. 5 Remote

#### A. Ultrasonic Sensor and PIR Sensor



Fig. 6 Object right



Fig. 7 Object left



Fig. 8 Object in front



Fig.9 Moving objects detected

### B. GPS-GSM System

An emergency button is being mounted over the stick in case of any emergency. When this button is pressed, sos message is sent to the already saved 10-digit mobile number. The message will have the latitude and longitude value of that particular location (fig 10). This particular latitude and longitude value can be traced using Google map (fig 11) and the location of the visually impaired person can be tracked by his/her relatives.

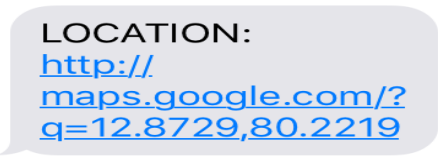


Fig. 10 Value of latitude and longitude

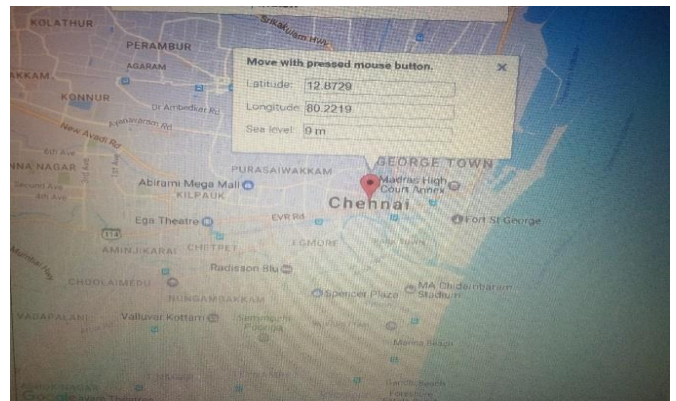


Fig. 11 Tracking the location using Google map

The proposed work incorporates interfacing an array of ultrasonic sensors and a PIR sensor unit with Mega Arduino (ATmega2560) in the advancement of the wise adviser to identify both static and dynamic articles. The ultrasonic sensor is utilized to identify static articles while the PIR sensor is utilized to recognize dynamic items. The presented model is planned with improved versatility and precision for visually debilitated individuals. An Emergency alert system is additionally being introduced by utilizing the GPS-GSM framework. In case of emergency, if the visually impaired person presses the emergency button, a message regarding his location is sent to his/her relative and the location can be tracked using google map by his/her relatives. Further to improve the movement of the visually debilitated individual, the direction of the hindrance is also considered i.e., this stick can tell the specific direction where the obstruction lies by giving a discourse cautioning message. When an obstacle is sensed by the sensor, an emergency speech warning message is given to the visually impaired person. In addition to this, a water sensor is appended to the framework to recognize if there is any presence of water in the way of the visually debilitated individual. In the event that the stick is lost, it can be discovered utilizing the remote that switch on the buzzer and subsequently determine the location of the stick. The remote comprises of Rf transmitter and the stick comprises of Rf receiver and remote transmission is being executed between the remote and the stick.

The walking stick with obstacle detection technology for visually challenged peoples has been developed by different researchers across the world. But still, the visually challenged people face some sort of limitations to lead their life. Rupali Kale et al. [6] recommended that the Global Positioning System (GPS), the discovery of hurdle and item evasion advancements focus to make the visually impeded individual completely liberated in all perspectives. Sabarish et al. [7] have portrayed the evolution of a navigation aid to help visually hindered individuals traverse productively, effortlessly, and firmly by distinguishing any hindrances. The proposed framework works based on the microcontroller with synthetic speech output. It incorporates two vibrators, two ultrasonic sensors which are mounted on the visually weakened client's shoulders, and the other one is coordinated into the stick. Rohit Sheth et al. [8] introduced the force of vibrations as a sign of the closeness of a hurdle in the walking pathway of the client. Four different voice

messages were given for down-step, up-step, an obstacle in front, and obstacle overhead. Mohammad Hazzaz Mahmud et al. [9] mentioned a non-contact distance measurement about the obstacle in about 2cm-3m. Obstacle and hole were determined easily by sensor readings. The framework uses two sets of ultrasonic sensors that identify the hindrances underneath the stick, in front and wet slippery terrain also recognized by similar sensors. To change the vibration design for various scopes of obstructions, these sensors additionally utilize Pulse Width Modulation.

Wahab et al. [10] sketched the evolution of the Smart Cane for the recognition of articles and give precise directions for the traversal of the visually hindered individual. The framework comprises of different sensors to manage the visually debilitated individuals. It utilizes ultrasonic sensors, fuzzy regulator, and servo engines to recognize the hurdles before the visually impeded individual and gives instructions through voice messages. Ayat Nada et al. [11] slogged on the location of flight of stairs, direction, and other hindrances present in the track of the within the scope of 2m. Prasun Shrivastava et al. [12] introduced his work for both indoor and outside use. At whatever point an individual holding the stick has a high heartbeat rate or low blood pressure rate or any troublesome condition, a message will be shipped off the concerned individual utilizing a GPS connected to it. Gayathri et al. [13] created smart walking sticks utilizing different sensors for obstructions, potholes, and dampness detection. The drawback of the stick is that it can't distinguish dynamic snags that are not directly in the scope of the ultrasonic sensor. Elderly and blind people used to navigate easily indoor and outdoor with the help of different sensors mounted on a walking stick [14]. Dada Emmanuel Gbenga et al. [15] suggested the GPS, and GSM modules to find the position of the user to communicate the location of a relative and/or care taker. The walking sticks also accommodate wide range of grips for flexible handling to the users.

#### VIII CONCLUSIONS

The proposed model was developed with great precision so that the visually impeded individual can readily move to any place independently. In the event that such a framework is created, it will go about as a fundamental stage for the generation of more such gadgets for the visually disabled shortly which will be economical. The results acquired were promising to a degree that guarantees the security and speed of the portability of the deprived user. The proposed model helps the user to move from one place to another easily and independently without assistance. GPS and GSM system identify latitude and longitude value which can be traced using Google map and the location of the user could be tracked by relatives. The model can be improved by using VLSI technology, in which a single chip can be made that comprises all the components so that the stick becomes lightweight and is easy to carry.

#### ACKNOWLEDGMENT

The authors are grateful to Sathyabama Institute of Science and Technology, School of Bio and Chemical Engineering, Department of Biomedical Engineering, Chennai, Tamil Nadu, India for furnishing all the required provisions. The fulfillment that goes with the

accomplishment of my mission would be deficient without the participation of the administration of Sathyabama Institute of Science and Technology.

#### REFERENCES

- [1] Parth Dhall , Pankaj Sharma , Shristi Thakur, Rishav Agarwal and Shubham Rastogi, A. REVIEW PAPER ON ASSISTIVE SHOE & CANE FOR VISUALLY IMPAIRED PEOPLE, *International Journal of Scientific Research and Management Studies (IJSRMS)*, 3 (2), 2010, pp: 113-117.
- [2] World Health Organization. Visual Impairment and Visually impaired personness. Available online: <http://www.Awho.int/mediacentre/factsheets/fs282/en/> (accessed on 24 January 2016).
- [3] J. Liu, L. Xu, and W. Jin, Electronic travel aids for the visually impaired person based on sensory substitution, In Proceedings of the 5th International Conference on Computer Science and Education (ICCSE), Hefei, China, 24–27 August 2010.
- [4] J. Sanchez, and M. Elias, Guidelines for designing mobility and orientation software for visually impaired person children, In Proceedings of the IFIP Conference on Human-Computer Interaction, Janeiro, Brazil, 10–14 September 2007.
- [5] Shashank Chaurasia and K.V.N. Kavitha, An electronic walking stick for visually impaired persons. *International Conference on Information Communication and Embedded Systems (ICICES2014)*, 2014 pp 1 - 5, DOI: 10.1109/ICICES.2014.7033988, IEEE Conference Publications.
- [6] Rupali Kale and A.P. Phatale, Design of a GPS based Virtual Eye for the Visually impaired person People, *International Journal of Current Engineering and Technology*. 4 (3), 2014, 2277 – 4106.
- [7] S. Sabarish, Navigation Tool for Visually Challenged using Microcontroller. *International Journal of Engineering and Advanced Technology*, 2(4), 2013, 2249 – 8958.
- [8] Rohit Sheth, Surabhi Rajandekar, Shalaka Laddha and Rahul Chaudhari, Smart White Cane – An Elegant and Economic Walking Aid, *American Journal of Engineering Research*, 3(10), 2014, 84-89.
- [9] Mohammad Hazzaz Mahmud, Rana Saha and Sayemul Islam. Smart walking stick - an electronic approach to assist visually disabled persons, *International Journal of Scientific & Engineering Research*, 4(10), 2013, 2229-5518.
- [10] A. Wahab, M. Helmy, A.A. Talib, H.A. Kadir, A. Johari, A. Noraziah, R.M. Sidek, A.A. Mutalib, Smart Cane: Assistive Cane for Visually-impaired People. *Int. J. Comput. Sci.* 8(4), 2011.
- [11] A. Ayat Nada, A. Mahmoud Fakhr, and F. Ahmed Seddik, Assistive Infrared Sensor Based Smart Stick for Visually impaired person People. Department of Computers and Systems Electronics Research Institute, 2015.



- [12] Prasun shrivastava, Akash Singh, Praveen Anand and Sagar.V. Medico stick an ease to visually impaired person and deaf. IEEE Sponsored 2nd International Conference on Electronics and Communication System, 2015, 978-1-4788-7225-8/15.
- [13] G. Gayathri, M. Vishnupriya, R. Nandhini, and M. Banupriya, SMART WALKING STICK FOR VISUALLY IMPAIRED. *International Journal of Engineering and Computer Science*, 3(3), 2014, 4057-4061.
- [14] Shalini Singh and Balwinder, Singh Intelligent Walking Stick for Elderly and Blind People. *International Journal of Engineering Research & Technology (IJERT)*, 9 (03), 2020, 19-22.
- [15] Dada Emmanuel Gbenga, Arhyel Ibrahim Shani and Adebimpe Lateef Adekunle, Smart Walking Stick for Visually Impaired People Using Ultrasonic Sensors and Arduino. *International Journal of Engineering and Technology (IJET)*, 2017, 9(5), 3435-3447.

# Acoustic Based Separation of Blood Cells Using Microfluidic Device

Caffiyar Mohamed Yousuff  
Department of Electronics and  
Communication Engineering  
C.Abdul Hakeem College of  
Engineering and Technology  
Melvisharam, India  
cmd.yousuff@gmail.com

Pavithra devi. E  
Department of Electronics and  
Communication Engineering  
C.Abdul Hakeem College of  
Engineering and Technology  
Melvisharam, India  
pavithrayadav666@gmail.com

Swetha.C  
Department of Electronics and  
Communication Engineering  
C.Abdul Hakeem College of  
Engineering and Technology  
Melvisharam, India  
swethanaidu1308@gmail.com

Ramya. R  
Department of Electronics and  
Communication Engineering  
C.Abdul Hakeem College of  
Engineering and Technology  
Melvisharam, India  
ramyaramesh2699@gmail.com

**Abstract**— A frequency-dependent simulation of 2D microparticle separation using the acoustophoresis approach in the COMSOL Multiphysics program is presented in this paper. The new microfluidic platform distinguishes itself by having different size separation using density-based separation of cell orientation and positioning using external sheath fluid. In this study, the simulation was created for various microparticles with sizes ranging from  $2\mu\text{m}$  to  $30\mu\text{m}$  by applying various voltages that result in acoustophoretic forces corresponding to different channel widths.

**Keywords**—COMSOL, acoustic radiation force, IDTs, microfluidic device, blood cell separation, multiple particles.

## I. INTRODUCTION

Cell separation and sorting can be used in a variety of biological and biomedical applications, like cell biology, diagnostics, and therapeutics. For e.g., the diagnosing and treating of HIV disease involve the separation of T-lymphocytes (CD4+) from blood [1]. Blood checks, which extract cells from a patient's blood sample, are the most common tool for diagnosing diseases. Manual blood checks will occasionally result in erroneous test findings, which can lead to patients getting the wrong drug, causing serious damage or allergies [2], [3]. It may lead to death. Filtration, centrifugation, acoustics, optics, and electrophoresis are only a handful of the methods that can be used to distinguish cells depending on their physical properties [4], [5].

Acoustic-based methods have the advantage of compressibility, biocompatibility, and label-free design among these strategies. The words "acoustophoresis" and "phoresis" (migration) refer to the movement's executors, i.e., "phoresis" (migration) and "acousto" (sound waves). Electric forces move particle in electrophoresis[6], and magnetic forces move particle in magnetophoretic [7]. Acoustophoresis is a non-contact, label-free form of communication. This technology is used more commonly in bioanalytical and therapeutic cell handling and manipulation applications. The Royal Society of Chemistry recently published a comprehensive tutorial series on acoustophoresis [8].

Acoustic tweezer performs exceedingly error-free cell simulations, are highly promising for cell separation applications and provide additional benefits in terms of ease to use and flexibility [9]. The most popular method to accomplish acoustic isolation is to set up a standing desk. An acoustic field occurs inside the channel. When there is static acoustic force populates the liquid medium will be

compelled to places the less amount of acoustic radiation under duress (pressure nodes) [10].

Particles of varying size and physical properties will be exposed to various types of acoustic radiation pressures move to the pressure nodes at different intervals, resulting in distinct identifiers for separation. The fluid flow direction of all current acoustic separation methods is collateral with the standing acoustic wave [11]. Acoustic separation can be an effective method for label-free separation of cells in biology science, detection of diseases, and clinical practice.

There are totally three inlets, the two inlets are symmetrically tilted, and a centre inlet makes up the device's geometry. Using a tilted-angle with three inlets, we suggest a previously inaccessible approach to acoustic isolation of cells and particles. Surface acoustic waves (SAW) cause pressure nodal lines that are inclined at a particular angle to the flow of direction [12]. In the world of digital electronics, SAW device play a vital role. Researchers have identified different types of SAW instruments, including actuators, filters, oscillators, resonators, and sensors, which are used in many industries and electronic equipment [13].

The interdigitated transducer (IDT), is fabricated on a microfluidic channel to create the Surface acoustic wave. IDT electrodes are typically made of copper metal. The IDT in the microfluidic channel decreases the propagation process velocity of surface waves. To overcome the shortcomings of the BAW-based methodology, the size-dependent particle separation using surface acoustic wave (SAW) has recently been investigated [14]. SAW-based technique works with the microfluidic channel material and it can be conveniently implemented into a multi-functional system. In today's SAW-based particle separation techniques, external sheath flows are used to separate the particles or oriented particle against the target of channel's central area [15].

In this novel acoustic configuration, we proposed a microfluidic system capable of splitting several particles ranging from  $2\mu\text{m}$  to  $20\mu\text{m}$ . We conducted detailed parametric analyses of key variables affecting the research platform's output to see how these parameters influence separation outcomes. As just a result, cells were isolated from blood cells with greater efficiency than previously conceivable. Finally, we used our acoustic radiation force to isolate clinical samples at high throughput and were successfully identified.

## II. MATERIALS AND METHODS

### A. Acoustophoresis

Acoustophoresis is the migration of particles or cells caused by the application of acoustic pressure. Nonlinear terms in the governing equations allow momentum to be transferred from an acoustic field to particles. The acoustic radiation force is the product of this net force acting on the particles.

$$F_{rad} = -2\pi r_p^3 \left[ \frac{1}{3} K_s Re(f_0^{sl} p^* \nabla p) - \frac{1}{2} \rho Re(f_1^{sl} u^* \cdot \nabla u) \right]$$

Whereas,

Radiation force **Frad** for single spherical particles of radius refers as 'r', density refer as 'ρ', the compressibility of liquid refer as 'Ks', acoustic velocity refer as 'u', acoustic pressure as 'p', where f0 and f1 are pre-factors is given by

$$f_0^{sl} = 1 - \tilde{k}_s, \quad \text{where} \quad \tilde{k}_s = \frac{k_{s,p}}{k_s}$$

$$f_1^{sl} = \frac{2(\tilde{\rho} - 1)}{2\tilde{\rho} + 1}, \quad \text{where} \quad \tilde{\rho} = \frac{\rho_p}{\rho}$$

$$k_{s,p} = \frac{1}{\rho_p (c_{p,p}^2 - \frac{4}{3} c_{s,p}^2)}, \quad \text{where} \quad k_s = \frac{1}{\rho c^2}$$

Creating a standing acoustic field within a flow channel is a common way to achieve acoustic separation. As a standing acoustic field occurs in a liquid medium, particles in the fluid are forced into pressure nodes with the acoustic radiation pressure. Particles with different physical properties and different size can emit acoustic radiation in different ways.

### B. Gravity force

$$F_g = m_p g \frac{\rho_p - \rho}{\rho_p}$$

where the force of gravity is balanced by the acoustic radiation pressure, that will exert sound wave on the surface of the medium.

### C. Geometry of the separation device

This paper represents three inlets, two outlets, and a middle area that generate acoustic radiation force with the help of IDT ranging in sequence will make the particles flow by altering the path. At the entry, blood containing a combination of cells will pump the particles of different sizes to outlet exits. Inlet 2 is used to inject blood mixture

cells, while the angularly inclined angle inlet is used to inject buffer solution. In the separation area, the width of the channel is 280 μm and the fluid flow is in laminar type. The diagram of the two-dimensional modelled geometry as seen in Figure 1. Inlet 1 has an inflow velocity of (800 μm/s), inlet 2 has a velocity of (134 μm/s), and inlet 3 has a velocity of (1000 μm/s) to centre the particles at appropriate outlets. The IDT is set so that micron-sized particles are subjected to adequate acoustic force. As seen in fig. 1, the IDT is attached in a series to provide positive and negative potentials, and the overall aspect creates a fitting non-uniform IDT to generate SAW, which results in acoustophoretic radiation force.

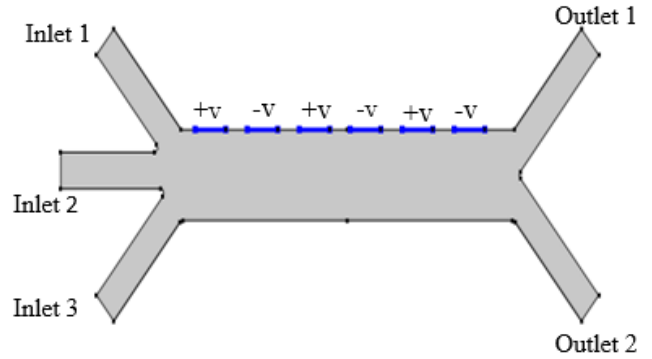


Fig. 1. Geometrical representation with three inlets and two outlet.

### D. Numerical Analysis

Computational equations were validated to test each convection, and acoustophoretic forces were found to be sensible on the cells. The 2D finite element model was studied using the COMSOL Multiphysics software kit, which returned the electric potential distribution and flows within the device's geometry.

In our model, the following physical phenomena are used:

- (1) Illustrate fluid movement with a creeping flow.
- (2) Measure the direction of different components of blood under the control of acoustophoretic force and particle pull is occurred using the particle tracing method.
- (3) The force of gravity is balanced by sound flowing through a fluid.

TABLE 1. DIAMETER AT WHICH PARTICLES ARE SEPARATE

Cells	Diameter (μm)	Density (kg/m <sup>3</sup> )
Cancer	21	1050
RBCs	6	1050

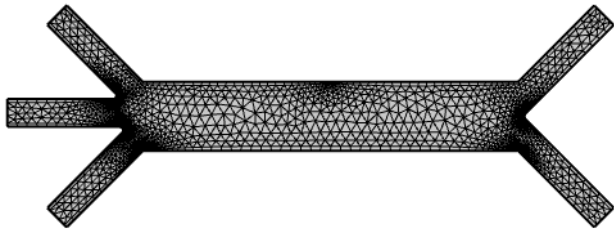


Fig. 2. The meshing stage was then finished using the default method.

### III. RESULT

#### A. Effect of velocity

The effect of flow velocity on particle separation was observed and we identified the optimum velocity for effective particle separation. The particle separation effects under various velocities are presented in Table 2 to Table 4

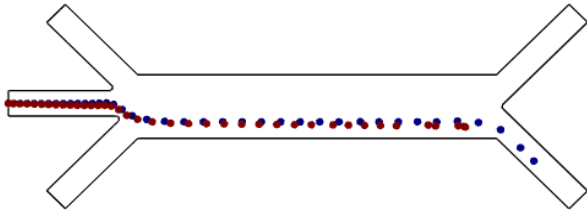


Fig. 3. It shows that higher velocity in inlet 1 and low velocity in inlet 2 and inlet 3 pushes the particles downwards(outlet 2).

TABLE 2. VELOCITIES AT WHICH PARTICLE FOCUS TO THE BOTTOM

	Inlet 1	Inlet 2	Inlet 3
Velocity Applied ( $\mu\text{m}$ )	1000	150	150

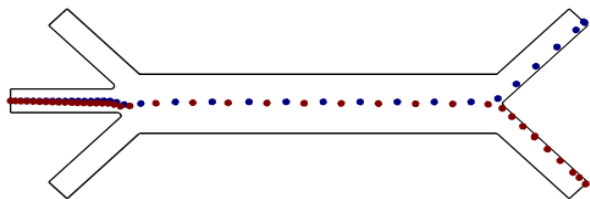


Fig. 4. It shows that higher velocity in inlet 3 and low velocity in inlet 1 and inlet 2 pushes the particles upwards (outlet 1).

TABLE 3. VELOCITY AT WHICH EFFECTIVE SEPARATION OCCURS

	Inlet 1	Inlet 2	Inlet 3
Velocity Applied ( $\mu\text{m}$ )	800	134	1000

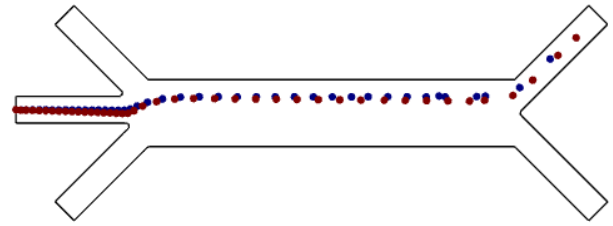


Fig. 5. It shows that higher velocity in inlet 1 and inlet 3, low velocity in inlet 2 pushes the particles to align in centre. Henceforth the particles got separated(outlet 1 and outlet 2).

TABLE 4.EFFECTIVE SEPARATION OF DIFFERENT BLOOD CELLS AT SPECIFIC INLET VELOCITIES

	Inlet 1	Inlet 2	Inlet 3
Velocity Applied ( $\mu\text{m}$ )	134	134	1000

#### B. Acoustic pressure

At a velocity range of 800-134-1000 $\mu\text{m}$ , the particle gets perfectly separated with an applied frequency of 50MHz, however,a frequency less than 50 MHz, the particles lead to imperfect isolation at the collection outlet. The effect of acoustic field generation under the influence of various frequency is shown in the figure below.

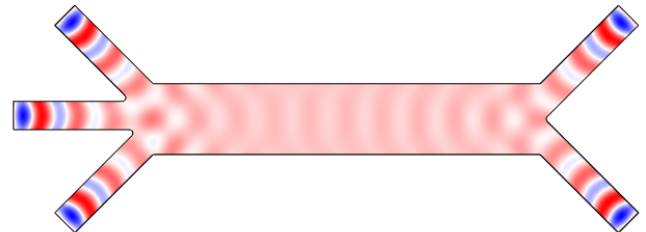


Fig. 6. Acoustic pressure occurs at 30MHz frequency.

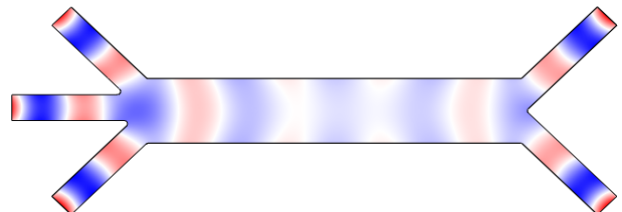


Fig. 7. Acoustic pressure occurs at 10MHz frequency.

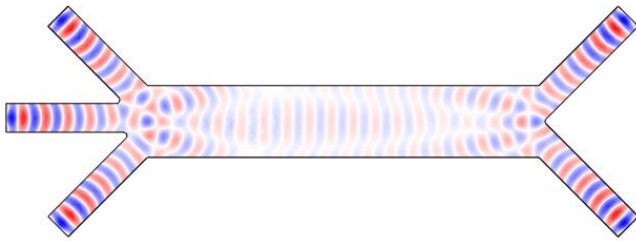


Fig. 8. Acoustic pressure occurs at 50MHz frequency.

C. Various sized particle separation

The effect of different blood cells separation is also further in our research. In our device, we observed that the blood cells of different size can be separated by just tuning the inlet velocity. We considered the size difference 5 μm and 21 μm which corresponds to the size of platelets and cancer cells, 8 μm and 25 μm which mimics the size of RBCs and Cancer cells and 15-25 μm that corresponds to the size of WBC and Cancer cells. Table 5 shows the effective separation of different blood cells combinations at specific velocities

TABLE 5.EFFECTIVE SEPARATION OF DIFFERENT BLOOD CELL AT SPECIFIC INLET VELOCITIES

Cells	Particle size	Inlet flow			Frequency (Hz)
		Inlet1 (μm)	Inlet 2 (μm)	Inlet 3 (μm)	
Platelets - Cancer cells	5-21	800	240	1000	8000
RBC- Cancer cells	8-25	800	334	1000	6000
WBC- Cancer cells	15-23	800	210	1200	8000

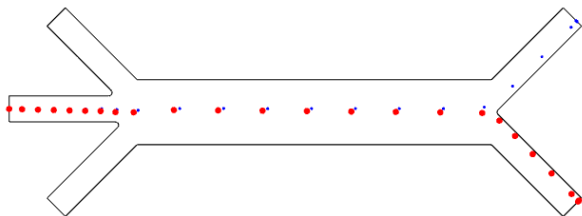


Fig. 9. The particles are emitted simultaneously and pass in the same direction. The Platelets are blue, and the cancer cells are red.

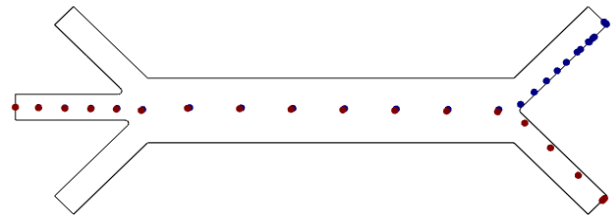


Fig. 10. The particles are emitted simultaneously and in the same direction. WBCs are shown in blue, while cancer cells are shown in red.



Fig. 11. The particles are emitted simultaneously and in the same direction. RBCs are shown in green, while cancer cells are shown in red.

IV. DISCUSSION

a) In this article, a two-dimensional consol simulation in which blood cells are sorted using SSAWs is investigated. The modelling is generated using a finite element simulation.

b) The Helmholtz and Navier–Stokes equations, as well as the equation of motion are solved.

c) Angular frequency and blood inlet velocity were discovered to be the optimum separation parameters.

In the COMSOL Multiphysics programme, simulations were observed.

The impact of cell displacement on flow velocity at the top, middle and bottom inlets was also studied. The data were analysed at different frequencies. The top and bottom inlets have a higher flow rate than the middle inlet, which serves to decrease the number of blood cells organised and required to flow in the centre fig.5. Higher flow rates were used to optimise the movement of the blood cell. The results of their tests revealed a very high rate of blood cell differentiation, implying perfect particle segregation with maximum efficiency. The acoustic-based separation system demonstrated here has a range of advantages for CTC separation applications, including, automated operation ,a label-free, contactless design, high biocompatibility, all in a compact and low-cost package.

V. CONCLUSION

This paper uses acoustophoresis to isolate various blood components using a channel configuration with a sequence arrangement of IDT. IDT configurations of sidewall make velocities deviate the cell separation on the outlet of the tube-based on the size. The Particle tracing method is used to model the displacement of the micro-particles for fluid flow interface into the channel due to drag force, radiation forces. The equilibrium of drag and

radiation governs the movement of particles. This is dependent on the particle size 'r' as well as the fluid and material parameters.

#### REFERENCES

- [1] S. O. Catarino, R. O. Rodrigues, D. Pinho, J. M. Miranda, G. Minas, and R. Lima, "Blood cells separation and sorting techniques of passive microfluidic devices: From fabrication to applications," *Micromachines*, vol. 10, no. 9, 2019, doi: 10.3390/mi10090593.
- [2] J. Ma *et al.*, "Paper Microfluidics for Cell Analysis," *Adv. Healthc. Mater.*, vol. 8, no. 1, 2019, doi: 10.1002/adhm.201801084.
- [3] B. I. M. Z. Ansar *et al.*, "Simulation Guided Microfluidic Design for Multitarget Separation Using Dielectrophoretic Principle," *Biochip J.*, vol. 14, no. 4, pp. 390–404, 2020, doi: 10.1007/s13206-020-4406-x.
- [4] Z. T. F. Yu, K. M. Aw Yong, and J. Fu, "Microfluidic blood cell sorting: Now and beyond," *Small*, vol. 10, no. 9, pp. 1687–1703, 2014, doi: 10.1002/sml.201302907.
- [5] M. Z. Ansar, M. Y. Caffiyar, and A. Mohammed Kashif, "Microfluidic device for Multitarget separation using DEP techniques and its applications in clinical research," in *2020 6th International Conference on Bio Signals, Images, and Instrumentation, ICBSII 2020*, 2020, doi: 10.1109/ICBSII49132.2020.9167655.
- [6] J. Zhu and X. Xuan, "Particle electrophoresis and dielectrophoresis in curved microchannels," *J. Colloid Interface Sci.*, vol. 340, no. 2, pp. 285–290, 2009, doi: 10.1016/j.jcis.2009.08.031.
- [7] C. Iliescu, G. Tresset, and G. Xu, "Dielectrophoretic field-flow method for separating particle populations in a chip with asymmetric electrodes," *Biomicrofluidics*, vol. 3, no. 4, 2009, doi: 10.1063/1.3251125.
- [8] M. C. Jo and R. Guldiken, "Active density-based separation using standing surface acoustic waves," *Sensors Actuators, A Phys.*, vol. 187, pp. 22–28, 2012, doi: 10.1016/j.sna.2012.08.020.
- [9] S. Nagrath *et al.*, "Isolation of rare circulating tumour cells in cancer patients by microchip technology," *Nature*, vol. 450, no. 7173, pp. 1235–1239, 2007, doi: 10.1038/nature06385.
- [10] L. Chen *et al.*, "Dielectrophoretic separation of particles using microfluidic chip with composite three-dimensional electrode," *Micromachines*, vol. 11, no. 7, 2020, doi: 10.3390/mi11070700.
- [11] X. Ding *et al.*, "Cell separation using tilted-angle standing surface acoustic waves," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 111, no. 36, pp. 12992–12997, 2014, doi: 10.1073/pnas.1413325111.
- [12] A. Lenshof and T. Laurell, "Continuous separation of cells and particles in microfluidic systems," *Chem. Soc. Rev.*, vol. 39, no. 3, pp. 1203–1217, 2010, doi: 10.1039/b915999c.
- [13] C. M. Yousuff, E. T. W. Ho, K. Ismail Hussain, and N. H. B. Hamid, "Microfluidic platform for cell isolation and manipulation based on cell properties," *Micromachines*, vol. 8, no. 1, pp. 1–26, 2017, doi: 10.3390/mi8010015.
- [14] C. M. Yousuff, N. H. B. Hamid, K. I. Hussain, and E. T. W. Ho, "Numerical modelling and simulation of dielectrophoretic based WBC sorting using sidewall electrodes," in *International Conference on Intelligent and Advanced Systems, ICIAS 2016*, 2017, pp. 1–5, doi: 10.1109/ICIAS.2016.7824090.
- [15] M. Punjiya, H. R. Nejad, J. Mathews, M. Levin, and S. Sonkusale, "A flow through device for simultaneous dielectrophoretic cell trapping and AC electroporation," *Sci. Rep.*, vol. 9, no. 1, 2019, doi: 10.1038/s41598-019-48198-x.

**2021 IEEE Seventh  
International Conference on  
Bio signals, Images and  
Instrumentation  
ICBSII 2021  
SESSION IX  
RESEARCH  
PAPERS**

# APPARENT DIFFUSION COEFFICIENT VALUES WITH NORMALIZATION IN CORRESPONDENCE WITH WHO GRADING OF GLIAL TUMOURS OF THE BRAIN

Gaurav Mishra  
Department of Radiodiagnosis  
Jawaharlal Nehru Medical College;  
Datta Meghe Institute of Medical  
Sciences  
Wardha, India  
[drgvmishra@gmail.com](mailto:drgvmishra@gmail.com)

Akhilesh Agrawal  
Research & Development  
Jawaharlal Nehru Medical College;  
Datta Meghe Institute of Medical  
Sciences  
Wardha, India  
[agrawalakhilesh77@gmail.com](mailto:agrawalakhilesh77@gmail.com)

Punit Fulzele  
Department of Pedodontics &  
Preventive Dentistry  
Jawaharlal Nehru Medical College;  
Datta Meghe Institute of Medical  
Sciences  
Wardha, India  
[punitr007@gmail.com](mailto:punitr007@gmail.com)

Ashutosh Bagde  
Research & Development  
Jawaharlal Nehru Medical College;  
Datta Meghe Institute of Medical  
Sciences  
Wardha, India  
[bagde.ashu@gmail.com](mailto:bagde.ashu@gmail.com)

## Abstract

**Aim:** To evaluate the Apparent diffusion co-efficient in grading of Glial tumors with histo-pathological grading

**Method:** Cross sectional hospital based observational study conducted over a time period of 2 years. All the cases registered with Acharya Vinoba Bhave Rural Hospital Sawangi, Wardha, diagnosed on histo-pathological findings as Glial tumours were included in this present study. All patients were examined on GE Brivo MRI machine with 1.5Tesla magnetic field strength in the Department of Radiodiagnosis. Using b value of 1000 s/sq. mm, slice thickness 5mm, interslice gap 2mm applied in the x, y and z axes, central as well as peripheral portions of the tumor were manually sampled, preferably getting rid of any cystic or necrotic areas within them. Histo-pathological diagnostic acumen was augmented with radiological features of the brain tumors upon routine MRI SEQUENCES coupled with ADC values.

**Result:** In the present study, out of the total of 142 patients included, the glial tumour cases were 54 in number. Amongst this group of glial tumours, the tumours diagnosed were pleomorphic xanthoastrocytoma, pontine glioma, and thalamic glioma, low grade glioma at other sites, high grade gliomas, grade 3 astrocytomas, glioblastoma multiformae, gliomatosis cerebri and gliosarcoma. The mean ADC for Pleomorphic xanthoastrocytoma, pontine glioma, thalamic glioma and Low grade gliomas were **0.98, 0.82, 0.95 and 0.89** respectively. While, the mean normalized ADC for Pleomorphic xanthoastrocytoma, pontine glioma, thalamic glioma and Low grade gliomas were **1.23, 1.03, 1.19 and 1.12** respectively.

**Conclusion:** The research gap analysis towards which research question was framed stands filled up by the generated new knowledge in terms of 'Apparent Diffusion Coefficient' and 'Normalised Apparent Diffusion Coefficient' values with identified cut off / demarcating

values for gradation of Glial Tumors into Low Grade / High Grade tumors

**Keywords—** Glioma, MRI, Apparent Diffusion Coefficient, histopathological grading

## I. INTRODUCTION

In our nation, which is a developing one making rapid progress, cancer still has its claws deep within, engulfing its very foundations. According to the available literature sources regarding epidemiological aspects of brain tumours, 3.4 per 100,000 population is the crude incidence of primary brain tumour for males and for females it is reported as 1.2 per 100,000 population. This constitutes less than 1% of the total tumours that stand identified. Last decade is marked by evident increase in primary brain tumours as a whole<sup>1</sup>. It is pertinent to note that the Population Based Cancer Registry (PBCR) in Wardha district, which is 2% of total area of Maharashtra state and with an estimated population as 15,18,995 with male to female ratio of 1000 : 946, reflected 54 cases of brain tumours like gliomas in the year 2013-2014<sup>2</sup>. The assigned arena for diagnostics emerged rapidly since the paramount discovery of X-rays by William Conrad Roentgen, dating back to 8th November, 1895. He was honoured with the coveted Nobel Prize in Physics in 1901 for his landmark contribution that change the face and facet of diagnostics in the interest of men and mankind. The 1970's stood out highlighting the rapid advances made in the stream of radiology. Sir Godfrey N. Hounsfield devised a "game changing" invention in the form of Computed Tomography (CT) machinery<sup>3</sup>. The greatest advantage of CT equipment proved to be delineating the calcification within the soft tissues as well as properly enlightening about the involvement of the skeleton. The features like grading of the tumour, its status regarding its severity, occurrence of intratumoral changes such as haemorrhage, necrosis, cystic degeneration were overlooked as CT could not properly deliver upon these counts, in spite of being the primary modality of detection back then<sup>4</sup>. Moving ahead on the



timeline, the diagnostics went through a paradigm shift as Magnetic Resonance Imaging (MRI) came into reckoning towards the end of the 1980's<sup>5,6</sup>. These comprise mainly of magnetic resonance spectroscopy (MRS) and Diffusion weighted imaging (DWI), Apparent Diffusion Co-efficient values etc.<sup>4</sup>. Taking into picture, the pre-operative diagnosis along with staging and grading of glioma tumors, Magnetic resonance spectroscopy (MRS) is considered to be the radiological gold standard<sup>7</sup>. However, the overall gold standard of tumor grading is fulfilled by performing a biopsy along with either an open or stereotactic neurosurgical procedure in order to decipher the histo-pathologic diagnosis of the notorious tumour<sup>8,9</sup>. Cellularity and grading of gliomas have been correlated with Apparent Diffusion Coefficient (ADC) values. Gliomas having an increased amount of cellularity as well as higher grades tend to show lower rate of diffusion of water molecules across the tumoral tissue resulting into correspondingly lower ADC values<sup>10,11</sup>. Thus, the glioma tumour cellularity can be said to be inversely proportional to its corresponding ADC value. The future advances hold the key to increasingly less time consuming and swifter tumour diagnosis and correspondingly their respective treatment protocols as technology moves ahead and more and more. Therefore the aim of this study was to evaluate the Apparent diffusion coefficient in grading of gliomas with histo-pathological grading”

## II. MATERIALS AND METHODS

### A. Study Design :-

Cross sectional hospital based observational study

### B. Study Setting

⊙ All patients were examined on GE Brivo MRI machine with 1.5Tesla magnetic field strength in the Department of Radiodiagnosis, Jawaharlal Nehru Medical College, constituent unit of Datta Meghe Institute of Medical Sciences (Deemed to be University), Sawangi Meghe, Wardha.

### C. Study Period

⊙ 2 years - 12<sup>th</sup> April, 2016 to 11<sup>th</sup> April, 2018.

### D. Participants

⊙ All the cases registered with Acharya Vinoba Bhave Rural Hospital attached to Jawaharlal Nehru Medical College, Sawangi, Wardha, diagnosed on histo-pathological findings as Glial tumours, Meningial tumours, and Tumours of the Sellar region were included in this present study.

### E. Tumour groups studied

Glial tumors (n=54)

### F. Sample size

Sample size availed in the present study was 142, out of which glial tumour cases were 54.

### G. Inclusion criteria

1. All patients aged 4 years and above histopathologically diagnosed as
  - a) Glial tumors
  - b) Meningial tumors
  - c) Tumors of the sellar region
2. All patients with MRI data of satisfactory imaging quality for the purpose of correlation.
3. Those giving informed consent

### H. Exclusion criteria

1. Patients aged less than 4 years of age (MR spectra of metabolites are difficult to distinguish between due to rapid metabolic changes until age of 4 years of age)
2. Patients with poor quality MRI data – due to artefact or contamination by the unwanted metabolites or magnetic inhomogeneities.
3. Patients with pacemakers.
4. Patients with cochlear implants.
5. Patients with claustrophobia.
6. Patients not giving consent.

### I. ADC (Apparent Diffusion coefficient)

Using b value of 1000 s/sq. mm, slice thickness 5mm, interslice gap 2mm applied in the x, y and z axes, central as well as peripheral portions of the tumor were manually sampled, preferably getting rid of any cystic or necrotic areas within them.

For normalization of the ADC values, we took the actual ADC value in the numerator and took the normal value of ADC in normal grey matter within the brain as 0.8 in the denominator. The output of this fraction calculation was taken as the normalized ADC value.

### J. Biopsy and histo-pathological correlation

- Formalin-fixed paraffin mounted specimen of the diseased tissue - open biopsy - experienced neurosurgeon.
- **Specimens** - examined by an multiple experienced pathologists - for mitotic figures, necrotic component, abnormal microvasculature, nuclear pleomorphism.
- The **histo-pathological grading** of the diseased tissue was done as per the **WHO grading of brain tumours, whereby**
- Low grade (class I and II) and
- High grade tumours (class III and IV).

- **Histo-pathological diagnostic acumen** was augmented with **radiological features** of the brain tumors upon routine **MRI SEQUENCES** coupled with **ADC values**

### III. RESULT

HPD	Mean value	Mean normalized ADC value
Pleomorphic xanthoastrocytoma	0.98	1.23
Pontine glioma	0.82	1.03
Low grade glioma	0.95(0.05)	1.19(0.06)
Thalamic glioma	0.89	1.12
<u>Balos et al</u> <sup>12</sup>	1.45	-

In the present study, out of the total of 142 patients included, the glial tumour cases were 54 in number. Amongst this group of glial tumours, the tumours diagnosed were pleomorphic xanthoastrocytoma, pontine glioma, thalamic glioma, low grade glioma at other sites, high grade gliomas, grade 3 astrocytomas, glioblastoma multiformae, gliomatosis cerebri and gliosarcoma.

In this study, the mean ADC for Pleomorphic xanthoastrocytoma, Pontine glioma, Thalamic glioma and Low grade gliomas were **0.98, 0.82, 0.95 and 0.89** respectively. While, the mean normalized ADC for Pleomorphic xanthoastrocytoma, Pontine glioma, Thalamic glioma and Low grade gliomas were **1.23, 1.03, 1.19 and 1.12** respectively. These findings match with the study done by Balos et al<sup>12</sup>.

High Grade astrocytoma Gr 3- 4 (Histopathologically diagnosed)		ADC Value x10 <sup>-3</sup>	
		Mean (sd)	Mean nADC(sd)
In our study	Glioblastoma Multiforme	0.76(0.05)	0.96(0.06)
	Gliomatosis cerebri	0.98	1.22
	Gliosarcoma	0.66(0.04)	0.83(0.05)
	Astrocytoma grade III	0.79	-
Dawoud et al <sup>13</sup>	High grade 3-4	0.953	-
Sherbeny et al <sup>14</sup>	GBM	0.9	-
Yamasaki et al <sup>15</sup>	GBM	1.079	-
Darweesh et al <sup>16</sup>	Anaplastic Astrocytoma Gr-3	0.6 -1.0	-
	GBM	0.2-0.8	-
<u>Balos et al</u> <sup>12</sup>	Astrocytoma Gr -3	1.14	-

### IV. DISCUSSION

Conventional MRI, being limited by its number of imaging sequences, has seemingly limited scope regarding differentiation of brain tumours with respect to their benignancy or malignancy. Apparent diffusion coefficient value is derived as an extremely valuable adjunct to the same front as it adds in the form of a coefficient which lets the radiologist know regarding the diffusion of water molecules across the tumour in the yardsticks of a number, which diffusion weighted imaging sequence could direct whether there is restriction in diffusion or not within the tumour<sup>17,18</sup>. In the present study, the mean ADC for Pleomorphic xanthoastrocytoma, Pontine glioma, Thalamic glioma and Low grade gliomas were 0.98, 0.82, 0.95 and 0.89 respectively. While, the mean normalized ADC for Pleomorphic xanthoastrocytoma, Pontine glioma, Thalamic glioma and Low grade gliomas were 1.23, 1.03, 1.19 and 1.12 respectively. These findings match with the study done by Balos et al<sup>12</sup>. In the present study in terms of the ADC values, a conclusion can be derived that significantly lower values of ADC as well as normalized ADC are the features of high grade malignant tumours as against the lower values of ADC as well as normalized ADC being attributable to low grade tumours. Thus it results in differentiation between a low grade glioma from that of a malignant high grade one. This finding is in conformity with the one reported by Sherbeny et al<sup>14</sup>, Yamasaki et al<sup>15</sup> and Darweesh et al<sup>16</sup>.

The findings in the present study in regard to ADC and nADC are in tune with the findings of Darweesh et al<sup>16</sup> and Bulakbasi et al<sup>19</sup> to the effect that significant lower ADC values were attributable to high grade tumours which are much lower in the benign low grade tumours. In the present study, calculated

ADC values showed the lowest values in high grade tumours as against the benign low grade tumours. This indicates that low ADC values are indicative of high grade tumours whereas high ADC values are indicative of low grade tumours. Differentiation of tumour sub-type within the grade was not observed which is akin to the observations reported by Sherbeny et al<sup>14</sup>, Dawoud et al<sup>13</sup>, Darweesh et al<sup>16</sup> and Bulakbasi et al<sup>19</sup>. As such, it could be safely concluded that ADC calculation can provide additive valuable information helping in tissue characterization of intra-cranial tumours leading to improved diagnostic grading and resultant management and prognosis. This can be said to be an important finding in the present study, which is in tune with the findings reported by Sherbeny et al<sup>14</sup>, Dawoud et al<sup>13</sup>, Darweesh et al<sup>16</sup> and Bulakbasi et al<sup>19</sup>. Thus, it can be hypothesized that higher the value of ADC, lower the tumour grade and vice versa pointing towards an inversely proportional relationship in between the two.

### V. CONCLUSION

Upon application of 'logistic regression analysis', ADC values and Normalized ADC values were calculated for low and high grade glial tumors respectively. For low grade tumors, the mean ADC value was observed to be 1.19 (p<0.001). For high grade glial tumors, the mean ADC

value was observed to be 0.88( $p < 0.001$ ). The normalized ADC values were observed to be as 1.49 for low grade glial tumors( $p < 0.001$ ). The normalized ADC values were observed to be as 1.10 for high grade glial tumors ( $p < 0.001$ ). The ADC 'cutoff' value for low grade glial tumour came out to be more than or equal to 1.029 for low grade glial tumors. The normalized ADC cut off value was observed to be more than or equal to 1.286 for low grade glial tumours. The ADC 'cutoff' value for high grade glial tumour came out to be less than 1.029. The normalized ADC cut off value was observed to be less than 1.286 for high grade glial tumours.

The cut off values obtained using logistic regression analysis can serve as adjunct findings in predicting the histopathological grade of the tumour upon ADC application on MRI giving a clue or a direction to the clinician, radiologist as well as the histopathologist in predicting the grade and behaviour of the tumour in question. Provision of a value of ADC as well as normalized version of the same within suitable intratumoural area can point towards the possible grade of the tumour in conjunction with the gold standard histopathological examination in vogue.

#### REFERENCES

1. Mahalakshmi M. Hierarchical Self-Organizing Map (HSOM) based Segmentation of Brain Tumour from Brain MRI Images. *IJSRD*. 2015; 3(03):1010-1013.
2. Gagane N. Population based cancer registry, wardha. Mahatma gandhi institute of medical sciences, sevagram. Individual Registry Write Up: 2010-2011:140-141.
3. Castillo M. History and evolution of brain tumor imaging: insights through radiology. *Radiology*. 2014 Nov;273 (2 Suppl):S111-25.
4. Kieffer SA, Brace JR. Intracranial neoplasm. In: Haaga JR, Dogra VS, Forsting M, Gilkeson RC, Ha HK, Sundram M, editors. *CT and MRI of the whole body*, 5<sup>th</sup> ed. Philadelphia: Mosby Elsevier; 2009. Vol 1. pp. 49-144.
5. Beckmann EC. CT scanning the early days. *BJR*. 2006;79(937):5-8.
6. Wilkins RH. *Neurosurgical Classics*. U.S.A.: American Association of Neurological Association; 1965
7. Kaddah RO, Khalil ME. Malignant focal brain lesions. Value of MRS tumour biomarkers in preoperative prediction of grades of malignancy. *The Egyptian Journal of Radiology and Nuclear Medicine*. 2014 DEC; 45(4): 1201-1208.
8. Shokry A. MRS of brain tumors: Diagrammatic representations and diagnostic approach. The Egyptian Journal of Radiology and Nuclear Medicine. 2012 Dec;43(4):603-612
9. Gugulothu, S.S., R.S. Choudhary, and S. Telrandhe. "CMOS Amplifiers Design for Neural Recording System." *Journal of Advanced Research in Dynamical and Control Systems* 11, no. 8 Special Issue (2019): 3071-76.
10. Drevelegas A, Papanikolaou N. Imaging Modalities in Brain Tumors. In: Drevelegas A, editors. *Imaging of Brain Tumors with Histological Correlations*, 2<sup>nd</sup> ed. Greece: Springer; 2012. pp. 1333.
11. Fulzele, P., Z. Quazi, A. Sirsam, S. Khobargade, Y. Chitriv, K. Singh, and S. Choudhary. "Methods for Early Detection of Postoperative Infectious Review." *Journal of Advanced Research in Dynamical and Control Systems* 11, no. 8 Special Issue (2019): 3155-67.
12. Balos DR, Gavrilović S, Lavrić S, Vasić B, Macvanski M, Damjanović D, et al. Proton magnetic resonance spectroscopy and apparent diffusion coefficient in evaluation of solid brain lesions. *Vojnosanit Pregl*. 2013 Jul;70(7) :637-44.
13. Dawoud MAE, Sherif MF, Eltomey MA. Apparent diffusion coefficient and Magnetic resonance spectroscopy in grading of malignant brain neoplasms. *The Egyptian Journal of Radiology and Nuclear Medicine*. 2014 Dec; 45(4), 1215-1222.
14. Sherbeny A, El-Shafey M, Biomy S, Shakal A, Hefeda M, Seiam A. Diagnostic yield of combined magnetic resonancespectroscopy and diffusion weighted imaging in intracranial neoplasms. *The Egyptian Journal of Radiology and Nuclear Medicine* . 2014 Sep; 45(3):849- 858
15. Yamasaki F, Kurisu K, Satoh K, Arita K, Sugiyama K, Ohtaki M, et al. Apparent Diffusion Coefficient of Human Brain Tumors at MR Imaging. *Radiology*. 2005 Jun; 235 (3):985-991.
16. Yamasaki F, Takayasu T, Nosaka R, Amatya VJ, Doskaliyev A, Akiyama Y, et al. Magnetic resonance spectroscopy detection of high lipid levels in intraaxial tumors without central necrosis: a characteristic of malignant lymphoma. J Neurosurg. 2015 Jun;122(6):1370-9.
17. Patil, A.R.B., S. Charde, R. Waghare, S. Warbhe, A. Sarve, P. Fulzele, and S. Choudhary. "Thoughts to Movements-Motor Imagery Based Brain Computer Interface." *Journal of Advanced Research in Dynamical and Control Systems* 11, no. 8 Special Issue (2019): 3168- 77.
18. Patil, A.R.B., S. Charde, R. Waghare, S. Warbhe, A. Sarve, P. Fulzele, and S. Choudhary. "Thoughts to Movements-Motor Imagery Based Brain Computer Interface." *Journal of Advanced Research in Dynamical and Control Systems* 11, no. 8 Special Issue (2019): 3168- 77.

19. Bulakbasi N, Kocaoglu M, Ors F, Tayfun C, Uçöz T. Combination of single-voxel proton MR spectroscopy and apparent diffusion coefficient calculation in the evaluation of common brain tumors. AJNR Am J Neuroradiol. 2003 Feb;24(2):225-33.

# Osteosit: A Dataset for Exploration and Classification of Osteoarthritis

\*<sup>1</sup>K.Rajakumar, \*<sup>2</sup>M.Parisa Beham, \*<sup>3</sup>R.Tamilselvi and \*<sup>4</sup>D.Radhika

\**Department of ECE, \*Sethu Institute of Technology, Tamilnadu-626 115.*

<sup>1</sup>rajakumar@sethu.ac.in, <sup>2</sup>parisabeham@sethu.ac.in, <sup>3</sup>tamilselvi@sethu.ac.in and <sup>4</sup>radhik.1993@gmail.com

**Abstract**—Osteoarthritis (OA) is considered the most common form of arthritis and is characterized by cartilage degradation and bone changes. Knee is most common part affected by this OA. Knee OA detection consists of classifying a knee radiograph into healthy, early stage or OA. The medical evaluation of X-ray images or CT images is currently performed by skilled physicians which is time consuming and prone to error. Considering this fact, recently more image based diagnosis system is evolving for the early detection of OA. The success of this diagnosis system mostly relies on large number of OA/Non OA images to be trained. Thus it is necessary to have a standard OA dataset for the researchers so as to take the evaluations to the advance state by properly analyzing the radiograph images. The proposed dataset, named as ‘Osteosit’, represents an initial attempt to provide a set of OA radiograph images of knee bones. It also consist OA affected images in both X-ray and CT modality. To promote the analysis of many uncertain problems in OA, Osteosit dataset will be available to the biomedical research community which can be viewed and downloaded at the following web address: <http://www.sethu.ac.in/Osteosit/>.

**Keywords**—Osteoarthritis (AO), Knee (AO), Osteosit, OA Dataset, common complex disorder

## I. INTRODUCTION

Arthritis disease is a joint disorder featuring inflammation of one or more joints. There are more than hundred types of arthritis in the world. It affects all people in spite of age, sex and race and leads to major disabilities at the later stage [1]. Some of the types of inflammatory arthritis are: Psoriatic arthritis, Ankylosing spondylitis, and Reactive arthritis. The most common type of arthritis are osteoarthritis and rheumatoid arthritis. Osteoarthritis is related to wear and tear of cartilage in the bone joints whereas the later related to the misdirected immune system. Osteoarthritis (OA) is one of the most frequent chronic illnesses, affects 40–60% of people with degenerative diseases. OA brings joint pain accompanied by various functional limitations. Knees, hips, hands, shoulder, elbow, spine and big toes are the generally affected peripheral joints in the human body. This disease places a huge burden on health-care services and the cost of OA to these services is expected to double by 2020 and again by 2030 [2]. Fig.1 shows an example of healthy knee joint and osteoarthritis affected knee joint. It is generally defined as “common complex disorder” with multiple risk factors. These risk factors are broadly categorized into: genetic factors, constitutional factors and biomechanical risk factors [2].

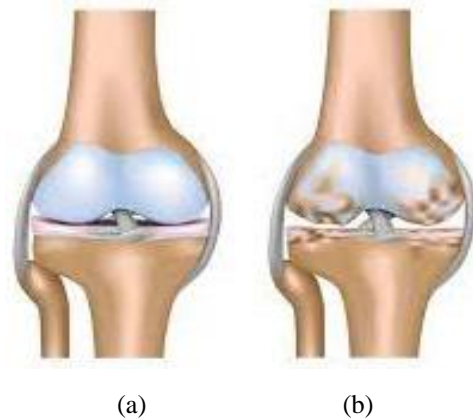


Fig.1. Knee bone Joints (a) Health knee joints (b) Osteoarthritis  
[Image courtesy: Practical Pain Management]

Such a life threatening disease has to be treated at an early stage so as to avoid misfunctionality of bone parts. Many researchers have considered this issue and provided many solutions to identify osteoarthritis at the earliest. Commonly, X-rays are used to diagnose osteoarthritis, typically revealing a loss of cartilage, bone spurs, and in extreme cases, bone rubbing against bone [3]. In most of the cases symptoms of osteoarthritis may arise even before the X-rays reveal the damages. For this reason, radiologists recommend the more perceptive MRI ((magnetic resonance imaging) and CT (computed tomography) forms of imaging for detecting early osteoarthritis. Compared to X-ray, MRI is a very sensitive imaging modality that can reveal delicate changes in bony and soft tissues whereas CT scans are excellent for showing bone spurs accurately and also shows how the bones are affected by the adjacent soft tissues [3].

Accurate diagnosis of OA is a challenging task. The diagnosis of osteoarthritis is usually based on visual perception of radiologist from the X-ray or CT or MRI. The early detection of OA as it has been suggested that the initial transformation occur in the subchondral bone before the reduction of joint space width [9]. Hence, early detection and accurate diagnosis are highly required to overcome this extremely disabling disease. As early-stage treatment of OA can be prevented the breakdown of cartilage and bone. Thus there is a growing interest among the researchers in developing a low-cost and non-invasive image-based diagnosis system for the early detection of knee OA.

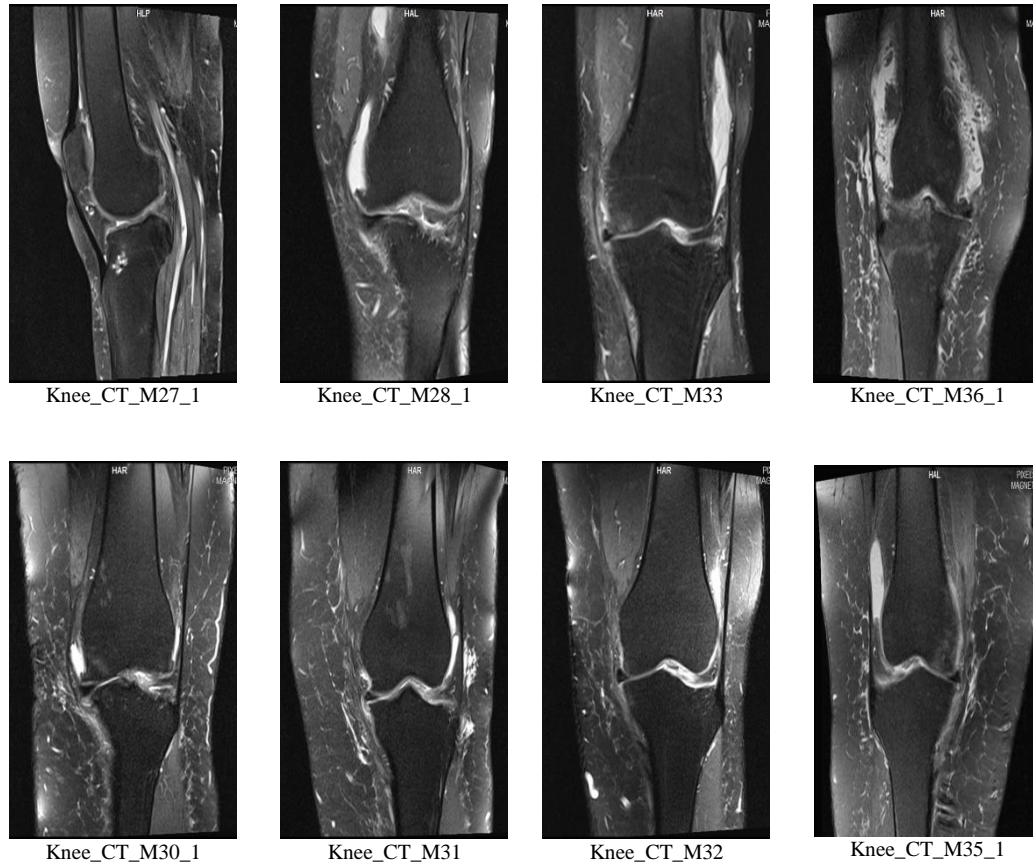


Fig. 2. Sample images of Knee in CT from Osteosit dataset

The important inference in the development of image based diagnosis system in OA is that, unavailability of suitable biomedical datasets in large numbers in various modalities. To facilitate research in Osteoarthritis, OAI (Osteoarthritis Initiative) [4] introduced a natural history dataset for osteoarthritis that include clinical evaluation data, radiological (X-ray and Magnetic resonance) images, and a bio specimen repository from 4796 men and women ages 45-79 enrolled between 2004 and 2006. Another public sharing dataset is MOST (Multicenter Osteoarthritis Study) [4] which is a longitudinal, prospective, observational study of knee OA, enrolled 3,026 study participants. Radiological data (X-Ray and MRI) for these subjects are collected and publicized for OA researchers. In both datasets, images are only in tabulated form and both are not publically available dataset.

Getting biomedical images in large number with proper ethical clearance is a major challenge in OA system. Even though some datasets are available in public which are not recent and there is no availability of high resolution X-ray and CT images. Keep all these facts in mind, in this work we proposed to build a new dataset of OA images in X-ray and CT modality, which we named it as 'OSTEOSIT Dataset' for the benefit of biomedical research community. Thus our main contributions in this work are:

- Develop a public domain research resource to facilitate the scientific evaluation of biomarkers for Osteoarthritis.
- Create a new OSTEOSIT dataset, involves 42 OA images in X-ray and 47 OA images in CT.
- Provide 10 pair of Knee OA images in X-ray modality.
- Present the annotation of the entire 50 subject's biological data with image modality.

## II. OSTEOSIT DATASET

In the biomedical research domain, unavailability of suitable biomedical dataset for advance research and development in OA motivated us to create a new and rare dataset, called as Osteosit. This Osteosit dataset is primarily a benchmark for the detection of arthritis at an early stage. Osteosit is built from the X-ray and CT radiograph bone images of arthritis. These images are collected from the standard scan centers in Tamilnadu after ethical clearance. This dataset consists of 42 OA images in X-ray and 47 images in CT modalities which are collected from 50 subjects. Among the 50 subjects, Osteosit consist of 44 female and 6 male subjects. Each subject comprises of one knee bone X-ray or CT image whereas few subjects consist of left and right knee bone images.



Fig. 3. Sample images of Knee in X-ray from Osteositis dataset

All the images are also labeled uniquely to identify easily the gender and modality of the subjects. An example of X-ray scan and CT scan images of 8 subjects from the Osteositis dataset are shown in Fig.2 and Fig. 3 respectively. Fig.4 shows the sample of knee bone images of both left and right leg captures in X-ray modality.

Osteositis dataset is developed through following steps: 1) Construction and composition 2) Labeling the images 3) Annotation. Each step is described in detail in the following subsections.

## 2.1. Construction and Composition Details

### A. Osteositis\_CT

CT scanning combines special X-ray equipment with sophisticated computers to produce multiple images of the internal structure of the body [5]. CT is typically used for examining joints that are deep in the body and difficult to assess with normal x-ray in the diagnosis of arthritis. The Osteositis dataset is constructed from the CT scan OA images of Indian people. From the 37 Indian subjects, totally 47CT Knee images have been collected. All the images are cropped manually and saved as separate images in 'jpg' (Joint Photographic Group) format. The size of all the images are 105 KB with Dimensions of 1105 x 659. The

categorization of the dataset, based on the bone images, motivate the researchers to interpret and analyze the risk condition of the knee bones accurately.

### B. Osteositis\_X-ray

A radiograph (x-ray) uses a low quantity of radiation to generate images of internal bone structures. X-rays illustrate the bones and the way they interrelate with each other at the joints. They are useful to evaluate the amount of cartilage at the ends of the bones and deformities that may lead to arthritis [5]. In the Osteositis dataset all the images are labeled perfectly for the ease of researchers. From the label itself one can easily identify the subject ID, modalities, bone part and gender of the subject. The size of all the images are 1.25 MB with Dimensions of 2000 x 2510. For example, the label for an X-ray scan image of subject is given as: Knee\_Xray\_F1.jpg; here Knee refers to bone part, X-ray refers to modality of the image captured, F refers Female, the gender of the subject and 1 is the subject ID.

### C. Osteositis\_Knee Pair

OA often affects small and large joints on both sides of the body (symmetrical), such as both knees, both hands and both wrists.



Fig. 4. Sample images of Knee Pair in X-ray from Osteositis dataset.

The size of all the images are 1.26 MB with Dimensions of 2510 x 2000. In the Osteositis\_Knee Pair images, the label for an X-ray scan image of subject is given as: KneeP\_Xray\_M12.jpg; here, KneeP refers to Knee pair, X-ray refers to modality of image, M refers Male, the gender of the subject and 12 is the subject ID.

### 2.2 Labeling the image


In the Osteositis dataset all the images are labeled perfectly for the ease of researchers. From the label itself one can easily identify the subject ID, CT or X-ray, bone part and gender of the subject. For example, the label for an Osteositis scan image of a subject is given as: Knee\_Xray\_F32.jpg and Knee\_CT\_M12.jpg. Here, in the first case Knee refers to bone part, X-ray& CT refers to the modality of image, and the later F32 refer of Female subject and 32 is the subject ID.

### 2.3 Annotation

Osteositis provides a detailed annotation through a careful analysis of each and every Osteositis scan image. We manually annotated the following attributes for each bone image.

- Unique ID of the subject
- Gender
- Age and
- Bone part
- Modality of the image
- Size of the image with Dimension

Fig.5. shows the example of annotation made for one of the images from Osteositis,



***Biological Data*** Subject  
 ID: F01  
 Gender: Female  
 Age: 67  
 Modality: X-ray  
 Resolution: 2000x2510  
 Size : 1 MB

Fig.5. Sample of annotation of X-ray OA images



## ETHICAL APPROVAL

The Osteosit dataset created in this paper is provided by Pixel scans, Trichy. The ethical committee of Pixel scans has reviewed and approved to conduct research using this dataset and publish papers based on the results using those biomedical images.

## III. CONCLUSION

In this paper, we have developed a new medical image dataset called OSTEOSIT, a collection of osteoarthritis scan images for healthcare application oriented research. It is introduced with the aim of providing a benchmark for orthopedic research and related development in detecting the arthritis at an early stage. This dataset can also be explored to identify the bone mineral density and calcium contents to classify the osteoporosis risk conditions. The main characteristics of this OSTEOSIT dataset are: a) 47 CT Knee images as *Osteosit\_CT*, b) 42 X-Ray knee scan images as *Osteosit\_X-ray* c) 10 Knee pair X-Ray images named as *Osteosit\_Knee Pair*. By providing this dataset available to the Osteoarthritis research community, we hope this Osteosit will serve as a public domain research resource in healthcare. This Osteosit dataset along with radiograph images with proper annotations will also be publically made available for research. This dataset can be viewed in the following web address: <http://www.sethu.ac.in/Osteosit/>.

## ACKNOWLEDGEMENT

The authors thank to Dr. B.R.J. Sathis Kumar, BRJ Ortho Centre, Coimbatore, TN, India for his help and suggestions in building this database. The authors also immensely thank Dr. Iayaraja Venkatachalam, Radiologist, Pixel Scans, Trichirappalli, India for providing X-Ray, CT scan images with all clinical interpretation to create this Osteosit database.

## REFERENCES

- [1] Arthritis Foundation (AF). Arthritis By The Numbers, 2019, <https://www.arthritis.org/getmedia/e1256607-fa87-4593-aa8a-8db4f291072a/2019-abtn-final-march-2019.pdf>
- [2] Jamshidi A, Pelletier JP, Martel-Pelletier J. Machine-learning-based patient-specific prediction models for knee osteoarthritis. *Nat Rev Rheumatol.* 2019 Jan;15(1):49-60. doi: 10.1038/s41584-018-0130-5. PMID: 30523334.
- [3] Roemer, F. W., Kwok, C. K., Hayashi, D., Felson, D. T., & Guermazi, A. (2018). The role of radiography and MRI for eligibility assessment in DMOAD trials of knee OA. *Nature reviews. Rheumatology*, 14(6), 372–380. <https://doi.org/10.1038/s41584-018-0010-z>
- [4] National Clinical Guideline Centre (UK). Osteoarthritis: Care and Management in Adults. London: National Institute for Health and Care Excellence (UK); 2014 Feb. PMID: 25340227.
- [5] Brahim A, Jennane R, Riad R, Janvier T, Khedher L, Toumi H, Lespessailles E. A decision support tool for early detection of knee OsteoArthritis using X-ray imaging and machine learning: Data from the OsteoArthritis Initiative. *Comput Med Imaging Graph.* 2019 Apr; 73:11-18. doi: 10.1016/j.compmedimag.2019.01.007. Epub 2019 Jan 29. PMID: 30784984.
- [6] Paixão T, DiFranco M, Goetz C, Ljuhar R, Meier P, Nehrer S. Machine learning predicts rate of cartilage loss: data from the osteoarthritis initiative (OAI) and the multicenter osteoarthritis (most) studies. *Orthopaedic Journal of Sports Medicine.* September 2020. doi:10.1177/2325967120S00530
- [7] Boesen M, Ellegaard K, Henriksen M, et al. Osteoarthritis year in review 2016: imaging. *Osteoarthritis Cartilage.* 2017;25(2):216-226. doi:10.1016/j.joca.2016.12.009
- [8] Lesca M. Holdt and Daniel Teupser. Genetic background of atherosclerosis and its risk factors. May 2015. doi:10.1093/med/9780199656653.003.0002.
- [9] Li, G., Yin, J., Gao, J. et al. Subchondral bone in osteoarthritis: insight into risk factors and microstructural changes. *Arthritis Res Ther* 15, 223 (2013). <https://doi.org/10.1186/ar4405>

# Voice Controlled Home Automation System

Nisha A<sup>1</sup>, Aneesh R P<sup>2</sup>, Ribil Mary Roy<sup>3</sup>, Blessy Sabu<sup>4</sup>

<sup>1</sup>Sivaji College of Engineering and Technology, Tamil Nadu, India

<sup>2,3,4</sup>IHRD Regional Centre, Thiruvananthapuram, Kerala, India

<sup>1</sup>1998nishanarayanan@gmail.com, <sup>2</sup>aneeshprakkulam@gmail.com,

<sup>3</sup>ribilmary27@gmail.com, <sup>4</sup>blessybben1997@gmail.com

**Abstract**—Advancements in the technologies are higher, especially for IoT, this can be integrated well with home automation to enhance daily lifestyle. Remote controlled automation systems are very helpful for bedridden patients and physically handicapped persons. Gesture controlled systems are being developed nowadays. These need more memory and processing speed. In this paper, a new system is proposed for home automation for supporting patients. This low power, low cost system works with voice commands without any other physical movements. Here a prototype is presented to implement the IoT based portable automation system. Google Assistant and Blynk are available applications for voice controlling. The demonstration of this system was performed with controlling the light. This system is simple and user friendly especially for elders and children. Hardware system is designed with the ESP32 Microcontroller which has an inbuilt Wi-Fi module. As the Wi-Fi connectivity is provided, the system can be controlled even from a large distance. Use of the latest technology and low cost, easy to operate system will make it accessible for every household. The use of home automation is now extended even to factories, various industries for the automatic ON/OFF of various machineries and electronic appliances. This system can efficiently monitor the power consumption of building automation systems like HVAC.

**Keywords**—Home automation, Google Assistant, Blynk, Wi-Fi, Light, Buzzer, Alert.

## I. INTRODUCTION

Home automation, especially in this technology era, performs a very important role where the single corner of the world is of network connectivity. The main aim of the technology is to increase efficiency and decrease the effort in human lives. Home automation is the automatic control of any electronic device in buildings and factories. Here all the devices are controlled by wireless communication facilities such as Wi-Fi, Bluetooth, etc. When these devices are connected via the internet; they can be controlled and monitored remotely. Home automation can be extended from light controlling to air-conditioning controlling and even to a smart security system [1].

Remote home monitoring permits people to manage and control various aspects of the home. These include motion detection, water leakage detection, monitoring temperature and finding burglary and fire, and controls for lights, fans from a laptop or a tablet, or a smartphone. The household activities can be automated by the development of specific appliances such as water heaters to reduce the time taken to heating water for bathing and automatic washing machines to reduce work for washing clothes. In developed countries, homes are wired for electricity, calling bells, Televisions, and phones by keeping home

automation in mind. The door lock system can be implemented with help of RFID cards [2].

Another application of home automation includes, when a person enters the room, the light turns on. In improved technology, the room can sense the presence of the person. Taking in mind the day of the week, time of the day and other factors, it can also set suitable lighting timing, temperature levels, television channels and music type. In the case of a smoke detector, when fire or smoke is detected, the lights in the entire house will start blinking to alert the user about the fire. In the case of a home theatre, the home automation system can avoid disturbances by locking other audio and video components and also make an announcement. The home automation system can also call up the house owner on their mobile phone to alert them [3].

Every appliance must be connected and communicate with each other. The basic aim of home automation is to control or monitor data from different appliances [4]. A smartphone or a web browser is used to control or monitor the home automation system. The household activities such as food preservation and preparation are automated with the help of pre-packaged food or pre-made food as such in a bread toaster. Automation of handling the food is available only for standardized products [4].

Other automated activities include the air conditioner, for making it to an energy saving setting when the house is empty and get back to the normal functioning when the resident is returning back home. The system can also store the list of products, records of its usage through bar codes or an RFID tag and replaces the order automatically when it is about to exhaust. Many people today prefer smart devices which can be controlled remotely by the Internet rather than the manual control to improve the standard of living [5].

The internet of Things (IoT) is the network of any device, including home appliances, vehicles. The IoT is a combination of software, sensors, actuators, and network connectivity that helps these objects to connect and exchange data. Here, each thing can be identified by their unique identity and also able to interoperate within the existing internet infrastructure. In short, IoT is a network of the interconnection of hardware and software to transfer the data [6].

Any objects that can be given an IP address and have the ability to transfer data over the Internet bring ways to get out of many problems and to connect devices from any of the remote places [6]. The Internet of Things technology is used to bring out an innovative result and large development of smart homes to improve the living

standards of human life. The growth of the Internet of Things will change many sectors like healthcare, automation, energy, transportation, etc. Cloud computing can be used to implement the IoT infrastructure that is arranged with sensors and actuators to monitor and control devices from anywhere in the world.

An Embedded system is a computer hardware system having software in it. It can be an individual system or can be a part of another bigger system. An embedded system can be either a microprocessor or a microcontroller designed to perform a specific task [7].

Applications of home automation include,

- Lighting control
- Irrigation control
- Electronic appliances control like HVAC
- Security system
- Power consumption monitoring

The main aim of this work is to create a home automation system using voice commands, smartphone and to produce an alert alarm. This system will be useful for elderly people, children, disabled people and bedridden patients where they can control devices without any physical movements.

## II. LITERATURE SURVEY

Many systems were designed based on home automation. A smart home security system is developed using the Arduino Uno board and ESP8266 module. This system can be used for home appliances controlling using a smartphone application and to provide home security safety from unknown people. A framework for domestic automation using Node MCU and Firebase was designed to control domestic apparatus through smartphone applications. MIT app is used to control domestic apparatus [8].

Chad Davison designed an Internet of Things (IoT) based smart home security system. This system is an RF-based smart door opening system by using Elegoo Mega 2560 microcontroller board, Raspberry Pi [9]. Sandeep Mishra and Jagruti proposed a home automation system using Node MCU and Blynk app. Here the devices are controlled using mobile applications [10].

Most of the existing system uses microcontrollers such as Raspberry Pi, Arduino UNO, ARM7, NodeMCU. These are connected to devices via GSM, Ethernet, Bluetooth, or Wi-Fi. Some systems work based on the value sent by the sensors and others work based on a predefined timing or by the human command by the use of a mobile phone or a PC. These are used for various applications like irrigation control, doorbell system, security system. Most of the system works on the GSM module, which is not possible to make connections in all situations [11].

Nowadays, everything is dependent on the internet, even not any exception for our home lights and appliances. The trend of urbanization is increasing to reduce human work. Also, home automation is extending from lights to appliances and from home to factories. As smart home automation is gaining more popularity, it is

essential for making home automation user-friendly and low cost .

## III. PROPOSED SYSTEM

The rapid emergence of the Internet-of-Things (IoT) based technologies redesigned almost every aspect including Home.Home automation makes human daily life easier, it helps to conserve energy reduces human effort, where lights can be controlled with just a touch . It even helps elderly and disabled people.

The figure 1 shows the block diagram of a voice-controlled home lighting automation system. This system controls the LED and buzzer. The Microcontroller board used is NodeMCU. The Wi-Fi connectivity can be provided by NodeMCU itself.

In this system, the main components are NodeMCU and smartphones. Here, it uses Google assistance to hear and convert the voice commands and is sent to IFTTT, and it interprets them and passes them to the Blynk app from where data is sent to Node MCU and it controls the appliance based on it. In this system, LED and buzzer are controlled to demonstrate the operation of the system. Here one light can be switched ON and OFF using voice command, the remaining other light-controlled using Android mobile app and in case of any emergency, the alert can be switched ON using voice command.

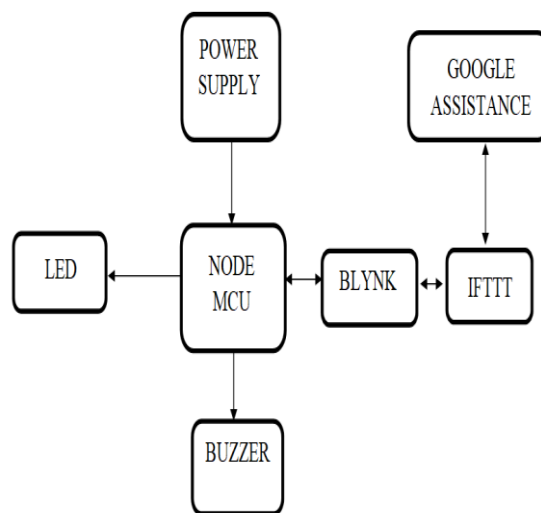


Figure 1: Block diagram

## IV. SYSTEM ARCHITECTURE

The first step is setting up the Blynk controls. Blynk is an application used to make virtual buttons which can be turned ON/OFF depending on the commands given to the Google Assistant and the next step is connecting the microcontroller and the final step is connecting to Google Assistant through IFTTT.

IFTTT is a website used to create conditional statements like if-else statements. For IF statement in IFTTT Google Assistant is selected and for THAT statement Webhooks selected. Webhooks are messages sent from applications automatically whenever any changes occur. When an event happens on trigger application that is, Google Assistant, the data about the

event is sent to Webhook URL from the action application that is, Blynk application.

## V. IMPLEMENTATION OF THE SYSTEM

### A. SOFTWARE IMPLEMENTATION

a) Arduino IDE :The system is controlled by the Arduino IDE (Integrated Development Environment) program. It consists of two functions such as Setup () and loop (). Arduino Integrated Development Environment (IDE) supports a simplified version of the C and C++ programming languages. Arduino IDE is open-source software. It can be used in different operating systems like Windows, Mac OS and Linux. They can be stored at extension of .ino.

b) Blynk application: Blynk application can control hardware, display sensor data, and store data. Blynk Server is responsible for all the interchanges between the Smartphone and devices. In this paper, this app controls home appliances like buzzer and light remotely by just clicking on and off from the app.

c) IFTTT (If This Then That) : IFTTT is a facility that is used to make simple conditional statements called applets [12]. The commands given to Google Assistant trigger the IFTTT. IFTTT interprets the data from the Google Assistant and from that IFTTT connects and controls various applications and devices .

d) Google Assistant : Google Assistant is Google's voice assistant software that is used to give voice commands [13]. It is an Artificial Intelligence based Virtual Assistant software. The users can give the voice commands through Google Assistant to access devices and applications. It is a convenient method for users to automate their devices .

e) Webhooks : Webhooks are messages sent from one app to another app that is sent automatically whenever some events occur. They can be messages or data and are sent to a specific URL which can be the app's phone number or address. They are user-specified callbacks with HTTP.

### B. HARDWARE IMPLEMENTATION

a) NodeMCU : NodeMCU is an IoT platform, Wi-Fi board[15] .They have firmware runs based on ESP8266 Wi-Fi System-on-chip and hardware based on ESP-12 module and ESP32 was formed when 32-bit MCU was added.The programming can be done using Lua language, Arduino IDE. The protocols followed by NodeMCU areIpv6, TCP, UDP, HTTP and FTP. It consists of 4MB ROM of flash storage. They can support micro-USB connection and OTA wireless uploading.It has both the features of Wi-Fi access point station and microcontroller. It uses BRT (Bias Resistor Transistor). They are easy to make into flash mode. They are widely used in automation systems because of its Wi-Fi compatibility .

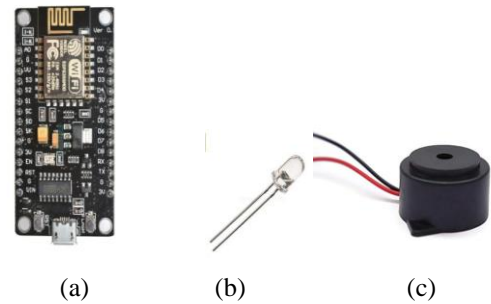


Figure 2: (a) NodeMCU (b) Two-terminal LED (c) Two-terminal Piezo buzzer

- Features of Node MCU
- Operating voltage: 3.3V
- 17 GPIO pins
- 1 analog pin
- 11 x I/O digital pins
- 1 x ADC pin
- 10-bit ADC
- Operating current: 80 mA
- Uses latest firmware version
- Frequency range: 2.4 to 2.5 GHz
- All pins have interrupt / PWM / I2C / one wire supported (except for D0)
- Good stability
- 4 Mbps communication speed

b) LED : Light Emitting Diode is a semiconductor diode. It is based on spontaneous emission. It emits light when current is passed through it. The band gap determines the energy of the light particles that are emitted by the LED. Different semiconductor materials with different band gaps produce different colors of light. In this paper, RED, BLUE, WHITE, GREEN LEDs are used. Here, long (positive) terminal is connected to the digital pin of NodeMCU and the short (negative) terminal is connected to the 1kohm resistor which is connected to the ground.

c) Buzzer :Buzzer is an audio-output producing device. This system uses a piezoelectric buzzer. They are normally used in alarm, timers and alert systems. Piezo buzzer generates beeps and tones by using piezo crystal. In this work, a buzzer is used to provide alert sound in case of any emergency situations. This buzzer can be controlled by voice comments and Blynk app.The buzzer consists of a piezo element which has a central ceramic disc surrounded with metal vibration disc. When current is given, causing the ceramic disk to expand or contract. This change in the ceramic disk causes the surrounding metal disc to expand and the sound gets produced.Here, red wire (positive) is connected to the digital pin of NodeMCU and black wire (negative) is connected to the 1k ohm resistor which is connected to the ground.

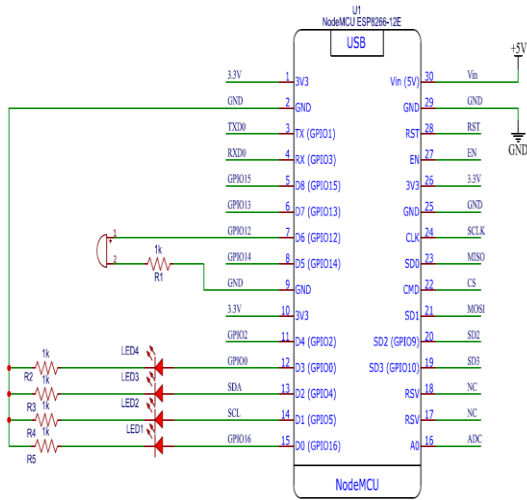


Figure 3: circuit diagram

V. RESULTS AND DISCUSSIONS

Home automation can be extended from smart doorbell, smart speaker, smart security camera, smart lock, smart smoke detector, robot vacuums cleaners, smart coffee machine, automated window to control outdoor lights and these can be controlled through a smartphone or tablet. Thus, home automation becomes integration for control of entertainment systems, climate and applications. For making home appliances fully automatic the Internets of Things (IoT) is promoted to provide improvements in smart home appliances. The internet technology is growing minute by minute and the internet connection is accessible everywhere. Human detection is achieved by the PIR sensor for security systems. The IoT is now ruling the world.

At first Blynk app is downloaded and a new project is created for four LEDs and one buzzer. Blynk and ESP 8266 libraries are downloaded and Program is created on Arduino IDE and uploaded to NodeMCU and all the connections are set up.

Then necessary applets are created in IFTTT for communication with Google assistance. Using the keyword "Hey Google/Ok Google" communication can be started. Google assistance receives the voice command, interprets it into data and checks whether the command is for IFTTT. If the data is meant for IFTTT then IFTTT receives the data. If the command is "Turn on light one", Google assistant interprets it as light one=ON. Then IFTTT receives light one = ON. IFTTT interprets the data like, if light one = ON then LED1= ON.

The interpreted data from IFTTT is then sent to the Blynk Server. From there the data is displayed on the dashboard as well as the data is sent to NodeMCU like LED1= ON. After NodeMCU receives the data, NodeMCU looks for the appliance connected to that point and the LED gets turned on automatically.

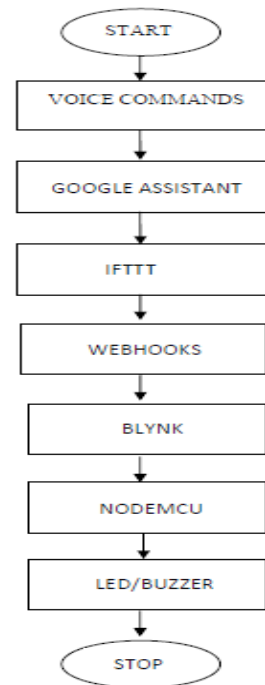


Figure 4: Flowchart

In this system Figure 5 and Figure 6, the LED and buzzer can be controlled using Google Assistance and Blynk app. The voice commands “turn on light one” and “turn on alert” are given by the user and the Google Assistant responds to those commands as shown in Figure 7. The Blynk dashboard displays LED/buzzer ON by green color and OFF by white color as shown in the Fig 8. The IFTTT applets used for communication with the Google Assistant are shown in Figure 9..



Figure 5: Simulation of Voice controlled home automation system with all light and buzzer ON

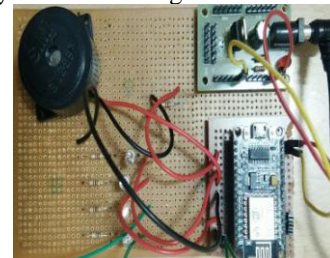


Figure 6: Prototype of Voice controlled home automation system

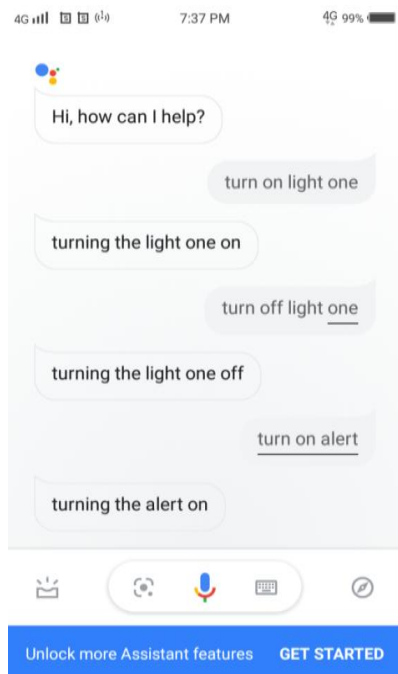


Fig 7: Google assistant responses

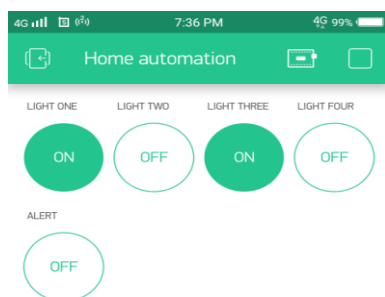


Fig 8: Blynk dashboard

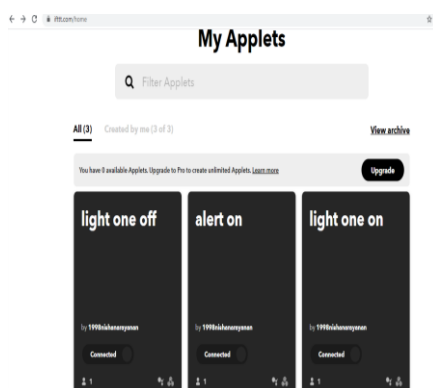


Fig 9: IFTTT Applets

The power consumption can be monitored and controlled through the data communicated with this. An adaptive scheduling algorithm can be incorporated to function equipments automatically. This will manage the energy[16].

## VI. CONCLUSION AND FUTURE SCOPE

In this paper, the proposed system provides an effective home automation system. This system brought the advantage of NodeMCU and Google assistant. The alert control help the elderly person and bedridden people to control the appliances without any physical movement just with voice comments and smartphone. The alert alarm can provide through this system.

As a part of future works, sensors like DHT11 for temperature and humidity values, an ultrasonic sensor can be included that can be used to detect people entering the building and can be used with the doorbell, security system along with the webcam based face recognition algorithm and automatic opening of doors and can also include a voice recognition algorithm. Another development that can be done is to provide an alert notification via e-mail or message.

## REFERENCES

- [1] Sharda R. Katre and Dinesh V. Rojatar, "Home Automation: Past, Present and Future", International Research Journal of Engineering and Technology (IRJET), Volume: 04, Issue: 10, Oct -2017.
- [2] A. M. Nitu, M. Jahid Hasan and M. S. Alom, "Wireless Home Automation System Using IoT and PaaS," 2019 1st International Conference on Advances in Science, Engineering and Robotics Technology (ICASERT), Dhaka, Bangladesh, 2019, pp. 1-6, doi: 10.1109/ICASERT.2019.8934676.
- [3] D. Q. R. Elizalde, R. J. P. Garcia, M. M. S. Mitra and R. G. Maramba, "Wireless Automated Fire Detection System on Utility Posts Using ATmega328P," 2018 IEEE 10th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management (HNICEM), Baguio City, Philippines, 2018, pp. 1-5, doi: 10.1109/HNICEM.2018.8666329.
- [4] Cuong Pham, Thomas Ploetz, Patrick Olivier, "Real-Time Activity Recognition for Food Preparation", IEEE International Conference on Computing and Communication Technologies, January 2010
- [5] Dileesh S, "Intelligent Air Conditioning System for Energy Management Using Wireless Control", International Journal of Innovative Science, Engineering & Technology, Vol. 1 Issue 8, October 2014.
- [6] F. Shanin et al., "Portable and Centralised E-Health Record System for Patient Monitoring Using Internet of Things(IoT)," 2018 International CET Conference on Control, Communication, and Computing (IC4), Thiruvananthapuram, 2018, pp. 165-170.
- [7] R. P. Aneesh, D. P. Vajja, P. P. Pramod, P. S. A. Kumar and A. Peethambaran, "Spacecraft command and data system simulator for the payload chaste in Chandrayaan-2 mission," 2017 IEEE International Conference on Circuits and Systems (ICCS), Thiruvananthapuram, 2017, pp. 97-102.
- [3] Amrutha R, Geetha D. Devanagavi, Prabhat Kumar Panda, "Face and Voice Assistant Based Smart Home Security and Home Automation System", International Journal of Science and Research (IJSR), Volume 9, Issue 6, June 2020.
- [4] P Bhaskar Rao and S.K. Uma, "Raspberry Pi Home Automation with Wireless Sensors using Smart Phone", International Journal of Computer Science and Mobile Computing (IJCSMC), Vol. 4, Issue. 5, pp. 797-803, May 2015.
- [5] Chinmay Bepery, Sudipto Baral, Animesh Khashkel & Farhad Hossain, "Advanced Home Automation System Using Raspberry-Pi and Arduino", International Journal of Computer

Science and Engineering (IJCSE), Vol. 8, Issue 2, Feb - Mar 2019.

[6] Satish Palaniappan, Naveen Hariharan, Naren T Kesh, Vidhyalakshimi S and Angel Deborah S, "Home Automation Systems - A Study", International Journal of Computer Applications, Volume 116, April 2015.

[7] Bhavna and Dr. Neetu Sharma, "Smart Home Security Solutions based on Internet of Things (IOT) using WIFI Interface", International Journal of Engineering Sciences & Research Technology (IJESRT), pp. 389- 397, May 2018.

[8] Shweta Singh, Shikha Verma, Surendra Kumar, Satish Kumar Singh, and Permendra Verma, "Home Automation Using Node MCU, Firebase & IOT", International Journal of Scientific Research and Review (IJSRR), Vol. 7, Issue.3, pp.1289-1294, March 2019.

[9] Mohammad Asadul Hoque and Chad Davidson, "Design and Implementation of an IoT-Based Smart Home Security System", International Journal of Networked and Distributed Computing (IJNDC), Vol.7, Issue.3, April 2019.

[10] Sandeep Mishra and Jagruti J. Mishra, "Home Automation Using Node MCU-32S and Blynk App", International Journal for Research Trends and Innovation (IJRTI), Vol.4, Issue.7, 2019.

[11] Georgewill M. Onengiye, Ezeofor J. Chukwunazo, "Design and Implementation of GSM-Based Automation of Household Appliances", International Journal of Scientific & Engineering Research (IJSER), Volume 7, Issue 5, May 2016.

[12] S. Heo, S. Song, J. Kim and H. Kim, "RT-IFTTT: Real-Time IoT Framework with Trigger Condition-Aware Flexible Polling Intervals," 2017 IEEE Real-Time Systems Symposium (RTSS), Paris, France, 2017, pp. 266-276, doi: 10.1109/RTSS.2017.00032.

[13] A. Lazić, M. Z. Bjelica, D. Nad and B. M. Todorović, "Google Assistant Integration in TV Application for Android OS," 2018 26th Telecommunications Forum (TELFOR), Belgrade, Serbia, 2018, pp. 420-425, doi: 10.1109/TELFOR.2018.8612143.

[14] A. Ghosh, S. Sinha, S. Pal and P. K. Sarkar, "Voice Over Appliance Management System," 2018 Fourth International Conference on Research in Computational Intelligence and Communication Networks (ICRCICN), Kolkata, India, 2018, pp. 188-192, doi: 10.1109/ICRCICN.2018.8718713.

[15] Espressif Systems, "ESP8266EX Datasheet", Version 6.6, 2020.

[16] Sharon N M, Aneesh R P, Rony B C, Sreedharan Embrandiri, "Adaptive Scheduling Algorithm for Centralized Building Energy Management System", IEEE 2018.

# Automatic Anesthesia Control System

Shelishiyah Raymond

*Bio medical dept*

*Vel Tech Rangarajan Dr. Sagunthala*

*R&D Institute of Science and*

*Technology*

Chennai, India

shelishiya@veltech.edu.in

Susmitha Edagottu

*Bio medical dept*

*Vel Tech Rangarajan Dr. Sagunthala*

*R&D Institute of Science and*

*Technology*

Chennai, India

susmithasreeramulu1806@gmail.com

Lasngewhun Mawblei

*Bio medical dept*

*Vel Tech Rangarajan Dr. Sagunthala*

*R&D Institute of Science and*

*Technology*

Chennai, India

lahun19maw@gmail.com

Minhajul Ahmed

*Bio medical dept*

*Vel Tech Rangarajan Dr. Sagunthala*

*R&D Institute of Science and*

*Technology*

Chennai, India

minhaj10645@gmail.com

**Abstract**—Providing painless surgery and delivering an accurate dose of anesthesia to the patient plays a very crucial role in any major surgeries. Failing in providing an accurate dose to the patient may show adverse effects and postoperative complications. In case of major surgeries which could take a longer period, the complete dosage of anesthesia could not be administered in a single dose to the patient since it may show lethal complications. Administering less dose of anesthesia makes the patient regain consciousness during the surgery. It is not easy for the anesthetist to deliver an accurate dose of anesthesia at regular intervals of time. To overcome such complications during surgeries a computer-controlled syringe is designed to deliver an accurate dose of anesthesia at regular intervals of time with constant speed. Therefore this project aims to introduce an automatic anesthesia control system integrating with monitoring parameters using Arduino Uno. This Embedded system uses a syringe pump to deliver the right amount of anesthesia to the patient. The anesthetist can set the desired amount of anesthesia that can be given to the patient with the help of a switch panel. Once the Arduino Uno receives the signal it activates the motor driver to drive the syringe pump at the preset intervals. The anesthesia is delivered to the patient according to the rotation of the stepper motor. After administration of anesthesia, the vital parameters like Temperature, Exhalation breath temperature, and Pulse are monitored side by side. If they are under the normal state then the second dose of anesthesia will be injected. On the onset of abnormality the doctor will be notified through a buzzer and anesthesia delivery would continue only if everything is under normal. Additionally, these parameters are checked by corresponding sensors. This integration of monitoring parameters increases the patient's safety and keeps the anesthesiologists at ease.

**Keywords**—Automatic, Embedded system, Arduino, Exhalation Breath Temperature (EBT), Vital parameters, computer controller.

## I. INTRODUCTION (ANESTHESIA)

Anesthesia is a medical drug that can be given to the patient to avoid pain during surgeries. It is the greatest discovery in the medicine which provides comfort to both patient and doctor while performing surgeries. Anesthesia is divided into three types General, Regional and Local Anesthesia. General anesthesia<sup>[10][11]</sup> is mainly used for all major surgeries to make the patient unconscious, Local<sup>[11]</sup>

can be given to the numb small area of the body whereas Regional<sup>[11]</sup> can be given to block pain in an area of the body such as an arm or leg. Delivery of anesthesia can be given to the patients in two ways either through inhalation or intravenous. Giving intravenous anesthesia<sup>[6][7]</sup> shows a quick response in patients. Anesthesiologists are overloaded with multiple tasks like to deliver anesthesia continuously to maintain unconsciousness till the end of surgery and also need to monitor physiological parameters side by side. So the anesthetist may fail to administer the accurate dose of anesthesia to the patient. Inducing accurate dose plays a crucial role because high or low dose may show adverse effects in patients like coma in case of high and coming to conscious during surgery in case of low. The main aim of this project is to deliver an accurate dose<sup>[9]</sup> of anesthesia along with physiological parameters. Programmed input is given to the Arduino to make the stepper motor rotate in desired speed. Once the stepper motor initiates it makes the Syringe pump move in forward and backward directions to inject the anesthesia into the patient's body. Body parameters are monitored side by side with specified sensors to overcome complications in vital parameters while delivering anesthesia. Physiological parameters need to be monitored according to American society of Anesthesiologists after giving anesthesia such as Temperature, SpO<sub>2</sub>, Heart rate, Blood pressure and providing adequate ventilation. The rotary motion of the stepper motor initiates the Syringe Pump to move. An electric buzzer is connected to these parameters to give alert to the anesthesiologist if any abnormal changes occur in these vital parameters. The main reason of automatic anesthesia control system is to avoid carrying extra monitoring equipment in any emergency situations or rural areas. So the proposed work integrates monitoring parameters with anesthesia system. One of the major advantages of Total intravenous anesthesia<sup>[1][8]</sup> over inhalation<sup>[1][8]</sup> anesthesia is that emergency surgeries in situations like COVID19<sup>[2]</sup> or any air-borne diseases it helps to prevent the spreading of the disease.

## II. MATERIALS AND METHODS

### A. Materials

Arduino UNO: Arduino UNO shown in figure 1 is a standard microcontroller board based on microchip



ATmega328P. It is equipped with the sets of digital and analog input/output (I/O) pins. It has 14 digital input/output pins, 6 Analog pins, USB connector, Power port, Microcontroller, Reset Switch, Crystal Oscillator, USB(Universal Serial Bus) interface chip, TX(Transmit) RX(Receive) LEDs. It can be connected to a computer with a USB cables. It can operate on an external supply from 6 to 20 volts.



Fig. 1. Arduino UNO

LM35 Temperature Sensor: LM35 Temperature sensor shown in figure 2 is used to detect the body temperature<sup>[5]</sup>. It can measure temperature in the range of -55 degree Celsius to +150 degree Celsius. It is an integrated circuit temperature sensor where the output voltage is linearly proportional to the temperature in Celsius. It operates from 4 to 30 volts and it can be operated by single power supplies. The LM35 temperature sensor is appropriate for remote applications and is cost-effective.

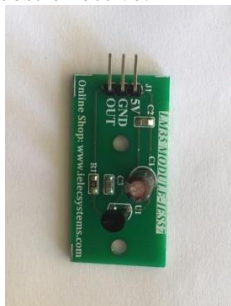


Fig. 2. LM35 Temperature sensor

NTC (Negative temperature coefficient) thermistor: The thermistor used to measure Exhalation breath temperature (EBT)<sup>[4]</sup> which is the Negative temperature coefficient (NTC) thermistor shown in figure 3. In this type of thermistor, the resistance decreases with an increase in temperature. Generally, these NTC (Negative temperature coefficient) thermistors are used for low-temperature measurement and are typically suitable for a temperature range between -55°C to +150°C. They are reliable and responds quickly and inexpensive. This device detects the change in temperature when the patient breathes through the mask. It comprises a buzzer which beeps when the patient's temperature is less than 34°C.



Fig. 3. NTC Thermistor

Pulse sensor: Pulse sensor shown in figure 4 measures Beats per minute (BPM) by placing a finger on the led. It is bilateral, on one side the LED (Light emitting diode) is placed and on the other side, noise elimination circuitry is placed. The finger is placed over the LED and the light emits from the LED will fall directly on the vein which will monitor the flow of blood as well as the heartbeats. The circuitry is responsible for noise cancellation and amplification. It works with either a 3V or 5V.



Fig. 4. Pulse sensor

Stepper motor (NEMA 17): Stepper motor<sup>[13][16]</sup> drives a full rotation into a number for steps which helps to deliver the drug in the required amount with the smooth motion for the syringe pump. Stepper motor rotates with steps. This works on the principle of electromagnetism. Using the driver in different step mode and with a capacitor of 100µF for decoupling and 12 v for powering the motor. Wire A and C of the stepper motor connected to pins 1A and 1B and the B and D wires to the 2A and 2B pins. Before connecting the motor the current limit of the motor driver<sup>[17]</sup> is set so that the current is within the current limit of the motor. Figure 5 shows stepper motor.



Fig. 5. Stepper motor

B. Methodology

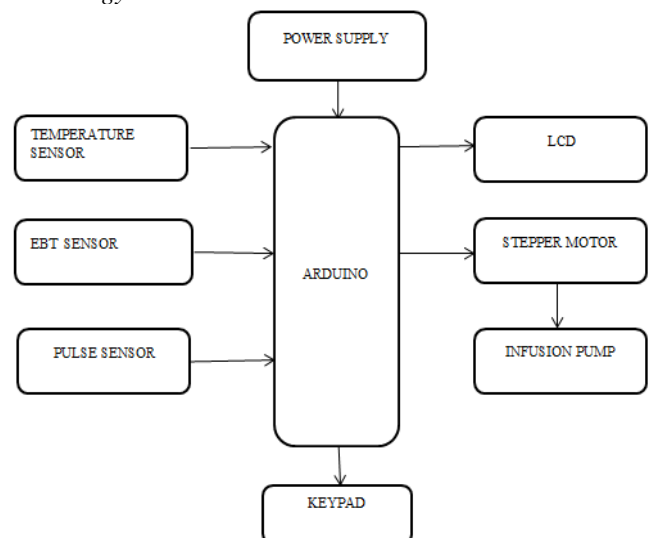


Fig. 6. Generalized block diagram of Automatic anesthesia control system



## V. OBJECTIVE

- To integrate monitoring system along with the anesthesia machine
- To make it applicable in emergency situations
- To develop a cost-effective delivery system
- To design a system for delivering anesthesia at constant rate
- To reduce the human error

## VI. DISCUSSION

Physiological readings have been obtained using different sensors. LM35 module which is a temperature sensor measure temperature by placing it under arm. Readings obtained from LM35 is 2-3 degree Celsius different from Medtech Digital thermometer. Thermistor which is a resistor whose resistance strongly depends on temperature. So thermistor is used to measure the temperature of respiration while inhaling and exhaling by placing it in the N95 mask. Pulse of the patient plays very crucial role when patient is under anesthesia. It is one of the vital parameter that changes easily when there occurs any changes in the internal environment of the patient. Pulse sensor is used to measure pulse reading. A buzzer is connected to these sensors to beep if there is any abnormality in the patient's vital parameters while delivering anesthesia. Physiological parameters of non-healthy subjects are yet to be measured.

## VII. FUTURE ENHANCEMENT

This project can be improved further by integrating few more monitoring parameters like SpO<sub>2</sub>, Blood pressure sensor. It can also be improved to control the motor with anyone of the parameter by making it stop if any abnormality occurs in patient's vital signs.

## VIII. CONCLUSION

From the overall conduct of the project, monitoring patient's parameters plays very crucial role when patient is under anesthesia. It's not easy for anesthesiologists to focus on multiple tasks. So Embedded patient monitoring system has been designed with different sensors to give alert to anesthesiologist if any abnormality occurs while delivering anesthesia to the patient via syringe pump which in turn controlled by stepper motor.

## REFERENCES

- [1] David Miller, Sharon R Lewis, Michael W Pritchard, Oliver J Schofield-Robinson, Cliff L Shelton, Phil Alderson, Andrew F Smith, "Intravenous versus inhalational maintenance of anaesthesia for postoperative cognitive outcomes in elderly people undergoing non-cardiac surgery", *Cochrane Database Syst Rev.*, vol. 8, PMID:PMC6513211, August 2018.
- [2] Tushar Chokshi, Shivakumar Channabasappa, Davies C Vergheese, Sukhminder Jit Singh Bajwa, Bhavna Gupta, Lalit Mehdiratta, "Re-emergence of TIVA in COVID times", *Indian journal of Anaesthesia*, vol. 64, pp. 47-53, May 2020.
- [3] S. Krishnakumar, Bethanney Janney, "Automatic anaesthesia regularization system (AARS) with patient monitoring modules", *International Journal of Engineering and Technology*, vol. 7, pp. 48-52, January 2018.
- [4] Todor A Popov, Tanya Z Kralimarkova, Marina Labor and Davor Plavec, "The added value of exhaled breath temperature in respiratory medicine", *Journal of Breath Research*, vol. 11, 034001, August 2017.
- [5] Barkha Bindu, Ashish Bindra and Girija Rath, "Temperature management under general anesthesia: Compulsion or option", *J Anaesthesiol Clin Pharmacol*, vol. 33, pp. 306-316, Jul-Sep 2017.
- [6] K. S. Aujla, Manbir Kaur, Ruchi Gupta, Sukhjinder Singh, Bhanupreet, and Tavleen, "A study to compare the quality of surgical field using total intravenous anesthesia (with propofol) versus inhalation anesthesia (with isoflurane) for functional endoscopic sinus surgeries", *Anesth Essays Res.*, vol. 11, pp. 606-610, Jul-Sep 2017.
- [7] A. F. Nimmo, A. R. Absalom, O. Bagshaw, A. Biswas, T. M. Cook, A. Costello, S. Grimes, D. Mulvey, S. Shinde, T. Whitehouse, M. D. Wiles, "Guidelines for the safe practice of total intravenous anesthesia (TIVA)", *Joint Guidelines from the Association of Anesthetists and the Society for Intravenous Anesthesia*, vol. 74, pp. 211-224, February 2019.
- [8] Abd-Elazeem Elbakry, Wesam-Eldin Sultan, Ezzeldin Ibrahim, "A comparison between inhalational (Desflurane) and total intravenous anesthesia (Propofol and Dexmedetomidine) in improving postoperative recovery for morbidly obese patients undergoing laparoscopic sleeve gastrectomy: A double-blinded randomised controlled trial", *J Clin Anesth*, vol. 45, pp. 6-11, March 2018.
- [9] Lina Jia, Jiachen Hou, Haibo Zheng, Lihua Sun, Yingying Fan, Xu Wang, Mingyue Hao, Yue Li, Tongwei Yang, "Study of the rational dose of propofol in elderly patients under bispectral index monitoring during total intravenous anesthesia", *Medicine(Baltimore)*, vol. 99:e19043, Jan 2020.
- [10] Guerin Smith, Jason R. D'Cruz, Bryan Rondeau, Julie Goldman, "General anesthesia for surgeons", *PMID:29630251*, Review, August 2020.
- [11] Christian Seger, Maxime Cannesson, "Recent advances in the technology of anesthesia", *F1000Res.*, vol. 9, Rev-375, May 2020.
- [12] Elias Mansour, Rotem Vishinkin, Stephane Rehet, Walaa Saliba, Falk Fish, Patrice Sarfati, Hossam Haick, "Measurement of temperature and relative humidity in exhaled breath", *Sensors and Actuators B:Chemical*, 127371, vol. 304, February 2020.
- [13] Lili Zhang, Li Liu, Jing Shen, Jiahui Lai, Kesen Wu, Zhao Zhang, Jian Liu, "Research on stepper motor motion control based on MCU", *IEEE, Chinese Automation Congress (CAC)*, pp. 3122-3125, Oct 2017.
- [14] Nour Merhi, Nour Mohamad, George HajjMoussa, Ahmad ElSayed, Saeed H. Bamashmos, Lara Hamawy, Mohamad HajjHassan, Mohamad Abou Ali, and Abdallah Kassem, "An Intelligent Infusion Flow Controlled Syringe Infusion Pump", *31<sup>st</sup> International Conference on Microelectronics(ICM)*, pp. 48-52, December 2019.
- [15] Akash K, Sangavi S, Venkatesan M, "Double acting syringe pump using a Rack and Pinion mechanism - Simulink model", *IEEE International Conference on Computational Intelligence and computing Research(ICCIC)*, December 2016.
- [16] Zhang, Deode, Wang, Jingqi, Qian, Lei, Yi, Jun, "Stepper motor open-loop control system modelling and control strategy optimization", *Archives of Electrical Engineering*, vol. 68, pp. 63-75, 2019.
- [17] Md. Mahedi Hasan, Md. Rokonzaman Khan, Abu Tayab Noman, Humayun Rashid, Nawsher Ahmed, S.M. Taslim Reza, "Design and Implementation of a Microcontroller based low cost computer numerical control(CNC) plotter using motor driver controller", *International Conference on Electrical, Computer and Communication Engineering (ECCE)*, April 2019.

# Design of a Medical Prototype Robot for Nurse Assistance

Steeve Shibu Chempolil  
Dept. of Biomedical  
Engineering  
Vel Tech Rangarajan Dr.  
Sagunthala R & D Institute  
of Science and Technology  
Chennai, India  
stevechempolil1@gmail.com

Renie Melvinia  
Basaiaiwmoit  
Dept. of Biomedical  
Engineering  
Vel Tech Rangarajan Dr.  
Sagunthala R & D Institute  
of Science and Technology  
Chennai, India  
reniemelvinia@gmail.com

Sneha Saji  
Dept. of Biomedical  
Engineering  
Vel Tech Rangarajan Dr.  
Sagunthala R & D Institute  
of Science and Technology  
Chennai, India  
snehasaji555@gmail.com

Karthik Raj V  
Dept. of Biomedical  
Engineering  
Vel Tech Rangarajan Dr.  
Sagunthala R & D Institute  
of Science and Technology  
Chennai, India  
karthickrajv@veltech.edu.in

**Abstract**— Robotics in the medical field is an ongoing trend in both research and commercial sectors. Robots are used in every hospital department for assistance in delivering things, surgery, checking vital signs, telepresence, etc. Medical prototype robot is a scenario-based assistive robot with a customized design to help the hospital staff fight against Covid-19 (Coronavirus disease) outbreak, ensuring social distancing. It has basic features like delivering medicine and small handheld devices, remote temperature sensing using IR (infrared) and UVC (ultraviolet type C) disinfection unit. The main aim of this prototype is to make the nurse not to handle the devices which was handled by the patients in which we can convey the information through an audio system (which is already available in the hospital) or a nurse will be assisting the initial instructions required (by ensuring the social distance) that is in the isolated ward so that the patient can do the task properly. For this prototype, we are using the basic microcontroller, that is, Arduino UNO. We were successful in taking readings with the help of a temperature sensor and were able to supply power to the UVC lamp in which it sterilized the objects inside the unit when it was exposed for 2-3 minutes. And finally, the robot was able to move successfully with the help of Arduino and Bluetooth setup.

**Keywords**— Robotics, Covid-19, telepresence, Arduino, Bluetooth, IR, UVC.

## I. INTRODUCTION

Medical prototype robot is an assistive robot; the need for this robot is to reduce the hospital staff workload such as doctors, nurses, cleaning staff and other services [1]. The services include delivering hospital documents to various departments, food and medicine to patients, blood samples to the laboratory and even checking the vital signs of patients, as shown in Fig. 1.

This is a scenario-based robot specially designed for the Covid-19 (Coronavirus disease) pandemic as it ensures non-contact support and assistance to patients to avoid the spread of the virus. Viruses such as COVID-19 become a threat to humankind's health globally [2]. The wheels attached enables mobility and the UV-C (ultraviolet type C)

sterilization unit ensures the disinfection of probes and small hand-handled device. Additional vital sign measurements include temperature sensing, which follows a non-contact infrared temperature sensor for body temperature measurement preferred in most hospitals [3]. Even though temperature measurement is the oldest method to diagnose, but it still plays a vital role as an indicator for diagnosing disease [4].

Several studies in recent years have investigated robots' social acceptance in various close contact situations with humans. For instance, the deployment of robots in the healthcare workflow across the continuum of care goes from prevention, screening, and diagnosis to treatment and home care [5]. The Psychosocial Aspects in Robotics (PAIR) Lab aims to study psychological and social factors relevant to robots' application, such as the views of stakeholders across the globe [6].

We designed this medical prototype robot based on the hospital's challenges in this pandemic situation of Covid-19. One of the most critical areas is really challenging inside the quarantine buildings or houses and isolation wards in hospitals [7]. As you can see from the Fig. 1, when we give inputs from the controlling device, our medical robot moves towards the patient and guides the patient accordingly. For instance, if we want to get the patient's temperature reading, the robot will go near and get the readings without any direct contact with the patient helping the doctors treat the patient accordingly. In this medical prototype robot, we also incorporated a UVC sterilization unit to sterilize small disposables or handheld devices. UVC light is widely used in healthcare facilities for disinfecting healthcare equipment, probes, surfaces, and operating rooms [8]. The ultraviolet light of wavelength 280 nm (nanometer) damages the DNA (deoxyribonucleic) and RNA (ribonucleic acid) of the virus and prevents them from multiplying further [9].

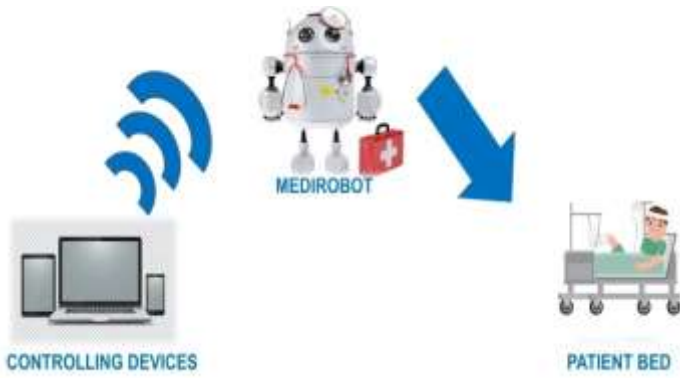


Fig. 1. Basic diagram of medical prototype robot.

## II. MATERIALS AND METHODOLOGY

### A. Block Diagram

Fig. 2 shows the overall connection from Arduino board to the IR temperature sensor, Bluetooth module, motor and battery.

It communicates with USART (Universal Synchronous Asynchronous Receiver Transmitter) at a 9600 baud rate. Using this Bluetooth, we can move the robot even without using Bluetooth beacons for wayfinding [10]. Bluetooth can be used only within some distance since the range for communication is of short-range [11].

The Arduino is used for debugging the code into the board for the mobility of the robot. It is also used to interface with the temperature sensor, where the code for the temperature measurement is debugged into the Arduino to measure the temperature based on the object as well as the ambient temperature.

The sensor is powered from the Arduino, which is 3 volts, but since it has a regulator, it can be used with 5 volts, which will get regulated to 3 volts.

The UVC is mainly used for disinfecting the probe to prevent contamination. UVC can be harmful when a person is exposed to it [12]. Therefore, the nurse has to carefully check that the patient keeps the probe inside the robot for the disinfection process.

In the UVC sterilization unit, either we can use LEDs (Light emitting diodes) or lamps that produce UVC light. Power consumption is lesser in UVC LEDs when compared to lamps. Additionally, LEDs require a battery which makes them portable, while the UVC lamps have to be plugged into the power source.

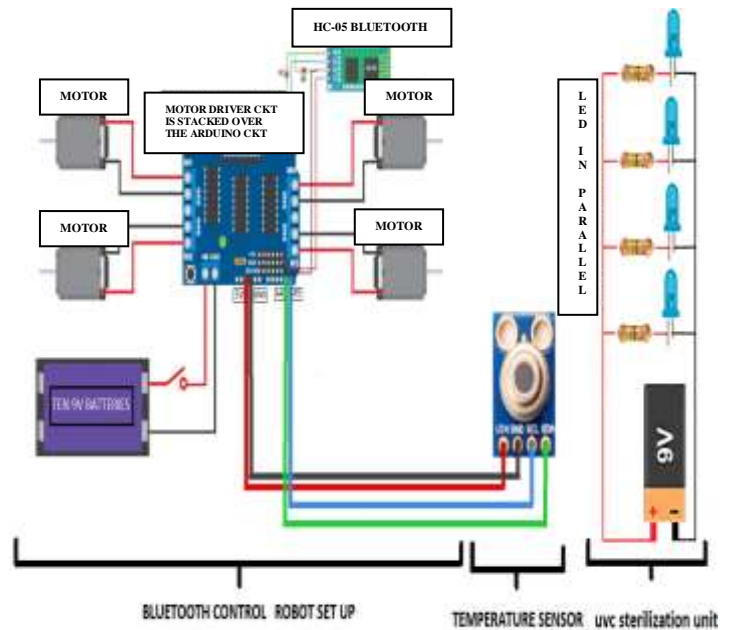


Fig. 2. The detailed block architecture of the system.

The IR temperature sensor is mainly used to measure the body temperature of a patient, where it will be attached to the hand of the robot and for taking the measurements, the robot will have to move closer to the patient. An internal state machine controls the temperature sensor.

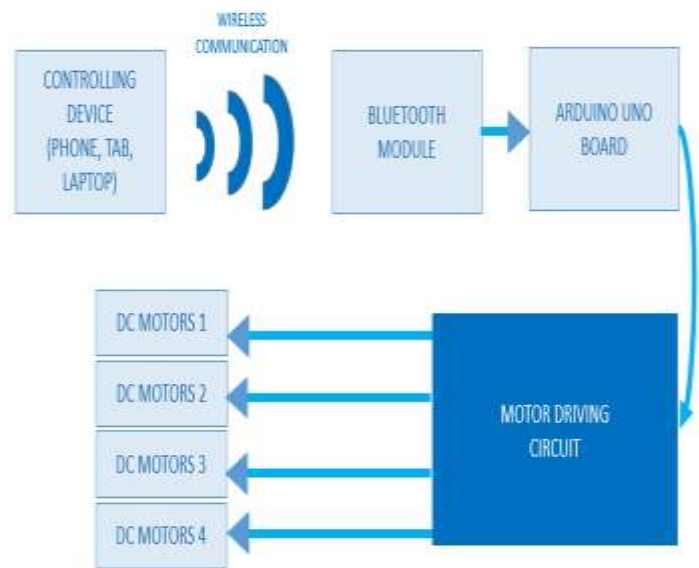


Fig. 3. Block diagram of remote controlling.

The movement of the robot is controlled with the help of the tablet, PC and laptop. The controlling device will be interfaced with Arduino in which the coding is fed into the Arduino and stack with the motor shield, as shown in Fig. 3.

B. Flowchart of Software Part

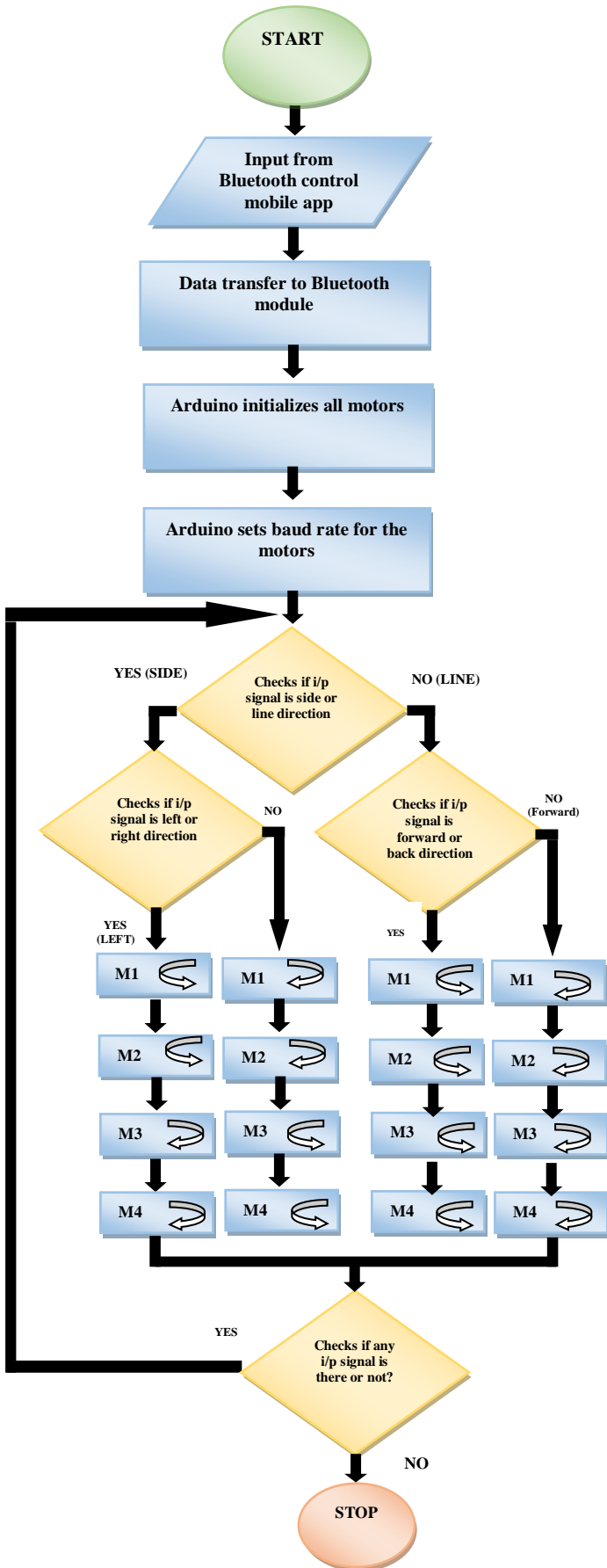


Fig. 4. Flow chart for remote control movement.

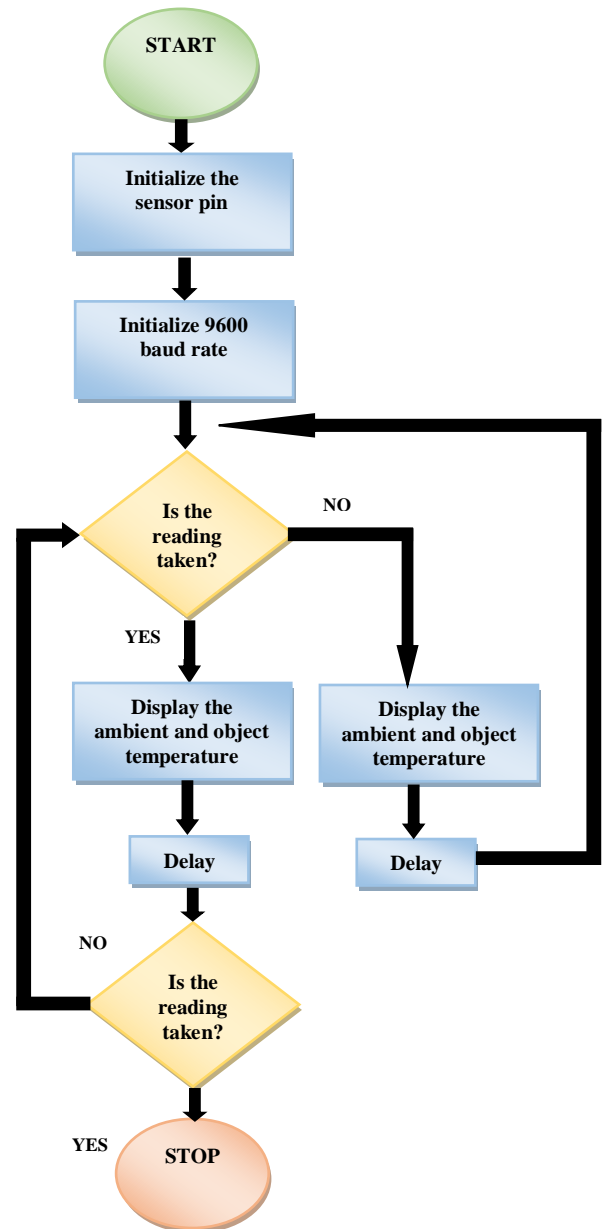


Fig. 5. Flow chart for temperature sensor.

Fig. 4 shows the outlook of how the remote control movement works based on the Arduino code, where M1, M2, M3, M4 represent the four motors of the robot. The code first checks whether the input signal is either line or sidewise direction. If it is side direction, it again checks whether it is left or right direction and enables each motor to move either clockwise or anti-clockwise as shown in Fig. 4. In similar manner it checks for line direction and enables each motor to move either clockwise or anti-clockwise direction depending on input signals as shown in Fig. 4. Fig. 5 shows the outlook of how the sensor works based on the Arduino code starting from sensor's initializing to the displaying of the reading taken by the sensor. And if the sensor takes the reading, it will display it and check whether the procedure is complete or not; if yes, then it will stop the procedure.

### III. RESULT

#### A. Temperature Sensor

TABLE I. REAL TIME READINGS OF TEMPERATURE SENSOR

No. of sub	Sensor Reading at Different Distance (cm)						Thermometer Reading (°C)	Error (°C)
	1 cm	2 cm	4 cm	6 cm	8 cm	10 cm		
1	32.97	32.85	28.45	27.67	26.55	26.21	35.09	2.12
2	33.93	33.91	30.05	28.65	27.79	27.27	35.05	1.12
3	32.75	31.41	29.43	26.24	25.16	24.13	36.09	3.12
4	32.59	32.21	29.35	26.27	25.45	24.3	35.05	2.46
5	33.21	33.15	28.95	26.35	25.35	25.53	35.03	1.82
6	31.57	31.57	26.31	23.45	22.91	21.99	35.09	3.52

Table I represents the sensor temperature readings taken from the forehead of the subject (sub) in which the sensor reading varies according to the distance between the sensor and the subject. At the same time, the clinical thermometer is measured from the underarm. Compared with the clinical thermometer, the sensor reading shows an error value of  $\pm 3$  degrees Celsius.

#### B. Ultraviolet Type C

The UVC light is germicidal and could sterilize with an exposure time of 2 minutes. We know that the dose relates to the irradiance as shown in (1):

$$H = E * t \quad (1)$$

where H is the UV dose in joule per square centimeter ( $J/cm^2$ ), E is the irradiance in watt per square centimeter ( $W/cm^2$ ), and t is the time in seconds. Therefore, the exposure time for sterilizing the probes inside the box of dimension 25 x 20 x 18 centimeter with four lamps of 15 watts each can be calculated by rearranging (1),

$$t = H / E \quad (2)$$

Taking UV dose to irradiate Covid-19 as  $6600 J/cm^2$  and irradiance as  $60 W/cm^2$ , we get an exposure time of 110 seconds approximately, 2 minutes [13]. Fig. 6 shows the setup of the ultraviolet type C sterilizing unit.



Fig. 6. Ultraviolet type C sterilizing unit.

#### C. REMOTE CONTROL ROBOT

The Arduino based robot was able to move around successfully with the help of controlling devices (for example, phone, tab and laptop) wirelessly through the Bluetooth module. Fig. 7 shows the pin connections with the Bluetooth module of the remote control robot. Fig. 8 shows the front panel of the mobile app for controlling the robot. Fig. 9 shows prototype setup of the final remote control robot.



Fig. 7. Pin connections with Bluetooth module of the remote control robot.

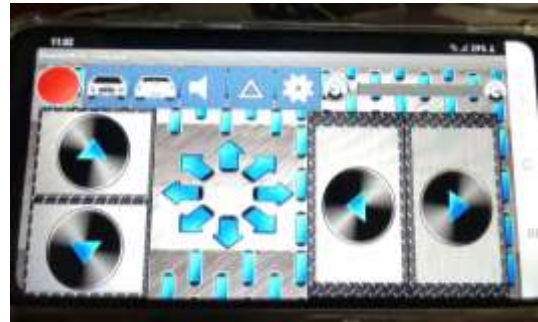


Fig. 8. Mobile app for controlling the robot.

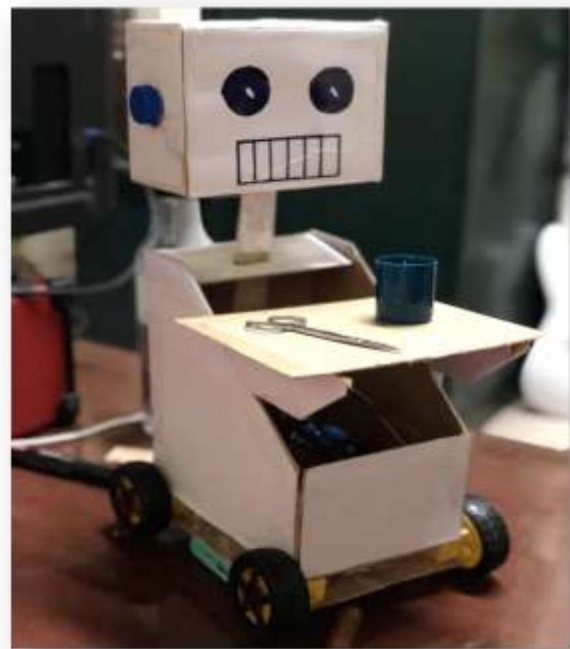


Fig. 9. Prototype setup of the final remote control robot.

#### IV. DISCUSSION

This prototype is an integration of 3 main features, that is, the temperature reading, UVC and robot movement. The temperature sensor we designed was a non-contact type in which we could successfully take the readings from different subjects. The sensor temperature reading was found to have an error of  $\pm 3$  degrees Celsius difference from that of the clinical thermometer. It was observed that the sensor could give a precise reading only when it is less than 3 cm away from the object that needs to be measured. In most of the existing designs, the temperature sensing robot has a drawback by entering the readings manually [14].

Ultraviolet type C has a wavelength from 100 to 280 nm, which was found to be more effective for sterilization of small handheld devices. The sterilization unit is incorporated into the robot without having a separate unit for sterilization. The UVC lamp could sterilize the probes inside the UVC unit. The existing designs of UVC robots only serve the purpose of sterilization while we have the sterilization unit incorporated inside the robot [15].

And finally, the robot could be moved remote controllably with the help of Bluetooth module with integrated features. But in most of the existing designs the robots are mainly used only for spillage cleaning and for delivering items but had no vital sign measurement unit or sterilization unit [16].

#### V. CONCLUSION

This is a low-cost medical prototype robot that is designed using readily available components from the Indian market. This robot can help the nurse in this pandemic situation by gathering the temperature information (which is a vital signal used for screening the Covid-19 patients) and sterilizing small handheld devices to prevent the virus from spreading. In the future, more vital signs can be integrated with this system, such as SpO<sub>2</sub>; its probe can be easily sterilized using our inbuilt feature.

#### ACKNOWLEDGMENT

We want to thank the faculty and the Department of Biomedical Engineering, Vel Tech Rangarajan Dr. Sagunthala R & D Institute of Science and Technology, for their constant support, guidance and resources. We express our gratitude to Vel Tech Rangarajan Dr. Sagunthala R & D Institute of Science and Technology for providing the facilities and infrastructure needed for the performance of the research work.

#### REFERENCES

- [1] A. Amal, A. Jory, A. Sadeem, A. Shahad, A. Sara, and B. Ahmed, "Design and implementation of a nurse robot," in *Proceedings of the International Conference on Industrial Engineering and Operations Management*, pp. 1309-1319, March 2020.
- [2] Z. Zeng, P. J. Chen, and A. A. Lew, "From high-touch to high-tech: COVID-19 drives robotics adoption," *Tourism Geographies*, vol. 22, pp. 724-734, May 2020.
- [3] M. Zukowski, K. Matus, E. Pawluczuk, M. Kondratiuk and L. Ambroziak, "Patients temperature measurement system for medical robotic assistant," in *AIP conference proceedings*, vol. 2029, pp. 020084-1-020084-8, Oct 2018.
- [4] N. H. Wijaya, Z. Oktaviahandani, K. Kunal, E. T. Helmy, and P. Nguyen, "Tympani thermometer design using passive infrared sensor," *Journal of Robotics and Control (JRC)*, vol. 1, pp. 27-30, Jan 2020.
- [5] A. Di Lallo, R. Murphy, A. Krieger, J. Zhu, R. H. Taylor and H. Su, "Medical Robots for Infectious Diseases: Lessons and Challenges from the COVID-19 Pandemic," *IEEE Robot. Autom. Mag.*, Jan 2021, in press.
- [6] E. A. Makky, *New Ideas Concerning Science and Technology*, 1st ed., vol. 3. Book Publisher International, 2020, pp. 26-33.
- [7] J. Willan, A. J. King, K. Jeffery, and N. Bienz, "Challenges for NHS hospitals during covid-19 epidemic," *BMJ*, vol. 368, pp. 1-2, March 2020.
- [8] C. S. Heilingloh, et al. "Susceptibility of SARS-CoV-2 to UV irradiation," *Am. J. Infect. Control*, vol. 48, pp. 1273-1275, Oct 2020.
- [9] P. Chanprakon, T. Sae-Oung, T. Treebupachatsakul, P. Hannanta Anan and W. Piyawattanametha, "An Ultra-violet sterilization robot for disinfection," in *2019 5th International Conference on Engineering, Applied Sciences and Technology (ICEAST)*, pp. 1-4, July 2019.
- [10] Yao-Jen Chang, Yen-Yin Chu, Chen-Nien Chen and Tsen-Yang Wang, "Mobile computing for indoor way finding based on bluetooth sensors for individuals with cognitive impairments," in *2008 3rd International Symposium on Wireless Pervasive Computing*, pp. 623-627, May 2008.
- [11] J. C. Haartsen and S. Mattisson, "Bluetooth-a new low-power radio interface providing short-range connectivity," in *Proceedings of the IEEE*, vol. 88, pp. 1651-1661, Oct 2000.
- [12] K. Cresswell and A. Sheikh, "Can Disinfection Robots Reduce the Risk of Transmission of SARS-CoV-2 in Health Care and Educational Settings?," *J. Med. Internet Res.*, vol. 22, pp. e20896, Sep 2020.
- [13] A. Baluja, J. Arines, R. Vilanova, J. Cortiñas, C. Bao-Varela, and M. T. Flores-Arias, "UV light dosage distribution over irregular respirator surfaces. Methods and implications for safety," *J. Occup. Environ. Hyg.*, vol. 17, pp. 390-397, Sep 2020.
- [14] B. Hassan and X. Xiaolan, "A decision framework for operation management of reconfigurable mobile service robots in hospitals," *IFAC Proceedings Volumes*, vol. 42, pp. 151-156, June 2009.
- [15] M. Guettari, I. Gharbi, and S. Hamza, "UVC disinfection robot," *Environ. Sci. Pollut. Res.*, vol. 98, pp. 434-441, Oct 2020.
- [16] E. Broadbent, J. R. Orejana, H. S. Ahn, J. Xie, P. Rouse and B. A. MacDonald, "The cost-effectiveness of a robot measuring vital signs in a rural medical practice," in *2015 24th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, pp. 577-581, Sep 2015.



# A Survey on Diabetic Retinopathy Disease Detection and Classification using Deep Learning Techniques

Valarmathi S

Department of Computer Science  
Avinashilingam Institute of Home Science and Higher Education  
for Women  
Coimbatore, India  
[valarmathis313@gmail.com](mailto:valarmathis313@gmail.com)

Vijayabhanu R

Department of Computer Science  
Avinashilingam Institute of Home Science and Higher Education  
for Women  
Coimbatore, India  
[vijayabhanu.r@gmail.com](mailto:vijayabhanu.r@gmail.com)

**Abstract**—Diabetes is the most commonly found chronic disease seen in many people of different age groups with poor insulin production, which causes high blood sugar. Diabetes, when left untreated, can lead to the development of several diseases across the body. Diabetic Retinopathy (DR) is an asymptomatic eye disease induced by diabetes that results in damaged retinal vessels. Many automatic diagnostic systems have been developed in the literature in which conventional handcrafted features were used. With the development of Deep Learning (DL), particularly in medical imaging, more accurate and potential results are produced, as it performs automatic feature extraction. Convolutional Neural Networks (CNNs) are the most widely used deep learning method in medical image analysis. In this paper, several Deep Learning-based diabetic retinopathy disease detection and classification techniques are analyzed and reviewed for better understanding.

**Keywords**—Convolutional Neural Network (CNN), Deep Learning, Diabetes, Diabetic Retinopathy, Medical image analysis

## I. INTRODUCTION

Diabetes is one of the highly prevalent global diseases, inducing adverse effects on various human parts, which shows micro or macrovascular changes. Nearly 382 million people were identified to be diagnosed with diabetes, and it may rise up to 592 million as per reports. Diabetes over period can lead to various ocular diseases such as Diabetic Retinopathy (DR), Glaucoma, Diabetic macular edema, Cataracts, and so on. Diabetic patients are greatly susceptible to DR which harms the retinal vessels, which in turn over period may cause complete vision loss. By proper screening and regular check-up, nearly 90% of patients can be diagnosed, and future consequences can be minimized. The massive problem lies here is that DR is mostly an asymptomatic eye disease which does not show unique symptoms until an end most stage is reached. However, manual inspections of retinal image features are difficult and demanding tasks. To overcome this problem, various automated diagnostic systems were developed in recent past, which supports ophthalmologists in examining retinal abnormalities.

This paper is organized as: the section 2 provides a brief note on Diabetic Retinopathy, types and its symptoms. Deep Learning and CNN concepts are explored in section 3. In section 4, the literature survey has been done to get a clear knowledge on the existing research works. The section 5

concludes the paper and section 6 provides a direction with the examined inferences.

## II. DIABETIC RETINOPATHY

Diabetic patients widely suffer from Diabetic Retinopathy (DR), which mostly does not show visible symptoms at the earlier stage and may result in blindness. DR affects the eye sight and harms the blood vessels in and around the retinal area. The Normal retina and DR-affected retina are shown in Fig. 1 and Fig. 2 respectively.

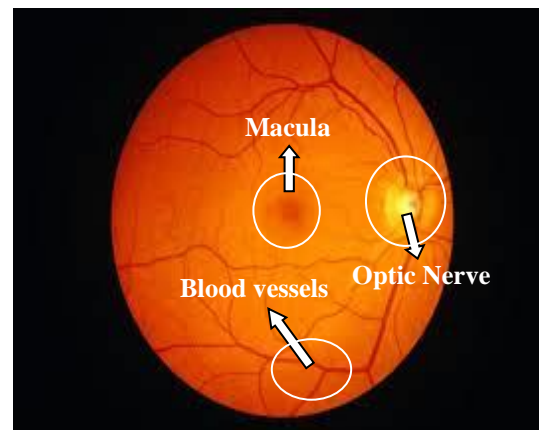


Fig. 1 . Normal Retina

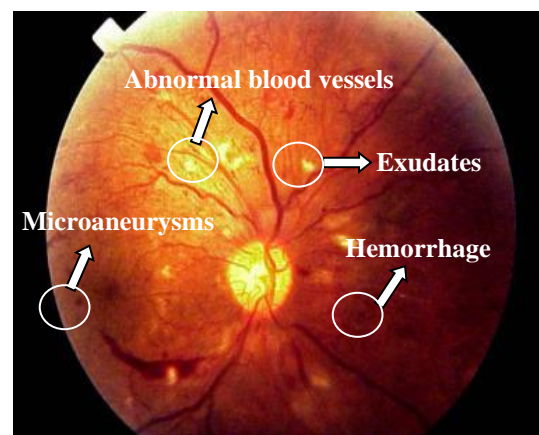


Fig. 2 . DR-affected Retina

According to the morphological variations in the color fundus images, Diabetic Retinopathy has been categorized into two types, Non-proliferative Diabetic Retinopathy (NPDR) and Proliferative Diabetic Retinopathy (PDR).

Non-proliferative diabetic retinopathy (NPDR) happens to be a primary stage DR retinal disease which exhibits some of the symptoms such as,

- **Microaneurysms (MA)** are one of the primary signs of DR, commonly found in the macular area, which emerges as red spots and less than  $125\mu\text{m}$  in size having sharp margins.
- **Hemorrhages (HM)** happen to be larger spots and greater than  $125\mu\text{m}$  in size. Blot HM and Superficial HM are the hemorrhage kinds that happen owing to capillary leakages.
- **Hard exudates** happen to be bright yellow spots in the macular area with sharp margins which occurs due to the plasma leakage.
- **Cotton wool spots** or **soft exudates** happen to be white spots on the retinal region which occurs due to the swollen nerve fibre.

Proliferative Diabetic Retinopathy (PDR) happens to be an advanced stage DR retinal disease which shows the following symptoms,

- **Neovascularization** primarily happens in the PDR stage, where new abnormal vessels are developed at the optic disc or elsewhere in the retinal region.
- **Vitreous hemorrhage** - The abnormal retinal vessels may proliferate inside or around the vitreous body.

### III. DEEP LEARNING

Deep Learning (DL) has become a robust and more powerful weapon which acts as a subset of Machine Learning (ML) and an optimal technique to the ML. Deep Learning model consists of a hierarchical-based architecture with a multilayered structure. In medical image analysis, DL plays a dominant role in classifying, localizing, segmenting, and detecting medical images. In Diabetic Retinopathy disease detection and classification, DL provides more impressive and potential results with several methods. Few DL-based methods are Convolutional Neural Networks (CNN), Deep Boltzmann Machines (DBM), Auto encoders, Deep Neural Networks (DNN), Recurrent Neural Networks (RNN), Deep Belief Networks (DBN), and Generative Adversarial Networks (GAN). More number of training data increases the model performance as both low, and high-level features are automatically extracted and learned. Convolutional Neural Networks (CNNs) are globally used by many researchers in medical imaging than other DL methods. The CNN architecture contains three common layers: convolutional layers, pooling layers, and fully connected layers.

As per researcher's requirement, CNNs size, layers, and filter count can vary. In the convolutional layers, several filters integrate to extract the image features in order to produce feature maps. Secondly, in the pooling layers, the dimensions of the feature maps are reduced frequently using the average or max-pooling method. Thirdly, fully connected layers are used, which produces the overall image feature set. Lastly, the classification is done by any of the two activation functions, sigmoid (binary classification) and softmax (multi-classification). In common, the DL-based DR detection and classification tasks can be performed, as shown in Fig. 3. Initially, the data can be acquired, preprocessed to enhance the image details, augmentation can be performed if required when the data samples are less. Then, the data can be fed into the model that extracts the image features and classify them based on the severity levels.

### IV. LITERATURE SURVEY

Many DL-based automatic DR detection systems were developed in the recent times. Some of the existing research works have been discussed in this section.

Zago et al. [1] developed a method where 2 CNNs (pre-trained VGG16 and CNN) were used to diagnose Diabetic Retinopathy or non-DR fundus images according to the probability of lesion patches. DIARETDB1 dataset was used for training. IDRiD, Messidor, Messidor-2, DDR, DIARETDB0, and Kaggle dataset were used for testing purposes. The Messidor dataset achieved best results with a sensitivity of 0.94 and an AUC of 0.912.

Jiang et al. [2] developed a model where 3 CNNs (Inception-v3, ResNet152, and Inception-ResNet-v2) were utilized to classify the fundus image dataset as referable DR or non-referable DR. Prior to CNN training, the images were resized, enhanced and augmented, and then Adaboost technique were used to integrate. Adam optimizer was used for updating the network weights and the system has achieved 88.21% of accuracy and AUC of 0.946.

Wang et al. in [3] have used 3 separate CNNs (pre-trained VGG16, AlexNet, and Inception-v3) to determine the five-stage DR using the Kaggle fundus dataset and the performance comparison of the individual CNNs are done. The fundus images were resized to various sizes for all 3 pre-trained models and produced an accuracy of 63.23%, 50.03%, and 37.43% in Inception-v3, VGG16, and AlexNet respectively. Some of the pros and cons of the DR detection and classification methods are presented in Table I.

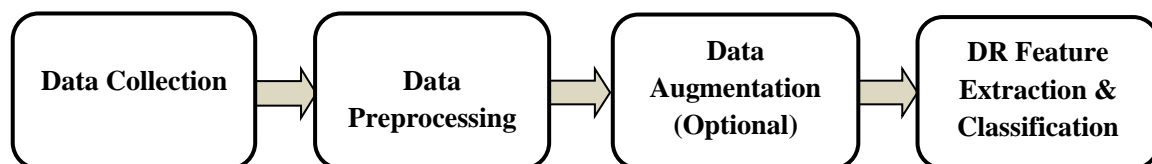


Fig. 3 . DR detection and classification process using DL method

TABLE I. Pros and Cons of DR detection and classification methods

DL Method & Ref.	Datasets used	Pros	Cons	Performance metrics
CNN [5]	Messidor-2 (1748 images)	<ul style="list-style-type: none"> <li>o IDX-DR was integrated with CNN</li> </ul>	<ul style="list-style-type: none"> <li>o They examined mild-DR images as no-DR</li> <li>o The five DR levels were not determined</li> </ul>	<ul style="list-style-type: none"> <li>o <b>AUC:</b> 0.980</li> <li>o <b>Specificity:</b> 87.0%</li> <li>o <b>Sensitivity:</b> 96.8%</li> </ul>
CNN (Custom, AlexNet, VGG-16, SqueezeNet) [6]	Messidor (1200 images)	<ul style="list-style-type: none"> <li>o The images were preprocessed by applying Histogram Equalization (HE) and other techniques.</li> <li>o 4 CNN architectures have been proposed</li> </ul>	<ul style="list-style-type: none"> <li>o To assess the work, only one dataset used</li> <li>o No DR lesions were determined</li> </ul>	<ul style="list-style-type: none"> <li>o <b>Accuracy:</b> 98.15%</li> <li>o <b>Specificity:</b> 97.87%</li> <li>o <b>Sensitivity:</b> 98.94%</li> </ul>
CNN (AlexNet, ResNet, GoogLeNet, VGGNet) [7]	Kaggle (35,126 images)	<ul style="list-style-type: none"> <li>o Transfer learning was used to reduce the training time</li> <li>o The FC layer and hyperparameter alone was tuned</li> <li>o VGGNet produced best results</li> </ul>	<ul style="list-style-type: none"> <li>o To assess the work, only one dataset used</li> <li>o No DR lesions were determined</li> </ul>	<p>For VGGNet,</p> <ul style="list-style-type: none"> <li>o <b>AUC:</b> 0.9786</li> <li>o <b>Accuracy:</b> 95.68%</li> <li>o <b>Specificity:</b> 97.43%</li> <li>o <b>Sensitivity:</b> 90.78%</li> </ul>
CNN (pre-trained AlexNet) [8]	1. Training - Kaggle 2. Testing - IDRiD	<ul style="list-style-type: none"> <li>o The AlexNet and handcrafted features are integrated</li> </ul>	<ul style="list-style-type: none"> <li>o To assess the work, only one dataset used</li> <li>o No DR lesions were determined</li> </ul>	<ul style="list-style-type: none"> <li>o <b>Accuracy:</b> 90.07%</li> </ul>
R-FCN [9]	1. Messidor 2. Private dataset	<ul style="list-style-type: none"> <li>o Feature pyramid network and 5 region proposal networks have been included to improve R-FCN method</li> </ul>	<ul style="list-style-type: none"> <li>o To assess the work, only one dataset used</li> <li>o No exudates detected, only HM and MA are detected</li> </ul>	<p><b>Sensitivity</b></p> <ul style="list-style-type: none"> <li>o <b>Messidor:</b> 92.59%</li> <li>o <b>Private dataset:</b> 99.39%</li> </ul>

Pratt et al. [4] developed a method where CNN was used with 10 convolutional layers, 8 max-pooling layers, 3 fully connected layers, and a softmax classifier to classify the Kaggle fundus images into 5 classes based on the DR severity levels. The color fundus images are normalized, and resized. In order to reduce overfitting, L2 regularization and dropout methods were used. The model has produced 95% specificity, 75% accuracy, and 30% sensitivity.

Hua et al. [10] extracted the retinal blood vessels in the DRIVE dataset images. The author has used a ResNet-101 pre-trained network to choose 4 feature maps, and the individual feature maps were integrated to produce a single map. The fundus images were augmented before CNN processing. The best feature maps were combined to achieve an accuracy of 0.951, a sensitivity of 0.793, AUC of 0.9732, and a specificity of 0.9741.

Wu et al. [11] developed a CNN, for extracting the retinal blood vessels from 3 standard datasets: STARE, DRIVE, and CHASE. The RGB images were converted into grayscale images, normalized, augmented, and the image contrast has been enhanced using CLAHE in the preprocessing phase. CNN built of encoder-decoder structure, which consists of convolutional, normalization, concatenation, dropout layers, and skip connections. The model has achieved 98.75%, 98.30%, and 98.94% AUC, and an accuracy of 96.72%, 95.82%, and 96.88% for the STARE, DRIVE, and CHASE databases respectively.

Oliveira et al. [12] developed a complete CNN model which extracts the blood vessels from the DRIVE, STARE, and CHASE\_DB1 datasets. The color fundus images were

preprocessed initially. After green channels extraction, Stationary Wavelet Transform (SWT) was applied and the dataset images are normalized. Finally the patches were extracted and it was augmented prior to CNN processing. The model has achieved AUC of 0.9821, 0.9905, and 0.9855 in DRIVE, STARE, and CHASE\_DB1 databases respectively.

Chudzik et al. [13] developed a CNN model which consists of 18 convolutional layers, batch normalization layers, 3 max-pooling layers, and up-sampling layers, with 4 skip connections. The author has worked with 3 datasets: E-ophtha, DIARETDB1, and ROC to determine microaneurysms from the fundus images. The color images were preprocessed prior to CNN processing. The green planes are extracted, cropped, resized, to produce a mask, Otsu thresholding has been applied and morphological functions are used to achieve a ROC of 0.355.

Wang et al. [9] developed a model that detects exudate lesions (hard) by integrating the handcrafted features with the CNN features using Random Forest (RF) classifier in HEI-MED and E-ophtha datasets. The preprocessing consists of following operations: cropping, normalizing, performing morphological operations, and thresholding (dynamic) have been used to determine the candidates. The CNN was made using 3 convolutional and pooling layers, and one FC layer has been used to determine the features. The model has produced an AUC of 0.9323 and 0.9644, and a sensitivity of 0.9477 and 0.8990 in HEI-MED and E-ophtha databases, respectively.

Yan et al. [14] developed a method with 2 CNN architectures (U-net and improved LeNet) to detect DR lesions from the DIARETDB1 database images by integrating the traditional handcrafted features, and the improved LeNet features using RF classifier. The green channels were cropped, CLAHE was used to enhance the image contrast, Gaussian filter has been used to remove the noise, and the morphological operations also performed in the preprocessing phase. To segment the blood vessels, U-net architecture was used, and the LeNet architecture has been improved with 4 convolutional layers, 3 max-pooling layers, and a fully connected layer to produce 48.71% sensitivity, when detecting red lesions.

Zhang et al. [15] proposed an automatic system called DeepDR that used an ensemble of Pretrained networks, Resnet and Inception V3. It has achieved an AUC of 97.7%, sensitivity of 97.5% and specificity of 97.7%.

## V. CONCLUSION

Automated Diabetic Retinopathy detection systems are cost-efficient; it reduces the detection time that helps ophthalmologists in diagnosing retinal abnormalities and timely treatment can be provided to avoid future consequences. These DR detection systems play a key role in diagnosing diseases more accurately. Initially, in this paper, Diabetic Retinopathy, DR types, and symptoms are discussed. Deep Learning and CNN concepts were explored. In the literature survey, some of the state-of-the-art techniques were reviewed. In most of the research works, CNNs have been used for its ability to provide more potential results and its efficiency, which outperform the other methods. More work to be done on this part to overcome all the drawbacks and to improve performance of the system.

## VI. FUTURE DIRECTIONS

Some studies in the literature have proposed that while combining the conventional handcrafted and CNN-based features, the system performance has been improved. In future, CNN variants can be combined to extract more relevant and salient image features which eventually increase Diabetic Retinopathy detection and classification rate.

## REFERENCES

- [1] Zago, GT, Andreao, RV, Dorizz, B and Teatini Salles, EO, "Diabetic retinopathy detection using red lesion localization and convolutional neural networks. Computers in Biology and Medicine", 116, 103537, 2020.
- [2] Jiang, H, Yang, K, Gao, M, Zhang, D, Ma, H and Qian, W, "An interpretable ensemble deep learning model for diabetic retinopathy disease classification", 41st Annual International conference of the IEEE engineering in medicine and biology society (EMBC), pp. 2045–2048, 2019.
- [3] Wang, X, Lu, Y, Wang, Y and Chen, WB, "Diabetic retinopathy stage classification using convolutional neural networks", International Conference on information Reuse and Integration for data science, p. 465–71, 2018.
- [4] Pratt, H, Coenen, F, Broadbent, DM, Harding, SP and Zheng, Y, "Convolutional neural networks for diabetic retinopathy", *Procedia Comput Sci.*, 90:200–5, 2016.
- [5] Abramoff, M. D, Lou, Y, Erginay, A, et al., "Improved Automated Detection of Diabetic Retinopathy on a Publicly Available Dataset through Integration of Deep Learning", *Invest Ophthalmol Vis Sci.*, 57(13):5200-5206, 2016.
- [6] Mobeen-Ur-Rehman, Khan, S. H, Abbas, Z, Danish Rizvi, S. M, "Classification of diabetic retinopathy images based on customised CNN architecture", In: *Proceedings-Amity International Conference on artificial intelligence*, p. 244–8, 2019.
- [7] Wan, S, Liang, Y, Zhang, Y, "Deep convolutional neural networks for diabetic retinopathy detection by image classification", *Comput Electr Eng.*, 72:274–82, 2018.
- [8] Harangi, B, Toth, J, Baran, A and Hajdu, A, "Automatic screening of fundus images using a combination of convolutional neural network and hand-crafted features", 41st annual International Conference of the IEEE Engineering in Medicine and biology society (EMBC), p. 2699–702, 2019.
- [9] Wang, J, Luo, J, Liu, B, Feng, R, Lu, L and Zou, H, "Automated diabetic retinopathy grading and lesion detection based on the modified R-FCN object-detection algorithm", *IET Comput. Vis.*, 14(1):1–8, 2020.
- [10] Hua, C. H, Huynh-The, T and Lee, S, "Retinal vessel segmentation using round-wise features aggregation on bracket-shaped convolutional neural networks", *Proceedings of the annual International Conference of the IEEE Engineering in Medicine and biology society, EMBS*, p. 36–9, 2019.
- [11] Wu, Y, Xia, Y, Song, Y, Zhang, Y and Cai, W, "NFN+: a novel network followed network for retinal vessel segmentation", *Neural Network*, 126:153–62, 2020.
- [12] Oliveira, A, Pereira, S and Silva, C.A, "Retinal vessel segmentation based on fully convolutional neural networks", *Expert Syst. Appl.*, 112:229–42, 2018.
- [13] Chudzik, P, Majumdar, S, Caliva, F, Al-Diri, B and Hunter, A, "Microaneurysm detection using fully convolutional neural networks", *Comput Methods Progr Biomed*, 158:185–92, 2018.
- [14] Yan, Y, Gong, J and Liu, Y, "A novel deep learning method for red lesions detection using hybrid feature", *Proceedings of the 31st Chinese Control and decision conference*, p. 2287–92, 2019.
- [15] Zhang, W.; Zhong, J.; Yang, S.; Gao, Z.; Hu, J.; Chen, Y.; Yi, Z. Automated identification and grading system of diabetic retinopathy using deep neural networks. *Knowl. Based Syst.*, 2019, 175, 12–25.

# Machine-Learning-Scheme to Detect Choroidal-Neovascularization in Retinal OCT Image

Venkatesan Rajinikanth  
*Department of Electronics and  
 Instrumentation Engineering  
 St. Joseph's College of Engineering  
 Chennai 600119, India  
 email: v.rajinikanth@ieee.org*

David Taniar  
*Faculty of Information Technology  
 Monash University  
 Clayton, Victoria 3800, Australia  
 email: David.Taniar@monash.edu*

Seifedine Kadry  
*Faculty of Applied Computing and  
 Technology, Noroff University College,  
 Kristiansand, Norway  
 email: skadry@gmail.com*

Hafiz Tayyab Rauf  
*Centre for Smart Systems, AI and  
 Cybersecurity, Staffordshire University,  
 Stoke-on-Trent, United Kingdom  
 email: hafiztayyabrauf093@gmail.com*

Robertas Damaševičius  
*Faculty of Applied Mathematics  
 Silesian University of Technology  
 44-100 Gliwice, Poland  
 email: robertas.damasevicius@polsl.pl*

**Abstract**— Eye is a fundamental sensory organ and any disease in eye will severely affect the sensory signal evaluation and conclusion making capability of the brain. The Choroidal-Neovascularization (CNV) is one of the harsh eye diseases in which a new blood-vessel grow from the choroid. Usually, the major cause of CNV is due to wet Age-Related-Macular-Degeneration (ARMD) and the formed new vessel will cause a leak in fluid which makes the retinal wet. The untreated CNV will lead to vision loss. In this research, detection of CNV using Optical-Coherence-Tomography (OCT) is presented using 484 images (242 Healthy and 242 CNV). In this work, a Machine-Learning-Scheme (MLS) is developed to examine the resized OCT of 256x256 pixels and the stages of this MLS includes; pre-processing, feature extraction, Mayfly-Optimization-Algorithm (MFA) based feature reduction, and two-class classification. The experimental outcome of this technique confirmed that the Fine-Gaussian-SVM (SVM-FG) classifier helped to accomplish an improved classification accuracy (>92%) compared to the alternative classifiers of this study.

**Keywords**— OCT image, Choroidal-Neovascularization, Retinal OCT, Mayfly-Optimization-Algorithm, Classification.

## I. INTRODUCTION

In human physiological system, a considerable number of sensory organs are available to accumulate the crucial information from the environment and these signals are collected and processed by the Central-Nervous-System (CNS) for assessment and decision making. Among these sensory organs, the eye plays a vital role to exchange the light into visual information and the illness in eye will severely affect the whole decision making process of the brain.

In humans, the disease in eye is due to; (i) Ageing, (ii) Infection and (iii) Accident. The untreated eye disease will lead to a lot of troubles, including loss of vision. Hence, a number of procedures are supported for the clinical level examination of the eye.

The common examination procedures followed in eye centres includes; (i) Visual check by a skilled ophthalmologist, (ii) Fundus image supported examination [1-4], (iii) Fluorescein-angiography [5] and (iv) Optical-Coherence- Tomography (OCT) [6-8]. All these approaches are non-invasive imaging procedures and check/record the essential section of the eye for the disease detection and treatment planning task.

In the proposed research, the retinal images recorded with OCT is considered for the assessment and this approach

records the essential section of the retina using cross-sectional imaging; which further supports the cross-sectional and longitudinal studies [9, 10]. The initial OCT scanner was invented in 1996 and the modern version of the OCT scanner was released in 2001 [11]. This technique is very efficient in capturing the retinal layers with specified thickness and the assessment helps to achieve a better result on a class of retinal images, which cannot be diagnosed using the conventional fundus-pictures.

The proposed research implemented a conventional Machine-Learning-Scheme (MLS) to classify the considered OCT images into healthy/CNV class. The various stages involved in the proposed technique includes; (i) Collection of the essential images for the assessment, (ii) Implementing image resizing to regularize the picture dimension, (iii) Image pre-processing to enhance the features, (iv) Extraction of handcrafted features, such as GLCM, Hu moments and LBP, (v) Feature reduction with Mayfly-Optimization-Algorithm (MOA) and (vi) Two-class classifier working and justification.

This effort is executed using the MATLAB software and the important images (484 numbers) are collected from [12]. The existing image with dimension 512x496 is reduced to 256x256 pixels to reduce the complexity in the assessment. The resized image is evaluated with two pre-processing method; (i) Saliency based enhancement, morphological segmentation and feature extraction using GLCM and HU, and (ii) LBP with Global weight parameter. The features extracted from each technique are separately selected with MOA and the reduced features are serially concatenated and used to prove the performance of executed MLS.

The overall performance of any MLS depends on the final outcome and in this research; a two class classification is implemented and the results attained with the Fine-Gaussian-SVM is good compared to other SVM and Decision-Tree (DT) variants considered in this research. The main contribution of this research includes; (i) implementing the chosen handcrafted-feature extraction, (ii) MOA based feature reduction and (iii) two-class classification with SVM and DT variants. For every approach, a 5-fold cross validation is executed and the finest value among these trials is chosen as the best outcome.

Other parts of this research is pre-arranged as below; Section 2 shows the context, Section 3 presents the methodology, Section 4 and 5 illustrates investigational results and conclusion of this research.

## II. CONTEXT

Assessment of the retinal abnormality is a significant task to detect and cure various retinal abnormalities. The fundus-picture supported retinal abnormality assessment is widely addressed to identify a class of retinal diseases. The earlier research also confirms the need of fundus-pictures to appraise the eye abnormalities using various image assisted techniques [1-4, 13, 14].

The retinal OCT is one of the new imaging modalities, widely in use from the year 1996 and this modality helps to evaluate various retinal abnormalities with better accuracy and hence, most of the clinical study suggests the need for the retinal OCT assessment [11]. Further, this technique will help the ophthalmologist in evaluating the various layers of the retinal layers and the optic nerve section in a live eye.

The work of Lake et al. [9] measured the retinal shape irregularity using the retinal OCT. Alamouti and Funk [10] presented a research to confirm the reduction in retinal thickness based on age. Drexler and Fujimoto [15] presented a detailed assessment procedure for the OCT appraisal. Drexler and Fujimoto [16] presented a detailed examination and identification of various section in retinal OCT picture. Bogunović et al. [6] presented a work on OCT fluid detection using the benchmark image database. The work of Pekala et al. [17] presented the Deep-Learning (DL) supported segmentation of retinal layer from the OCT image. Kermany et al. [8] presented a DL scheme to classify the OCT into various classes and their database is publicly shared to promote the OCT related research work.

This research aims to implement a MLS scheme to detect the CNV from the considered OCT images of Kermany et al. [8] with improved accuracy.

## III. METHODOLOGY

The aim of the research is to develop a suitable MLS to detect the CNV from the retinal OCT image. The proposed MLS is illustrated in Figure 1. Initially, the collected images are resized into 256x256 pixels and the converted imagery are then considered for the assessment. The proposed MLS scheme is having two OCT examination pipeline; (i) To extract the OCT section to detect the GLCM and Hu moments and, (ii) Detection of the LBP features from the enhanced images. The features of this pipeline are separately reduced using the MOA and the reduced features are serially combined based on the feature rank and then a classification process is executed using a 5-fold cross-validation.

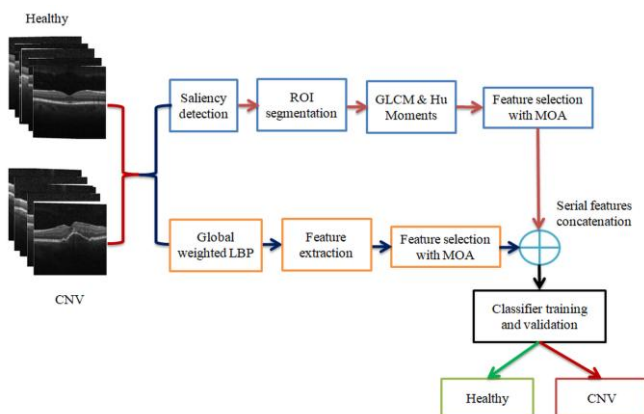


Fig.1. The structure of the proposed CNV detection system

The greatest outcome achieved with the 5-fold cross justification is then used as the final result which helps to categorize the considered retinal OCT into healthy and CNV class.

### A Image Database

The test pictures considered in this research are collected from [12]. This dataset consist the test images with a dimension 512x496 pixels and during the assessment, every test picture is resized to 256x256 pixels and the sample images considered in this work is depicted in Figure 2 and Table 1.

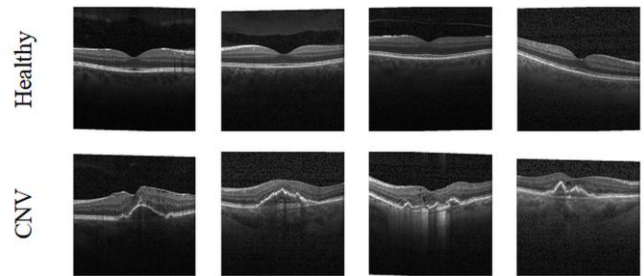


Fig.2. Sample trial images of healthy and CNV class

TABLE I. CONSIDERED RETINAL OCT IMAGES FOR RESEARCH

Image class	Dimension	Modified size	Retinal OCT		
			Total	Training	Testing
Healthy	512x496	256x256	242	200	42
CNV	512x496	256x256	242	200	42

Table 1 depicts the information on the test pictures considered in this research for training and testing the proposed MLS.

### B. Image pre-processing

This part of the research presents both the pre-processing pipelines separately.

#### 1. Saliency Enhancement and Morphological Segmentation

The saliency technique helps to enhance the essential information (image pixel with better intensity) in the considered test image and after the enhancement; the segmentation of the retinal layer is extracted using the morphological segmentation scheme. The earlier works shows the information on saliency scheme and morphological segmentation [18-20]. The proposed pipeline is an automatic segmentation procedure and mines the retinal layer (binary image) for further assessment.

The GLCM and Hu moments are extracted from the binary form of the retinal OCT and the extracted features are presented in Eqns. (1) and (2) and the total features extracted are shown in Eqn. (3) ;

$$F_{GLCM} (1 \times 25) = GLCM(1,1), \dots, GLCM(1,25) \quad (1)$$

$$F_{Hu} (1 \times 7) = Hu(1,1), \dots, Hu(1,7) \quad (2)$$

$$F_{Total} (1 \times 32) = F_{GLCM} (1 \times 25) + F_{Hu} (1 \times 7) \quad (3)$$

A feature reduction process is to be implemented to reduce the features in order to avoid over fitting problem during the automated detection [21-23].

2. Global Weighted LBP supported Enhancement

The second pre-processing pipeline involves in the creation of the LBP images with varied weights (W=1,2,3,4) and from every image, 59 LBP features are extracted as depicted below. For each test image, the number of LPB features extracted will be; 59x4=236 features and this is depicted in Eqn. (4) ;

$$F_{LBP}(1 \times 236) = F_{LBP1}(1 \times 59) + F_{LBP2}(1 \times 59) + F_{LBP3}(1 \times 59) + F_{LBP4}(1 \times 59) \quad (4)$$

In this work, feature reduction process is to be employed to reduce the existing features to a lower value.

D. MOA based feature reduction

In this research, the dominant feature vector is selected using the recently developed Mayfly-Optimization-Algorithm (MOA) [24]. The MOA is a hybrid heuristic procedure; invented by integrating the best features of Firefly, Particle-Swarm and Genetic-algorithm [24, 25]. The concept of the MOA is simple and having various stages as depicted in Figure 3.

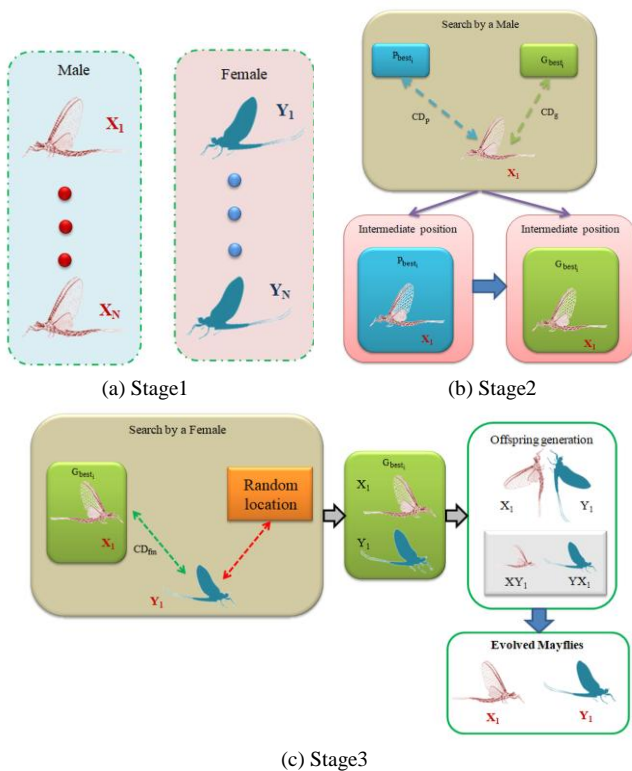


Fig.3. Different stages existing in MOA (a) Initialization of the MOA with chosen male- and female-Mayfly, (b) Finding the optimal value by the male mayfly, and (c) Finding the best male Mayfly by female-Mayfly and offspring generation

The working mechanism of the MOA is similar to the FA and the task of the Mayfly is to find the optimal solution for the chosen problem based on the assigned cost value. All the Mayflies in the considered MOA will follow the Levy flight search strategy as existing in the FA algorithm [26]. In this

work, the MOA is employed to reduce the considered image features presented in Eqn. (3) and Eqn. (4) and after the feature reduction, the reduced features are serially combined based on the rank as discussed in the earlier work [22].

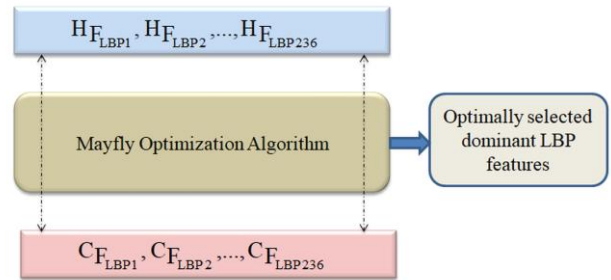


Fig.4. MOA based feature reduction implemented for LBP features

Figure 4 depicts the feature selection implemented to select the optimal LBP features from the initial feature vector of dimension 1x236 features. During this process, the MOA computes the Hamming-distance between the healthy and CNV features based on the distance; which provided the reduced features with dimension 1x44. Similar procedure is employed for the GLCM and Hu and a reduced feature value of 1x13 is attained. After finding the optimal features, a feature ranking and serial concatenation is implemented to get the final feature vector of dimension 1x57 ((1x44) + (1x13)) is attained and these features are then used to evaluate two-class classifier employed. The MOA is initialized as; number of agents = 30, search dimension = number of features to be selected, objective value = maximal Hamming-distance, maximum iterations (Iter\_max) = 5000 and terminating criteria = Iter\_max.

E. Classification and Validation

Classification is the final stage in the proposed MLS and in this works the MOA optimized handcrafted features are adopted to test the presentation of the implemented classifiers. In this scheme, the SVM modification such as, linear (SVM-L), quadratic (SVM-Q), cubic (SVM-C), Fine Gaussian (SVM-FG), Medium Gaussian (SVM-MG), Coarse Gaussian (SVM-CG) are initially considered and later the Decision-Tree variants, such as Coarse Tree (CT), Medium Tree (MT) and Fine Tree (FT) are also considered and the results are presented [18].

The performance of proposed MLS is authorized via 42 trial images and the performance of proposed structure is validated by computing the crucial performance-values (PV), such as True-Positive (TP), True-Negative (TN), False-Positive (FP) and False-Negative (FN), From these PV, additional measures, such as accuracy (AC), precision (PR) sensitivity (SE), specificity (SP) and F1-Score (FS) and based on these values, the performance of the developed MLS is confirmed.

IV. RESULT AND DISCUSSION

This sector shows the investigational result attained using the MLS for the retinal OCT dataset and this work is executed using the MATLAB software.

The proposed technique is implemented using the retinal OCT images of healthy/CNV class and the sample picture is depicted in Figure 5. In this figure, the change in retinal layer

can be seen in the CNV case compared to the healthy class OCT. The initial task in the MLS is to segment the retinal fragment from the OCT using the preferred segmentation practice.

Figure 6 depicts the segmentation process implemented to extract the retinal layer. Fig 6(a) and (b) depicts the chosen test picture and the saliency assisted enhancement, Fig 6(c) and (d) depicts the morphological enhancement and the segmentation of the retinal region from the chosen test image respectively. Fig 6(d) is then considered to extract the GLCM and Hu moments from the binary part of the image.

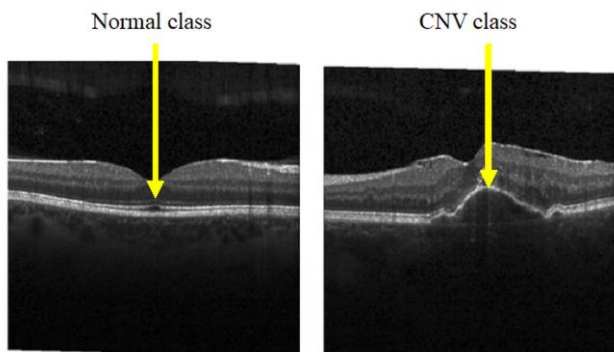


Fig.5. Variation in retinal layer of healthy/CNV class OCT picture

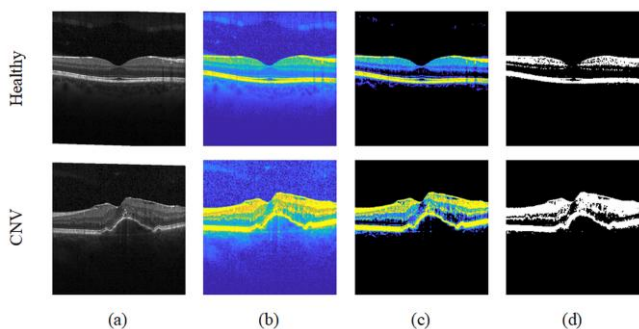


Fig.6. Segmentation result attained using the first pre-processing pipeline (a) Trial picture, (b) Saliency based enhancement of retinal region, (c) Morphological enhancement, (d) Extracted binary picture

Figure 7 presents the LBP enhanced retinal OCT for various weighting parameters and from these images, it is clearly seen that the change in weights will improve the visibility of the retinal pattern. Fig 7(a) to (d) depicts the attained LBP pattern for various weights, such as  $W=1$  to 4 respectively. The related change in the LBP histogram is depicted in Figure 8, in which Fig 8(a) and (b) shown the histogram for healthy and CNV class images, respectively.

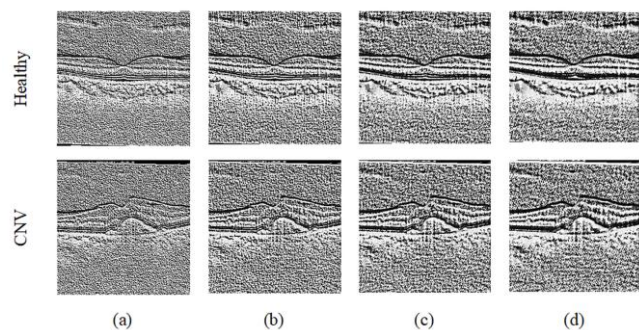


Fig.7. LBP enhanced retinal OCT picture for various weights (a) to (d) denotes the LBP for  $W=1$  to 4

These LBP patterns are then considered to extract the LBP features of dimension  $1 \times 59$  from each image and helped to achieve a feature vector of dimension  $1 \times 236$  from all four images [27]. These features are reduced using the MOA and the reduced feature with dimension  $1 \times 57$  is then used to train and authorize the considered two-class classifiers using a 5-fold cross validation process. In this work, various classifiers are considered for the assessment and the best result attained among the 5-fold cross justification is chosen as the final result as depicted in Table II.

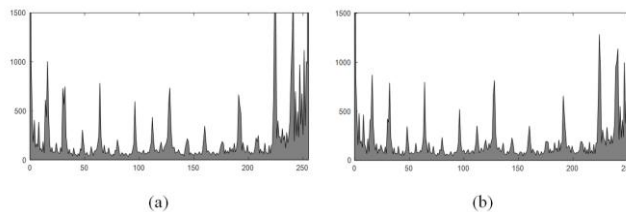


Fig.8. The LBP histogram pattern (a) Healthy class, (b) CNV class

TABLE II. PERFORMANCE VALUES ATTAINED WITH TWO-CLASS CLASSIFIERS

Method	TP	FN	TN	FP	AC (%)	PR (%)	SE (%)	SP (%)	FS (%)
SVM-L	37	5	34	8	84.52	82.22	88.09	80.95	85.06
SVM-Q	34	8	40	2	88.09	94.44	80.95	95.24	87.18
SVM-C	38	4	37	5	89.28	88.37	90.47	88.09	89.41
SVM-FG	38	4	40	2	92.86	95.00	90.47	95.24	92.68
SVM-MG	39	3	38	4	91.67	90.69	92.86	90.47	91.76
SVM-CG	38	4	39	3	91.67	92.68	90.48	92.86	91.57
CT	37	5	40	2	91.67	94.87	88.09	95.24	91.36
MT	39	3	37	5	90.47	88.63	92.86	88.09	90.69
FT	40	2	36	6	90.48	86.96	95.24	85.71	90.90

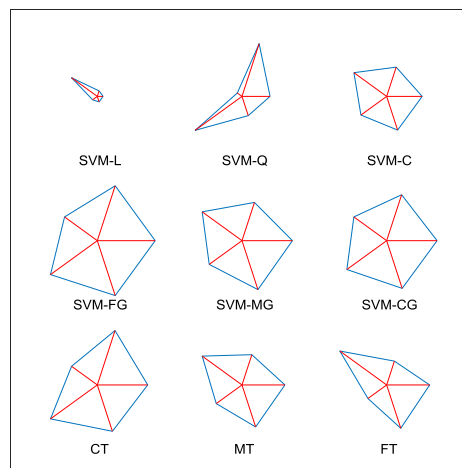


Fig.9. Glyph-diagram to represent the overall performance

Figure 9 presents the overall performance of the chosen classifiers and from Table II and Fig 9, it can be noted that the classification performance attained with the SVM-FG is better compared to other methods.

In the proposed work, only limited images are considered for the assessment (242 images per case) and in future, the number of images can be increase to further improve the classification accuracy of the proposed MLS. The performance of the MLS can be confirmed with a chosen DL



method. Further, the main future scope of this study include the replacing the considered MOA with the recently invented Red-Fox-Optimization [28] method to optimize the handcrafted feature value.

## V. CONCLUSION

Examination if the eye is essential for the elderly people and the age related eye disease needs to be accurately diagnosed to fix the vision problem. A MLS framework is planned to examine the CNV class retinal abnormality using the retinal OCT images. In this work, two different image pre-processing pipeline is applied to improve the trial image and to extract the features, such as GLCM, Hu and LBP and the collected features are optimized using the MOA. The optimized features (1x57) is then considered to test and validate the two-class classifier and the attained result of this study substantiate that the MLS with SVM-FG classifier helped to achieve a better accuracy (>92%) contrast to other techniques.

## REFERENCES

- [1] N. Kowsalya, A. Kalyani, C. J. Chalcedony, R. Sivakumar, M. Janani, and V. Rajinikanth, "An approach to extract optic-disc from retinal image using K-means clustering," In 2018 Fourth International Conference on Biosignals, Images and Instrumentation (ICBSII), IEEE, pp. 206-212, 2018.
- [2] G. Kavitha and S. Ramakrishnan, "An approach to identify optic disc in human retinal images using ant colony optimization method", *Journal of Medical Systems*, vol. 34, no. 5, pp. 809-813, 2010.
- [3] T.D.V. Shree, K. Revanth, N.S.M. Raja, Raja and V. Rajinikanth, "A hybrid image processing approach to examine abnormality in retinal optic disc", *Procedia Computer Science*, vol. 125, pp. 157-164, 2018.
- [4] K. R. Remya, and M. N. Giriprasad, "Computer based Technique to Examine Diabetic Retinopathy in Fundus Retinal Images," In 2020 5th International Conference on Computing, Communication and Security (ICCCS), IEEE, pp. 1-8, 2020.
- [5] M. A. Klufas, S. N. Patel, M. C. Ryan et al., "Influence of fluorescein angiography on the diagnosis and management of retinopathy of prematurity," *Ophthalmology*, vol. 122, no. 8, pp.1601-1608, 2015. Doi: 10.1016/j.ophtha.2015.04.023.
- [6] H. Bogunović, F. Venhuizen, S. Klimscha et al., "RETOUCH: the retinal OCT fluid detection and segmentation benchmark and challenge. " *IEEE transactions on medical imaging*, vol. 38, no. 8, pp.1858-1874, 2019. Doi: 10.1109/TMI.2019.2901398.
- [7] J. Tian, B. Varga, E.Tatrai et al., "Performance evaluation of automated segmentation software on optical coherence tomography volume data," *Journal of biophotonics*, vol. 9, no. 5, pp.478-489, 2016. Doi: 10.1002/jbio.201500239
- [8] D. S. Kermany, M. Goldbaum, W. Cai et al., "Identifying medical diagnoses and treatable diseases by image-based deep learning," *Cell*, vol. 172, no. 5, pp.1122-1131, 2018. <https://doi.org/10.1016/j.cell.2018.02.010>.
- [9] S. Lake, M. Bottema, K. Williams, and K. Reynolds, "The correlation between optical coherence tomography retinal shape irregularity and axial length," *PloS one*, vol. 14, no. 12, pp. e0227207, 2019.
- [10] B. Alamouti, and J. Funk, "Retinal thickness decreases with age: an OCT study," *British journal of ophthalmology*, vol. 87, no. 7, pp.899-901, 2003.
- [11] J.G. Fujimoto, C. Pitris, S. A. Boppart, and M. E. Brezinski, "Optical coherence tomography: an emerging technology for biomedical imaging and optical biopsy," *Neoplasia*, vol. 2, no. 1-2, pp. 9-25, 2000.
- [12] <https://www.kaggle.com/paultimothymooney/kermany2018>
- [13] E. Priya, and V. Rajinikanth, *Signal and Image Processing Techniques for the Development of Intelligent Healthcare Systems*. Springer, 2020.
- [14] V. Rajinikanth, E. Priya, H. Lin, and F. Lin, *Hybrid Image Processing Methods for Medical Image Examination*. CRC Press, 2021.
- [15] W. Drexler, and J. G. Fujimoto, eds. *Optical coherence tomography: technology and applications*. Springer Science & Business Media, 2008.
- [16] W. Drexler, and J. G. Fujimoto, "State-of-the-art retinal optical coherence tomography," *Progress in retinal and eye research*, vol. 27, no. 1, pp.45-88, 2008.
- [17] M. Pekala, N. Joshi, T.Y. A. Liu, N. M. Bressler, D. C. DeBuc, and P. Burlina, "Deep learning based retinal OCT segmentation," *Computers in biology and medicine*, vol. 114, pp.103445, 2019.
- [18] M.A. Khan et al., "Computer-aided gastrointestinal diseases analysis from wireless capsule endoscopy: A framework of best features selection," *IEEE Access*, vol. 8, pp.132850-132859, 2020. Doi: 10.1109/ACCESS.2020.3010448.
- [19] X. Hou, and L.Zhang, "Saliency detection: A spectral residual approach," In *2007 IEEE Conference on computer vision and pattern recognition*, pp. 1-8. IEEE, 2007.
- [20] F. Meyer, and S. Beucher, "Morphological segmentation," *Journal of visual communication and image representation*, vol. 1, no. 1, pp.21-46, 1990.
- [21] N. Dey et al., "Social-Group-Optimization based tumor evaluation tool for clinical brain MRI of Flair/diffusion-weighted modality," *Biocybernetics and Biomedical Engineering*, vol. 39, no. 3, pp. 843-856, 2019. <https://doi.org/10.1016/j.bbe.2019.07.005>.
- [22] A. Bakiya, K. Kamalanand, V. Rajinikanth, R.S. Nayak, and S. Kadry, "Deep neural network assisted diagnosis of time-frequency transformed electromyograms," *Multimedia Tools and Applications*, vol. 79, no. 15, pp.11051-11067, 2020.
- [23] N. Dey, Yu-Dong Zhang, V. Rajinikanth, R. Pugalenth, and N.S.M. Raja, "Customized VGG19 architecture for pneumonia detection in chest X-rays," *Pattern Recognition Letters*, vol.143, pp.67-74, 2021. <https://doi.org/10.1016/j.patrec.2020.12.010>.
- [24] K. Zervoudakis, and S. Tsafarakis, "A mayfly optimization algorithm," *Computers & Industrial Engineering*, vol. 145, pp.106559, 2020. <https://doi.org/10.1016/j.cie.2020.106559>.
- [25] Z-M. Gao, J. Zhao, S-R. Li, and Y-R. Hu, "The improved mayfly optimization algorithm," In *Journal of Physics: Conference Series*, vol. 1684, no. 1, pp. 012077. IOP Publishing, 2020.
- [26] N.S.M. Raja, K. S. Manic, and V. Rajinikanth, "Firefly algorithm with various randomization parameters: an analysis," In *International Conference on Swarm, Evolutionary, and Memetic Computing*, pp. 110-121. Springer, Cham, 2013. [https://doi.org/10.1007/978-3-319-03753-0\\_11](https://doi.org/10.1007/978-3-319-03753-0_11).
- [27] A. Gudigar et al., "Global weighted LBP based entropy features for the assessment of pulmonary hypertension," *Pattern Recognition Letters*, vol.125, pp.35-41, 2019. <https://doi.org/10.1016/j.patrec.2019.03.027>.
- [28] D. Połap, and M. Woźniak, "Red fox optimization algorithm," *Expert Systems with Applications*, vol. 166 pp.114107, 2021. <https://doi.org/10.1016/j.eswa.2020.114107>.

# AN OPTIMIZATION OF FEATURE SELECTION FOR CLASSIFICATION USING META-HEURISTIC WHALE OPTIMIZATION ALGORITHM

V. Yaraswini<sup>1</sup>, Dr. Santhi Baskaran<sup>2</sup>

1. Research Scholar, Computer Science and Engineering Department, Pondicherry Engineering College, Puducherry, India.
2. Professor & Head, Information Technology Department, Pondicherry Engineering College, Puducherry, India

**Abstract.** Data mining is the action of searching the large existing database in order to get new and best information. In this paper Feature selection which is a part of Data mining is performed to do classification. We propose a technique to get the optimized feature selection to perform classification using Meta Heuristic algorithms. We applied new and recent advanced optimized algorithm named Whale Optimization algorithm on UCI datasets that showed comparatively equal results with best performed existing firefly but with less number of features selected. The work is implemented using JAVA and the Medical dataset (UCI) has been used. Classification is done using J48 classifier in WEKA tool. We demonstrate the comparative results of the presently used algorithm with the existing algorithms like Ant colony, Firefly, Cuckoo Search and Harmony Search thoroughly.

**Index Terms:** Optimization, Meta-heuristic, f-Measure, Complexity.

## 1. Introduction

Data Mining [1] is the way of searching important information from the huge present all over in the repository. Data Mining falls in to two ways namely Association and Classification analyzing methods.

Optimization algorithm provides a systematic way of developing and leveling new solutions to gain an optimal result. The optimization process must only be used in those problems where there is a specific need of accomplishing a quality or a competitive work. It is expected that the solution obtained through an optimization method is better than other results in terms of the selected objective.

This paper shows the Bat algorithm and Modified Bat algorithm accuracy rates when compared to existing algorithms namely Firefly, Cuckoo search and Harmony Search algorithms that showed almost equal results of the best accuracy rates in existing work. There are various applications with respect to data mining and optimization techniques in different fields. This method proves the better analysis which gives the best results and improved accuracy. The following are the different field of applications.

1. Network Security
2. Computer Vision and Processing
3. Nature Inspired fields.
4. Medical Fields
5. Transition Probabilities for Radio Systems
6. Intrusion Detection
7. Education
8. Financial Banking

## 2. OVERVIEW ON DATAMINING

Data mining process involves the following stages.

**a) Problem Definition.** In this stage the analysis of the problem in the business problem is done and tries to get the clear idea of the problem to be solved. This

takes some time to make an exact definition of the problem and it does not require any data tools.

**b) Exploration of Data.** In this stage data is explored by identifying quality problem to understand the metadata meaning. It is next level of problem definition stage which frequently exchanges the data.

**c) Preparation of Data.** In this stage data model is built after the exploration of data. Data is collects, clear the unwanted data and arrange the data in a format like tables and records.

**d) Data Modeling.** At this stage after preparation of information, different mining functions are applied to the same kind of data. A high quality of mining model is prepared based on the changes in the parameters until we get optimal data model. Finally the good quality model is built and evaluated.

**e) Evaluation of the Model.** In this stage the evaluated model is checked and tested whether the quality is good or not and objective is satisfied or not?

**f) Deployment.** In this stage after the evaluation of data, the exporting of the data is done and the results are checked into database tables.

## 2. Algorithms

### 2.1-FIREFLY ALGORITHM:

Firefly Algorithm (FA)[9] being a Nature Inspired algorithm works based on flashing nature of the fireflies. The main reason for its flash is to move as a system which provides signal to absorb the other fireflies towards itself. The algorithm was implemented to perform Feature Selection (FS) for Image Processing and Eigen Value Optimization problem along with other related domains and has been performed so that better results are obtained. In order to achieve best optimal feature subset increases the predictive accuracy of the classifier. In this algorithm, along with considering the brighter firefly to obtain the predictive accuracy of the dataset we have also considered a comparatively brighter firefly and the predictive accuracy of that method have also been calculated.

- Choosing a Brightest Firefly
- Choosing a Comparatively Brighter Firefly

It includes previously chosen firefly solution and the newly selected brighter firefly solution in our computation. The latest solution is found by solving the below formulation

$$\text{New } X_i = X_i + \beta_0 \exp(-\gamma(r_{ij})^2) (X_i - X_j) + \alpha \epsilon_i$$

Where,

### Parameter Settings of Firefly Algorithm

$X_i$	The solution pointed by the current firefly (Classification Accuracy)
-------	---

$X_j$	The solution pointed by the brightest firefly (First Method) The solution pointed by the comparatively brighter firefly (Second Method)
$\beta_0$	Between 0 and 1
$\gamma$	Between 0.1 and 10
$r_{ij}$	Distance is fixed to 1
$\alpha$	Parameters selected within the range [0,1] randomly

**2.1 EXISTING WORK ALGORITHM- CUCKOO SEARCH ALGORITHM:**

Cuckoo Search Algorithm (CSA) [10] being idealized by its breeding behavior was tested on engineering optimization and embedded design problems. The results being obtained by this algorithm for engineering optimization problems are quite convincing in its results. Hence forth, Cuckoo Search Algorithm for FS is implemented. In this Cuckoo Search Algorithm addition of three constraints such as Eviction, Abandon and Survival is done. This type of method uses its historical memories for the location and status of the eggs being laid by the cuckoos.

This Algorithm Mainly Concentrates On Replacing Not Good Nests With The Potentially Good Nests. The position of the egg replaces the position of the random new eggs in another nest in three cases. If the eggs have been evicted, if the cuckoos abandon the nest or if the eggs have been hatched resulting to its survival.

- Hatching of eggs – Survival of the Cuckoo
- Abandoning the nest – Host bird abandons its nest and migrates to some other place to build another nest.
- Evicting the eggs – The host bird throws the cuckoo bird’s eggs.

Equation has been used as the main formula for the computation of Cuckoo Search Algorithm.

$$F_{ij} = \frac{\{(\alpha[I_i(next) - I_i(bp)]) * iter\}}{maxcuckoo}$$

Where, **Parameter settings for Cuckoo Search Algorithm**

$F_{ij}$	Fitness function used to find the alpha value of the cuckoo
$I_i(next)$	Accuracy of the cuckoo bird being selected
$I_i(bp)$	Accuracy of the host cuckoo bird
Iter	Fixed to 1
$\alpha$	Between 0 and 4
Maxcuckoo	Maximum number of cuckoos in a particular dataset

**3.3 EXISTING WORK ALGORITHM- HARMONY SEARCH ALGORITHM:**

The underlying principle behind this HSA algorithm [11] is has been that this algorithm might face the search on the grounds of Pitch, Amplitude and Timbre producing a perfect harmony. The initial random solutions that are considered may be far away from the feasible solutions. To get closer to the feasible and promising solutions, the exact solutions can be obtained by choosing only the amplitude of the tone. In this method, the best accuracy of Tunes is found with the frequency value of Tunes which varies with time (t) and computation is done for the new solution.

In HSA, Individual music player (variable which takes decision) runs a bit of music (value) which finds the best harmony (global optimum) at the last. Based on the Lambda value two constraints will be considered. Noise (high pitch) and Melody (low pitch). Equation has been used as the main formula for the computation of Amplitude in Harmony Search Algorithm

$$\lambda = \frac{c}{h_i[next] * h_i[music]}$$

Where, **Parameter Settings for Harmony Search Algorithm**

$\lambda$	Lambda Value (Amplitude)
$C$	Velocity of Light
$h_i[next]$	Accuracy of newly selected tune
$h_i[music]$	Accuracy of the existing tune that needs to be compared with the new tune

**4. Proposed work Algorithm- Whale Optimization Algorithm**

The whale optimization algorithm (WOA) is a novel meta- heuristics algorithm and it is a population based method. The main concept of the Whale optimization algorithm (Encircling prey) is about Humpback whales a. Encircling prey b. Exploitation phase: (bubble-net attacking method). c. Exploration phase: search for prey.

**a. Encircling prey:**

The Humpback whales know the location of prey and encircle them. They consider the current best candidate solution is best obtained solution and near the optimal solution. After assigning the best candidate solution, the other agents try to update their positions towards the best search agent as shown in the following equation

$$D = |C \cdot X^*(t) - X(t)|$$

$$X(t+1) = X^*(t) - A \cdot D$$

In this equation A, C, D, X are vectors. Where t is the current iteration, A and C are coefficient vectors, X\* is the position vector of the best solution, and X indicates the position vector of a solution, | | is the absolute value. The vectors A and C are calculated as follows:

$$A = 2a \cdot r - a$$

$$C = 2 \cdot r$$

Where components of a is linearly decreased from 2 to 0 over the course of iterations and r is random vector in [0; 1]

**b) Exploitation phase (bubble-net attacking method):**

The humpback whales attack the prey with the bubble-net mechanism. This mechanism is mathematical formulated as follow:

*Shrinking encircling mechanism:*

In this mechanism, the value of A is a random value in interval [-a, a] and the value of a is decreased from 2 to 0 over the course of iterations as shown

*Spiral updating position mechanism:*

In this mechanism, the distance between the whale location and the prey location is calculated then the helix-shaped movement of humpback is created as shown in the following equation

$$X(t + 1) = D' \cdot e^{bl} \cdot \cos(2\pi l) + X^*(t)$$

Where  $D' = |X^*(t) - X(t)|$  is the distance between the prey (best solution) and the  $i$ th whale,  $b$  is a constant,  $l$  is a random number in  $[-1, 1]$ .

The humpback whales used both mentioned two mechanisms when they swim around the prey. The mathematical model of these two mechanisms, we assume that there is a probability of 50% to choose between these two mechanisms to update the position of whales as follow

$$X(t + 1) = X^*(t) - A \cdot D \text{ when } p < 0.5$$

$$X(t + 1) = D' \cdot e^{bl} \cdot \cos(2\pi l) + X^*(t) \text{ when } p \geq 0.5$$

Where  $p$  is a random number in  $[0; 1]$

**c) Exploration phase:**

In the exploration phase, the humpback whales (search agents) search for prey (best solution) randomly and change their positions according the position of other whales. In order to force the search agent to move far away from reference whale, we use the  $A$  with values  $> 1$  or  $< -1$ . The mathematical model of the exploration phase is as follows

$$D = |C \cdot X_{rand} - X|$$

$$X(t + 1) = X_{rand} - A \cdot D$$

Where  $X_{rand}$  is a random position vector chosen from the current population.

**Whale Optimization Algorithm[13]**

1. Initialize the whales population  $X_i(i = 1, 2, \dots, n)$  [13]
2. Calculate the fitness of each search agent
3.  $X^*$  = the best search agent while ( $t < \text{maximum number of iterations}$ )
4. for each search agent
5. Update  $a, A, C, l$ , and  $p$
6. if ( $p < 0.5$ )
7. if ( $|A| < 1$ ) Update the position of the current search agent
8. elseif ( $|A| \geq 1$ ) Select a random search agent ( $X_{rand}$ )  
Update the position of the current search agent
9. end if
10. elseif ( $p \geq 0.5$ ) Update the position of the current search
11. end if
12. end for  
Check if any search agent goes beyond the search space and amend it  
Calculate the fitness of each search agent
13. Update  $X^*$  if there is a better solution
14.  $t = t + 1$
15. end while
16. return  $X^*$

**The WOA algorithm Explains**

Step 1. Firstly initialize the objective function of the algorithm and assign the best solution based on the size  $n$  of the whale.

Step 2. Start the count of iteration  $t$ .

Step 3. The starting count  $n$  is assigned randomly and individual agent  $X_i$  in the count is computed based on the fitness  $f(X_i)$ .

Step 4. Assign the best search agent  $X$ .

Step 5. The following steps are repeated until the termination criterion Satisfied.

Step 5.1. the iteration counter is increasing  $t = t + 1$ .

Step 5.2. All the parameters  $a, A, C, l$  and  $P$  are updated.

Step 5.3. the exploration and exploitations are applied according to the values of  $p$  and  $|A|$

Step 6. The best search agent  $X$  is updated.

Step 7. The overall process is repeated until termination criteria satisfied.

Step 8. Produce the best found search agent (solution) so far  $X$ .

This explains about the implementation of Whale Optimization algorithm and its behavior towards the Feature Selection process which helps in reducing the count of features in the data and gives the best accuracy compared to the existing algorithm.

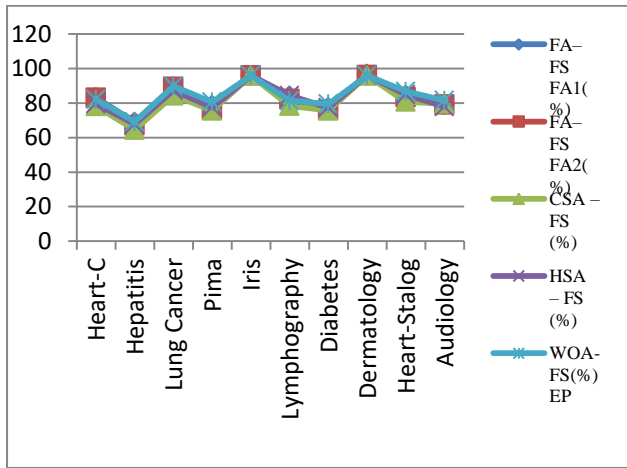
**5. EXPERIMENTAL SETUP AND RESULT ANALYSIS**

Fourteen standard datasets drawn from the UCI collection were used in the experiments. These datasets were chosen due to nominal class features. The number of attributes, instances and number of classes vary in the chosen dataset to represent different combinations. All the features in 10 fold cross validation is done through Weka tool The classifier used for evaluating the feature subsets generated is J48, Naïve Bayes and Logistic. Feature subset(FS) generation by Firefly Algorithm(FA), Cuckoo Search Algorithm(CSA) and Harmony Search Algorithm(HSA) is implemented using Net Beans IDE in the existing work. Feature subset generation by Whale Optimization algorithm (WOA) has been implemented using Net Beans IDE while the UCI dataset of Medical field is run through WEKA tool to get the classification rate which is processed through J48 classifier in the proposed work.

**Table.1 Comparison of all the accuracy of Existing with Proposed Algorithms**

Datasets	FA-FS FA1(%)	FA-FS FA2(%)	CSA - FS (%)	HSA - FS (%)	WOA-FS(%) EP
Heart-C	83.15	83.07	78.217	79.53	82.172
Hepatitis	69.03	67.00	64.516	67.74	<b>69.182</b>
Lung Cancer	89.50	89.33	84.375	87.37	<b>89.872</b>
Pima	78.70	76.10	75.651	77.47	80.610
Iris	<b>96.00</b>	<b>96.00</b>	<b>96.00</b>	<b>96.00</b>	<b>96.00</b>
Lymphography	<b>84.10</b>	82.25	78.378	84.43	81.404
Diabetes	77.47	76.00	75.65	77.21	<b>79.430</b>
Dermatology	<b>96.72</b>	96.17	95.901	96.07	95.985
Heart-Stalog	84.44	83.39	80.74	84.81	<b>86.775</b>
Audiology	79.051	79.201	79.646	77.87	<b>81.528</b>

It is clear from the table 1 the accuracy of the Whale Optimization Algorithm shows better results compared to other algorithms. EP\* Encircling Prey



**Fig.1 Graphical Representation of Accuracy of all these algorithms**

From Fig.1, we can observe that the Whale Optimization algorithm which is obtained for 10 UCI datasets showed better accuracy when compared to other algorithms.

**Table. 2 Comparison of all the features selected in Existing with Proposed Algorithms**

Datasets	FA-FS	CSA-FS	HSA-FS	WOA-FS
Heart-C	6	7	5	5
Hepatitis	8	10	9	9
Lung Cancer	14	20	18	12
Pima	3	3	4	3
Iris	2	3	2	2
Lymphography	8	7	10	7
Diabetes	4	5	6	5
Dermatology	23	22	25	22
Heart-Stalog	7	6	7	6
Audiology	53	55	49	47

In this table 2, it is clear that the number of features selected for Whale Optimization algorithm given is better when compared to other algorithms.

**6. INFERENCE**

From the table.1, We can infer that Whale Optimization Algorithm applied for UCI dataset for FS with respect to Accuracy gives better as well equal accuracies

We can infer that the features for this Algorithm get reduced as well as gets increased in some of the datasets in other Algorithms

By comparing Firefly Algorithm, Cuckoo Search Algorithm and Harmony Search Algorithm for FS with respect to Features: we can infer that the features for Whale Optimized Algorithm gets reduced.

**7. CONCLUSION**

The proposed system for FS Optimization to perform Classification has been implemented and tested using 10 datasets. The datasets are taken from UCI repository and Table 1 describes the 10 datasets that we have used. This paper has attempted to give accuracy results of the popular Meta-Heuristic algorithms. The optimized feature selection to perform classification using Meta Heuristic algorithms like ABC, Firefly, Cuckoo search and Harmony search algorithms in

which Firefly algorithms showed good results compared to other algorithm by giving the best accuracy rate with less number of features selected and also fine f-Measure value is noted. In proposed new and recent advanced optimized algorithms like Whale Optimization algorithm that showed comparatively better results with existing but with less number of features selected. We demonstrate the comparative results of the presently used algorithms with the existing algorithms thoroughly.

**References.**

- [1]. Tan, Steinbach, Kumar. (2005). "Introduction to Data Mining".
- [2]. Hassan AbouEisha et.al, (2018) "Extensions of Dynamic Programming for Combinatorial Optimization and Data Mining"
- [3]. Sunil Kawale, "Datamining and Optimization Techniques" International Journal of Statistika and Matematika", ISSN. 2277- 2790, E-ISSN. 2249-8605, Volume 6, Issue 2, 2013 pp 70-72
- [4]. Nidhi Tomar and Prof. Amit Kumar Manjhvar "A Survey on Data mining optimization Techniques" International Journal of Science Technology & Engineering | Volume 2 | Issue 06 | December 2015 ISSN (online). 2349-784X
- [5]. Basturk B, Karaboga D (2006) "An artificial bee colony (ABC) algorithm for numeric function optimization". IEEE Swarm Intelligence Symposium, 12–14 May, Indianapolis
- [6]. Bergh F, Engelbrecht AP (2006) "A study of particle swarm optimization particle trajectories". Inf Sci 176. 937–971.
- [7]. Rao, R. Venkata. "Teaching Learning Based Optimization Algorithm. And Its Engineering Applications". Springer, 2015.
- [8] Rao, R. Venkata, and V. D. Kalyankar. "Parameter optimization of modern machining processes using teaching-learning-based optimization algorithm". Engineering Applications of Artificial Intelligence 26, no. 1 (2013). 524-531.
- [9]. Shunmugapriya .P and Kanmani S, P.Sindhuja, G.Koperundeivi, V.Yasaswini, "Firefly Algorithm Approach for the Optimization of Feature Selection to Perform Classification", International Conference on Advances in Engineering & Technology, IEEE-ICAET 2014.
- [10]. Xin-She Yang, Suash Deb, "Cuckoo Search Via Levy Flights", World Congress On Nature and Biologically Inspired Computing (NaBIC 2009)
- [11]. Xin-She Yang and X. He. "Bat algorithm: Literature review and applications". International Journal of Bio-Inspired Computation, 5(3):141-149, 2013.
- [12]. Richardson, P.: The secrete life of bats. <http://www.nhm.ac.uk>
- [13]. S. Mirjalili and A. Lewis, *The Whale Optimization Algorithm*. Advances in Engineering Software, 95, (2016), pp.51-67.
- [14]. Dasgupta D, Zbigniew M, editors. "Evolutionary algorithms in engineering applications". Springer Science & Business Media; 2013.
- [15]. Jihong Liu,"A Hybrid Feature Selection Method for Data Sets of thousands of Variables" IEEE, 2010.

# Dental Plaque Segmentation in Digital Photographs using K-Means Clustering in L\*a\*b\* Color Space

Dr.R.Karthika devi <sup>1</sup>

Professor /ECE

Sethu Institute of Technology

mail-id: karthikadevicee@sethu.ac.in

D.Radhika<sup>2</sup>

Sethu Institute of Technology, Tamilnadu

**Abstract** - The objective of this work is to address and measure a novel method of quantitative measurement of plaque with improved sensitivity and reproducibility, using color models in digital dental plaque photographs. The computer-aided plaque detection and segmentation by K-means clustering in L\*a\*b\* color space is proposed. The proposed K-means clustering in L\*a\*b\* method for dental plaque detection is automated, quantitative, and objective, whereas the traditional methods using plaque indices are manual, semi-quantitative, and subjective. The execution time, the memory required and over-segmentation are reduced when compared to the existing method. The results obtained are producing reduced inter and intra-operator repeatability.

**Keywords** - Dental plaque segmentation, Digital photograph, L\*a\*b\* color space, K-means clustering

## I. INTRODUCTION

The proper and efficient care of the mouth is very crucial to a person's general health and overall appearance. The untreated dental disease can create tooth loss, gum problems and many other dental problems and dental issues. Dentists and physicians also trust that there could be a connection between dental disease and heart disease, premature births, diabetes, stroke, and low birth weight.

The dental diseases are generally classified into two primary groups. The first one includes those diseases of the teeth like the development of cavities and tooth decay. The tooth decay takes place when plaque is allowed to linger on the tooth's surface for a long period. The plaque pertains to a gluey substance, which contains bacteria that feed on the sugars consumed by a person. These bacteria created acid and absorbed through the tooth's surface, which creates the formation of cavities. Segmentation and detection methods are specifically valuable in the area of computer-aided medical diagnosis applications where visualization of the parts which are difficult to identify by human vision is the critical component [1].

## II. DENTAL PLAQUE

Dental plaque is a mass of bacteria that propagates on the tooth surfaces. At first, it is a very sticky and colorless deposit. It is mostly pale yellow or brown. It is mostly found in between the teeth, on the front of teeth, behind teeth, on chewing surfaces, along the gum line. The dentists can diagnose the dental plaque clinically. However, the traditional clinical diagnosis may have some limitations. The limitations include:

1. They are semi-quantitative and subjective.
2. They have less sensitivity and reproducibility in measuring the amount of plaque[1][2]

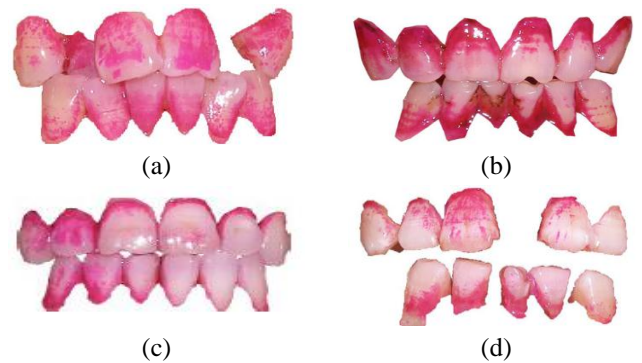


Fig. 1. Dental Plaque photograph

Using Computer Aided Detection or Diagnosis (CAD)[3][4] and designed image analysis system, a new method of fully quantitative and automatic measurement can be developed and evaluated with increased sensitivity and reproducibility[3][4]. Medical image analysis mostly deals with diagnostic issues, associate diagnosticians in their decisions. However, dentistry provides at least one more application for some image analysis techniques, which is Computer-Aided Diagnosis (CAD) systems for image enhancement and semi-automated image segmentation, but these systems establish one of the most important trends in modern dentistry, and this makes them one of the top dental image analysis applications. Computer-Aided diagnostics is

the second primary use of dental image analysis[4][5]. Generally, the dental plaque is colorless, so it is stained prior to assessment usually. The dental plaque is in red color after applying disclosed dyeing solution for 10s followed by a thorough mouth rinse. This makes the color of dental plaque obviously dissimilar from that of the other parts on the tooth surface as well as the background [4]. Based on the color reduction algorithm, the scheme of the dental plaque segmentation and quantification has five consecutive stage [4][5].

### III. LITERATURE SURVEY

The already existing methods are listed out in the table.

Table 3.1 Literature Survey

S. No.	Author	Work Type	Method	Color space	Imaging Modality/ view
1.	Smith <i>et al.</i> 2001	Semi-automatic	Analysed using Adobe Photoshop	RGB	Digital photographs/ anterior teeth
2.	Carter <i>et al.</i> (2004)	Automatic	Statistical analysis in HSI color space	HSI	Digital photographs/ anterior teeth
3.	Pretty <i>et al.</i> (2005)	Automatic	Planimetric techniques	RGB	Quantitative Light-fluorescence (QLF) images
4.	Kang Jiayin <i>et al.</i> (2006)	Automatic	Cellular Neural Network (CNN) associated with histogram analysis.	RGB	Digital photographs/ anterior teeth
5.	Kang Jiayin <i>et al.</i> (2007)	Automatic	FCM based classification in HSI space	HSI	Digital photographs/ anterior teeth
6.	Kang Jiayin <i>et al.</i> (2010)	Automatic	Modified Kernelized FCM	RGB	Digital photographs/ anterior teeth
7.	Kang Jiayin <i>et al.</i> (2010)	Automatic	Histogram Aided FCM	Gray scale image	Digital photographs/ anterior teeth
8.	Jiayin kang <i>et al.</i> (2010)	Automatic	Mean Shift based Algorithm	HSI	Digital photographs/ anterior teeth
9.	Oi-Hong Tung <i>et al.</i> (2008)	Automatic	Two-photon fluorescence imaging technique	Gray scale image	fluorescence image/single tooth

From the above literature survey there is no combination of K-means[6][7] with HSI color space is not proposed to find the dental plaque. So it is implemented and compared the response with the already existing method.

### IV. DENTAL PLAQUE SEGMENTATION USING K-MEANS CLUSTERING IN L\*a\*b\* COLOR SPACE

To detect the dental plaque region, the first step is to compute the Sample Colors in L\*a\*b\* Color Space for Each Region to visually distinguish the background, tooth region and plaque region colors from one another [8]. The CIE L\*a\*b\* supports quantification of these visual differences. The CIELAB color space is obtained from the CIE XYZ tri-stimulus values. The CIELAB have luminosity 'L\*' or brightness layer, chromaticity layer 'a\*' showing where color falls along the red-green axis, and chromaticity layer 'b\*' showing where the color falls along the blue-yellow axis. Categorize the Colors in 'a\*b\*' Space using K-Means Clustering and label every pixel in the image using the results from K-Means.

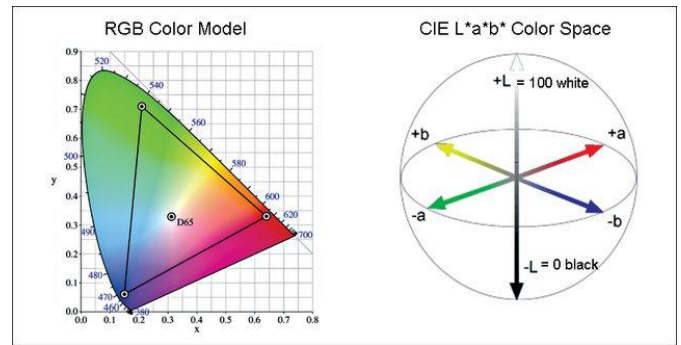


Fig. 2 RGB Color Model and CIE L\*a\*b\* Color space

### Steps in Dental Plaque Segmentation using K-means Clustering in L\*a\*b\* Color Space

1. Read the input plaque image.
2. Convert the RGB colors into L\*a\*b colors.
3. Separate the Colors in 'a\*b\*' Space Using K-Means Clustering is number of clusters to be grouped[9][10].

#### Algorithm 1 k-means algorithm

- 1: Specify the number  $k$  of clusters to assign.
- 2: Randomly initialize  $k$  centroids.
- 3: **repeat**
- 4:     **expectation:** Assign each point to its closest centroid.
- 5:     **maximization:** Compute the new centroid (mean) of each cluster.
- 6: **until** The centroid positions do not change.

4. Label every pixel in the plaque image using the results from K-means clustering.
5. Display the labeled images.

#### Flowchart for K-means Clustering in L\*a\*b\* Color Space

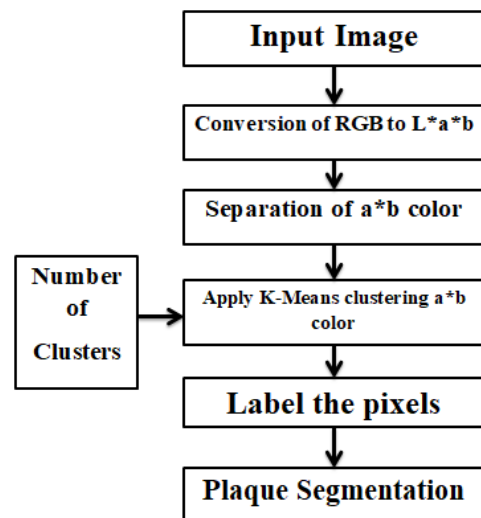


Fig. 3 Flow Chart of Dental plaque segmentation using K-means clustering in L\*a\*b\* color space

V. RESULTS AND DISCUSSION

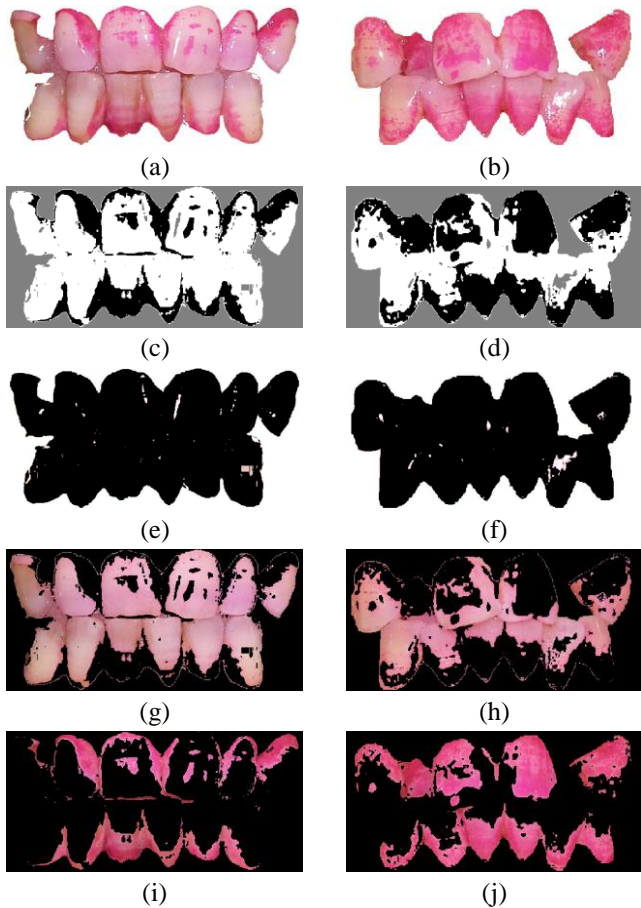


Fig. 4 Dental plaque segmentation using K-means clustering in L\*a\*b color space. (a,b) Original image (c,d) K-means clustered output in L\*a\*b space, (e,f) Cluster 1 (Background), (g,h) Cluster 2 (Tooth region), (i,j) cluster 3 (plaque region).

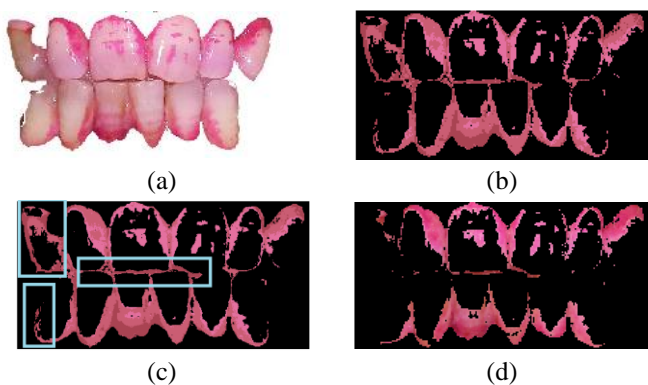


Fig. 5 Comparison of dental plaque detection algorithms (a) Original Image (b) Plaque Detection Using K-means Clustering in HSI color Space (c) Over segmentation in K-means Clustering in HSI color space (d) Plaque segmentation Using K-means Algorithm in L\*a\*b color space

Table 2 Dental plaque segmentation using K-means clustering in L\*a\*b color space

Parameters \ Images	Plaque area (pixels)-A	Full tooth area (pixels)-B	(A/B)*100 PPI(%)	Time (sec)
Teeth_1	24698	94573	26.11528	1.584475
Teeth_2	41387	85289	48.5256	1.585014
Teeth_3	3202	10640	30.09398	0.961103
Teeth_4	2400	9377	25.59454	1.020547
Teeth_5	9759	34085	28.63136	1.241804
Teeth_6	16930	35887	47.17586	1.198275

Table 3 Comparison of PPI with K-means with HSI color space

Teeth	PPI in K - Means + HSI	PPI in K - Means + L*a*b
Teeth_1	30.73	26.11
Teeth_2	58.26	48.52
Teeth_3	27.76	30.09
Teeth_4	18.22	25.59
Teeth_5	24.60	28.63
Teeth_6	52.47	47.17

Table 4 The Performance evaluation of Dental plaque detection Methods for Teeth\_1 image.

Method	Time (seconds)	Memory (bytes)	PPI (%)
FCM in HSI color space (Kang Jiayin <i>et al.</i> 2007)[10]	3.124	5.1740e+09	15.68
Adaptive K-means clustering (Zheng, Xin <i>et al.</i> 2018)[11]	2.710	5.1355e+09	25.76
K-means in HSI Color space (Proposed)	2.255	5.1762e+09	30.73
K-means in L*a*b color space(Proposed)	1.584	5.1421e+09	26.11



The main key for plaque segmentation is thresholding the S-plane. That disadvantage is eliminated by this K-means in  $L^*a^*b$  color space. The K-means clustering in  $L^*a^*b$  color space requires only the number of clusters. It does the operation on  $a^*b^*$ . The over-segmentation is eliminated and also faster than the K-means clustering in HSI color space. The plaque Percentage index is calculated and compared. To reduce the process time and memory requirement, the color reduction based algorithm is proposed. From the above comparison results Table 5.3, the execution time for K-means in  $L^*a^*b$  color space is less when compared to other methods and the plaque percentage index is more similar to the ground truth value.

## VI. CONCLUSION

The Dental plaque segmentation using K-means clustering in  $L^*a^*b$  color space is presented. There is a disadvantage in the Dental plaque segmentation using K-means clustering in HSI color space that is, the over-segmentation of plaque. Even though there is no plaque in the tooth, it shows that there is a plaque. The plaque Percentage index is calculated and compared with existing methods. The execution time for K-means in  $L^*a^*b$  color space is less when compared to other methods and the plaque percentage index is more similar to the ground truth value.

## REFERENCES

- [1] Chen, Q., Jin, X., Zhu, H., Salehi, H. S., & Wei, K. (2020). 3D distribution of dental plaque on occlusal surface using 2D-fluorescence-image to 3D-surface registration. *Computers in Biology and Medicine*, 123, 103860.
- [2] Smith, R. N., Brook, A. H., & Elcock, C. (2001). The quantification of dental plaque using an image analysis system: reliability and validation. *Journal of clinical periodontology*, 28(12), 1158-1162.
- [3] Carter, K., Landini, G., & Walmsley, A. D. (2004). Automated quantification of dental plaque accumulation using digital imaging. *Journal of dentistry*, 32(8), 623-628.
- [4] Kang, J., & Ji, Z. (2010, July). Dental plaque segmentation and quantification using histogram-aided fuzzy c-means algorithm. In *Proceedings of the 29th Chinese Control Conference* (pp. 3068-3071). IEEE.
- [5] Kang, J., Li, X., Luan, Q., Liu, J., & Min, L. (2006). Dental plaque quantification using cellular neural network-based image segmentation. In *Intelligent computing in signal processing and pattern recognition* (pp. 797-802). Springer, Berlin, Heidelberg.
- [6] Górriz, J. M., Ramírez, J., Lang, E. W., & Puntonet, C. G. (2006). Hard C-means clustering for voice activity detection. *Speech communication*, 48(12), 1638-1649.
- [7] Zheng, X., Lei, Q., Yao, R., Gong, Y., & Yin, Q. (2018). Image segmentation based on adaptive K-means algorithm. *EURASIP Journal on Image and Video Processing*, 2018(1), 68.
- [8] Cheng, H. D., Jiang, X. H., Sun, Y., & Wang, J. (2001). Color image segmentation: advances and prospects. *Pattern recognition*, 34(12), 2259-2281.
- [9] Chetrus, V, Ion, IR, Marsh, PD & Bradshaw, DJ 2013, 'Dental Plaque - Classification Formation and Identification', *Int. J. Med. Detistry*, vol. 3, pp. 139-143.
- [10] Jia-yin kang, Le-quan min, Qing-xian Luan, Xiao Li & jin-zhu 2007, 'Dental plaque quantification using FCM-based classification in HSI color space', *Liu Proceedings of the 2007 International Conference on Wavelet Analysis and Pattern Recognition*, pp. 2-4.
- [11] Zheng, X., Lei, Q., Yao, R. et al. Image segmentation based on adaptive K-means algorithm. *J Image Video Proc.* 2018, 68 (2018). <https://doi.org/10.1186/s13640-018-0309-3>
- [12] R. K. Devi, M. S. Dawood, R. Murugan, R. Lenamika, S. Kaviya and L. Vasini K., "Fuzzy based Regional Thresholding for Cyst Segmentation in Dental Radiographs," 2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS), Madurai, India, 2020, pp. 544-549, doi: 10.1109/ICICCS48265.2020.9121104.
- [13] Smith, RN, Brook, AH & Elcock, C 2001, The quantification of dental plaque using an image analysis system: reliability and validation', *Journal of Clinical Periodontology*, vol. 28, pp. 1158-1162.
- [14] Smith, RN, Rawlinson, A, Lath, D, Elcock, C, Walsh, TF & Brook, AH 2004, 'Quantification of dental plaque on lingual tooth surfaces using image analysis: reliability and validation', *Journal of Clinical Periodontology*, vol. 31, pp. 569-573.
- [15] Smith, RN, Rawlinson, A, Lath, DL & Brook, AH 2006, 'A digital SLR or intra-oral camera: preference for acquisition within an image analysis system for measurement of disclosed dental plaque area within clinical trials', *Journal of Periodontal Research*, vol. 41, pp. 55-61.
- [16] Kang, J., & Ji, Z. (2010, July). Dental plaque segmentation and quantification using histogram-aided fuzzy c-means algorithm. In *Proceedings of the 29th Chinese Control Conference* (pp. 3068-3071). IEEE.
- [17] Pretty, W, Edgar, M & Smith, PW 2005, 'Quantification of dental plaque in the research environment', *J. of dentistry*, vol. 33, pp. 193-207, 2005.
- [18] Jia-yin kang, Le-quan min, Qing-xian Luan, Xiao Li

- & jin-zhu 2007, 'Dental plaque quantification using fcm-based classification in HSI color space', Liu Proceedings of the 2007 International Conference on Wavelet Analysis and Pattern Recognition, pp. 2-4.
- [19] Jiayin Kang, ZJ 2010, 'Dental Plaque Quantification using Mean-shift-based Image Segmentation', IEEE International Symposium on Computer, Communication, Control and Automation
- [20] Kang Jiayin, JZ 2010, 'Dental Plaque Segmentation and Quantification using Histogram-aided Fuzzy C-Means Algorithm', Proceedings of the 29th Chinese Control Conference, pp. 29-31.
- [21] Kang, JY, Li, X, Luan, QX et al. 2006, 'Dental plaque quantification using cellular neural network-based image segmentation', Lecture notes in control and information sciences, vol. 345, pp. 797-802

## *Our Sincere Thanks to Technical Sponsors*



**Dr. N. Kumarappan**  
Chairman, IEEE Madras Section.



**Dr. S. Devaraj**  
Secretary, IEEE Madras Section.



**Dr. Joseph Gladwin**  
Treasurer, IEEE Madras Section.



**Dr. P. Sakthivel**  
Chairman, IEEE Computer Society,  
Madras Chapter.



**Dr. K.B. Jayanthi**  
Chairman, IEEE EMBS, Madras Section.



**Dr. B. Banu Rekha**  
Secretary, IEEE EMBS, Madras Section



# GLIMPSES OF 2020



## PROCEEDINGS OF 2021 IEEE SEVENTH INTERNATIONAL CONFERENCE ON BIOSIGNALS, IMAGES AND INSTRUMENTATION (ICBSII 2021)

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 Seventh International Conference on Bio Signals, Images and Instrumentation (ICBSII-2021) 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.

Research papers were received from diverse areas such as Physiological Modeling, Medical Imaging, Medical Robotics, Biomechanics, Biomedical Instrumentation and Nano-materials amounting to more than 200 papers. After a rigorous review process by an expert review committee, one third of the received papers that displayed quality in idea and work were selected for final presentation at the conference.

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 center, which works unanimously towards materializing it and they were instrumental in its success.

The Department of Biomedical Engineering, since its inception in 2005, has been a pioneer in the field of biomedical technology, instrumentation, and administration. The department has excellent infrastructure, experienced faculty members and motivated students. Department also has foreign collaborations which includes Birmingham City University, UK, Drexel University Philadelphia, and several industries such as L&T Medical System, Sri Ramachandra Medical College, Indian Biomedical Skill Consortium (IBSC), NIEPMD. Department has successfully conducted a retreat on “How to make India ready for 21st Century Medical device revolution?” with International and National delegates presenting their views on the topic. To add feather to the crown, the department has conducted six International conferences (ICBSII) in 2013, 2015, 2017, 2018, 2019, 2020 and two national conferences (NCABES) in 2014, 2016.

