

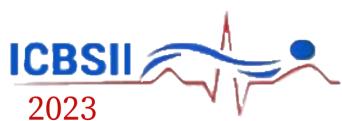
PROCEEDINGS OF THE
NINTH INTERNATIONAL CONFERENCE ON BIOSIGNALS,
IMAGES, AND INSTRUMENTATION

March 16 & 17, 2023

ICBSII 2023

Department of Biomedical Engineering

Sri Sivasubramaniya Nadar College of Engineering
Old Mahabalipuram Road, Kalavakkam - 603110



Proceedings of the 9th International Conference on Biosignals, Images and Instrumentation (ICBSII 2023)

Department of Biomedical Engineering

Sri Sivasubramaniya Nadar College of Engineering



In association with

Centre for Healthcare Technologies



March 16 - 17, 2023

Editorial Board

Chief Editor

Dr. A. Kavitha, Prof & HoD/BME

Co- Editors

Ms. B. Divya, AP/BME

Dr. S. Arun Karthick, ASP/BME

Contents

Message from the Chief Patron	xii
Dr. Kala Vijayakumar, President, Sri Sivasubramaniya Nadar Institutions	
Message from the Patron	xiii
Dr. V. E. Annamalai, Principal, SSNCE	
Message from the Conference Chair	xiv
Dr. A. Kavitha, Professor & Head, BME	
Message from the Organizing Chairs	xv
Dr. L. Suganthi, Associate. Prof/BME	
Dr. M Dhanalakshmi, Assistant Prof/BME	
Dr. R Nithya, Assistant Prof/BME	
Conference Organizing Committee	xviii
Technical Advisory Committee	xx
Review Committee	xxi

Chief Guest

Prof. Yuichi Kurita..... xxiv
Professor, Biological System Engineering Lab, Hiroshima University, Japan

Keynote Speakers

Dr. Rajeswari Aghoram.....	xxiv
Department of Neurology, JIPMER, Pondicherry	
Dr. Arthur McClelland.....	xxv
Centre for Nanoscale Systems, Harvard University, USA	
Dr. Yuvaraj Rajamanickam.....	xxvi
NTU, Singapore	
Dr. Sudhir Ganesan,	xxvii
Sri Ramachandra Institute of Higher Education and Research, Chennai	
Dr. David Belo.....	xxviii
Fraunhofer Portugal AICOS, Portugal	

Research Paper

Track 1

1. Variability of E-field in Dorsolateral Prefrontal Cortex Upon Change in Electrode Parameters in tDCS.	
Utkarsh Pancholi, Vijay Dave
	1
2. Factors Affecting the Efficacy of Retinal Implants	
Abdulrahman Alqahtani
	9
3. Detection of Atherosclerotic Plaque using Laser-Induced Breakdown Spectroscopy	
Srinidi S, Lohitaa J, Mercy A, Pauline John, Madan Mohan B, Nilesh J Vasa
	14
4. 2D Fluid-Structure Model of Aortic Valve Using a Derived Muscular Model Equation	
Yousef Alharbi
	18
5. Designing of IoT Based Portable Vital Health Parameter Monitoring System	
Krishnakumar. S, Najeeba. M, Fazila. A, Bethaney Janney J, Sindu divakaran, Sabarivani. A
	22
6. Design of low power & high-speed Multiway Merge Sorting Networks for Median Filter	
V. Mariselvam, S. R. Vishal, T. Surya, S. Vishnuram
	28

Research Paper

Track 2

1. Cuffless BP Measurement Using Single Site Photoplethysmography Ramakrishnan Maharajan	33
.....	
2. Simulation of a Non-Contact Triple-Wavelength Reflectance Pulse Oximeter to Measure Oxygen and Carbon Monoxide Saturations Swathi Veerabathraswamy, Pauline John, Nilesh J. Vasa	38
.....	
3. Smart Crop Protection System From Animals Using AI K Dharanipriya, S Sathyageetha, K S Sowmia, J Srinidhi	46
.....	
4. Digital Stethoscope For Instant Monitoring For Cardiac Auscultation R Chitra, N Jayapreetha, D Swetha, S Swetha	53
.....	
5. A Novel Approach In Web Based 3D Virtualization For Healthcare B Banu Rekha, K Harish Babu, S Charushree, V Vasunthra, P Divyadharshan	59
.....	
6. Classification Of Cultivars Employing The Alexnet Technique Using Deep Learning K. Gurunathan, V. Bharathkumar, M. Haji Ali Meeran, K. Hariprasath, R. Jidendiran	65
.....	
7. Eye Blink Based Biometric Authentication System G Umashankar, G. Hari Krishnan, T. Sudhakar, G Mohandass, J R. Ramya, V. Devika, Shaina Banu S	72
.....	

Research Paper

Track 3

1. Analysis of Raw 3D Images of Stages of Alzheimer's Disease using Deep Learning Thamizhvani T R, R J Hemalatha, Rachel Cynthia V, Swetha S	76
2. A Deep Learning Framework for Semantic Segmentation of Nucleus for Acute Lymphoblastic Leukemia Detection Prasanna A, Saran S, Manoj N M, Alagu S	81
3. Vision Based Real-time Active Protection System Using Deep Convolutional Neural Network Rubin Bose S, Varun Sharma K, Karrthik Kishore V, Tharunraj S, Nikhil Srinivas G, Regin R	89
4. Transfer Learning Approach for Epizootic Ulcerative Syndrome and Ichthyophthirius Disease Classification in Fish Species Vijayalakshmi M, Sasithradevi A, P Prakash	95
5. Prediction of New Causing Proteins for Alzheimer's, Parkinson's, and Huntington's diseases: Protein – Protein Interaction Analysis Sai Gopala Swamy Gadde, Srinivas Kudipudi, Bhargavi Vaka, Sai Kumar Vadapalli	100
6. Segmentation of Cell Objects in IIF Images using Split-Bregman Global Convex Optimization Kanchana Devanathan, Yasavvi Prasad K and Sai Amarnath G	107

Research Paper

Track 4

1. Facial Expression Recognition using Convolutional Neural Network Alben Richards M J, Prakash P, Kaaviya Varshini E, Kasthuri P, Diviya N	113
2. Infrared Thermograms for Diagnosis of Dry Eye: A Review Persiya J, Sasithradevi A, S Mohamed Mansoor Roomi	118
3. Design of an Automated Healthcare System Using Hand Gestures for Paraplegic People Bethanney Janney J, Renny Valentina, Aakesh U, S Krishnakumar, Yaswant Rajasekaran, T Sudhakar	124
4. Development of a Fiber-Optic Spectral Domain based Optical Coherence Tomography in Near Infrared Range for Biomedical Applications Richa Parihar, Nilesh J Vasa, Joy Mammen, Pauline John	129
5. An Analysis of the Design and Evolution of Braces for Lower Back Pain Bethanney Janney J, Ponni S, Kishore R, Meenatchi S, Keerthivasan V, Saraswathi B	133
6. Sitting posture Analysis using CNN and RCNN Nishitha R, N R Krishnamoorthy, T Sudhakar, Dhanalakshmi K, Bethanney Janney J, Vigneshwaran S	137

Research Paper

Track 5

1. Parkinson's Disease Detection and Classification of Stages from Drawing Patterns Using Deep Learning Karthikeyan B, Jerolin Futina J B, Pandia Abinaya P, Shanmuga Priya R	142
2. Convolutional Neural Networks for Image Emotion Recognition by fusing Differential and Supplementary Information Shashank Rapolu, Aman Singh, Ayush Dhingra, Ajaz Hussain and Tushar Sandhan	149
3. A Novel Design for Automatic Measurement of Reaction Time for Audio Visual and Muscular Stimulus E.Sathish, V. Kiruthika, Keerthika.N	154
4. Active Feedback using Riemannian Features for Motor Imagery Classification Kurusetti Vinay Gupta, Rishi Mittal, Durgesh Ameta	160
5. Statistical and Deep Convolutional Feature Fusion for Emotion Detection from Audio Signal Durgesh Ameta, Vinay Gupta, Rohit Pilakkottil Sathian, Laxmidhar Behera, Tushar Sandhan	166
6. A Review on Preprocessing of EEG Signal Suveetha, Dhanaselvam P, Nadia Chellam C	173

Research Paper

Track 6

1. Wearable Device for Early Detection of Muscle Cramp with Heating Pad

Prasanna K M, Raagini R, Sadhana K, Dr Revathi J

..... 180

2. Segmentation and Severity Classification of Dementia in Magnetic Resonance Imaging using Deep Learning Networks

Sarath Vignesh A, Denicke Solomon H, Dheepan P, Kavitha G

..... 185

3. Extraction, Processing and Analysis of Surface Electromyogram Signal and Detection of Muscle Fatigue using Machine Learning Methods

D V Pravin, A J Ragavkumar, S Abinesh, G Kavitha

..... 191

4. Resampling-free Fast Particle Filtering with Application to Tracking Rhythmic Biomedical Signals

Mohammed Ashik, Ramesh Patnaik M, Praveen Babu Choppala

..... 199

5. Design of an IP Core for Motion Blur Detection in Fundus Images Using an FPGA-based Accelerator

Rohit Jacob George, Charaan S, Swathi R, S. P. Joy Vasantha Rani

..... 205

6. Design of DCT Hardware Accelerator using FPGA for Medical Image Authentication

Swetha V, Evangeline Divya Sagayee G, Brintha J S, Karthika K, Joy Vasantha Rani S P

..... 211

Dedicated to

All the Staff and Students

of

Department of Biomedical Engineering

From the Chief Patron

Dr. Kala Vijayakumar

President, SSN Institutions



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

Biomedical Engineering is a multi-disciplinary branch of study which brings together healthcare and technology, the front runners in the modern world. I congratulate the Department of Biomedical Engineering for organizing the Ninth International Conference on Biosignals, Images and Instrumentation (ICBSII 2023), 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 worldwide. 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.

Dr. Kala Vijayakumar,
Chief Patron,
ICBSII – 2023

From the Patron

Dr. V. E. Annamalai

Principal, SSN College of Engineering



I am pleased that the department of Biomedical Engineering is organizing the Ninth International Conference on Biosignals, Images and Instrumentation (ICBSII 2023), in association with the Centre for Healthcare Technologies of SSNCE, in a manner befitting the biomedical engineer's community.

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 involvement in a wide range of activities with students and faculty has strengthened it, allowing students to get a thorough understanding of industrial requirements to the fullest extent possible.

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

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

Dr. V. E. Annamalai
Patron,
ICBSII – 2023

From the Conference Chair

Dr. A. Kavitha

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



Educating the youth aids the growth of literacy and avoids the spread of misinformation. It helps people learn about the community as a whole and develop informed opinions on many issues around the world. With such opinions, we can support anyone in need and create a positive impact on our communities. That being said, it gives me immense pleasure to present the Ninth International Conference on Bio-signals, Images, and Instrumentation (ICBSII 2023).

Biomedical engineering is a discipline that bridges many fields such as biology, physical sciences, and engineering in order to improve medicine and human health. This benefits society by undertaking research that creates quantitative correlations across scales in the human body and then applies that knowledge to the development of new instruments to improve human health. The impact and significance of findings keeps growing progressively. Workshops, seminars, project exhibitions, and guest lectures on current trends relevant to the core and multidisciplinary disciplines in biomedical engineering are often offered by the department to help students obtain a thorough understanding of industry requirements.

The 9th International Conference on Biosignals, Images, and Instrumentation is being organised by the Centre for Healthcare Technologies, a multidisciplinary research initiative focusing on research through innovation in healthcare, in collaboration with the Department of Biomedical Engineering, with the goal of instilling research aptitude in students and providing a great platform for researchers to showcase their work in various domains of healthcare. It is rightly said, “A group becomes a team when each member is sure enough of himself and his contribution to praise the skills of others.” – Norman Shidle.

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

Dr. A. Kavitha
Conference Chair,
ICBSII – 2023

From the Organizing Chair

Dr. L. Suganthi

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



The advancements in biomedical engineering have revolutionized the way we approach healthcare, and the ICBSII-2023 conference presents a fantastic opportunity to bring together experts and professionals in this field to exchange knowledge, share ideas, and collaborate on future innovations. Technology has transformed the health business and made numerous advances to our lives, including a decrease in the death rate from sickness and the ability to replicate an organ. It's crucial to make sure that modifications are maintained because they occur frequently.

It's fascinating to see how technology will advance, adapt, and develop in the healthcare industry. Collaboration with specialists and academics from various fields of biomedical engineering is possible at this conference. The multidisciplinary nature of this conference promises to promote a holistic approach to healthcare, which will ultimately benefit patients worldwide.

I extend my sincere gratitude to the management of SSN College of Engineering, our president Dr. Kala Vijayakumar SSN Institutions, and Dr. V. E. Annamalai sir - Principal, SSN College of Engineering for granting the department a delightful opportunity to organize this brilliant event that enables us to grow on a global level. I would like to thank our department Head and organizing chair Dr. A. Kavitha, and my fellow organizers of the conference, and all my colleagues in the ICBSII-2023 co coordinating committee for their support, constructive feedback, timely, tireless, and meticulous efforts in conducting this online event. 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. I appreciate the efforts of everyone involved in making this conference possible. It takes a lot of dedication and hard work to organize an event of this magnitude. I am confident that this conference will be a great success and will pave the way for future developments in biomedical engineering.

Dr. L. Suganthi
Organizing Chair
ICBSII – 2023

From the Organizing Chair

Dr. R. Nithya

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



The Ninth International Conference on Biosignals, Images, and Instrumentation is a significant event in the field of Biomedical Engineering, and I am excited about the opportunities it offers to the participants. This conference provides a platform for researchers, academicians, and industry professionals from around the world to share their ideas and expertise, exchange knowledge, and collaborate to drive innovation in healthcare. It is a remarkable achievement to bring together such a diverse group of professionals under one roof and create an atmosphere that promotes learning, networking, and collaboration.

The scientific tracks cover the latest developments and trends in Biomedical Engineering, which will benefit the attendees, especially students and researchers. The workshops, plenary lectures, keynote lectures, and invited lectures by renowned experts will provide valuable insights into the advancements in healthcare research.

I would like to extend my gratitude to Dr. Kala Vijayakumar - President, SSN Institutions and Dr. V. E. Annamalai - Principal, SSN College of Engineering for the encouragement and support offered to our department to organize this delightful event. I would like to thank our Head and organizing chair Dr. A. Kavitha for her belief in me and providing an opportunity to serve as a member of the organizing committee. My heartfelt thanks to my colleagues who are fellow organizers and committee members, supporting staff and student volunteers of the Biomedical Engineering department whose shared responsibilities and teamwork has paved for a smooth organization of this prestigious event.

Dr. R. Nithya
Organizing Chair,
ICBSII – 2023

From the Organizing Chair

Dr. M. Dhanalakshmi

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



The healthcare industry has seen a growing focus on the field of Biomedical Engineering, as new advancements aim to improve individuals' health and well-being. These innovations have led to futuristic technologies that are transforming healthcare, from diagnosis and analysis to treatment and recovery. With the constant evolution of healthcare needs, an international forum that facilitates the exchange of ideas and information is crucial. It is with great pleasure that we announce the eighth international conference on the latest innovations in Bio-signals, Images, and Instrumentation (ICBSII 2023), to be held from March 16-17, 2023 at SSN College of Engineering.

The conference agenda includes seminars, plenary talks, keynote lectures, and invited lectures from national and worldwide authorities in healthcare research. In order to give students, researchers, academic professionals, and industrialists the best chance to improve their understanding of the most recent multidisciplinary techniques in biomedical engineering, parallel technical sessions have also been planned. Also, the conference offers a helpful forum for establishing new connections in the field of biomedical engineering, with several opportunities for networking with sector leaders.

We are delighted to have received an overwhelming response from the research community in India and abroad to participate in our conference. We received over 100 research articles, and only close to 40% of the submissions were accepted after a rigorous review process extended by national and international reviewers with unmatched expertise in relevant fields. This was done to ensure the quality of the presented papers. We are excited to announce that the selected papers will be published in the IEEE, adding one more feather to our cap.

I would like to express my gratitude to all the reviewers and authors who contributed to the papers. We could not have organized such a successful event without their valuable contributions. I also extend my thanks to Dr. Kala Vijayakumar, President of SSN Institutions, and Dr. V. E. Annamalai, Principal of SSN College of Engineering, for their encouragement and support in organizing this event. Additionally, I would like to thank our Head and organizing chair Dr. A. Kavitha for her belief in us and providing an opportunity to serve as a member of the organizing committee. Finally, I would like to thank our colleagues who are fellow organizers and committee members, supporting staff, and student volunteers of the Biomedical Engineering department, whose shared responsibilities and teamwork have paved the way for the smooth organization of this prestigious event.

Dr. M. Dhanalakshmi
Organizing Chair,
ICBSII – 2023

Conference Organizing Committee

Chief Patron

Dr. Kala Vijayakumar, President, SSN Institutions

Patron

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

Conference Chair

Dr. A. Kavitha, Professor and HOD/BME

Organizing Chairs

Dr. L. Suganthi

Dr. M. Dhanalakshmi

Dr. R. Nithya

Conference Treasurer

Dr. R. Subashini

Information Contact

Dr. S. Saranya

Dr. Sachin Gaurishankar Sarate

Publication Chair

Ms. B. Divya

Dr. S. Arun Karthick

Technical Program Chair

Dr. N. Venkateswaran

Dr. K. Nirmala

Dr. S. Pravin Kumar

Dr. B. Geethanjali

Dr. J. Vijay

Supporting Staff

Ms. J. Suganthi

Mr. E. Sivananthan

Mr. A. Udayachandran

Mr. T. R. Sasidharan

Ms. Gayathri

Technical Advisory Committee

International Advisory Committee Members

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

National Advisory Committee Members

- **Dr. S. Ramakrishnan**, IIT Madras
- **Dr. Nilesh J Vasa**, IIT Madras
- **Dr. Renu John**, IIT Hyderabad
- **Dr. N. Sujatha**, IIT Madras
- **Dr. K. Mohanavelu**, DEBEL, DRDO, Bangalore
- **Dr. G. Sudhir**, Orthopaedic Spine Surgeon, SRMC, Chennai
- **Dr. Sasikala**, Anna University, Chennai
- **Dr. Venugopal**, APJ Abdul Kalam Technological University - Kerala
- **Mr. Balamurugan**, HCL
- **Dr. B. Banu Rekha**, PSG College of Technology, Coimbatore.
- **Dr. Hariharan**, NIT Uttarakhand
- **Dr. Vaithiyanathan Dhandapani**, NIT Delhi
- **Dr. K K Deepak**, AIIMS Delhi
- **Dr. Gnanou Florence Sudha**, Puducherry Technological University
- **Dr. C. Arunachalaperumal**, Chairman, IEEE Photonics society
- **Dr. K. Porkumaran**, Chairman, IEEE Madras section
- **Dr. K. B. Jayanthi**, Chairman IEEE EMBS, K.S. Rangasamy College of Technology

Review Committee

- **Dr. Arun Karthick S**, Associate Professor, SSN College of Engineering, Chennai
- **Dr. M C Jobin Christ**, Professor, Rajalakshmi Engineering College, Chennai
- **Dr. B Geethanjali**, Associate Professor, SSN College of Engineering, Chennai
- **Dr. C Vinoth Kumar**, Associate Professor, SSN College of Engineering, Chennai
- **Dr. Dhanalakshmi M**, Assistant Professor, SSN College of Engineering, Chennai
- **Dr. R Subashini**, Assistant Professor, SSN College of Engineering, Chennai
- **Dr. S Pravin Kumar**, Associate Professor, SSN College of Engineering, Chennai
- **Dr. K. Palani Thanaraj**, Associate Professor, St. Joseph's College of Engineering, Chennai
- **Dr. Bethanney Janny**, Assistant Professor, Sathyabama Institute of Science and Technology, Chennai
- **Dr. Josephine Julina**, Assistant Professor, SSN College of Engineering, Chennai
- **Mr. Umashankar G**, Assistant Professor, GRT Institute of Engineering & Technology, Tiruttani
- **Dr. Sachin G**, Assistant Professor, SSN College of Engineering, Chennai
- **Dr. Harikrishnan G**, GRT Institute of Engineering & Technology, Tiruttani
- **Dr. Durga G**, Associate Professor, SSN College of Engineering, Chennai
- **Mr. Manikandan K**, Assistant Professor, Sona College of Technology, Salem
- **Dr. Mariammal K**, Assistant Professor, Anna University, Chennai
- **Dr. A K Gnanasekar**, Professor, Rajalakshmi Institute of Technology
- **Dr. Suveetha Dhanaselvam**, Associate Professor, Velammal College of Engineering & Technology, Madurai
- **Ms. Angel Deborah**, Assistant Professor, SSN College of Engineering, Chennai
- **Dr. Vaithyanathan D**, Assistant Professor, National Institute of Technology, Delhi
- **Ms. Divya B**, Assistant Professor, SSN College of Engineering, Chennai
- **Dr. E. Priya**, Professor, Sri Sairam Engineering College, Chennai

- **Dr. A. K. Jayanthi Ph.D.**, Professor, SRM University, Chennai
- **Dr. N Venkateshwaran**, Professor, SSN College of Engineering, Chennai
- **Dr. S Saranya**, Assistant Professor, SSN College of Engineering, Chennai
- **Dr. K Nirmala**, Associate Professor, SSN College of Engineering, Chennai
- **Dr. Nithya R**, Assistant Professor, SSN College of Engineering, Chennai
- **Dr. Suganthi L**, Associate Professor, SSN College of Engineering, Chennai
- **Dr. Vijay J**, Associate Professor, SSN College of Engineering, Chennai
- **Dr. Pauline John**, Postdoctoral Associate, New York University, Abu Dhabi
- **Dr. Punitha N**, Assistant Professor, SSN College of Engineering, Chennai
- **Dr. Joe Louis**, Associate Professor, SSN College of Engineering, Chennai
- **Dr. Anbuselvi M**, Associate Professor, SSN College of Engineering, Chennai
- **Dr. Moorthi M**, Professor & HOD, Saveetha Engineering College
- **Dr. Suchetha M**, Professor, Vellore Institute of Technology, Chennai
- **Dr. Padmapriya N**, Associate Professor, SSN College of Engineering, Chennai
- **Dr. Priyadarshini R**, Associate Professor, SSN College of Engineering, Chennai
- **Dr. Manikandan S**, Professor, PSR Engineering College, Sivakasi
- **Dr. Sasirekha S**, Associate Professor, SSN College of Engineering, Chennai
- **Dr. Saraswathi S**, Associate Professor, SSN College of Engineering, Chennai
- **Dr. Sheeba Santhosh**, Assistant Professor, Panimalar Engineering College, Chennai
- **Dr. Mercy Theresa**, Associate Professor, Saveetha School of Engineering
- **Dr. Ivan Corazza**, Assistant Professor, Diagnostic and Specialty Medicine - DIMES, Bologna, Italy
- **Dr. Aparna Neetiyyath**, Fullbright, Nehru fellow, Harvard Medical School, USA

**Ninth International Virtual
Conference on Biosignals,
Images and Instrumentation
(ICBSII 2023)**

KEYNOTE SPEAKERS' PROFILE

PROFILE:



Prof. Yuichi Kurita

Professor, Biological System Engineering Lab, Hiroshima University, Japan

Dr. Yuichi Kurita is a professor at Hiroshima University's Graduate School of Advanced Sciences and Engineering. He earned both his master's and doctorate degrees from Nara Institute of Science and Technology. His research interests include human informatics, perceptual information processing, biomedical engineering, rehabilitation sciences, assistive robotics, and haptics. He is a member of several academic organizations, including the Society of Instrument and Control Engineers, the Robotics Society of Japan, the Virtual Reality Society of Japan, the Japan Society of Mechanical Engineers, the Information Processing Society of Japan, and IEEE. He has approximately 240 articles to his name, including six book volumes. He has also given several oral and poster presentations. His accomplishments include honorable mentions at the international conferences on advanced mechatronics. He has also served as an editor and reviewer for a number of internationally renowned peer-reviewed publications. He has been meticulous in planning a variety of academic events and conferences, including robotic symposia, augmented humans 2020 and superhuman sports design competition



Dr. Rajeshwari Aghoram

Associate Professor, Department of Neurology, JIPMER, Pondicherry

PROFILE:

Dr. Rajeshwari Aghoram is an associate professor in the department of neurology at the Jawaharlal Institute of Postgraduate Medical Education and Research Institution (JIPMER) in Pondicherry. She has almost 10 years of clinical expertise in neurology. Her research interests include clinical neurology, epilepsy, neuro epidemiology, EEG electrophysiology, and medicinal modeling. She has almost 30 papers to her name, including two Cochrane reviews. She has presented over 100 talks at different national and

PROFILE:

Dr. Arthur Mc Clelland is a principal scientist at the Harvard University supporting researchers across disciplines including physics, chemistry, material sciences, electrical engineering, molecular biology, cell biology and zoology with material characterisation completed his Bachelor of Science in engineering physics from University of Pittsburgh and Master of science in Electrical Engineering from University of Michigan. His doctorate research work was on the spectroscopic investigation of buried polymer and organic surfaces. He completed his postdoctoral research from North Eastern University. He has served as a physics instructor for 3 years at WISE – Girls in Science and Engineering summer program. He was one of the 30 technical staff at the Center for Nanoscale Systems, the largest core lab at the Harvard University from 2012 – 2016 where he supported around 450 active users on scientific instruments and experimental design. He also has been the lecturer at the Harvard university extension school from 2013 to 2016 where he played a key role in delivering lectures on optical microscopy. Currently, he is the principal scientist at the center of CNS and is responsible for providing technical training about the information that can be collected from a sample through optical microscopy.



**Dr. Arthur
McClelland**

*Principal Scientist, Center for
Nanoscale Systems, Harvard
University*

PROFILE:



**Dr. Yuvaraj
Rajamanickam**

*Education Research Scientist, NTU,
Singapore*

Dr. Yuvaraj Rajamanickam is a Research Scientist in the National Institute of Education (NIE) at Nanyang Technological University (NTU), Singapore. He completed his undergraduate in Electronics and Communication Engineering and his post-graduation in Biomedical Engineering from Anna University. He was also the recipient of a gold medal during his post-graduation. He was conferred the doctoral degree in Biomedical Electronics from University Malaysia Perlis in 2015. Before joining NIE, he was a Research Fellow in the School of Electrical and Electronic Engineering at NTU and the Neural Systems Lab at University of Kentucky, USA. His primary research interests are machine learning algorithms, with applications to EEG signal analysis that include affective computing, brain-machine interface, neurological disorder diagnosis, and cognitive neuroscience. His current research works focus on affective computing in education using neurophysiological signals. He has received several innovative research awards, medals, and certificates on excellent publications and research. He has published over 30 research papers in refereed international journals and also has 2 patents credited to his name. He serves as a reviewer of IEEE Transactions on Neural Systems and Rehabilitation Engineering, NeuroImage, Neuroscience Letters, IEEE Journal of Biomedical and Health Informatics, and IEEE Transactions on Affective Computing journals. Furthermore, he is a fellow of IEEE, Society for Neuroscience, Brain-Computer Interface Society, and Cognitive Neuroscience Society.

PROFILE:

Dr. Sudhir Ganesan is a senior consultant spinal surgeon in Sri Ramachandra Institute of Higher Education and Research. He is specialized to deal with all types of spinal pathologies like degenerative spinal disorders, spinal deformities, infections, tumors and fractures of spine. He has received MBBS, DNB(Ortho), MNAMS, FNB(Spine Surgery) Spine Surgery Training in QMC, Nottingham, UK and Endoscopic Spine Surgery Fellowship in Germany. As part of his research, he has examined Type 2 diabetes induce early senescence and degeneration in human intervertebral disc - A tissue biomarker evaluation; The role of Platelet rich plasma in spinal fusion surgeries; Analysis of factors & their correlation with ligamentum flavum hypertrophy in lumbar spine; Clinical & radiological outcome of delta fixation vs reduction fixation in high grade spondylolisthesis. His high academic record includes many distinctions and gold medals like University second in pathology, top rank in anatomy, medicine, surgery, distinction in ophthalmology and pharmacology. He received the Dr. Balu Sankaran Gold Medal for securing All India First Rank in DNB Orthopedics. He excelled as the Best Outstanding Fellow in Sir Ganga Ram Hospital, from where he pursued his FNB in Spine Surgery. He has presented many papers, posters and has participated in various State, National and International Conferences.



Dr. Sudhir Ganesan,

*Consultant Spine Surgeon,
SRIHER*

PROFILE:



Dr. David Belo

*Senior Scientist - Machine Learning
for Time Series Fraunhofer Portugal
AICOS, Portugal*

David Belo, PhD is a Senior Scientist in Machine Learning for Time Series at Fraunhofer Portugal AICOS and Coordinator of the AISYM4MED Horizon Europe project. He has a Ph.D. in Biomedical Engineering from the NOVA University of Lisbon, obtained in 2021, and his specialties include Deep Neural Networks, Keras, TensorFlow, Biosignal Processing, Machine Learning, and Project Management. David is a Portuguese Biomedical Engineer with a passion for technology, development, health, and the human body. He has a vast and multidisciplinary experience in both academia and industry, and his career is focused on developing innovative solutions for healthcare using machine learning. In addition to his academic and industry experience, David participated in the NASA Frontier Development Lab in 2019, where he helped develop AI tools for aiding the Astronaut Health program. As a Senior Scientist, David manages and develops R&D projects for applications using machine learning. He has also gained expertise in AI architectures, AWS solutions architecture, and leadership, among others. David is a creative and persistent individual with strong leadership skills, who believes that everyone in a team has a role to play.

TRACK 1

**Ninth International Conference in Biosignals, Images and Instrumentation,
SSNCE, Chennai, March 16 – 17, 2023**

Variability of E-field in dorsolateral prefrontal cortex upon a change in electrode parameters in tDCS.

Utkarsh Pancholi

Research Scholar, Biomedical Engineering Department, L.D.
College of Engineering, Ahmedabad-380015
Gujarat Technological University, Gujarat, India
utkarshpancholi6387@gmail.com

Vijay Dave

Research Supervisor, Biomedical Engineering Department,
Government Engineering College, Sector-28, Gandhinagar-382028
Gujarat Technological University, Gujarat, India
vijaydave12@gmail.com

Abstract - Transcranial direct current stimulation (tDCS) is a part of the transcranial electrical stimulation method widely used for treating patients with neurological and psychological abnormalities, along with application in cognitive improvements. With a simple design and operating procedure, tDCS is considered a safe and effective therapy choice. With predefined treatment protocols, it is possible to achieve the required electric field within the inner structures of the brain to excite and inhibit neuronal activity and its outcomes. The generated electric field shows variation among individuals due to anatomical and functional changes in the brain tissues. In-situ modeling of therapeutic procedures can help to assure the probabilistic outcome of tDCS. In this study, we have obtained results for electric field strength variability in a cognitively normal subject. We have simulated the subject with variation in stimulating electrode size and shape, a combination of electrode-Gel and electrode-sponge with SimNIBS Ver (3.2.6), and measured electric field strength and focality. Simulated results show less dependence on gel or sponge thickness and more reliance on electrode size and shape for E-field and focality. The increasing size of electrodes reduces electric field strength and focality with asymmetrical E-field distribution, whereas decrement generates a more symmetrical and focused E-field with higher strength.

Index Terms – Transcranial direct current stimulation, Computational Modelling, Electric field strength, Focality.

I. INTRODUCTION

External non-invasive electrical stimulation on the human body has varying stimulation modalities, i.e., electrotherapy [1], electro-convulsive therapy [2], functional electrical stimulation [3], transcranial electrical stimulation (Transcranial alternative current stimulation [4] and Transcranial direct current stimulation [5]), transcranial magnetic stimulation [6] and vagus nerve stimulation [7]. Other invasive therapy includes cortical and subcortical therapeutic intervention through implantable electrodes and intrathecal delivery pumps or micro-pumps [8]. Invasive modalities require minimal surgery for inserting stimulation electrodes, whereas non-invasive methods do not require surgery or implantation. In non-invasive therapeutic intervention, tDCS and tACS are considered safe and easy to use methods used for neuromodulation applications. Transcranial alternating stimulation uses a small alternating current of frequency < 100 Hz applied externally on the subject's scalp to alter or enhance the brain's function [9].

Transcranial direct current stimulation uses a minimal amount of direct current (1-2mA) [10] on the pre-specified positions on the subject's scalp for a fixed duration of time. Application of current generates a local electric field on a target beneath the stimulation electrode due to resistance offered by the brain tissues, i.e., white matter, grey matter, cerebrospinal fluid, blood vessels, skin, scalp, and skull bone. Electric fields here are a vector quantity that possesses both magnitude and direction. These quantities vary in different subjects as individuals carry different brain tissue conductivities. This difference encourages researchers to find the outcome and after-effects of tDCS stimulation using in-situ platforms to avoid unnecessary exposure of electric field to the tissues with non-interest. Simulation tools are available to analyze electric field distribution on a given target with the measure of strength and focality of the field. As the therapy outcome changes subject to subject with the same treatment protocol, it must be analyzed with different stimulation protocols.

Stimulation parameters, i.e., direct current intensity, electrode size, electrode shape, and combination of electrode-sponge and electrode-gel, should be considered before the application on a target location. These selections will accordingly change the behavior of neuromodulation in a subject. The combination of electrode-sponge or electrode-gel induces depolarization and repolarization in the cerebral cortex using anodal and cathodal stimulation, respectively. The two-electrode stimulation method is the bipolar method currently used by available tDCS devices in the market. The placement of anode and cathode is decided by 10–20 EEG electrode placement method [11]. Current delivered by anode and cathode changes concentration of calcium (Ca^{2+}) ions intracellularly which changes behaviour of neurons and their excitability [12]. Generated electric field does not produce action potentials rather alters characteristics of neurons under the target location.

Gaurav sharma et al. studied the effects of conventional bipolar stimulation and high-definition stimulation (HD-tDCS) with 4×4 , 3×3 , and $2\times 2 \text{ cm}^2$ electrode sizes and five circular HD-tDCS configurations [13]. They set the thickness of the sponge pad electrode and gel fixed at 2 mm and 0.50 mm, respectively. Kevin A. Caulfield et al. found a novel electrode position method that enhances on-target stimulation magnitude twofold in 3000 head models [14]. The effects of change in

thickness, shape and size together with sponge or gel are not reported till date in any literature.

In this study, we made head model with cognitively normal subject from the human connectome project (HCP). We used the available freeware software SimNIBS ver. 3.2.6 to generate 3D head mesh, image processing, and electric field analysis. We have analyzed and compared results obtained from multiple combination of size, shape and Gel/Sponge which can be used as an optimized standard for new electrode fabrication. The left dorsolateral prefrontal cortex (LDLPFC) and right dorsolateral prefrontal cortex (RDLPFC) were considered as an anodal and cathodal target, respectively positioning F3 (anode) and F4 (cathode) based on 10-20 EEG placement method.

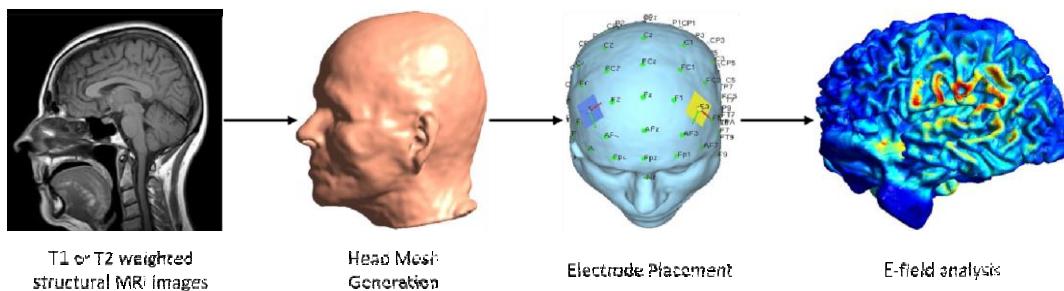


Fig. 1: Process flow for E-field analysis

Structural MRI images: We have used T2 weighted structural MRI image from Human Connectome Project (HCP) [15] of healthy adult. The image was pre-processed to remove facial features and to perform corrections. Imaging protocols used with Siemens 3T scanner: 3D-T2 space, TR/TE-3200/561 (ms), FOV-224×224 (mm), voxel size- 0.7 mm isotropic, BW-744, iPAT-2, acquisition time- 6:48 (min: sec).

The radiologist reviewed scans to ensure the absence of clinical abnormalities in any brain region. Imaging data were processed with software tools in FreeSurfer v5.3.0 and FSL v5.0. **Head mesh generation:** We have used the headreco pipeline (SimNIBS 3.2.6) to reconstruct tetrahedral mesh from structural MRIs. Headreco relies on statistical parametric mapping (Ver. SPM12) and computation anatomy toolbox for SPM (CAT12). Mesh created with headreco consists of 0.5 nodes per mm², called vertex density. Mesh resolution and manual editing can be made with other functionalities available in the tool. **Electrode placement and simulation criteria:** The placement of stimulating tDCS electrodes

II. METHODOLOGY: The main objective of this study is to analyze the variability of electric field strength and focality (Area affected by an electric field) on a given target. Electric field strength (norm-E) is calculated with the finite element method (FEM). The first step is to have a T2-weighted structural MRI image database for a subject in a standard format, i.e., the Neuroimaging Informatics Technology Initiative (NifTI) format. Finite element mesh (3D mesh) was generated with headreco processing pipeline part of volume conductor modeling enabled with SimNIBS. The three-dimensional mesh was ready for simulation setup (tDCS parameter selection). In the final step, electric field analysis was done for e-field strength (norm-E) and focality. Fig.1 shows a process flow for E-field analysis.

follows the standard 10-10/10-20 electrode placement method used for electroencephalograph (EEG). Anode and cathode were placed on F3 and F4 to target LDLPFC. To analyze the effects of electrode+gel and electrode+sponge thickness, we have divided the study into two major sections: 1) Rectangular electrode with varying shape and thickness of electrode+gel and electrode+sponge 2) Elliptical electrode with varying shape and thickness of electrode+gel and electrode+sponge. Tables 1 shows the stimulation parameters for both the groups. In both cases, the current intensity kept similar for simplification and comparison, i.e., 1mA. It is important to note that conductivities values change subject to subject as each individual has different conductivities for different tissues. We have assigned standard conductivities values to brain tissues: 1) White matter- 0.126 2)Grey matter- 0.275 3) Cerebrospinal fluid- 1654 4) Bone- 0.01 5) Scalp- 0.465 6) Eyeballs- 0.5 6) Compact bone- 0.008 7) Spongy bone- 0.0258) Blood- 0.6 8) Muscle- 0.16 9) 1.0 [16],[17],[18],[19].

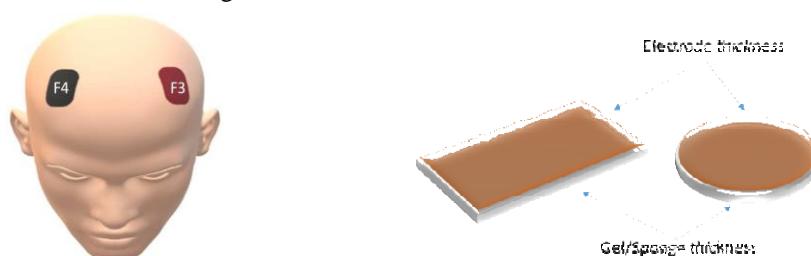


Fig. 2: a) 3D illustration of bipolar electrode placement on a human head (Red electrode F3-Anode and black electrode F4-Cathode) and b) Electrode shape with a conceptual illustration of electrode and gel/sponge thickness.

Simulation protocols for group-1 and group-2					
Group-1 (Shape)	Group-2 (Shape)	Size	Thickness		Current intensity
			Electrode + Gel (EG)	Electrode + Sponge (ES)	
Rectangular	Elliptical	1×1 cm ²	1 mm + 5 mm	1 mm + 5 mm	1 mA
		1×1 cm ²	2 mm + 5 mm	2 mm + 5 mm	1 mA
		1×1 cm ²	3 mm + 5 mm	3 mm + 5 mm	1 mA
		1×1 cm ²	4 mm + 5 mm	4 mm + 5 mm	1 mA
		1×1 cm ²	5 mm + 5 mm	5 mm + 5 mm	1 mA
		3×3 cm ²	1 mm + 5 mm	1 mm + 5 mm	1 mA
		3×3 cm ²	2 mm + 5 mm	2 mm + 5 mm	1 mA
		3×3 cm ²	3 mm + 5 mm	3 mm + 5 mm	1 mA
		3×3 cm ²	4 mm + 5 mm	4 mm + 5 mm	1 mA
		3×3 cm ²	5 mm + 5 mm	5 mm + 5 mm	1 mA
		5×5 cm ²	1 mm + 5 mm	1 mm + 5 mm	1 mA
		5×5 cm ²	2 mm + 5 mm	2 mm + 5 mm	1 mA
		5×5 cm ²	3 mm + 5 mm	3 mm + 5 mm	1 mA
		5×5 cm ²	4 mm + 5 mm	4 mm + 5 mm	1 mA
		5×5 cm ²	5 mm + 5 mm	5 mm + 5 mm	1 mA

Table 1: Simulation protocol for group-1 and group-2 study. For both groups, only the shape of the electrodes changes, and other parameters are considered the same.

E-field analysis: Transcranial direct current stimulation (tDCS) is a form of non-invasive brain stimulation that uses a low-intensity direct current to modulate brain activity. A critical aspect of tDCS is the analysis of the electric field distribution within the brain. This can be done using numerical modeling techniques, such as the finite element method, which can simulate the flow of electricity through the skull and tissue. The results of these simulations can be used to optimize the placement of the electrodes, the current intensity, and the duration of the stimulation to achieve the desired effects on brain activity. The E-field equation is typically represented as $E = -\text{grad } (V)$, where E is the electric field vector, grad is the gradient operator, and V is the electric potential. This equation can calculate the electric field at any point within the brain based on the distribution of electric potential (V) across the scalp and the skull. The electric potential is determined by the position and configuration of the electrodes, the current's intensity, and the tissue's conductivity. The E-field equation can be solved numerically using techniques such as the finite element method (FEM) or boundary element method (BEM) to create detailed models of the electric field distribution within the brain. These models can then be used to optimize the placement of the electrodes, the current intensity, and the duration of the stimulation to achieve specific effects on brain activity.

Generated electric fields have three characteristics 1) Field strength measured in V/m, 2) Direction of the field 3) Focality, i.e., grey matter volume having an electric field $\geq 50\%$ or 70% of the peak values [20]. Focality in transcranial direct current stimulation (tDCS) refers to the ability to target specific regions of the brain with the stimulation selectively. This is important because different areas in the brain are responsible for various functions, and selectively targeting a particular part can lead to specific changes in brain activity and behavior. There are several ways to achieve focality in tDCS, including using multiple electrodes, varying the position of the electrodes on the scalp, and using different types of current waveforms. One of the most common methods for achieving

focality in tDCS is the use of a "bipolar" configuration, where two electrodes are used, one of which is placed over the target brain region (the "anode") and the other is placed over some distance away (the "cathode"). This creates a "field of polarization" that selectively targets the anodal electrode region. Additionally, more recent methods such as High-Definition tDCS (HD-tDCS), which uses many smaller electrodes placed in a specific pattern, allow for more fine-grained stimulation of particular brain regions.

The direction of the field plays a significant role in deciding the possibility of depolarization or hyperpolarization of the neurons under therapy. It is also possible to use a "bilateral" tDCS, where the current flows in both directions simultaneously, with one anode and one cathode on each hemisphere. This is thought to have more global effects on the brain and is less specific than anodal or cathodal tDCS. It's important to note that the current direction in tDCS is not always a straightforward concept, as different factors such as the type of electrode montage, the current intensity and duration, the underlying neural networks, individual differences, and other parameters may influence the outcome. In this study, we have reported field strength and its focality on both simulation protocols with varying electrode size, shape, and conductive medium combinations. Specifically, LDLPFC is considered for the study assuming the most vulnerable target in cognitive decline and associated neurological diseases.

III. RESULTS

Electric field analysis and focality

We have simulated different shapes of electrodes, i.e., rectangular and circular, with fixed current intensity varying electrode thickness in a combination of gel and sponge. Electrode gel and electrode sponges are conductive materials in transcranial direct current stimulation (tDCS) experiments. The main difference between the two is their consistency and conductivity. Electrode gel is a liquid or semi-liquid substance that is applied to the surface of the electrode to improve conductivity and reduce skin impedance. It is typically more

conductive than electrode sponges, but it can be messier and may require more frequent reapplication during the experiment. Electrode sponges, on the other hand, are made of a moistened, compressed foam material that is placed on the electrode. They are less conductive than electrode gel but less messy and easier to use. They may also be more comfortable for the participant and require less frequent reapplication. Both electrode gel and electrode sponges can be used in tDCS experiments, but the choice will depend on the experiment's specific requirements and the researcher's preferences. We simulated both gel and sponge and got the results for both combinations of stimulation protocols, i.e., simulation protocol-1 and 2. Table.2 shows Simulation results of protocols for group-1. Fig.3 shows the electric field generated due to the variety of electrodes and gel, whereas Fig.4 shows the electric field generated due to the combination of electrode and sponge. It is observed that the thickness of the electrode, in the range of 1-5 mm, does not show a significant change in electric field strength with the exact size of the electrode. Also, gel shows good conductivity over sponge which typically ranges from 0.01-0.02 when applied on a rectangular electrode shape. The average electric field for $1 \times 1 \text{ cm}^2$ is 3.08e-01 V/m in 99.9 % of the total peak electric field within the grey matter of a subject. Focality shows that 8.92e+03 mm³ volume is exposed with an electric field for 75% of the peak electric field. The average electric field for $3 \times 3 \text{ cm}^2$ is 3.01e-01 V/m in 99.9 % of the total peak electric field within the grey matter of a subject. Focality shows that 7.65e+03 mm³ volume is exposed with an electric field for 75% of the peak electric field. The average electric field for $5 \times 5 \text{ cm}^2$ is 2.64e-01 V/m in 99.9 % of the total peak electric field within the grey matter of a subject. Focality shows that 6.41e+03 mm³ volume is exposed to an electric field for 75% of the peak electric field. Sponge shows a minimal change of conductivity decline compared with electrode gel. The average electric field for $1 \times 1 \text{ cm}^2$ is 3.09e-01 V/m in 99.9 % of the total peak electric field within the grey matter of a subject. Focality shows that 8.92e+03 mm³ volume is exposed with an electric field for 75% of the peak electric field. The average electric field for $3 \times 3 \text{ cm}^2$ is 3.02e-01 V/m in 99.9 % of the total peak electric field within the grey matter of a subject. Focality shows that 7.60e+03 mm³ volume is exposed with an electric field for 75% of the peak electric field. The average electric field for $5 \times 5 \text{ cm}^2$ is 2.62e-01 V/m in 99.9 % of the total peak electric field within the grey matter of a subject. Focality shows that 6.32e+03 mm³ volume is exposed with an electric field for 75% of the peak electric field.

exposed with an electric field for 75% of the peak electric field. The average electric field for $3 \times 3 \text{ cm}^2$ is 3.01e-01 V/m in 99.9 % of the total peak electric field within the grey matter of a subject. Focality shows that 7.65e+03 mm³ volume is exposed with an electric field for 75% of the peak electric field. The average electric field for $5 \times 5 \text{ cm}^2$ is 2.64e-01 V/m in 99.9 % of the total peak electric field within the grey matter of a subject. Focality shows that 6.41e+03 mm³ volume is exposed to an electric field for 75% of the peak electric field. Sponge shows a minimal change of conductivity decline compared with electrode gel. The average electric field for $1 \times 1 \text{ cm}^2$ is 3.09e-01 V/m in 99.9 % of the total peak electric field within the grey matter of a subject. Focality shows that 8.92e+03 mm³ volume is exposed with an electric field for 75% of the peak electric field. The average electric field for $3 \times 3 \text{ cm}^2$ is 3.02e-01 V/m in 99.9 % of the total peak electric field within the grey matter of a subject. Focality shows that 7.60e+03 mm³ volume is exposed with an electric field for 75% of the peak electric field. The average electric field for $5 \times 5 \text{ cm}^2$ is 2.62e-01 V/m in 99.9 % of the total peak electric field within the grey matter of a subject. Focality shows that 6.32e+03 mm³ volume is exposed with an electric field for 75% of the peak electric field.

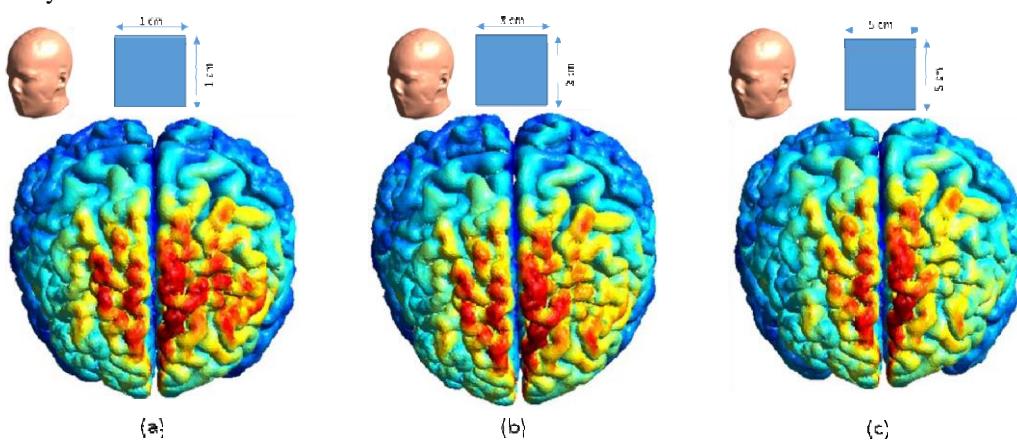


Fig.3 Variability of E-field strength upon size of electrodes (E+G) (Stimulation protocols: Current intensity- 1 mA, shape of electrode- Rectangular)

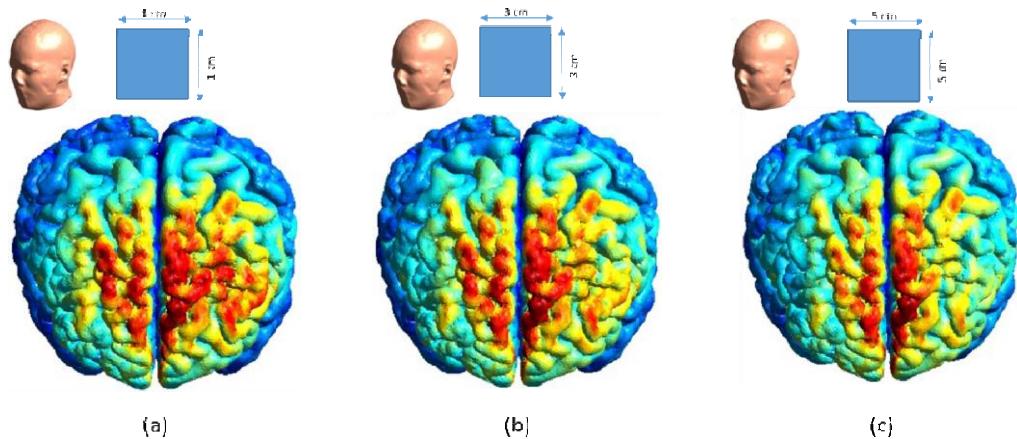


Fig.4 Variability of E-field strength upon size of electrodes (E+S) (Stimulation protocols: Current intensity- 1 mA, shape of electrode- Rectangular)

Simulation results of protocols for group-1 (Shape –Rectangular, Current intensity-1 mA)						
Size (cm ²)	Thickness (mm) (E+G)	E-field strength (99.9%) (V/m)	Focality (75%) (mm ³)	Thickness (mm) (E+S)	E-field strength (99.9%) (V/m)	Focality (75%) (mm ³)
1×1	1+5	3.08e-01	8.92e+03	1+5	3.09e-01	8.92e+03
1×1	2+5	3.08e-01	8.92e+03	2+5	3.09e-01	8.92e+03
1×1	3+5	3.08e-01	8.92e+03	3+5	3.09e-01	8.92e+03
1×1	4+5	3.08e-01	8.92e+03	4+5	3.09e-01	8.92e+03
1×1	5+5	3.08e-01	8.92e+03	5+5	3.09e-01	8.92e+03
3×3	1+5	3.01e-01	7.65e+03	1+5	3.02e-01	7.60e+03
3×3	2+5	3.01e-01	7.65e+03	2+5	3.02e-01	7.60e+03
3×3	3+5	3.01e-01	7.65e+03	3+5	3.02e-01	7.60e+03
3×3	4+5	3.01e-01	7.65e+03	4+5	3.02e-01	7.60e+03
3×3	5+5	3.01e-01	7.65e+03	5+5	3.02e-01	7.60e+03
5×5	1+5	2.64e-01	6.40e+03	1+5	2.62e-01	6.32e+03
5×5	2+5	2.64e-01	6.40e+03	2+5	2.62e-01	6.32e+03
5×5	3+5	2.64e-01	6.40e+03	3+5	2.62e-01	6.32e+03
5×5	4+5	2.64e-01	6.41e+03	4+5	2.62e-01	6.32e+03
5×5	5+5	2.64e-01	6.41e+03	5+5	2.62e-01	6.32e+03

Table.2: Simulation results of protocols for group-1 (Shape –Rectangular, Current intensity-1 mA) (E+G- Electrode+gel, E+S- Electrode+ sponge)

Table.3 shows Simulation results of protocols for group-2. Fig.5 shows the electric field generated due to the combination of electrode and gel. In contrast, Fig. 6 shows the electric field generated due to the combination of electrode and sponge in a circular electrode type. The gel shows good conductivity over a sponge which typically ranges from 0.09-0.24 when applied to the circular shape of the electrode. The average electric field for $1 \times 1 \text{ cm}^2$ is 3.06e-01 V/m in 99.9 % of the total peak electric field within the grey matter of a subject. Focality shows that 8.97e+03 mm³ volume is exposed with an electric field for 75% of the peak electric field. The average electric field for $3 \times 3 \text{ cm}^2$ is 2.97e-01 V/m in 99.9 % of the total peak electric field within the grey matter of a subject. Focality shows that 7.99e+03 mm³ volume is exposed with an electric field for 75% of the peak electric field. The average electric field for $5 \times 5 \text{ cm}^2$ is 2.76e-01 V/m in 99.9 % of the total peak electric field within the grey matter of a subject. Focality

shows that 7.04e+03 mm³ volume is exposed to an electric field for 75% of the peak electric field. Sponge shows a minimal change of conductivity decline compared with electrode gel. The average electric field for $1 \times 1 \text{ cm}^2$ is 3.05e-01 V/m in 99.9 % of the total peak electric field within the grey matter of a subject. Focality shows that 8.95e+03 mm³ volume is exposed with an electric field for 75% of the peak electric field. The average electric field for $3 \times 3 \text{ cm}^2$ is 2.96e-01 V/m in 99.9 % of the total peak electric field within the grey matter of a subject. Focality shows that 7.94e+03 mm³ volume is exposed with an electric field for 75% of the peak electric field. The average electric field for $5 \times 5 \text{ cm}^2$ is 2.71e-01 V/m in 99.9 % of the total peak electric field within the grey matter of a subject. Focality shows that 7.00e+03 mm³ volume is exposed with an electric field for 75% of the peak electric field.

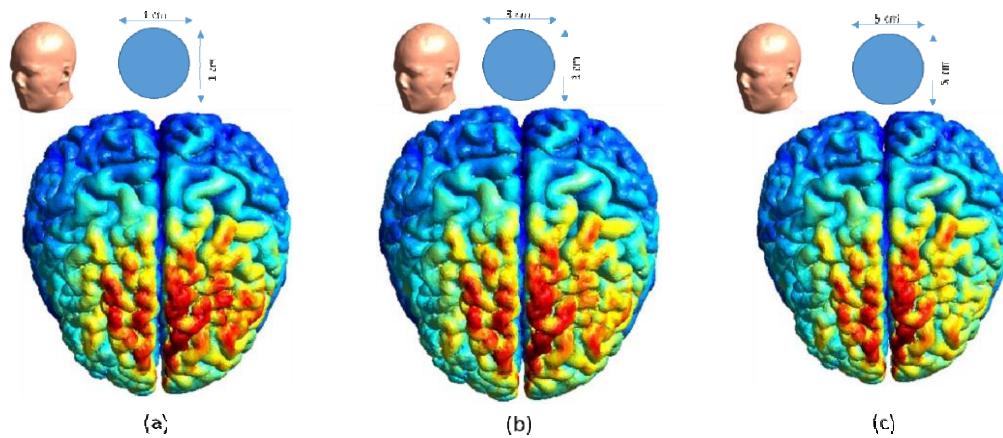


Fig.5 Variability of E-field strength upon the size of electrodes (E+G) (Stimulation protocols: Current intensity- 1 mA, shape of electrode- Circular

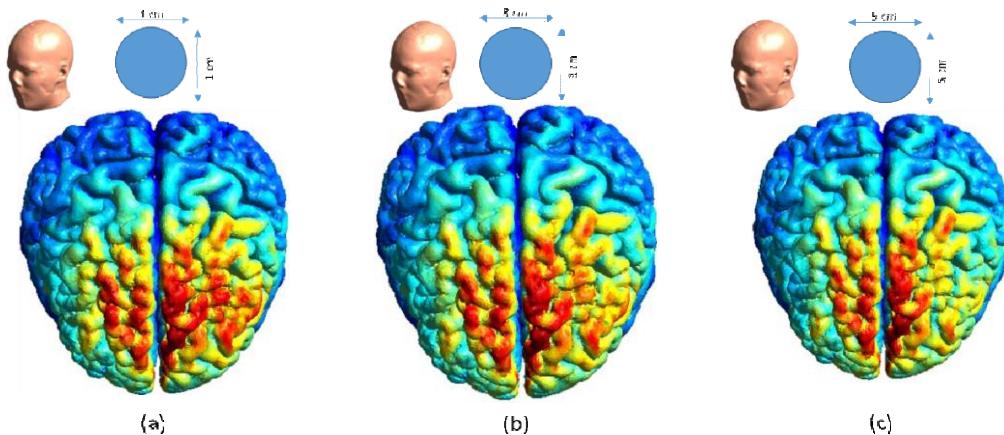


Fig.6 Variability of E-field strength upon the size of electrodes (E+S) (Stimulation protocols: Current intensity- 1 mA, shape of electrode- Circular)

Simulation results of protocols for group-2 (Shape –Elliptical, Current intensity-1 mA)						
Size (cm ²)	Thickness (mm)	E-field strength (99.9%) (V/m)	Focality (75%) (mm ³)	Thickness (mm)	E-field strength (99.9%) (V/m)	Focality (75%) (mm ³)
	(E+G)			(E+S)		
1x1	1+5	3.06e-01	8.97e+03	1+5	3.05e-01	8.95e+03
1x1	2+5	3.06e-01	8.97e+03	2+5	3.05e-01	8.95e+03
1x1	3+5	3.06e-01	8.97e+03	3+5	3.05e-01	8.95e+03
1x1	4+5	3.06e-01	8.97e+03	4+5	3.05e-01	8.95e+03
1x1	5+5	3.06e-01	8.97e+03	5+5	3.05e-01	8.95e+03
3x3	1+5	2.97e-01	7.99e+03	1+5	2.96e-01	7.94e+03
3x3	2+5	2.97e-01	7.99e+03	2+5	2.96e-01	7.94e+03
3x3	3+5	2.97e-01	7.99e+03	3+5	2.96e-01	7.94e+03
3x3	4+5	2.97e-01	7.99e+03	4+5	2.96e-01	7.94e+03
3x3	5+5	2.97e-01	7.99e+03	5+5	2.96e-01	7.94e+03
5x5	1+5	2.76e-01	7.04e+03	1+5	2.71e-01	7.00e+03
5x5	2+5	2.76e-01	7.04e+03	2+5	2.71e-01	7.00e+03
5x5	3+5	2.76e-01	7.04e+03	3+5	2.71e-01	7.00e+03
5x5	4+5	2.76e-01	7.04e+03	4+5	2.71e-01	7.00e+03
5x5	5+5	2.76e-01	7.04e+03	5+5	2.71e-01	7.00e+03

Table.3: Simulation results of protocols for group-2 (Shape –Elliptical, Current intensity-1 mA) (E+G- Electrode+gel, E+S- Electrode+ sponge)

IV. DISCUSSION: Results of both simulations show similar electric fields and focality when compared to the electrode+gel and electrode+sponge combination. The difference between the parameters is negligible, concluding that it is more convenient to use electrode+sponge for stimulation as it is more comfortable for users or physicians for the application and firm adhesion to the subject's skin. Also, the thickness of the electrode in the range of 1-5 mm does not show any significant change in electric field strength when used with the same electrode size, i.e., for 1x1 cm², 3x3 cm², and 5x5 cm². But change in the electrode's size and shape results in measurable change for electric field and focality both. The average electric field strength of 1x1 cm² is 3.08E-01 V/m and 8.92E+03 mm³ focality in the targeted grey matter of the subject's brain. It reduces to 2.27 %, i.e., 3.01E-01 V/m and

7.65E+03 mm³ when a 3x3 cm² rectangular electrode is combined with electrode+gel. Further, it reduces to 14.28 %, i.e., 2.64E-01 V/m (electric field strength) and 6.41E+03 (Focality) in 5x5 cm² compared to 1x1 cm². In case of electrode+sponge, the electric field strength in 1x1 cm², 3x3 cm² and 5x5 cm² are 3.09E-01 V/m, 3.02E-01 V/m and 2.62E-01 V/m respectively and focality 8.92E+03 mm³, 7.60E+03 mm³ and 6.32E+03 mm³ respectively. Fig.7 and 8 show the Variability of E-field strength (a) and focality (b) upon the size of electrodes (E+G) (Stimulation protocols: Current intensity- 1 mA, shape of electrode- Rectangular) and Variability of E-field strength (a) and focality (b) upon the size of electrodes (E+S) (Stimulation protocols: Current intensity- 1 mA, shape of electrode- Rectangular).

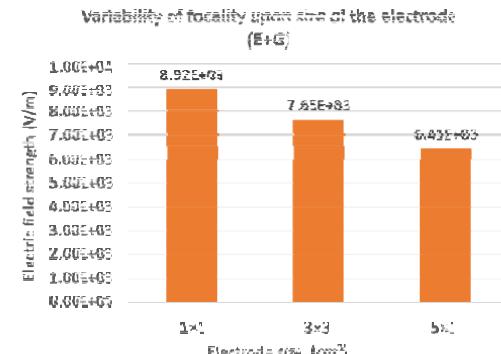
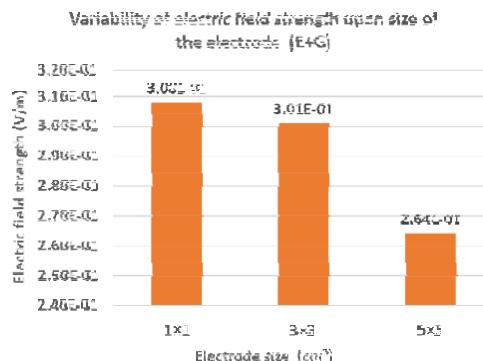


Fig.7 Variability of E-field strength (a) and focality (b) upon the size of electrodes (E+G) (Stimulation protocols: Current intensity- 1 mA, shape of electrode- Rectangular)

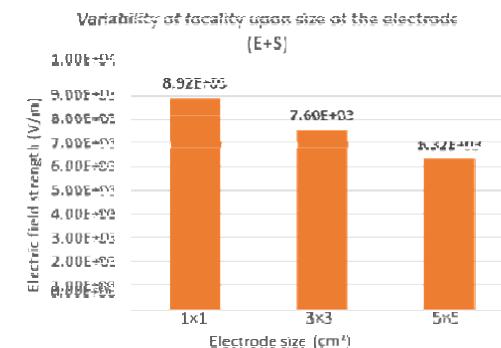
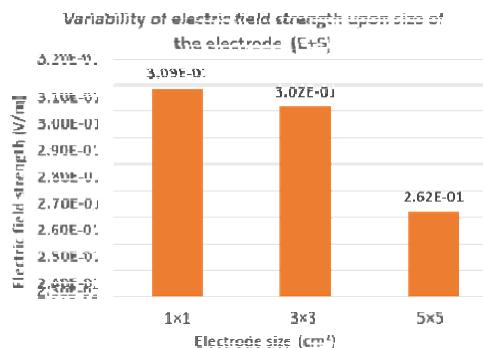


Fig.8 Variability of E-field strength (a) and focality (b) upon the size of electrodes (E+S) (Stimulation protocols: Current intensity- 1 mA, shape of electrode- Rectangular)

For circular electrodes, the average electric field strength of $1 \times 1 \text{ cm}^2$ is $3.06 \times 10^{-1} \text{ V/m}$ and $8.97 \times 10^3 \text{ mm}^3$ focality in the targeted grey matter of the subject's brain. It reduces to 2.27 %, i.e., $2.97 \times 10^{-1} \text{ V/m}$ and $7.99 \times 10^3 \text{ mm}^3$, when a $3 \times 3 \text{ cm}^2$ circular electrode is combined with electrode+gel. Further, it reduces to 14.28 %, i.e., $2.76 \times 10^{-1} \text{ V/m}$ (electric field strength) and 7.04×10^3 (Focality) in $5 \times 5 \text{ cm}^2$ compared to $1 \times 1 \text{ cm}^2$. In case of electrode+sponge, the electric field strength in $1 \times 1 \text{ cm}^2$, $3 \times 3 \text{ cm}^2$ and $5 \times 5 \text{ cm}^2$ are $3.05 \times 10^{-1} \text{ V/m}$, $2.96 \times 10^{-1} \text{ V/m}$

and $2.71 \times 10^{-1} \text{ V/m}$ respectively and focality $8.95 \times 10^3 \text{ mm}^3$, $7.94 \times 10^3 \text{ mm}^3$ and $7.00 \times 10^3 \text{ mm}^3$ respectively. Fig.9 and 10 show the Variability of E-field strength (a) and focality (b) upon the size of electrodes (E+G) (Stimulation protocols: Current intensity- 1 mA, shape of electrode- Circular) and Variability of E-field strength (a) and focality (b) upon the size of electrodes (E+S) (Stimulation protocols: Current intensity- 1 mA, shape of electrode- Circular).

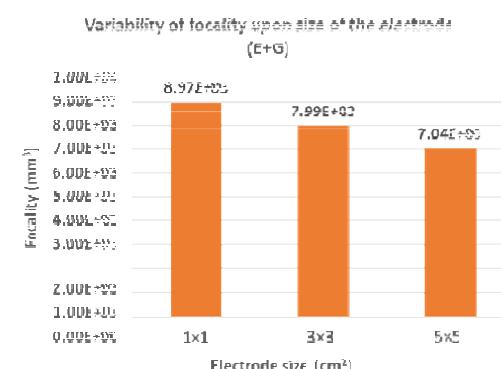
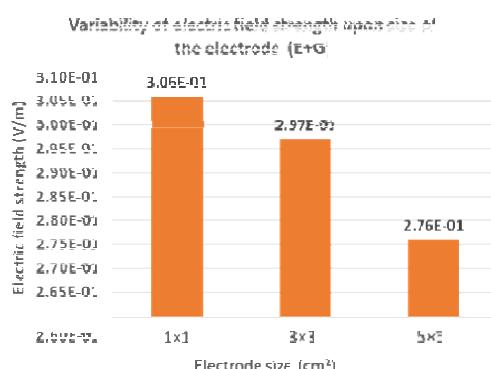


Fig.9 Variability of E-field strength (a) and focality (b) upon size of electrodes (E+G) (Stimulation protocols: Current intensity- 1 mA, shape of electrode- Circular)

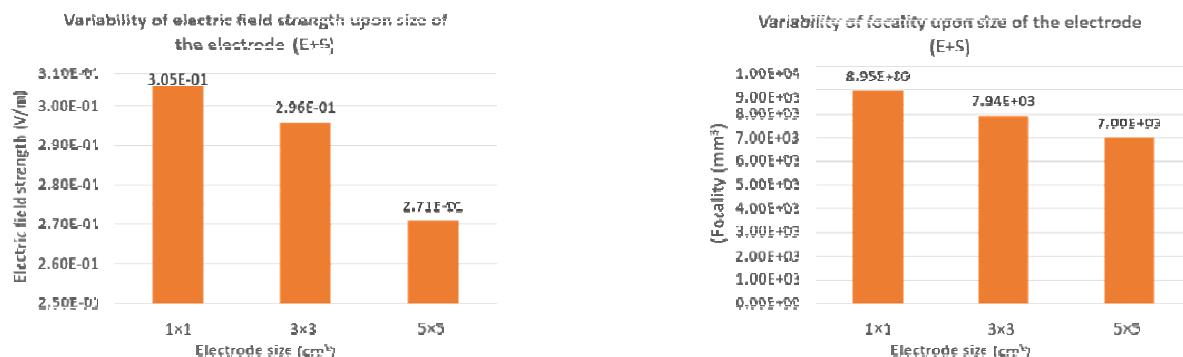


Fig.10 Variability of E-field strength (a) and focality (b) upon the size of electrodes (E+S) (Stimulation protocols: Current intensity- 1 mA, shape of electrode- Circular)

V. CONCLUSION:

The size and shape of the electrodes used in transcranial direct current stimulation (tDCS) can affect the distribution of the electric current flowing through the brain. A larger electrode size can result in a more diffuse current distribution, potentially affecting a wider brain area. On the other hand, a smaller electrode size can result in a more focal distribution of current, potentially affecting a specific brain region. A circular electrode will have a more symmetrical current distribution, while a rectangular electrode will have a more asymmetrical distribution. The electrode size and shape choice can depend on the specific target area of the brain and the desired outcome of the tDCS session. A more focal stimulation will target a smaller area, whereas a less focal stimulation will have a more widespread effect. The focality of tDCS can be adjusted by varying the size and shape of the electrodes and the current density. We have considered the uniform thickness of the sponge and conductive gel. In future studies, one can vary the thickness of the conductive medium and check the effect on electric field strength and focality.

REFERENCES

- [1] K. Pieber, M. Herceg, and T. Patemostro-Sluga, "Electrotherapy for the treatment of painful diabetic peripheral neuropathy: A review," *J. Rehabil. Med.*, vol. 42, no. 4, pp. 289–295, Apr. 2010, doi: 10.2340/16501977-0554.
- [2] S. V. Eranti and D. M. McLoughlin, "Electro-convulsive therapy – state of the art," *Br. J. Psychiatry*, vol. 182, no. 1, pp. 8–9, Jan. 2003, doi: 10.1192/BJP.182.1.8.
- [3] D. N. Rushton, "Functional Electrical Stimulation and rehabilitation—an hypothesis," *Med. Eng. Phys.*, vol. 25, no. 1, pp. 75–78, Jan. 2003, doi: 10.1016/S1350-4533(02)00040-1.
- [4] A. Antal and W. Paulus, "Transcranial alternating current stimulation (tACS)," *Front. Hum. Neurosci.*, vol. 7, no. JUN, p. 317, Jun. 2013, doi: 10.3389/FNHUM.2013.00317/BIBTEX.
- [5] M. A. Nitsche *et al.*, "Transcranial direct current stimulation: State of the art 2008," *Brain Stimul.*, vol. 1, no. 3, pp. 206–223, Jul. 2008, doi: 10.1016/J.BRS.2008.06.004.
- [6] J. P. Lefaucheur *et al.*, "Evidence-based guidelines on the therapeutic use of repetitive transcranial magnetic stimulation (rTMS): An update (2014–2018)," *Clin. Neurophysiol.*, vol. 131, no. 2, pp. 474–528, Feb. 2020, doi: 10.1016/J.CLINPH.2019.11.002.
- [7] R. H. Howland, "Vagus Nerve Stimulation," *Curr. Behav. Neurosci. reports*, vol. 1, no. 2, p. 64, Jun. 2014, doi: 10.1007/S40473-014-0010-5.
- [8] M. M. Bottros and P. J. Christo, "Current perspectives on intrathecal drug delivery," *J. Pain Res.*, vol. 7, pp. 615–626, Nov. 2014, doi: 10.2147/JPR.S37591.
- [9] A. Antal, K. Boros, C. Poreisz, L. Chaib, D. Terney, and W. Paulus, "Comparatively weak after-effects of transcranial alternating current stimulation (tACS) on cortical excitability in humans," *Brain Stimul.*, vol. 1, no. 2, pp. 97–105, 2008, doi: 10.1016/J.BRS.2007.10.001.
- [10] A. J. Woods *et al.*, "A technical guide to tDCS, and related non-invasive brain stimulation tools," *Clin. Neurophysiol.*, vol. 127, no. 2, pp. 1031–1048, Feb. 2016, doi: 10.1016/J.CLINPH.2015.11.012.
- [11] T. L. Rich and B. T. Gillick, "Electrode Placement in Transcranial Direct Current Stimulation—How Reliable Is the Determination of C3/C4?," *Brain Sci.*, vol. 9, no. 3, Mar. 2019, doi: 10.3390/BRAINSCI9030069.
- [12] S. Das, P. Holland, M. A. Frens, and O. Donchin, "Impact of transcranial direct current stimulation (tDCS) on neuronal functions," *Front. Neurosci.*, vol. 10, no. NOV, p. 550, Nov. 2016, doi: 10.3389/FNINS.2016.00550/BIBTEX.
- [13] G. Sharma and S. Roy Chowdhury, "Enhancement in Focality of Non-Invasive Brain Stimulation through High Definition (HD) Anodal Transcranial Direct Current Stimulation (tDCS) Techniques," *2019 IEEE Conf. Comput. Intell. Bioinforma. Comput. Biol. CIBCB 2019*, 2019, doi: 10.1109/CIBCB.2019.8791447.
- [14] K. A. Caulfield and M. S. George, "Optimized APPS-tDCS electrode position, size, and distance doubles the on-target stimulation magnitude in 3000 electric field models," *Sci. Reports* 2022 12, vol. 12, no. 1, pp. 1–15, Nov. 2022, doi: 10.1038/s41598-022-24618-3.
- [15] "Human Connectome Project | Mapping the human brain connectivity." <http://www.humanconnectomeproject.org/> (accessed Jan. 10, 2023).
- [16] T. A. Wagner, M. Zahn, A. J. Grodzinsky, and A. Pascual-Leone, "Three-dimensional head model simulation of transcranial magnetic stimulation," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 9, pp. 1586–1598, Sep. 2004, doi: 10.1109/TBME.2004.827925.
- [17] A. Opitz, W. Paulus, S. Will, A. Antunes, and A. Thielscher, "Determinants of the electric field during transcranial direct current stimulation," *Neuroimage*, vol. 109, pp. 140–150, Apr. 2015, doi: 10.1016/j.neuroimage.2015.01.033.
- [18] A. Roda, G. Serra-Mir, L. Montoliu-Gaya, L. Tiessler, and S. Villegas, "Amyloid-beta peptide and tau protein crosstalk in Alzheimer's disease," *Neural Regen. Res.*, vol. 17, no. 8, p. 1666, Aug. 2022, doi: 10.4103/1673-5374.332127.
- [19] G. B. Saturnino, A. Antunes, and A. Thielscher, "On the importance of electrode parameters for shaping electric field patterns generated by tDCS," *Neuroimage*, vol. 120, pp. 25–35, Oct. 2015, doi: 10.1016/J.NEUROIMAGE.2015.06.067.
- [20] G. B. Saturnino, O. Puonti, J. D. Nielsen, D. Antonenko, K. H. Madsen, and A. Thielscher, "SimNIBS 2.1: A Comprehensive Pipeline for Individualized Electric Field Modelling for Transcranial Brain Stimulation," *bioRxiv*, pp. 1–13, Dec. 2018, doi: 10.1101/500314.

Factors Affecting the Efficacy of Retinal Implants

Abdulrahman Alqahtani

Department of Medical Equipment Technology

Majmaah University

Al-Majmaah, Saudi Arabia

a.alqahtani@mu.edu.sa

Abstract—In this study, a realistic rabbit OFF retinal ganglion cells (RGCs) model was utilized with various return electrode configurations to determine the optimal activation approach for retinal prosthesis. Monopolar and hexapolar return electrode configurations placed epiretinally were utilized. Four types of waveforms were used to deliver the current through the stimulating electrode: biphasic cathodic first, biphasic anodic first, monophasic cathodic, and monophasic anodic pulses. The RGC response was computed at three locations. When the RGC was placed directly below the stimulating electrode, no discernible difference in the required current thresholds for activating RGCs was found between monopolar and hexapolar return electrodes configurations irrespective of the used current stimulus waveform. Moreover, when the RGC was situated remotely, the variation in current thresholds between monopolar and hexapolar was evident. In addition, monophasic anodic was more efficient for monopolar configuration. Additionally, monophasic cathodic stimulation stimulates the RGCs closest to the stimulating electrode.

Index Terms—RGC, Retinal implants, Hexapolar, Monopolar, Electrical stimulation

I. INTRODUCTION

Retinitis Pigmentosa (RP) is a hereditary condition that affects roughly 1 in 4000 people individuals globally [1]. The disease is characterized by the gradual loss of photoreceptors within the retina and is accompanied by severe impaired vision. Laser, gene, and pharmacological therapies are being investigated as potential treatments for various degenerative vision disorders [2]. Retinal prostheses devices are proposed to restore vision in patients with vision impairment by stimulating surviving retinal neurons electrically. Four types of retinal prostheses were proposed: epiretinal, subretinal, suprachoroidal and transscleral. Regardless of the type employed, the objective is to electrically excite intact retinal ganglion cells in order to generate percepts that carry visual scene information. Clinical trials have demonstrated that blind patients can perceive and identify simple shape objects after retinal prosthesis implantation and rehabilitation training [1]. However, the resolution of these devices remain needing more improvement.

The performance of retinal implants is associated with operational and safety issues. For example, the injected charge should be within the safe charge limit for electrodes [1]. Neurological damage and electrode corrosion can be avoided by keeping the charge density below $300 \mu C cm^{-2}$, which is the safe limit for platinum electrodes [13]. Moreover, the length, type and symmetry of pulse duration have been found

to have great impact on neural electrical stimulation in different studies [4]–[6]. The activation of passage axons in retinal implants, which has a substantial impact on the resolution of the perceived image, has been successfully avoided using, for instance, shorter pulses (0.1 ms) or bigger pulses (25 ms) [5]. Additionally, the asymmetry of pulses has been shown to minimize the threshold for activating retinal neurons. Therefore, investigating effective stimulation paradigms will be useful for enhancing the performance of retinal prostheses.

The utilized return electrodes configuration during electrical stimulation has an impact on the threshold and spatial activation, according to studies on cochlear and visual implants [1], [3]. Various return electrodes are proposed for on

cochlear and visual prosthesis such as monopolar, bipolar, tripolar and hexapolar configurations. These studies showed generally monopolar configuration activated neurons with low charge injection where other configurations provide a confined activation [1], [3].

Investigating the effects of return electrode arrangements on retinal electrical stimulation is the goal of this study. This investigation was conducted using a model of morphologically realistic RGCs with epiretinal stimulation. The effects of two return electrodes with four types of pulse duration were simulated and compared.

II. METHODS

The morphologically accurate rabbit OFF RGC model created by Guo et al. was successfully recreated using COMSOL Multiphysics finite element simulation software [7]. A semi-ellipsoid region that resembled the extracellular medium was used to contain the morphologically realistic OFF RGC cell. The soma, axon, hillock, axon initial segment (AIS), and dendrites make up this particular type of cell (Fig. 1a). Two of the current return electrodes used in retinal implants were investigated: hexapolar and monopolar. First, the hexapolar configuration was set up in a hexagonal layout (Fig. 1b). It consists of six return electrodes that are circular, 200 μm in diameter, and spaced 400 μm apart from one another. According to the illustration in (Fig. 1b), the central electrode of the hexagon acted as the stimulating electrode, while the six nearby local electrodes served as returns connected to the ground. Second, the return electrode in the monopolar setup is situated far away, at the ellipsoid's bottom boundaries (Fig. 1c).

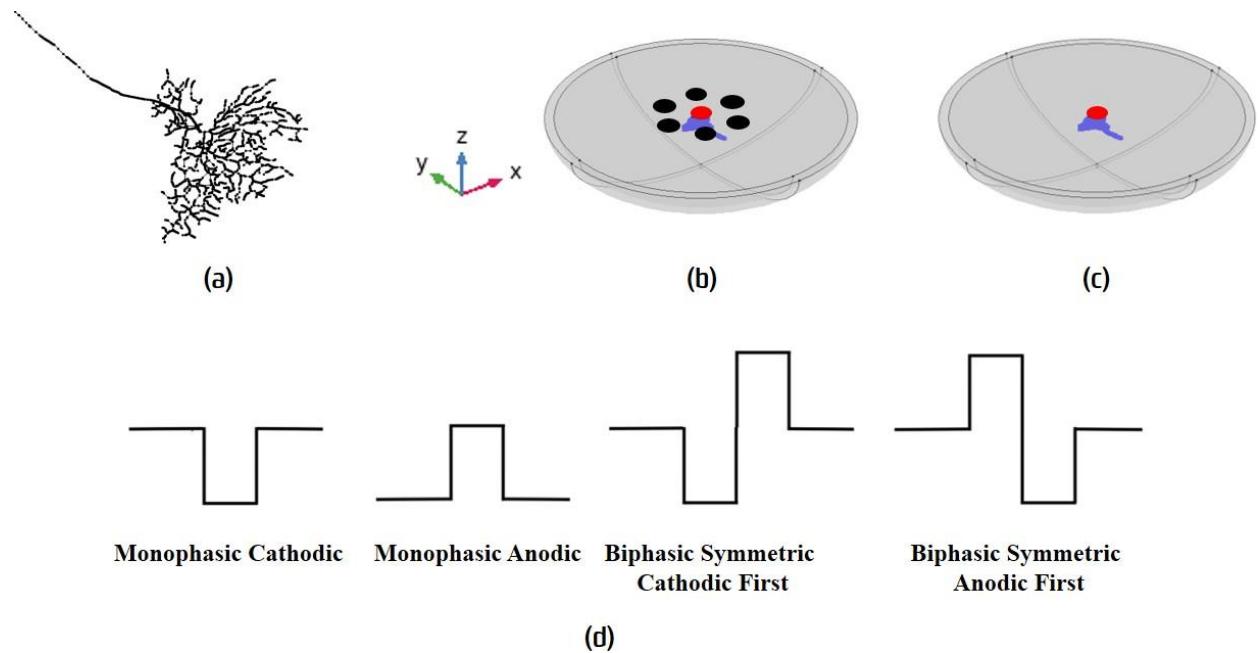


Fig. 1. The epiretinal stimulation models settings (a) the morphologically realistic rabbit OFF RGC (b) hexapolar (c) monopolar return electrodes configurations. The return electrodes are shown by black circles, whereas the stimulating electrodes are depicted by red circles. (d) the four stimulus waveforms utilized in this investigation.

Four kinds of waveforms with periods of 0.5 ms and no interphase intervals were used: biphasic cathodic first, biphasic anodic first, monophasic cathodic, and monophasic anodic pulses (Fig. 1d). To map the required thresholds for activating RGCs, the RGC was positioned at 120 μm below the stimulating electrode. The other two sites of measurements were 400 μm on the left and the right of the plane of the stimulating electrode at the same vertical distance (120 μm) .

A. Mathematical Formulations

The extracellular voltage distribution was governed by Poisson's equation in our OFF RGC discrete model (V_e):

$$\nabla \cdot (-\sigma_b \nabla V_e) = I \quad (1)$$

where σ_b is the vitreous layer's electric conductivity of 1.28 (S/m) [8] and I is the volumetric current density (A/m^3) due to cell membrane current flow.

The varying-radius cable equation [9], which takes the variation in the radius of RGC compartments into account, was used to calculate the RGC activation as follows:

$$\frac{\partial}{\partial x} \left[\frac{r^2}{2\rho_i} \frac{\partial V_m}{\partial x} \right] = r(C_m \frac{dV_m}{dt} + i_{ion}) \quad (2)$$

where the membrane potential is represented by V_m , x is the axial cable distance, r describes the RGC compartment's radius (μm), ρ_i the intracellular resistivity and was set to 110 $\Omega \cdot \text{cm}$, and membrane capacitance (C_m) expressed as

a function of membrane area ($\mu\text{F}/\text{cm}^2$). The total ionic current through the RGC membrane, expressed as a function of membrane area, is denoted by the symbol i_{ion} .

All gating variables and the detailed ionic current formulations, conductance and kinetic parameter values of each rate and each RGC compartment can be found in more detail in [7].

III. RESULTS AND DISCUSSION

A. The activation of RGCs below the stimulating Electrode

Using either a monopolar or a hexapolar return electrode design, we assessed the necessary current to activate RGCs during epiretinal electrical stimulation. Our simulations (Fig. 3) showed that, when the RGC was positioned exactly beneath the stimulating electrode for any of the four stimulus waveforms, the necessary current thresholds for activating RGCs do not statistically differ between the monopolar and hexapolar return electrodes arrangements. These results seem to contradict the majority of *in vitro* and *in vivo* studies [1], [3], [10], which revealed that hexapolar configuration requires a current threshold that is at least two times greater than monopolar. On the other hand, our simulations are in line with a prior study that used *in vitro* testing and computer modeling to demonstrate the necessary current thresholds for activating RGCs during suprachoroidal electrical stimulation do not significantly differ between monopolar and multipolar configurations [11]. Furthermore, as shown in our simulations,

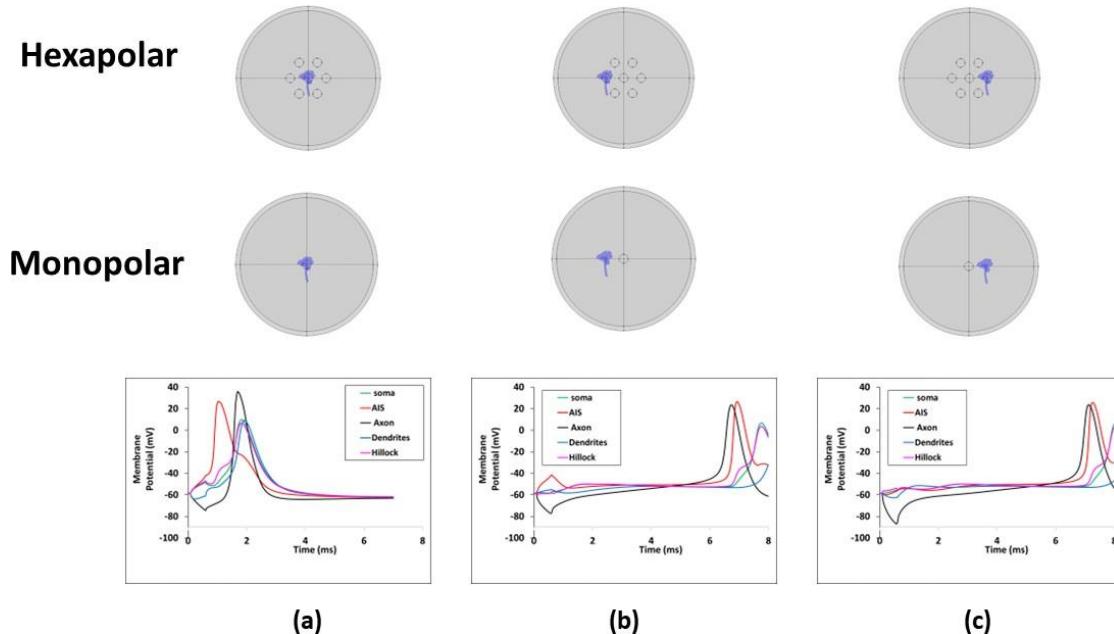


Fig. 2. Action potentials for various RGC compartments during using monopolar and hexapolar configurations with monophasic cathodic at (a) below the stimulating electrode and at a $400\ \mu\text{m}$ distance on (b) left and (c) right of the stimulating electrode.

recent computational studies have shown that when the electrode diameter is larger than the distance between electrodes and targeted tissue, the variation between monopolar and hexapolar in current thresholds for activating RGCs is minimal [12]–[14].

B. The activation of RGCs far away from the stimulating Electrode

The variation in current thresholds between monopolar and hexapolar return electrode configurations has been clearly seen when the RGC was located far away from the stimulating electrode. For example when the RGC was positioned on the left side by $400\ \mu\text{m}$ distance (Fig. 2b), the hexapolar configuration used higher current thresholds than monopolar with all the used pulse durations with a range of 13% to 80% (Fig. 3). Moreover, the range of variation in the current thresholds between the two return electrodes was obvious when the RGC is located on the right side by the same distance, which was between 4% to 95% (Fig. 3). That means there is a significant variation between the two kinds of return electrodes for only those RGCs located far away from the active electrode even if the distance between the active electrode and RGC was less than the electrode's diameter. These significant findings explain and combine many *in vivo* and *in vitro* experiments [1], [3], [10], which pointed to monopolar configuration usually needing lower current thresholds compared to hexapolar, and there is no discernible difference between hexapolar and monopolar return electrodes when the distance between the active electrode and the targeted

RGCs is less than the electrode diameter, according to several *in vitro* [11] and modeling studies [12]–[14].

C. The influence of pulse duration on RGC activation

In retinal prostheses, the most common type of pulse is a biphasic charge-balanced pulse with a cathodic first [16]–[18]. It is known that cathodic phase depolarizes neurons close the activating electrode and generates an action potential, whereas the anodic phase hyperpolarizes proximal neurons [16]. Consequently, a cathodic pulse needs a lower charge to generate an action potential when the activating electrode is placed near neural cells [17]. However, several *in vitro* investigations utilizing subretinal prostheses found that anodal stimulation needed lower charge than cathodal stimulation to generate an action potential [16], [18]. Moreover, neuronal responses rely on the placement and proximity of the activating electrode to retinal neurons [13], [18]. Our simulations revealed valuable insights on the impact of pulse duration on the epiretinal electrical stimulation of RGCs. Firstly, monophasic anodic pulse duration was the lowest current threshold for monopolar configuration for the three locations of RGC and for the center RGC location during hexapolar configuration, which is an agreement with recent studies [4], [6], [16], [18]. Secondly, biphasic cathodic first pulse was the more efficient charge injection for hexapolar configuration with exception for those RGCs located directly below the stimulating electrode which is consistent with experimental study [15]. Thirdly, in contrast to all other pulse durations that target the RGC axon, monophasic cathodic pulse duration was the effective pulse for activating

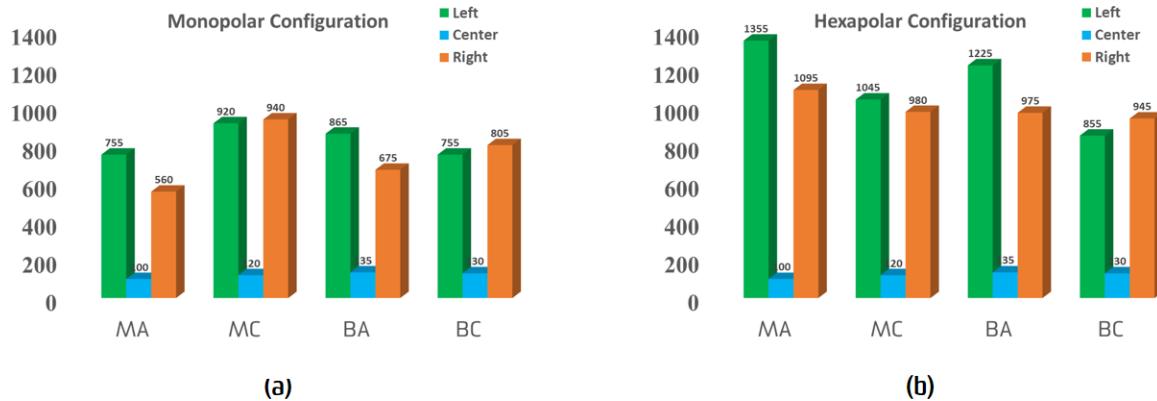


Fig. 3. Current threshold for activating RGCs for either using (a) monopolar or (b) hexapolar configurations with using the four pulse durations: biphasic cathodic first (BC), biphasic anodic first (BA), monophasic cathodic (MC), and monophasic anodic (MA) pulses.

those RGCs located below the stimulating electrode (Fig. 2a), which targets the soma or axon initial segment (AIS) of the RGC not its distal axon and that in agreement with [6]. When the RGC located away from the stimulating electrode, the axon was activated before the soma or AIS regardless the used pulse duration (Fig. 2b & c). This finding will be helpful for preferable retinal activation, which is considered as a one of the challenges associated with retinal implants.

IV. CONCLUSION

The impact of return electrodes configuration in retinal electrical stimulation was investigated using four types of current pulses. Both kinds of return electrodes configurations: hexapolar and monopolar need similar current thresholds to activate RGCs when the distance between the stimulating electrode and RGC is less than the electrode diameter (RGC is directly below the stimulating electrode). Monopolar return configuration exhibited lower current threshold when RGC was situated remotely either on left or right of the stimulating electrode. Monophasic anodic was the more efficient current waveform for monopolar (all positions) and hexapolar (center position) configurations while biphasic cathodic was the most efficient for hexapolar configuration when RGC was positioned far away either left or right of the stimulating electrode.

REFERENCES

- [1] M. N. Shivedasani et al., "Evaluation of stimulus parameters and electrode geometry for an effective suprachoroidal retinal prosthesis," *Journal of neural engineering*, vol. 7, no. 3, p. 036008, 2010.
- [2] R. D. Jager, W. F. Mieler, and J. W. Miller, "Age-related macular degeneration," *N Engl J Med*, vol. 358, pp. 2606-17, Jun 12 2008.
- [3] Matteucci PB, Chen SC, Tsai D, Dodds CW, Dokos S, Morley JW, Lovell NH, Suaning GJ. Current steering in retinal stimulation via a quasimonopolar stimulation paradigm. *Investigative ophthalmology & visual science*. 2013 Jun 1;54(6):4307-20.
- [4] Tong W, Meffin H, Garrett DJ, Ibbotson MR. Stimulation strategies for improving the resolution of retinal prostheses. *Frontiers in neuroscience*. 2020 Mar 26;14:262.
- [5] Weitz AC, Nanduri D, Behrend MR, Gonzalez-Calle A, Greenberg RJ, Humayun MS, Chow RH, Weiland JD. Improving the spatial resolution of epiretinal implants by increasing stimulus pulse duration. *Science translational medicine*. 2015 Dec 16;7(318):318ra203-.
- [6] Chang YC, Ghaffari DH, Chow RH, Weiland JD. Stimulation strategies for selective activation of retinal ganglion cell soma and threshold reduction. *Journal of neural engineering*. 2019 Feb 4;16(2):026017.
- [7] G. Tianruo, T. David, W. M. John, J. S. Gregg, K. Tatiana, H. L. Nigel, et al., "Electrical activity of ON and OFF retinal ganglion cells: a modelling study," *Journal of Neural Engineering*, vol. 13, p. 025005, 2016.
- [8] A. Alqahtani, A. Al Abed, T. Guo, N. H. Lovell, and S. Dokos, "A continuum model of electrical stimulation of multi-compartmental retinal ganglion cells," in *Engineering in Medicine and Biology Society (EMBC), 2017 39th Annual International Conference of the IEEE*, 2017, pp. 2716-2719: IEEE.
- [9] Alqahtani A, Abed AA, Alharbi Y, Bakouri M, Lovell NH, Dokos S. A varying-radius cable equation for the modelling of impulse propagation in excitable fibres. *International Journal for Numerical Methods in Biomedical Engineering*:e3616.
- [10] P. B. Matteucci, A. Barriga-Rivera, C. D. Eiber, N. H. Lovell, J. W. Morley, and G. J. Suaning, "The effect of electric cross-talk in retinal neurostimulation," *Investigative ophthalmology & visual science*, vol. 57, no. 3, pp. 1031-1037, 2016.
- [11] C. D. Eiber, S. Dokos, N. H. Lovell, and G. J. Suaning, "Multipolar field shaping in a suprachoroidal visual prosthesis," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 25, no. 12, pp. 2480-2487, 2017.
- [12] G. K. Moghaddam, N. H. Lovell, R. G. Wilke, G. J. Suaning, and S. Dokos, "Performance optimization of current focusing and virtual electrode strategies in retinal implants," *Computer methods and programs in biomedicine*, vol. 117, no. 2, pp. 334-342, 2014.

- [13] Abramian M, Lovell NH, Habib A, Morley JW, Suaning GJ, Dokos S. Quasi-monopolar electrical stimulation of the retina: a computational modelling study. *Journal of neural engineering*. 2014 Feb 21;11(2):025002.
- [14] Alqahtani A, Al Abed A, Lovell NH, Dokos S. Optimizing stimulation strategies for retinal electrical stimulation: A modelling study. In: 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) 2019 Jul 23 (pp. 2872-2875). IEEE.
- [15] Hadjinicolaou AE, Savage CO, Apollo NV, Garrett DJ, Cloherty SL, Ibbotson MR, O'Brien BJ. Optimizing the electrical stimulation of retinal ganglion cells. *IEEE Transactions on neural systems and rehabilitation engineering*. 2014 Oct 17;23(2):169-78.
- [16] John SE, Shivedasani MN, Williams CE, Morley JW, Shepherd RK, Rathbone GD, Fallon JB. Suprachoroidal electrical stimulation: effects of stimulus pulse parameters on visual cortical responses. *Journal of Neural Engineering*. 2013 Aug 8;10(5):056011.
- [17] Liang T, Zhao L, Sui X, Zhou C, Ren Q, Qi Y. Threshold suprachoroidal-transretinal stimulation current required by different-size electrodes in rabbit eyes. *Ophthalmic Research*. 2011;45(3):113-21.
- [18] Jensen RJ, Rizzo III JF. Thresholds for activation of rabbit retinal ganglion cells with a subretinal electrode. *Experimental eye research*. 2006 Aug 1;83(2):367-73.

Detection of Atherosclerotic Plaque using Laser-Induced Breakdown Spectroscopy

Srinidi S*

*Dept. of Biomedical Engineering
Sri Sivasubramaniya Nadar College of
Engineering
Chennai, India
srinidi1857@bme.ssn.edu.in*

Lohitaa J

*Dept. of Biomedical Engineering
Sri Sivasubramaniya Nadar College of
Engineering
Chennai, India
lohitaa1834@bme.ssn.edu.in*

Mercy A

*Dept. of Biomedical Engineering
Sri Sivasubramaniya Nadar College of
Engineering
Chennai, India
mercy1837@bme.ssn.edu.in*

Pauline John

*Dept. of Engineering
New York University Abu Dhabi
Abu Dhabi, UAE
paulinejhn@gmail.com*

Madan Mohan B

*Dept. of Cardiology
MGM Healthcare
Chennai, India
dr.madan.dm@gmail.com*

Nilesh J Vasa

*Dept. of Engineering Design
Indian Institute of Technology Madras
Chennai, India
njvasa@iitm.ac.in*

Abstract— Atherosclerosis is one of the major causes of mortality worldwide. Identification of the plaque components in the coronary vessels is necessary to estimate the risk of cardiovascular disease and treatment at the early stage. Laser-induced breakdown spectroscopy (LIBS), a robust elemental analysis, was performed to obtain the elemental composition of normal and calcified plaque. For this process, phantoms mimicking the atherosclerotic plaques were fabricated using chemical constituents that highly match the optical properties of normal and calcified plaque. Further, optical transmission studies were performed in these phantoms around the wavelength of 1300 nm. A transmittance percentage of ~ 44.8% and ~ 0.75% were measured in the non-calcified and calcified plaque phantom, respectively. Studies based on LIBS provided prominent calcium peaks in calcified plaque at 400 nm, 590 nm and 780 nm, compared to non-calcified plaque. This can serve as a basic tool for the optical characterization of plaque and also lay a foundation for quantitative and elemental analysis of atherosclerotic plaque.

Keywords—Laser-induced breakdown spectroscopy, atherosclerotic plaque, calcified, phantoms

I. INTRODUCTION

Cardiovascular disorders tend to affect the normal functioning of the heart. Atherosclerosis is a disease affecting the coronary arteries, which are the blood vessels supplying blood to the muscle mass of the heart [1]. It is a buildup of fats and cholesterol that results in the thickening of the intimal layer of arteries. Development of the plaques start to evolve with the buildup of small LDL cholesterol crystals in the intima (internal layer of artery) [2]. The plaque progresses with the proliferation of fibrous tissues, lowering the blood flow. The connective tissue is made up of the fibroblast. Calcium gets deposited in the blood vessels which ends up in the narrowing of the arteries, causing sclerosis [3]. Excising atherosclerotic plaque from the human body to study its optical properties is an invasive technique, so phantoms were designed to mimic the optical properties of human arteries and plaque which were used for studying the feasibility of using optical techniques. Existing imaging technologies such as computed tomography (CT), intravascular ultrasound (IVUS), magnetic resonance imaging (MRI), catheter-based contrast angiography, etc.,

pose limitations of limited temporal and spatial resolution, long scanning time, motion artifacts, risk of exposure to radiation, requirement of contrast agent, inability to distinguish the arterial wall layers, calcified and noncalcified plaque [4].

In this proposed work, the optical properties of atherosclerotic plaque were analyzed using transmission spectroscopy and the feasibility of characterizing elemental composition of atherosclerotic plaque phantoms was performed using laser induced breakdown spectroscopy (LIBS). LIBS is a method which helps in identifying the elemental composition of all types of materials independent of their state and it also helps in measurement of multiple elements [5]. Our results compare and evaluate the performance of optical techniques for further extension towards diagnostic and therapeutic applications in cardiology.

II. MATERIALS AND METHODS

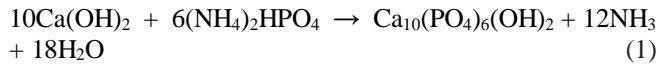
Phantom preparation and optical analysis were performed according to the workflow process shown in Fig. 1. Phantoms were prepared using gelatin, India ink and intralipid 20% in different compositions to mimic the atherosclerotic plaque. Gelatin was used as the base material for mimicking the plaque phantoms as it is water-soluble at high temperatures greater than 50° C. It can also mimic the optical properties of tissue when used with a scattering agent. India ink (Dr. Ph Martin's) was used as the absorbing agent in mimicking the plaque phantoms [6]. Intralipid 20% (Fresenius Kabi) was chosen to be the scattering agent as it is biologically similar to lipid membrane of cells and its organelles of the body [6].



Fig. 1. Schematic of the workflow process.

Samples were prepared using different concentrations of water, gelatin, ink and intralipid to mimic the normal artery and plaque phantoms. Non-calcified phantoms were prepared using the aforementioned constituents. Hydroxyapatite

(HAP) is a calcium phosphate substance that is similar to hard tissue in terms of morphology and composition which was used to mimic the calcified plaque [7]. Since hydroxyapatite is expensive, hydroxyapatite crystals were prepared using 7.42 g of calcium hydroxide and 7.92 g of diammonium hydrogen phosphate in 100 ml of water, as shown in Eq. (1) [8].



III. EXPERIMENTAL SETUP

A. Transmission Spectroscopy Setup

Transmission studies were performed using the transmission spectroscopy to analyze the optical properties of the atherosclerotic plaque phantoms. Transmission studies were performed for a wavelength of 1300 nm using a supercontinuum light source (Leukos – SM250-IR) which produces light ranging from wavelength 900 nm to 2800 nm. A band pass filter (Edmund Optics) of wavelength range from 1280 nm to 1330 nm was used to limit the bandwidth of the output signal to the required optical window to perform transmission studies. The light transmitted through the sample that is mounted on the sample holder is detected by the optical spectrum analyser (AQ6370B, Yokogawa) via an optical fiber. The schematic of the transmission spectroscopy setup is shown in Fig. 2.

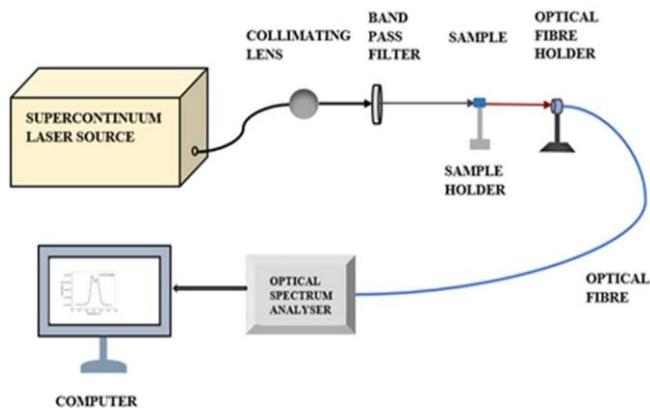


Fig. 2. Schematic of transmission spectroscopy setup.

B. Laser-Induced Breakdown Spectroscopy (LIBS) Setup

A Q-switched Nd³⁺: YAG laser source was used to generate pulses of the laser in a 10 ns duration at a wavelength of 1064 nm. The energy of the laser source was fixed at 30 mJ. The focusing lens L1 was placed after the laser source, with a focal length of 25 cm. It was followed by a sample holder where the sample was mounted. Followed by the sample holder was the focusing lens L2 which was placed at a focal length of 100 cm. The collecting fiber was used to collect and transmit the signal to the Ocean optics (USB2000+UV-VIS-ES) spectrometer which was connected to a computer where the data was collected and plotted using the Origin software. A schematic of the LIBS setup is shown in Fig. 3.

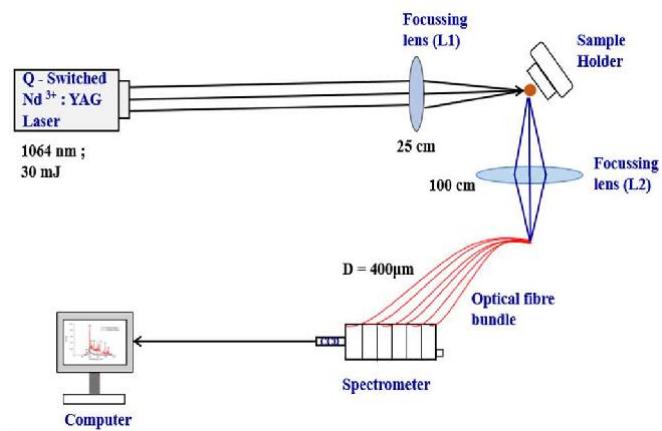


Fig. 3. Schematic of LIBS setup.

IV. RESULTS AND DISCUSSION

A. Transmission Studies

The transmitted intensity is governed by Beer Lambert's law which is represented by the following Eq. (2),

$$I = I_0 e^{-\alpha l} \quad (2)$$

where,

I = Intensity of light transmitted through the sample

I_0 = Intensity of incident light

α = Absorption coefficient (cm^{-1})

l = Optical path length (cm)

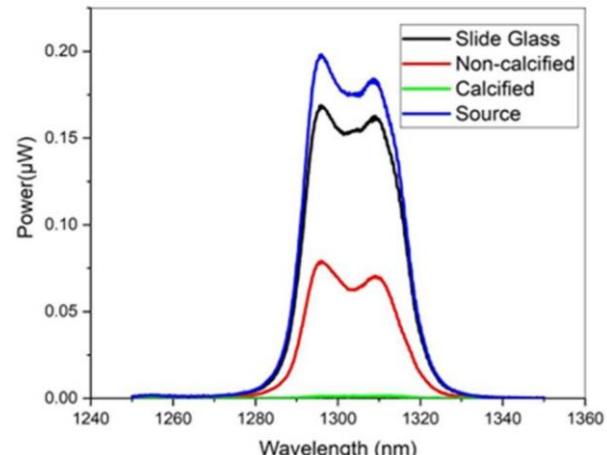


Fig. 4. Spectrum of source and transmission spectra of the samples (substrate - slide glass, calcified and noncalcified plaque) at the center wavelength of 1300 nm.

Figure 4 shows the spectrum of the source (blue trace), transmission spectrum of slide glass without the sample (black trace) and the two samples calcified (green trace) and non-calcified plaque phantom (red trace), in the wavelength range of 1240 nm to 1360 nm, with the center wavelength of 1300 nm. Intensity of calcified plaque phantom decreases compared to the calcified plaque phantom. From Eq (2), absorbance and transmittance are calculated for non-calcified and calcified plaque samples as tabulated in Table I.

TABLE I. TRANSMITTANCE (%), ABSORPTION COEFFICIENT AND ABSORBANCE CALCULATED FOR NON-CALCIFIED AND CALCIFIED PLAQUE AT WAVELENGTH 1300 NM

Phantoms	Transmittance (%)	Absorption coefficient (mm^{-1})	Absorbance
Non-calcified	44.8	0.8	0.348
Calcified	0.748	4.89	2.126

B. LIBS Studies

i. Non-Calcified plaque phantom

At a wavelength of 500 nm, carbon was found to be prominent and at a wavelength of 780 nm, oxygen was obtained and at a wavelength of 680 nm and 490 nm, hydrogen was found to be prominent, in non-calcified plaque phantom. Trace amounts of phosphorus, sodium and nitrogen were also obtained as shown in Fig. 5. These elements are the constituents of the materials like gelatin, ink and which was used to mimic the non-calcified plaque phantom. The results were verified using the NIST Database [9].

ii. Calcified plaque phantom

At a wavelength of 400 nm calcium was found to be prominent and from the range of 400 nm to 900 nm many traces of calcium were found as shown in Fig. 6, in calcified plaque phantom. This is due to the presence of hydroxyapatite which was coated on the plaque sample to mimic the calcified plaque. The results were verified using the NIST Database.

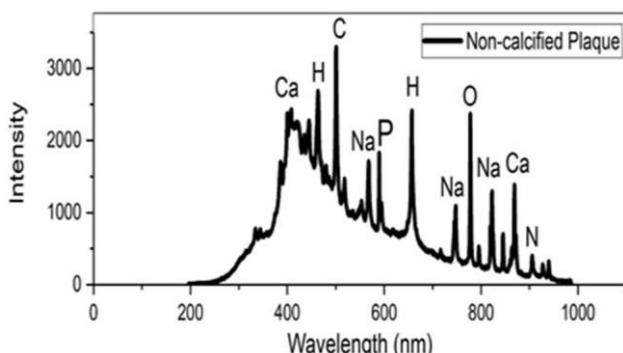


Fig. 5. LIBS analysis of non-calcified plaque.

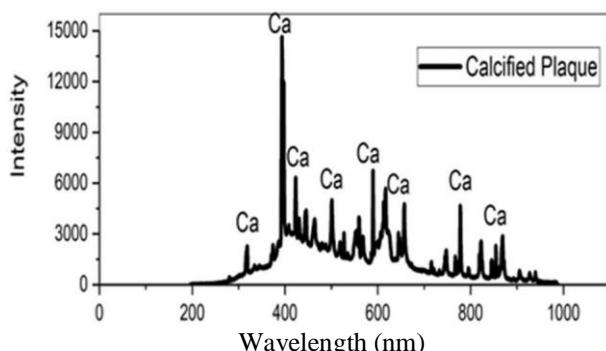


Fig. 6. LIBS analysis of calcified plaque.

The elements at the prominent peaks in the different samples and their corresponding wavelengths are tabulated in Table II.

TABLE II. EXPERIMENTAL RESULTS OBTAINED USING LIBS

Samples	Corresponding wavelengths (nm) of the elements	Elements denoted by prominent peaks
Non-calcified plaque	500 680 790	Carbon Hydrogen Oxygen
Calcified plaque	400 590 780	Calcium

IV. CONCLUSION

In transmission studies it was clearly observed that the transmittance percentage decreases for calcified plaque compared to non-calcified plaque because of the increase in the attenuation coefficient of calcified plaque compared to non-calcified plaque. Change in the absorption characteristics of calcified, non-calcified plaque phantom and normal blood vessels can also help in the diagnosis of atherosclerosis. From the results obtained using transmission spectroscopy studies, absorption coefficient was estimated using Beer Lambert's law and was found to be 8 cm^{-1} for non-calcified plaque phantom and 48.9 cm^{-1} for calcified plaque phantom which was verified from the literature. In the case of LIBS studies, at a wavelength of 400 nm Ca peaks were prominent for calcified plaque and for non-calcified plaque elements C, H, O were prominent in the wavelength range of 500 – 800 nm. This method could help to investigate the severity and progression of atherosclerotic plaque. Timely treatment could be provided by diagnosing the plaque formation at an early stage, thereby, avoid surgeries and also reduce mortality rates.

One of the major advantages of LIBS is its easy detection of the trace elements and their spectral emission lines. It is a rapid, non-destructive method that makes it possible to analyze all types of materials regardless of their state. The other major advantage of LIBS is its non-ionizing radiation. LIBS can be used in real time diagnosis of atherosclerosis plaque using a portable version of the setup where the entire unit can be modified into a compact and simple optical fiber-based setup, incorporated with a catheter that could be used by the doctors to detect plaque in the patients. These optical studies also lay a foundation for quantitative and elemental analysis of atherosclerotic plaque. Besides diagnosis of early-stage atherosclerosis, this technique can also be extended towards therapeutic application in cardiology for the removal of plaque with appropriate laser parameters.

ACKNOWLEDGMENT

The authors would like to acknowledge the support provided by Prof. R. Sarathi, Department of Electrical Engineering, IIT Madras, in accessing the lab facilities for performing LIBS experiments, Dr. K. Aravindh, Research Centre, SSN College of Engineering and Ms. Anbuthangam Ashokan, Department of Biotechnology, IIT Madras, for the support in preparing phantoms.

REFERENCES

- [1] Bruning, R.S. and Sturek, M., 2015. Benefits of exercise training on coronary blood flow in coronary artery disease patients. *Progress in Cardiovascular Diseases*, 57(5), pp.443-453.
- [2] Ouwens, D.M., Sell, H., Greulich, S. and Eckel, J., 2010. The role of epicardial and perivascular adipose tissue in the pathophysiology of cardiovascular disease. *Journal of Cellular and Molecular Medicine*, 14(9), pp.2223-2234.
- [3] Giachelli, C.M., 2004. Vascular calcification mechanisms. *Journal of the American Society of Nephrology*, 15(12), pp.2959-2964.
- [4] Zhang, M., Xie, Z., Long, H., Ren, K., Hou, L., Wang, Y., Xu, X., Lei, W., Yang, Z., Ahmed, S. and Zhang, H., 2022. Current advances in the imaging of atherosclerotic vulnerable plaque using nanoparticles. *Materials Today Bio*, p.100236.
- [5] Singh, V.K. and Rai, A.K., 2011. Prospects for laser-induced breakdown spectroscopy for biomedical applications: a review. *Lasers in Medical Science*, 26(5), pp.673-687.
- [6] Calfon, M.A., Vinegoni, C., Ntziachristos, V. and Jaffer, F.A., 2010. Intravascular near-infrared fluorescence molecular imaging of atherosclerosis: toward coronary arterial visualization of biologically high-risk plaques. *Journal of Biomedical Optics*, 15(1), p.011107.
- [7] Dorozhkin, S.V., 2009. Calcium orthophosphates in nature, biology and medicine. *Materials*, 2(2), pp.399-498.
- [8] Candidato Jr, R.T., Sokołowski, P., Pawłowski, L. and Denoirjean, A., 2015. Preliminary study of hydroxyapatite coatings synthesis using solution precursor plasma spraying. *Surface and Coatings Technology*, 277, pp.242-250.
- [9] <https://physics.nist.gov/PhysRefData/ASD/LIBS/libs-form.html>

2D Fluid-Structure Model of Aortic Valve Using a Derived Muscular Model Equation

Yousef Alharbi

Department of Biomedical Technology
College of Applied Medical Sciences,
Prince Sattam bin Abdulaziz University
 Al-Kharj 11942, Saudi Arabia
 y.alharbi@psau.edu.sa

Abstract— Fluid-structure interaction (FSI) models provide promise for the study of normal and disordered heart function, particularly the effect of fluid dynamics on the motion of aortic valve (AV) leaflet. The objective of this study is to examine the impact of AV normal leaflets and AV with stenosis on blood flow by developing 2D FSI computational model to simulate the AV leaflets when interacting with blood motion. The AV leaflets was modelled using a spring-like derived formulation. Furthermore, the blood flow was modeled as an incompressible Navier-Stokes fluid, and the mesh deformation of the fluid and leaflets due to trans-valvular pressure on AV borders and the ensuing leaflets movement was calculated using the Arbitrary Lagrangian Eulerian (ALE) technique. The 2D model was able to reproduce AV leaflets opening profile as a result of the transvalvular pressure on AV. The maximum velocity simulated at the peak inlet velocity (at 0.2 s) was 1.47 m/s, and this value agreed with the reported range (1.1 – 1.7 m/s) from previous studies. Additionally, the simulated maximum flow velocity of the AV model with calcification was 2.58 m/s and it classifies as AV with mild stenosis based on the reported range from the literature. The developed FSI model provides insights into the impact fluid dynamics on the AV leaflets movement using a spring-like derived formulation.

Keywords—aortic valve, fluid-structure interaction, calcification, transvalvular pressure, linear elastic

I. INTRODUCTION

Coronary artery diseases and aortic valve stenosis (AVS) is a major contributor to death and illness worldwide [1]. Aortic valve stenosis occurs due to the buildup of calcium on the valve leaflets during the cardiac cycle. Clinical studies have shown that this stenosis not only affects the flow of blood in the aortic root but also has a major impact on coronary artery flow and may lead to the development of coronary artery disease. The prevalence of AVS affects approximately 2-7% of the population over 65 years of age [2, 3]. Consequently, this disease can lead to chest pain, shortness of breath, and fainting, and can increase the risk of heart failure and stroke [2].

Using computational fluid dynamics (CFD) simulations can provide a precise and accessible way to study outcomes that cannot be determined through clinical estimation or current experimental techniques. These simulations consider the unique interaction between the aortic valve (AV) and the structure of the left heart [4, 5]. In addition to improving the precision and reach of Transcatheter aortic valve implantation (TAVI), CFD modeling shows promise as a technique for guiding structural cardiac intervention [6, 7].

The objective of this study is to examine the impact of AV calcification on blood flow. The AV leaflets was modelled using a derived sprig-like model equation. The interaction between the fluid and AV leaflets was strongly coupled using ALE formulations. The simulations were performed using COMSOL Multiphysics 5.6 (COMSOL AB, Sweden). Such models might give insight into the interaction between AV function and blood flow that would not be easily obtained from clinical or experimental data alone.

II. METHODS

A. Model Geometry

The geometry and dimensions of the 2D model are shown in Fig. 1 (a). The geometrical parameters (Table. 1) was adopted from [8] with some modifications.

TABLE I. 2D AV MODEL GEOMETRICAL PARAMETERS.

Parameter	Value (mm)	Description
a	22.2	Ventricular side length (inlet)
b	20.36	AV height
c	23.5	Ascending aorta length (outlet)
d	65	The model length
e	13.5	AV leaflets length

B. Fluid Properties and Formulation

In this model, the incompressible Navier-Stokes equations were used to govern the fluid motion. We used a $k-\epsilon$ RANS model to simulate turbulent flow.

$$\rho \frac{\partial \vec{u}_f}{\partial t} + \rho(\vec{u}_f \cdot \nabla_x)\vec{u}_f = \nabla_x \cdot [-p\mathbf{I} + (\mu + \mu_T)(\nabla\mu + (\nabla\mu)^T)] + \mathbf{F} \quad (1)$$

$$\rho \nabla_x \cdot \vec{u}_f = 0 \quad (2)$$

$$\rho \frac{\partial k}{\partial t} + \rho(\vec{u}_f \cdot \nabla_x)k = \nabla_x \cdot \left[\left(\mu + \frac{\mu_T}{\sigma_k} \right) \nabla_x k \right] + P_k - \rho \epsilon \quad (3)$$

$$P_k = \mu_T [\nabla_x \vec{u}_f : (\nabla_x \vec{u}_f + (\nabla_x \vec{u}_f)^T)] \quad (4)$$

$$\rho \frac{\partial \epsilon}{\partial t} + \rho(\vec{u}_f \cdot \nabla_x)\epsilon = \nabla_x \cdot \left[\left(\mu + \frac{\mu_T}{\sigma_\epsilon} \right) \nabla_x \epsilon \right] + C_{\epsilon 1} \frac{\epsilon}{k} P_k - C_{\epsilon 2} \rho \frac{\epsilon^2}{k} \quad (5)$$

Where \vec{u}_f and p denote the fluid velocity and pressure, respectively. \mathbf{I} is the identity matrix, ∇_x denotes the spatial gradient, ϵ is the turbulent kinetic energy dissipation rate, and $\mu_T = \rho C_\mu \frac{k^2}{\epsilon}$ is the turbulent viscosity. Moreover, σ_k , σ_ϵ , $C_{\epsilon 1}$, $C_{\epsilon 2}$ and C_μ are the model parameters. μ_f and ρ denote the viscosity and density of the blood and were set to 4.71×10^{-3} Pa s and 1060 kg/m^3 , respectively.

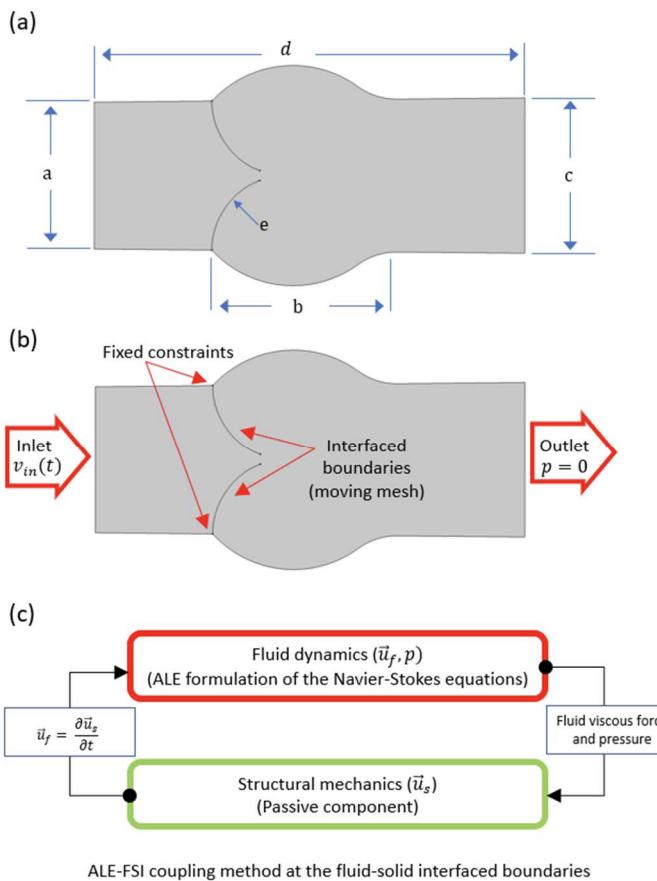


Fig. 1. (a) Model geometry, (b) boundary conditions, and (c) Schematic diagram of AV model workflow.

C. Muscular Model

AV leaflets were modelled as linear spring (see Eq. (6) and (7)). The linear spring model was validated with Laplace law.

$$\begin{cases} \nabla_X \cdot \Gamma_1 = F_x \\ \nabla_X \cdot \Gamma_2 = F_y \end{cases} \quad (6)$$

$$\begin{cases} \Gamma_1 = k \nabla_{T_X} u_{s_x} \\ \Gamma_2 = k \nabla_{T_X} u_{s_y} \end{cases} \quad (7)$$

The ∇_X and ∇_{T_X} are the material frame spatial derivatives and tangential gradients, respectively. Moreover, F_x and F_y are the total fluid forces along x and y directions, respectively. The u_{s_x} and u_{s_y} are the displacement of the leaflets along x and y directions, respectively. The elastic stiffness of AV, k , was set to 2.5 N for normal leaflets and 25 N for AV with stenosis. The stiffness for normal AV leaflets was adjusted to These values were chosen based on the reported young's modulus values by [9].

D. Boundary Conditions

As shown in Fig. 1 (b), a velocity profile was applied at the inlet boundary as expressed in Eq. (8), while zero pressure was set at the outlet boundary.

$$v_{in} = \begin{cases} 0.5 (1 - \cos(5\pi t)) & t \leq 0.4 \\ 0 & 0.4 \leq t \leq 0.6 \end{cases} \quad (8)$$

A strong coupling framework was used to examine the interaction between fluid and elastic boundaries (see Fig. 1 (c)). The AV elastic leaflets edges were subjected to stress caused by fluid viscous stress and pressure. The displacement of the edges was then utilized as a boundary condition to determine the deformation at the interface. The velocity of the fluid on the leaflets was made equal to the wall velocity, and the ALE method was used to solve the fluid interior mesh.

E. Mesh Convergence

A mesh convergence analysis was conducted to determine the

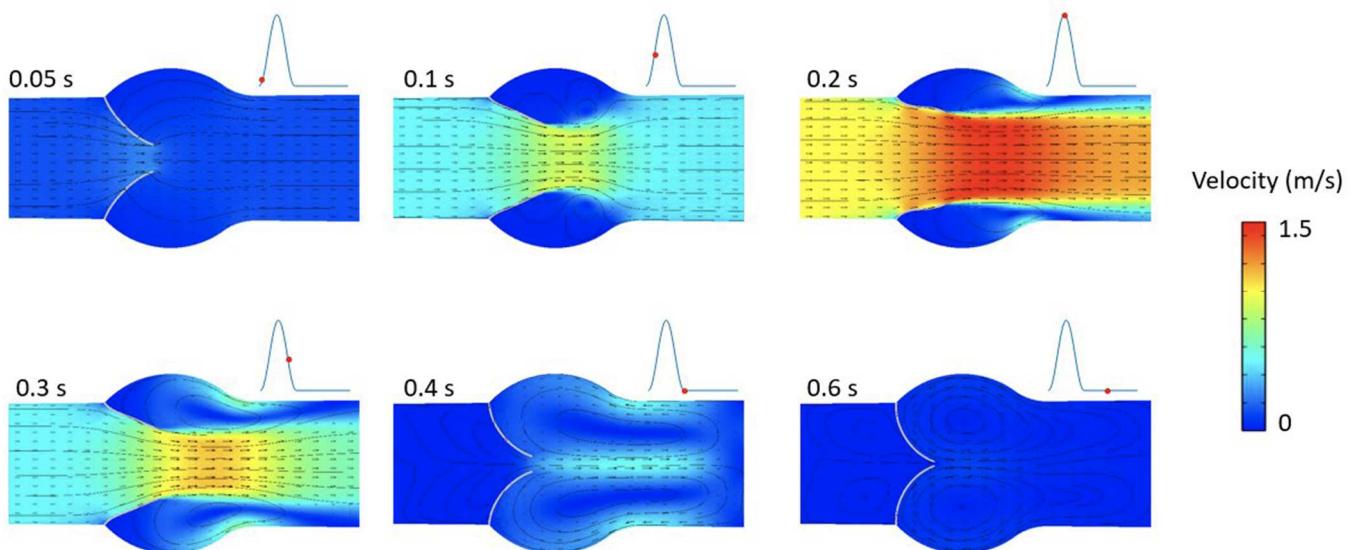


Fig. 2. Velocity and streamlines for normal AV. The time points are shown with red dots on the cardiac cycle.

optimal mesh refinement. The simulation results were stored every 1 ms with a maximum time step of 0.1 ms. Three distinct meshes were simulated with identical model settings, as indicated in the table. 2, and the differences between these mesh refinement options was calculated using Eq. (9). The differences were less than 1% during the different measurements of maximum of velocity, pressure, and leaflets displacement. The total number of elements in the selected mesh was 36.562×10^3 .

$$\text{difference (\%)} = \frac{|M_1 - M_2|}{M_2} \times 100\% \quad (8)$$

TABLE II. MESH CONVERGENCE ANALYSIS.

Number of elements			
Mesh 1	Mesh 2	Mesh 3	
16.284×10^3	36.562×10^3		50.991×10^3
Differences			
Variable	Meshes (1,2)	Meshes (1,3)	Meshes (2,3)
Peak velocity (m/s)	0.013	0.034	0.04
Maximum pressure (mmHg)	0.75	0.96	0.2
Maximum leaflets displacement (mm)	0.31	0.28	0.03

III. RESULTS AND DISCUSSION

A. Flow Velocity

Fig. 2 illustrates the blood flow through the AV leaflet. The maximum velocity simulated at 0.2 s (the peak profile) was 1.47 m/s. The normal peak velocity is in agreement with the reported range (1.1 – 1.7 m/s) [10]. However, calcified AV simulation (see Fig. 3) shows higher flow velocity in comparison with the normal AV. The maximum velocity at 0.2 s was 2.58 m/s, and this value is highly agreed with the reported range (2.5 – 2.9 m/s) [11]. In addition, the maximum flow velocity during the whole simulation is shown in Fig. 4 (a) for both cases.

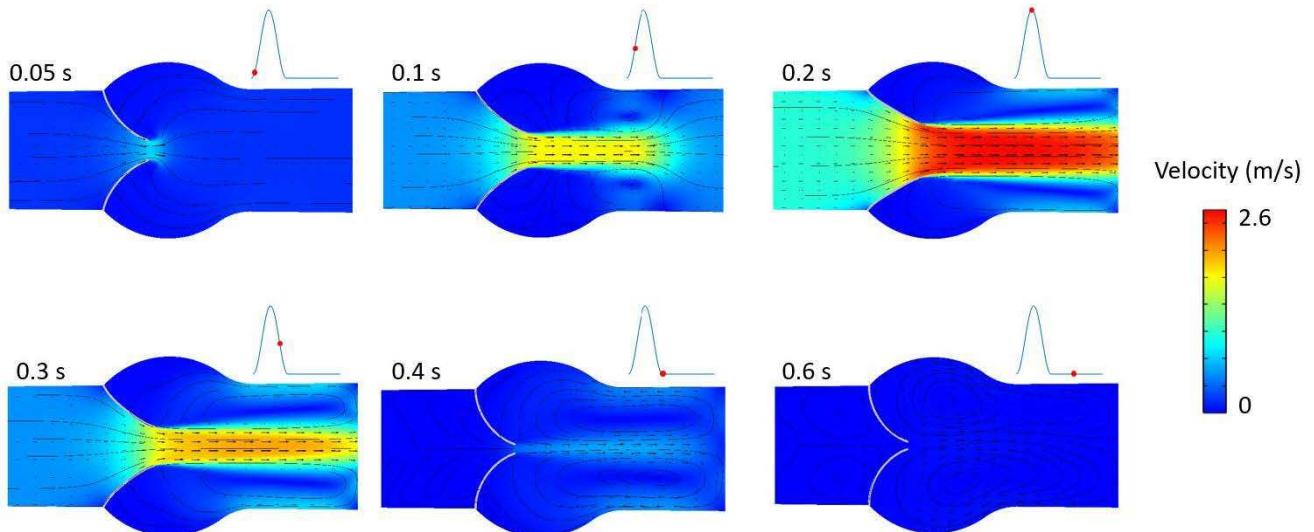


Fig 3. Velocity and streamlines for AV with stenosis.

B. Transvalvular Pressure Gradients and Leaflets Maximal Openning

The transvalvular pressure gradient (TPG) of the AV leaflets was measured by taking the difference of the surface averaged pressure on both sides of the leaflets (ventricular and aortic root sides). The changes of TPG on AV leaflets during the simulation are shown in Fig. 4 (b).

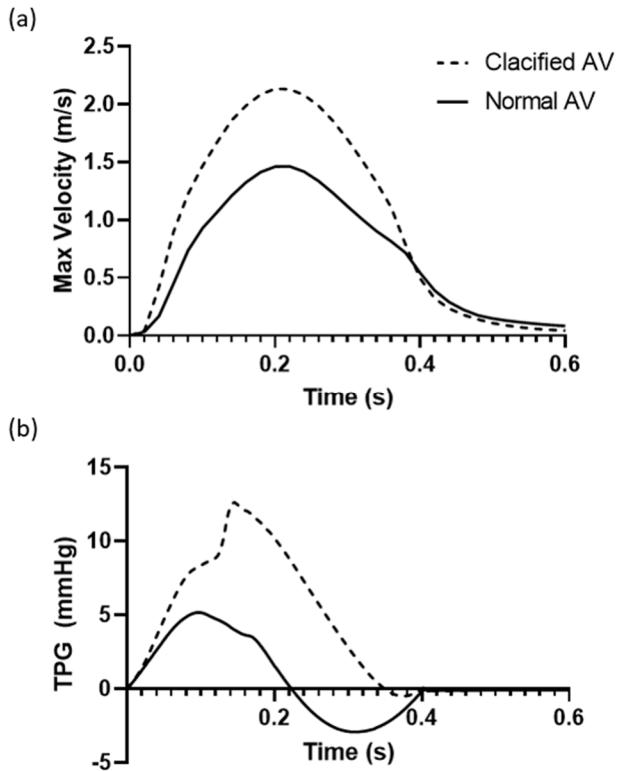


Fig 3. (a) Maximum flow velocity and (b) the transvalvular pressure on AV leaflet.

The maximum value of TPG was 5.16 mmHg for normal AV simulation. This value in a good agreement with the reported value (6.47 mmHg) [12]. On the other hand, maximum TPG for AVS simulation was 12.6 mmHg and this

value highly agreed with previous stated range by [13] (11 – 19 mmHg).

Additionally, there was a noticeable variation in the maximal distance between the AV leaflets for both normal and diseased simulations. The leaflet maximum opening in normal simulation was 15.6 mm, whilst it was 8.7 mm for AV with stenosis. These values well agreed with the simulations results published previously by [14].

The extended work is to develop a 3D AV model coupled with pre- and after- load lumped-parameter models in order to incorporate realistic simulation and simulate aortic stenosis, systemic hypertension, and congestive heart failure. In addition, the effect of left ventricular outflow tract obstruction (LVOTO) on the AV can be investigated.

IV. CONCLUSION

The developed ALE-FSI model in this paper was able to examine the effect of AV calcification on blood flow. The AV was modelled using a derived spring-like model and the interaction between the fluid and AV elastic boundaries was fully coupled using ALE formulations. The simulations results show good agreement with the reported ranges from the literature regarding maximum velocity and TPG values. Such models might give insight into the interaction between AV function and blood flow that would not be easily obtained from clinical or experimental data alone.

REFERENCES

- [1] A. H. Rapp, R. A. Lange, J. E. Cigarroa, E. C. Keeley, and L. D. Hillis, "Relation of pulmonary arterial diastolic and mean pulmonary arterial wedge pressures in patients with and without pulmonary hypertension," *The American journal of cardiology*, vol. 88, no. 7, pp. 823-824, 2001.
- [2] B. Iung and A. Vahanian, "Epidemiology of valvular heart disease in the adult," *Nat Rev Cardiol*, vol. 8, no. 3, pp. 162-72, Mar 2011, doi: 10.1038/nrccardio.2010.202.
- [3] A. T. F. Members et al., "Guidelines on the management of valvular heart disease (version 2012) The Joint Task Force on the Management of Valvular Heart Disease of the European Society of Cardiology (ESC) and the European Association for Cardio-Thoracic Surgery (EACTS)," *European heart journal*, vol. 33, no. 19, pp. 2451-2496, 2012.
- [4] S. Wald, A. Liberzon, and I. Avrahami, "A numerical study of the hemodynamic effect of the aortic valve on coronary flow," *Biomechanics and modeling in mechanobiology*, vol. 17, pp. 319-338, 2018.
- [5] B. Su et al., "Numerical simulation of patient- specific left ventricular model with both mitral and aortic valves by FSI approach," *Computer methods and programs in biomedicine*, vol. 113, no. 2, pp. 474-482, 2014.
- [6] A. A. Basri et al., "Fluid structure interaction on paravalvular leakage of transcatheter aortic valve implantation related to aortic stenosis: A patient- specific case," *Computational and mathematical methods in medicine*, vol. 2020, 2020.
- [7] S. Pasta et al., "Transcatheter heart valve implantation in bicuspid patients with self- expanding device," *Bioengineering*, vol. 8, no. 7, p. 91, 2021.
- [8] H. G. Bahraseman, E. M. Languri, N. Yahyapourjalaly, and D. M. Espino, "Fluid- structure interaction modeling of aortic valve stenosis at different heart rates," *Acta of bioengineering and biomechanics*, vol. 18, no. 3, pp. 11--20, 2016.
- [9] A. R. Kivi et al., "Fluid structure interaction modelling of aortic valve stenosis: Effects of valve calcification on coronary artery flow and aortic root hemodynamics," *Comput Methods Programs Biomed*, vol. 196, p. 105647, Nov 2020, doi: 10.1016/j.cmpb.2020.105647.
- [10] R. Hälvä et al., "Peak flow measurements in patients with severe aortic stenosis: a prospective comparative study between cardiovascular magnetic resonance 2D and 4D flow and transthoracic echocardiography," *Journal of Cardiovascular Magnetic Resonance*, vol. 23, pp. 1-10, 2021.
- [11] L. Chaothawee, "Diagnostic approach to assessment of valvular heart disease using magnetic resonance imaging, part II: a practical approach for native and prosthetic heart valve stenosis," *Heart Asia*, vol. 4, no. 1, pp. 171-5, 2012, doi: 10.1136/heartasia-2012-010124.
- [12] H. Saisho et al., "Ex vivo evaluation of the Ozaki procedure in comparison with the native aortic valve and prosthetic valves," *Interact Cardiovasc Thorac Surg*, vol. 35, no. 3, p. ivac199, Aug 3 2022, doi: 10.1093/icvts/ivac199.
- [13] A. L. Castel, S. Maréchaux, J. Laaouaj, D. Rusinaru, Levy, and C. Tribouilloy, "Relationship between Cutoff Values of Peak Aortic Valve Velocity and Those of Other Doppler Echocardiographic Parameters of Severity in Patients with Aortic Stenosis and Normal Flow," *Echocardiography*, vol. 29, no. 10, pp. 1150-1156, 2012.
- [14] N. Pil, A. G. Kuchumov, B. Kadyraliev, and V. Arutunyan, "Influence of Aortic Valve Leaflet Material Model on Hemodynamic Features in Healthy and Pathological States," *Mathematics*, vol. 11, no. 2, p. 428, 2023.

Designing of IoT Based Portable Vital Health Parameter Monitoring System

Krishnakumar. S*

Department of Biomedical Engineering,
Sathyabama Institute of Science and
Technology,
Chennai, India
drkrishnakumar_phd@yahoo.com

Bethanney Janney J

Department of Biomedical Engineering,
Sathyabama Institute of Science and
Technology,
Chennai, India
jannydoll@gmail.com

Najeeba. M*

Department of Biomedical Engineering,
Sathyabama Institute of Science and
Technology,
Chennai, India
fazah.y@gmail.com

Fazila. A

Department of Biomedical
Engineering,
Sathyabama Institute of Science and
Technology,
Chennai, India.
fazilaahmed217@gmail.com

Sabarivani. A

Department of Biomedical Engineering,
Sathyabama Institute of Science and
Technology,
Chennai, India
sabarivanibme@gmail.com

Abstract— The mandated project is concerned with coming up with a sophisticated and cutting-edge way for average people to assess their health without a doctor's or lab technician's assistance. They can use the suggested equipment to independently monitor themselves. Sensors like the LM35 temperature sensor and blood pressure sensor are used in health monitoring systems. The technology has been modified for a wireless emergency telemedicine system, and in this prototype, it aids the blinded by converting text into speech so they can hear the output. Remote health monitoring raises the standard of care, cuts down on attention, and gives patients more control. The system includes an Arduino, data cable, non-invasive glucometer, digital sphygmomanometer, pulse oximeter sensor, heartbeat sensor, A/D device, signal learning circuit, and mobile. It is a reasonably priced, detachable device. Blood pressure, oxygen saturation, temperature, pulse rate, glucose level, and voice communication will all be displayed as a result of this gadget. In this project, an Arduino Uno is used to demonstrate IoT based e - health monitoring system. The system tracks the patient's critical health metrics and broadcasts the observed data on a specific IP address (Wi-Fi). To demonstrate the effectiveness of the suggested system, a prototype has been created. In order to avert problems and provide care for patients at the appropriate moment, the current research enables people periodically evaluate their health metrics.

Keywords— IoT, E-health monitor, Wi-Fi, sensors, Arduino, Bluetooth, voice communication

I INTRODUCTION

A healthcare system is one that assists society in protecting and enhancing the health of its citizens. Medical professionals and engineers are working to make healthcare and biomedical technologies more advanced, computerized, and portable in order to meet the demands and requirements of patients [1]. As a consequence of rapid medical innovation, contemporary civilizations health-care systems have become significantly more mature and efficient. Individual wearable monitoring systems and portable technology are replacing traditional in hospital systems [2]. Many ideas for generating a faster and more accurate pre-diagnosis including examples for laypeople. The monitors under consideration had to be able to measure any or all of the vital signs such as heart rate, blood glucose level, oxygen saturation, blood pressure, and temperature in an outpatient setting with minimum restrictions on the patient's typical lifestyle. Monitoring blood glucose levels non-invasively using near-infrared spectroscopy. Monitoring

blood glucose levels is essential to preventing diabetic complications and organ damage. Invasive glucose monitoring is painful and can harm nerves, hence non-invasive glucose testing is employed instead [3]. The amount of light that is perceived after passing through the finger can be used to assess blood glucose levels. An Android application (app) is created to display and preserve the measured glucose levels, together with the date and time, in a text file that can be accessed whenever necessary [4]. Exploring the use of voice message input by visually impaired people on mobile devices is difficulty of nonvisual text entering on mobile devices has been the subject of a lot of recent research [5]. The researcher has tried to overcome this problem using gestures, we look into a different mode of voice communication. Visually challenged persons utilized speech more frequently and input lengthier texts than sighted people, according to our findings. Regular blood pressure checks are crucial to prevent severe harm. For patients with diabetes and those just getting out of the hospital, maintaining and monitoring the blood glucose saturation level is essential [6]. To avoid serious consequences, they must get their blood pressure checked on a regular basis. Maintaining and tracking the blood glucose level saturation level is essential for patients who have recently been discharged from the hospital as well as diabetics. When we travel and there is an emergency the portable para monitor will come in handy and affordable. A smartphone application based on Android that communicates with Arduino software and a Bluetooth Module will assist visually challenged people by transforming text to audio [7]. The new technology allows patients in self-isolation or self-quarantine during pandemic of COVID 19 to report daily health parameters and is very difficult to doctors via through mobile phones is a key contribution of this study.

This project is for the common people to check their health parameters regularly. Here for the blind people, we are providing a loud speech output for them and also to give them a compact portable device which is easy to carry. While comparing to other devices in the market we are giving the best quality and a cost-effective device. The device will help in current COVID pandemic situation people don't need to go to any clinics or hospitals for their regular health checkup they can check the vital health parameters in their home itself. Therefore, we gathered the main four health parameters such as SpO₂, BP, temperature

and glucose are measured and displayed text message and voice communication. In this context the present research demonstrates develop a portable vital parameter monitor system with an IP address, Bluetooth, and voice communication module.

II MATERIALS AND METHODS

In order to build the prototypic device, a variety of methods were employed to connect a variety of sensors (sensing body parameters) to an Arduino UNO and interface the Arduino UNO with an LCD 2*16 to review the results. The results will also be displayed on the Android app, which has a text-to-speech converter to make them audible to blind people. All of this device's data is kept on a smart phone, on which someone may install an app to access their medical records whenever they choose. For this reason, it won't be necessary to carry around additional equipment or have visiting doctors check these values. We've added some coding to our programming so that it can calculate associated parameters. This device was used to measure the body temperature, SPO2, blood pressure, and glucose level. Basic medical record by seeing these readings on the smartphone app and LCD.

III BLOCK DIAGRAM

Fig.1 displayed block diagram of a potential health monitoring gadget. In order to properly monitor the patient, we have created a project in which the MAX30100 sensor is used continuously to measure the blood oxygen saturation levels and heart rate. These two variables are shown on an LCD screen. When a patient's blood oxygen saturation level reaches a critical level, the oxygen supply is opened for the patient, and when it returns to normal, it is shut off. Therefore, this project aids in providing care for patients at the appropriate moment and avoiding problems.

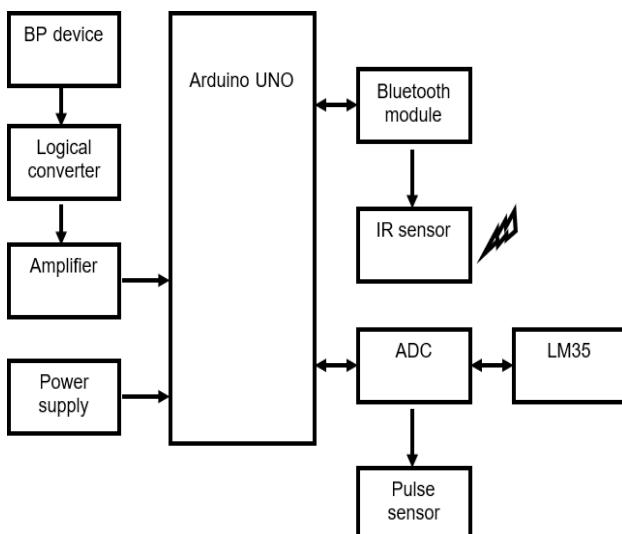


Fig.1. Block diagram of a potential health monitoring gadget

IV HARDWARE USED

A. Power supply module

A step down transformer in the power supply circuit reduces 230 volts into 12 volts. In this circuit, a bridge rectifier made of 4 diodes produces pulsed dc voltage that is then passed to a capacitor filter, which removes any remaining A.C components even after refinement. The regulator receives filtered DC voltage and outputs steady

DC voltage of 12 volts. Although the required power is 5V DC, 230V AC power is turned to 12V AC. As a result, 17V AC power must first be converted to DC power before stepped down to 5V DC.

Arduino uno kit

The Arduino Uno is a open source microcontroller board created by Arduino that is based on Microchip ATmega 328P microprocessor. The board has input/output (I/O) pins for both digital and analogue circuits that can be connected to expansion boards (shields) and other boards. The Arduino IDE (Integrated Development Environment) and a type B USB connector can be used to programme the board's 14 digital I/O pins (six of which can be used to generate PWM output) and 6 analogue I/O pins. It can operate with voltages between 7 and 20 volts and can be fuelled by a USB cable or an external 9-volt battery.

Memory

The flash memory where the Arduino sketch is stored (programmable memory). Variables are created and updated in SRAM when the sketch runs (static random-access memory). Programmers can use EEPROM as a form of long-term memory. Flash memory and EEPROM are examples of non-volatile memory types (the information persists after the power is turned off). SRAM turns volatile and is lost when the power is switched on and off.

Temperature sensor

A number of procedures must be followed to ensure the best sensing accuracy. Self-heating has the potential to decrease accuracy, just like any other temperature-sensing gadget. The lowest current suitable for the application should be utilised with the LM135. At the maximum operating temperature, enough current must be supplied to power the calibration pot, the sensor, and any other external loads. Self-heating mistakes can be eliminated if the sensor is utilised in an environment with consistent thermal resistance. If the device is powered by a current that is temperature stable is feasible. The relationship between the temperature and the heating will depend on the zener voltage. Self-heating and scale factor errors are hence inversely related to absolute temperature.

MAX30100 sensor

MAX30100 combines a heart rate monitor and a pulse oximeter. It is an optical sensor that uses an infrared and red LED to emit two different light wavelengths before measuring the absorbance of pulsing blood using a photodetector. The device's 16-deep FIFO stores the digital output data, which is entirely programmable via software registers. I2C is a digital interface that connects it to the host microcontroller for communication. The pulse oximetry subsystem in the MAX30100 consists of an exclusive discrete temporal filter, 16-bit sigma delta ADC, and ambient light cancellation (ALC). Software can shut down the MAX30100, which runs on 1.8V and 3.3V power sources, with very minimal standby current.

2x16 LCD Display

Use of 16x2 standard alphanumeric LCD display, which is widely available and offer a simple method for displaying status messages for your project. A liquid crystal display, or LCD, screen is a type of electronic display module that has several uses. A 16x2 LCD display is a basic module that is used in a wide range of devices and circuits. It is possible to display 224 characters and symbols on the

16 x 2 intelligent alphanumeric dot matrix display. The command register of the display contains numerous commands. registries of data as a result, both the top polarising filter's angle and the angle of light passing through the polarised glass molecule alter.

B. IR Sensor

A relay is a switch that is electrically controlled. Relays that operate on a solid-state basis are one type, although many relays employ an electromagnet to mechanically operate a switch. Relays are used when a single signal has to control numerous circuits or when many low-power signals need to control different circuits individually. Relays with calibrated operating characteristics and occasionally several operating coils are used to protect electrical circuits from overload or flaws; in modern electric power systems, digital instruments still go by the term of "protective relays."

C. BP sensor with pump

Solid state blood pressure monitor may be used at home is extremely important for people with hyper tension. It is convenient to use and transport, making it especially useful in remote areas with few medical services. The pressure sensor in the chuff serves as the system's primary sensing component. This pressure sensor should be carefully chosen to provide an accurate and trustworthy measurement. One of the pressure sensors utilised in this system is the Honeywell 26 PC SMT pressure sensor. This sensor can measure higher pressure levels and is small and affordable. This sensor can measure pressure more rapidly and precisely since it is directly connected to the printed circuit board.

D. Blue tooth module

Bluetooth module that can transmit in both ways of HC-05. Consequently, it is full duplex. Most microcontrollers are compatible with it. Because it makes use of the Serial Port Protocol, it performs this (SSP). The module communicates at a baud rate of 9600 using a USART, while it also supports different baud rates. This module can therefore be connected to any microcontroller that supports USART.

V SOFTWARE USED

A. Visual Studio

Microsoft Visual Studio is used as an IDE. Simple developer features like IntelliSense code completion and debugging in Visual Studio Code are mixed with the simplicity of a source code editor. The heart of Visual Studio Code is a lightning-fast source code editor that's perfect for regular use. Syntax highlighting, bracket matching, auto-indentation, box selection, snippets, and other features are available in VS Code, which supports hundreds of different languages. Simple developer features like IntelliSense code completion and debugging in Visual Studio Code are mixed with the simplicity of a source code editor. The best web, native, and language-specific technology is included into the architecture of Visual Studio Code.

B. Arduino Software

A useful tool for acquiring new abilities is Arduino. Anyone can begin experimenting by following the step-by-step instructions in a kit or discussing ideas with other Arduino members online while utilising C programming for the AVR on Arduino. A knowledgeable circuit designer can create their own version of the module by extending and

updating the Arduino board blueprints, which are made available under a Creative Commons licence. This makes the hardware open source and upgradeable.

VI RESULTS AND DISCUSSION

The data from each of the project's sensors is measured using the Arduino software. The measured parameters are displayed on an LCD screen, and the outputs are transmitted via Bluetooth to the mobile device where a text to speech converter reads them aloud.

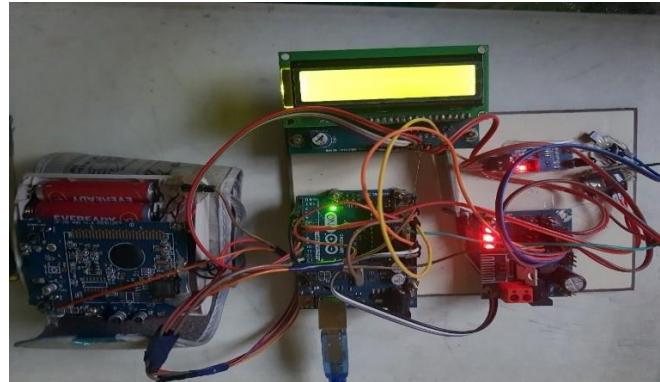


Fig. 2. Hardware set up of health monitoring device

The prototype of health monitoring device is shown in the fig. 2. Which includes an LCD display that will show the output along with an Arduino, power supply, IR sensor, temperature sensor, pulse oximeter sensor, and blood pressure kit. A photodetector, two LEDs, better optics, low-noise analogue signal processing, and pulse oximetry and heart rate signals are employed to detect signals. In this study blood pressure sensor and an IR sensor used to monitor blood glucose levels. Using Arduino software, the data from each sensor in this project is measured. Users keep a finger on the sensor module to check their body temperature, SPO2 and glucose levels are also measured, and a blood pressure band is fastened to the wrist to measure blood pressure. The sensor output is displayed on the 2 x 16 LCD screen.

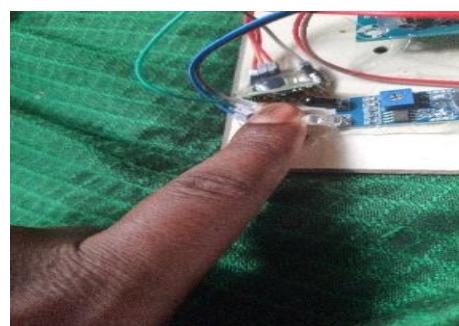


Fig. 3. Sensor Module

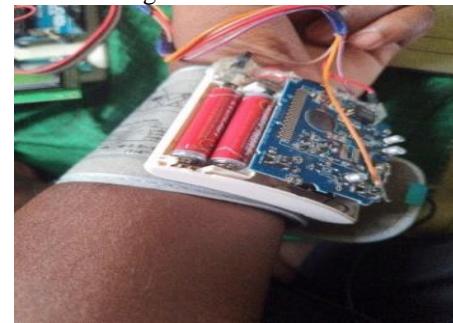


Fig. 4. BP Parameter

Figs. 3 and 4 place users in their finger on the sensor module to check their body temperature. SPO2 and glucose levels are also recorded, and a blood pressure band is attached to the wrist to gauge blood pressure. Fig. 5 shows all vital health Parameters output in display.

H- HIGH BLOOD PRESSURE

S-SPO2 LEVEL

T-TEMPERATURE

G-GLUCOSE LEVEL

D-DETECTING

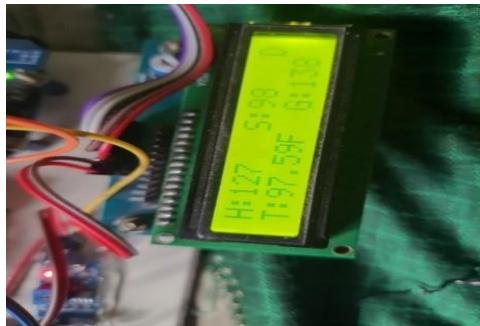


Fig. 5. All the Parameters Output in display

Within 10 seconds of executing the procedure, the results are displayed on the LCD screen. Fig. 6. displayed output converted into speech communication message which can be identified by the visually impaired persons.



Fig. 6. Displayed Output Converted into Speech

Earlier created a similar device that continuously tracks the patient's health signs and is semi-portable. This wearable smart ECG has basic capabilities that patients can use to monitor and perform in real time. Healthcare professionals can use their smartphones to instantly and easily access patient data. Similarly, Checkme, a vital signs detector device, was created to reduce testing time, variations, and incorrect inputs throughout the Electronic Health Record in order to improve routine patient monitoring (EHR). People who are visually impaired can use this programme to translate text to speech by connecting a Bluetooth module to Arduino software. The results show that sensitivity was increased by introducing correlation features to show the natural relationship between the key parameters. The test results demonstrate the system's stability and outstanding measurement accuracy.

Lei Ru et al. [8] reported that human health monitoring systems using the sensitive, precise, and accurate BP300 sensors from the Internet of Things. Along with capacitors, resistors, and air pumps. The pressure sensor converts the blood flow pressure signal in the inflated bandage into a voltage signal with a voltage amplitude that is similar to that of the microcontroller. In order to make health monitoring more individualised and timelier, this project seeks to combine Internet of Things (IoT) technology with health monitoring. Real-time monitoring is provided by incorporating the data using the Internet of Things for processing, networking, and computing. The monitoring system recorded an average temperature of 36.5, 36.4, and 36.5 ($^{\circ}$ C) for the three people whose temperatures were tested at 36.4, 36.7, and 36.5 ($^{\circ}$ C), respectively. The ECG acquisition system uses a clear and convenient display to show the user's ECG.

Reflective pulse oximeter is built into the back cover of a smart portable device, making it simple to monitor heart rate (HR) and SpO2 for use at home [9]. With no significant difference shown by paired t-tests our prototype oximeter can achieve equivalent performance to a clinical oximeter for quick and easy measurement of SpO2 for mobile healthcare. Daarani and Kavitha Mani [10] investigated the non-invasive monitoring of blood glucose levels using a near-infrared sensor. They suggested that non-invasive technique to measure blood sugar levels. The blood glucose level is determined using the variation in NIR light intensity that the photo detector detects after passing through the finger. Non-invasive Glucose Measurement and Correlation with Heart Rate Variability. To monitor blood glucose levels and develop treatment plans for diabetes, blood glucose levels must be monitored often. Glucose measurement techniques and provide various classification characteristics based on size, invasiveness, investigated material, sensor qualities, applied method, activation type, response latency, measurement time, and data access [11]. Internet of things is becoming more and more popular in the field of remote monitoring in health care system [12]. The technology is well advanced that can identify a patient's dangerous condition by analysing sensor data and notifying medical staff, including nurses, doctors and hospital administrators via push alerts. Shahanim Mohamad Hadis et al. [13] developed an Internet of Things-based patient monitoring system that uses sensors to track two key vital signs. Two basic vital signs, such as body temperature and respiration rate, were developed into a patient monitoring system that was disclosed. The monitoring system was constructed using an IoT platform and the ESP8266 Wi-Fi Module and Arduino Mega 2560. Each sensor module uses temperature to determine the amount of each vital sign, which is done using two sensor modules can identify risk to a patient. IoT based diagnosing of heart rate and SpO2 saturation level any patient can be detecting our body disease by an Electronics device. Max30100 electronics device with an oxygen level sensor and a pulse sensor provides the accurate readings, and all readings may be sent to the specific with the aid of a data base using the BLYNK mobile application or Think Speak [14].

Pulse oximeters are used to monitor the heart rate and arterial oxygen saturation was monitored by Madhan Mohan et al. [15]. Pulse oximeter photoplethysmography (PPG)

measurements are significantly impacted by the noise created by movement. The Normalized Least Mean Square is the foundation of the adaptive filter method (NLMS). A tri-axis accelerometer's data shows directionality on each of the three axes. The adaptive filter adds them all up and uses the result as reference noise signal. The recommended low-cost medical technology sensor system can be utilised as a wearable wireless sensor that can be used as a plug-and-play sensor with Arduino to track crucial human health indicators. [16]. The embedded systems, Arduino, GSM module, and other sensors, used in this study provide a straight forward monitoring method that does not require the use of a smartphone or internet connectivity. The recommended device is a portable system that records vital signs data and transmits text messages as warning signals in the event of a medical emergency [17]. The present study can be used for vital health signs monitoring system in smart hospital for COVID-19 patients and visually impaired patients is affordable.

VII CONCLUSIONS

According to the results, sensitivity was enhanced by including correlation characteristics to show the natural link between the crucial factors. The COVID-19 epidemic increased the number of patients who required hospital admission. As a result of growing workloads and insufficient resources, hospitals around the world encountered problems. A completely functional smart hospital or intelligent senior living facility could result from the successful proposed design and execution of a full scale smart vital health sign monitoring system. The monitoring system was completed, allowing it to be used for a variety of clinical applications with minimal cost and installation time for bedridden patients, and comforting care patients using IoT platform. It is concluded that the Internet of Things-based human health monitoring system may give individuals access to daily health management, which is critical to raising the standard and quality of healthcare services.

ACKNOWLEDGMENT

Authors are glad to Sathyabama Institute of Science and Technology, Dept. of Biomedical Engineering, Chennai, for providing all the required provisions to complete the project successfully. The fulfillment that goes to the participation and administration of Sathyabama Institute of Science and Technology.

REFERENCES

- [1] Haleem A, Javaid M, Singh RP, Suman R. Telemedicine for healthcare: Capabilities, features, barriers, and applications. *Sens Int*. 2021; 2:100117. doi: 10.1016/j.sintl.2021.100117. Epub 2021 Jul 24. PMID: 34806053; PMCID: PMC8590973.
- [2] Guk K, Han G, Lim J, Jeong K, Kang T, Lim EK, Jung J. Evolution of Wearable Devices with Real-Time Disease Monitoring for Personalized Healthcare. *Nanomaterials (Basel)*. 2019 May 29;9(6):813. doi: 10.3390/nano9060813. PMID: 31146479; PMCID: PMC6631918.
- [3] Tang L, Chang SJ, Chen CJ, Liu JT. Non-Invasive Blood Glucose Monitoring Technology: A Review. *Sensors (Basel)*. 2020 Dec 4;20(23):6925. doi: 10.3390/s20236925. PMID: 33291519; PMCID: PMC7731259.
- [4] Ventola CL. Mobile devices and apps for health care professionals: uses and benefits. *P T*. 2014 May;39(5):356-64. PMID: 24883008; PMCID: PMC4029126.
- [5] Al-Razgan M, Almoaiqel S, Alrajhi N, Alhumegani A, Alshehri A, Alnafiae B, AlKhamiss R, Rushdi S. A systematic literature review on the usability of mobile applications for visually impaired users. *PeerJ Comput Sci*. 2021 Nov 22;7:e771. doi: 10.7717/peerj-cs.771. PMID: 34901428; PMCID: PMC8627222.
- [6] Dhatariya K, Corsino L, Umpierrez GE. Management of Diabetes and Hyperglycemia in Hospitalized Patients. [Updated 2020 Dec 30]. In: Feingold KR, Anawalt B, Blackman MR, et al., editors. *Endotext* [Internet]. South Dartmouth (MA): MDText.com, Inc.; 2000-.
- [7] Busaeed, S.; Mehmood, R.; Katib, I.; Corchado, J.M. LidSonic for Visually Impaired: Green Machine Learning-Based Assistive Smart Glasses with Smart App and Arduino. *Electronics* 2022, 11, 1076. <https://doi.org/10.3390/electronics11071076>
- [8] Lei Ru, Bin Zhang, Jing Duan, Guo Ru, Ashutosh Sharma, Gaurav Dhiman, Gurjot Singh Gaba, Emad Sami Jaha, and Mehedi Masud, A Detailed Research on Human Health Monitoring System Based on Internet of Things, Deepak Gupta, (2021)
- [9] Zhiyuan Lu, Xiang Chen, Zhongfei Dong, Zhangyan Zhao, Xu Zhang, A Prototype of Reflection Pulse Oximeter Designed for Mobile Healthcare, *IEEE Journal of Biomedical and Health Informatics*, (2015),
- [10] Daarani.P & Kavithamani.A, BLOOD GLUCOSE LEVEL MONITORING BY NONINVASIVE using near infra-red SENSOR, (2017)
- [11] Marjan Gusev, Lidija Poposka, Gjoko Spasevski, Magdalena Kostoska, Bojana Koteska, Monika Simjanoska, Nevena Ackovska, Aleksandar Stojmenski, Jurij Tasic, and Janez Trontelj, Noninvasive Glucose Measurement Using Machine Learning and Neural Network Methods and Correlation with Heart Rate Variability. *Journal of Sensors* Volume 2020, Article ID 9628281, 13 pages <https://doi.org/10.1155/2020/9628281>
- [12] Javaid M, Khan IH. Internet of Things (IoT) enabled healthcare helps to take the challenges of COVID-19 Pandemic. *J Oral Biol Craniofac Res*. 2021 Apr-Jun;11(2):209-214. doi: 10.1016/j.jobcr.2021.01.015. Epub 2021 Jan 30. PMID: 33665069; PMCID: PMC7897999.
- [13] Shahanim Mohamad Hadis, Nor, Nazri Amirnazarullah, Muhammad, Mahdi Jafri, Muhammad, Abdullah, Samihah, IOt Based Patient Monitoring System using Sensors to Detect, Analyse and Monitor Two Primary Vital Signs, *Journal of Physics, Conference Series*, Volume 1535, Issue 1, (2020)
- [14] Sanjay Kumar, & Piyush Agrawal, IoT-based diagnosing of Heart rate and SpO2 saturation level, *IILM AHL College of Engineering & Technology*, (2014), 8-12.
- [15] Madhan Mohan P, Annie Nisha A, Nagarajan Velmurugan, Smiley Jeya Jothi E, Measurement of arterial oxygen saturation (SpO2) using PPG optical sensor, *International Conference on Communication and Signal Processing*, (2016)
- [16] Mian Mujtaba Ali, Shyqyri Haxha, Munna Alam M, Chike Nwibor, Mohamed Sake, Design of Internet of Things (IoT) and Android Based Low-Cost Health Monitoring Embedded System Wearable Sensor for Measuring SpO2,

Heart Rate and Body Temperature Simultaneously, Wireless

Personal Communications (2020) 111:2449–2463

[17] Rachel Austin, Fiona Lobo, Swarnalatha Rajaguru,
GSM and Arduino Based Vital Sign Monitoring System,
(2021), volume 15, 78-89.

DESIGN OF LOW POWER & HIGH-SPEED MULTIWAY MERGE SORTING NETWORKS FOR MEDIAN FILTER

V.MARISELVAM

*Electronics and Communication Engineering
M.Kumarasamy College of Engineering
Karur, India
mariselvamv@gmail.com*

S.R.VISHAL

*Electronics and Communication Engineering
M.Kumarasamy College of Engineering
Karur, India
vissamyraj@gmail.com*

T.SURYA

*Electronics and Communication Engineering
M.Kumarasamy College of Engineering
Karur, India
suryat1472001@gmail.com*

S.VISHNURAM

*Electronics and Communication Engineering
M.Kumarasamy College of Engineering
Karur, India
vishnuram6435@gmail.com*

Abstract- Sifting is an important activity in a wide range of applications, including image, video, and sensor fusion, ATM and network shifting, statistical modeling, forecasting, intelligent systems and robotics, and data gathering and databases. The proposed work's main objective is to develop a new data comparison technique that offers a limited technique for ordering or ranking networks based on speed, energy, and region. In this research, we offer a new area-efficient and comparative estimator for median filter application. This proposed system was written in Verilog HDL, simulated with Modelsim 6.4 c, then synthesized with the Xilinx tool. The suggested system is built on an FPGA Spartan 3 XC3S 200 TQ-144. The design of a low-power, high-speed multiway merge sorting network for median filter applications is presented in this study. To achieve a balance between power efficiency and sorting performance, the suggested network employs a hybrid sorting strategy that incorporates the benefits of both parallel and serial sorting algorithms. The network is built utilizing a hybrid of digital and analogue circuit approaches, resulting in substantial power savings while retaining good sorting performance. The suggested median filter is assessed using simulation and experimental data, which show that it is more successful than existing techniques at lowering power consumption and improving sorting speed.

Keywords- Median Filter; Field programmable gate arrays; Sorting; Sorting Networks; Xilinx; MATLAB; Image Processing; Multiway merge sorting; N-sorters.

I. INTRODUCTION

Computer science analog systems consist of preset wires that carry data and compare elements by number to link pairs of wires, switching the values on the lines if they are out of order, and are abstract devices. They are known as selecting networks when they are created with the intention of performing sorting on predetermined sets of values. Selecting channels differs from ordinary compared sorts in that they cannot handle inputs of any size and that the analogies they make follow a fixed order, regardless of the outcomes of earlier comparisons.

The adaptability of comparison sequences is advantageous for parallel processing as well as hardware implementation. Sorting nets are simplistic, but their theory is unexpectedly complex and deep. Armstrong, Nelson, and O'Connor examined sorting networks for the first time around 1954, and they later filed for a patent on the concept. Both hardware and software may be used to create sorting networks. Donald Knuth explains how binary linear comparators may be built as simple and direct three-state electrical devices. Batcher advocated their use in place of buses and the speedier but even more expensive crossbar switches to construct computer hardware switching networks in 1968. Sorting nets, notably bitonic merge sort, have been used by the GPGPU community to construct sorting algorithms for visual processors since the 2000s. Applications for filtering include data mining, databases, image, video, and data analysis; automation; data science; scheduling; ATM and telecommunications shifting; and automation and machine intelligence.

Filtering is typically performed in hardware for higher-performance applications, employing application-specific integrated circuits or specialized gate sets. The hardware sorting devices are designed for some applications that use image processing (like median filtering), the quantity of inputs might range from nine to tens of thousands. The source of data can be floating-point numbers with a precision of 4 to 256 bits, integers, or binary values. The main problems with hardware and software are hardware costs and power usage. In many applications, there are levels for the total chip area. Keeping chip temperatures low is a critical thing as fabrication methods progress because leakage current rises as temperature increases. The minimum amount of power must be used. An essential objective is to provide hardware-based sorting methods that are inexpensive and power-efficient. The typical method involves updating CAS (review and switch) devices within a network known as a batcher (or bitonic) chain. Hardware-based solutions exceed sequential software-based items due to their parallel nature. Hardware costs and battery life are based on the number of CAS units and the price of each CAS unit.

II. EXISTING WORK

In Existing System presented a dynamic analysis based on streams of equally spaced random bits. In this computing paradigm, each digit is given the same weight. the likelihood of seeing a one in the stream as different to a "0" is used to encode numbers, which are limited to the [0, 1] interval. To demonstrate Conventional Design is used for the CAS Network is Designed.

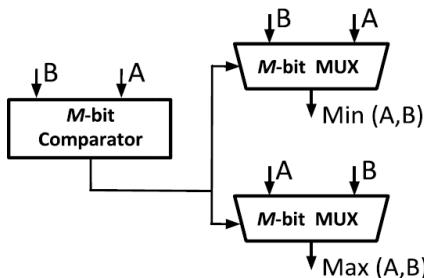


Fig.1.Existing System

A predetermined number of these data and comparator units are used to connect pairs of wires, swapping values if they are not in the correct sequence. Filtering networks are networks that are obviously designed to do sorting on a set of data. Sorting networks differ from traditional comparison sorts in that they cannot handle very large inputs, and their assessment process is predetermined regardless of the results of earlier comparisons. For data and hardware implementation, this independence of comparison sequences is useful. Sorting is typically done in hardware for high-performance applications, employing FPGAs or application-specific embedded processors. Depending on the targeted uses, the electronics sorting devices are constructed differently.

Today's hardware sorting networks employ three sorters that work together at every stage of the fusion process. Our mono devices have been demonstrated to be significantly quicker than previous cutting-edge 3-sorter networks for sorting a modest number N total of unsorted source values in the range of 3 to 16. In order to quickly sort N total > 16 inputs with a high proportion of unsorted inputs using Kenneth Batcher's systemic filtering network. Methods recommend multiple rows and columns that have been integrated, which combine more than three into a single sorted result list made from sorted lists. These multiple-way fusion methods employ just three sorters during the whole merge process and don't outperform Batcher's three merge networks in terms of speed. The UCMS algorithm shown here seems to have no connection to two of these multiway merging techniques.

One group of academics presented SS-M(k), an algorithm for a systematic multiway merge sorting system. Recently, a multiway merging method for CPUs that seems to be analogous to the SS-M(k) was developed. The main goal of the Kolte group was to find the median of pixel windows, but they also provided whole sorting networks for multiway merge sorting for both odd and even rows. Their method is thought to require four more sorters to complete sorting a 4*4 rectangle, while the VV/UCMS approach only needs two additional 3 sorters. It is unknown how they construct their sorting networks.

It is a challenging issue due to the competing objectives of low power consumption and quick sorting speed. Multiway merge sorting is a popular sorting approach for median filters since it is efficient and simple to apply. The trade-off between power consumption and sorting speed is one of our project's most difficult problems. This involves careful design of the sorting network, which includes selecting appropriate sorting components and optimizing the circuit topology. Median filters sometimes necessitate the simultaneous sorting of many data streams, increasing the complexity of the sorting network. To obtain the greatest potential performance, the design of our median filters necessitates a mix of advanced circuit design approaches, optimization algorithms, and rigorous trade-off analyses.

The development of this method is a critical topic of study in computer science and electrical engineering. One of the most difficult aspects of constructing median filters is striking a balance between power usage and sorting speed. To address these issues, researchers have devised a number of approaches for improving the performance of multiway merge sorting networks. Overall, the design requires a mix of hardware and software modifications to attain the needed performance.

III. PROPOSED WORK

The proposed system has a technique for designing a digital-stream three-cell sorter that uses logic circuits with the suggested variables and muxes in high-value double time. Next, we look at a three-cell sorter median filter, which has three rows and three columns of cell sorters on both sides. This filter acts as a median and maintains minimization and maximisation after the process. Finally, we also use a VLSI part that reads text files. We can also use MATLAB as a part that also reads the text file into MATLAB because it should be on a text file because it converts pixels into an image and displays it as an image. Another system path is to acquire a jpg picture, convert it to RGB, then to grey-level data, then to source images, and finally to a text document. The MATLAB part illustrates the conversion of RGB to binary picture flow. The proposed data comparator includes two inputs and two outputs as its fundamental elements. Data signals have an internal structure made up of logic circuits that provide a signal that drives two multiplexers, which generate a high or low value based on the signal to identify which of the two inputs is greater.

The basic action of the comparator is to subtract bits. A network of compare and swap (CAS) devices is frequently used. is set up, with hardware expenses and power requirements changing depending on the number and the price per CAS block. Our study indicates a labelling, space, and energy-efficient "shear sorting exploiting the "Three Cell Sorter" sorting method for networks. Two inputs and two outputs make up the proposed information comparator's basic structure. The data comparators' internal structure consists of digital logic that generates a signal that drives two multiplexers, which, depending on the signal, produce a high or low value to indicate which of the two inputs is higher. Fig. 1 provides a block schematic of the data comparators' fundamental layout. Subtracting two digits is the comparator's fundamental operation.

Block Diagram

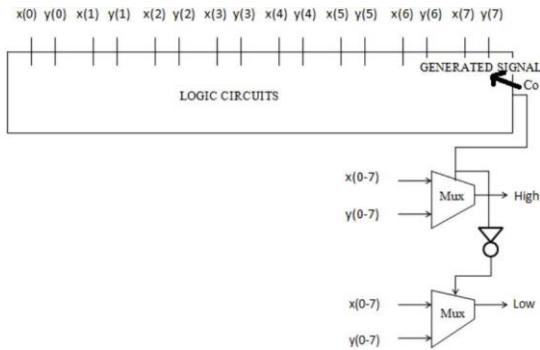


Fig.2. Block Diagram

A. APPLICATION OF UNARY DESIGN: MEDIAN FILTER:

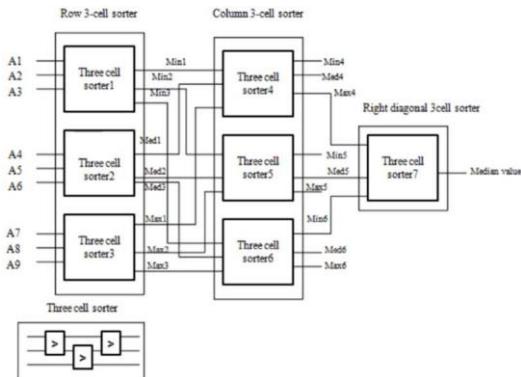


Fig.3. Median Filter

B. PROPOSED SYSTEM FLOW:

RGB to binary picture conversion workflow
(MATLAB Part)

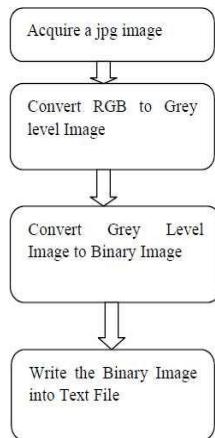


Fig.4. System Flow

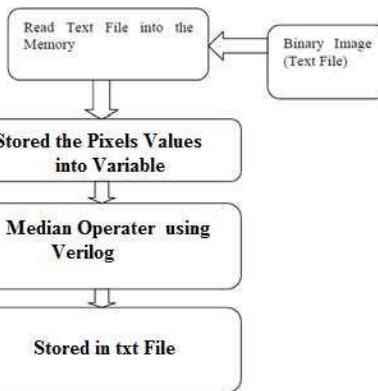


Fig.5. VLSI Part

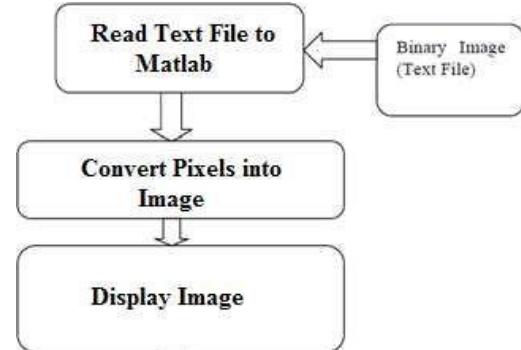


Fig.6. MATLAB PART

C. FIELD-PROGRAMMABLE GATE ARRAY (FPGA)

A Spartan-3 FPGA kit with a median filter can be used to achieve multiway merge sorting. By breaking up the data into smaller subarrays and sorting each one separately before merging them back together, the multiway merge sorting method can handle big datasets. The full dataset may be sorted by repeating this process. The size of the dataset and the quantity of RAM that is readily accessible often dictate the number of subarrays. A technique for digital signal processing called a median filter is frequently used to eliminate noise from a signal. It functions by taking the median value obtained from a collection of samples inside a sliding window and substituting it for the current value. This may aid in noise reduction and signal smoothing. You could design the sorting algorithm in a hardware description language like VHDL or Verilog and then implement multiway merge sorting on an FPGA.

The sorting process may be greatly accelerated by programming the FPGA to carry out the processes in parallel. The finished product can either be delivered to an output device or kept in external memory. You may use a sliding window strategy to get the median value of the subarrays and then add a median filter to the sorting process. A hardware implementation of a median filter algorithm, such as a quick-select method or a bubble-sort algorithm, can be used to accomplish this. Each subarray can receive the median filter before being merged back together. Overall, putting multiway merge sorting with a median filter into practise on a Spartan-3 FPGA kit can be difficult and call for a solid grasp of digital signal processing and hardware design. Nevertheless, a high-performance sorting algorithm that is appropriate for large datasets may be created with careful preparation and execution.

IV. METHODOLOGY

The essential point is that we will implement the VLSI code in Matlab, as well as receive the output in the FPGA Spartan kit and show it on the LED, as well as simulate it in the Xilinx ISE. In a sorting network, a median filter may be used to identify the median value in a sequence of integers. A An algorithm known as a sort channel can arrange a group of integers in a certain pattern by comparing and exchanging them. The median filter compares and swaps nearby items in the sequence until the centre element, or median, is in its proper location. This median value may then be applied to image or signal processing tasks like noise reduction or edge identification. The development of our project required the use of sophisticated algorithms and strategies to maximise the performance of the sorting network. The primary goal is to reduce power usage while maintaining fast sorting rates. This is accomplished by the employment of techniques like pipelining, parallel processing, and the application of specialised sorting algorithms. Furthermore, sophisticated circuit design approaches like dynamic power management and voltage scaling can be used to cut power usage even further. Overall, the objective is to build a sorting network that can effectively and swiftly sort massive volumes of data while consuming as little power as possible.

The design of a low-power, high-speed multiway merge sorting network for a median filter using the three-shell sorter technique involves the following methodology: The first method is to define the problem statement and the objectives of the project. In this case, the goal is to create a low-power, high-speed multiway merge sorting network for a median filter using the three shell sorter technique. Conduct extensive research on the three-shell sorter technique, multiway merge sorting networks, and median filters. This will provide a better understanding of the concepts and techniques involved in the design process. Choose the appropriate design tools, such as CAD software, simulators, and programming languages, to develop and test the design. Select an appropriate algorithm for the multiway merge sorting network. The algorithm should be efficient and optimised for low power and high-speed operations, and then the circuit should be designed for the multiway merge sorting network using the selected algorithm.

The circuit should be optimised for low power consumption and high-speed operations to Using simulators to simulate the design and test its functionality The simulation should be performed under various conditions to ensure the design is robust and reliable. And to optimise the design to achieve the desired performance metrics, such as low power consumption, high-speed operation, and accuracy. Implement the design on a suitable hardware platform, such as an FPGA or ASIC. Test and validate the design to ensure it meets the design requirements and objectives. In our document, we describe the design process, including the design specifications, algorithms, circuit diagrams, simulations, optimizations, and test results.

V. SIMULATION IMPLEMENTATION

A. HARDWARE DESCRIPTION LANGUAGE (HDL)

Digital systems with millions of gates may now be implemented on a single silicon chip because of a sharp rise in the logic density of silicon chips. Two well-liked HDLs are VHDL and VERILOG, which can give a thorough and precise description of a design. The U.S. Department of Defense financed the development of VHDL, which stands for VHSIC hardware description language, in the early 1980s. It offers clear representations of designs at various abstraction levels and can provide gate-level implementations of designs from their VHDL descriptions. VERILOG and ADA are comparable; however, VERILOG differs from ADA in a few keyways. This tutorial's goal is to introduce the modelling language VHDL. The standard language for describing the design and operation of integrated circuits is called VHDL (ICs). The Very High-Speed Integrated Circuits (VHSIC) programme of the US government created it, and the Institute of Electrical and Electronic Engineers (IEEE) approved it in 1987. It is stated using VHDL-93 and later versions' syntax and is subject to revision at least every five years. Differences in syntax shouldn't be a problem.

B. VERILOG

One of the two primary Hardware Description Languages (HDL) utilized by academic and industrial hardware designers is Verilog. As most electrical and computer engineers learn C in college, Verilog is quite similar to C and is well-liked by them. Gateway Design System Company, now a division of Cadence Design Systems, Inc., first released Verilog in 1985. Verilog HDL was a Cadence-only language until May 1990, when Open Verilog International (OVI) was established. In the hope that the language would gain more recognition and the market for software products connected to Verilog HDL would expand more quickly, Cadence was encouraged to make the language available in the public domain.

VI. RESULTS

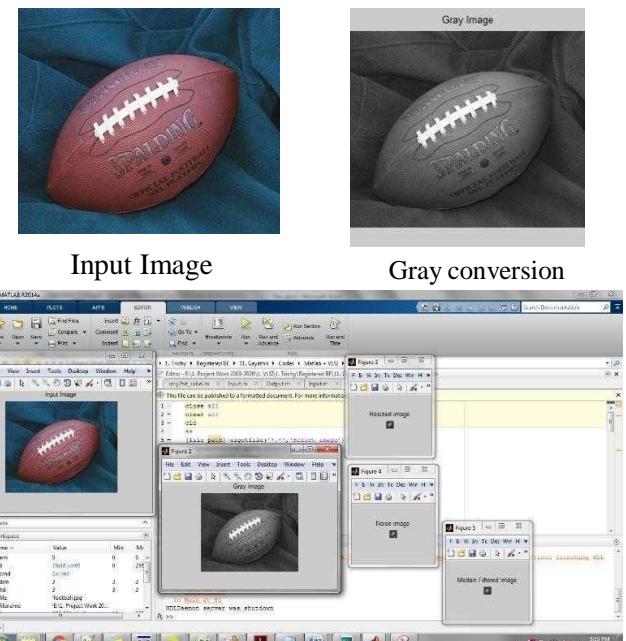


Fig.7. OVERALL OUTPUT

Method name	Area in number of LUT			Delay		
Name	LUT	Gate Count	Slices	Delay	Gate or logic delay	Path or route delay
Conventional	432	311	217	Total 58.688 Ns	29.693 Ns	28.995 ns
Proposed	511	3849	259	54.946 Ns	28.442 Ns	26.504 ns

Fig.8. AREA AND DELAY COMMPARISON TABLE

VII. CONCLUSION

Some applications have made significant use of batcher sorting network. They are valued by communication systems and communication switching networks due to their linear shape. Yet, a tags data type construction of a bigsorting network is costly since it requires so many CAS units. With higher input data resolution, the VLSI cost rises noticeably. Such networks' limited use is a result of their expensive hardware and high energy usage. This study indicates an architecture of sorting networks based on substring computing that is both space and power efficient. The handling of data has an impact on the basic gates that make up the required processes. The cost of converting datafrom/to binary is the only significant overhead in the technique. When compared to the expenditures of a traditional stacked binary system, more than only space and power benefits are shown.

REFERENCES

- [1] R. B. Kent and M. S. Pattichis, "Design of High-Speed Multiway Merge Sorting Networks Using Fast SingleStage N-Sorters and N-Filters" IEEE Access, vol. 10, pp. 2169-3536, 2022.
- [2] R. B. Kent and M. S. Pattichis "Design, implementation, and analysis of high-speed single-stage N-sorters and N filters," IEEE Access, vol. 9, pp. 2576-2591, 2022.
- [3] S. W. A.-H. Baddar and B. A. Mahafzah, "Bitonic sort on a chainedcubic tree interconnection network," J. Parallel Distrib. Comput., vol. 74, no. 1, pp. 1744-1761, Jan. 2014.
- [4] A. Alaghi, W.-T. J. Chan, J. P. Hayes, A. B. Kahng, and J. Li, "Trading accuracy for energy in stochastic circuit design," J. Emerg. Technol. Comput. Syst., vol. 13, no. 3, p. 47, May 2017.
- [5] A. Alaghi and J. P. Hayes, "Survey of stochastic computing," ACM Trans. Embedded Comput. Syst., vol. 12, no. 2s, pp. 92:1-92:19, 2013.
- [6] A. Alaghi, C. Li, and J. P. Hayes, "Stochastic circuits for real-time image-processing applications," in Proc. 50th ACM/EDAC/IEEE DAC, May 2013, pp. 1-6.
- [7] R. Chen and V. K. Prasanna, "Computer generation of high throughput and memory efficient sorting designs on FPGA," IEEE Trans. Parallel Distrib. Syst., vol. 28, no. 11, pp. 3100-3113, Nov. 2017.
- [8] W. Qian, X. Li, M. D. Riedel, K. Bazargan, and D. J. Lilja, "An architecture for fault-tolerant computation with stochastic logic," IEEE Trans. Comput., vol. 60, no. 1, pp. 93-105, Jan. 2011.
- [9] A. Farmahini-Farahani, H. J. Duwe, III, M. J. Schulte, and K. Compton, "Modular design of high-throughput, low-latency sorting units," IEEE Trans. Comput., vol. 62,no. 7, pp. 1389-1402, Jul. 2013.
- [10] N. Guo et al., "Energy-efficient hybrid analog/digital approximate computation in continuous time," IEEE J. Solid-State Circuits, vol. 51, no. 7, pp. 1514-1524, Jul. 2016.
- [11] M. H. Najafi, D. J. Lilja, M. Riedel, and K. Bazargan, "Power and area efficient sorting networks using unary-processing," in Proc. IEEE 35th Int. Conf. Comput. Design (ICCD), Nov. 2017, pp. 125-128.
- [12] E. Nikahd, P. Behnam, and R. Sameni, "High-speed hardware implementation of fixed & runtime variable window length 1-D median filters," IEEE Trans. Circuits Syst. II, Exp. Briefs, vol. 63, no. 5, pp. 478-482, May 2016.
- [13] V. A. Kumar, V. Nandalal, K. Sheikdavood and S. J. Kumaresan, "Smart Application Based Circuit Breaker with Secure Passcode," 2022 IEEE 2nd International Conference on Mobile Networks and Wireless Communications (ICMNWC), Tumkur, Karnataka, India, 2022, pp. 1-5, doi: 10.1109/ICMNWC56175.2022.10031851.
- [14] V. Mariselvam, A. Angelin Peace Preethi, Akilarasu Ganeshan, "Etched Triangle Resonator Dual Band Microstrip Band Pass Filter", July 2022,Wireless Personal Communications 125(12):1-8, DOI:10.1007/s11277-022-09620-2.
- [15] C. Vivek, S. Palanivel Rajan, "Z-TCAM : An Efficient Memory Architecture Based TCAM", Asian Journal of Information Technology, ISSN No.: 1682-3915, Vol. No.: 15, Issue : 3, pp. 448-454, 2016.
- [16] S. Palanivel Rajan, K. Sheik Davood, "Performance Evaluation on Automatic Follicles Detection in the Ovary", International Journal of Applied Engineering Research, ISSN No.: 0973- 4562, Vol. 10, Issue 55, pp. 1-5, 2015.
- [17] S. Palanivel Rajan, T. Dinesh, "Statistical Investigation of EEG Based Abnormal Fatigue Detection Using LabVIEW", International Journal of Applied Engineering Research, ISSN: 0973-4562, Vol. 10, Issue43, pp.30426-30431, 2015.
- [18] S. Palanivel Rajan, M. Paranthaman, "Characterization of Compact and Efficient Patch Antenna with single inset feeding technique for Wireless Applications", Journal of Applied Research and Technology, ISSN: 1665-6423, Vol. 17, Issue 4, pp. 297-301, 2019.
- [19] S. Palanivel Rajan, et.al., "Performance Evaluation of Mobile Phone Radiation Minimization through Characteristic Impedance Measurement for Health-Care Applications", IEEE Digital Library Xplore, ISBN : 978- 1-4673-2047-4, IEEE Catalog Number: CFP1221T-CDR,2012.
- [20] S. Palanivel Rajan, R. Sukanesh, S. Vijayprasath, "Analysis and Effective Implementation of Mobile Based Tele-Alert System for Enhancing Remote Health- Care Scenario", Health MED Journal, ISSN No. : 1840- 2291, Vol. No. 6, Issue7,pp.2370-2377,2012.

TRACK 2

**Ninth International Conference in Biosignals, Images and Instrumentation,
SSNCE, Chennai, March 16 – 17, 2023**

Cuffless BP Measurement Using Single Site Photoplethysmography

Ramakrishnan Maharajan
Department of Electronics and
Communication Engineering
SRM University, AP
Guntur, India
mrama2@gmail.com

Abstract— In this work, a blood flow-based method has been proposed to estimate beat to beat blood pressure parameters from the photoplethysmography (PPG) signal from a single PPG sensor. PPG signal represents the changes in the blood volume. The first derivative of it reflects the blood flow rate. In this work, the features are extracted from the blood flow rate reflected by the first derivative of PPG signal. The proposed method has been validated using Clinical data available in MIMIC II database. The validation shows that the systolic blood pressure and diastolic blood pressure estimated from a single site PPG signal has the mean error \pm SD as 0.95 ± 5.14 mmHg for the beat-to-beat Pulse Pressure (PP) and 0.402 ± 4.85 mmHg for beat-to-beat Systolic Blood Pressure (SBP).

Keywords—Cuffless Blood Pressure Measurement, Photoplethysmography, Stroke Volume, Pulse Pressure.

I. INTRODUCTION

Hypertension is the major reason for cardiovascular diseases (CVD), which is the major cause for 32% deaths worldwide [16]. In India, cardiovascular diseases were the main cause for more than 2.1 million deaths in 2015 [1]. And 24% of Acute Myocardial infarctions, 21% of Peripheral Vascular Disease and 29% of strokes are directly associated with chronic hypertension conditions [2]. Detection of hypertension in the early stages itself and appropriate medical interventions, lifestyle changes will drastically reduce the mortality rate due to CVDs.

Currently Blood pressure is measured with cuff wrapped around the upper arm or wrist and the blood pressure is measured using sphygmomanometer. In this measurement, the cuff is inflated to give enough pressure to occlude the radial artery. Once the blood flow in radial artery is stopped completely, the pulse cannot be sensed in the forearm. Now the pressure on the cuff is reduced slowly. The pressure at which the pulse is sensed at the forearm/wrist is Systolic Blood Pressure(SBP). And the cuff pressure is further reduced and the pressure at which the Korotkoff sounds are observed through stethoscope is taken as Diastolic Blood Pressure (DBP). This measurement method is the golden standard in clinical practice.

Oscillometric type blood pressure measurement is widely used nowadays in home and clinical environments. This measurement also uses the inflatable cuff. The cuff pressure is increased to 200mmHg and then reduced slowly. At a particular pressure, where the blood flow starts, there is some pressure oscillations produced at the cuff which can be sensed to calculate SBP and DBP.

Cuffless BP measurement have been studied extensively in the literature [11,12,14,15]. Most of the works reported, calculate the blood pressure using Pulse Transit Time (PTT) [17] measured using PPG and ECG sensors.

Pulse Transit Time (PTT) is the indirect measure of the Pulse Wave Velocity (PWV) from which the blood pressure can be estimated. There are two categories of cuffless BP measurement. One is cuffless BP measurement with calibration and another is without calibration. In the cuffless BP measurement using calibration, BP is measured using the standard measurement device along with the BP estimating parameters like PTT, Photoplethysmography Intensity Ratio (PIR) [18] etc to establish the relationship between the parameters/features of the measurement techniques and the blood pressure.

Calibration free cuffless BP measurement doesn't require the BP measurement using the standard measurement device for each subject. Such techniques try to model the blood pressure in terms of the measured features with data from huge population [19]. Machine learning techniques are used widely for blood pressure estimation using the features extracted from the physiological signals. Cuffless BP measurement using the BCG [20] and SCG [21] and impedance analysis [22] were also reported in the literature. Few works also have been done to measure the blood pressure using single PPG sensor [5-12].

II. METHODS

A. Photoplethysmography

Photoplethysmography is an optical technique used to measure the blood volume changes in blood vessels in a non-invasive way. The light wave of a particular wavelength is passed through the skin surface. Part of which get absorbed or reflected that corresponds to the blood volume changes [3]. PPG sensor has a light source and photo detector. It is the inexpensive and non-invasive method to detect the cardiovascular pulse propagated throughout the body. Figure 1 shows the setup and working of transmissive and reflective PPG sensor.

Figure 1 also shows PPG waveform. The morphology of PPG signal is similar to that of blood pressure signal. Like BP signal, PPG signal also has systolic point, diastolic point and dicrotic notch which specifies the end of systolic cycle. The blood pressure estimation method proposed uses simple signal processing techniques so that the computational complexity of the method will be less. This enables the method to be used in wearable devices which uses computationally constraint low power microcontrollers. The PPG waveform reflects blood volume changes in blood vessels.

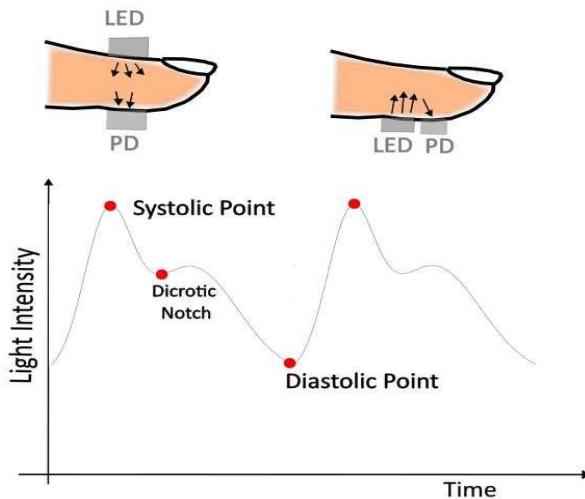


Figure 1. Transmissive and Reflective Type PPG Sensor and PPG Signal

The first derivative of the blood volume will give blood flow rate.

B. Blood Pressure from blood flow rate

Typical PPG signal and its first derivative that reflects blood flow [4] are given in Figure 2. The proposed method uses the first derivative of the PPG signal for finding the stroke volume and peak flow which are proportional to the Pulse Pressure and the Systolic Blood Pressure (SBP) respectively. The Arterial compliance (C) is the change in blood volume (dV) for a given change in the pressure(dP). The arterial compliance can be calculated as a ratio of total volume of blood ejected during a cardiac cycle (stroke volume) to maximum pressure difference (Pulse Pressure).

$$C = \frac{SV}{PP} \quad (1)$$

Where SV is Stoke Volume, and the PP is the Pulse Pressure.

The proposed method, considers the arterial compliance as constant which makes the Pulse Pressure directly proportional to the Stroke Volume. The linear relationship between the pulse pressure and the stroke volume is derived from calibration. The measured Stroke Volume (SV) for multiple cardiac cycles will be taken and calibrated against the actual Pulse Pressure (PP) measured using ABP signal. Here, the outliers are removed to get the better calibrated stroke Volume.

The linear relationship between the systolic Blood Pressure (SBP) and the peak flow is derived using the two element windkessel model as shown in equation (4). The two-element wind Kessel model of arterial tree is a lumped model where the impedance offered by the arterial network to the blood flow is modelled as electrical parallel RC network. R represents systemic vascular resistance and C represents total arterial compliance. Figure 3 shows the 2-element wind Kessel model [5].

Two element windkessel models the low frequency components of arterial network well though it cannot predict the blood pressure wave shape accurately [4]. This makes the two-element windkessel model suitable for predicting the systolic, diastolic and pulse pressure accurately.

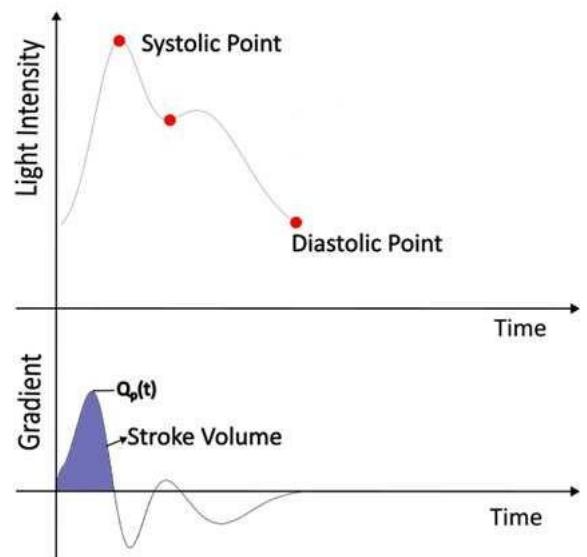


Figure 2. PPG Waveform and its first derivative

For the two-element wind Kessel model, the flow rate is given by the following equation.

$$Q(t) = \frac{P(t)}{R} + C \frac{dP(t)}{dt} \quad (2)$$

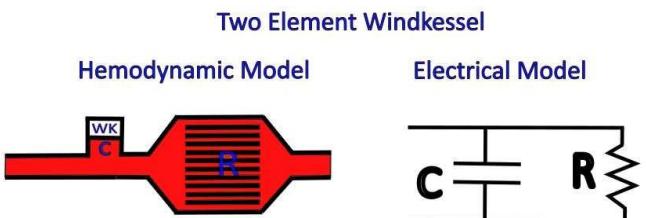


Figure 3. Two element Wind Kessel Model

At the systolic point, the pressure change with respect to time is not significant, the above equation reduces to

$$Q(t) = \frac{P(t)}{R} \quad (3)$$

Assuming the systemic vascular resistance also is a constant, at systolic points,

$$Q(T_s) \propto P(T_s) \quad (4)$$

The calibration is done with the known Systolic Pressure values and the relationship between the peak flow rate and the systolic pressure is derived for the subject.

The proposed method has been applied on the clinical waveform data available on UCI machine learning repository. This contains pre-processed data from Multi-parameter Intelligent Monitoring in Intensive Care (MIMIC) II database [13]. Reference [14] gives the details of the pre-processing done on MIMIC II data. The dataset contains ECG data collected through bipolar limb lead II configuration, PPG from the patient's finger and the Intra Arterial Blood Pressure (IABP) signal acquired from invasive cannula needle in the patient's artery. The clinical waveform data

containing the PPG and the ABP waveform are taken. Figure 4 shows the PPG waveform, its first derivative and the corresponding Arterial Blood Pressure (ABP). The proposed method is applied to the data of 25 subjects. The BP estimation algorithm to estimate the Pulse Pressure and Systolic Blood pressure is given below.

BP Estimation Algorithm:

1. Load the mat file containing PPG, ABP data.
 2. Take PPG and ABP data for i^{th} subject.
 3. Calculate the first derivative of PPG signal (flow signal) using MATLAB ‘gradient’ function.
 4. Find the peaks and its locations using MATLAB ‘findpeaks’ function.
 5. Find the zero crossing of the first derivative of PPG signal.
 6. Find the area under the first derivative of PPG signal between the successive zero crossing points for each cardiac cycle. This gives the value proportionate to the stroke volume.
 7. Apply this stroke volume value on the linear expression relating the stroke volume and pulse pressure. (Derived through calibration)
 8. Apply the peak of 1st derivative of PPG signal on the linear expression relating the maximum flow and systolic pressure. (Derived through calibration)
 9. Find the Actual SBP (peak point) and DBP (valley point) of the ABP signal for each cardiac cycle.
 10. Find the error between the estimated Pulse Pressure, SBP to the actual Pulse Pressure and SBP measured from ABP signal.
 11. Find the mean error and standard deviation of PP and SBP for i^{th} subject.
 12. Go to next subject’s signal. Repeat from step 2 until total number of subjects becomes 25.
 13. Find the mean of the mean PP, SBP of all the 25 subjects.
-

III. RESULTS

A. Validation with Intra Arterial Blood Pressure

Multi point calibration is done to get the best relationship between the pressure parameters and the PPG signal features. For the Pulse Pressure calibration and Systolic BP calibration, the Pulse Pressure/SBP from the invasive ABP waveform and the corresponding measure of stroke volume/Peak flow derived from PPG signal for 50 cardiac cycles have been taken. Few Cycles of PPG signal, its first derivative and the corresponding Arterial Blood pressure is shown in Figure 4. The peak of the first derivative of the PPG, which reflects the peak blood flow is correlated to the SBP and the area under the systolic flow curve which is the area above the zero-crossing line of the first derivative PPG curve. This area reflects the Stroke Volume (SV) which is correlated with the Pulse Pressure.

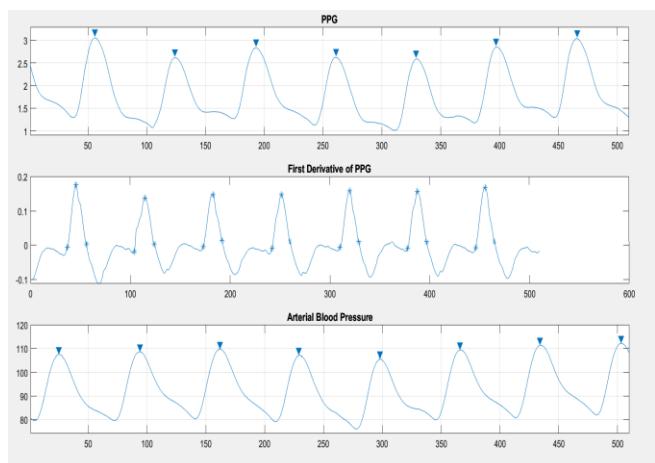


Figure 4. PPG, 1st Derivative of PPG and ABP Signal

Multi-point calibration is done to find the slope(M) and the intersect(C) for the linear relationship. Once the calibration is done for a subject, the relationship derived from the calibration is used to find the beat to beat and the mean of the PP and SBP for 100 cardiac cycles. Figure 5 shows the bland-Altman plot of the beat to beat PP/SBP prediction.

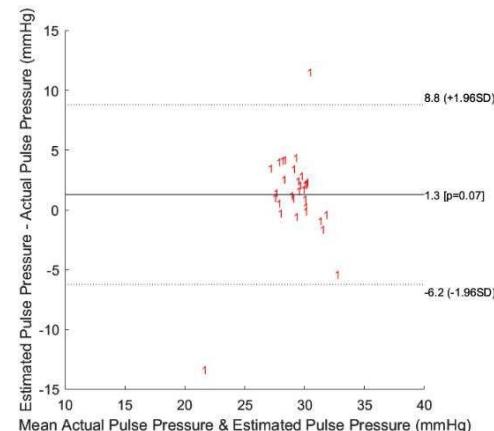


Figure 5. Beat to Beat Pulse Pressure

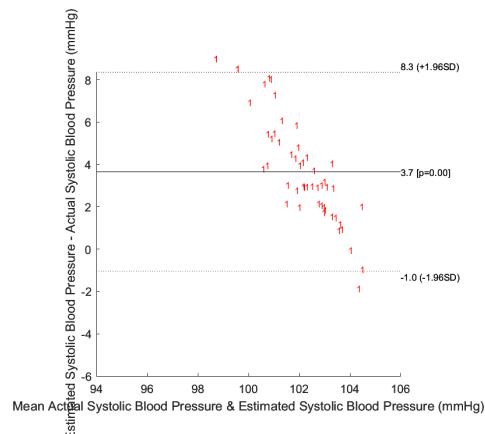


Figure 6. Beat to Beat SBP

IV. DISCUSSION

The results of the flow-based blood pressure measurement have been compared with the work done earlier. Table 1

shows the works done to extract the blood pressure parameters from the single site, finger PPG [15]. Though the combined mean error for the pulse pressure and the systolic blood pressure is lower for this work, the standard deviation for the pulse pressure is higher. And the correlation coefficient which we get ($r = 0.29$ for PP and $r = 0.12$ for SBP) shows that the predicted blood pressure follows the actual blood pressure accurately within certain minimal range. But it should be noted that in this work, the beat-to-beat BP is measured. It is observed that there is less variations in the stroke volume calculated (area under the first derivative of PPG) to the higher pulse pressure variations which can be attributed to the larger deviations in the pulse pressure. In future, the method can be used to measure the average BP to evaluate the performance of proposed method. And the two PPG features can be used with the machine learning and deep learning frameworks to determine the prediction performance of such models.

TABLE I. COMPARISON OF RESULTS OF BP ESTIMATION USING SINGLE SITE PHOTOPLETHYSMOGRAPHY

Study	Participants	Gold Standard	Signal type (Number of Features)	Error (mmHg)
Raichle et al. (2018)[5]	all F, pregnant	ABP	PPG(N/R)	Ms = 5±15
Hsu et al. (2018) [6]	94 (45:M, 49:F)	N/R	PPG(2)	N/R
Alex et al. (2018) [7]	7(3:M, 4:F)	Volume Clamp	PPG(2)	RMSEs < 7
Zadi et al. (2018) [8]	15(18: M, 7:F)	Volume Clamp	PPG(2)	RMSEs < 8
Lin et al. (2018) [9]	22	Volume Clamp	PPG(5), VPG(8),APG (6)	MES = 4±9 MED = 4±5
Chandrasekhar et al. (2018) [10]	35	ABP	PPG(4)	MES = 9 MED = 8
Dey et al. (2018) [11]	205 (90:M, 115:F)	Manual	PPG/VPG/APG(233) + Demographics (3)	MAES = 7±9 MAED = 5±6
Acciaroli et al. (2018) [12]	8 (7:M, 1:F)	Invasive	PPG(10)	RMSED = 7±2
This work	25	Invasive	PPG(1)	MEPP = 0.95 ± 5.14 MES = 0.402 ± 4.853

V. CONCLUSION

In this work, the two features, area under the 1st derivative of PPG curve during systolic phase and the peak of the first derivative of PPG signal are used to derive the Pulse Pressure and the SBP respectively. The models based on simplified expressions are used to estimate the PP and SBP by considering the total arterial compliance and the total peripheral resistance as constant. This approach reduces the computational complexity which aids the implementation of

the proposed method in wearable devices. The combined mean error \pm SD for the 25 subjects is 0.95 ± 5.14 and 0.402 ± 4.853 for beat-to-beat Pulse Pressure and the Systolic Blood Pressure respectively.

REFERENCES

- [1] Ke, C., Gupta, R., Xavier, D., Prabhakaran, D., Mathur, P., Kalkonde, Y., Kolpak, P., Suraweera, W. & Jha, P. Articles Divergent trends in ischaemic heart disease and stroke mortality in India from 2000 to 2015: a nationally representative mortality study. (2018), www.thelancet.com/lancetgh
- [2] Department of Health Research - Ministry of Health & Family Welfare, India, Hypertension: the silent killer. (https://main.icmr.nic.in/sites/default/files/press_realease_files/Hypertension.pdf), Accessed: 2023-01-18
- [3] Elgendi, M. (2012). On the Analysis of Fingertip Photoplethysmogram Signals. *Current Cardiology Reviews*, 8(1), 14–25. <https://doi.org/10.2174/157340312801215782>
- [4] Westerhof, N., Lankhaar, J.-W., & Westerhof, B. E. (2009). The arterial Windkessel. *Medical & Biological Engineering & Computing*, 47(2), 131–141. doi:10.1007/s11517-008-0359-2
- [5] Raichle, C.J.; Eckstein, J.; Lapaire, O.; Leonardi, L.; Brasier, N.; Vischer, A.S.; Burkard, T. Performance of a blood pressure smartphone app in pregnant women: The I PARR Trial (iPhone app compared with standard RR measurement). *Hypertension* **2018**, *71*, 1164–1169. [CrossRef]
- [6] Hsu, P.-C.; Wu, H.-T.; Sun, C.-K. Assessment of subtle changes in diabetes-associated arteriosclerosis using photoplethysmographic pulse wave from index finger. *J. Med Syst.* **2018**, *42*, 43. [CrossRef]
- [7] Alex, R.M.; Zhang, R.; Watenpaugh, D.E.; Behbehani, K. Mathematical Modeling of Arterial Blood Pressure Using Photoplethysmography Signal in Breath-hold Maneuver. In Proceedings of the 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Honolulu, HI, USA, 18–21 July 2018; pp. 2711–2714.
- [8] Zadi, A.S.; Alex, R.; Zhang, R.; Watenpaugh, D.E.; Behbehani, K. Arterial blood pressure feature estimation using photoplethysmography. *Comput. Biol. Med.* **2018**, *102*, 104–111. [CrossRef]
- [9] Lin, W.H.; Wang, H.; Samuel, O.W.; Liu, G.; Huang, Z.; Li, G. New photoplethysmogram indicators for improving cuffless and continuous blood pressure estimation accuracy. *Physiol. Meas.* **2018**, *39*, 025005. [CrossRef] [PubMed]
- [10] Chandrasekhar, A.; Kim, C.S.; Naji, M.; Natarajan, K.; Hahn, J.O.; Mukkamala, R. Smartphone-based blood pressure monitoring via the oscillometric finger-pressing method. *Sci. Transl. Med.* **2018**, *10*, eaap8674. [CrossRef] [PubMed]
- [11] Dey, J.; Gaurav, A.; Tiwari, V.N. InstaBP: Cuff-less Blood Pressure Monitoring on Smartphone using Single PPG Sensor. In Proceedings of the 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Honolulu, HI, USA, 18–21 July 2018; pp. 5002–5005.
- [12] Acciaroli, G.; Facchinetto, A.; Pillonetto, G.; Sparacino, G. Non-Invasive Continuous-Time Blood Pressure Estimation from a Single Channel PPG Signal using Regularized ARX Models. In Proceedings of the 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Honolulu, HI, USA, 18–21 July 2018; pp. 3630–3633.
- [13] A.L. Goldberger, L.A. Amaral, L. Glass, J.M. Hausdorff, P.C. Ivanov, R.G. Mark, J.E. Mietus, G.B. Moody, C.-K. Peng, H.E. Stanley, Physiobank, physiotoolkit, and physionet: components of a new research resource for complex physiologic signals, *Circulation* **101** (23) (2000) e215–e220.
- [14] M. Kachuee, M.M. Kiani, H. Mohammadzade, M. Shabany, Cuff-less High-accuracy calibration-free blood pressure estimation using pulse transittime, *IEEE International Symposium on Circuits and Systems (ISCAS'15)*(2015).
- [15] Hosanee M, Chan G, Welykholowa K, Cooper R, Kyriacou PA, Zheng D, Allen J, Abbott D, Menon C, Lovell NH, Howard N, Chan W-S, Lim K, Fletcher R, Ward R, Elgendi M. Cuffless Single-Site Photoplethysmography for Blood Pressure Monitoring. *Journal of*

- Clinical Medicine. 2020; 9(3):723. doi: <https://doi.org/10.3390/jcm9030723>
- [16] WHO Cardiovascular Diseases -Fact Sheet, [https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-\(cvds\)](https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)). Accessed on 30.1.2013
- [17] R. Mukkamala et al., "Toward Ubiquitous Blood Pressure Monitoring via Pulse Transit Time: Theory and Practice," in IEEE Transactions on Biomedical Engineering, vol. 62, no. 8, pp. 1879-1901, Aug. 2015, doi: 10.1109/TBME.2015.2441951.
- [18] X. -R. Ding, Y. -T. Zhang, J. Liu, W. -X. Dai and H. K. Tsang, "Continuous Cuffless Blood Pressure Estimation Using Pulse Transit Time and Photoplethysmogram Intensity Ratio," in IEEE Transactions on Biomedical Engineering, vol. 63, no. 5, pp. 964-972, May 2016, doi: 10.1109/TBME.2015.2480679.
- [19] Yamanaka Syunsuke, Morikawa Koji, Morita Hiroshi, Huh Ji Young, Yamamura Osamu, " Calibration-Free Cuffless Blood Pressure Estimation Based on a Population With a Diverse Range of Age and Blood Pressure," in Frontiers in Medical Technology, vol. 3, 2021, doi: 10.3389/fmedt.2021.695356.
- [20] C. -S. Kim, A. M. Carek, R. Mukkamala, O. T. Inan and J. -O. Hahn, "Ballistocardiogram as Proximal Timing Reference for Pulse Transit Time Measurement: Potential for Cuffless Blood Pressure Monitoring," in IEEE Transactions on Biomedical Engineering, vol. 62, no. 11, pp. 2657-2664, Nov. 2015, doi: 10.1109/TBME.2015.2440291.
- [21] M. Das, T. Choudhary, L. N. Sharma and M. K. Bhuyan, "Noninvasive Accelerometric Approach for Cuffless Continuous Blood Pressure Measurement," in IEEE Transactions on Instrumentation and Measurement, vol. 70, pp. 1-9, 2021, Art no. 4008109, doi: 10.1109/TIM.2021.3122182.
- [22] B. Ibrahim and R. Jafari, "Cuffless blood pressure monitoring from a wristband with calibration-free algorithms for sensing location based on bio-impedance sensor array and autoencoder," Scientific Reports, vol. 12, no. 1. Springer Science and Business Media LLC, Jan. 10, 2022. doi: 10.1038/s41598-021-03612-1.

Simulation of a Non-Contact Triple-Wavelength Reflectance Pulse Oximeter to Measure Oxygen and Carbon Monoxide Saturations

Swathi Veerabathraswamy

*Dept. of Engineering Design
Indian Institute of Technology Madras
Chennai, India
swathiv1309@gmail.com*

Pauline John

*Department of Bioengineering
New York University
Abu Dhabi, UAE
paulinejhn@gmail.com*

Nilesh J. Vasa

*Dept. of Engineering Design
Indian Institute of Technology Madras
Chennai, India
njvasa@iitm.ac.in*

Abstract—This paper relates to the design of a novel reflectance pulse oximeter, which can measure not only the arterial oxygen saturation, but also that of carbon monoxide, in a non-contact manner. Such an instrument is especially relevant in the monitoring of patients who have a history of smoking, or those who have been exposed to carbon monoxide poisoning, such as from fires. The remote manner of measurement ensures that the risk of contact contamination is eliminated, and that the instrument can be used on infants as well, especially those born prematurely and having sensitive skin. A novel algorithm was devised which uses light sources of three different wavelengths to calculate the above mentioned saturations. A thorough sensitivity analysis was performed to determine the most suitable wavelengths for the purpose, which were determined to be 660 nm, 640 nm, and 940 nm. A system was designed to leverage the above mentioned algorithm, and the same was simulated using Simulink. The simulation gave satisfactory results, with maximum errors in calculation of oxygen and carbon monoxide saturations being 0.36% and 0.66%, respectively. These results are discussed along with measures that can be used to improve accuracy.

Keywords—Arterial blood, carboxyhemoglobin, deoxyhemoglobin, oxyhemoglobin, sensitivity

I. INTRODUCTION

Pulse oximeters were first developed in 1941, and have been used extensively ever since to measure the pulse rate and oxygen saturation of blood in a non-invasive manner. This is done using photoplethysmography (PPG), an optical technique used to detect volumetric changes in blood in peripheral circulation. Typically, two different wavelengths of light (660 nm and 940 nm) are used to quantify the composition of arterial blood [1].

Pulse oximeters have also been adapted to measure other vital parameters including breathing rate and perfusion index [2], [3]. A cutoff frequency of 10 - 35 Hz is recommended for the extraction of breathing rate from the PPG signal [3]. In addition, the PPG waveform obtained can be analyzed to qualitatively understand other variations in the blood volume, including changes in blood pressure [2]. Table I gives a list of vital parameters that can be measured or analyzed using pulse oximetry, along with their normal values [4].

TABLE I: List of vitals that can be assessed using a pulse oximeter

Vital parameter	Normal range	Medical conditions
Oxygen saturation	95-100%	Hypoxemia (< 90%) - can cause complications in organs and tissues
Pulse rate (resting)	1 - 1.6 Hz	Arrhythmia - improper beating of the heart (Bradycardia - <1 Hz; Tachycardia - >1.6 Hz)
Breathing rate	0.16–0.33 Hz	Asthma, anxiety, pneumonia, congestive heart failure or lung disease
Perfusion index	0.02% - 20%	Peripheral artery diseases, diabetes, obesity or blood clots
Blood pressure	Systolic: 90 - 120 mmHg Diastolic: 60 - 80 mmHg	Hypotension (lower than 90/60 mmHg) or hypertension (higher than 140/90 mmHg)

Pulse oximeters can be of two types - transmittance and reflectance - based on whether they measure the intensity of light transmitted through the body part or that reflected from the body part. Commercially, transmittance pulse oximeters are the most widely used, but are limited by their application to extremities like fingertips and earlobes. Their accuracy also depends on factors such as vasoconstriction, probe placement, movement, and gravity artifacts. In recent years, however, reflectance oximetry has gained popularity, as they can be used in diverse locations such as the fingertip, wrist, chest, forehead, and earlobe. Studies, however, have shown that the fingertip is still the most preferred measurement site [5]- [9]. The finger, forehead and wrist reflectance values were found to have correlation coefficients of 0.8867, 0.8477 and 0.6358 against the reference values, leading to the conclusion that reflectance PPG signals obtained from the finger were higher in quality than those obtained on the wrist or forehead [6]. In particular, the reflectance PPG obtained on the chest was found to be highly corrupted by breathing artifacts. The mean amplitude of PPGs were reported to be 1.95 V, 0.21 V, 0.46 V, 0.26 V, 0.67 V and 0.73 V from the finger, upper wrist, lower wrist, arm, earlobe, and forehead, respectively [7].

Both transmittance and reflectance oximeters have to be worn on or placed in contact with the measurement site, which

may present risks related to the transmitting of contagious diseases or infection, and inaccuracies related to the placement of the sensor or movement artifact. For this reason, non-contact oximeters are being studied, but have not been successful enough to be made commercially available. In addition, non-contact pulse oximeters have mostly been designed using cameras and not light sources [10], [11]. This method, however, takes a much longer measurement time compared to conventional light based oximeters (with each session taking between 35 minutes to 3.5 hours) [10], [11], and has a higher cost due to the presence of cameras. This paper attempts to use a reflectance pulse oximeter technique with different source and detector configurations, in conjunction with collimated light sources to measure the vitals in a non-contact manner.

Though pulse oximeters are widely used, they are not without shortcomings and inaccuracies. Due to this reason, many doctors still prefer to use arterial blood gas analysis in critical conditions where accurate measurements are required, which is accepted as the gold standard in measuring oxygen saturation. One of the major causes of pulse oximeters outputting falsely normal or high oxygen saturation is carbon monoxide poisoning [12]. Carbon monoxide (CO) is an odorless, colorless, and tasteless gas, which can cause poisoning by limiting the supply of oxygen to organs. Every year, at least 430 people die in the U.S. from accidental CO poisoning. Approximately 50,000 people in the U.S. visit the emergency department each year due to accidental CO poisoning [13]. The CO concentration in blood is usually between 0.5% and 1.5% in non-smokers. A concentration of 10%-20% usually causes nausea, fatigue, and confusion. As it increases to 21%-30%, it is accompanied by headache, visual impairment and decreased sensory perception. A higher concentration (31%-50%) can cause dizziness, fainting, and vomiting, whereas a very high concentration (>51%) can trigger seizures, coma, and death [14]. To counter this issue, co-oximeters have been developed that use lights of multiple wavelengths and are capable of simultaneously measuring carbon monoxide and methemoglobin saturations as well [15]- [17]. However, these instruments use multiple light sources (up to 12), and are not capable of remote measurement through reflectance.

This work differentiates from the above mentioned approaches by using a novel algorithm using only three light sources, to combine the benefits of a reflectance pulse oximeter with that of a co-oximeter, to measure the saturations of oxygen and carbon-monoxide in a non-contact manner. The operational principles governing such a device are explained in Section II, and the algorithm is then introduced in Section III. Our model uses three light sources, of wavelengths 640 nm, 660 nm, and 940 nm, which were selected after a theoretical sensitivity analysis, which is described in Section IV. There are models that evaluate the accuracy of pulse oximeter algorithms [18], but none for co-oximeters or reflectance oximeters. Hence, a Simulink based model was developed to evaluate the success of this wavelength selection, the details of which are discussed in Section V. Such a simulation will also be useful for preliminary testing of new algorithms,

without the hardware and implementation costs associated with fabricating the same. The results of the same, along with measures that can be taken to improve accuracy, are discussed in Section VI.

II. OPERATIONAL PRINCIPLES

This section discusses the operational principles required to understand the working of our proposed pulse oximeter - Beer Lambert's law, and the composition of arterial blood. Finally, the working of a reflectance pulse oximeter is also explained in brief.

A. Beer-Lambert's Law

Beer-Lambert's law, also known as Beer-Lambert-Bouguer law, states that the absorbance of light of a wavelength λ , $A(\lambda)$, in a homogeneous medium is proportional to the path length, l , and concentration of the absorbing species, c [1], and the proportionality constant $\epsilon(\lambda)$ is called the molar absorption coefficient or extinction coefficient of the substance at wavelength λ . The absorbance can then be calculated as a ratio of the logarithms of the intensity of light incident on the absorbing species ($I_0(\lambda)$) and the intensity of light transmitted after absorption by the species ($I(\lambda)$). In particular, if n absorbing species are present in the solution, then the total absorbance $A_t(\lambda)$ is given by

$$A_t(\lambda) = \sum_{i=1}^n c_i \cdot \epsilon_i \cdot l_i. \quad (1)$$

B. Composition of Arterial Blood

The major species of hemoglobin present in arterial blood are oxyhemoglobin (HbO_2), carboxyhemoglobin ($HbCO$), methemoglobin (MetHb), and reduced or deoxyhemoglobin (Hb). The relative compositions of HbO_2 , $HbCO$, and MetHb in a normal subject are 95-100%, 2-5% (or upto 9% in heavy smokers) and 0-3%, respectively [19]- [20]. The extinction coefficients of each of these components is shown in Fig. 1.

Conventional oximeters assume that only HbO_2 and Hb are present. Thus, they use two wavelengths of light, usually 660 nm and 940 nm, to calculate their saturations using equation (1) with $n = 2$. However, since the extinction coefficients of HbO_2 and $HbCO$ are almost equal at 660 nm, this algorithm detects $HbCO$ as HbO_2 . This gives us falsely high readings of HbO_2 , especially in patients who have a history of smoking or have been exposed to CO poisoning. To counter this, light of three different wavelengths are used to calculate the percentages of HbO_2 , $HbCO$, and Hb .

This measurement is aided by the fact that arterial blood is pulsatile in nature. Because of this, the light transmitted or reflected from the finger or body part will contain an AC and a DC component, due to arterial blood and the surrounding tissues, respectively. Taking the ratio of these AC and DC components ensure that absorption of light due to surrounding tissues, pigments, or water do not factor into the algorithm.

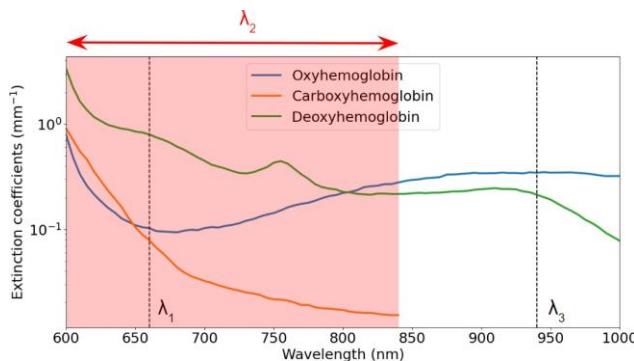


Fig. 1: Extinction coefficients of various components of arterial blood. The black dotted lines (λ_1 and λ_3) show the two wavelengths of light used by conventional oximeters, 660 nm and 940 nm, respectively. The wavelength λ_2 will be chosen from the region marked in red (600 nm - 840 nm) based on a sensitivity analysis.

C. Design of a Reflectance Pulse Oximeter

A reflectance pulse oximeter contains LEDs and a photodetector, all placed on the same side of the body part (most commonly, the fingertip). Light from the LEDs are then alternatively pulsed, and the intensity of light reflected by the finger (after absorption according to Beer-Lambert's law) is measured by the photodetector. This measured waveform is then used to calculate the oxygen and carbon monoxide saturations. In such oximeters, it is important to ensure that the distance between the LEDs and the photodetector is sufficient so that light from the source reaches the arteries before being reflected. The effective penetration depth (z_{max}) can be approximated to be proportional to the square root of the distance between source and detector, d , with a proportionality constant $2/4$ [5], [21].

III. PROPOSED METHODOLOGY

This section describes the novel algorithm we devised to simultaneously measure the saturations of both oxygen and carbon monoxide.

The proposed algorithm makes use of three sources of light (of wavelengths λ_1 , λ_2 and λ_3) instead of two as in conventional pulse oximeters. We now define six constants, a_1 through c_3 , as

$$a_1 = \epsilon_{3d}\epsilon_{2c} - \epsilon_{2d}\epsilon_{3c} \quad (2)$$

$$a_2 = \epsilon_{3c}\epsilon_{1d} - \epsilon_{3d}\epsilon_{1c} \quad (3)$$

$$a_3 = \epsilon_{1c}\epsilon_{2d} - \epsilon_{1d}\epsilon_{2c} \quad (4)$$

$$b_1 = \epsilon_{3d}\epsilon_{2o} - \epsilon_{2d}\epsilon_{3o} \quad (5)$$

$$b_2 = \epsilon_{3o}\epsilon_{1d} - \epsilon_{3d}\epsilon_{1o} \quad (6)$$

$$b_3 = \epsilon_{1o}\epsilon_{2d} - \epsilon_{1d}\epsilon_{2o} \quad (7)$$

$$c_1 = \epsilon_{3c}\epsilon_{2o} - \epsilon_{2c}\epsilon_{3o} + \epsilon_{2d}\epsilon_{3o} - \epsilon_{3d}\epsilon_{2o} + \epsilon_{3d}\epsilon_{2c} - \epsilon_{3c}\epsilon_{2d} \quad (8)$$

$$c_2 = \epsilon_{3o}\epsilon_{1c} - \epsilon_{3c}\epsilon_{1o} + \epsilon_{3d}\epsilon_{1o} - \epsilon_{3o}\epsilon_{1d} + \epsilon_{3c}\epsilon_{1d} - \epsilon_{3d}\epsilon_{1c} \quad (9)$$

$$c_3 = \epsilon_{1o}\epsilon_{2c} - \epsilon_{1c}\epsilon_{2o} + \epsilon_{2d}\epsilon_{1c} - \epsilon_{2c}\epsilon_{1d} + \epsilon_{1d}\epsilon_{2o} - \epsilon_{1o}\epsilon_{2d}. \quad (10)$$

where ϵ_{io} , ϵ_{ic} and ϵ_{id} are the extinction coefficients of HbO_2 , HbCO , and Hb respectively at wavelength λ_i [22]- [23].

When light of wavelength λ_i is reflected from the finger, the photodetector receives a noise-added waveform, from which we extract the AC and DC components, $(AC)_i$ and $(DC)_i$. Using these, we define the ratios-of-ratios R_1 and R_2 as

$$R_1 = \frac{(AC)_1/(DC)_1}{(AC)_3/(DC)_3} \quad (11)$$

$$R_2 = \frac{(AC)_2/(DC)_2}{(AC)_3/(DC)_3}. \quad (12)$$

We then calculate the saturations of oxygen and carbon monoxide as

$$[\text{HbO}_2] = \frac{a_1 R_1 + a_2 R_2 + a_3}{c_1 R_1 + c_2 R_2 + c_3} \quad (13)$$

$$[\text{HbCO}] = \frac{b_1 R_1 + b_2 R_2 + b_3}{c_1 R_1 + c_2 R_2 + c_3}. \quad (14)$$

Since the absorption coefficients of skin pigments such as melanin (10 mm^{-1}), surrounding tissues (0.1 mm^{-1}), and fat molecules (0.001 mm^{-1}) is fairly constant throughout the range of light used (600-1000 nm), the use of the ratios R_1 and R_2 ensure that absorption of light by these substances does not have to be factored in the algorithm. In addition, the external optical noise due to the non-contact geometry is also taken care of using the ratio-of-ratios method.

IV. SENSITIVITY ANALYSIS

Commercial pulse oximeters use light sources of wavelength 660 nm (λ_1) and 940 nm (λ_3) to measure the saturation of oxygen. There have also been models using organic green and red LEDs (532 nm and 626 nm), along with organic photodetectors [24]. But in order to measure the saturation as carbon monoxide as well, an additional third light source is required. In this section, the analysis performed to find the optimal wavelength of light λ_2 , which will make the instrument more sensitive to small changes in the vitals of the patient, is described.

The sensitivity of an instrument can be defined as the change in input parameters required to produce a given change in output parameters. In case of the described pulse oximeter, the input parameters are R_1 and R_2 (as these are calculated directly from the photodetector), and the output parameters are the saturations of oxygen and carbon monoxide ($[\text{HbO}_2]$ and $[\text{HbCO}]$, respectively). Thus, we define the sensitivity of the pulse oximeter as the change in the ratios R_1 and R_2 required to produce a given change (5% or 10% change) in $[\text{HbO}_2]$ or $[\text{HbCO}]$.

First, the values of the ratios R_1 and R_2 for various concentrations of $[\text{HbO}_2]$ and $[\text{HbCO}]$ are calculated, as the wavelength λ_2 is varied from 600 to 840 nm. This range is chosen as it is the intersection between the ranges where absorption of light by other components of the body are minimal,

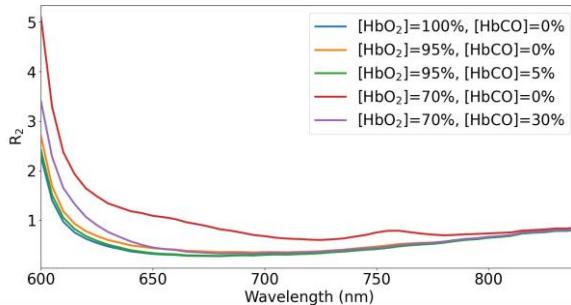


Fig. 2: R_2 values for different values of $[HbO_2]$ and $[HbCO]$

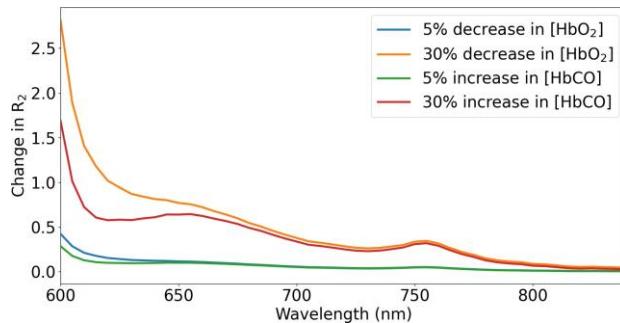


Fig. 3: Change in R_2 with change in $[HbO_2]$ and $[HbCO]$

and where the absorption of light by carboxyhemoglobin is considerable (as shown in Fig. 1). Using this, changes in R_1 and R_2 required to obtain various changes in $[HbO_2]$ and $[HbCO]$ (5% decrease in $[HbO_2]$, 30% decrease in $[HbO_2]$, 5% increase in $[HbCO]$, and 30% increase in $[HbCO]$) are calculated, for different values of λ_2 . Ideally, a high change in R_1 and R_2 is preferred, even for a small change in $[HbO_2]$ and $[HbCO]$. Such a wavelength will be chosen as the optimal wavelength, λ_2 . However, it is to be noted that R_1 remains constant while changing λ_2 , as it depends only on λ_1 and λ_3 , as shown in (11). Thus, the sensitivity analysis was performed using R_2 only. Fig. 2 shows a plot of R_2 for various values of λ_2 . Thereafter, the change in R_2 values are plotted, which is shown in Fig. 3.

Figure 3 was used to choose an ideal value of λ_2 , where R_2 has a high change even for a small change in $[HbO_2]$ and $[HbCO]$. It is observed that values of λ_2 below 650 nm provide high sensitivity. From Fig. 1 and Fig. 3, we focus on a few typical wavelengths in particular:

- 600 nm - Since the sensitivity of the instrument measured is very high at this wavelength, we try to understand the feasibility of using such a light source. However, it can be seen that at wavelengths close to 600 nm, the extinction coefficients of oxyhemoglobin and carboxyhemoglobin are very close (difference between extinction coefficients of oxyhemoglobin and carboxyhemoglobin at 600 nm is 0.12 mm^{-1}). Thus, such a wavelength may not be ideal.
- 805 nm - Since the extinction coefficients of oxyhemoglobin and deoxyhemoglobin are equal at that point, and the extinction coefficient of carboxyhemoglobin is

much lower than that of the other two, we initially hypothesized that the sensitivity might be high using a light source of this wavelength. However, since the extinction coefficient of carboxyhemoglobin itself is very low (0.0165 mm^{-1}) at such a wavelength, the sensitivity turned out to be low as well.

Table II shows the changes in R_2 obtained for four different wavelengths, for different changes in $[HbO_2]$ and $[HbCO]$. From the above discussion and the values in Table II, the ideal wavelength should be less than 650 nm, but not very close to 600 nm. Thus, wavelengths between 630 nm and 640 nm are ideal for λ_2 . In further calculations and analysis, we use a light source of wavelength 640 nm.

TABLE II: Change in R_2 values for various changes in $[HbO_2]$ and $[HbCO]$

2*Change in saturations	Change in R_2 for various values of λ_2			
	600 nm	630 nm	640 nm	805 nm
5% decrease in $[HbO_2]$ (from 100% to 95%, with $[HbCO] = 0\%$)	0.42	0.13	0.12	0.01
30% decrease in $[HbO_2]$ (from 100% to 70%, with $[HbCO] = 0\%$)	2.81	0.87	0.81	0.08
5% increase in $[HbCO]$ (from 0% to 5%, with $[HbO_2] = 95\%$)	0.28	0.09	0.09	0.01
30% increase in $[HbCO]$ (from 0% to 30%, with $[HbO_2] = 70\%$)	1.69	0.57	0.61	0.06

V. SYSTEM DESIGN AND SIMULATION

The proposed system design is explained in this section, along with a simulation to model the working of the same.

A. System Design

The oximeter probe will contain three LEDs (of wavelengths 640, 660 and 940 nm), which are used upon collimation using a pinhole, to ensure that scattering is minimized. The light reflected from the finger or any other body part will be detected by the photodetector TIOPT101. The TIOPT101 was selected as it has a good responsivity (greater than 0.4 A/W) over the required spectrum (600 nm - 1000 nm), and also contains an inbuilt amplifier [25]. The LEDs will be placed equidistant from the photodetector to ensure that the path length of light at all wavelengths is equal.

For non-contact measurement, the probe should be designed as a handheld device, which can be pointed at the site of interest (fingertip, wrist, earlobe, or forehead), from a distance of 20-50 mm. The ideal distance between the LEDs and the photodetector can be calculated from [21], where the required penetration depth from the skin is up to 20 mm (for arteries). Since the device will be placed at a distance of 50 mm from the skin, the required distance between the LED and photodetector was calculated to be greater than 25 mm using [5]. As light reflected from all depths less than 20 mm will be detected with such a setup, it can be used on infants and premature babies as well, where the required penetration depth is lower

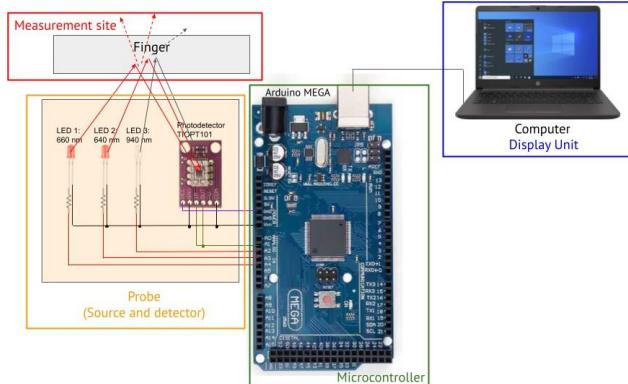


Fig. 4: A schematic of the described system design. The components used are - LEDs (640 nm, 660 nm, 940 nm), Photodetector (TIOPT101), Arduino Mega 2560 microcontroller, and a computer.

than in adults. In order to reduce motion artifact, replaceable transparent spacers can be used to mount the finger, which will not affect the reflectance or transmission of light, while aiding in the non-contact measurement. Furthermore, if necessary, optical fibers can be used to define a path from the probe to the tip of the finger and back, without touching the finger itself, such as in [26].

The signal obtained from the photodetector should be amplified if necessary, and passed through a low pass filter, to filter components having frequency greater than 50 Hz, which are not clinically significant and are most likely to be noise and motion artifacts. The filtered signal can then be passed through an analog to digital converter, after which a demultiplexer should be used to separate out the signals corresponding to the three different wavelengths. These steps can all be performed digitally, without the use of hardware, using a microcontroller such as Arduino Mega 2560. This can then be used to calculate the ratio-of-ratios R_1 and R_2 , after which the concentrations $[HbO_2]$ and $[HbCO]$ can be calculated using (13) and (14). Information including the saturations, waveform, pulse rate, breathing rate, and perfusion index can be displayed on the LCD unit or monitor as required. Fig. 4 shows the system design described above.

B. Simulation

To model and simulate the system described above, Simulink and MATLAB are used. The complete model is shown in Fig. 5a, the components of which are described below:

- Pulse generators - Three pulse generators are used, with a time period of 0.01 s, and a duty cycle of 20%. A phase difference of 0.002 s was also added between subsequent pulse generators to ensure that all LEDs are not powered simultaneously.
- LED Block - This block contains the three LEDs of wavelengths 640 nm, 660 nm, and 940 nm, which are each powered by a pulse generator.

- Finger block - The finger block, shown in Fig. 5b, calculates the absorption of light by blood using a MATLAB function block, which has the extinction coefficients programmed into it. The arterial blood volume (modeled using a wave generator block) is then multiplied, and the intensity of the remaining light is given as output from this block.
- Optical noise - Band-limited white noise is generated by this block, which is used to model the noise due to surrounding light, and the scattering of light by external sources apart from the finger. Band-limited white noise is used for this purpose since ambient light has components of light from the entire spectrum of the photodetector (visible and IR), which will contribute to noisy measurement. This is especially relevant for a non-contact geometry, as the distance between the probe and the finger increases scattering and external noises.
- Photodiode / Photodetector - The light intensity measured at the photodetector is modeled using this block, which takes in the light intensity reflected from the finger and that due to external light.
- Demultiplexer - The signal from the three different light sources are split by this block, which uses the output of the pulse generators for the same. The waveforms obtained after this step are visualized in Fig. 6a, for all three wavelengths.
- ADC - A zero-order hold is used to convert the continuous signals into discrete values.
- Buffer - A buffer is used to store the waveform, which is then sent to a MATLAB code for calculating the saturations.

The simulation was run and the oxygen and carbon monoxide saturations were calculated according to the algorithm described in Section III, for various values of SpO_2 (ranging from 70% to 100%) and $SpCO$ (ranging from 0% to 30%). Figure 6b shows the AC and DC signals in each of the three waveforms, and the results obtained are discussed in Section VI.

VI. RESULTS

Figure 7a shows the results obtained during various simulations as mentioned in Section V-B, where the true point and the calculated point are plotted on the $[HbO_2]$ vs $[HbCO]$ space. The errors in the calculated values in both $[HbO_2]$ and $[HbCO]$ are plotted in Fig. 7b. It was observed that the maximum errors in calculation of $[HbO_2]$ and $[HbCO]$ are 0.36% and 0.66% respectively. Further, error in $[HbO_2]$ calculation are all negative, the error in calculation of $[HbCO]$ are more evenly distributed. For comparison, most commercial oximeters have an accuracy rate of around 1.8% [27].

These errors are due to the optical noise added during simulation. In this simulation, since the errors are negligible, there was no need to improve accuracy using other analytical methods. However, bias correction can be performed to improve the accuracy of calculation if required. In a physical

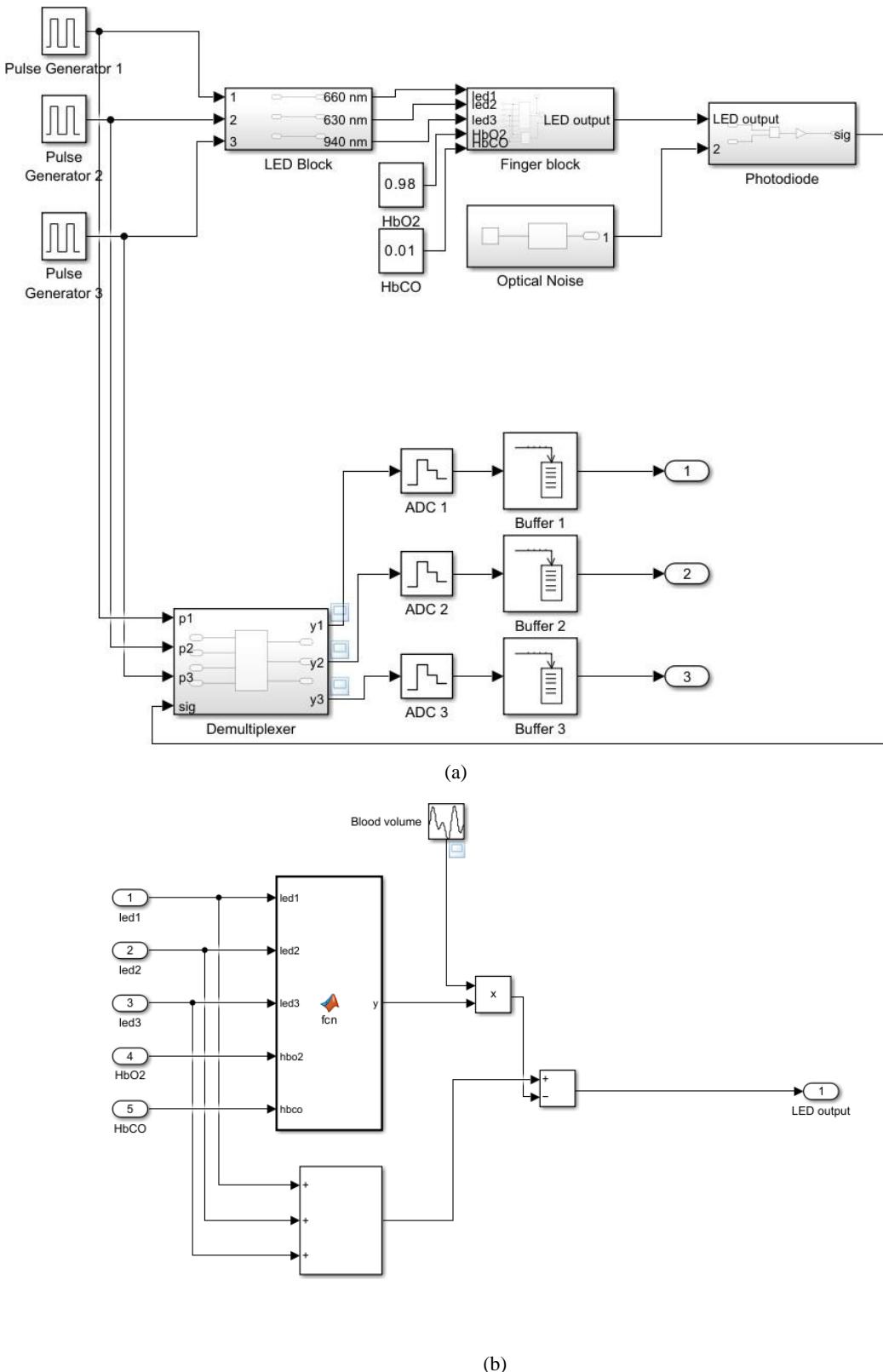


Fig. 5: Simulink model of the pulse oximeter (a) Complete oximeter setup, and (b) Finger block.

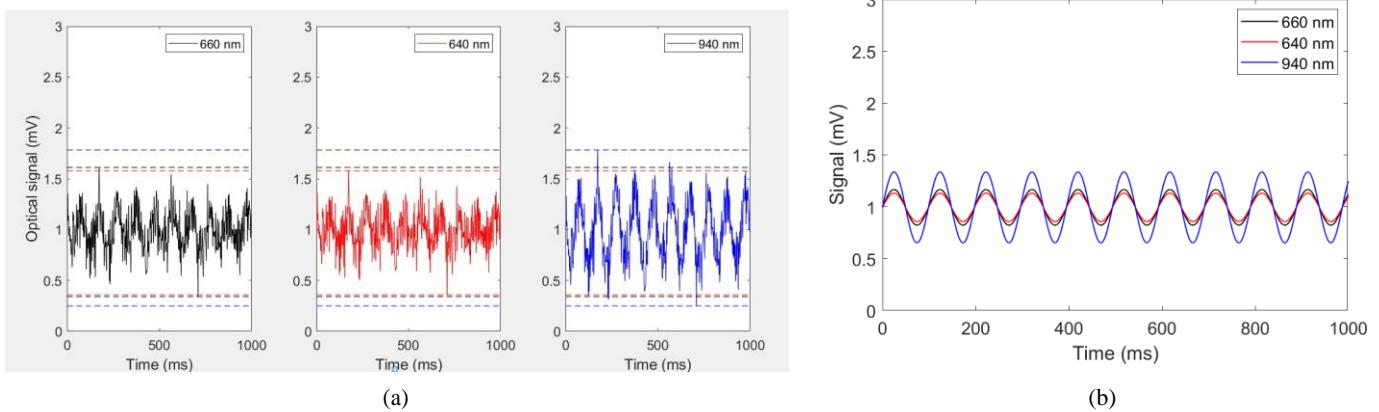


Fig. 6: Waveforms measured at different points in the simulation (a) After demultiplexing (the dotted lines show the difference in amplitudes measured at three wavelengths - 660 nm, 640 nm, and 940 nm), and (b) Waveforms extracted using FFT to calculate the AC and DC signals.

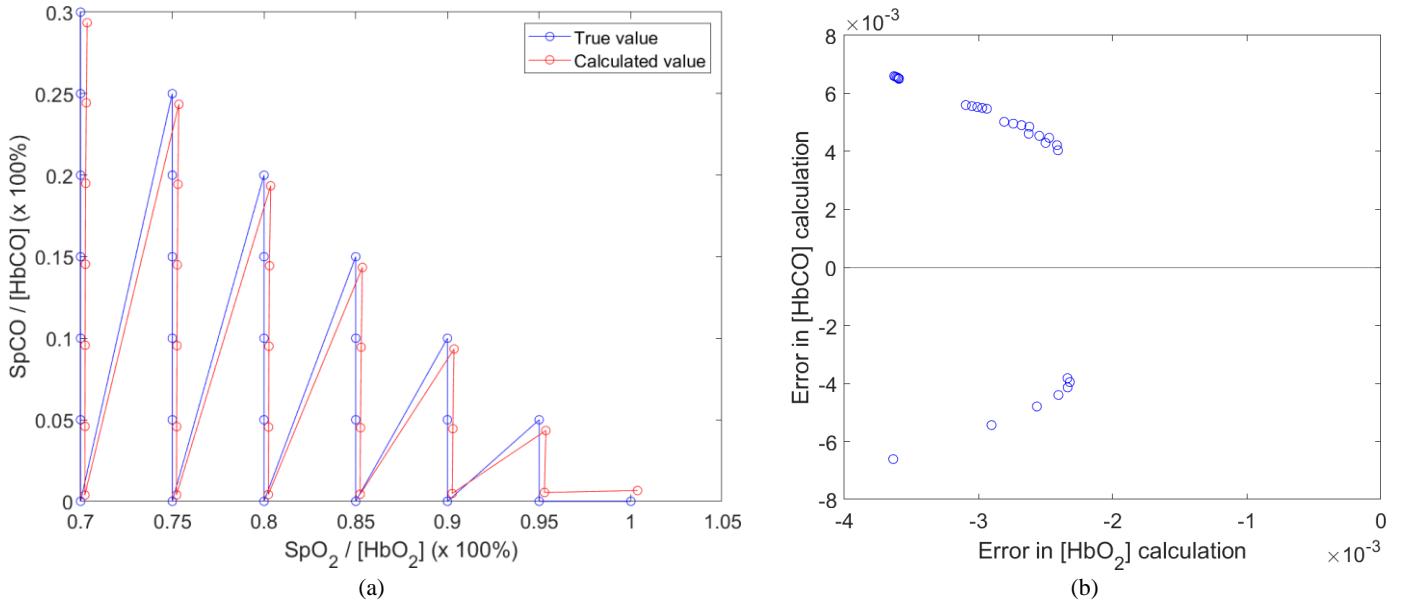


Fig. 7: Simulation results : (a) Various points ($[HbO_2]$ vs $[HbCO]$) simulated, with the true value and the calculated value shown, and (b) Error obtained during simulation

system, the probe can also be insulated such that optical noise from surrounding light sources is minimized.

VII. CONCLUSIONS

This paper proposes a novel algorithm and a system design for a non-contact reflectance pulse oximeter, which uses three wavelengths of light for measuring the arterial saturations of oxygen and carbon monoxide. The wavelengths have been selected after a thorough sensitivity analysis, and are 640 nm, 660 nm, and 940 nm. A simulation has also been performed on Simulink to evaluate the proposed algorithm, which gives satisfactory results. The errors obtained between the true and calculated saturation values in all simulations are small enough that they are of little clinical significance. Thus, the proposed algorithm is a reliable method of simultaneously measuring

oxygen and carbon monoxide saturations in arterial blood. This is a preliminary study based on simulation which requires further verification using in-vivo experiments on animals or humans.

The advantages of this proposed design are the ease of measurement, wherein a handheld device can be used to measure vitals in a non-contact manner. In addition, the ability to measure oxygen and carbon monoxide saturation simultaneously will allow the device to be used in critical situations such as on burn victims. The major considerations while designing the device is that since the light source is placed away from the skin, there might be interference with any ambient light. This can be prevented by using an insulating shield around the photodetector. In addition, since the device

is not in contact with the patient, motion artifact will be very high. This can be prevented by placing the patient's hand on a fixed mount, and also mounting the oximeter to prevent movement.

VIII. ACKNOWLEDGEMENTS

The authors would like to thank the Young Research Fellow Program, IIT Madras for the opportunity to pursue this research, and the IIT Madras alumni of 1979 for conceptualizing the same and providing the necessary funding. We are also grateful to Dr. R. Gowri Shankar, M.B.B.S., of the IIT Madras Hospital for his valuable inputs.

REFERENCES

- [1] G. Baura, "Pulse Oximeters", in *Medical Device Technologies: A Systems Based Overview Using Engineering Standards*, 2nd ed., Waltham, MA, USA: Elsevier, 2012, ch. 11, pp. 237-255.
- [2] A. A. Alian and K. H. Shelley, "Photoplethysmography: Analysis of the Pulse Oximeter Waveform," in *Monitoring Technologies in Acute Care Environments: A Comprehensive Guide to Patient Monitoring Technology*, New York, NY: Springer New York, 2013, pp. 165-178.
- [3] P. H. Charlton et al., "Breathing Rate Estimation From the Electrocardiogram and Photoplethysmogram: A Review," *IEEE Rev. Biomed Eng.*, vol. 11, pp. 2-20, Oct 2018.
- [4] F. H. Martini and E. F. Bartholomev, "Essentials of Anatomy and Physiology", 8th ed., Hoboken, New Jersey: Pearson Education, 2017.
- [5] Q. Cai, J. Sun, L. Xia and X. Zhao, "Implementation of a wireless pulse oximeter based on wrist band sensor," in *2010 3rd International Conference on Biomedical Engineering and Informatics*, Yantai, China, 2010, pp. 1897-1900.
- [6] K. A. et al., "Reflectance Pulse Oximetry for Blood Oxygen Saturation Measurement from Diverse Locations-A Preliminary Analysis," in *2018 IEEE International Symposium on Medical Measurements and Applications (MeMeA)*, Rome, Italy, 2018, pp. 1-6.
- [7] V. Hartmann et al., "Quantitative comparison of photoplethysmographic waveform characteristics: Effect of measurement site," in *Front. Physiol.*, vol. 10, pp. 1-8, Mar 2019.
- [8] S. Seifi et al., "Accuracy of pulse oximetry in detection of oxygen saturation in patients admitted to the intensive care unit of heart surgery: Comparison of finger, toe, forehead and earlobe probes," in *BMC Nurs.*, vol. 17, no. 1, pp. 1-7, Apr 2018.
- [9] H. Lee et al., "Reflectance pulse oximetry: Practical issues and limitations," in *ICT Express*, vol. 2, no. 4, pp. 195-198, Dec 2016.
- [10] A. R. Guazzi et al., "Non-contact measurement of oxygen saturation with an RGB camera," in *Biomed. Opt. Express*, vol. 6, no. 9, pp. 3320-3338, Aug 2015.
- [11] M. Villarroel et al., "Non-contact vital-sign monitoring of patients undergoing haemodialysis treatment," in *Sci. Rep.*, vol. 10, no. 1, pp. 1-21, Oct 2020.
- [12] E. D. Chan et al., "Pulse oximetry: Understanding its basic principles facilitates appreciation of its limitations," in *Respir. Med.*, vol. 107, no. 6, pp. 789-799, Mar 2013.
- [13] Center for Disease Control and Prevention, *Carbon Monoxide (CO) Poisoning Prevention* [Online]. Available: <https://www.cdc.gov/nceh/features/coposicing/index.html>.
- [14] C. Mattiuzzi and G. Lippi, "Worldwide epidemiology of carbon monoxide poisoning," in *Hum. Exp. Toxicol.*, vol. 39, no. 4, pp. 387-392, Dec 2019.
- [15] D.A. Colguhoun et al., "Ability of the Masimo pulse CO-Oximeter to detect changes in hemoglobin," in *J. Clin. Monit. Comput.*, vol. 26, no. 2, pp. 69-73, Apr 2012.
- [16] D. R. Hess, "Pulse oximetry: Beyond SpO₂," in *Respir. Care*, vol. 61, no. 12, pp. 1671-1680, Dec 2016.
- [17] S. Vasudevan, "Simultaneous measurement of oxygen and carbon monoxide saturation using pulse oximeters," Masters thesis, Dept. Biomed. Eng., Worcester Polytechnic Institute, Worcester, MA, 2011.
- [18] M. Shokouhian et al., "Simulink based behavioural modelling of a pulse oximeter for deployment in rapid development, prototyping and verification," in *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, San Diego, CA, USA, 2012, pp. 5566-5569.
- [19] C. Higgins (2005, October), *Causes and Clinical Significance of Increased Carboxyhemoglobin* [Online]. Available: <https://acute-care-testing.org/en/articles/causes-and-clinical-significance-of-increased-carboxyhemoglobin>.
- [20] C. Higgins (2006, October), *Methemoglobin* [Online]. Available: <https://acute-care-testing.org/en/articles/methemoglobin>.
- [21] Maxim Integrated, "Understanding Penetration Depth vs. Wavelength for Biosensor Applications," Application Note 6433 [Online], 2014. Available: <https://cdn.sparkfun.com/assets/9/7/b/f/e/AN6433.pdf>.
- [22] Natural Phenomena Simulation Group, University of Waterloo, Human Blood Data [Online]. Available: <http://www.npsg.uwaterloo.ca/data/blood.php>.
- [23] S. Prahl, Optical Absorption of Hemoglobin [Online]. Available: <https://omlc.org/spectra/hemoglobin/index.html>.
- [24] Lochner, C.M. et al., "All-organic optoelectronic sensor for pulse oximetry," in *Nat. Commun.*, vol. 5, no. 1, Dec 2014.
- [25] Texas Instruments, "OPT101 Monolithic Photodiode and Single-Supply Transimpedance Amplifier," OPT101 datasheet, Jan. 1994 [Revised Jun. 2015].
- [26] L. Hernández-Quintanar et al., "Fiber-optic pulse oximeter for local oxygen saturation determination using a Monte Carlo multi-layer model for calibration" in *Comput. Methods Programs Biomed.*, vol. 187, Apr 2020.
- [27] J. C. da Costa et al., "Research: Comparison of the Accuracy of a Pocket versus Standard Pulse Oximeter" in *Biomed. Instrum. Technol.*, vol. 50, no. 3, pp. 190-193, Jun. 2016.

SMART CROP PROTECTION SYSTEM FROM ANIMALS USING AI

DHARANIPRIYA K

*Electronics and Communication
Engineering*
*M.Kumarasamy College of
Engineering Karur,India*
dharanipriyak.ece@mkce.ac.in

SATHYAGEETHA S

*Electronics and Communication
Engineering*
*M.Kumarasamy College of
Engineering Karur,India*
sathyasivaram05@gmail.com

SOWMIA K S

*Electronics and Communication
Engineering*
*M.Kumarasamy College of
Engineering Karur,India*
sowmia0602@gmail.com

SRINIDHI J

Electronics and Communication Engineering
M.Kumarasamy College of Engineering
Karur,India srnidhijks@gmail.com

Abstract--*Animal assaults on crops are one of the main risks to crop production reduction. Crop raiding is one of the most acrimonious disputes as farmed land encroaches on formerly uninhabited areas. Pests, natural disasters, and animal damage pose severe risks to Indian farmers, lowering productivity. Farmers' traditional tactics are ineffective, and it is not practical to hire guards to watch over crops and keep animals away. Since animal and human safety are equally important, it is crucial to safeguard the crops from harm brought on by creatures without causing any harm, as well as divert the animal. As a result, we employ deep learning to recognize animals that visit our farm utilizing the deep neural network idea, a branch of computer vision, in order to overcome the aforementioned issues and achieve our goal. In this project, we will periodically check in on the entire farm using a camera that will continuously record its surroundings. We are able to recognize when animals are entering with the use of a deep learning model, and then we use an SD card and speaker to play the right sounds to scare them away. The many convolutional neural network libraries and principles that were used to build the model are described in this research.*

Keywords--*Crop raiding, Safeguard, Camera, SD card, Speaker, Sound.*

I. INTRODUCTION

Agriculture, which also generates a range of industrial raw materials, provides for the food needs of the populace. Animal interference in agricultural areas

results in a significant loss of crops. Animal raiding damage to crops has grown significantly in recent years. Animals like the buffalo, the cow, and goat may be very destructive and occasionally even result in human casualties. In villages, potato and wheat crop yields have suffered significant overall losses. Due to the stringent rules, small farmers lose up to 40 to 50 percent of their produce to animals yet are unable to take any drastic actions. In India in particular, elephants are a class that is particularly prone to conflict, which has led to an increase in human-animal conflict. Therefore, a mechanism that can assist the farmers in driving these animals away as soon as they become aware of their intrusion is required[9].

Farmers that want to try this technique won't have to spend a lot of money because the current trend is toward strong crop returns with minimal investment. Each planting season begins with the expectation that the harvest will be good. Natural catastrophes including cyclones, floods, droughts, and other recurring events pose a major threat to farmers in the state. In addition to this, unexpected pest attacks and crop damage from domestic animals during harvest season results in significantly decreased or occasionally lost yields. Every harvest season, animals like buffalo, cows, and goats, among others, significantly damage surrounding crop fields in forested and hilly terrain.

Government representatives and farmers are unable to resolve this problem, and they are only able to attempt to drive these creatures back into forested areas by beating drums and exploding firecrackers, among other methods. Researchers and planners from all around the world are looking for alternate solutions to the farmer dilemma. A farmer who lives close to a forest

or a mountainous area worries about whether he will be able to feed his family with a strong crop production. While there are methods for protecting crops from animals, such as wire fences and electric fences, their prices are greater and are dependent on the size of the area. Additionally, it is only allowed to protect animal species that are in risk of extinction.

Electric fences occasionally endanger human life and are frequently positioned far from any potential human contact. Consequently, it is crucial to use any electrical automation that won't endanger the environment or put people or other animals at risk. Here, we are using artificial intelligence to detect animal encroachment in agricultural regions as part of our study. We employ sound waves to divert animals away from agricultural land without endangering both animals and people. We utilized an Arduino Uno microcontroller, an SD card, and an amplifier in our project to play sounds that would scare away animals.

If an animal is seen in the camera, the microcontroller will receive the data. Then compare the output data to the sounds. If an animal is found, audio will play and data can be shown on the LCD[7].

Passive Infrared (PIR) sensor is positioned so as to provide larger range of detection. The hardware of the new system is described after a brief introduction of an existing system. The goals of the system and the rationale behind the implementation of various components are described in the objectives and scope.

A. OBJECTIVES

- To study Image processing techniques and apply an algorithm for convolutional neural network.
- To design an automatic system to control animal crop detection using image processing.
- To offer a monitoring and deterrent system for animal crop attack using deep learning.

B. ADVANTAGES

1. Less Human Error Rates

One of artificial intelligence's primary advantages is its capacity to significantly reduce mistakes and enhance accuracy and precision. Each decision made by artificial intelligence is based on data that has already been gathered and a certain set of algorithms. These mistakes may be fully removed when correctly coded.

2. Zero Risks

By giving specific activities to AI robots, individuals may avoid various risks, which is another important advantage of AI. Metal-bodied machines are tough and able to endure harsh conditions, making them perfect for space travel, bomb defusing, and exploring the deepest parts of seas. They may also provide accurate work with more responsibility and durability.

3. 24x7 Availability

People only spend on average three to four hours a day working productively, according to numerous studies. People need breaks and vacation time in order to balance their personal and professional life. Contrarily, AI may operate continuously without rest periods. They multitask far more quickly than humans can, and they can reason quickly and precisely. With the aid of AI algorithms, they can even carry out difficult repetitive tasks with ease.

4. Digital Support

Nowadays, virtually all large companies engage with their clients via digital assistants, which reduces the need for on-site personnel. You may directly ask a Chabot what you need by conversing with one. Given how far chatbot technology has come, it may be difficult to tell if you are corresponding with a human or a Chabot.

5. Fresh Innovations

Practically every sector has benefited from the creation of novel ideas using AI to solve complex problems. A breakthrough discovery has helped physicians detect the early stages of breast cancer in women using cutting-edge AI-based technologies.

6. Unbiased Conclusions

Humans are guided by their emotions whether we like it or not. AI, in comparison, is emotionless and approaches problems in a very pragmatic and logical way. The elimination of prejudice is a key benefit of artificial intelligence that ensures more accurate decision-making.

C. SCOPE

1. To create a security system that would protect the safety of the farm
2. Limit how animals can enter the field.
3. Send us a warning using the GSM module.
4. Create a system that, when an animal tries to access the farm, activates solar animal repellents.
5. Streak light will shine on that side .

II. ANIMAL DETECTION USING DEEP LEARNING ALGORITHM

This proposed work develops an algorithm to detect the animals in wild life. Since there are many different animals manually identifying them can be a difficult task. This algorithm classifies animals based on their images so we can monitor them more efficiently. Animal detection and classification can help to prevent animal-vehicle accidents, trace animals and prevent theft. This can be achieved by applying effective deep learning algorithms[6].

III. RELATED WORKS

Utilizing two key algorithms: an algorithm for object recognition using artificial neural networks and an algorithm for motion detection using sensors. Animal movement near highways can be detected using a PIR sensor, and objects can be with the aid of ANN. When a motion is discovered, the object identification system determines if it was caused by a moving animal or by some other element. Following the detection of the motion, the object identification algorithm examines it to determine if the motion was brought on by an animal movement or by any other factors. If this happens as a result of an animal movement, alerts are sent over MQTT to an Android application that uses Google Maps to display them on the relevant area.

In a farm, an algorithm is utilized to identify the animals. The variety of animals makes it challenging to manually identify each one. In order to better effectively monitor animals, this system categorizes them based on their photographs. Animal tracing, theft prevention, and animal-vehicle accident prevention can all be aided by animal detection and classification[5]. Applying efficient deep learning techniques can help with this. In order to find animals in a given image, an algorithm is employed. A more accurate feature vector

with gradients oriented toward the histogram is called WCoHOG. Co-occurrence Histograms of Oriented Gradients is an extension of this (CoHOG). For high dimensional data, LIBLINEAR classifier is employed in order to improve accuracy[8]. Two benchmark datasets the animal and Camera Trap datasets were used for tests. Experimental results demonstrate that W-CoHOG outperform current state-of-the-art techniques.

The classification of animal behaviors from films captured by a fixed-point camera is done using a method. Animal detection and tracking are required in order to categorize animal behavior. Traditional methods for identifying moving objects rely on backdrop and frame removal. Since indoor animals are vulnerable to sunlight and shadow, conventional approaches are not ideal for their detection. We provide a technique for following animals and categorizing their activity using DeepLabCut's skeletal data. The results of the experiments demonstrate that the proposed method is better than the traditional way[10].

Even still, human-shaped scarecrows are used in fields of crops to keep birds and other animals away from growing crops and from disrupting and devouring them. Because of the many problems with these ideas, increasing agricultural security is currently a top priority. As a result, it focuses on recommending a system that can identify intruders, watch out for any harmful behavior, and then alert the system owner to it. It operates as an adaptable system that provides farmers with a practical means of ensuring the whole protection of their fields from any assaults or trespassing activities.

IV. SYSTEM DESIGN

PROPOSED SYSTEM

- The major objectives of our initiative are to prevent animal damage to crops and to divert animals safely.
- A system for detecting animals is intended to alert users of their presence. We tracked the animal's movement using a PIR sensor.
- As part of this project, we'll use a camera to periodically check in on the entire farm while it records its surroundings all day. Using a deep learning model, we detect the presence of animals

and then play the appropriate noises to scare them away.

- The paper describes many frameworks and convolutional neural network ideas that were applied to create the model.
- The PC will utilize a USB to UART converter to send the information to the microcontroller if an animal is discovered using the CNN method in an agricultural region.
- The microcontroller will read the file from the SD card, switch on the speaker, and direct the animal away from the agricultural field after getting the data from the PC.

A. ARCHITECTURE DIAGRAM

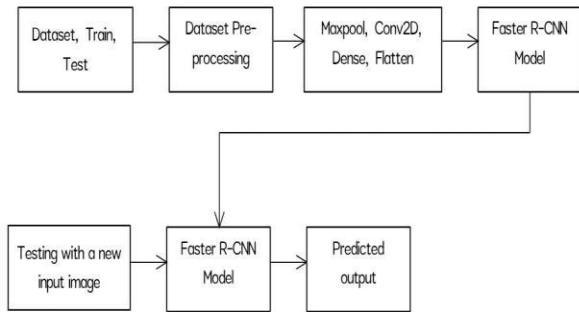


Fig 1 : Architecture Diagram

B. Proposed Block Diagram

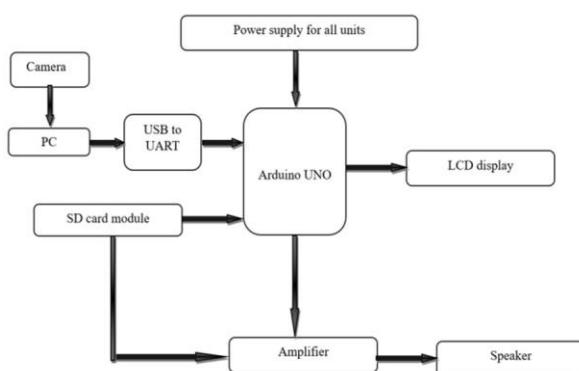


Fig 2 : Proposed Block Diagram

V. MODULES DESCRIPTION

- Image Dataset Collection
- Image Pre-processing
- Importing Modules
- Capturing the images of animals
- Camera Interfacing

1. IMAGE DATASET COLLECTION

For this assignment, we must collect each image of a buffalo, a cow, or goat that gives the impression of being an animal. This is the most important phase of the project. As a result, every image we see is derived from recent or archived CCTV footage. Following the receipt of the data, the following steps may be followed.

2. IMAGE PREPROCESSING

Pre-processing is needed after collecting all the photos. As a result, not all visuals can effectively transmit information. So that we may identify, rename, and prepare photographs by changing their sizes. We can train our deep learning model using the photographs once the process is complete.

3. IMPORTING MODULES

The next step is to import all of the necessary library files. Collections of functions and short execution codes are contained in library files. We can complete all the necessary processes of object, detection and picture processing with the help of these library files. In this project, we make use of crucial library files like Tensor Flow, opencv, keras, and others. These libraries will help to improve the effectiveness and adaptability of our deep learning model for processing real-time photos or videos.

4. CAPUTURING THE IMAGES OF ANIMALS

We need to collect every image that contains an animal crop for this project. This is the most important phase of the project. As a result, every image we see is derived from recent or archived CCTV footage. After receiving the three animal datasets, subsequent steps can be conducted.

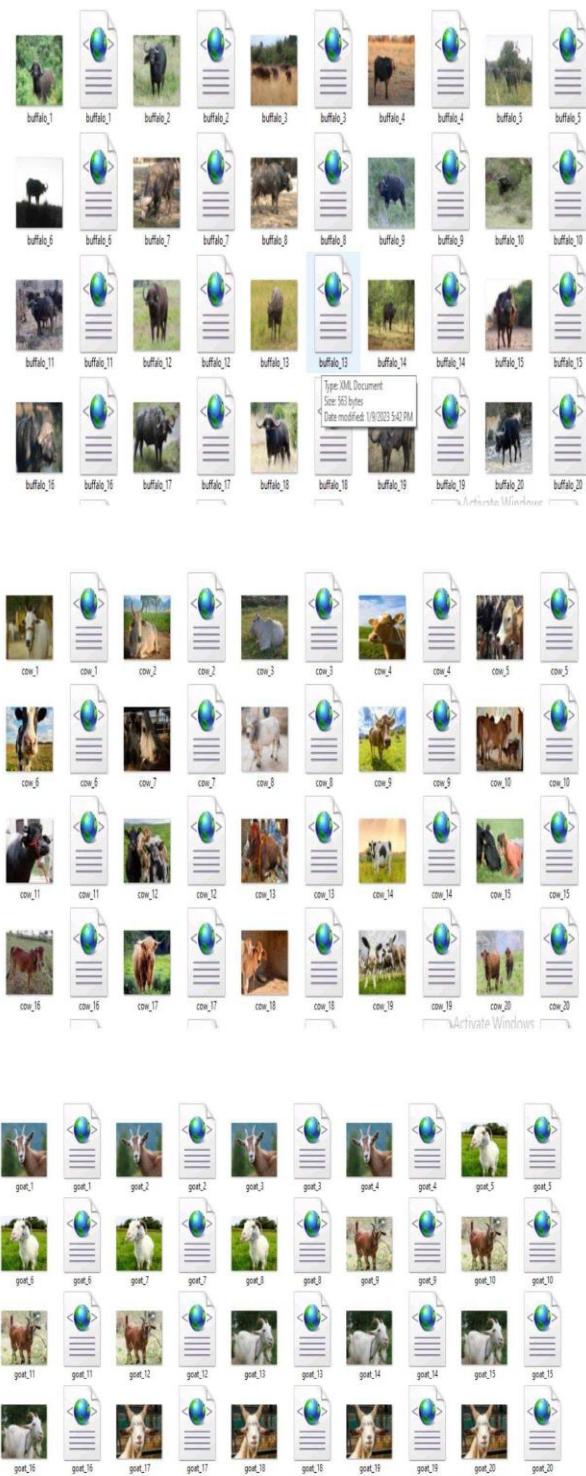


Fig 3 : Tested and Trained Images of Animals

TRAINING DATASET:

A large dataset of images that have been labeled with the appropriate class labels is input into CNN during training. The CNN network processes each image, assigning values at random, and then compares the results to the original image's class label.

5. CAMERA INTERFACING

One of the most important phases of photo processing is computer vision. Since the camera will make everything in the real environment visible to the computer, we must connect the camera to our deep learning model. A camera can operate any kind of media. If required, the captured animal can serve as the main source of proof.

VI. RESULT AND CONCLUSION

In agricultural fields next to forests and mountainous areas around the world, animals like buffalo, cows, and goats seriously harm the crops. Animals are responsible for 50% to 60%, and perhaps 100%, of crop damage. The villagers were harassed and began looking for alternatives to chase away animals out of the fields that were prepared for cropping the majority of their income and support came from their crops, which they desperately needed to save. Their crops may be saved using acoustic devices powered by Arduino. It was discovered that installing our equipment on the crop field would boost the farmers' financial gain. Additionally, the mechanism we have in place guarantees the safety of both crops and people, and it poses no threat to animals. In the present, agricultural vandalism caused by animals and fire has become a big societal problem. Since there hasn't been an efficient solution for this issue to yet, it needs serious consideration. As a result, its design has broad social application because it seeks to solve this issue. This design will aid farmers in protecting their fields and orchards, and persevere them for the protection of their lands and to prevent them from suffering big financial losses. Additionally, this will assist them in increasing crop yields, which will increase their profitability.



Fig 4 : Hardware Design

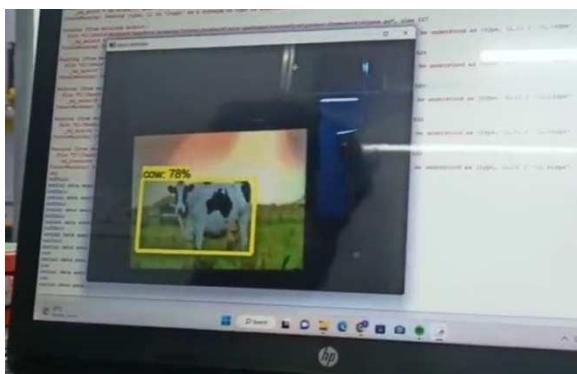


Fig 5 : Animal Detected

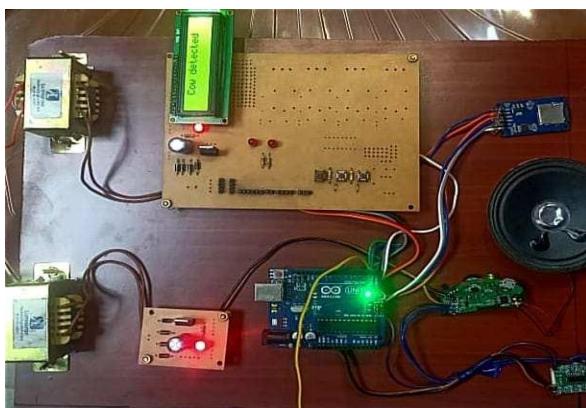


Fig 6 : Output of the Project

REFERENCES

- [1] Bayani, A., Tiwade, D., Dongre, A., Dongre, A. P., Phatak, R., & Watve, M. (2016). Assessment of crop damage by protected wild mammalian herbivores on the western boundary of Tadoba-Andhari Tiger Reserve (TATR), Central India. *PLoS one*, 11(4).
- [2] Rao, K. S., Maikhuri, R. K., Nautiyal, S., & Saxena, K. G. (2002). Crop damage and livestock depredation by wildlife: a case study from Nanda Devi Biosphere Reserve, India. *Journal of environmental management*, 66(3), 317-327.
- [3] Davide Adami; Mike O. Ojo; Stefano Giordano(2021).Design, Development and Evaluation of an Intelligent Animal Repelling System for Crop Protection Based on Embedded Edge-AI.
- [4] Bapat, V., Kale, P., Shinde, V., Neha, D., & Shaligram, A. (2017). WSN application for crop protection to divert animal intrusions in the agricultural land. *Comput. Electron. Agric.*, 133, 88-96.
- [5] Santosh Kumar, Sanjay Kumar Singh, Ravi Shankar Singh, Amit KumarSingh & Shrikant Tiwari (2016). Real-time recognition of cattle using animal biometrics.
- [6] S Meera ; R Sharmikha sre ; K Priyadarshin ; P.V. Varshitha; R. SaiCharitha(2022). Animal Detection Alert System.
- [7] Jinook Oh & W. Tecumseh Fitch (2016).CATOS (Computer Aided Training/Observing System): Automating animal observation and training.
- [8] M. Thenmozhi, M. Saravanan, K. Pradeep Mohan Kumar, S. Suseela & S. Deepan (2020). Improving the prediction rate of unusual behaviors of animal in a poultry using deep learning technique.
- [9] Patrick Baur(2022). When Farmers are Pulled in Too Many Directions: Comparing Institutional Drivers of Food Safety and Environmental Sustainability in California Agriculture.
- [10] Mahir Patel, Yiwen Gu, Lucas C. Carstensen, Michael E. Hasselmo & Margrit Betke(2022). Animal Pose Tracking: 3D Multimodal Dataset and Token-based Pose Optimization.
- [11] Kumar, N. M., Chopra, S. S., de Oliveira, A. K. V., Ahmed, H., Vaezi, S., Madukanya, U. E., & Castañón, J. M. (2020). Solar PV module technologies. In *Photovoltaic Solar Energy Conversion* (pp. 51-78). Academic Press.
- [12] Aghaei, M., Kumar, N. M., Eskandari, A., Ahmed, H., de Oliveira, A. K. V., & Chopra, S. S. (2020). Solar PV systems design and monitoring. In *Photovoltaic Solar Energy Conversion* (pp. 117145). Academic Press.

- [13] T. Mhammad, "Using Ultrasonic And Infrared Sensors For Distance Measurement" World Academy of Science, Engineering and Technology, pp. 293-298, 2009.
- doi:10.1109/ISSE.2007.4432828.
- [15] Senthil Kumar Kandasamy, S. Maheswaran, K. R. Kavitha, D. Karthikeyan, J. Indra, K. Sheikdavood, Y. Syamala, K. Srilakshmi, Chandrasekaran Arumugam, DhivyaBalamoorthy, "HighPerformance Graphene Oxide/Polypyrrole/Ziziphusjujuba/Prunus dulcis Ternary Composite Electrodes for Supercapacitor for Sensor Applications", Journal of Nanomaterials, vol. 2022, Article ID 9201910, 10 pages, 2022. <https://doi.org/10.1155/2022/9201910>
- [16] Sheikdavood, K., P. Surendar, and A. Manikandan. "Certain Investigation on Latent Fingerprint Improvement through MultiScale Patch Based Sparse Representation." *Indian Journal of Engineering* 13.31 (2016): 59-64.[16] K. Sheikdavood, K. Prabhu, S. Allirani. "PerformanceEvaluationon Category Classification of an Images Using Convnet". 2021. *Annals of the Romanian Society for Cell Biology*, June 2021, pp. 22028, <https://www.annalsofrscb.ro/index.php/journal/article/view/1661>
- [17] C.Vivek, S.PalanivelRajan, "Z-TCAM : An Efficient Memory Architecture Based TCAM", Asian Journal of Information Technology, ISSN No.: 1682-3915, Vol. No.: 15, Issue : 3, pp. 448454, 2016.
- [14] Discant, A. Rogozan, C. Rusu and A. Benshair, "Sensors For Obstacle Detection" 2007 30th International Spring Seminar on Electronics Technology (ISSE), Cluj-Na- poca, 2007, pp. 100- 105.
- [18] S.PalanivelRajan, K.ShekDavood, "Performance Evaluation on Automatic Follicles Detection in the Ovary", International Journal of Applied Engineering Research, ISSN No.: 0973-4562, Vol. 10, Issue 55, pp. 1-5, 2015.
- [19] S.PalanivelRajan, T.Dinesh, "Statistical Investigation of EEG Based Abnormal Fatigue Detection Using LabVIEW", International Journal of Applied Engineering Research, ISSN: 0973-4562, Vol. 10, Issue 43, pp.30426-30431, 2015.
- [20] S.PalanivelRajan, M.Paranthaman, "Characterization of Compact and Efficient Patch Antenna with single inset feeding technique for Wireless Applications", Journal of Applied Research and Technology, ISSN: 1665–6423, Vol. 17, Issue 4, pp. 297-301, 2019.
- [21] S.PalanivelRajan, et.al., "Performance Evaluation of Mobile Phone Radiation Minimization through Characteristic Impedance Measurement for Health-Care Applications", IEEE Digital Library Xplore, ISBN : 978-1-4673-2047-4, IEEE Catalog Number: CFP1221T-CDR, 2012.
- [22] S.PalanivelRajan, R.Sukanesh, S.Vijayprasath, "Analysis and Effective Implementation of Mobile Based Tele-Alert System for Enhancing Remote Health-Care Scenario", HealthMED Journal, ISSN No. : 1840-2291, Vol. No. 6, Issue 7, pp. 2370–2377, 2012.

DIGITAL STETHOSCOPE FOR INSTANT MONITORING FOR CARDIAC AUSCULTATION

Chitra R

Electronics and communication engineering
Sri Sairam Engineering college
Chennai,TamilNadu,India
chitra.ece@sairam.edu.in

Jayapreetha N

Electronics and communication engineering
Sri Sairam Engineering college
Chennai,TamilNadu,India
sec19ec132@sairamtap.edu.in

Swetha D

Electronics and communication engineering
Sri Sairam Engineering college
Chennai,TamilNadu,India
sec19ec020@sairamtap.edu.in

Swetha S

Electronics and communication engineering
Sri Sairam Engineering college
Chennai,TamilNadu,India
sec19ec014@sairamtap.edu.in

Abstract— The main problem that often arises for doctor is using stethoscope to detect lung sounds. The analysis of lung sound obtained by a stethoscope is challenging when the signal level is extremely low. As a consequence, the current acoustic stethoscope has to be replaced by a digital electronic stethoscope. The primary goal of this study is to design a digital stethoscope that monitor lungs sounds and identify potential illnesses using CNN model . The notification of the detected disease is sent by pushover application and graphical representation is shown in the THINKSPEAK application. Based on the analysis of the CNN model, the lung disease detection method achieved an average accuracy of 95%, which means it could be applied to diagnosis of lung disease in the real world.

Keywords—*Lung sound, digital stethoscope, CNN model, THINKSPEAK, Pushover*

I. INTRODUCTION

A major cause of death worldwide is respiratory illness, including pneumonia and upper respiratory tract infections [1]. A prompt and accurate diagnosis is essential for effective care, medical supervision, and the avoidance of a failing respiratory system. In 45% of the states, less than 1 doctor per 1000 people is present, according to World Health Organization statistics [2]. A simple respiratory diagnostic equipment that does not require any specific training to use is critical in cities with limited access to medical care. In addition, this method reduces the effort of the doctor and allows continuous monitoring of the patient's disease. Different types of digital stethoscopes have started to appear on the market recently, taking the place of traditional stethoscopes and the information must be processed for further process. Auscultation is one of the techniques used to measure lungs sounds. Techniques reveals the details regarding a range of biological functions resulting from heart and lung diseases So, only a trained medical professional who is proficient in breath sounds is likely to provide a correct diagnosis after listening to the heart beat sounds.

Lung volume can be expressed as heartbeats per minute. The

auscultation method is one of the techniques used to measure these lungs. These techniques reveal details regarding a range of bodily functions resulting from heart, lung diseases. A number of variables, such as the existence of background noise or even other sensory stimulation, may have an impact on one's ability to recognize heartbeat. As a result, only a qualified doctor who is adept in breath sounds is likely to provide a correct diagnosis after listening to the lung. The proposed system uses Electrodermal Activity sensor to get the lung signal as input and then detects various lung disease using CNN model and display the graphical representation in THINKSPEAK application. Additionally, a notification with detected disease is sent by pushover application. By the above proposed solution anyone can detect various lung disease.

II. LITERATURE REVIEW

The research of N. Jatupaiboon et al. intends to create and d a cost-efficient electronic stethoscope model with real-time digital signal processing, frequency filtering, and noise reduction capabilities. The system's disadvantage is that the signal is transmitted via wired communication [3].

Nipun Katyal et al idea is to use a microphone to record the heart sound, subject it to analogue filtering, and then store the result in a computer. The result enables to extract features from the input signal and store them for later use [4].

B. Malik et al discussed the construction of an electronic stethoscope that is utilized to listen to pulmonary and cardiac sounds. The noises are recorded from the skin's surface and then filtered and amplified so that the signal level exceeds the threshold for audibility. A loudspeaker is used to clearly hear the signals that have been processed. [5].

Y. -C. Wu et al. proposed using a Raspberry Pi 3B to create an electronic stethoscope by adding a single mike into the regular stethoscope record lung and heart audio. Amplification and filtering are performed on the acquired signals before the audio signals are transmitted to the Raspberry Pi via audio lines [6].

M Ramesha et al proposed work uses an electronic Stethoscope to analyze the acoustic features of cardiac sounds. The development and implementation of an IoT-based electronic stethoscope. The sensor data is sent into the microcontroller, which converts the analog impulses into digital signals that are subsequently transferred to the computer through bluetooth. The MATLAB program in the computer is used to analyze the signals which calculates and displays the heart rate [7].

Yi Ma et al research is based on LungSys, a system that consists of an android tablet computer running software and a commercial digital stethoscope. The digital stethoscope uses a built-in Bluetooth device to communicate electrical signals from an acoustic sound to a mobile tablet. The proprietary tablet software application analyses lung sounds in real time using the bi-ResNet(BRN) neural network model and notifies users of any unexpected respiratory sounds [8].

Yongpeng Liu et al. is to create a system for remote telemedicine and self-screening. A stethoscope was used as a tool to measure and analyze respiratory sounds with a proprietary device. [9].

Georgios Petmezas et al developed a solution to overcome the imbalance in training data, proposing a novel hybrid neural model that employs the focal loss (FL) function. A convolutional neural network (CNN) is used to collect features from short-time Fourier transform (STFT) spectrograms before being used by a long short-term memory (LSTM) network to identify four different lung sounds. Using this model's application to categorise the disorders, an accuracy of 74.57% was attained. [10].

Aykanat et al employ two types of machine learning techniques: MFCC characteristics in a support vector machine (SVM) and spectrogram images in a convolutional neural network (CNN). It used to overcome the outcome of the Svm classifier to analyze with CNN algorithm since employing A popular audio classification technique combines MFCC characteristics with an SVM algorithm.. For classify pulmonary sounds, created four sets of data for every CNN and Svm classifier. The results was accuaracy of 86% obtained for this model [11].

III. MODULES

A) NodeMCU

A NodeMCU device is similar to an Arduino. ESP8266 is the key element of it. The pins of this device can be programmed. NodeMCU has WiFi built-in and with the help of micro-USB port, the device can receive electricity. The price is reasonable. There are various programming environments available for it. The operating voltage for the ESP8266 NodeMCU is between 2.5 and 3.6 volts, the operating current is between 80 and 600 milliamperes, and the sleep mode current is 20 amperes. The on-board Micro USB port provides power to the ESP8266 NodeMCU. The ESP8266 NodeMCU has 200 Kb of flash memory, 80 Kb of DRAM, and 32 Kb of RAM. Digital pins D0 to D10, PWM pins 12 and an analogue pin A0 are all present in the ESP8266 NodeMCU. 5 Ground Pins, 3 3.3 V Pins, and 1 Vin Pin are provided, which allows to connect an external +5V supply that is not connected to the USB. There are a total of 17

GPIO pins on the ESP8266 NodeMCU. The pins can be utilized for peripheral functions such as two UART interfaces and a 10-bit ADC channel. The NodeMCU is shown in Fig. 1.

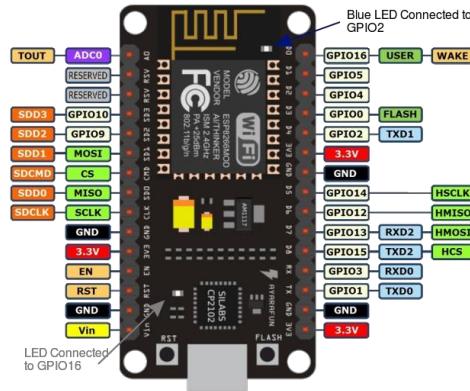


Fig. 1. NodeMCU

B) Power Supply

Power supply is another name for a source of electrical power. A power supply unit, also referred to as a PSU, is a device or a system that supplies a load or group of loads with electrical or other types of energy. Electrical energy sources are more commonly used than mechanical energy sources. Fig. 2 depicts the power supply.

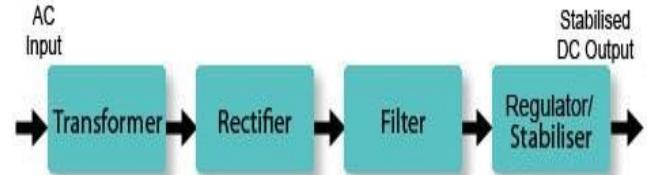


Fig. 2 Power Supply

C) THINGSPEAK Ascii

THINGSPEAK is a matrix laboratory that combines a fourth-generation programming language with a numerical computation environment . THINGSPEAK allows the user to handle matrices, display functions and data, construct algorithms and connect with programming languages such as C, C++, Java, and Fortran. There are data structure types in THINGSPEAK. Application development with graphical user interface features is supported by THINGSPEAK. GUIDE (GUI development environment) is a feature of THINGSPEAK for creating GUIs graphically. Moreover, graph-plotting features are firmly interwoven. It is a technical computing language with great performance. Only 10 characters, including description numbers, are supported by the THINGSPEAK user interface in ASCII. The NodeMCU Data Input signal will be converted into ASCII code at THINGSPEAK using this ASCII character for decimal number.

D) Electrodermal Activity (EDA):

The electrical properties of the skin, which alter, can be precisely measured using the bio signals and EDA sensor. The

changes are occurred due to altered sympathetic nervous system activity that alters sweat gland activity and sweat production. Due to the low-noise signal filtering and amplification circuit design, even the weakest electrodermal skin reaction events can be recognized with the highest performance. Fig. 3 displays an EDA sensor.



Fig. 3. EDA Sensor

E) LM358 Dual Op-Amp

Two operational amplifiers are powered by a single power source in the dual operational amplifier integrated circuit (IC) LM358. It comprises of two separate compensated operational amplifiers with high gain frequency and low power.

A single source of power may operate the LM358 over a broad variety of voltages. It is better suited for applications needing low voltage level AC and moderate voltage level DC. The LM358 is frequently used in practical applications including transducer amplifier and traditional op-amp circuit design because of its inexpensive price and small size.. Up to 20 mA per channel may be supplied by the LM358 IC, which can function between 3V and 32V DC. Fig. 4 shows the LM358 Dual Op Amp.

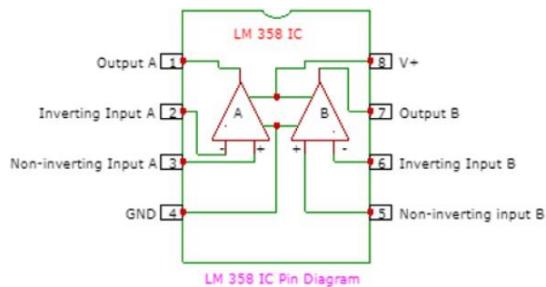


Fig. 4. Pin diagram of LM358

F) PushOver

Pushover is a service offering for companies sending messages to numerous users that includes a number of added capabilities such as user management and failover SMS message delivery. For all communication in Pushover, by using industry-standard TLS (HTTPS) encryption at every stage of the process, included between your servers and our API servers, those push servers to your devices, and our apps back to our servers.

G) Jupyter Notebook

JupyterLab is a highly modern web-based platform for programming, database, and notebook collaboration. The

adaptability of the interface allows users to design and arrange machine learning, predictive learning, and deep learning processes. A modular design invites additions that boost and enhance usefulness. Jupyter Notebook is used for CNN model implementation.

IV METHODOLOGY

The Electrodermal Activity (EDA) is a sensor that transmits the signal through Wi-Fi Pro in order to record the weak lung activity signal. Due to the fact that the lung signal are acoustic impulses with voltage ranging from 1 mV to 1.5 mV. To increase frequencies of the acoustic signal LM358 Dual Op-Amplifier is used in this study. NodeMCU will process the lungs signal and the result will be the detected disease such as Healthy Lung, Chronic Obstructive Pulmonary Disease, Pneumonia (often associated symptom with COVID19) and Upper Respiratory Tract Infection using the Convolutional neural network (CNN) Model. If the Voltage range from 70 to 90 Volt ,it is detected as pneumonia if Voltage range from 90 to 100 Volt then it is detected as COPD and for URTI values ranges from 100 to 120 V .THINKSPEAK has various fields for disease such as healthy, pneumonia, COPD and URTI . If the particular disease is detected , NodeMCU will send the value 1 to the specified field in THINKSPEAK software using the in-built Wi-Fi module. Hence THINKSPEAK will display a straight line in the disease field. Additionally, a notification is sent using pushover application. Fig. 5 depicts the block diagram of the digital stethoscope.

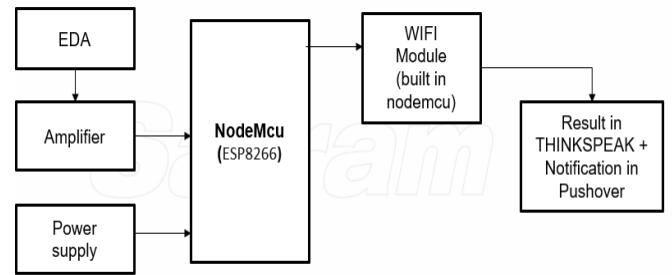


Fig. 5. Block Diagram of the Digital stethoscope

V. CNN Model Implementation

A. Dataset Information

The dataset utilised in this study was originally collected to assist the research challenge held in conjunction with the International Conference on Biomedical Health Informatics, or ICBHI 2017. The data set contain audio file came from 126 different patients which has 920 audio files with annotations, ranging in duration from 10 to 90 seconds. The recordings, which last for 5.5 hours and 6898 respiratory cycles, include 886 wheezes, 1864 crackles, and 506 wheezes and crackle. Patients of different ages are present including young people, adults, and senior citizens. The respiratory sounds in the dataset

are divided into various lung disease such as healthy, Chronic obstructive pulmonary disease (COPD), pneumonia and Upper respiratory tract infection (URTI). In this case, the term "healthy" refers to the respiratory sound recorded from a healthy individual, whereas others relate to the sound captured from a sick patient.

B. Preprocessing:

The dataset consists of annotated files which has respiratory cycle timings which is used to extract audio file segments that represent a patient's breathing cycle timing. Additionally, the length of each respiratory cycle varies from 10 to 90 seconds, with some of these variations being outliers. The Soundfile module is used to write to the output path of freshly produced audio files and the Librosa module to import audio files in order to trim all respiratory cycles to this mean duration.

C. Balancing: After preprocessing, the dataset will be biased towards one or more classes; various approaches should be used to correct the imbalance. The test and train datasets in this study are balanced on an equal distribution of classes using the stratify feature of the sklearn(machine learning library)train test split.

D. Feature Extraction :

The use of third-party libraries is quite beneficial for feature extraction because it may be challenging to extract features from audio samples .The three most essential features : Mel-frequency cepstral coefficients (MFCCs), CHROMA STFT, and MEL SPECTROGRAM, are extracted using Librosa.

VI. CNN MODEL ARCHITECTURE

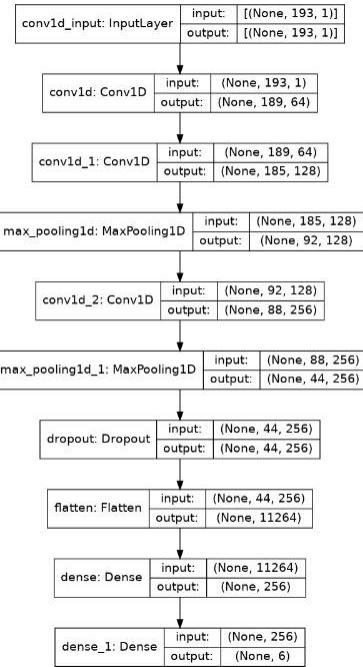


Fig. 6. CNN model architecture

Fig. 6 depicts the implementation of the CNN model architecture in this study. The spectrogram characteristics are

applied to the classifier for further categorization after feature extraction. Based on the MFCC features derived from the audio data, the classifier employed CNN to distinguish between sounds related to pneumonia, chronic obstructive pulmonary disease, the healthy group, and upper respiratory tract infection group. The classifier is composed of many convolution and max-pooling layers, followed by activation and fully linked layers. a sequential model that outputs the dense layer.

The convolutional layers are used to identify features. Multiply the input by a matrix as algorithms work, save the output in a feature map, and then slide a filter window over it. The forward pass involves convolution of the filters between the height and breadth of the inputs. The result is a 1D block made up of the height and breadth dot products. Another component of CNN is the pooling layer. In order to decrease the number of parameters, it shrinks the representation's spatial size. There are two forms of pooling: average pooling and maximum pooling. A dropout layer is employed to avoid overfitting. To introduce nonlinearity from the input to the output, each convolutional layer employs the Rectified Linear unit (ReLU) activation function. Two completely linked layers are at the top, and the output layers activation function is SoftMax. The next step is to flatten the pooled featured map after it has been collected. The pooled feature map matrix must be reduced to a single column as part of the flattening procedure. The flattened feature map is after that fed into a neural network. The stage components are the input layer, full linked layer, and output layer. The expected result is obtained in the output layer.

The sample height, sample breadth, and number of filters make up the input layer. The number of nodes for each layer is specified by the filter parameter. Each convolutional layer has a different number of filters, ranging from 16, 32, 64, and 128 overall. A 2x2 filter matrix is produced as a consequence of the kernel size parameter's specification of the kernel window's size. In the CNN layers, a dropout value of 20% is used. four neurons make up the last output layer, which represents the six classifications of pneumonia, COPD, URTI, and healthy.

VII. RESULT AND DISCUSSION

In this study, the CNN Model was trained with 920 audio files from different persons, and its overall accuracy was 95% and loss was 0.1564. Whereas the accuracy of the previously trained CNN model is 86%. As result, this system has better accuracy than the existing system. Fig. 7 and Fig. 8 represent the accuracy and loss of the CNN model. The accuracy of each lung disease is shown in the Table 1. The study comprises of NodeMCU for processing the lung signals. The NodeMCU sent the result to the THINGSPEAK software using the in built WIFI module. Fig. 9 shows the complete set up of the project.Fig. 10 represent a straight line if a particular disease detected , here healthy field is detected . Fig. 11 shows the notification sent by the Pushover application.

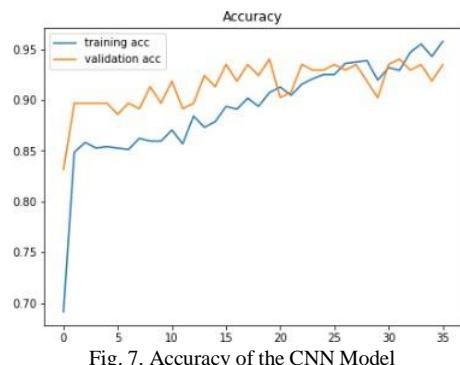


Fig. 7. Accuracy of the CNN Model

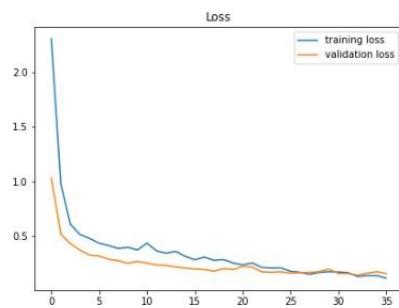


Fig. 8. Loss of the CNN Model

TABLE 1: ACCURACY FOR EACH LUNG DISEASE

Disease	precision
COPD	0.98
Healthy	0.80
URTI	0.67
Pneumonia	0.50



Fig. 9. Complete setup of the project



Fig. 10. graphical represtation in THINKSPEAK

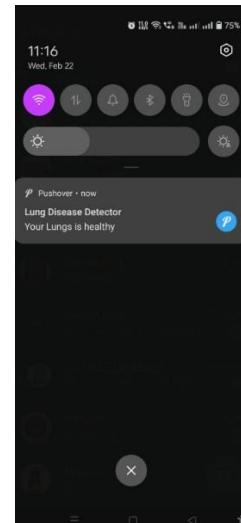


Fig. 11. Notification by Pushover

VIII. CONCLUSION

In this project, a brand-new technique for employing a digital stethoscope to assess the acoustic characteristics of lung sounds is proposed. A digital stethoscope based on IoT is created and put into use to address the shortcomings of traditional stethoscopes. The NODEMCU receives data from the EDA sensor as input, where the analog signals are digitized, then signals are evaluated and type of disease is identified and displayed in THINKSPEAK in terms of which signals plotted and digital value respectively. By capturing data samples from multiple candidates, the proposed device readings were found to be highly linked with the real rate. A low-cost digital stethoscope prototype with real-time digital signal processing, frequency filtering, and noise reduction has been developed. Future work will focus on making the prototype more portable, improving the sound quality of the sounds that are recovered,

and adding new features for wireless data transfer and sound recording.

REFERENCES

- [1] Gunjak, M., & Morty, R. E. (2022). World health day observances in November 2022: pneumonia, chronic obstructive pulmonary disease, preterm birth, and antimicrobial resistance in focus. *American Journal of Physiology-Lung Cellular and Molecular Physiology*, 323(5), L603-L610.
- [2] Indicator metadata registry details. In: World Health Organization. <https://www.who.int/data/gho/indicator-metadata-registry/imr-details/3107>. Accessed 13 Feb 2023
- [3] Jatupaiboon, N., Pan-Ngum, S., & Israsena, P. (2010, November). Electronic stethoscope prototype with adaptive noise cancellation. In 2010 Eighth International Conference on ICT and Knowledge Engineering (pp. 32-36). IEEE.
- [4] Katyal, N., Mathrani, M., & Hota, M. K. (2018, July). Preparation of a Digital Stethoscope. In 2018 IEEE International Conference on System, Computation, Automation and Networking (ICSCA) (pp. 1-4). IEEE.
- [5] Malik, B., Eya, N., Migdadi, H., Ngala, M. J., Abd-Alhameed, R. A., & Noras, J. M. (2017, September). Design and development of an electronic stethoscope. In 2017 Internet Technologies and Applications (ITA) (pp. 324-328). IEEE.
- [6] Wu, Y. C., & Chang, F. L. (2021, September). Development of an electronic stethoscope using raspberry. In 2021 IEEE International Conference on Consumer Electronics-Taiwan (ICCE-TW) (pp. 1-2). IEEE.
- [7] Ramesha, M., DankanGowda, V., Jeevan, K. M., & Sathisha, B. M. (2020, December). Implementation of IoT Based Wireless Electronic Stethoscope. In 2020 Third International Conference on Multimedia Processing, Communication & Information Technology (MPCIT) (pp. 103-106). IEEE.
- [8] Ma, Y., Xu, X., Yu, Q., Zhang, Y., Li, Y., Zhao, J., & Wang, G. (2019, October). Live demo: LungSys-Automatic digital stethoscope system for adventitious respiratory sound detection. In 2019 IEEE Biomedical Circuits and Systems Conference (BioCAS) (pp. 1-1). IEEE.
- [9] Liu, Y., Lin, Y., Zhang, X., Wang, Z., Gao, Y., Chen, G., & Xiong, H. (2017, August). Classifying respiratory sounds using electronic stethoscope. In 2017 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computed, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and SmartCityInnovation(SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI) (pp. 1-8). IEEE.
- [10] Petmezas, G., Cheimariotis, G. A., Stefanopoulos, L., Rocha, B., Paiva, R. P., Katsaggelos, A. K., & Maglaveras, N. (2022). Automated lung sound classification using a hybrid CNN-LSTM network and focal loss function. Sensors, 22(3), 1232.
- [11] Aykanat, M., Kılıç, Ö., Kurt, B., & Saryal, S. (2017). Classification of lung sounds using convolutional neural networks. EURASIP Journal on Image and Video Processing, 2017(1), 1-9.
- [12] Dataset link - <https://bhichallenge.med.auth.gr>

A Novel Approach in Web Based 3D Virtualization For Healthcare

Banu Rekha B

*Dept. of Biomedical engineering
PSG College of Technology
Coimbatore, India
bbr.bme@psgtech.ac.in*

Charushree S

*Dept. of Biomedical engineering
PSG College of Technology
Coimbatore, India
19d208@psgtech.ac.in*

Divyadharshan P

*Dept. of Biomedical engineering
PSG College of Technology
Coimbatore, India
19d213@psgtech.ac.in*

Harish Babu K

*Dept. of Biomedical engineering
PSG College of Technology
Coimbatore, India
19d217@psgtech.ac.in*

Vasunthra V

*Dept. of Biomedical engineering
PSG College of Technology
Coimbatore, India
19d245@psgtech.ac.in*

Abstract—3D Virtualisation refers to the process of creation of graphical contents using 3D softwares. These include images and animations that improve the communication and provide the users with a more realistic online experience. It also serves as a great tool to demonstrate either a prototype or a finished product to the stakeholders. In the present, the customers and learners prefer viewing and learning the virtual model of the product in order to efficiently understand the concept, theory and principle behind it. In this case, 3D virtualization provides an excellent opportunity for customers and learners to see the layout of the product design, texture and working model as animation.

Keywords—3D virtualization, Graphic Design, E-Learning

I. INTRODUCTION

Interactive web based 3D virtualization of healthcare based products are useful for medical education and for service based learning objectives of the medical devices. 3D virtualization plays a significant role in architecture and interior design of the product for training purposes. A lot of industrial sectors, such as product designing and development, entertainment and media have profited from this as a result of incorporating this technology. The field of healthcare is making its advancement in the technology driven world by integrating science along with the current available technology. Newer medical devices are being manufactured and developed according to the current needs. Conventional methods of describing the medical devices haven't been successful in acquiring the necessary information about it. Visualizing the medical device through two dimensional images does not give the needed information about it to an extent. Our goal is to create a 3D-based interactive virtual environment of medical equipments for education, business, and service training.

Our work focuses on designing the 3D virtual models of medical devices such as tracheostomy tube, bileaflet heart valves, mammogram and integrating them into a web based platform with interactive features, so that users can learn about the internal components, functionality, and specifics of how the device works. Our work aims to serve as a one-stop platform for learning, visualization, and demonstration of the medical equipment so that customers may engage with the virtual model of the device and understand the theoretical principles underpinning each device.

II. 3D VIRTUALIZATION

The technique of developing graphical content with 3D software is known as 3D virtualization. It is the process of using software to create graphics and renderings. It serves as a powerful tool for bringing design concepts to life. With the help of this technology, it is possible to produce visually appealing digital content in fields such as games, movies, commercials and music videos. It has been used in transforming 2D drawings into reality. 3D animation and virtualization provides the users a three-dimensional view with real-world effects that 2D effects cannot provide. It can handle designs in real-time, quickly transition from concept to concrete and explore various options, refinements and multiple design versions. 3D virtualization technology is implemented in many industries to create excellent graphic content including images, diagrams and animations that contribute to better communication or a more real-life digital experience. In the field of product design, manufacture and presentation, 3D virtualization is a great tool to demonstrate either a prototype or a finished product to stakeholders. It helps developers to visualize the to-be product in a real environment from every angle and gather data about potential flaws and risks so that they

could improve the developing process by coming up with new designs and head for innovations.

3D virtualization is also being implemented widely in the field of medicine. Medical 3D has become a revolutionary concept in the present, which offers visual representations of scanned body parts with the help of virtual models. This offers depth and details when compared to the conventional static two-dimensional medical images, therefore facilitating better diagnoses, reduction of surgical operations, learning curves, and minimization of operational costs.

Some of the key features of 3D Virtualization are listed down below:

- 3D virtualization offers better product presentation in a lighter, flexible, faster and safer way and it offers interactive and immerse product presentation.
- With the help of 3D technology, consumers can interact with the product, view it from a complete angle while rotating, zooming in, zooming out and analyzing each detail.
- It is also the best way to convey the ideas than photorealistic renderings, which requires much more input and is less flexible.
- 3D renders and virtualizations can improve communication between the vendor and the customer by providing a more accurate representation of ideas and concepts. Accurate models and imagery helps to improve comprehension and understanding of the product.
- It also helps to boost creativity by allowing users to experiment with different design possibilities.

III. LITERATURE SURVEY

Web-based 3D medical image visualization describes the framework and technology, integration as well as the implementation details. The developed system has been used as an educational tool prototype in the Biomedical engineering department. It was compared with a conventional 2D visualization application with regard to the system efficiency and user satisfaction. The 3D system generally demonstrated better performance and a higher level of satisfaction [1]. Web-based medical data visualization and information sharing towards application in distributed diagnosis includes color mapping and high-quality image production in a web-based raycasting algorithm that was created to give real-time data visualization on web browsers by utilizing the new 3D features of WebGL2 [2]. Interactive

virtual scanning electron microscope inspired by 3d game-design using the game production engine Unity 3D, a virtual 3D SEM was created in order to offer affordable hands-on instruction. Using 3ds Max, a programme for 3D modeling and rendering, it has accurate 3D models of the actual parts^[3]. Using the examples of volcanic areas with different geological features, how the potential of the approach allow users to be able to virtually map and measure remotely, and to collect data for research and teaching is well described in the paper Real world based immersive Virtual Reality for research, teaching and communication in volcanology^[4]. Communicating 3d data interactive 3d pdf documents for expert reports and scientific publications in the field of forensic medicine provides exemplary data for the most frequently used imaging techniques in postmortem imaging, three example cases were selected. Postmortem computed tomography (PMCT), postmortem magnetic resonance imaging (PMMR), and 3D surface imaging data sets were used to build 3D surface models^[5]. Interactive 3D PDFs offer the possibility to communicate 3D information to the reader while maintaining all the benefits of a regular 2D PDF. With Adobe Acrobat, the reader can interactively navigate through 3D data sets and create sufficient depth cues to generate a realistic 3D perception of the data^[6]. A computational approach to body mass index estimation from dressed people in 3d space examines the relationship between estimated body volumes and BMIs, the investigation of BMI estimation from three-dimensional (3D) visual data, and the subsequent development of an effective BMI computing method^[7]. The authors study BMI estimation from the three-dimensional (3D) visual data by measuring the correlation between the estimated body volume and BMIs, and then develop an efficient BMI computation method^[8]. A review of three-dimensional medical image visualization provides an overview of 3D visualization techniques for medical imaging with the intention of bridging the gap between medical experts and visualization researchers and the needed web links to free software tools for various medical imaging modalities that have been developed for a variety of purposes, including visualization, analysis, and segmentation^[9]. Interactive graphical applications bring the understanding to a higher level, where the users get actively engaged. The interactive process enhances the understanding, even if the user does not interact with a strategy but just uses his intuition. These applications can be also used to explain complex concepts to the general public^[11]. Recent technological innovations have created new opportunities for the increased adoption of virtual reality (VR) and augmented reality (AR) applications in medicine. While medical applications of VR have historically seen greater adoption from patient-as-user applications, the new era of VR/AR technology has created the conditions for wider adoption of clinician-as-user applications^[12].

IV. SOFTWARES USED

A. Blender

Blender is a free and open-source 3D computer graphics software tool set used for creating animated films, visual effects, sculpting, 3D-printed models, motion graphics, interactive 3D applications, and virtual reality. It supports every step of the 3D pipeline, including modeling, animation, simulation, rendering, compositing, motion tracking, and even video editing and game development.

Although there are a lot of free and open source softwares such as SketchUp, Rhino, Autodesk, Maya, 3ds Max available in the market for use, Blender is preferred over the latter because:

- It supports a wide array of file formats for importing and exporting, and external plugins to work in a more smoother way.
- It has a relatively user-friendly interface given the complexity of what it is capable of. Shortcuts can be used for various operations in Blender.
- It is also an excellent physics engine which makes it easy to create realistic animations and images.
- The animations can be easily generated with movie grade capabilities in Blender and it can also incorporate any other animations made in any other software.
- Blender serves as a great tool in rendering the models in a more realistic way. It has a deep bench of tools to make complex designs easier.

B. Verge3D

Verge 3D is a real-time renderer and a toolkit used for creating interactive 3D experiences running on websites. It is used to develop interactive animations, product configurators, captivating presentations, online businesses, explainers, e-learning materials, or even browser games using well-known design softwares like Blender, 3ds Max, or Maya. Puzzles, which are built using Google Blockly low code programming, are a strong yet simple tool used by Verge 3D to create behavior scenarios. The developer can make their content interactive and user-responsive with the aid of these puzzles.

IV. METHODOLOGY

The proposed work consists of integration of tasks performed in different softwares. Modeling the device parts,

post processing techniques and providing the animation for interaction are done using Blender along with Verge 3D. The information regarding the device and the graphical user interface, are integrated along with the complete model using Verge 3D, thus creating an interactive webpage for the device. The architecture of the proposed work is given in Figure 1.

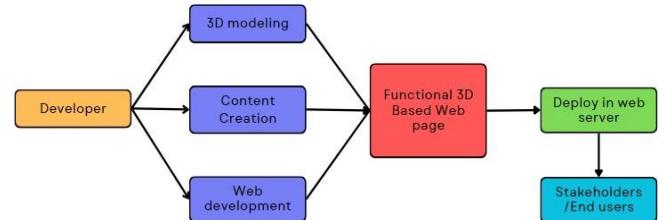


Fig 1(a):- Methodology of proposed work

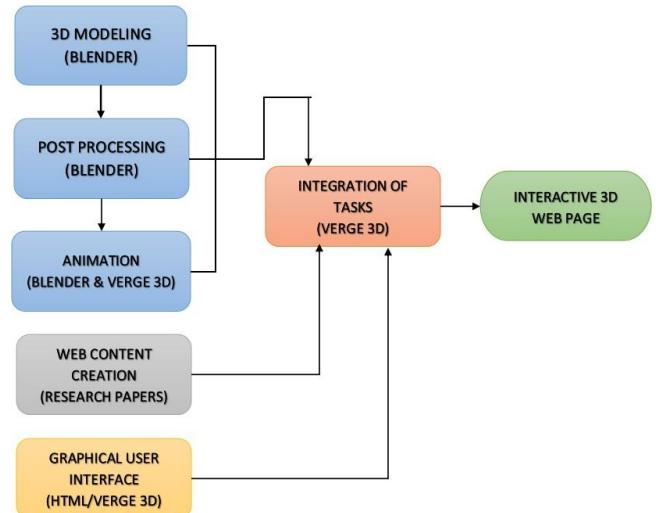


Fig 1(b):- Architecture of proposed work

A. Modeling

The modeling of real-time devices not only aids in conveying information about the product in great detail, but it also helps to set early expectations for the customer or learner so that they are aware of how a product works, what the respective sizes of each component are, and how the product's form, color, material, and finish change with each change in configuration.

B. Post processing

Post-processing 3D models improves the product's visual appeal and effectively gives the model surface color and texture to make it realistic like real-time model, as well as adding appropriate lighting to ensure brightness for a clear view of the model and background effects for display.

C.Animation

Animation is significant because it enables immersive interaction among viewers. By using 3D software, animators can produce computer-generated objects that appear to be 3D even when they are on a 2D surface. In contrast to 2D animation, 3D animation is used for complex themes and goods since it is easier to see them in 3D space and almost realistic, and it helps to better convey the features and functionality. This conveys a clear message that is simple for the audience to understand.

D. Webpage Content Creation

The purpose of content creation is to educate stakeholders about the purposes, functionality, and components of the product. Research articles and textbooks serve as the source material for the content creation. The content assists the user in comprehending the device's theoretical aspects. The user can actively learn the significance of each component of the device and become familiar with each component when the content is shown alongside the 3D model.

E. Graphical User Interface

One of the most crucial elements of a web page is the graphical user interface. The graphical user interface facilitates effective user communication with the website. With the aid of Verge 3D, which makes it simple to combine a 3D model and content with the web page structure on which it will be used. This procedure will ensure that the website is interactive and user-friendly. The overall user experience is directly impacted by the website's interactions.

F. Integration of tasks

The final phase includes combining different tasks accomplished on different platforms into the finished product. Integrating 3D modeling, HTML scripts, and Verge 3D puzzle blocks into one webpage is crucial for this.

V. IMPLEMENTATION AND RESULTS

A. Modeling

Blender software is used to create the 3D models of the device. The work consists of importing the physical device's related digital twin model, creating missing or ambiguous components in the current model, scaling to balance out the device's size, and dividing limited sections into distinct primary parts as needed for effective post-processing.

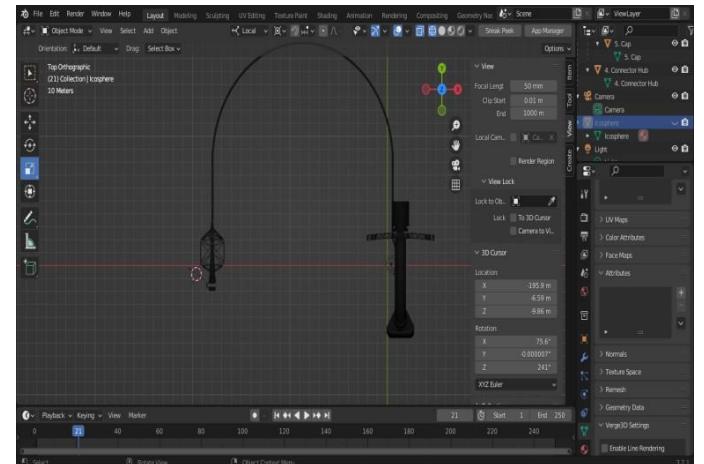


Fig 2:- Wireframe view of the tracheostomy tube in Blender

B. Post processing

To replicate the real model, object textures such as transparency, color, and surface are added to the device. The linked components were separated in order to offer color and texture that matched the real-time device. Background effects for presenting the device model in a web page and adding lighting and camera effects were also done using Blender software.

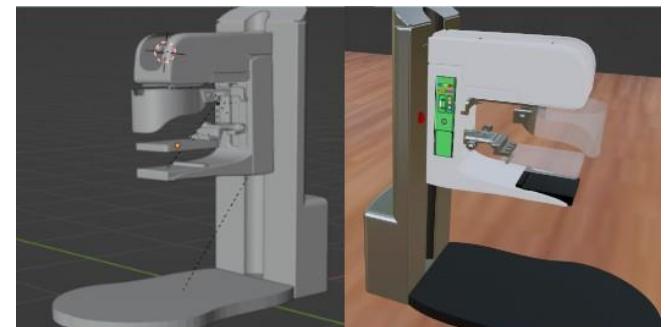


Fig 3:- Post processing works on Mammogram in Blender

C.Animation

The animation includes rendering a device part motion that corresponds to its original uses. Blender and Verge 3D were used to animate the devices' working components. Glowing effects were added to help identify certain components of the device. To give the controls and add annotations that when clicked, will identify the portion, Verge 3D puzzles are employed.

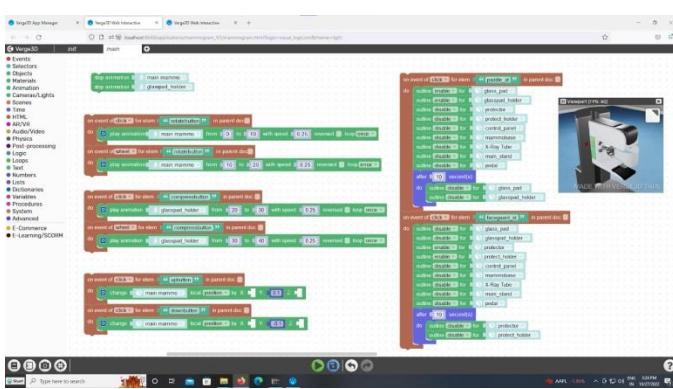


Fig 4:- Puzzles for Animation using Verge 3D

D. Webpage Content creation

Information regarding the purposes, functionality and parts were researched from various books and literatures. The website contents were created accordingly based on the inferences made.

E. Graphical user interface

Graphical user interface creation process entails designing the page layout for the website and developing interactive elements that are user-friendly using HTML and Bootstrap, such as buttons, cards, and accordions. Verge 3D functionality is offered for 3D model interaction.

F. Integration and final output

Verge 3D software is ultimately used to combine all of the aforementioned activities onto a single platform once they have all been completed using separate software. To design a responsive homepage, it also involves HTML scripts, 3D models, and verge puzzles.

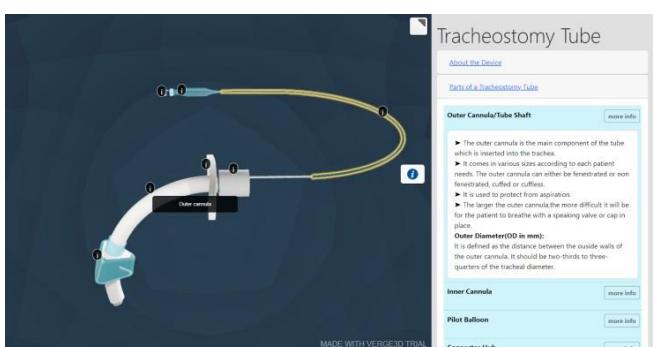


Fig 4:- Webpage of Tracheostomy tube

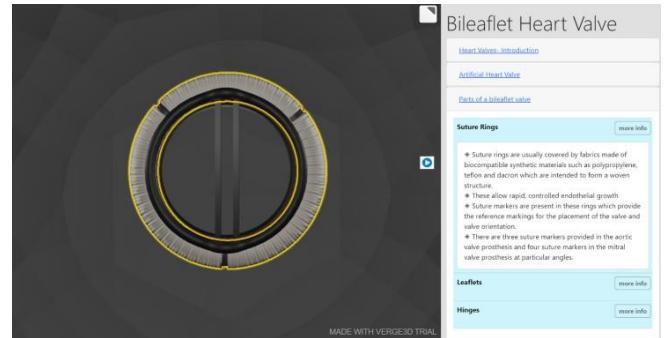


Fig 5:- Webpage of Bileaflet heart valve



Fig 6:- Webpage of Mammogram

G. Results

The final webpage developed in the figure 4,5 and 6 shows a complete interactive webpage with the features such as easy part identification of the medical devices, animation of working of the particular device, important features, learning objectives along with relevant textures and dimension of the device. Controls are also available for moving, zooming and identification of parts on clicking.

VI DISCUSSIONS

The paper “Three-Dimensional Display technologies for Anatomical Education” aims to augment current curriculum and improve students’ spatial knowledge of anatomy, many educators, anatomists, and researchers use three-dimensional (3D) visualization technologies^[10]. The development of 3D virtualization technology is definitely in the fast track of growth and is in a continuous state of flux. It also serves as a cost minimizing solution for various sectors wherein product demonstration is required. Technology is being used more and more in various industrial areas as new breakthroughs are put into practise. The features of this work includes, Easy part identification of complicated parts of medical devices, 3D animations illustrating the functionalities of the device, Learning information along with 3D model for educational purpose and Interactive web page facility.

VII FUTURE SCOPE

By utilizing immersive VR/AR technology for the demonstration of the device in an artificial environment that completely replaces the real-world surroundings, the web application will develop into a software-based product that will be accessible to the various stakeholders and will give users an even better learning experience and interaction with the device.

VIII CONCLUSION

Thus the e-learning platform for several healthcare devices is created using blender and Verge 3d software and they are made as interactive webpages for effective understanding purposes. This work will be more purposeful for many medical professionals in equipment servicing and in marketing. Web content creation for a wide variety of medical equipment and instruments consumes more time as their design, material, working, dimensions are more complicated.

REFERENCES

- [1] Sittapong Settapat, Tiranee Achalakul, Michiko Ohkura, “Web-based 3D medical image visualization framework for biomedical engineering education”, Computer application in engineering education, Volume 22, Issue 2, 2019.
- [2] Qi Zhang, “Web-based medical data visualization and information sharing towards application in distributed diagnosis”, Informatics in medicine unlocked, volume 14, 2018, pp.69-81.
- [3] Pnina AriGur, Igor Lapsker, Tyler William Bayne, “Interactive Virtual Scanning Electron Microscope Inspired by 3D Game-Design”, International Journal of Materials Science and Applications, Volume 11, Issue 1, pp. 37-41 2022.
- [4] A.Tibaldi, F.L.Bonali, F.Vitello, “Real world based immersive Virtual Reality for research, teaching and communication in volcanology”, Bulletin of volcanology, 2020.
- [5] Filipi Pires, Carlos Costa, Paulo Dias, “On the Use of Virtual Reality for Medical Imaging Visualization”, Journal of Digital Imaging, 4(4), pp. 1034-1048, 2021.
- [6] Soren Kottner, Patricia Mildred Flach, Dominic Gascho, “Communicating 3D data interactive 3D PDF documents for expert reports and scientific publications in the field of forensic medicine” International Journal of legal medicine, 2019.
- [7] Shiwei Yan, Wei Wang, Jianming Geng, “Research on 3D Visualization Simulation System of Deck Crew Based on OpenGL Technology”, Journal of Physics Conference series, Vol. 1550, Iss. 3, 2020, pp. 203-207
- [8] Min Jiang, Yuanyuan Shang, Guodong Guo, “A Computational Approach to Body Mass Index Estimation from Dressed People in 3D Space”, IET Image Processing, Volume 14, Issue 7, pp. 1248-1256, 2020.
- [9] Liang Zhou, Mengjie Fan, Charles Hansen, “A Review of Three Dimensional Medical Image Visualization”, Science partner journal, 2022.
- [10] Matthew Hackett, Michael D. Proctor, “Three-Dimensional Display technologies for Anatomical Education”, Journal of science education and technology, 2016.
- [11] Francesca Taddei, Benjamin Rodenberg, Christian Karpfinger, “Interactive Web Apps for Visualizations in Mathematics and Engineering”, Proceedings in Applied Mathematics and Mechanics, Volume 18, 2018.
- [12] Justin Sutherland, Jason Balec, Adnan sheik, “Applying Modern Virtual and Augmented Reality Technologies to Medical Images and Models”, Journal of Digital Imaging, pp.38-53, 2019.

CLASSIFICATION OF CULTIVARS EMPLOYING THE ALEXNET TECHNIQUE USING DEEP LEARNING

K.Gurunathan
Department
of Information Technology
M. Kumarasamy College of
Engineering
Karur, Tamilnadu
gurunathank.it@mkce.ac.in

K.Hariprasath
Department
of Information Technology
M. Kumarasamy College of
Engineering
Karur, Tamilnadu
hariprasath.hp.karur@gmail.com

V.Bharathkumar
Department
of Information Technology
M. Kumarasamy College of
Engineering
Karur, Tamilnadu
kumarbharat58867@gmail.com

R.Jidendiran
Department
of Information Technology
M. Kumarasamy College of
Engineering
Karur, Tamilnadu
jidendiransoft@gmail.com

M.Haji ali meeran
Department
of Information Technology
M. Kumarasamy College of
Engineering
Karur, Tamilnadu
hajiaslan111apnt@gmail.com

ABSTRACT--The recognition of hybrid fruits by humans is considered a challenging task as fruits exist in various colors, sizes, shapes, and textures. In the marketplace, where prediction of fruits, vegetables, and pulses by retailers and the general public is very difficult. In this paper, we have recognised two different classes of fruits, vegetables, and pulses and categorised them based on the stages of their existence and maturity in a procedure called fruit maturity classification. Crop distribution, good fruit counts, crop harvesting, crop disease detection, weed management, and production forecasting are all aspects of managing water and soil and are just a few of the smart agricultural applications that employ robust learning (DL). The finest deep learning is what this project seeks to produce in algorithms for estimating fruit quality and ripeness in order to forecast fruit shelf life.

Manual vegetable and fruit detection is a hard process when done in large numbers, but it becomes easy when done in small amounts. Automated detection of these is thus used. Fruit, crop, and pulse images served as the input for the first stage of processing, which included detection. Background removal, extraction of colour and texture properties, and categorization comprised

the three steps of the process. The K-means clustering method was used for background subtraction. Statistical attributes were used to pinpoint colour characteristics.

This research proposes a simple and effective method for detecting fruits and predicting their nutrition information using deep Alex networks (DAN). The datasets used in the investigation were obtained from the Fruit 360 library of image processing problems. Apples, berries, bananas, grapes, papaya, peaches, avocados, and various apple tastes are among the fruit groups. The trials are also carried out on a variety of additional fruit samples gathered from various Web archives. The network design consists of three fully linked layers, including the max pooling and RELU levels, and five convolution layers. $227*227*3$ photos should be used as input for 96 filters that are $11*11*3$ and have a stride length of 4.

Index Terms—Fruits, Vegetables, Hybrid, Machine Learning, Deep Learning

I.INTRODUCTION

Fruit, crop, and vegetable grading is necessary yet time-consuming given the significance of food in our everyday lives. Computerized methods of automatically grading are seen as the answer to this issue, reducing the need for human work. There is some evidence to suggest that as fruit ages, it goes through a series of metabolic changes that alter its physical characteristics and chemical make-up, including its nutritional content. There are two categories of fruit grading techniques: non-visual and visual. The main non-visual grading criteria are chemistry, scent, and tactile perception.

Only when the fruit is still clinging to the tree during development does it reach maturity, which is evidenced by a stop in cell division and a buildup of dry material. Along the postharvest value chain, the quality of all fruits and vegetables is significantly impacted by their ripeness at harvest. An efficient and effective automated model that can recognise and categorise the fruits based on their maturity degree in a short amount of time is desperately needed. Big data technologies and highly effective computers have given birth to DL technology, opening up crop management and crop harvesting prospects in the context of agricultural activities.

II.RELATED WORK

In [1], they have identified nine distinct fruit classes. Fruit picture datasets may be found online, and some photographs can be found simply by utilising a phone's camera. Pre-processing was used on these images to crop out the background and isolate the blob that represents the fruit. Fruits are shown, and their visual characteristics are captured through combinations of colour, shape, and texture factors. Multiclass SVM and KNN classifiers provide additional input for these feature datasets. The colour image is first converted to grayscale via GLCM (Gray Level Concurrence Matrix). The image is then changed to a binary

version. The biggest blob or item, which would also be regarded as fruit, is extracted from the picture using morphological processes, which are also utilised to fill in the gaps in the image. After cropping the largest lump, the binary numbers are reset to their original intensity levels. According to the research, combining colour, texture, and form produces outcomes that are superior to or on par with those obtained when using any two feature categories. The second inference that can be made is that KNN outperforms SVM in this scenario.

In [2], there are various stages of the training process, which include the following: Collecting fruit images is the first step. Then, using the FCH and MI methods, feature extraction is used to extract the fruit's characteristics, which are then converted into vector feature forms that can be stored in databases. On the vector of the database's image of fruits, a later clustering method called K Means Clustering is used to carry out the procedure. The following actions were taken throughout this study's testing phase: Find the fruit by opening the file picture. The next step is to extract the features from the facial picture, which are then converted into a vector feature form using the same training process. Following that, the Euclidian distance between the new fruit picture features and features already present in the database was calculated as part of the recognition process using the KNN approach, and the results were compared with the clustering findings.

In [3], Zhang et al. (2015) used a high-end dual camera setup to collect both visible (RGB) and infrared (IR) pictures. They gathered 1088 RGB+IR-matched pictures from six different sources. They named this dataset VAIS and made it available to the public for use. Their goal for employing infrared imaging is to improve nighttime performance. They used SIFT characteristics to train VGG-16 and Gnostic Fields. They achieved 87.4 percent daytime accuracy and 61.0 percent nighttime accuracy using those classifiers in combination.

Fruit identification was demonstrated by Patel, Jain, and Joshi [4] using an enhanced multiple feature-based approach. Effective feature extraction is trained into an image processing method in order to identify the fruit. The algorithm's design is to determine various weights for the input test image's properties, such as intensity, colour, orientation, and edge.

The approaches for fruit processing's sorting and grading were introduced by Nagganaur and Sannanki [5].

The machine begins the procedure by taking a picture of the fruit. Following that, the picture is sent to Matlab for feature extraction, categorization, and grading, all of which are accomplished using a fuzzy logic technique.

The literature review contains several recognition and classification systems that may automatically examine the fruits for illnesses, a maturity phase, category recognition, etc. The approach taken by [6] to categorise the bananas using the CIE Lab and hue channel. The fuzzy parameters were modified as part of the particle swarm optimization (PSO) process. [7] classified fruits using a kernel support vector machine with several classes (KSVM). SVMs were trained using the reduced feature vector and 5-fold stratified cross-validation. During the categorization process, a mix of color, texture, and form characteristics were employed. A unique multi-layered feed-forward unsupervised neural network called a convolutional neural network (CNN) was created to handle picture categorization. The feature extraction layer is another name for the convolution layer of a convolutional neural network (CNN).

III. EXISTING SYSTEM

The CNN algorithm was used by the authors to distinguish between yellowish-green, unripe, medium, and ripe bananas. Before submitting them to training, image noise was removed using a

bilateral filter, and for variants, data augmentation was used. In terms of accuracy (96.18%) and execution time, the recommended model surpassed NASNet Mobile. The authors experimented with various CNN hyper-parameters to sort ripe Medjool dates using CNN from scratch architectures, ResNet50, ResNet101, ResNet152, VGG16, VGG19, InceptionV3, etc. The effectiveness of CNN architectures in classifying the maturity of Medjool dates was rated in terms of accuracy and processing speed. The Adam optimizer's 0.01 learning rate with 128 batch sizes gave the VGG19 model a maximum accuracy of 99.32%.

IV. PROPOSED MODEL

An example of a deep learning system is a CNN, which is made up of neurons and uses trainable weights and biases to classify incoming images. A CNN may comprise tens or even hundreds of layers, each of which may be trained to recognise certain aspects of an image. A CNN may automatically and adaptively learn spatial hierarchies of information by using convolutional layers. For example, pooling layers and completely linking layers are a few of the building blocks. The output of each training picture's consolation is used as the input for the next layer after each image is convolved using a range of resolution filters.

With the input picture size set to 112*112*3 for the supplemented dataset and 227*227*3 for the original dataset, this study uses three convolutional layers with two max-pooling layers. The loss function used is the cross-entropy function, while Adam is selected as the optimizer, so that weight and offset modifications may be made more steadily thanks to Adam's method. To balance training and validation accuracy and loss, a 20% dropout was employed. The output layer uses the softmax activation function as a last step.

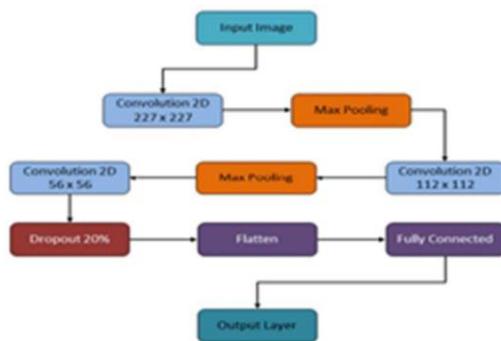


Fig 1: CNN ARCHITECTURE

A.ALEXNET

The AlexNet architecture is divided into eight layers, three of which are completely connected and five of which are convolutional. The convolution layer, which is the foundational component of the network, constitutes the top layer. The first layer of AlexNet's convolution window is 11 by 11. Objects in ImageNet data frequently occupy pixels that are times wider and higher because ImageNet photographs are eight times larger than MNIST photographs and include more visual information. We need a larger convolution window to catch the item. Electronics 4100, 2022 The convolution window shapes of 8 of the 13 are altered to 5*5 and then 3*3.

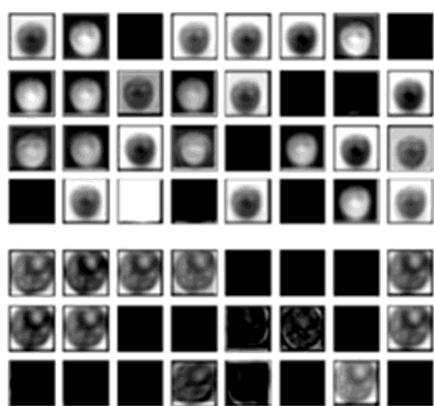


Fig 2: ALEXNET CLASSIFICATION

The final convolutional layer is immediately followed by two gigantic, fully connected layers with a total of 4096 outputs. The input picture,

which has pixel values of 227 for width and 227 for height, is transferred to the input layer along with further 3D colours that are RGB-saved images. The concealed layers are then transmitted with the image, where it is processed before moving on to the fully connected layers, which include convolutional layers of different filters, and lastly to the layer of output. The first complicated layer is reached after the input, and the picture is transmitted. is reduced in size, depending on whether padding is there or not. The pooling layer's input image will be 55*55*96 in size since the first layer's filter size is 96. The two formulas stated above are used to determine the values for each input size. When using the supplemented dataset, the picture size is decreased to 112 by 112 pixels since training a model with a large dataset consumes more RAM than is available. As a result, by scaling down the image to "112," the AlexNet model is trained, verified, and tested.



Fig 3: PROPOSED FRAMEWORK

B.PREPROCESSING

To extract the image's features, we used three different learning algorithms. Preprocessing is followed by image resizing and rgb to grey conversion, and three learning algorithms are available: preprocessing followed by SVM, bag of features, and custom-trained convolutional neural networks using transfer learning.

C.SEGMENTATION

A method for segmenting a collection of data into a preset number of groups is called "K-mean segmentation clustering." The most widely used method is k, which stands for clustering. It divides a group of items using the K-means clustering algorithm. It creates k separate clusters from a given amount of data. The K-means algorithm consists of two sections. After computing the k centroid in the first phase, each data point is assigned to the cluster with the centroid that is closest to it in the second phase data into a group of k numbers of data.

D.HYBRID METHOD CNN WITH SVM

Hybrid fruit picture categorization using SVM (support vector machines). A hybrid CNN-SVM model is recommended for the Maraval dataset's ship categorization. The recommended method offers the best of both worlds by combining SVM and CNN classifiers. Convolutional neural networks (CNNs), which are used for supervised learning, consist of multiple completely connected layers.. CNN is able to learn invariant local properties and functions similarly to humans. It can extract the most discriminating information from unprocessed ship photos. The suggested method extracts the most recognisable characteristics from the raw input photos using a 5x5 kernel/filter. The mm filter in the convolutional layer is convolved with the cnn input neurons in the input layer.

E.TENSORFLOW

Open source software for numerical computations is called Tensor flow. It was initially intended to be used for machine learning and deep neural network research. Users who wish to employ neural networks in various scenarios may find neural network topologies in Tensor Flow along with retraining scripts.



Fig 4: HYBRID FRUIT IMAGES

F.KERAS

To prepare, model, evaluate, and optimise neural networks, one uses the open-source Keras Python-based neural network library. It has the ability to run on top of TensorFlow. Given that the backend is responsible for high-level APIs utilise it for management.. It is designed for both the training process with a fit function and the development of a model. It is intended for low-level computing using tensors or TensorFlow and backend convolution. Preprocessing, modelling, optimization, testing and presentation python libraries imported.

V. EXPERIMENT RESULTS

The deep learning framework PyTorch has an inbuilt tensor data format (a multidimensional tensor). In this paper, MobileNet was applied to the Fruit Dataset to determine how well the network performs at classification. The Fruits dataset was used to generate these 1260 images, which are divided into 7 categories: 15% of these images are used to test the model, and 85% are used for training and overall 98.74%. The network is trained using a 14-batch size across 10 epochs. The results

demonstrate that the suggested paradigm operates well, when compared to established models and show promise for use in practical situations. This type of increased precision and accuracy will help increase the machine's overall fruit recognition effectiveness.



Fig 5: UPLOADING IMAGES

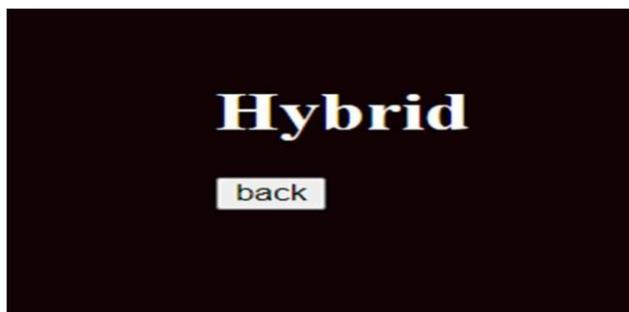


Fig 6: PREDICTION

VI. CONCLUSION

The fruit may be identified by properties like form, colour, and texture in the suggested project. This expands people's understanding of certain uncommon and unusual fruits. The project's major focus is on minimising human effort and simplifying human existence. Fruit identification might decrease the continuous issues that are present. It reduces misunderstandings about the particular fruit. Future work that might be done on this project includes the development of a web application. This software is available to users 24/7, from any location.

REFERENCES

- [1] Xiaoyang Liu, Dean Zhao, Weikuan jia, Wei Ji, Yueping Sun, "A Detection Method for Apple

Fruits Based on Color and Shape Features", IEEE Access, 22 May 2019

[2] Seng, Woo Chaw, and Seyed Hadi Mirisae. "A new method for fruits recognition system." Electrical Engineering and Informatics 2009. ICEEI'09. International Conference on. Vol. 1. IEEE, 2009.

[3] Nishat Tasnim, Md. Romyull Islam, and Shaon Bhatta Shuvo "A Convolution Neural Network Based Classification Approach for Recognizing Traditional Foods of Bangladesh from Food Images"

[4] Sahu, Dameshwari, and Chitesh Dewangan. "Identification and Classification of Mango Fruits Using Image Processing." Int. J. Sci. Res. Comput. Sci. Eng. Inf. Technol 2.2 (2017): 203-210.

[5] Shadman Sakib1, Zahidun Ashrafi2, Md. Abu Bakr Sidique3. "Implementation of Fruits Recognition Classifier using Convolutional Neural Network Algorithm for Observation of Accuracies for Various Hidden Layers

[6] Hannan M.W., Burks T.F., Bulanon D.M., "A Machine Vision Algorithm for Orange Fruit Detection", Agricultural Engineering International: the CIGR E journal, vol-XI,Pages:1-7, December-2009.

[7] A. Vyas et al, "Colour Feature Extraction Techniques of Fruits: A Survey", International Journal of Computer Applications (0975 – 8887) Volume 83 – No 15, December 2013.

[8] Ms. Snehal Mahajan , Prof. S. T. Patil "Optimization and Classification of Fruit using Machine Learning Algorithm". IJIRST – International Journal for Innovative Research in Science & Technology Vol.3- No 01 June 2016

[9] V. Leemans and M.-F. Destain, "A real-time grading method of apples based on features extracted from defects," J. Food Eng., vol. 61, no. 1, pp. 83_89, Jan. 2004.

[10] T.-N. Do and J.-D. Fekete, "Large scale classification with support vector machine algorithms," in Proc. 6th Int. Conf. Mach. Learn. Appl., Cincinnati, OH, USA, Dec. 2007, pp. 7_12.

[11] Bargoti, S. and Underwood, J., 2017, May. Deep fruit detection in orchards. In Robotics and

Automation (ICRA), 2017 IEEE International Conference on(pp. 3626–3633). IEEE

[12] Prabha, D. Surya, and J. Satheesh Kumar. "A study on image processing methods for fruit classification." Proc. Int. Conf. on Computational Intelligence and Information Technology, CIIT 2012.

Eye Blink Based Biometric Authentication System

G Umashankar
Department of Biomedical Engineering
GRT Institute of Engineering and
Technology
Tiruttani, India

G Hari Krishnan
Department of Electrical and
Electronics Engineering
Sree Vidyanikethan Engineering
College, Mohan Babu University
Tirupati, India
haris_eee@yahoo.com

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

G Mohandass
Department of Biomedical Engineering,
Agni Institute of Technology, Chennai,
India

T Devaraju
Department of Electrical and
Electronics Engineering
Sree Vidyanikethan Engineering
College, Mohan Babu University
Tirupati, India

V. Devika
Department of Biomedical Engineering
GRT Institute of Engineering and
Technology
Chennai, India

Shaina Banu S
Department of Biomedical Engineering
GRT Institute of Engineering and
Technology
Chennai, India

Abstract— Devices are locked and unlocked using biometric security based on eye blinks rather than PINs. Using devices in uncontrolled situations, on the other hand, makes you extremely vulnerable to spoofing by replay assaults. Stationary facial expression authentication is an instance of a biometric; hackers use the facial images of fictitious individuals to activate devices, leading to unprotected activities and the unintentional disclosure of personal data. In place of static face authentication, we addressed a biometric security system that tracks a person's eye blinking motions as well as their face. The suggested system gets a stream of images as information by instructing the user to follow an eye-blink pattern. The developed scheme then confirms the individual's identity using the proper eye-blink pattern.

Keywords—Biometric authentication, spoofing, eye blink and person's identity

I. INTRODUCTION

Every organization must prioritize security, and automatic service access is essential to enhancing security and privacy [1]. Security is the most important aspect of any organization. Biometric recognition's primary objective is to reliably distinguish between individuals and applications using one or more signals produced by physical or behavioral characteristics like fingerprints, faces, iris, voices, hands, or a written signature. Biometric technology has a number of benefits over conventional security measures [2]. A password or PIN that could be lost or stolen is not required to be remembered, nor is it necessary to carry a key or card with you at all times [3]. However, biometric systems have a number of disadvantages, including a lack of privacy because it is possible to access facial photographs from social media or be accidentally clicked, and anyone might obtain finger prints [4].

The proposed research work aims to create a prototype based on a biometric authentication system based on a person's eye blink. As a first stage, it analyzes the possibilities for constructing the authentication system before it can be designed for an authentication system. There is a lot of spoofing or assault in the field of biometric systems nowadays. To create a prototype biometric authentication system based on a person's eye blink, detect the eye blink with the help of a button change, and then confirm the authentication [5], [6].

These days, security is being discussed more and more in a wide range of industries, as well as the technological tools that can be used to combat it: computer access control, e-commerce, banking, and so forth [7] [8]. There are two conventional ways to establish someone's identity. The first approach is based on knowledge. A person's ability to activate a mobile phone is based on their knowledge of personal information, such as a PIN number. The second option is to base ownership on tokens. It might be as straightforward as a badge, a key, or another piece of identification. These two types of identification can be used together to boost security in situations like those involving bank cards [9, 10]. They do, however, each have shortcomings of their own. In the first example, the password is utilized; however, a third party could guess the password or forget it. In the second situation, the badge (or ID, or key) may be lost or stolen [11] [12].

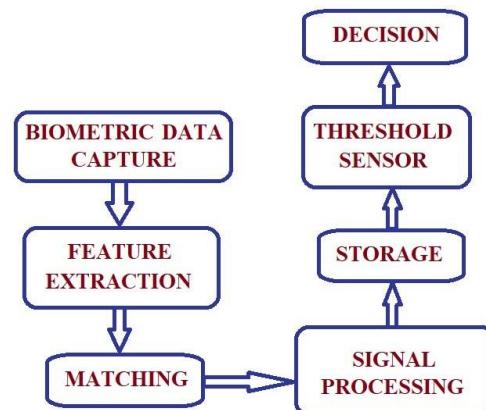


Figure .1 Architecture of the Generic Biometric System

As shown in the above figure 1, the major modules in the generic biometric authentication system are the capture module, which is the biometric system's entrance point and consists of capturing biometric data in order to extract a digital representation [13] [14]. In order to expedite the verification and recognition stages [15], signal processing modules enable the optimization of processing time and the digital representation collected during the enrollment phase

[16] [17]. To determine how comparable the two sets of biometric information are, the matching module compares the information retrieved by the extraction module with the information of the registered models [18], [19]. The judgment module evaluates whether the similarity index returned by the matching module is adequate to identify a specific person [20] [21]. Biometric systems can be used for a wide range of applications [22, 23]. Biometrics can help with security problems in transactions, making everyday life safer and more practical.

Biometric solutions are utilized to satisfy the following domain requirements: Biometric technology and law enforcement have a long history of collaboration, and as a result of this mutually beneficial relationship, some significant achievements in identity management have occurred. The police force's use of biometrics is now truly multimodal. Fingerprint, face, and voice recognition are all important tools for improving public safety and locating the people we're looking for. Biometric technology can be used in a variety of settings, including at the border. Border crossing procedures can be automated with the help of biometric technologies. Passenger screening programs that are reliable and automated, as well as automated SAS, help government agencies be more efficient while still keeping borders safer than ever. In the sphere of healthcare, biometrics provides an improved model. Medical records are among the most sensitive personal documents, and professionals need to be able to access them quickly and properly. A lack of security and proper accounting can distinguish between a quick and correct diagnosis and health fraud.

III. EXISTING BIOMETRIC AUTHENTICATION SYSTEMS

The main objective behind the facial recognition system is the certainty that every person has a unique face [10]. As we know that every person has a unique fingerprint, similarly, every individual face has unique features. Here we use features of the face of an individual. We can store the features of the faces of many individuals, and they can be identified according to their face features. Facial testimony and facial recognition are difficult and challenging pieces of work. For facial recognition systems to be authentic, they must work accurately and precisely. The facial recognition technique captured the image using the camera and mapped it for comparison with the images stored in the database. If the captured image matches any of the stored images, it displays face matched; otherwise, it displays face not matched. This paper elaborates in detail the entire process of a facial recognition system using the OpenCV library. For face detection, we use the OpenCV library and the Haar cascading algorithm.

This work provides a feasibility study of faraway interest degree estimation primarily based on eye blink frequency [11]. The proposed blink detection device was entirely based on convolution neural networks (CNNs). Systems and strategies for spoofing detection in photos are defined herein. A camera can produce a collection of photographs, a primary plurality of photographs within the collection of photographs, including an illustration of a person's frame, and a 2D plurality of photographs within the collection of photographs, including an illustration of the human's surroundings [12].

IV. MATERIALS AND METHODOLOGY

The "Internet of Things" (IoT) is the term used to describe the countless numbers of physical devices that are attached to the internet and are actively gathering and modifying data. Thanks to the introduction of reasonably priced laptop chips and the widespread accessibility of wi-fi networks, it is now possible to demonstrate the entirety of the Internet of Things, from a tablet to a jet. Gadgets that would normally be dumb are given a level of virtual intelligence by linking all of those different items and attaching sensors to them, enabling them to bring real-time data 33 without involving a user.

The digital and physical worlds are coming together through the Internet of Things to establish a smarter, more customer satisfaction for us. A sizable open-source library for computer vision, machine learning, and image processing, OpenCV currently plays a crucial role in real-time operations, which may be crucial in cutting-edge systems. In pictures and films, it has the ability to find objects, faces, or even human handwriting. The suggested eye blink biometric system block diagram and software execution steps are shown in Figure 2.

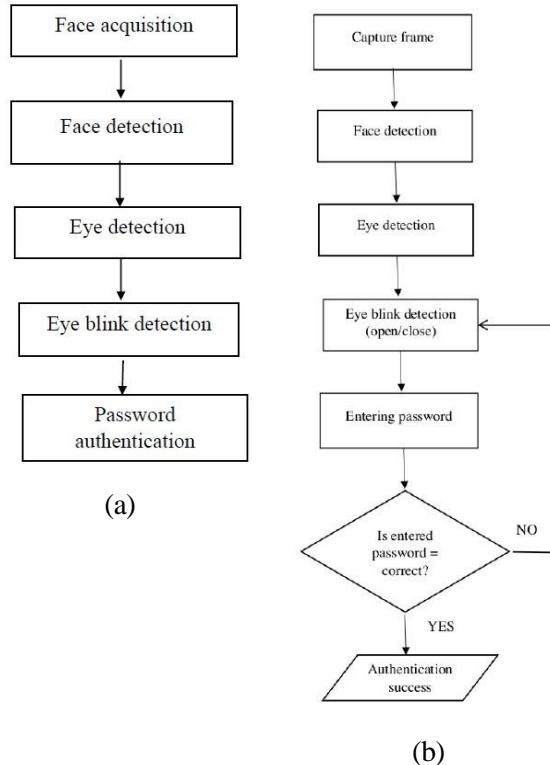


Figure 2 Proposed eye blink based biometric system (a) block diagram and (b) software flow

When used in conjunction with certain other modules, like NumPy, Python can systematize the OpenCV array shape for evaluation. In order to find visible styles and their various abilities, we rent vector areas and perform mathematical operations on those capabilities. The original OpenCV model was updated to version 1.0. OpenCV is free for both academic and commercial use because it is released under the BSD license. It provides C++, C, Python, and Java interfaces and supports Windows, Linux, Mac OS, iOS, and Android. The main objective of OpenCV's development was

to create real-time programs that were as environmentally friendly as possible.

OpenCV is a terrific device for photograph processing and PC vision. It's an open-supply package deal for duties such as face detection, objection tracking, landmark detection, and more. Python, Java, and C++ are a few of the languages supported. Hundreds of helpful functions and algorithms are included in the library, all of which are freely available to us. Some of these functions are quite broad and can be found in nearly every computer vision application. On the other hand, numerous of the functions are still unknown and have not gotten important attention. OpenCV's most outstanding features, which may be applied in a variety of operations, include the following:

To improve the application of screen unlocking by a human face, an eye-blink approach (see Figure 2) in conjunction with facial recognition is proposed. When a user controls a mobile device, facial recognition authentication is required to complete the user's identity identification before the user may access the device's system. As shown in Figure 2, face acquisition, face detection, eye detection, eye blink detection, and password authentication are the five blocks that make up the system architecture.

- Face acquisition: A face image is captured using the graphical user interface of an Open CV, the front-facing camera of a mobile device, or the webcam in a laptop.
- Face detection: It detects our faces while we are facing the camera on our laptop, mobile device, or webcam.
- Eye detection: The open CV detects the presence of eyes and indicates whether they are closed or open.
- Keyboard frame work with the help of an eye blink is done based on the password with the assistance of the GUI. The keyboard has keys from 0 to 9, and the keyboard frame works with the aid of eye blinks based on the password with the help of the GUI.
- Password authentication: When a button on the keyboard is changed with the aid of the eye.

V. RESULTS AND DISCUSSION

The authentication process was experimented with the help of 10 student volunteers, ages 20 to 24. Our proposed study is non-invasive, requiring only a non-contact webcam for authentication. Figure 3 shows Opening the program in an IDLE shell (Python 3.10, 64-bit) and running the program code with the help of the Run module After execution, a GUI tab and button change frame will open in the GUI; users' aliveness is detected, and it also shows the eye detected and eyes open; the next eye blink is detected; and the button change frame has a sequence of numbers from 0 to 9, which is helpful for entering the password. If the user enters the correct password, it will show authentication success, and finally, the output is authentication success. The password that was entered by the user is (1, 2, 3).

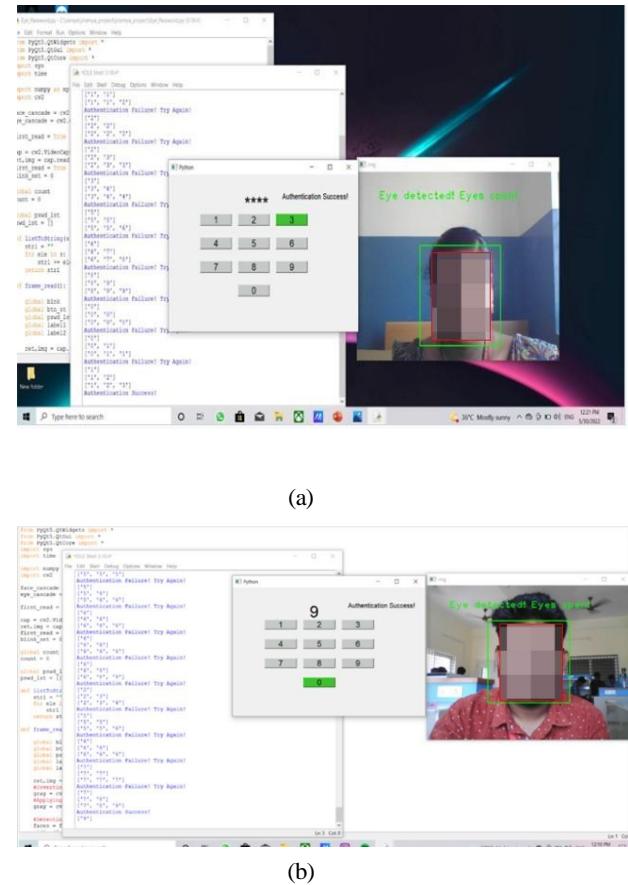


Figure .3 Real time biometric authentication and password entry for (a) Female volunteer and (b) Male volunteer.

Biometric authentication plays a major role in securing the data. The proposed prototype improves the biometric authentication and verification of humans using eye blinking movements. It runs on a laptop with a webcam. The proposed biometric authentication model has been discussed and implemented. This method is very convenient to use, and there is no need to operate manually. This method is very useful for securing personal data. To provide privacy and data security, we proposed biometric authentication with eye blinking detection. It can also be used by people who are wearing spectacles and sunglasses, but lightning conditions affect the performance of the proposed system.

VI. CONCLUSION

People tend to save vital information in their mobile phones due to the popularity of smart phones, such as bank information, photographs, customer data, and subscription information, among other things. When the information is stored on a mobile phone, however, it will require a strong security system to prevent the theft of associated data by others. Face recognition on mobile devices can be easily tricked by using color images of the user. As a result, the eye blink was proposed in this research as a way to improve eye blink detection. Our solution did not require any difficult procedures to verify the user's identification; instead, the user simply had to blink his or her eyes and enter the proper password to complete the authentication process.

VII. FUTURE WORK

Our proposed work includes developing an automated teller machine (ATM) based on eye blinks. The security system will be improved as a result of this. Another option is to build a person who is paralyzed but is still able to convey their thoughts, ideas, and needs. It will undoubtedly assist paralyzed people in communicating their thoughts through the system's provided language. The objective of this technology is to use an eye movement algorithm to lessen the effort required for disabled people to communicate their thoughts.

VIII. REFERENCES

- [1] Albelwi, S., & Mahmood, A. (2017). A framework for designing the architectures of deep convolutional neural networks. *Entropy*, 19(6), 242.
- [2] Jain, A.K., Ross, A., & Pankanti, S. (2006). Biometrics: A toll for information security. *IEEE Transactions on Information Forensics and Security*, 1(2), 125-143.
- [3] Kim, K. W., Hong, H.G., Nam, G.P., & Park, K. R. (2017). A study of deep CNN-based classification of open and closed eyes using a visible light camera sensor. *Sensors*, 17(7), 1-21.
- [4] Komulainen, J., Hadid, A., & Pietikainen, M. (2013). Context based face anti-spoofing. In *IEEE 6th international conference on biometrics: theory, applications and systems* (pp.1-4).
- [5] Maatta, J., Hadid, A., & Pietikainen, M. (2013). Face spoofing detection from single images using micro-texture analysis. In *International joint conference on biometrics (IJCB)* (pp. 1-7).
- [6] Noman, M., Bin, T., Ahad, M., & Rahman, A. (2018). Mobile-based eye-blink detection performance analysis android platform. *Frontiers*, 5,4.
- [7] Vivek Anand, Vimal Singh Parihar, Shubham Kumar Sharma, Vikas Singhal4 and Pramod Kumar Sethy, (2021) “Facial recognition system using opencv”, *International Journal of Creative Research Thoughts*,2320-2882.
- [8] Roberto Daza, Aythami Morales, Julian Fierrez, Ruben Tolosane (2021) “Feasibility study of attention level estimation via blink detection applied to e-Learning”, *AAAI Workshop on Artificial Intelligence for Education (AI4EDU)*, 2021.
- [9] Scott Pfursich, David L. Graumann, Rahul Deva Ghosh, Ansuya Negi, Ranjit S Narjala (2018) “Facial liveness detection in image biometrics” Patent.
- [10] Ratha, N. K., Connell, J. H., & Bolle, R. M. (2001). An analysis of minutiae matching strength. *Audio and Video-Based Biometric Person Authentication (AVBPA)*, 2091, 223-228.
- [11] Rogmann, N., & Lee, M.K. (2015). Liveness detection in biometrics. In *International conference of the biometrics special interest group (BIOSIG)* (pp. 1-14).
- [12] Singh, S., & Prasad, S. V. A. V. (2018). Techniques and challenges of face recognition: A critical review. *Procedia Computer Science*, 143, 536-543.
- [13] Tang, D., Zhou, Z., Zhang, K. (2018). Face Flashing: A secure liveness detection protocol based on light reflections. In *Network and distributed systems security (NDSS) symposium*.
- [14] V. Anand, G. Krishna, G. Mohandass, R. Hemalatha, and S. Sundaram (2010) Predicting grade of prostate cancer using image analysis software, *Trends in Inform. Sciences and Computing (TISC)*. IEEE, pp 122–124.
- [15] Sheeba Santhosh, A Vimala Juliet, G Hari Krishnan. “Bio Impedance Signal Analysis with Variation in Input Signal Frequency”, *International Journal of Advanced Trends in Computer Science and Engineering*, Vol.9(4), pp.5332-5336 (2020).
- [16] 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).
- [17] 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).
- [18] Ganesan, U., Paul, N.E.E., Krishnan, G.H., Aarthi, S., Swamy, I. K., “Detecting Diabetes Mellitus from Tongue Image Hybrid Features and Neural Network Classifier”, *Proceedings of 4th International Conference on Cybernetics, Cognition and Machine Learning Applications, ICCCMLA2022*, 2022, pp.425-427.
- [19] Ganesan, U., Vimala, J.A., Hari, K.G., “Elbow Joints for Upper-Limb Prosthesis: Analysis of Biomedical EEG Signals using Discrete Wavelet Transform”, *International Journal of Engineering Trends and Technology* , 2022, 70(7), pp. 190–197.
- [20] Sudhakar, T., Hari Krishnan, G., Prem Kumar, J., ...Devanesan, P.S., Shalini, S., “Inducement of Artificial Sleep using Low Strength Magnetic Waves”, *Journal of Physics: Conference Series* , 2022, 2318(1), 012028.
- [21] Sudhakar, T., Hari Krishnan, G., Krishnamoorthy, N.R., ...Pradeepa, M., Raghavi, J.P., “Sleep Disorder Diagnosis using EEG based Deep Learning Techniques”, *Proceedings of 2021 IEEE 7th International Conference on Bio Signals, Images and Instrumentation, ICBSII 2021*, 2021, 9445158
- [22] Santhosh, S., Juliet, A.V., Krishnan, G.H., “Impact of electrodes separation distance on bio-impedance diagnosis”, 2021, 14(1), pp. 141–146.
- [23] Santhosh, S., Juliet, A.V., Krishnan, G.H., “Predictive analysis of identification and disease condition monitoring using bioimpedance data”, *Journal of Ambient Intelligence and Humanized Computing*, 2021, 12(2), pp. 2955–2963.

TRACK 3

**Ninth International Conference in Biosignals, Images and Instrumentation,
SSNCE, Chennai, March 16 – 17, 2023**

ANALYSIS OF RAW 3D IMAGES OF STAGES OF ALZHEIMER'S DISEASE USING DEEP LEARNING

Thamizhvani TR*

Department of Biomedical Engineering
Vels Institute of Science, Technology
and Advanced Studies
Chennai, India
thamizhrajeswari94@gmail.com

Swetha S

Department of Biomedical Engineering
Vels Institute of Science, Technology
and Advanced Studies
Chennai, India
19615115@velsuniv.ac.in

RJ Hemalatha

Department of Biomedical Engineering
Vels Institute of Science, Technology
and Advanced Studies
Chennai, India
hemalatharj@velsuniv.ac.in

Rachel Cynthia V

Department of Biomedical Engineering
Vels Institute of Science, Technology
and Advanced Studies
Chennai, India
19615124@velsuniv.ac.in

Abstract— Alzheimer's disease is defined as a brain ailment that progressively impairs thinking, cognitive skills and the capacity to complete most basic tasks. Memory loss, cognitive changes, and other neuronal brain disorders are all symptoms of AD, a degenerative condition. Since risk awareness encourages patients to take preventative measures even before the onset of irreversible brain damage, an absolute diagnosis of Alzheimer's disease is vital. Total brain atrophy and hippocampal atrophy are considered to be the main diagnostic tests for the condition. For this condition, early identification is important, and automatic system design is required. Computer-assisted methods are implemented for the analysis of AD in several types of research and the outcomes are constrained due to the congenital findings. Early stages of AD can be diagnosed but not predicted because prediction is only useful before the disease manifests itself. Deep learning (DI) techniques are used to analyze the raw MRI 3D images to identify AD and its progressive stages. The efficiency of the deep learning networks is defined to be less, which is indefinite for diagnosing AD stages. Therefore, processing of the images is necessary for the detection and to predict the progression of AD.

Keywords— Deep learning, Magnetic Resonance Imaging, RES NET, Proposed DEEP NET.

I. INTRODUCTION

Alzheimer's disease (AD) is a disorder in the brain that develops a decline in cognitive activities [1]. This neuronal disorder worsens over time affecting the functions of the brain. AD is becoming a predominant cause of dementia in the aged population, affecting more than 30 million people worldwide. Commonly, 65 years of age or older population may have AD and the ratio is even higher in developing countries [2]. The early symptoms begin with the knowledge of recollecting the newly learned information because this disorder primarily affects the learning phase of the brain. As the disease advances in stage, the symptoms increase and affect regions related to disorientation, language, mood swings, loss of motivation, not managing self-care, and behavior. The primary symptoms formed in this disorder are always co-related with stress and age factor. With the decline in the patient's condition, they detach from social and emotional contacts and slowly leading to a gradual decrease in their physiological functions.

The pathology of the disorder and stages can be analyzed briefly using these advanced imaging modalities such as Computed tomography (CT), Magnetic Resonance Imaging (MRI), Single photon emission tomography (SPECT), and Positron emission tomography (PET). Potential biomarkers were defined and developed for the improved diagnosis and analysis of the disease. The pattern of evolution of stages of Alzheimer's can also be defined using different biomarkers like Cerebro Spinal Fluid (CSF), molecular biomarkers, hippocampus region biomarkers, and neuroimaging biomarkers. These biomarkers are specially defined with two advanced imaging modalities like MRI and PET. MRI describes the structural information of the brain with which the stages of Alzheimer's are analyzed from cognitively normal (CN) to AD. Imaging biomarkers for AD with neurodegeneration are widely available in the MRI which is non-invasive and more cost-effective compared to Positron Emission Tomography (PET) [3]. The computer-assisted design for the analysis of AD and progression is being under study and is necessary for early diagnosis and treatment.

Braak H et.al., (1997) defines the frequency-based stages in Alzheimer's disease depend on the intracellular metabolisms and structural degree of destruction of the brain atrophy [4]. Imaging defines the atrophy of the brain in a precise manner. L.shen et.al., (2020) describe the integrated analysis of the images using techniques of machine learning and deep learning. Brain imaging genomics is studied with the help of normal and machine learning techniques. Thus, algorithms related to machine and deep learning were more efficient and appropriate for brain imaging genomics analysis [5].

In the proposed model the database images of the stages of Alzheimer's disease (AD) are derived from the Alzheimer's disease neuroimaging initiative database. The images are subjected to the reorientation process. The analysis of the raw data for the detection of Alzheimer's disease was performed using different deep-learning algorithms or networks. The network proposed is specifically defined for the analysis of the brain with the increased convolution and regression layers. The efficiency of the network range is defined to be in the range of less for the effective description of progressive stages of AD. Hence t

processing of the 3D images is necessary for biomarker detection. A deep learning network is proposed for the analysis of raw 3D images to describe AD and its stages. The 3D raw MRI image of AD used for the process of classification is illustrated in figure 1.

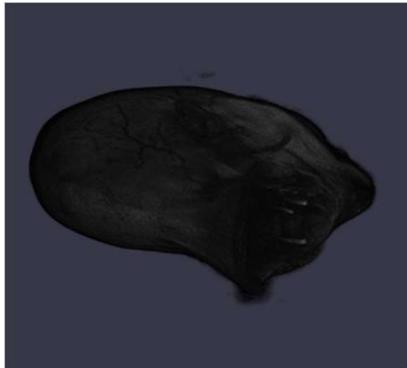


Fig.1. Raw 3D MRI image of Alzheimer's Disease

II. METHODOLOGY

Though there are many proposed algorithms to classify the different stages of Alzheimer's disease, into Normal Cognitive (CN), Mild Cognitive Impairment (MCI), Late Cognitive Impairment (LMCI), and Early Cognitive Impairment (EMCI) in the literature, and even a few approaches were made to determine the transition stages of Alzheimer's Disease. In the proposed method, the images are subjected to the reorientation process. The 3D MRI images (raw data) are analyzed using deep learning networks for the identification and grading of Alzheimer's disease stages. The efficiency of the networks in classifying and identifying the stages of AD is defined to be less than 90%. Therefore, image processing of 3D MRI images is necessary to define the different stages of AD and to determine the cognitive decline of the brain. The process flow is illustrated in figure 2.

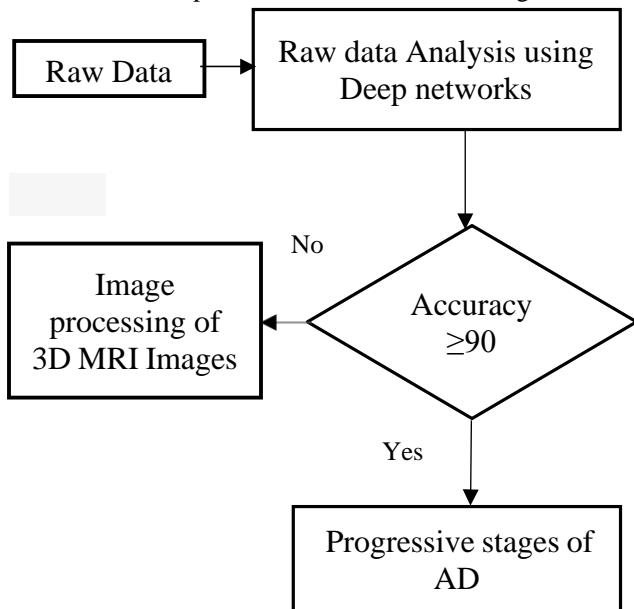


Fig. 2. Process flow of the proposed method

A. Image Acquisition

3D Magnetic Resonance images of different stages of AD used in the proposed method were derived from the Alzheimer's disease neuroimaging initiative (ADNI)* database. ADNI database is a data repository maintained by the laboratory of neuroimaging (LONI) at the university of southern California, the LONI image & data archive (IDA) with more than 2000 images of different brain imaging modalities like MRI, PET, DTI and so on. Images of different transitional stages such as CN, MCI, EMCI, LMCI and AD of the neurological disorder are available in the database. Magnetic resonance imaging (MRI) of brain structures obtained using 1.5 & 3 Teslas are stored in the repository for different age groups especially people above 45 [6].

B. Raw Data Analysis

In the proposed methodology, the source images were obtained from the ADNI database and where the images are subjected to the reorientation process to configure the images to the standard Montreal Neuro-Imaging (MNI) template. The raw data images without processing are provided as input to the deep learning networks for analysis. Different deep learning networks such as UNET, RES-NET-18, RES-NET-50, CNN, and LE-NET for the analysis of the raw data. The training of the data is performed with 100 raw 3D images of each stage of AD used for the process of analysis and classification. Batch processing and normalization of the data are performed for the multiclass classification process. Multiclass differentiation is performed in which four different stages of AD are under consideration for the classification process. The efficiency of the deep neural networks in the process of classification of different stages of AD is determined using ROC Curve and its features. Since the accuracy value of the different stages of AD is less than 65%, determining that the networks implemented form an indefinite system for the accurate difference of the progressive phases of AD. With the help of the results, image processing algorithms have to be implemented for the process of identification of stages of Alzheimer's Disease such as EMCI, LMCI, CN, and Severe AD. The raw 3D MRI images of the stages of AD are analyzed using various predefined deep-learning networks. A proposed deep-learning network is also defined for the determination of the progressive phases of AD. The implementation and layers of the networks used are described below.

Convolutional Neural Network (CNN)

Convolutional networks are comprised of many hidden units, input nodes, and output units. Contrary to a typical neural network, the layers of neurons of a convolutional network are arranged in three dimensions (width, height, and depth dimensions) [7]. The CNN can now transform a three-dimensional input volume into an output volume. The hidden layers are composed of convolution, pooling, normalizing, and fully connected layers. Three layers namely, convolutional, pooling, and fully connected layers are implemented for the deep learning network. CNN processes data using a grid-like architecture and belongs to the class of neural networks. The foundational component of CNN that is primarily responsible for computation is the convolution layer.

U-NET

The CNN architecture underwent minimal alterations to accommodate the U-NET convolutional neural network architecture. A bottleneck is intended to be built in the middle of 3D-expanding U-NETs and contracting paths by combining convolution and pooling procedures. Following this bottleneck, the image is rebuilt using a convolutional and up-sampling technique [8].

RES-NET- 18

ResNet18 has an architecture with 18 deep layers. The architecture of this network is to enable the efficient operation of several convolutional layers. There are two main variations in this model set [9]. The first formulation, known as mixed convolution (MC), uses 2D convolutions in the top layers and 3D convolutions exclusively in the network's initial layers. The results usually worsen when more deep layers are added to a network.

RES-NET -50

ResNet-50, a 50-layer network with 48 convolutional layers, a Max Pool layer, and an average pool layer, is a type of convolutional neural network. Artificial neural networks (ANNs) that use residual blocks to build networks are known as residual neural networks [10]. With the increasing layer, the RES NET networks can be designed and used for the process of classification.

LE-NET

There are seven layers in the LeNet-5 CNN architecture. The layer composition consists of three convolution layers, two subsampling layers, and two fully connected layers. Le-Net has the fundamental components of a convolutional neural network, including the convolution layers, pooling layer, and complete connection layer, establishing the groundwork for its future growth [11]. They used this architecture for recognizing the handwritten and machine-printed characters.

Proposed DEEP NET

The proposed deep learning network is defined with the 3D input layer with the image size to be depending on the input image in 3D. A layer of 3D convolution is defined next to the input layer with padding of voxels as a structural feature. After the convolution of padding, an average pooling layer in 3D for the voxel's description is designed. Each layer of the network is interlinked with a hyperbolic tangent layer designed for the coordinate functioning of the layers. A set of 25 convolution and average pooling is again designed with a factor in reduced axes size. The average pooling layer describes the voxel's features in the neighborhood and determines the functional changes in the entire region of the brain [12]. Then a convolution layer in 3D is described for the reduced voxel description for the determination of the units of the image. Two fully connected layers with the discrimination of the four classes are defined for the classification process. The output classification layer is defined with a four-way (multiclass) classification process. A new network designed specifically for the 3D MRI brain images of different stages of AD. Classification process efficiency is also determined using the ROC and its features.

The proposed deep network layers are defined below in figure 3.

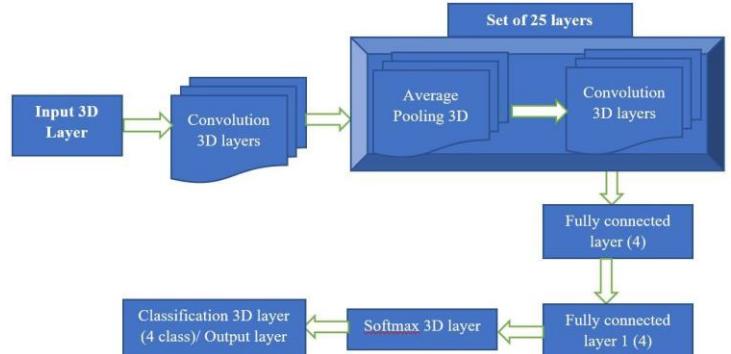


Fig. 3. Layers of proposed deep network

The efficiency of the networks is determined by the description of the Receiver operating characteristic curve (ROC) and the features derived from the ROC curve [13]. The network's training and testing progress is determined by the accuracy, precision, sensitivity, and specificity of the network [14]. The efficiency of the deep learning networks is defined to be less than 65% stating that any network design requires the processing of the images using imaging algorithms. For the process of classification of different stages of AD, descriptive features of the voxels in the regions of atrophy are definite. Therefore, 3D image processing of the raw MRI images of the different stages of AD is performed to determine the biomarker and to study progressive states of cognition in different stages of AD.

III. RESULTS AND DISCUSSION

In the proposed methodology, the raw images of the stages of Alzheimer's disease are analyzed and studied by comparing the hippocampus region segmented. The primary step in this methodology proposed the re-orientation of images to the standard MNI template. After reorientation, the unprocessed images in 3D are subjected to classification and analysis to detect the stages of AD. Deep learning networks are implemented for the determination of progressive phases of AD. The new net is proposed for the classification of raw 3D MRI images.

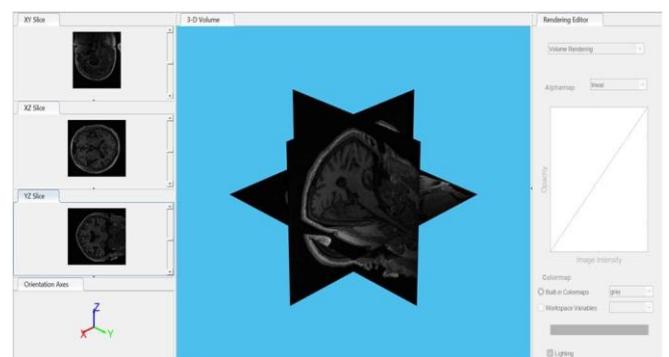


Fig. 4. Reorientation of the 3D MRI images

The orientation process describes the atrophy of the brain regions and biomarker, the hippocampus more effectively when compared with other standard templates for brain structures. The reoriented image is illustrated in above figure 4. Different deep-learning techniques are used for the description of stages of AD with the help of raw data. Predefined networks like RESNET, LENET, and UNET are used for the process of classification of raw images. The training progress and testing phase of the networks are determined with the help of the progression curve. The progression curve with epochs defined for the net proposed is defined below in figure 5.

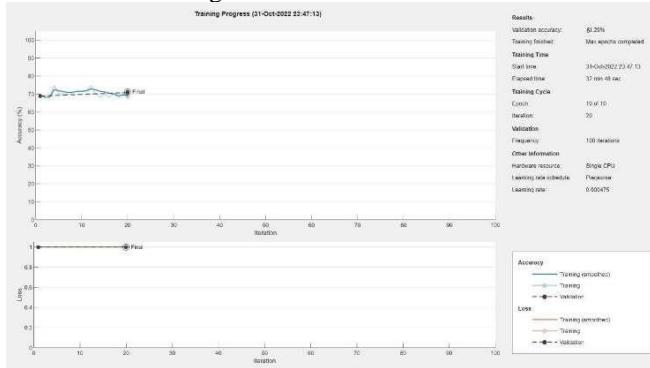


Fig. 5. Progression curve of proposed deep network

The ROC curve can be used to describe the classification and diagnosis of AD phases and evaluate the effectiveness of deep learning algorithms. The ROC is illustrated in figure 6 given below.

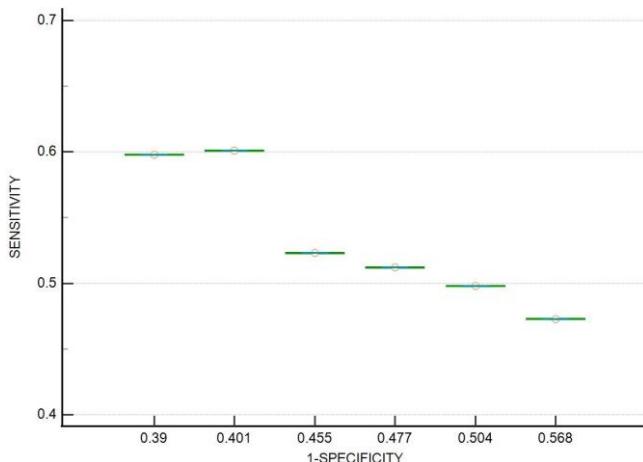


Fig. 6. ROC Curve for the different deep learning networks

The ROC defined for each network's training and testing process provides a clear vision of the efficiency and the probabilistic error of the networks used for the determination of the progression of AD and its stages. The features derived from the ROC curve are illustrated in the table given below in Table 1.

TABLE I. CLASSIFICATION FEATURES OF ROC CURVE FOR DIFFERENT DEEP LEARNING ALGORITHMS

RAW 3D IMAGES ANALYSIS (CLASSIFICATION FOR STAGES OF AD)					
Type of Deep learning algorithm	Accuracy (%)	Sensitivity	Specificity	Error	Precision
RESNET - 18	51.2	0.512	0.523	0.488	0.512
RESNET - 50	55.76	0.598	0.61	0.445	0.553
LENET	52.4	0.523	0.545	0.476	0.501
UNET	49.4	0.498	0.496	0.506	0.485
CNN	47.3	0.473	0.432	0.527	0.471
Proposed Net	61.29	0.601	0.599	0.388	0.612

The network error is defined to be more than the region-based classification because the entire 3D images with unprocessed voxel range values are set as input to the networks. The classification of stages of AD is not more definite due to the uneven distribution of voxels and inappropriate feature descriptions of the networks with the overall atrophy of the brain structures. Deep learning algorithms mainly consider descriptive features depending on the layers implemented. Even for a precise layer of networks, the feature description is not more efficient to classify the stages of AD and there is no definite descriptive structure for the deep network to analyze the atrophy of regions and identify the stages of AD.

The processing of the images is necessary as with the previous research and discussions, a computer-aided system or any biomarker definition requires the processing of images. In a few images, the features are clearly defined using deep learning networks with specific designs for the classification and identification of the state of the images. Sarraf et.al., (2019) describes the recognition of the stages of cognitive impairment with the images of descriptive fMRI and MRI comparison model using neural networks in topology structures [15]. Gupta Y et.al., (2019) defines the process of early diagnosis of AD using descriptive voxel-based features combined from the cortical and hippocampus regions, so the determination of these regions defines the state of AD [16]. The region description and analysis were not more definite when raw data from 3D MRI is produced as input for the deep learning networks. The atrophy of the region based on the feature description as a whole is not so very clear. In order to further identify the progression through the stages of AD and to anticipate the brain's cognitive state, image-processing techniques must be applied. Pre-processing, segmentation of the biomarker region and classification of the stages using transfer learning are the future works to be implemented for the early diagnosis of AD and the progression of the stages.

IV. CONCLUSION

The human brain is divided into several different areas, each of which has a specific function that enables the active functioning of the human body. Alzheimer's Disease (AD) is defined as a form of cognitive decline of the human brain. The brain undergoes both anatomical and functional changes. Atrophy of the structural regions in the brain describes the

presence of AD and helps in the categorization of stages of AD [17]. Three-dimensional magnetic resonance imaging is derived from the Alzheimer's disease Neuroimaging Initiative (ADNI) database (<http://adni.loni.usc.edu>). The proposed method is raw data analysis of the 3D MRI images using deep learning algorithms that help to discriminate the different stages of Alzheimer's disease such as cognitively normal (CN), Late mild cognitive impairment (LMCI), Early mild cognitive impairment (EMCI) and Severe AD. Different Deep learning networks or algorithms such as U-NET, RESNET-18, LENET, and proposed DEEP NET are implemented for the process of classification of AD and its stages. The Comparison of the epochs and analysis of efficiency is determined using the ROC curve and its features. With this process implementation, the analysis of raw data using deep learning is defined to be less significant, in classifying the stages of AD. Thus, the application of image processing techniques or algorithms is required for the determination of the stages of AD and to describe the progressive variations in brain cognition.

REFERENCES

- [1] Crane, P.K., Carle, A. and Gibbons, L.E., "Development and assessment of a composite score for memory in the Alzheimer's Disease Neuroimaging Initiative (ADNI)", *Brain Imaging and Behavior*, vol. 6, pp. 502–516, 2012. <https://doi.org/10.1007/s11682-012-9186-z>
- [2] Vemuri, P. and Jack, C.R., "Role of structural MRI in Alzheimer's disease", *Alz Res Therapy*, vol. 2, pp. 23, 2010. <https://doi.org/10.1186/alzrt47>
- [3] Beheshti I., Maiusa N., Daneshmand M., Matsuda H., Demirel H., and Anbarjafari G, "Classification of Alzheimer's disease and prediction of mild cognitive impairment conversion using histogram-based analysis of patient-specific anatomical brain connectivity networks", *J. Alzheimer's disease*, vol. 60, pp. 295–304, 2017.
- [4] Braak H. and Braak E, "Frequency of stages of Alzheimer-related lesions in different age categories", *Neurobiol. Aging*, vol. 18, pp. 351–357, (1997).
- [5] L. Shen and P. M. Thompson, "Brain Imaging Genomics: Integrated Analysis and Machine Learning, In Proceedings of the IEEE, vol. 108, pp. 125-162, Jan. 2020, doi: 10.1109/JPROC.2019.2947272.
- [6] E.E. Bron, et al., "Standardized evaluation of algorithms for computer-aided diagnosis of dementia based on structural MRI: the CAD dementia challenge", *Neuroimage*, vol. 111, pp. 562-579, 2015.
- [7] Junhao Wen, Elina Thibeau-Sutre, Mauricio Diaz-Melo, Jorge Samper-González, Alexandre Routier, Simona Bottani, Didier Dormont, Stanley Durrleman, Ninon Burgos and Olivier Colliot, "Convolutional neural networks for classification of Alzheimer's disease: Overview and reproducible evaluation", *Medical Image Analysis*, vol. 63, 2020, 101694.
- [8] Islam, M., Vibashan, V.S., Jose, V.J.M., Wijethilake, N., Utkarsh, U. and Ren, H., "Brain Tumor Segmentation and Survival Prediction Using 3D Attention UNet", In: Crimi, A., Bakas, S. (eds) Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. Lecture Notes in Computer Science, Springer, Cham, vol. 11992, pp.262-272, 2020. https://doi.org/10.1007/978-3-030-46640-4_25.
- [9] A. Ebrahimi, S. Luo and R. Chiong, "Introducing Transfer Learning to 3D ResNet-18 for Alzheimer's Disease Detection on MRI Images," 2020 35th International Conference on Image and Vision Computing New Zealand (IVCNZ), Wellington, New Zealand, 2020, pp. 1-6, doi: 10.1109/IVCNZ51579.2020.9290616.
- [10] Fulton LV, Dolezel D, Harrop J, Yan Y and Fulton CP, "Classification of Alzheimer's Disease with and without Imagery Using Gradient Boosted Machines and ResNet-50", *Brain Sciences*, vol. 9(9), pp. 212, 2019. <https://doi.org/10.3390/brainsci9090212>.
- [11] Taihao Li, Di Jin, Cuifen Du, Xiumei Cao, Haige Chen, Jianshe Yan, Na Chen, Zhenyi Chen, Zhenzhen Feng and Shupeng Liu, "The image-based analysis and classification of urine sediments using a LeNet-5 neural network", *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, vol. 8, pp. 109-114, 2020. DOI: 10.1080/21681163.2019.1608307
- [12] Tong, Z., Aihara, K. and Tanaka, G., "A Hybrid Pooling Method for Convolutional Neural Networks", In: Hirose, A., Ozawa, S., Doya, K., Ikeda, K., Lee, M., Liu, D. (eds) Neural Information Processing. ICONIP 2016, Lecture Notes in Computer Science, Springer, Cham, vol. 9948, pp. 454-46, 2016. https://doi.org/10.1007/978-3-319-46672-9_51
- [13] Tonozuka R, Itoi T, Nagata N, Kojima H, Sofuni A, Tsuchiya T, Ishii K, Tanaka R, Nagakawa Y and Mukai S, "Deep learning analysis for the detection of pancreatic cancer on endosonographic images: a pilot study", *J Hepatobiliary Pancreat Sci*, vol. 28(1), pp. 95-104, Jan 2021. doi: 10.1002/jhbp.825. Epub 2020 Oct 15. PMID: 32910528.
- [14] Litjens G, Sánchez CI, Timofeeva N, Hermans M, Nagtegaal I, Kovacs I, Hulsbergen-van de Kaa C, Bult P, van Ginneken B and van der Laak J., "Deep learning as a tool for increased accuracy and efficiency of histopathological diagnosis", *Sci Rep*, vol. 23(6), 26286, May 2016. doi: 10.1038/srep26286. PMID: 27212078; PMCID: PMC4876324.
- [15] S. Sarraf D. D. Desouza J. A. E. Anderson and C. Saverino, "MCADNNNet: Recognizing stages of cognitive impairment through efficient convolutional fMRI and MRI neural network topology models", *IEEE Access*, vol. 7, pp. 155584-155600, 2019.
- [16] Y. Gupta K. H. Lee K. Y. Choi J. J. Lee B. C. Kim and G. R. Kwon, "Early diagnosis of Alzheimer's disease using combined features from voxel-based morphometry and cortical-subcortical and hippocampus regions of MRI t1 brain images", *PLoS ONE*, vol. 14(10), e0222446, Oct. 2019.
- [17] Moradi E, Pepe A, Gaser C, Huttunen H and Tohka J, "Alzheimer's Disease Neuroimaging Initiative. Machine learning framework for early MRI-based Alzheimer's conversion prediction in MCI subjects", *Neuroimage*, vol. 104, pp. 398-412, 2015. doi: 10.1016/j.neuroimage.2014.10.002.

A Deep Learning Framework for Semantic Segmentation of Nucleus for Acute Lymphoblastic Leukemia Detection

Prasanna.A
Department of
Electronics Engineering
Madras institute of
Technology
Anna University
Chennai, India
ashokprasanna01@gmai
l.com

Saran.S
Department of
Electronics Engineering
Madras institute of
Technology
Anna University
Chennai, India
saran.sevvel2001@gma
il.com

Manoj.N.M
Department of
Electronics Engineering
Madras institute of
Technology
Anna University
Chennai, India
manojms4614@gmail.c
om

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

Abstract

Acute lymphoblastic leukemia is a form of blood cancer in which the bone marrow overproduces immature white blood cells. A novel semantic segmentation of nucleus for detection of Acute lymphoblastic leukemia is proposed here. The input images are obtained from public database “ALL-IDB2”. Resizing, SMOTE and Augmentation are carried out as preprocessing. After pre-processing, segmentation of nucleus is performed by SegNet and ResUNet. The performance of SegNet and ResUNet are compared. The segmented images are given as input to the classification models. Using Xception, Inception-v3 and ResNet50 models, the segmented images are classified as healthy and blast cells. It is found that Inception-v3 performs better than Xception and ResNet50 with an accuracy of 93.74%. This will be helpful to detect Acute lymphoblastic leukemia at the earliest.

Keywords: CNN, SegNet, ResUNet, Inception, Xception, ResNet-50

I. INTRODUCTION

Leukemia is a blood cancer that affects the reproduction of white blood cells (WBC) in the bone marrow. It causes an increase in the number of abnormal WBCs, which leads to a decrease in immunity [1]. Leukemia is divided into acute and chronic variants according to how quickly the disease progresses. Chronic leukaemia progresses

slowly over time, and the more mature leukocytes can perform some of their usual tasks. In contrast to acute leukaemia, which develops swiftly and shows a significant rise in leukemic cells, chronic leukemia develops over time. Leukemia is further split into myelogenous and lymphoid types depending on the type of afflicted cell from which the malignancy arises [2].

Proliferation of immature white blood cells and accumulation of abnormal cells cause ALL. This type of cancer progresses rapidly, replacing functional lymphocytes with inoperable leukemic cells. Accurate and early detection of these cells are important. It is a common cancer in children aged 0 to 14 years. The survival rate for children is only 65.3%.

The objective of the work is to propose a novel segmentation technique for separating the nucleus from acute lymphoblastic leukemia cells. Deep learning networks are suggested for classification of healthy and blast cells. The major contribution is segmenting the nucleus for better classification. The survival rate of cancer patients improved significantly if early detection of the cancer is done. The accurate predictions are of utmost priority.

The paper is organized as follows: Section 1 covers the detailed overview of Leukemia. Section 2 presents the literature survey that has been done for the purpose of this paper.

Section 3 covers the details of proposed methodology with image pre-processing, segmentation, and classification. Section 4 reveals the implementation and results obtained in this paper. The conclusion and future scope of the paper is given in section 5 and references are given in section 6.

II. LITERATURE REVIEW

Shallu Sharma [3] demonstrated a pre-trained Xception model to classify breast cancer histopathological images based on magnification, in contrast to manual methods. Stable performance was best for the Xception model and the SVM classifier with a "radial basis function" kernel.

Rohan Khandekar [4] proposed an AI model for the detection of all blast cells. The YOLOv4 algorithm is used for blast cell classification and detection from blood cell micrographs. Images from the two datasets ALL_IDB1 and C_NMC_2019 were used to train the model.

Xiaodong Chena [5] suggested a lung cancer detection system based on lung computed tomography (CT) image scores utilising the SegNet approach model. SegNet is used in the model to segment images, and it compares the Deeplab v3, VGG 19, and SegNet methods used for manual segmentation of lung cancer images for accuracy, sensitivity, total image segmentation time, specificity, and overlap rate.

Foivos I. Diakogiannis [6] proposed a solid basis for the effective resolution of the semantic segmentation issue posed by a single high-resolution aerial photograph. ResUNet, a brand-new deep learning architecture, and a brand-new loss function based on dice loss make up the framework. ResUNet reconstructs colored input images and locates boundaries. The ISPRS 2D Potsdam dataset is used to evaluate the simulation system's performance.

Ketan Joshi [7] proposed a work that used Inception V3 to provide a solid framework for classifying sports images according to their environment and related environments. The performance of the neural network is contrasted with that of other classifiers such as Random Forest, KNN, and SVM.

Chayan Mondal [8] proposed a weighted ensemble model for ALL classification. In the preliminary test set, has produced a weighted F1-score of 88.6%, a balanced accuracy of 86.2%, and an AUC of 0.941 by using the ensemble candidates' kappa values as their weights.

III. PROPOSED METHOD

An effective approach for ALL cell detection is proposed in this paper. The ALLIDB2 collection contains 130 microscopic pictures of healthy white blood cells and ALL cells [9]. The block diagram of the proposed work is shown in Fig. 1. The performance of the proposed system is analysed for both segmentation and classification. The methodology of the proposed work is discussed below.

A. Image pre-processing

Pre-processing performs series of operations to transform or change data. It is done to make raw data suitable for the learning model and prevents the model from overfitting. In the proposed system, image resizing, SMOTE and augmentation are done. The input images are resized to 256 x 256 dimension as the network accepts input of this much size. SMOTE is a data augmentation technique by randomly selecting a minority class data point and its k nearest minority class neighbours, and then generating new samples by interpolating between these points in the feature space.

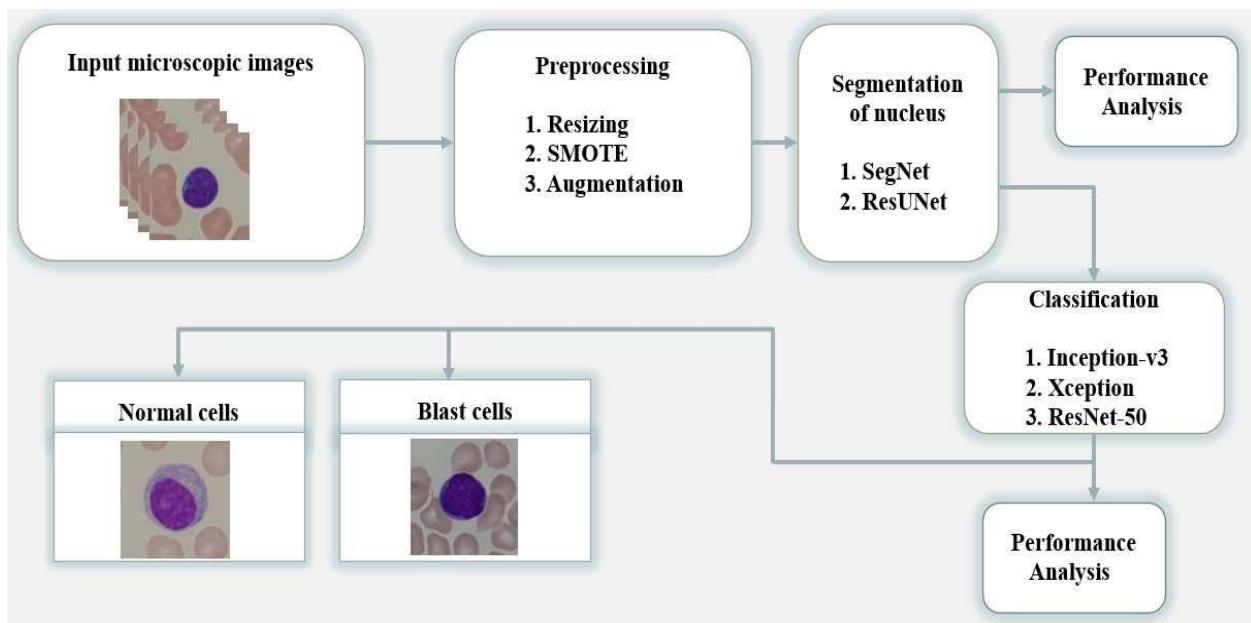


Fig. 1. Block diagram of the proposed work

Augmentation is done to prevent the model from overfitting where the results from SMOTE is used to create a new set of images by rotating clockwise by 90 degree and flipping it horizontally.

B. Image segmentation

Segmentation of cell nucleus is done by SegNet and ResUNet. SegNet is a semantic segmentation model that consists of an encoder and decoder network, then a classification layer at the pixel level. A deep learning architecture called ResUNet is developed using residual connections and UNet. To reduce the number of parameters and enhance segmentation accuracy, the ResUNet architecture incorporates residual connections into UNet.

C. Classification

Inception V3 model, Xception model and ResNet-50 are used for classification of ALL cells. Inception-v3 model is made up of both symmetric and asymmetric building blocks. The inception modules used in each stage reduces the number of parameters and channels by replacing filters at each stage by a filter with smaller size. Xception is an extension of the Inception architecture that

uses depthwise separable convolutions to reduce the number of parameters and computational complexity of the network. ResNet-50 includes 50-layers trained using residual connections, which are connections between layers of a neural network that help prevent the network from becoming too deep and complex.

IV. RESULTS AND DISCUSSION

In the proposed work, the acquired database is increased by pre-processing techniques. The nucleus is segmented using deep learning technique and the obtained results are carried out to classification for the detection of cancer cells. The required images are obtained from ALL-IDB2 database consists of two classes such as Cancer cells and healthy cells with 130 microscopic blood cell images in each class. Each has image dimension of 257 x 257. The Fig. 2 displays the sample images used in this project. By pre-processing the undesired pixels were removed. Pre-processing the images include Resizing, SMOTE, Augmentation. Image resizing is necessary to ensure uniformity and to fit the images suitable for segmentation model the images are resized to 256 x 256 dimension.

Type Of Cells	Sample Images		
Healthy Cells			
Cancer Cells			

Fig. 2. Sample images from ALL IDB2 dataset

SMOTE generates synthetic images by oversampling the minority class samples and generates minority class images to match the majority class. Synthetic images are artificial images generated by SMOTE. Augmentation is done with Rotation by 90 and horizontal flip. The new set of images are generated. Further the results are given into segmentation. The Table 1 shows the number of images after each pre-processing stages. After SMOTE, the number of healthy and blast images have been increased from 130 to 243. After Augmentation, the number of healthy and blast images have been increased from 243 to 729. Therefore, the total number of images is 1458 is utilized for the proposed work. The segmented results of both Healthy and Cancer cells by SegNet and ResUNet with comparison of ground truth images are tabulated and shown in the Fig. 3. In predicted image, the white region is classified as the nucleus of the cell by the model. The input image is merged with the

predicted image to the form the segmented image. The comparison between the SegNet and ResUNet are better understandable with the values of the performance metrics is tabulated in Table 2. Dice coefficient as in (1) used to evaluate the performance of a segmentation model ranges between 0 and 1, where P is the predicted result and G is the ground truth result.

$$\text{Dice coefficient} = \frac{2 * |P \cap G|}{|P| + |G|} \quad (1)$$

The hyperparameters such as learning rate of 0.00001, epoch of 10, batch size of 16 are used for SegNet and a learning rate of 0.00001, epoch of 15, batch size of 16 are used for ResUNet. ResUNet model have attained a maximum accuracy of 98.36 % while SegNet attained accuracy of 96.99%. The training and validation accuracy plot is shown in Fig.4. The segmentation is also compared by evaluating the dice coefficient values which showed that ResUNet model gives better results. The results from ResUNet are further used for classification purposes. ResUNet is designed to get high performance with fewer parameters. SegNet can conserve memory by switching the pooling indices from the compression path to the expansion path.

Table 1. Number of images after pre-processing

Type	Sample Images	After SMOTE	After Augmentation
Healthy	130	243	729
Blast	130	243	729

Table 2. Performance comparison between SegNet and ResUNet

Model	Accuracy	Validation Accuracy	Loss	Validation Loss	Dice Coefficient
SegNet	97.3%	96.99%	6.6%	7.49%	97.4%
ResUNet	98.26%	98.36%	3.6%	3.3%	98.2%

TYPE OF CELL	INPUT IMAGES	GROUND TRUTH	PREDICTED IMAGES BY SegNet	SEGMENTED IMAGES BY SegNet	PREDICTED IMAGES BY ResUNet	SEGMENTED IMAGES BY ResUNet
HEALTHY CELLS						
HEALTHY CELLS						
CANCER CELLS						
CANCER CELLS						

Fig. 3. Segmented images by SegNet and ResUNet

As opposed to ResUNet, which moves entire feature maps from the compression path to the expansion path. ResUNet's segmented images are fed into the ResNet-50, Inception-v3, and Xception models. The Table 3 shows the comparison between Inception-v3, Xception and ResNet-50 in terms of the accuracy and loss. The results convey that Inception-v3 has better results compared to Xception and ResNet-50 attaining an accuracy of 93.74%. The hyperparameters such as learning rate of 0.00001, epoch of 30, batch size of 32 are used for Inception-v3, Xception and ResNet50. The training and validation accuracy plot of Inception-v3, Xception and ResNet50 are plotted in Fig. 5. Xception can

be seen saturated after 5 epochs and Inception-v3 has been slightly increasing on each epoch. The Table 4 depicts the performance of Inception-v3, Xception and ResNet-50. Precision, Recall, and F1 score are utilised as performance metrics for comparison. According to the performance measures of Inception V3, Xception, and ResNet-50, Inception V3 performs better than ResNet-50 and Xception. As Inception-v3 has deeper networks compared to ResNet-50. It is computationally less expensive as it uses auxiliary classifiers as regularizer. ResNet-50 network is too shallow as the learning might be very inefficient and the detection of errors becomes difficult.

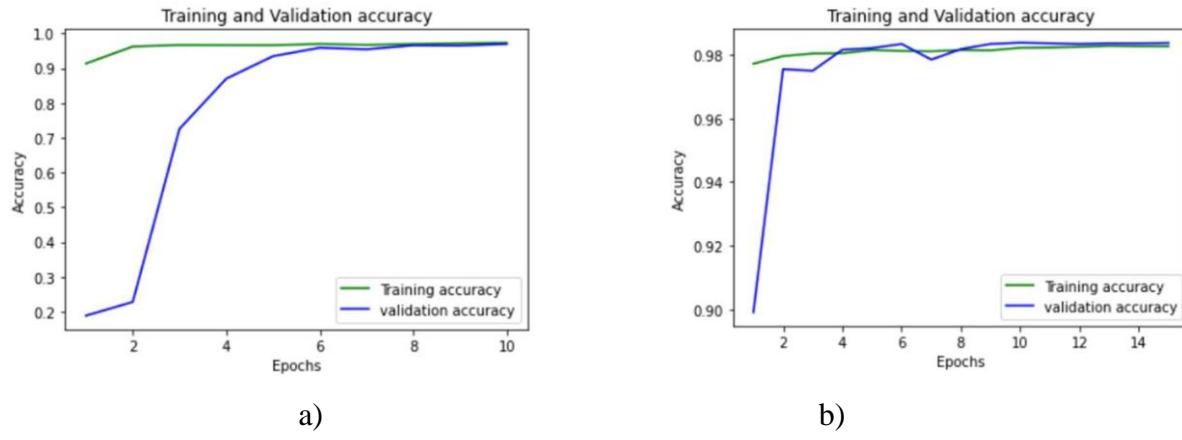


Fig. 4. Training and validation accuracy of a) SegNet and b) ResUNet

Table 3. Performance comparison between Inception-v3, Xception and ResNet-50

MODEL	ACCURACY	VALIDATION ACCURACY	LOSS	VALIDATION LOSS
Inception-v3	93.74%	90.74%	25.77%	28.80%
Xception	90.39%	89.38%	26.97%	28.81%
ResNet-50	72.81%	71.92%	56.42%	55.59%

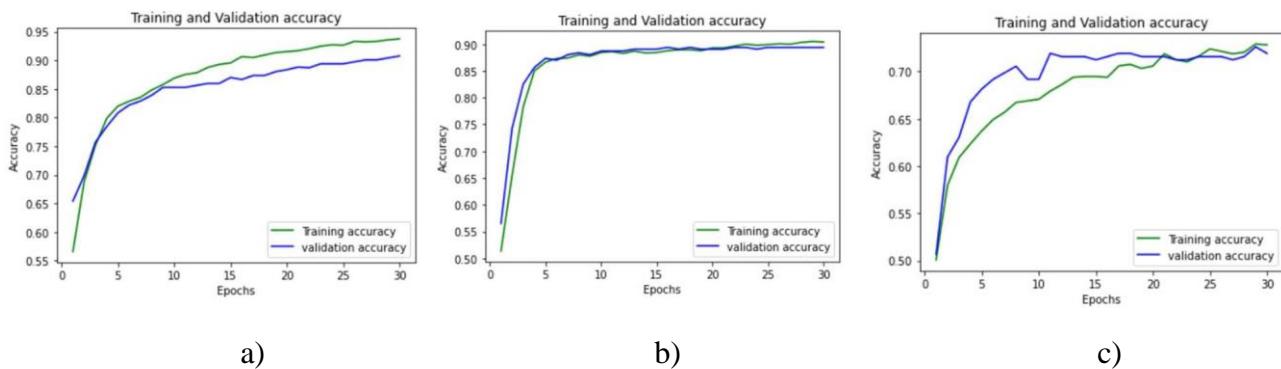


Fig. 5. Training and validation accuracy of a) Inception-v3, b) Xception and c) ResNet50

The confusion matrix of Inception-v3, Xception and ResNet50 are shown in Fig. 6. Inception-v3 model has attained a maximum accuracy of 93.74% while Xception attained 90.39% and ResNet-50 attained 72.81%. The performance measured used to evaluate the efficiency are precision, sensitivity, F1 score and specificity. Sensitivity as defined

in (2) is used to measure the ability of a model to correctly identify positive cases. Specificity as in (3) is a measure of the ability of a model to correctly identify negative cases, where W as true positive, X as true negative, Y as false positive and Z as false negative.

$$\text{Sensitivity} = \frac{W}{W+Z} \quad (2)$$

$$\text{Specificity} = \frac{X}{X+Y} \quad (3)$$

Inception V3 model overall performance in all metrics is comparatively higher than Xception and ResNet-50. The proposed work will be useful to identify the blast cells more accurately.

Table 4. Performance analysis of Inception-v3, Xception and ResNet-50

Model	Precision	Sensitivity	F1 score	Specificity
Inception-v3	93.1%	93.5%	93.3%	93.14%
Xception	88.6%	92.3%	90.4%	88.2%
ResNet-50	73.55%	71.33%	72.42%	74.3%

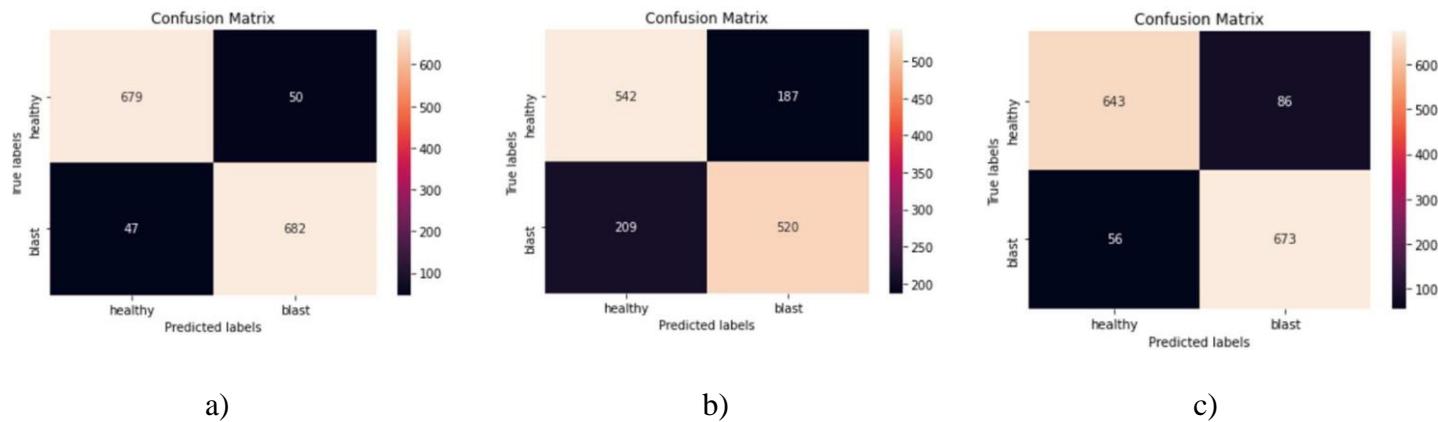


Fig. 6. Confusion matrix of a) Inception-v3, b) Xception and c) ResNet50

Table 5. Performance comparison of proposed method with some existing method

Authors	Methodology	Accuracy
Sahol etal., [10]	Texture + EHO-NN	91.8%
Sonali Mishra etal., [11]	DCT + SVM-L	89.76%
Chayan Mondal etal., [8]	Xception, DenseNet-121, MobileNet, VGG-16 and InceptionResNet-V2	88.8%
Abdeldaim etal., [12]	Shape, color, and texture + SVM(RBF)	93.41%
Proposed method	Inception-v3, Xception and ResNet-50	93.74%

Shape, color and texture features are extracted with a classification accuracy of 91.8% [10]. the color conversion and

thresholding based segmentation method and different classifiers are used to get best classification accuracy of 93.41% [12]. The proposed method has high performance with an accuracy of 93.74% over other existing methods is listed in Table 5.

V. CONCLUSION

The most prevalent type of malignancy in white blood cells is acute lymphoblastic leukaemia. The immature cancer cells will progress rapidly. The timely prognosis and treatment are need of time. The main objective of the work is to create a highly accurate diagnostic approach for acute lymphocytic leukaemia. For this purpose, around 1458 images are utilized. All the images are pre-processed by using different methods such as Resizing, SMOTE and

Augmentation. All the pre-processed results are given to segmentation which is carried out by SegNet and ResUNet. The ResUNet outperforms then SegNet with an accuracy of 98.26%. Using Inception-v3, Xception, and ResNet-50, three deep learning approaches are used to classify the segmented data. And their performance is compared to find the better model. With an accuracy of 93.74%, the Inception-v3 model performed better than the Xception and ResNet-50 models. The proposed work will be useful for the clinicians to take decisions on Acute lymphoblastic leukemia at the earliest.

VI. REFERENCES

- [1] Pradeep Kumar Das, Diya V A, Sukadev Meher, Rutuparna Panda and Ajith Abraham, “A Systematic Review on Recent Advancements in Deep and Machine Learning Based Detection and Classification of Acute Lymphoblastic Leukemia”, IEEE Access,2022.
- [2] Jan Zidek, Alexandra Bodzas and Pavel Kodytek, “Automated Detection of Acute Lymphoblastic Leukemia from Microscopic Images Based on Human Visual Perception”, Frontiers in Bioengineering and Biotechnology, Vol.8, Aug 2020.
- [3] Shallu Sharma and Sumit Kumar, “The Xception model: A potential feature extractor in breast cancer histology images classification”, ICT Express, Vol.8, March 2022.
- [4] Rohan Khandekar, Prakhyaa Shastry et.al., “Automated blast cell detection for Acute Lymphoblastic Leukemia diagnosis”, Elsevier, 2021.
- [5] Xiaodong Chen, Qiongyu Duan et.al., “Segmentation of lung computed tomography images based on SegNet in the diagnosis of lung cancer”, Elsevier, 2021.
- [6] Foivos I. Diakogiannis, François Waldner, Peter Caccetta and Chen Wu, “ResUNet-a: A deep learning framework for semantic segmentation of remotely sensed data”, ISPRS Journal of Photogrammetry and Remote Sensing, vol.162, April 2020.
- [7] Ketan Joshi, Vikas Tripathi, Chitransh Bose and Chaitanya Bhardwaj, “Robust Sports Image Classification Using InceptionV3 and Neural Networks”, Procedia Computer Science, Vol.167, 2020.
- [8] Chayan Mondala,d, Md. Kamrul Hasana,1,* , Md. Tasnim Jawada , Aishwariya Duttab , Md. Rabiul Islama , Md. Abdul Awalc , Mohiuddin Ahmada, “Ensemble of Convolutional Neural Networks to diagnose Acute Lymphoblastic Leukemia from microscopic images”, Informatics in Medicine Unlocked, Elsevier, Vol.27, 2021.
- [9] ALL IDB2 DATASET – <https://homes.di.unimi.it/scotti/all/>.
- [10] Ahmed T. Sahlol, Fatma Helmy Ismail2, Ahmed Abdeldaim3, and Aboul Ella Hassanien4, “Elephant Herd Optimization with Neural Networks: A \$ase 4study on Acute Lymphoblastic Leukemia Diagnosis”, 12th International Conference on Computer Engineering and Systems (ICCES), 2017.
- [11] Sonali Mishra, Lokesh Sharma, Bansidhar Majhi and Pankaj Kumar Sa, “Microscopic Image Classification Using DCT for the Detection of Acute Lymphoblastic Leukemia (ALL)”, Proceedings of International Conference on Computer Vision and Image Processing, Springer, 2016.
- [12] Ahmed M. Abdeldaim, Ahmed T. Sahlol, Mohamed Elhoseny and Aboul Ella Hassanien, “Computer Aided Diagnostic System for Detection of Leukemia Using Microscopic Images”, Procedia Computer Science, Elsevier, Vol.70, 2015.

Vision Based Real-Time Active Protection System Using Deep Convolutional Neural Network

Rubin Bose S^{*}

Department of Electronics and
Communication Engineering
SRM Institute of Science and
Technology, Ramapuram,
Chennai, India.
rubinbos@srmist.edu.in

Varun Sharma K¹, Karrthik Kishore V²,

Tharunraj S³, Nikhil Srinivas G⁴
^{1,2,3,4}, Department of Computer Science
and Engineering
SRM Institute of Science and
Technology,
Rampuram, Chennai, India.

Regin R

Department of Computer Science
and Engineering
SRM Institute of Science and
Technology, Ramapuram,
Chennai, India.
reginr@srmist.edu.in

Abstract—The health of children and women are highly affected due to conflicts or war. The effects of war create terrible emotional consequences and physical consequences. The well-being and development of nation are also ensured by an intelligent defense system. Threats faced by tanks and other armored vehicles on the battlefield are getting more complicated. The proposed vision based active protection system installed in the tank is capable of recognizing the hostile targets precisely and destroy targets in the air before entering into the territory. This real-time Active Protection System can save the life of the civilians during warfare. The proposed model integrates a vision-based image processing technique with ultrasonic sensor for the real-time active protection system. The model utilizes lightweight deep CNN model (YOLOv5s architecture) on a Raspberry-Pi¹ processor to recognize the hostile targets. Then, the predicted data is transferred from Raspberry-Pi¹ processor to the cloud. Raspberry-Pi² processor receives the information from the cloud and controls the missile operation of the tank in real-time. The Raspberry Pi processor is a low-power computing device, and YOLOv5s is familiar for its light weight and timely recognition. The proposed YOLOv5s model obtained an Average Precision of 93.10%, Average Recall of 89.50%, and F1-score of 91.26%. The Prediction time of the model is 4.1ms on Google Colab and 405ms on Raspberry-Pi processor.

Keywords—Active protection system, YOLOv5, Raspberry-pi, Radar.

I. INTRODUCTION

The health and prosperity of nations are severely damaged by war. According to the world congress on public health, the psychological effects of war are terrible. It includes an increase in the prevalence of somatoform disorders, post-traumatic stress disorder, anxiety, and depression. War also leads to an upsurge in death rates and disabilities, the collapse of social and economic structures, famine, multiple interruptions to healthcare system, the breakdown of medical supply chains, the migration of health practitioners, and fatal disease outbreaks. As per the recent studies it is revealed that, in 2017 at least 16% of children and 10% of women globally were lived in dangerously close proximity to war zones or made homeless as a result of conflicts. Moreover, this exposes them to harassment, sexual assault, isolation, early marriage, and exploitation.

Every nation has a defence team to protect the country from foreign aggression and internal challenges, and to maintain security and peace within its territory. The countries make this possible with the help of modern protecting equipment's. Armoured vehicles and military tanks have become crucial weapons on the battlefield since their invention. With the proliferation of intelligent anti-armour

weapons, tanks and armoured vehicles have been subjected to all-around and 3-dimensional attacks both on the ground and from the air, putting the crew's survivability to the test like never before. To meet this challenge, many nations have built tanks and armoured vehicles with protection as the primary function. However, the protection of tanks and armoured vehicles are strengthened with the introduction of an active protection system.

The purpose of the active protection system is to recognize the approaching hostile targets before they are attacked and mislead, intercept, or eradicate them. Threats to tanks and armoured vehicles are challenging on the battlefield. Tanks are man operated in the battlefield. In the midst of war and battle, all the protective measures are done by the tank operators. The simultaneous task of detecting the hostile targets and actively protecting themselves against the enemy is a time-consuming, and tedious task.

Numerous inventions in the field of artificial intelligence and vision-based technologies made the real-time active protecting system simpler. The tasks like object classification and detection have achieved impressive milestones because of improvements in deep learning and current machine technologies. The recognition of anti-military vehicles like aircrafts, helicopters, tanks, drones, UAVs, and trucks in an unconstrained environment, dynamic backgrounds, different illumination condition, various viewpoints and rapid motion is a challenging and time-consuming task.

In this paper, a deep CNN based model YOLOv5s along with camera and an ultrasonic sensor setup is presented to recognize the targeted vehicle in real-time. The proposed YOLOv5s is a quick and precise object detector algorithm that is perfect for dynamic computer vision applications. Therefore, the hostile target is actively recognized by using the YOLOv5s model. The YOLOv5s model is loaded into the Raspberry-Pi¹ processor. Then the prediction information is transmitted to cloud and from the cloud, the data are transferred to the Raspberry-Pi² to initialize the missile operations.

The vision based active protection system increases the tank protection without compromising manoeuvrability. The proposed mechanism also intelligently recognize the hostile targets and destroy them in the space before entering into our territory. This system can save millions of life from the disaster. The performance of the proposed YOLOv5s model is trained and tested in real-time.

II. LITERATURE REVIEW

Samir et al. [1] suggested a hostile target destroying system using a robotic platform. The combo of stepper motor

with an ultrasonic sensor and a laser missile launcher is utilized to recognize the moving target and attack the target autonomously. The author used Atmel Mega AVR series 8-bit microcontroller. It has a high-performance, low-power ATmega32 processor. The author utilized a cannon laser to fire the closest identified target. The distance between the target and the tank is identified using ultrasonic sensor.

Nagakishore et al. [2] presents an intelligent sonar-based object-tracking system along with DC-gearied motor-powered firing system is utilized for real time protection. An ultrasonic sensor is used to identify the targets at various distances and in any type of lighting conditions (day or night). The author used Atmel 89C52 microcontroller as a control mechanism. When a target is found, the control unit instructs the firing mechanism to destroy it.

Pooja et al [3], suggested an active protection unit by equipping the robot platform with a missile detection system and auto-destroying system. The author utilized an ATmega16, 8-bit, RISC-based microcontroller with 16 MHz is used as a processing and control unit. It has an internal analog to digital converter (ADC) unit and it utilizes low power. The usage of a stepper motor and an ultrasonic sensor enhances to rotate the complete setup to 360 degrees. Therefore, the stepper motor focuses the target and the laser setup fires the recognized hostile target. The author measured the distance between the hostile target and active protection system using an ultrasonic sensor.

Anbalagan et al. [4] suggested an active protection technique that outlines the use of an ultrasonic sensor to protect against a missile that is approaching the target. The controller is interfaced in such a way that the stepper motor used is to turn the ultrasonic sensors 360 degrees. The ultrasonic sensors are utilized to predict the distance between the protection system and the hostile target. The launcher spins towards the target and shoots after receiving the target's information from the control room. The author utilized PIC microcontroller as a processing unit.

Yang et al. [5] suggested an approach to evaluate the efficiency of active Protection System. A complete simulation system is used to determine the effects of added faults. By figuring out how to estimate the impact of errors or, alternatively, by estimating the ideal stand-off distance and counter measure/target distance, the designers may use the model to perform a system-level trade-off with the current stand-off range for intercepting inbound targets.

Likun et al. [6] presents a counterattack component to perform as an active protection system. Its primary duty is to carry out the goal of destroying the incoming targets. The suggested systems will not function independently on the field, rather, they will work in teams or tandem with other weapons, usually with the aid of troops.

Bose et al. suggested various single stage CNN architectures like SSD MobileNet-V2 [7], RetinaNet-DSC [8] and Hybrid SSR models [9], and Sreekar et al. [10] presents YOLOv2 model to perform real-time object detection.

Eslamiat et al [11] compared the performance of the YOLOv3, and YOLOv4, architectures. The author utilized aerial image datasets to train, test and validate the different YOLO algorithms. The author obtained better performance in terms of accuracy and prediction time for the latest version than the other algorithms. The YOLO based architectures are

preferred to be the best choice for the real-time recognition systems.

Rahman et al [12] utilizes the YOLO based CNN model to identify the faulty insulators in the electric pole. The author trained and validated the YOLOv4 with YOLO5 models using faulty insulator datasets. The author obtained an F1-score of 81.00 % to detect the faulty insulators.

Ge et al [13] presents and compare the contrast versions of YOLO 3, 4, and 5. The YOLO architectures are used to perform real-time object detection. The author utilized the MS COCO dataset to test and train the architectures. It is identified that YOLOv5 outperformed YOLOv4 and YOLOv3 in terms of accuracy and prediction time.

Based on the extensive literature survey, every model has its pros and cons. The authors [7], [11], [12] and [13] are suggested the YOLO architecture for the real-time object detection. In this paper, a lighter version of YOLOv5 (YOLOv5s) CNN model with camera and ultrasonic sensor is used for a vision based active protection system. YOLOv5s is utilized to perform the real-time hostile target recognition and the ultrasonic sensor setup is used for distance measurement between the hostile target and tank. The lighter version of YOLOv5 is trained and evaluated using the military datasets. The performance of the YOLOv5s model is also analysed.

III. METHODOLOGY

An integrated deep learning-based hostile target recognition technique with ultrasonic sensor is proposed to perform a real-time active protection in the battlefield tanks. The overview of the Active Protection System (APS) is illustrated in fig1. The hostile target recognition is performed using a lightweight model of YOLOv5 (YOLOv5s). The ultrasonic sensor setup is used to confirm the detection made by the deep learning module and estimates the distance between the hostile target and tank.

The YOLOv5s [14] is an efficient algorithm in recent times used in a wide range of applications. The YOLOv5s is a single-stage object identification model that detects objects using a convolution neural network. YOLO can recognize an object in a neural network in a single stage, unlike the other deep learning models. There are several available versions of YOLO, and this work uses YOLOv5s.

The proposed YOLOv5s model is trained using a custom dataset. This custom-created dataset contains 6 classes of various Military vehicles (Military Aircraft, Military Helicopter, Military Tank, Drones, UAV, Military trucks) and 200 samples per class. The dataset is split as 80:20 for training and testing to avoid overfitting. The training dataset has 960 samples and the testing dataset consists of 240 sample images of 6 different classes.

The custom dataset is used to train a YOLOv5 (You Only Look Once v5) as it is the fastest detecting model with good accuracy. The model is then used as an inference to detect the class of foreign attacks (through military tanks, trucks, aircrafts and drones) in real time with a camera. This model detects susceptible vehicles from the live feed with localization and the category of the vehicles are identified with a high confidence score.

The YOLOv5 architecture consists of three modules (1) CSPDarknet module (2) PANet and (3). The model

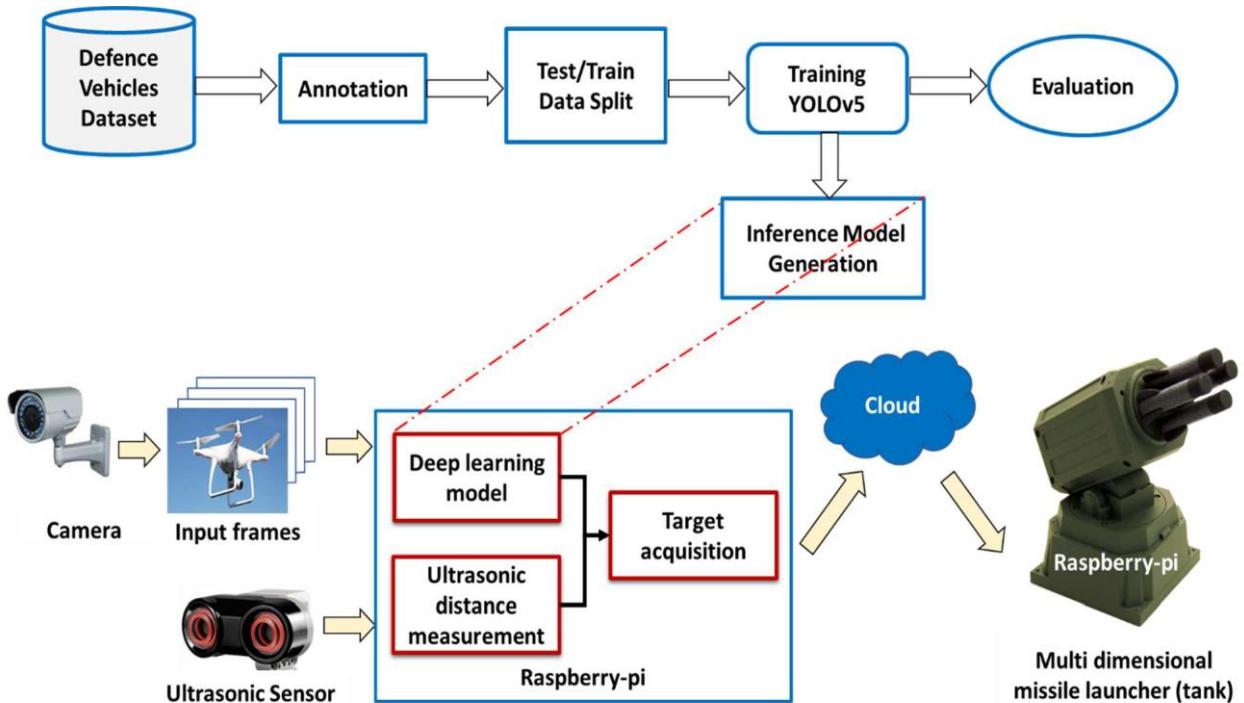


Fig. 1. Functional block diagram of vision based active protection system

parameters are routed into a CSPDarknet to extract features, then into a PANet to combine features, and finally via a YOLO layer to identify output (class, score, location, and size)..

The CSP Darknet53 is utilized as a backbone feature extraction network. It addresses the recurring gradient data in large backbones and combines the gradient change into a feature map. The reduction in the training parameters reduces the inference time, enhance accuracy, and shrink the model size.

A. YOLOv5s Architecture

The YOLOv5s architecture has three main structural blocks i) Backbone, ii) Neck and iii) Head. Fig.2 demonstrates the lightweight YOLOv5s architecture.

- YOLOv5s Backbone-** It uses the cross-stage partial network-based CSPDarknet as the foundation for extracting features from photos.
- YOLOv5s Neck -** It creates a feature pyramid network using PANet to aggregate the features before sending them to the Head for predictions.
- YOLOv5s Head Layers:** Layers that produce predictions using anchor boxes for object recognition.

YOLOv5s returns three outputs: the classes of the detected objects, their bounding boxes and the objectness scores.

Transfer learning, a valuable technique for quickly retraining a model on fresh data without the need to retrain the entire network, was used to train this model. This enhances training times and uses fewer resources than standard training. During the iteration, the ratio between the predicted and true values is calculated using the loss function.

$$L_T = L_{cls} + L_{cnf} + L_{box} \quad \dots\dots (1)$$

where L_T is the total loss, L_{cls} is represented as a classification loss and expressed in equation 2. L_{cnf} is denoted as confidence loss and expressed in equation 3, and L_{box} is the bounding box loss.

$$L_{cls} = \sum_{i=0}^{x^2} l_i^{obj} \sum_{j=0}^R [(P_i(c) - \hat{P}_i(c))^2] \quad \dots\dots (2)$$

$$L_{cnf} = \sum_{i=0}^{x^2} \sum_{j=0}^R l_i^{obj} [(C_i - \hat{C}_i)^2] + \beta_{noobj} \sum_{i=0}^{x^2} \sum_{j=0}^R l_i^{noobj} [(C_i - \hat{C}_i)^2] \quad \dots\dots (3)$$

where, $P_i(c)$ is denoted as a probability to be an object. l_i^{obj} and l_i^{noobj} are represented as the indicator function. C_i is denoted as the objectness.

The performance of the YOLOv5s model is determined using various parameters like Precision, Recall, F1-score and Prediction time. The various parameters are expressed in equation 4, 5, and 6.

$$\text{Precision} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}} \quad \dots\dots (4)$$

$$\text{Recall} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}} \quad \dots\dots (5)$$

$$\text{F1-Score} = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \quad \dots\dots (6)$$

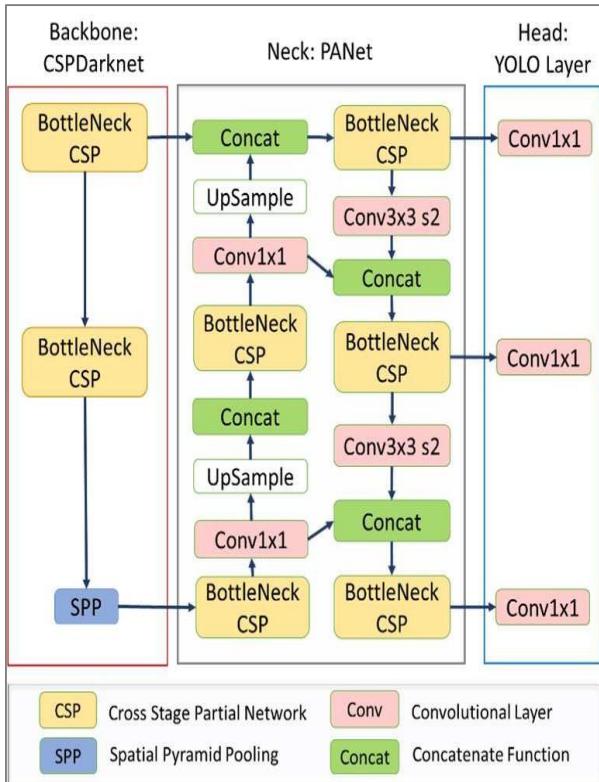


Fig. 2. YOLOv5s architecture

IV. EXPERIMENTAL SETUP

A. Training

The pre-trained weights from the COCO dataset are used to train the YOLOv5s model. The training and testing phases employ the customised dataset. The gradient decent optimization approach is used to improve the training model. With a batch size of 8, the model has been trained for epochs 100, 150, and 200 with a learning rate of 0.0002. Using Python tools from the Deep Learning Toolbox and the Pytorch framework, Google Colab is used to train the model. Python modules like Pandas, Matplotlib, Numpy, and OpenCV packages are used in the experimental analysis. The utilization of Google Colab enhances the training speed as it offers free GPU runtime.

B. Evaluation and Prediction

Test samples are used to put the trained models to the test, and their performance is assessed using measures such as mean average Precision, Precision, Recall, and Prediction Time for an Intersection over Union (IoU) value. The accuracy of hostile attack detection is determined by IoU. When the projected bounding box perfectly overlaps the real-world bounding box, IoU equals one. As long as the IoU is 0.5, the forecast is deemed to be correct, and for higher values of IoUs, the prediction is precise. An IoU value of 0.65 is used in the evaluation of the suggested model.

C. Implementation

Active protective system is implemented using two Raspberry-Pi controllers using the python framework. The Raspberry Pi² takes input from the coordinates given by the Raspberry Pi¹ and controls the servo motors in the multi-

dimensional missile launcher. Pulse width modulation (PWM) signals are used to operate the missile launcher's servo motors. To share the coordinates with the Raspberry Pi² commanding the servo motors, the MQTT broker is hosted on the devices. The targets are first acquired by the image processing and their distance is measured using an ultrasonic sensor. After localization of the target, the Raspberry-Pi¹ send information to the Raspberry-Pi² to fire the target precisely.

V. RESULTS AND DISCUSSIONS

The custom-built dataset is used to train and test the YOLOv5 model. Using a stochastic gradient descent (SGD) optimizer, an eight-batch deep learning-based hostile attack detection model is trained. The SGD has a 0.0002 learning rate, a 0.937 momentum, and a 0.0005 decay rate. The YOLOv5s proposed model is trained using a transfer learning technique with pre-trained weights. The proposed YOLOv5s model is a lightweight and time-efficient detecting model with high precision. The model is then used as an inference to detect the class of hostile attack through military tanks, trucks, aircrafts, UAVs and drones in real-time using a vision-based setup. The proposed model recognizes the suspicious attack from the live feed and identified its location and category with a high confidence score.

Metrics including Average Precision, Average Recall, F1-Score, and Prediction Time are used to evaluate the effectiveness of the YOLOv5s model. With an IoU threshold of 0.65, the performance is assessed for the models at various epochs (100, 150, and 200). The prediction time is calculated during the hostile attack detection process.

TABLE I. PERFORMANCE METRICS OBTAINED USING YOLOv5s.

CNN model	Average Precision (%)	Average Recall (%)	F1-Score (%)
YOLOv5s (100epochs)	91.90	84.20	87.88
YOLOv5s (150 epochs)	93.10	89.50	91.26
YOLOv5s (200 epochs)	93.10	89.50	91.26

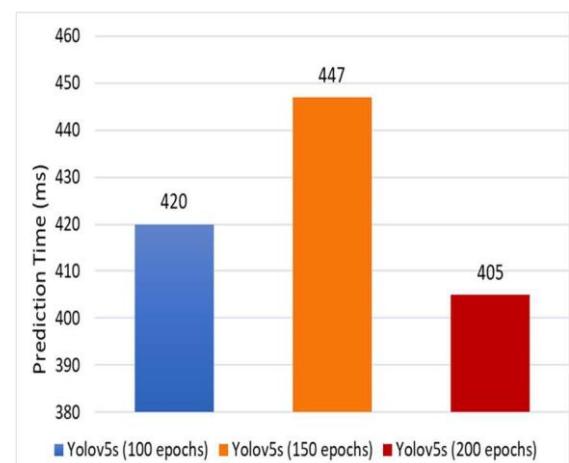


Fig. 3. Prediction time using YOLOv5s model on Raspberry-Pi processor.

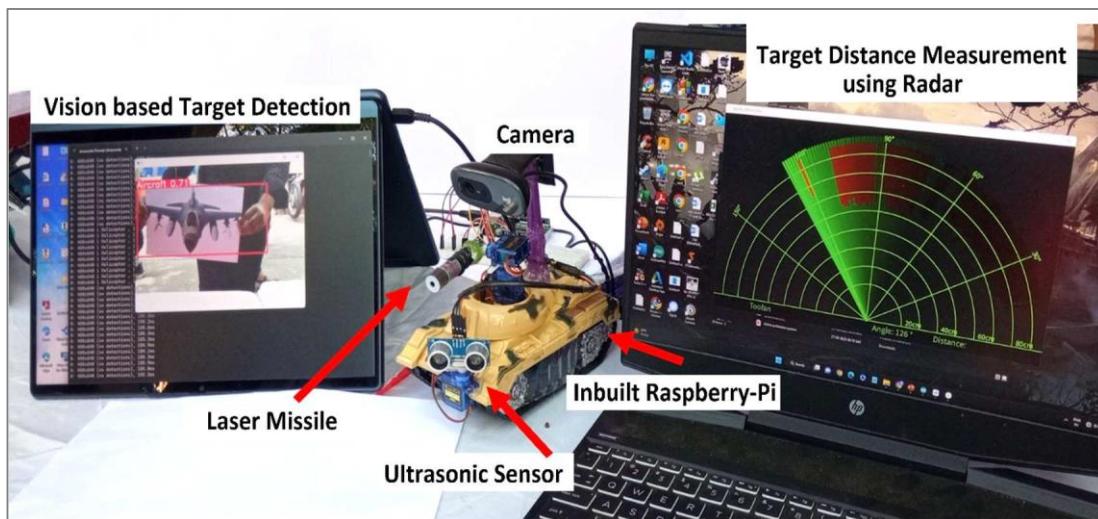


Fig.4. Setup of vision based active protection system



Fig. 5. Sample frames of hostile target recognition

TABLE II. COMPARISON OF THE COMPUTATIONAL TIME ON GOOGLE COLAB

CNN model	Speed (ms)			
	Pre-process	Inference	NMS per image	Overall Prediction time per Image
Yolov5s (100epochs)	0.2	3.8	1.7	5.7
Yolov5s (150 epochs)	0.1	3.7	2.8	6.6
Yolov5s (200 epochs)	0.1	2.0	2.0	4.1

Table I illustrates the performance metrics obtained using YOLOv5s. The suggested YOLOv5s model obtained an average precision of 91.90%, average recall of 84.20% and F1-score of 87.88% for 100 epochs. The performance of the YOLOv5s model is also estimated for epochs 150 and 200.

The F1-score, Average Precision, and Average Recall are 93.10%, 89.50% and 91.26% for the epochs of 150 and 200 respectively. There is no significant variation in the performance metrics for epochs 150 and 200. The model trained for 150, and 200 epochs perform better than the model trained for 100.

The prediction time of the suggested model is analysed on both Google Colab and Raspberry-pi processors. The YOLOv5s model obtained a prediction time of 5.7ms, 6.6ms, and 4.1ms for the epochs 100, 150 and 200 respectively on Google Colab. Table II demonstrates the comparison of the computational time on Google Colab. The model trained for an epoch of 200 predicts with less prediction time (4.1ms) than the other models. The prediction time of 420ms, 447ms, and 405ms are obtained for the epochs 100, 150 and 200 respectively on Raspberry-pi. Figure 3 illustrates the prediction time obtained using the YOLOv5s model on a Raspberry-Pi processor.

The object detection model with the highest average recall and precision is a very effective object detector. High average

accuracy and average recall were attained by the suggested YOLOv5 model. The suggested model is therefore ideal for the real-time practical implementation of an Active Protective System since it has high average accuracy, high average recall, and short prediction times.

TABLE III. COMPARISON WITH EXISTING METHODS

Author	Methodology	
	Ultrasonic Sensor Based	Computer Vision Based
Proposed	✓	✓
Samir et al. [1]	✓	-
Nagakishore et al. [2]	✓	-
Pooja et al. [3]	✓	-
Anbalagan et al. [4]	✓	-
Yang et al. [5]	✓	-

Table III illustrates the comparison of the proposed model with the existing methods. The propose model is unique and it integrates the computer vision based deep learning model with ultrasonic sensor to perform a real-time active protection. Whereas, the existing models Samir et al. [1], Nagakishore et al. [2], Pooja et al.[3], Anbalagan et al. [4], Yang et al. [5] utilize only the sensor based setup for active protection. The proposed model is more precise and intelligent than the existing approaches.

TABLE IV. COMPARISON OF PREDICTION TIME WITH OTHER MODELS

Author	Model	Prediction Time
Proposed	YOLOv5s (200 epochs)	4.1 ms
Sreekar et al. [1]	SSD Lite Mobilenet-V2	10 ms
Bose et al. [7]	YOLOv2	12 ms
Yan et al. [14]	YOLOv5	15 ms

Table IV lists the comparison of prediction time with the other models. The proposed YOLOv5s model trained for 200 epochs obtained a less prediction time of 4.1 ms on google colab. Whereas, Sreekar et al. [1] obtained 10 ms on SSD Lite MobileNet-V2 model, Bose et al. [7] achieved 12 ms on YOLOv2 model and Yan et al. [14] reported 15 ms of prediction time on YOLOv5 model. The proposed model has less prediction time than the other existing models.

Figure 4 demonstrates the complete setup of hostile target recognition using the proposed YOLOv5s model. Fig. 5 illustrates the sample frames of hostile target recognition like military aircraft, military helicopters, military tanks, drones, UAVs, and military trucks. The predicted hostile targets are predicted with a high confidence score. The higher confidence scores enhance the accurate prediction of the hostile target detection.

The ultrasonic transmitter and receiver detect obstacles and display distance between the hostile target and tank on the monitor. The ultrasonic transmitter and receiver are rotated and controlled by a servo motor setup. All the information collected from the sensor is processed by the Raspberry-Pi and displayed.

VI. CONCLUSION

The nation's ability to prosper and maintain its health are severely devastated by war. The vision based Active

Protection System in conflict zones actively recognizes the approaching hostile target and develops a counterattack in the air. Therefore, the Active Protection System protects the human beings and society from a huge disaster. The proposed framework integrates a vision based deep learning model to perform hostile target recognition with an ultrasonic sensor to perform a real-time active protection. The vision-based active protection system utilizes YOLOv5s deep CNN model for the recognition of hostile attacks done by hostile military tanks, trucks, aircrafts, UAVs and drones. The distance between the hostile target and the tank are determined using the ultrasonic sensor. Then the missile is launched to destroy the hostile target from attack. The performance of the model is evaluated using the test samples of the custom dataset. The proposed model obtained an Average Precision of 93.10%, Average Recall of 89.50%, F1-score of 91.26% and a Prediction time of 4.1ms. The proposed technique with a high configuration cameras and ultrasonic sensors can be applied in various real-time applications in the defence field. It can also be implemented in remote monitoring and tracking frameworks.

REFERENCES

- [1] Samir Chopra, Suman Bharti, Tarun Singh Negi, P. D Kulkarni, "Missile Detection by Ultrasonic and Auto Destroy System," (IJESRT) May 2014.
- [2] S Nagakishore Bhavanam, Acharya Nagarjuna, Microcontroller Based Missile Detection and Destroying System. (ICIECE), July 2014.
- [3] Ms. Palwe Pooja Balasaheb, Ms. Shinde Tejasree Anil, Ms. Sonawane Chaitali Shivajirao, and S. M. Bhilegaonkar, "Missile Detection and Auto Destroy System on a Robot Platform. (IJSRD), 2015.
- [4] K. Anbalagan, V. Divakar, Y. Sathik Basha, and M. Senthil Kumar, "Automatic Mystery Detection and Destroy Using Embedded Systems," (IJRE), 2016.
- [5] Yang, Y & Song, H & Wang, P & Zhang, Y. (2020). A simulation method of active protection system defeat probability. Journal of Physics: Conference Series. 1507. 082020. 10.1088/1742-6596/1507/8/082020.
- [6] Likun Yang and Jiagang Xu "Analysis on the Development of Active Protection System for Tanks and Armored Vehicles" J. Phys.: Conf. Ser. 1855 012034, 2021.
- [7] Bose, S. R., Kumar, V. S., "Precise Recognition of Vision-Based Multi-hand Signs Using Deep Single Stage Convolutional Neural Network". In: Singh S.K., Roy P., Raman B., Nagabhusan P. (eds), *Computer Vision, and Image Processing. CVIP 2020. Communications in Computer and Information Science*, vol 1377. Springer, Singapore. https://doi.org/10.1007/978-981-16-1092-9_27.
- [8] Bose, S. R., Kumar, V. S., 'In-situ Identification and Recognition of Multi-hand Gestures Using Optimized Deep Residual Network'. Journal of Intelligent & Fuzzy Systems, vol. 41, no. 6, pp. 6983-6997, 2021.
- [9] Bose, S. R., Kumar, V. S., 'In-situ recognition of hand gesture via Enhanced Xception based single-stage deep convolutional neural network'. *Expert Systems with Applications*, Volume 193, 116427. 2022
- [10] Sreekar, C., Sindhu, V. S., Bhuvaneshwaran, S., Bose, S. R., and Kumar, V.S., "Positioning the 5-DOF Robotic Arm using Single Stage Deep CNN Model," *Seventh International conference on Bio Signals, Images, and Instrumentation (ICBSII)*, pp. 1-6, 2021 doi: 10.1109/ICBSII51839.2021.9445124.
- [11] Nepal, U.; Eslamiat, H. Comparing YOLOv3, YOLOv4 and YOLOv5 for Autonomous Landing Spot Detection in Faulty UAVs. Sensors 2022, 22, 464. <https://doi.org/10.3390/s22020464>
- [12] Rahman, E.U.; Zhang, Y.; Ahmad, S.; Ahmad, H.I.; Jobaer, S. Autonomous vision-based primary distribution systems porcelain insulators inspection using UAVs. Sensors 2021, 21, 974
- [13] Ge, Z.; Liu, S.; Wang, F.; Li, Z.; Sun, J. Yolox: Exceeding yolo series in 2021. arXiv 2021, arXiv:2107.08430
- [14] Yan B, Fan P, Lei X, Liu Z, Yang F. A Real-Time Apple Targets Detection Method for Picking Robot Based on Improved YOLOv5. Remote Sensing. 2021; 13(9):1619. <https://doi.org/10.3390/rs13091619>.

Transfer Learning Approach for Epizootic Ulcerative Syndrome and Ichthyophthirius Disease Classification in Fish Species

Vijayalakshmi M

School of Electronics Engineering
Vellore Institute of Technology
Chennai, India
vijayalakshmi.m2021@vitstudent.ac.in

Sasithradevi A

Centre for Advanced Data Science
Vellore Institute of Technology
Chennai, India
sasithradevi.a@vit.ac.in

P Prakash

Department of Electronics Engineering
Madras Institute of Technology
Chennai, India
prakashp_mit@annauniv.edu

Abstract— A substantial contributor to the national economy in terms of food, nutrition and economic progress is the Indian Fish industry. Along with satisfying domestic demand, this sector has significantly increased its contribution to foreign exchange revenue by export. The fish industry is currently developing in a direction that will increase the commercialization and specialization of aquatic species. The major bottleneck in the development of fish industry is fish diseases as it affects the productivity rate heavily. Thus, it is essential to predict fish disease at early stage. In this paper the computer vision technique is implemented for automatic detection and classification of fish disease in fish industry. A transfer learning strategy is proposed for automatic fish disease detection. Healthy fish, fish with EUS disease, and fish with Ichthyophthirius disease are the three classes considered for the analysis. The dataset is trained by proposed VGG16 model and is compared with AlexNet. Low level features are better extracted using VGG16 which plays essential role in classification. AlexNet provides accuracy of 64%, while VGG16 produces accuracy of 91%.

Keywords— *Fish disease, Computer Vision, Classification, Transfer Learning*

I. INTRODUCTION

The penultimate idea behind Blue Revolution [1] is to use the water resources to the fullest extent, possible for the development of fisheries in a sustainable manner while taking environmental concerns into consideration. This will benefit the nation's economy, fisher men and also contribute to food and nutritional security. The fish industry is a major contributor to global economy, food production and also major supply of protein rich food for the world population. It is stated that there are approximately 34,000 fish species living in the salt and fresh water. The demand for food production is continuously increasing in order to meet the growing world's population [8]. From last decades the demands the consumption of fish food consumption is considerably raising worldwide. Hence fish industry is becoming more and more prominent in regular supply of fish to match the world-wide food demand. This clearly illustrates the necessity of fish production and the vital role of the fishes in fish farming. Fish farming is influenced by a variety of variables, including climatic change, the underwater environment, biological factors, etc. The main factors affecting the management of fisheries [9] are water body pollution, change in environmental conditions and overfishing. The biggest challenge for the fish farming is the prevalence of various diseases in these fishes, which immediately reduces the pace of production. Thus, the research on fish species is paying more attention not only in feeding system [11], fish production [12], fish classification [13], behavioral analysis [14] etc., but also for disease control [15] in fish species.

The major factor that effects in the spreading of fish diseases are water turbidity, temperature, pH concentration, climatic change, oxygen level and other microorganism present in aquatic environment [16]. A method of evaluating water quality is necessary because it is the prominent cause of diseases brought on by the environment. Due to sudden change in temperature, the nitrogen level in water changes which leads to poor environment condition for aquatic species resulting in various disease. Since fish diseases spread quickly in an aquatic environment, it is crucial to predict them in well advance. Live detection and monitoring diseased fish are embedded with numerous complexities and that it requires manual inspection with direct human visualization. The accuracy level also depends on the skilled level of the individual person and is quite unpredictable. Therefore, automatic detection of fish diseases is proven to be an essential episteme in smart fish industry and is noteworthy to identifying environmental liabilities, comprehending disease symptoms, and enhancing the welfare and health of fish. With the current technological revolution, computer vision [17] has been widely adopted in disease analysis, particularly in humans, aquatic species and floras, and also in assisting human experts to render the appropriate therapy.

Many studies are conducted in fisheries sectors for disease control and prediction in fish species. Wide reviews are conducted for developing smart fisheries sector based on various technologies. Several classification techniques, including statistical methods, rule-based expert systems, machine learning, deep learning, and hybrid methods [18] were published. These cutting-edge methods for fish industry were subsequently proposed with a prime focus to conduct early and precise disease identification.

Waleed et al. [2] suggested that fish farms can use an automatic method to identify the diseases like Ichthyophthirius (Ich), Columnaris, and Epizootic Ulcerative Syndrome (EUS) [19]. Yang et al. [23] compared the effectiveness of applying various Convolutional Neural Network (CNN) architectures to the underwater image dataset in various color spaces. From the comparative study on deep net architecture, it was inferred that Alex Net acquired good accuracy in classifying the fish disease. Juel Sikder et al., [3] developed a smart model to automatically identify and categorise fish infections in freshwater, particularly in Bangladesh's Rangamati Kaptai Lake and Sunamganj Hoar region. The Fuzzy based K-means clustering and Fuzzy C-means [3] fuzzy clustering algorithms are used for fish disease detection to segment the filtered image. The features are extracted from the segmented regions using Gabor's Filters with combination of the Gray Level Co-occurrence Matrix [4]. Finally, the test image is classified using Multi-class Support Vector Machines

(MSVM). The K-means and C-means fuzzy logic are implemented for the classification. Divinely et al., (2022) [4] implemented a Probabilistic Neural Network (PNN) to detect fish infections (EUS ulcers, a fungus disease). Images for the database and input are gathered from a variety of sources and online sites, respectively. After converting RGB to grey, images are preprocessed to remove undesirable distortions and enhance some image attributes that are important to the entire processing pipeline. For effective detection, a variety of extraction techniques have been used, the fish diseases taken for consideration are ammonia toxicity, worm on fins like camallanus and dropsy. Perumal N et al.,[5] surveyed the aquaculture fish cultivation in Ramanathapuram district Tamil Nadu and proposed a method to detect the EUS disease in various fish species. For the purpose of identifying and categorizing EUS fish disease, the author employs the K Star algorithm. K Star uses an entropy-based distance function and is an instance-based classifier. Distance is measured using the K* technique, which is also used to establish constant characteristics and missing values. This method focuses on instance-based learning approach for classification and evaluation. The highest level of accuracy is achieved using K* algorithm when compared to other algorithms.

Transfer learning is one of the most popular approaches implemented in deep network models. Transfer learning approach is critically important due to its remarkable results in a variety of computer vision applications [6], including segmentation, target recognition, classification and many more. As this requires less pre-processing and results in high accuracy it has become very popular among researchers. Depending on the various deep models [10] the number of layers and the parameters vary accordingly.

Therefore, in this research, transfer learning approach is implemented for fish diseases detection. Particularly in this paper the Deep models implemented are VGG 16 and Alex Net. This paper follows a methodical format as Section 2 focuses on Methodology followed by description of collected data and Network architecture in Section 3. Section 4 discusses in detailed analysis and results and conclusion is stated in Section 5.

II. METHODOLOGY

The overall methodology for disease detection in fish species is illustrated in Fig. 1. The proposed methodology implements the transfer learning approach and suitable optimization operations. The foremost step is in digital image processing is the image acquisition followed by the preprocessing techniques. The images are acquired and in order to increase the datasets the augmentation algorithm is implemented. A necessity of applying augmentation algorithm is to increase the diseased fish images with the help of following operations like flipping, scaling, zooming, right shift and left shift. The collected images are labelled as good, red_spot and white_spot, thus subjected to three class classification. These images are then resized to 224*224 dimensions. The collected images are split into train and test then examined in the predefined deep models using transfer-based technique. The deep models taken for consideration are AlexNet and VGG-16. The model is trained and tested for our own datasets of fish species of three classes. The performance of the deep models are evaluated using the performance metrics.

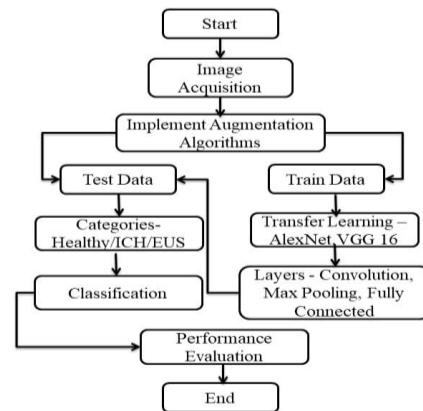


Fig. 1. Proposed workflow for disease detection in fish species

III. MATERIALS AND METHODS

A. Dataset - Fish Disease

The categories taken for consideration are healthy fish, EUS disease and Ich disease in fish. The parasite infestation ICH, commonly known as white spot syndrome, is produced on by the ciliated protozoan Ichthyophthirius multifiliis and affects mostly the freshwater cultivated fish [7]. One of the most typical infections found fish in tropical region is Ich [19]. Its symptoms include microscopic white spots that resemble salt granules on the body section and also gill of fish. Some common symptoms include frequent body rubbing against environmental items, appetite loss and unusual hiding activity. The oomycete infection is also termed as Epizootic ulcerative syndrome (EUS) otherwise called as red spot disease. This type of disease is mostly present in fish cultivated in tropical water environment and in very low temperature regions. The brackish water and freshwater cultivated fishes [8] in continents like Asia and Australia are highly affected by this infection. Fish initially get red blotches on their skin. These lesions grow to become ulcers and large erosions packed with mycelium and necrotic tissue. Then slowly Granulomas begin to develop in the internal organs. The images of three categories of fish are shown in Fig. 2.



Fig. 2. Sample images of three classes of fishes i) Healthy fish ii) EUS
iii) ICH

The fish disease images are acquired from fish farms and from internet sources. The data augmentation is applied to increase the number of images due to limited resources of fish diseased images [20]. The various augmentation techniques employed are rotation, zooming, and cropping, noise addition, vertical and horizontal flip, left and right shift. After performing the augmentation 1000 images per categories are taken into consideration. The train test ratio implemented is 80:20.

B. Network Architecture-VGG 16

VGG-16 model was introduced by Karen Simonyan and Andrew Zisserman [21] ILSVRC challenge and it is demonstrated on ImageNet dataset and won the first place in

object detection and second place in object classification. The input sizes of the image are fixed and have RGB Channels, the image dimensions (224, 224, 3). This model consists of 16 layers with total number of parameters 138.4 million approximately which consists of 13 convolution layers stacked with 5 max pooling layers combined with 3 fully connected layers. This model classifies the 1000 classes of objects that are from ImageNet datasets. In AlexNet the dimension of the kernels is different for each convolution layer. But in case of VGG-16 all convolution layer has a uniform kernel size of 3×3 and all max pooling layer have kernel size of 2×2 throughout the layer. The convolution layer is activated using activation function called Rectified Linear Unit (ReLU) with non-linear property. The 13 convolution layer and the 3 fully connected layers uses tuneable parameters so totally the 16 layers hence this model is named as VGG-16 model and shown in Fig. 3. The convolution layer uses same padding which indicates that the resolution of the image before performing the convolution operation is same after performing the convolution operation.

The input image dimension 224x 224 including RGB Channels are passed through the stacks of convolution layer. The first two convolution layer consists of 64 channels of kernel size 3×3 . The formula for calculating the output layer is shown in the below equation.

$$O = \left[\frac{(I - K + 2P)}{S} \right] + D \quad (1)$$

Where, O is the output dimension of convolution

I is the Input image dimensions

K is the Kernel size

P is the Padding

S is the stride

D is the depth of the feature map

The output is passed to the max pooling layer of the stride size 2×2 and the image dimension becomes $112 \times 112 \times 64$ after performing the max pooling operation. The combination of two convolution layer along with max pooling layer forms the first stack. Then the output was passed through the next stack consisting of two convolution layers along with max pooling layer with 128 filters instead of 64 filters the output dimension becomes $56 \times 56 \times 128$. This is followed by third stack consisting of three convolutional layers and max pooling layer with 256 filters and this gives the output image dimension $28 \times 28 \times 256$. Same process was performed for next two stacks and then connected to three fully connected layers.

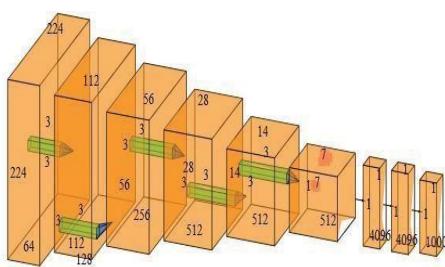


Fig. 3. Architecture of VGG16. The outer box represents each layer and the inner box represents the kernel size

C. Network Architecture-AlexNet

AlexNet [22] model was introduced by Alex Krizhevsky in an object recognition challenge (ILSVRC) and this model is implemented in ImageNet dataset and won the challenge in 2012. This architecture consists of five convolution layers along with the max pooling layer and the 3 fully connected layers with tuneable parameters. The activation function used is ‘ReLU’ activation function as the training process speed is much better than other activation function. The dropout layer is also implemented to avoid the overfitting issue. The entire architecture is clearly shown in Fig. 4.

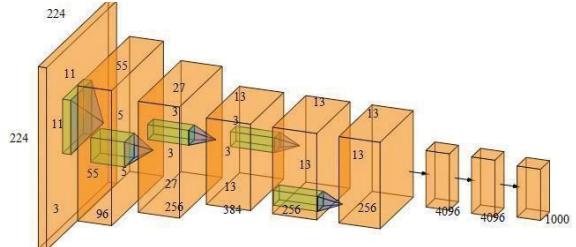


Fig. 4. Architecture of AlexNet. The dimensions details are mentioned for each layer and inner box represents the kernel size

IV. RESULT AND DISCUSSION

The three categories of fish are taken for disease classification and implemented the transfer learning approach using VGG-16 model and AlexNet [23] model. The collected images are augmented and the number of sample fish is expanded to 1000 images for respective categories. Then as a preliminary process of pre-processing all the images are resized to respective image dimensions depending on deep models with implementing one hot encoding. Then the datasets are split into train and test data with 800 images and 200 images of each category respectively. The kernel size, strides and the activation functions [25] for deep models are assigned to respective deep models. Deep network models are used to train and test the datasets. The performance parameters Precision, Recall, and F1-score [24] for EUS category are inferred to be 0.69, 0.66, and 0.67 for Alex-Net, respectively. VGG-16 achieved 0.92, 0.94, 0.93 in terms of Precision, Recall and F1 score respectively for same category. The results obtained are tabulated in Table 1 and Table 2. The overall accuracy of AlexNet is 64% and that of VGG-16 is 91%. The sample output images of predicted category and original category for VGG 16 and AlexNet are shown in Fig. 5 and Fig. 7 respectively. The confusion matrix for VGG 16 is shown in Fig. 6. The drawback of this transfer learning strategy is that the layers cannot be modified, which makes it difficult to denoise because few low-level characteristics are extracted. So, a novel model to be developed in the future while taking the denoising effect into account.

TABLE I. TABULATION FOR ALEXNET – RESULT SUMMARY

Fish Class	AlexNet – 62.3 million Parameters – 240 MB		
	Precision (%)	Recall (%)	F1-Score (%)
Healthy	0.58	0.73	0.65
EUS	0.69	0.66	0.67
ICH	0.69	0.52	0.60
Accuracy – 64 %			

TABLE II. TABULATION FOR VGG 16 – RESULT SUMMARY

Fish Class	VGG-16 – 138.4 million Parameters – 528 MB		
	Precision (%)	Recall (%)	F1-Score (%)
Healthy	0.85	0.91	0.88
EUS	0.92	0.94	0.93
ICH	0.96	0.86	0.91
Accuracy – 91 %			

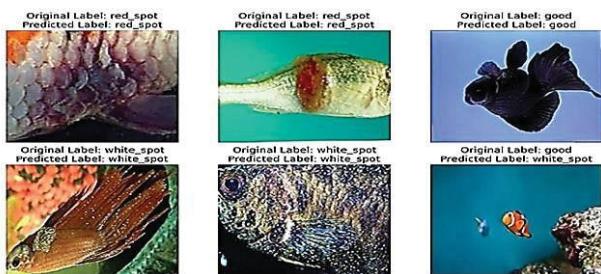


Fig. 5. Result Obtained using VGG 16

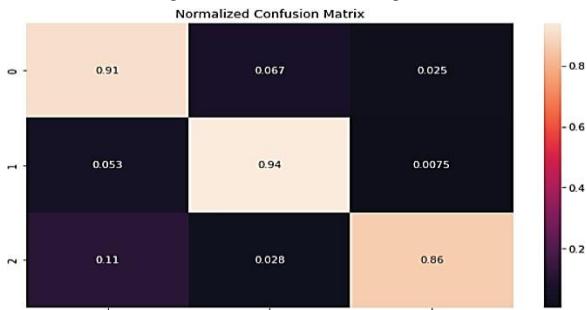


Fig. 6. Confusion Matrix for VGG16

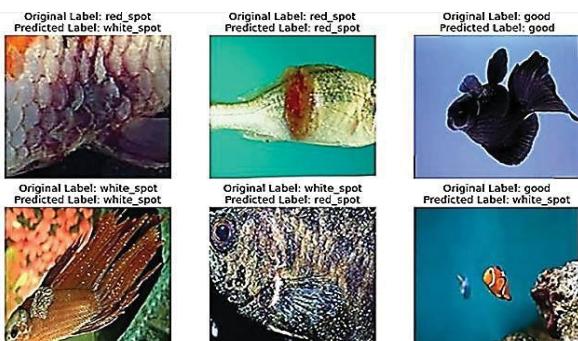


Fig. 7. Result Obtained using AlexNet

V. CONCLUSION

Fish disease prevention has traditionally been considered a difficult endeavor that necessitates the implementation of early disease prediction and detection. The three classes of fishes for disease detection are detected automatically through the lens of computer vision technology to trace its corresponding behavior. Consequently, in this domain, transfer learning has been adapted to study the pathogenic variables of Ichthyophthirius (Ich) and Epizootic ulcerative syndrome (EUS). Furthermore, the Automatic diagnosis of diseased fish have scrutinized through the deep models of AlexNet and VGG-16. Hence, on a comparative basis, the VGG 16 model provides promising results and in future the novel deep model is to be developed for various fish diseases.

REFERENCES

- [1] Garlock, T., Asche, F., Anderson, J., Bjørndal, T., Kumar, G., Lorenzen, K., ... & Tveterås, R. (2020). A global blue revolution: Aquaculture growth across regions, species, and countries. *Reviews in Fisheries Science & Aquaculture*, 28(1), 107-116.
- [2] Waleed A Medhat H Esmail M Osama K Samy R and Ghanim T M 2019 Automatic recognition of fish diseases in fish farms Proc. Int. Conf. Comput. Eng. Syst. pp. 201–06 doi: 10.1109/ICCES48960.2019.9068141
- [3] Sikder, Juel & Sarek, Kamrul & Das, Utpol. (2021). Fish Disease Detection System: A Case Study of Freshwater Fishes of Bangladesh. *International Journal of Advanced Computer Science and Applications*. 12. 10.14569/IJACSA.2021.01206100.
- [4] Divinely, S. J., Sivakami, K., & Jayaraj, V. (2019). Fish diseases identification and classification using Machine Learning. *Int. J. Adv. Res. Basic Eng. Sci. Technol.(IJARBEST)*, 5, 46-51
- [5] Perumal, N. V., Rajkumar, M., Perumal, P., & Rajasekar, K. T. (2009). Seasonal variations of plankton diversity in the Kaduviyar estuary, Nagapattinam, southeast coast of India. *J. Environ. Biol.*, 30(6), 1035-1046.
- [6] Crowl, S., Jordan, B. T., Ahmed, H., Ma, C. X., & Naegle, K. M. (2022). KSTAR: An algorithm to predict patient-specific kinase activities from phosphoproteomic data. *Nature communications*, 13(1), 4283.
- [7] Ahmed, M. S., Aurpa, T. T., & Azad, M. A. K. (2022). Fish disease detection using image based machine learning technique in aquaculture. *Journal of King Saud University-Computer and Information Sciences*, 34(8), 5170-5182.
- [8] Fabic, J. N., Turla, I. E., Capacillo, J. A., David, L. T., & Naval, P. C. (2013, March). Fish population estimation and species classification from underwater video sequences using blob counting and shape analysis. In 2013 IEEE international underwater technology symposium (UT) (pp. 1-6). IEEE.
- [9] Pearlman, J., Garello, R., Delory, E., Castro, A., del Río, J., Toma, D. M., ... & Zielinski, O. (2014, September). Requirements and approaches for a more cost-efficient assessment of ocean waters and ecosystems, and fisheries management. In 2014 Oceans-St. John's (pp. 1-9). IEEE.
- [10] Chen, T., Zhong, L., Zhou, N., & Hoppe, D. (2021, December). Catch weight prediction for multi-species fishing using artificial neural networks. In 2021 20th IEEE International Conference on Machine Learning and Applications (ICMLA) (pp. 1545-1552). IEEE.
- [11] Hu, W. C., Chen, L. B., Huang, B. K., & Lin, H. M. (2022). A computer vision-based intelligent fish feeding system using deep learning techniques for aquaculture. *IEEE Sensors Journal*, 22(7), 7185-7194.
- [12] Thai, M. T., Chu-Ha, T. N., Vo, T. A., & Do, T. H. (2022, December). An Approach to Recommend Fishing Location and Forecast Fish Production by Using Big Data Analysis and Distributed Deep Learning. In 2022 RIVF International Conference on Computing and Communication Technologies (RIVF) (pp. 322-327). IEEE.
- [13] Dey, K., Hassan, M. M., Rana, M. M., & Hena, M. H. (2021, July). Bangladeshi indigenous fish classification using convolutional neural networks. In 2021 International Conference on Information Technology (ICIT) (pp. 899-904). IEEE.
- [14] Kalaiselvi, V. K. G., Ranjani, J., & SM, V. K. (2022, March). Illegal Fishing Detection using Neural Network. In 2022 International Conference on Communication, Computing and Internet of Things (IC3IoT) (pp. 1-4). IEEE.
- [15] Ahmad, A., Saraswat, D., & El Gamal, A. (2022). A survey on using deep learning techniques for plant disease diagnosis and recommendations for development of appropriate tools. *Smart Agricultural Technology*, 100083.
- [16] Hitesh C Rituraj P and Prodip D 2015 Image processing technique to detect fish disease *Int. Jour. of Com. Sci. & Secur.* 9 (2) p. 121-31.
- [17] Yasruddin, M. L., Ismail, M. A. H., Husin, Z., & Tan, W. K. (2022, May). Feasibility study of fish disease detection using computer vision and deep convolutional neural network (dcnn) algorithm. In 2022 IEEE 18th International Colloquium on Signal Processing & Applications (CSPA) (pp. 272-276). IEEE.

- [18] Vo TTE, Ko H, Huh J-H, Kim Y. Overview of Smart Aquaculture System: Focusing on Applications of Machine Learning and Computer Vision. *Electronics*. 2021; 10(22):2882. <https://doi.org/10.3390/electronics10222882>
- [19] Sankar, M. M., Rao, C. N., Sailaja, G., Bhuvaneswary, N., & Gunasekhar, P. (2013). White spot syndrome virus detection in shrimp images using image segmentation techniques. *International Journal of Advanced Research in Computer Science and Software Engineering*, 3(9).
- [20] Aquino, H. L., Concepcion, R. S., Dadios, E. P., Sybingco, E., & Bandala, A. A. (2020, December). Trend forecasting of computer vision application in aquaponic cropping systems industry. In 2020 IEEE 12th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management (HNICEM) (pp. 1-6). IEEE.
- [21] Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.
- [22] Yuan, Z. W., & Zhang, J. (2016, August). Feature extraction and image retrieval based on AlexNet. In Eighth International Conference on Digital Image Processing (ICDIP 2016) (Vol. 10033, pp. 65-69). SPIE.
- [23] Yu, W., Yang, K., Bai, Y., Xiao, T., Yao, H., & Rui, Y. (2016, June). Visualizing and comparing AlexNet and VGG using deconvolutional layers. In Proceedings of the 33 rd International Conference on Machine Learning.
- [24] Canbek, G., Taskaya Temizel, T., & Sagiroglu, S. (2021). BenchMetrics: A systematic benchmarking method for binary classification performance metrics. *Neural Computing and Applications*, 33(21), 14623-14650.I.
- [25] Rasamoelina, A. D., Adjailia, F., & Sinčák, P. (2020, January). A review of activation function for artificial neural network. In 2020 IEEE 18th World Symposium on Applied Machine Intelligence and Informatics (SAMI) (pp. 281-286). IEEE.

Prediction of new causing proteins for Alzheimer's, Parkinson's, and Huntington's diseases: Protein-Protein Interaction Analysis

Sai Gopala Swamy Gadde
Dept. of CSE

V R Siddhartha Engineering College
Vijayawada, India
gadde.saigopal@gmail.com

Dr. K. Srinivas
Dept. of CSE

V R Siddhartha Engineering College
Vijayawada, India
vrdrks@gmail.com

Bhargavi Vaka
Dept. of CSE

V R Siddhartha Engineering College
Vijayawada, India
vakabhargavi15@gmail.com

Sai Kumar Vadapalli
Dept. of CSE

V R Siddhartha Engineering College
Vijayawada, India
vadapallisaikumar33@gmail.com

Abstract—neurodegenerative diseases, including Alzheimer's disease, Parkinson's disease, and Huntington's disease, result in the dysfunction or death of cells within the central nervous system. These Neurodegenerative diseases can be diagnosed but not predicted at their early stages, as prediction is only applicable before the diseases manifest themselves. Evolutionary and Protein-Protein interaction analysis plays a vital role in identifying new-causing proteins and their functionalities in these chronic diseases directly or indirectly. The primary objective of this paper is to discover newly associated proteins that contribute to the development of these neurodegenerative diseases collectively. This framework consists of two modules. The primary module includes applying evolutionary analysis to identify the top proteins of neurodegenerative diseases. The secondary module includes the implementation of the Protein-Protein interaction (PPI) analysis to find a relation among the identified proteins by building a PPI network. This study reveals that APP, PSEN1, MAPT, GIG25, PINK1, PARK7, SNCA, HTT, HIP1, and HAP1 are the proteins having a relation directly and indirectly among Alzheimer's, Parkinson's, and Huntington's diseases.

Keywords—Neurodegenerative diseases, Alzheimer's, Parkinson's, and Huntington's, Evolutionary analysis, Protein-Protein interactions (PPIs).

I. INTRODUCTION

The central nervous system cells stop functioning or die as a result of neurodegenerative disorders like Alzheimer's disease, Huntington's disease, and Parkinson's disease. Degenerative nerve diseases have an impact on numerous bodily processes such as balance, movement, speech, breathing and heart function. Alcoholism, a tumour, a stroke, or some other conditions like Toxins, chemicals, and viruses are the causes of neurodegenerative diseases. Among all these diseases the top diseases are Alzheimer's, Parkinson's, and Huntington's disease. The degeneration of neurons in the brain that occurs in neurodegenerative diseases can be observed in various levels of neuronal circuitry, ranging from the molecular level to the systemic level. These diseases are

regarded as incurable because there is no known mechanism to stop the ongoing destruction of neurons [1]. The Novelty of this study includes the combination of all three diseases together and the findings stating the relationship between the proteins causing these three diseases together.

A. Alzheimer's Disease

In Alzheimer's disease, the brain shrinks and brain cells start to die as a result of the neurologic degenerative disorder. According to current theories, the abnormal protein buildup in and around brain cells is the main cause of Alzheimer's disease. The disease causes dementia means losing thinking ability. Memory loss, forgetting things and losing the ability to concentrate are symptoms of these diseases [2].

B. Parkinson's Disease

Parkinson's disease affects both the nervous system and the body parts that are controlled by the nerves. The problem worsens with time. The loss of nerve cells in the substantia nigra, a region of the brain, causes Parkinson's disease. At first, the symptoms start as a little tremor in one hand. Although tremors are one of the illness's typical symptoms, you might also stiffen up or move more slowly. The slowness of movement, poor balance and coordination, as well as shaking in the hands, arms, legs, jaw, or head, are among the symptoms [3].

C. Huntington's Disease

Brain nerve cells gradually deteriorate due to Huntington's disease, an extremely rare inherited condition. The physical, cognitive, and psychological problems that Huntington's disease frequently brings on have a significant impact on a person's functioning abilities. Concentration problems, depression, and mood swings are symptoms caused by Huntington's disease [4].

D. Phylogenetic tree

A species' connections, including where it is supposed to have originated from, which includes the closest related species and other details, are represented according to phylogeny. The tree is built using the neighbour-joining algorithm. Edge lengths may be used to calculate time, and each node in the phylogenetic tree with descendants denotes the predicted most recent common ancestor of those descendants [19].

E. Protein-Protein Interaction (PPI)

PPIs are the actual physical connections between two or more proteins and the representation of intricate biological processes. Currently, PPIs are utilized to build PPI networks to analyze complex pathways and uncover the activities of unidentified proteins. With the help of PPI, the relation and the influences of other proteins are easily identified on the considered proteins. Nowadays PPI interactions are made easy by using online tools, which helps in saving time.

II. LITERATURE SURVEYS

Srinivas et al [5] discussed about the Protein-Protein Interaction (PPI) analysis Approaches and methods. The PPI detection methods which they applied and used include In vitro, In silico, and In vivo among these methods, in vitro has limitations like cost, and time, and using In silico for Protein-Protein interaction can make fast predictions for a large set of compounds.

Hussain et al [6] discussed about the identification of proteins involved in neurological disorders like Parkinson's, Huntington's, and Alzheimer's disease development are closely related to each other. The similarities are identified by constructing a Phylogenetic analysis among the proteins.

Karbalaei et al [7] discussed about the study of Protein-Protein interactions of Non-alcoholic fatty liver disease (NAFLD) and Alzheimer's disease based on systems biology methods. The genes that are common in both diseases are identified and constructed in a network using a String database. finding the hub and bottleneck proteins are used for drug targeting.

Srinivasa Rao et al [8] discussed about the improvement of pharmacological aspects requires a better knowledge of the molecular pathways underlying the disease. For this construction, a PPI network from the dataset includes the interactions taken from different databases that include Intact, HPRD, Bio Grid, MINT, DIP, INNATEDB, and BIND. After the PPI network, to rank and analyze the disease proteins, support ratings were assigned.

Hernandez et al [9] discussed about the molecular mechanism behind aSyn aggregation and its associated pathology in Parkinson's disease (PD), they hypothesized modifications to these interactions and essential components. The modifications in the associates may stop aSyn accumulating, and counteract synucleinopathies harmful consequences, therefore providing fresh opportunities for pharmaceutical targets in the management of the disease.

Botelho et al [10] discussed about the identification of genes encoding proteins with variations highly related to PD and periodontitis and proposed Using these genes, a study of the Protein-Protein interactions (PPI) network, Genome-Wide Association Studies (GWAS) database was used to collect relevant genes. Then, using the Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) database, they carried out a sensitivity analysis with a confidence cutoff of 0.7 and a protein interaction analysis with a confidence cutoff of 0.9.

Goñi et al [11] discussed about the determination of the degree and betweenness, two crucial network theory measures, which are traits that set nodes that are implicated in MS and AD (seed-proteins) and nodes that are not (neighbours). The researchers conducted an evaluation of four networks, which included the MS-blood network, the MS-brain network, the AD-blood network, and the AD-brain network, using empirically-confirmed Protein-Protein Interaction (PPI) data.

Rakshit et al [12] discussed about Protein-Protein Interaction (PPI) databases, microarray-based gene expression data, post-mortem brain tissue samples from Parkinson's disease patients, and PPI networks of differentially expressed (DE) genes were constructed. Both the frontal cerebral cortex and the substantia nigra yielded samples. To create two Query-Query PPI (QQPPI) networks, two sets of DE genes were chosen from the microarray data using 2-tailed t-tests and Significance Analysis of Microarrays (SAM).

Chakraborty et al [13] discussed about PPI networks by creating a network that displays proteins that were associated with the network that was found in both the MS and HD datasets. Significant relationships between the DEGs and the microRNAs, particularly the highest degree microRNA, USA-mir-155-5p, were discovered. The Nod-like receptor signalling pathway was revealed to be the highest-enriched KEGG pathway overall. Therefore, this study may serve as a springboard for future academic investigations that would elucidate more insightful data on both NDs.

Heinz et al [14] discussed about MID1 protein. Explained that mutant RNA transcript is also involved in cellular dysfunction and described how the mutant HTT transcript interacts abnormally with a protein complex that contains the MID1 protein. They demonstrated that MID1 expression is abnormally elevated in the brains of HD patients and HD mice. The idea that inhibiting the interaction between MID1 and mutant HTT, mRNA to prevent mutant HTT translation would be an effective therapeutic approach is further supported by our discovery.

Xie et al [15] discussed about to suggest that treating HD may be possible by using overexpression to increase striatal BDNF supply. They discovered that some BTg and YAC. BTg mice experienced seizures by the age of 16 months. They were expressed in epileptogenic areas such as the entorhinal cortex and hippocampus, which may be the

reason why BTg and YAC; BTg animals experienced more seizures.

Nopoulos et al [16] discussed about the path etiology of Huntington's disease. They explained the DNA changes that cause Huntington's disease which include the unique feature of a DNA trinucleotide (triplet) repeat where Huntington is caused by the 36-39 CAG repetitions. The CAG repetition is said to be common if the CAG repetition is from 10-35. The causes and symptoms like depression, and mood swings are explained in this study.

Proskura et al [17] discussed about HTT Protein-Protein interaction (PPI). The author considered a set of proteins that cause Huntington's disease and constructed a PPI network to identify the relationships among the considered proteins which share the most common relationship with the proteins. The PPI network is examined to identify the relation between the proteins among themselves. The complexes responsible for moving freshly produced GluR2-AMPARbu were described, and it was determined whether the vesicular movement was in the direction of the positive or negative end of the microtubule. From the study, the authors brought up new innovative techniques for the directed verification of diverse data and the diagnosis, avoidance, and treatment of the associated Huntington's disease.

III. METHODOLOGY

A. Data Collection:

Compile a list of proteins associated with neurodegenerative diseases such as Alzheimer's, Parkinson's, and Huntington's from Uniport (<https://www.uniprot.org/>), Gene cards (https://www.gene_cards.org/), NCBI (<https://www.ncbi.nlm.nih.gov/>), and WWW (World wide web). From that collection, the protein sequence of the related top 10 proteins is in the form of FASTA (fast alignment). Fig 3.1 shows the data of the proteins taken from the Uni-Port when searching for a specific protein and Fig 3.2 shows the FASTA data of a protein when searched in NCBI.

UniProtKB 6,694 results

Entry	Entry Name	Protein Names	Gene Names	Organism
Q9BX50	COPA1_HUMAN	Collagen alpha-1(XXV) chain[...]	COL25A1	Homo sapiens (Human)
P29216	A4_MACMU	Amyloid-beta precursor protein[...]	APP	Macaca mulatta (Rhesus macaque)
Q12830	BPTF_HUMAN	Nucleosome-remodeling factor subunit BPTF[...]	BPTF, FAC1, FALZ	Homo sapiens (Human)
O94985	CSTN1_HUMAN	Calsyntenin-1[...]	CLSTN1, CS1, KIAA0911	Homo sapiens (Human)
Q99MQ5	COPA1_MOUSE	Collagen alpha-1(XXV) chain[...]	Col25a1	Mus musculus (Mouse)
Q28757	A4_SHEEP	Amyloid-beta precursor protein[...]	APP	Ovis aries (Sheep)

Figure 3.1: Uniport proteins sample image.[19]

RecName: Full=40-kDa huntingtin-associated protein; Short=HAP40; AltName: Full=CpG island protein; AltName: Full=Factor VIII intron 22 protein

UniProtKB/Swiss-Prot: P23610.2

GenPept [Identical Proteins](#) [Graphics](#)

```
>sp|P23610.2|HAP40_HUMAN RecName: Full=40-kDa huntingtin-associated protein; Short=HAP40;
AltName: Full=CpG island protein; AltName: Full=Factor VIII intron 22 protein
MAAAAAGGGGAGGPGPEAGFLARYRLVSNIKKRFLRKPINVAEAGEQFGQLGRELRAQECLPYAAWCQ
LAVARCQQALFHGPGEALALTEAARLFLRQERDARQLRVCPAAYGEPQLQAAAASALGAAVRLHLELGQPA
AAALCCLELAALRLDQPAAGAGHFRQAQQLQPLAALQALGEAASCQLARDYTGALAVFTRIQL
AREHGSHPVQSLRPPPPPAPOPGPGATPALPAALLPPNSGSAAPSPAAALGAFSDLVRCESVRLLLLL
QPPPAKLLPEHAQTLEYNSWEAFDSHQESSGQQLPEELFLLLQSLVWATHEKOTEAIKSLOVEMPLTA
EQJHLHLHLVLTETTSPSGQV
```

Figure 3.2: FASTA data of the HAP40 protein.

B. Methodology:

The Methodology diagram in Figure 3.3 represents an outline of the different stages involved in the current research project. The research mainly consists of two steps: Evolutionary analysis and Protein-Protein Interaction (PPI) analysis. The Evolutionary analysis was conducted on proteins obtained from Uni-Port and Gene-card databases. The analysis identified other proteins that have an impact on Alzheimer's, Parkinson's, and Huntington's disease, after combining the results of Evolutionary and PPI analysis.

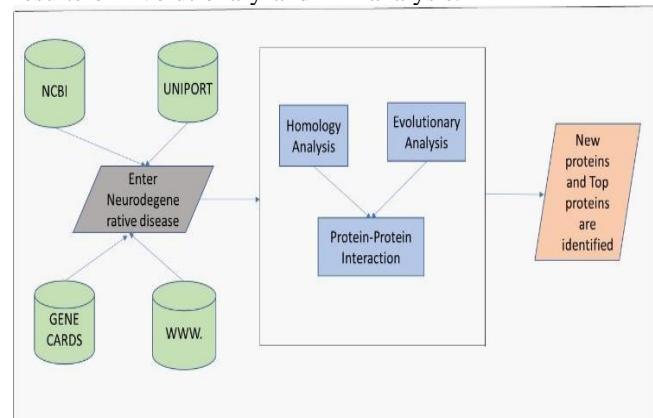


Figure 3.3: Proposed System Diagram

i. Evolutionary analysis:

To learn how the program for multiple sequence alignment can accurately and effectively align 30 sequences, FASTA (fast alignment) of 30 sequences is entered into the cluster omega tool (<https://www.ebi.ac.uk/Tools/msa/clustalo/>). Multiple-sequence alignments that are significant to biology link disparate sequences. Identification of Evolutionary linkages can be done using phylogenograms and cladograms. Phylogenetic trees should be constructed based on the similarities between the most common variables. How similar the proteins are to one another is indicated by the distance between them. To understand the evolution of the proteins, use the phylogenetic tree to locate the most recent predecessors.

Steps Used:

Step 1: Open Cluster omega tool.

Step 2: Enter the FASTA data of all the considered proteins one by one in the supported format.

Step 3: Select the Homo Sapiens category in the proteins.

Step 4: After the results are loaded click on the Phylogenetic tree.

Step 5: Study the Phylogenetic tree and identify similar proteins.

ii. Protein-Protein interaction (PPI) analysis:

Constructing a Protein-Protein Interaction (PPI) network enables the identification of interactions between proteins. The PPI network provides a comprehensive understanding of the relationships between the proteins being analyzed. The PPI network gives us a clear idea of direct and indirect relations among the proteins. To do Protein-Protein Interaction analysis, an online tool known as the string tool (https://stringdb.org/cgi/input?sessionId=bY6X0Yf5DpXN&input_page_show_search=on) has been used.

Steps Used:

Step 1: Open String-db online tool

Step 2: Click on multiple proteins in the list.

Step 3: Enter the names of the similar proteins identified in the previous step.

Step 4: Select the Homo Sapiens in the organism list

Step 5: Review the list of proteins in the Homo Sapiens and click continue.

Step 6: Click on settings and select the confidence as 0.7 and select hide the indirect nodes.

Step 7: Study the PPI network and identify the relationships among the proteins.

IV. RESULTS

The result includes all the 2 analyses the Evolutionary analysis, the Protein-Protein Interaction (PPI) analysis for every disease, and the combination of three diseases to identify the new protein which causes the three diseases

A. Alzheimer's disease:

Evolutionary analysis:

A phylogenetic tree is constructed by taking 10 Alzheimer's disease-causing proteins. The proteins are histamine h2 receptor, Presenilin-1, Alpha-1-antichymotrypsin (precursor), Beta-site amyloid precursor protein cleaving enzyme 1, Cathepsin D, Amyloid-beta precursor protein (APP), Microtubule-associated protein tau, Calsenilin, Apolipoprotein, Apoptosis-antagonizing transcription factor (AATF) and these 10 proteins share a similarity of 0.42417, 0.39183, 0.40782, 0.35004, 0.34837, 0.40187, 0.38599, 0.40064, 0.36023, and 0.36479. From the similarity, Amyloid-beta Precursor Protein (APP),

Microtubule-associated protein tau, Presenilin-1, and Alpha-1-antichymotrypsin (precursor) proteins are considered.

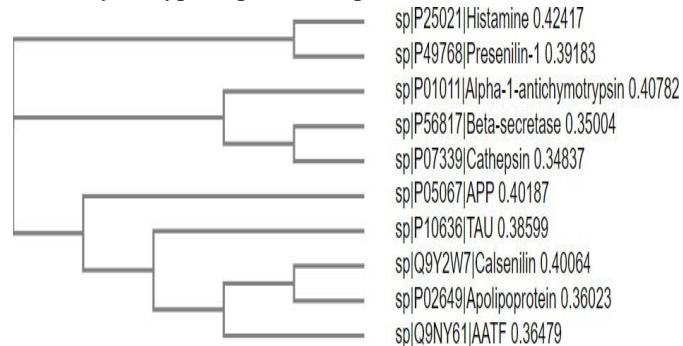


Figure 4.1: Evolutionary analysis of 10 Alzheimer's proteins.

Protein-Protein interaction (PPI) analysis:

Amyloid-beta Precursor Protein (APP), Microtubule-associated protein tau (MAPT), Presenilin-1, and Alpha-1-antichymotrypsin (precursor) are the four proteins to which Protein-Protein Interaction (PPI) is applied to give an idea of how all four proteins interact with one another using the String tool, the PPI network is created for the 4 proteins and identified relationships among them.

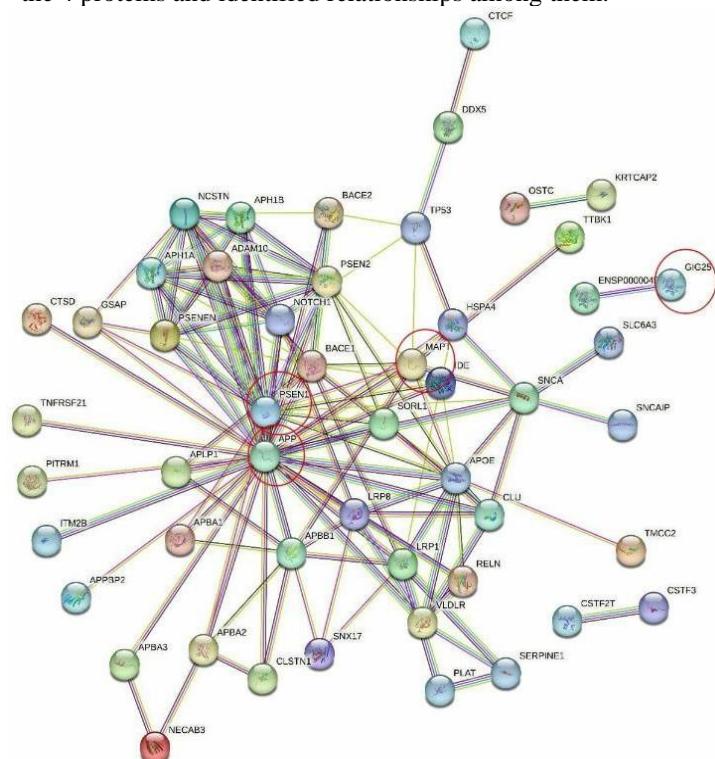


Figure 4.2: protein-protein interaction (PPI) analysis of similar proteins in Alzheimer's disease.

B. Parkinson's disease:

Evolutionary analysis:

A phylogenetic tree is constructed by taking 10 Parkinson's disease-causing proteins. They are PINK1 (Serine/threonine-protein kinase PINK1), AAKG2 (activated protein kinase subunit gamma-2), DDT4 (DNA damage-inducible transcript 4 protein), FBX7 (F-box only protein 7), CERU (Ceruloplasmin), PARK7 (Parkinson disease protein 7), TERA (Transitional endoplasmic reticulum ATPase),

RNF11 (RING finger protein 11), HTRA2 (Serine protease HTRA2), SNCAP (Synphilin-1), CHM2B (Charged multivesicular body protein 2b) and these 10 proteins share a similarity of 0.3841, 0.40696, 0.38932, 0.37513, 0.40991, 0.43011, 0.36111, 0.40504, 0.35847, 0.36044, 0.44145. From the similarity PINK1, PARK7, and SNCA proteins are considered.

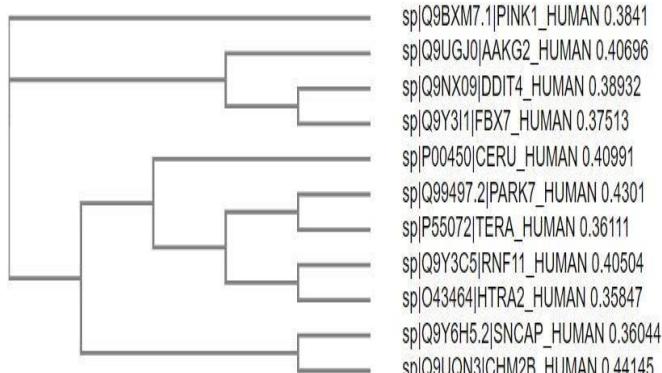


Figure 4.3: Evolutionary analysis of 10 Parkinson's proteins.

Protein-Protein interaction (PPI) analysis:

Protein-Protein Interaction (PPI) is applied to PINK1, PARK7, and SNCA to provide an understanding of the interactions that the three proteins have with one another using the String tool, the PPI network is created for the 3 proteins and identified relationships among them.

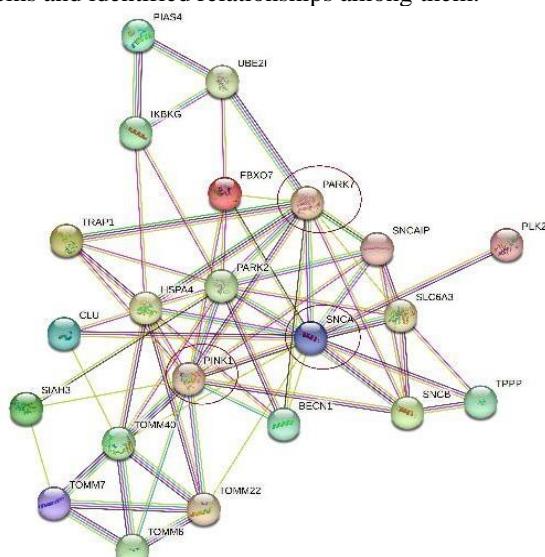


Figure 4.4: protein-protein interaction (PPI) analysis of similar proteins in Parkinson's disease.

C. Huntington's disease:

Evolutionary analysis:

A phylogenetic tree is constructed by taking 10 Huntington's disease-causing proteins of the Prothrombin, Alpha-1-antitrypsin, Nucleolar protein 14, Complexin-2, Huntington-associated protein 1, Fibroblast growth factor receptor 3, Junctophilin-3, Huntington, Huntington-interacting protein 1, Histone-lysine N-methyltransferase and these 10 proteins share a similarity of -0.11122, 0.11122, 0.2197, 0.33045, 0.3684, 0.38869, 0.40244, 0.40252, 0.40868, 0.3707. From the similarity Huntington-associated

protein 1, Huntington protein and Huntington-interacting protein 1 are considered.

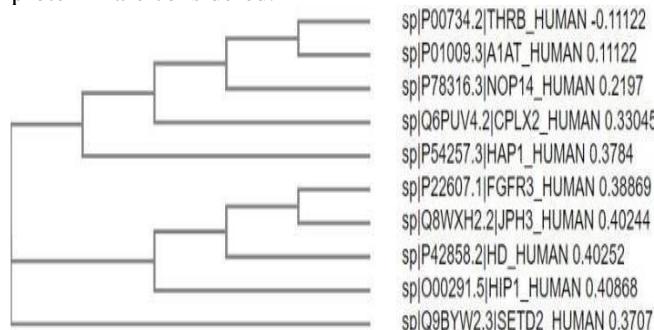


Figure 4.5: Evolutionary analysis of 10 Huntington's proteins.

Protein-Protein interaction (PPI) analysis:

The Protein-Protein Interaction (PPI) method is used for the three proteins Huntington-associated protein 1 (HAP1), Huntington-interacting protein 1 (HIP1), and Huntington (HTT) to provide an understanding of how each protein interacts with the other. Using the String tool, the PPI network is generated for the three proteins and interactions between them are established.

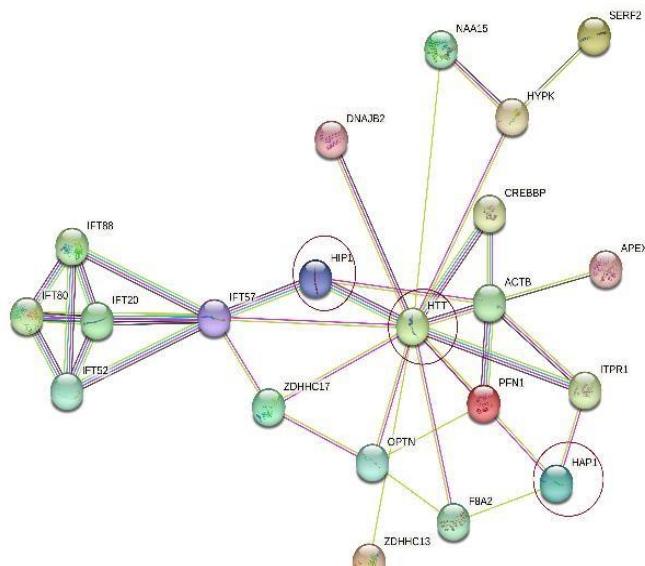


Figure 4.6: protein-protein interaction (PPI) analysis of similar proteins in Huntington's disease.

D. Combination of 3 Diseases (Alzheimer's, Parkinson's, and Huntington's Diseases):

Evolutionary analysis of Alzheimer's, Parkinson's, and Huntington's Diseases:

By combining all the 30 proteins of the 3 diseases Alzheimer's, Parkinson's, and Huntington's Diseases a Phylogenetic tree is constructed and similar proteins among all three diseases are identified. This Evolutionary helps to identify the new proteins which have an involvement in all three diseases together and have an influence on causing the three diseases combinedly.

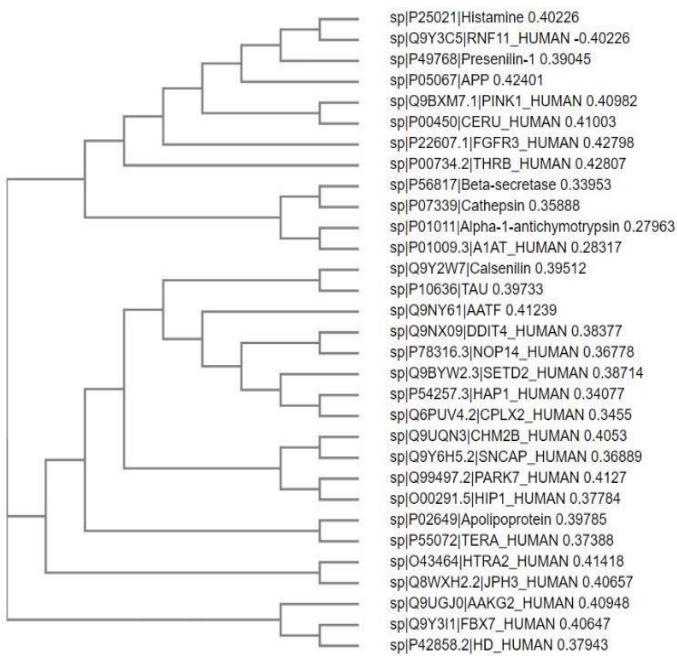


Figure 4.7: Evolutionary analysis of 30 proteins.

Protein-Protein interaction (PPI) analysis of Alzheimer's, Parkinson's, and Huntington's Diseases:

For the proteins under consideration, Amyloid-beta Precursor Protein (APP), Microtubule-associated protein tau (MAPT), Presenilin-1, Alpha-1-antichymotrypsin (precursor), PINK1, Synuclein Alpha (SNCA), PARK7 Huntingtin (HTT), Huntingtin-Associated Protein 1 (HAP1), Huntingtin Interacting Protein1 (HIP1) network has been created that illustrates the interactions between the proteins. From the network, it is identified that HTT, the main Huntington's disease-causing protein is related to SNCA, PINK1, PARK7, the Parkinson's disease-causing proteins directly, and APP, the Alzheimer's disease-causing protein directly. It is also identified that APP has a direct relation with SNCA, a Parkinson's disease-causing protein. So, it is observed that Alzheimer's, Parkinson's, and Huntington's diseases have some similarities among them.

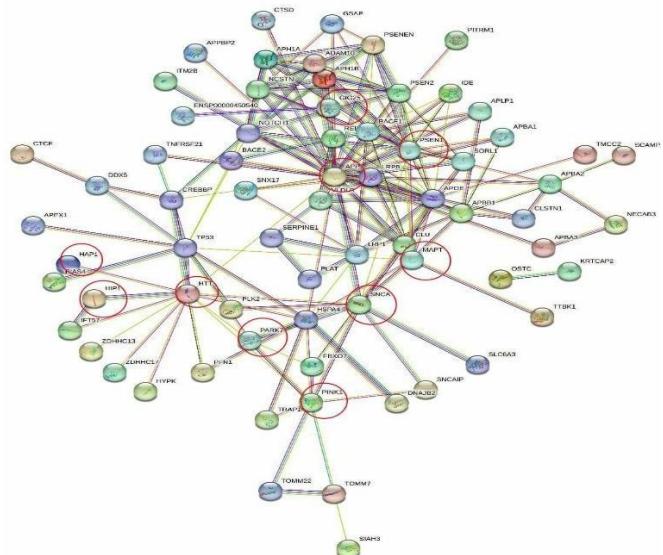


Figure 4.8: Protein-Protein interaction (PPI) analysis of neurodegenerative diseases (Alzheimer's, Parkinson's, and Huntington's Diseases).

V. CONCLUSION

The focus of this study is on the three most prominent neurodegenerative diseases, namely Alzheimer's, Parkinson's, and Huntington's diseases. For each disease 10 proteins were considered from the literature surveys and the score of the gene card, then the 10 proteins from each disease were examined using evolutionary analysis separately and found 4 proteins from Alzheimer's, 3 proteins from Parkinson's, and 3 proteins from Huntington's disease are having high similarity. The 4 proteins from Alzheimer's disease are Amyloid-beta Precursor Protein (APP), Microtubule-associated protein tau (MAPT), Presenilin-1, Alpha-1-antichymotrypsin (precursor), 3 proteins from Parkinson's disease are PINK1, Synuclein Alpha (SNCA), PARK7 and 3 proteins from Huntington's disease are Huntingtin (HTT), Huntingtin-Associated Protein1 (HAP1), Huntingtin-Interacting Protein1 (HIP1). Similarly, Evolutionary analysis Protein-Protein Interaction analysis was applied to these similar proteins of 3 diseases and built a PPI network to know the direct or indirect connections between the proteins and revealed that APP, PSEN1, MAPT, GIG25, PINK1, PARK7, SNCA, HTT, HAP1, and HIP1 are the proteins having relation among themselves and may be the reason for causing these three diseases i.e. Alzheimer's, Parkinson's, and Huntington's diseases. The Future work includes Further verification and Domain-Domain analysis to find the new proteins that cause the disease and apply the structure level analysis.

REFERENCES

- [1] Uttara B, Singh AV, Zamboni P, Mahajan RT, Oxidative stress and neurodegenerative diseases: a review of upstream and downstream antioxidant therapeutic options. *Curr Neuropharmacol*, 2009 Mar;7(1):65-74.
- [2] Kumar A, Sidhu J, Goyal A, Alzheimer Disease, [Updated 2022 Jun 5]. In: StatPearls [Internet], Treasure Island (FL): StatPearls Publishing, 2022 Jan.
- [3] Kouli A, Torsney KM, Kuan WL, Parkinson's Disease: Etiology, Neuropathology, and Pathogenesis. In: Stoker TB, Greenland JC, editors. Parkinson's Disease: Pathogenesis and Clinical Aspects [Internet]. Brisbane (AU): Codon Publications; 2018 Dec 21, Chapter 1.
- [4] Ajitkumar A, De Jesus O, Huntington Disease [Updated 2022 Oct 7]. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2022 Jan.
- [5] Srinivasa Rao V, Srinivas K, Kumar GN, Sujin GN, Protein interaction network for Alzheimer's disease using a computational approach. *Bio information*. 2013 Dec;6(9):968-72.
- [6] Hussain B, Khalid H, Nadeem S, Sultana T, Aslam S, Phylogenetic and chronological analysis of proteins causing Alzheimer's, Parkinson's and Huntington's diseases. *Int. J. Bio automation* 2012 16, 165–178.
- [7] Karbalaei R, Allahyari M, Rezaei-Tavirani M, Asadzadeh-Aghdaei H, Zali MR, Protein-protein interaction analysis of Alzheimer's disease and NAFLD based on systems biology methods unhide common ancestor pathways. *Gastroenterol Hepatol Bed Bench*. 2018 Winter;11(1):27-33.
- [8] V. Srinivasa Rao, K. Srinivas, G. N. Sujini and G. N. Sunand Kumar, Protein-Protein Interaction Detection: Methods and Analysis. *Proteomics* 2014, 341–365 (2014).
- [9] Hernandez SM, Tikhonova EB, Karamyshev AL, Protein-Protein Interactions in Alpha-Synuclein Biogenesis: New Potential Targets in Parkinson's Disease. *Front Aging Neurosci*. 2020 Mar 17; 12:72.

- [10] Botelho J, Mascarenhas P, Mendes J, J & Machado V, Network protein interaction in Parkinson's disease and periodontitis interplay: A preliminary bioinformatic analysis. *Genes (Basel)*. 2020, Vol:11, pp:1–51.
- [11] Goñi J, Esteban FJ, de Mendizábal NV, Sepulcre J, Ardanza-Trevijano S, Agirrezabal I, Villoslada P, A computational analysis of protein-protein interaction networks in neurodegenerative diseases. *BMC Syst Biol*. 2008 Jun 20; 2:52.
- [12] Rakshit, Hindol & Rathi, Nitin & Roy, Debjani, Construction and Analysis of the Protein-Protein Interaction Networks Based on Gene Expression Profiles of Parkinson's Disease. *PloS one*. 9. e103047.
- [13] Chakraborty, Souvik & Maiti, Tarasankar, Identification of potential proteins and microRNAs in Multiple Sclerosis and Huntington's diseases using in silico methods (2021).
- [14] Heinz A, Schilling J, van Roon-Mom W & Krauß S, The MID1 Protein: A Promising Therapeutic Target in Huntington's Disease. *Genet*, 2021, Vol:12, pp: 1–7.
- [15] Xie Y, Hayden MR, Xu B, BDNF overexpression in the forebrain rescues Huntington's disease phenotypes in YAC128 mice. *J Neurosci*, 2010 Nov 3;30(44):14708-18.
- [16] Nopoulos PC, Huntington disease: a single-gene degenerative disorder of the striatum. *Dialogues Clin Neurosci*, 2016 Mar;18(1):91-8.
- [17] Proskura AL, Vechkapova SO, Zapara TA, Ratushniak AS, [Protein-Protein Interactions of Huntingtin in the Hippocampus]. *Mol Biol (Mosk)*, 2017 Jul-Aug;51(4):734-742. Russian.
- [18] Munjal G, Hanmandlu M, Srivastava S, Phylogenetics Algorithms and Applications. *Ambient Communications and Computer Systems*. 2018 Dec 10; 904:187–94.
- [19] 'UniProt', 'Find Your Protein', <https://www.ncbi.nlm.nih.gov/>, Accessed on 13-03-2023

Segmentation of Cell Objects in IIF Images using Split-Bregman Global Convex Optimization

Yasasvi Prasad. K

Dept. of Biomedical Engineering

SRM Institute of Science and Technology

Chennai, India

kp5977@srmist.edu.in

Sai Amarnath. G

Dept. of Biomedical Engineering

SRM Institute of Science and Technology

Chennai, India

gg6594@srmist.edu.in

Suhail Parvaze. P

R & D Consultant

Philips Innovation Campus

Bengaluru, India

suhailsp@gmail.com

Kanchana. D

Dept. of Biomedical Engineering

SRM Institute of Science and Technology

Chennai, India

kanchand2@srmist.edu.in

Abstract—In this work, an attempt has been made to segment the cell objects in homogeneous pattern in Indirect Immunofluorescence (IIF) images. Homogenous pattern is indicative of the several autoimmune diseases but it is most commonly observed in lupus or Sjogren's syndrome. The IIF images for this study are obtained from a publicly available dataset. The images are preprocessed using various intensity normalization techniques such as Histogram Equalization (HE), Adaptive Histogram Equalization (AHE), Local Contrast Enhancement (LCE) and Intensity Adjustment (IA). The preprocessing methods are validated by obtaining image intensity profile of the preprocessed images. Comparison of intensity profile indicates that the IA technique has better contrast enhancement than the other three methods. After preprocessing, segmentation is carried out in the intensity adjusted IIF images. Three different segmentation techniques namely Watershed Segmentation (WS), Super pixel based Fast-Fuzzy C-Means Segmentation (SPFFCMS) and Split-Bregman Global Convex Segmentation (SP-GCS) are implemented. The segmentation results are validated with cell object extraction in the binary images. Results indicate that SP-GCS method has extracted more cell objects than WS and FFCMS algorithm. The binary masks extracted from SP-GCS method are then multiplied with the original images to obtain the segmented regions. The images could be further utilized for classification of homogenous pattern in IIF imaging. From this study, it can be concluded that IA method has better contrast enhancement for homogenous pattern and SP-GCS method has better segmentation results. Hence, this study could assist the clinicians in the identification of homogenous pattern for the diagnosis of lupus.

Keywords—*Indirect Immunofluorescence, Homogenous pattern, Lupus, Sjogren's syndrome, Histogram Equalization, Adaptive Histogram Equalization, Local Contrast Enhancement, Image Intensity Adjustment, Watershed Segmentation, Split-Bregman Global Convex Segmentation*

I. INTRODUCTION

Human immune system protects the body from several infectious diseases by attacking the pathogens. The condition wherein the body's immune system is attacking the own healthy cells and tissues of the body is called autoimmune disease. There are around eighty different types of autoimmune diseases [1, 2]. Each autoimmune disease has specific clinical indication depending on the region where it is affected [3].

Around 4% of the population in the world is affected by one or more autoimmune disease known to the scientific community. The third most common type of autoimmune

disease found in India is Lupus which affects around 9 out of 10 women [4].

Indirect Immunofluorescence (IIF) imaging is the most common method of screening for autoimmune diseases. This method was initially used to diagnose cardinal autoantibodies. Over the last two decades, a lot of progressive research has taken place in the diagnosis of autoimmune diseases [5, 6]. There are various immunometric methods that are designed to identify specific autoantibodies directed against the specific autoantigens. Currently, IIF imaging is the gold standard for the diagnosis of autoimmune diseases. This method uses Human Epithelial Type-2 (HEp-2) cell as a substrate for the detection of specific antigens [7, 8].

According to the recent statement issued by American College of Rheumatology (ACR), the IIF method is able to detect the Antinuclear Antibodies (ANA) for the antigens specific for autoimmune diseases. There are multiple ANA patterns associated with various diseases [9, 10]. The most commonly observed ANA patterns are homogenous, speckled, nucleolar, centromere, nuclear membrane and mitotic spindle. Each ANA pattern is specific for one or more autoimmune diseases. The biomedical industry has developed scientific solutions to automate IIF screening procedure from preparation of slides, substrates to microscopic observations [11, 12].

Contrast enhancement is the process by which the image quality is improved by increasing the brightness of the objects in the image. There are two different approaches in contrast enhancement namely direct and indirect methods [13]. In direct method, the image contrast is enhanced by using an inbuilt function whereas in indirect method, there is no specific term for increasing the contrast. On the basis of image transformation techniques, these contrast enhancement methods are further divided into spatial domain and frequency domain methods [14].

Several researchers have explored the contrast enhancement methods. Histogram Equalization is a method in which the pattern is improved by eliminating the effect of stain. A gray level transform is applied which flattens the histogram of the image resulting in contrast enhancement [15]. Image Enhancement (IE) is one of the most commonly used preprocessing method used in digital as well as medical image processing [16]. In several deep learning-based image classification methods, IE is used for

normalizing the image intensity. Among all the IE methods, Contrast stretching or Intensity Adjustment (IA) is the most commonly used technique to increase the input image intensity range to the expected intensity range [17].

Image segmentation is the process in which similar pixels are grouped together. In mathematical analysis, this method is defined as clustering and it is used in several image processing applications. The primary algorithm behind segmentation is either region dividing or merging depending on the type of clustering used. Optimization based segmentation has been introduced to segregate objects based on edges, dissimilar pixels and consistency in between regions [18].

In HEp-2 cell segmentation, there are several challenges because of the nature of the variability in cell images. The most common challenges are as follows: (i) variations in staining pattern and image intensity (ii) presence of mitotic cells (iii) image artifacts due to variation in illumination. To develop a fully automated IIF imaging system, segmentation plays a key role in identifying the cell structures and patterns [19].

II. METHODOLOGY

The proposed approach consists of two modules: Pre-processing and segmentation. The block diagram representation of the proposed approach is shown in the figure 1.

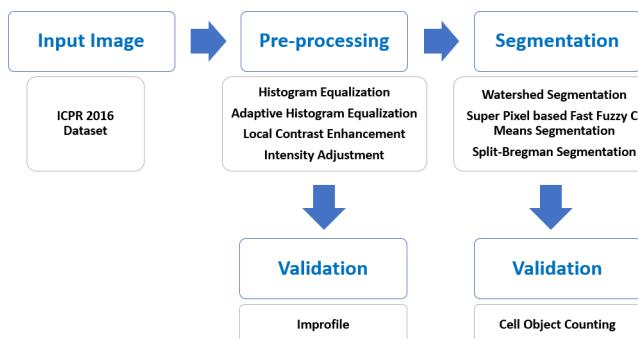


Fig. 1. Workflow for Segmentation of IIF Images

In the pre-processing module, the input images obtained from the dataset are subjected to contrast enhancement using various pre-processing techniques and the results are validated by taking the intensity profile of the enhanced images. In the second module, segmentation of the pre-processed images is carried out to extract the cell objects. The segmentation results are validated by counting the number of cell objects before and after segmentation.

A. Database Description

The HEp-2 specimen images are considered from the publicly available International Conference on Pattern Recognition (ICPR) 2016 database [20]. This dataset was procured at Sullivan Nicolaides Pathology Laboratory, Australia in 2013. It was gathered from 1008 patient sera with ANA test confirming positive. Each sample was diluted in the ratio of 1:80 and a microscope containing monochrome camera was used to photograph the slide specimen. Each sample was imaged at four different locations rendering four images per specimen.

B. Pre-processing

Preprocessing is one of the preliminary steps in image processing where the input image is modified with appropriate preprocessing methods. This step improves the image features that are relevant such that it can be used for further processing. The most common preprocessing methods are filtering and contrast enhancement [7, 19]. Contrast enhancement technique improves the overall image intensity in two ways: Local contrast enhancement and Global contrast enhancement. Local contrast enhancement methods operate in neighboring pixel regions whereas global contrast enhancement methods operate in the whole image. Four contrast enhancement algorithms are utilized in this study.

Histogram Equalization (HE)

In this method, the intensity values of the image are transformed such that the histogram values match a specific histogram value. This method of contrast enhancement is generally used to adjust the overall image intensity [10].

Adaptive Histogram Equalization (AHE)

In this method, the grayscale images are enhanced by transforming the values of the image. This method uses local image enhancement by operating in the neighborhood pixels of the image [11].

Local Contrast Enhancement (LCE)

This method enhances the image contrast by detecting the boundaries in an image. Two parameters namely amplitude and amount of enhancement are defined to perform this operation. For an image, A is the minimum amplitude required to enhance the image contrast and B is the amount of contrast enhancement.

Intensity Adjustment (IA)

This method enhances the contrast of the image by specifying a range within which image enhancement will be performed. The minimum and maximum threshold are set to perform contrast adjustment.

C. Segmentation

Segmentation is the process of dividing image into multiple partitions to analyze the characteristics of the image. It mainly involves separating the foreground objects from background bases on the differences in terms of color, shape and texture. The image segmentation process is widely utilized in image processing for facial recognition and medical image classification [7]. In this study, three different segmentation algorithms are used: Watershed Segmentation (WS), Super pixel based Fast Fuzzy C-Means Segmentation (SPFFCMS) and Split-Bregman Global Convex Segmentation (SP-GCS).

Watershed Segmentation (WS)

In Watershed segmentation, the foreground objects are detected by extracting the boundaries and splitting it into markers for creating watershed. The process separates the foreground and background [7, 15].

Super Pixel based Fast Fuzzy C-Means Segmentation (SPFFCMS)

In this method, the fast fuzzy C-means clustering is incorporated into the conventional super pixel segmentation.

The regions obtained by super pixel partition are converted into histograms. Then Sobel operator is applied to generate gradient images. To obtain image boundaries of the super pixel regions, WS and multiscale morphological gradient reconstruction are performed [21].

Split-Bregman Global Convex Segmentation (SP-GCS)

In SP-GCS method, the L1-regularized problem based on convex optimization is solved during segmentation process. This method is mainly used for denoising images while preserving the image features such as edges [6, 18]. The algorithm for SP-GCS segmentation is as follows:

1. The edges are detected in image by the following equation

$$g(\xi) = \frac{1}{1 + \beta|\xi|^2}, \quad (\text{Eq.1})$$

2. Using the flow equation, the curve evolution is calculated as

$$\varphi_t = (g\nabla \cdot \frac{\nabla\varphi}{|\nabla\varphi|} - \langle \nabla g, \nabla\varphi \rangle) |\nabla\varphi|. \quad (\text{Eq.2})$$

3. The optimization problem can be expressed as

$$\min_{Q, c_1, c_2} \text{Per}(Q) + \mu \int_Q (c_1 - f)^2 + \mu \int_Q (c_2 - f)^2 \quad (\text{Eq.3})$$

4. The general L1-regularized problems is given by the equation

$$\arg \min_u \|\Phi u\|_1 + \frac{\mu}{2} \|Au - fk\|^2 \quad (\text{Eq.4})$$

5. The SP-GCS equation is given by

$$\min_{0 \leq u \leq 1} \|\nabla u\|_g + \mu \langle u, r \rangle \quad (\text{Eq.5})$$

$$\text{where } r = (f - c_1)^2 - (f - c_2)^2$$

6. The updation for optimization is given by

$$\begin{aligned} u^{k+1} &= GS_{GCS}(r^k, \bar{d}^k, \bar{b}^k) \\ \bar{d}^{k+1} &= \text{shrink}_g(\nabla u^{k+1} + \bar{b}^k, \lambda) \\ \bar{b}^{k+1} &= b^k + \nabla u^{k+1} - \bar{d}^{k+1} \end{aligned} \quad (\text{Eq.6})$$

Based on the above equations, SP-GCS method is applied to the IIF images to obtain the cell objects in the foreground.

D. Validation

Intensity Profile of Images

The preprocessed images are validated using image intensity profile. This method draws a line segment along the length of the image and calculated the intensity profile at regularly spaced intervals. The improfile function calculates the image intensity profile and plots the values along the line segment along the image.

Cell Object Count

To validate the process of image segmentation, cell objects are counted after extraction of the cell boundaries. The foreground cell objects extracted with binary masks are counted using this method and based on the number of

objects, the segmentation approach with better results can be identified. This process of segmentation validation is helpful whenever the ground truth is not available in the dataset. The images segmented are subjected to cell counting and the method which identifies most cell objects is considered to be the best among the ones utilized for this study.

III. RESULTS AND DISCUSSION

The representative homogenous IIF images and the results of preprocessing are shown in the figure 2. From the figure it can be observed that the intermediate intensity images have less contrast when compared to positive intensity images. Appropriate preprocessing has to be identified to extract cell object information. The preprocessed images using the various methods are shown below. From the results it can be observed that the HE method enhance the background as well as the cell structures. The intensity of the images is very high when compared to original images.

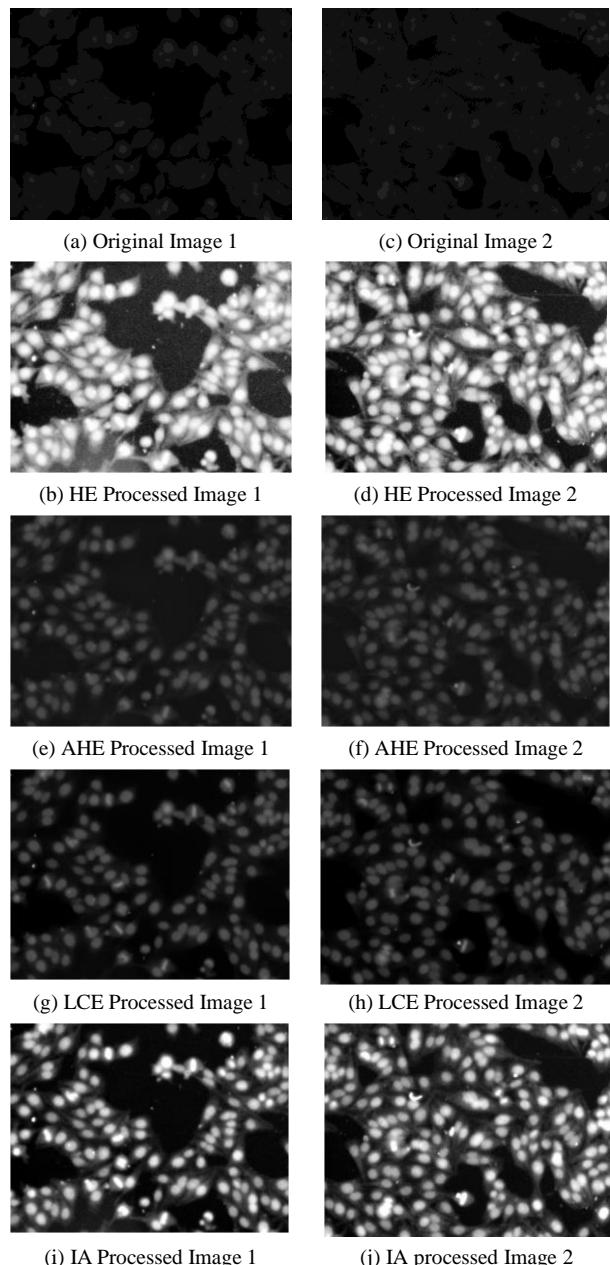


Fig.2. Preprocessed images

Background enhancement along with the cell structures could cause false positive in the classification process. In AHE method, the cell objects as well as the background are enhanced slightly when compared to the original image. However, the cell texture is not clear in this method of preprocessing. In LCE method, the background information is suppressed and the cell structures are enhanced. However, the amount of enhancement and the threshold has to be identified for optimal contrast enhancement. In IA method the cell objects are clear and the intensity is maximum. The background structures enhanced during preprocessing have to be removed by appropriate segmentation technique.

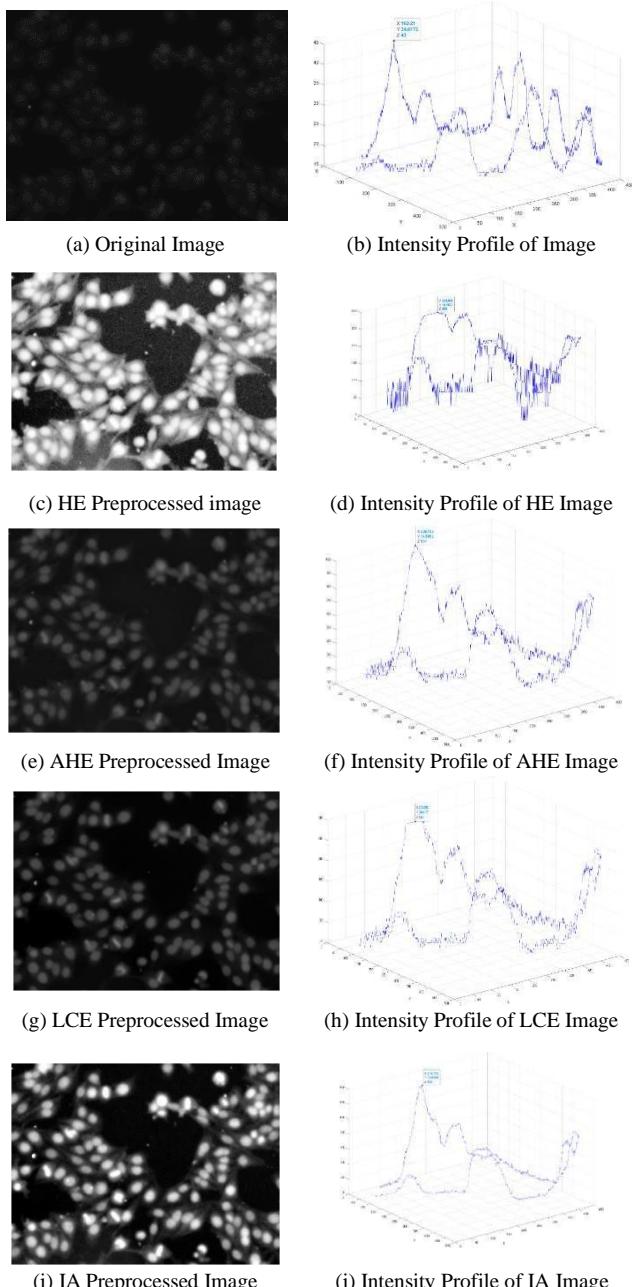


Fig. 3. Intensity Profile Validation

The preprocessing methods are validated by taking intensity profile of the images. The intensity profile of original image and the preprocessed images are shown in the figure 3. From the figure, it could be observed that the intensity variations are profoundly high after contrast

enhancement in all the four methods. Smooth lines are observed for IA processed images when compared to the other three methods indicating that the preprocessing has smoothening effect in the image.

In WS segmentation the images are partitioned into ridges based on the texture. The edges are detected and the cell structures are extracted. The segmented region and cell object boundaries are shown in the figure 4. From the images, it is observed that the cell boundaries are extracted but the image partitions based on texture fails to capture the exact boundary details. The IIF images have different textures based on the patterns and WS is not the ideal method for the extraction of the cell boundary extraction.

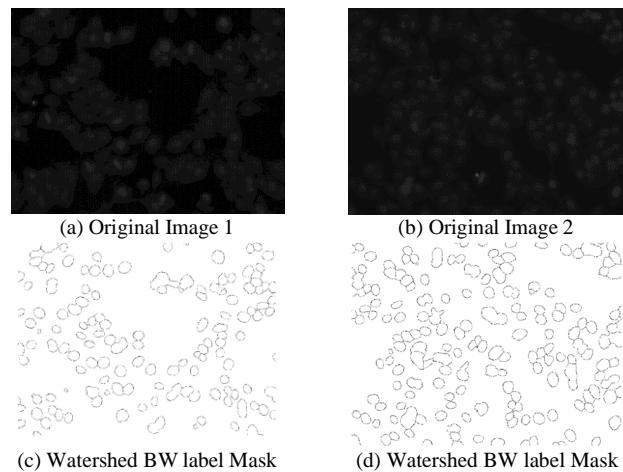


Fig. 4. WS Segmentation Images

In SPFFCMS method, the images are segmented using super pixel extraction and partition using histogram. The extracted boundaries of the final output are shown in the figure 5. It is visible that the cell boundary extraction is not uniform in the images. The blurriness in the boundaries is dependent on the segmentation parameters of the SPFFCMS algorithm. Optimization of the segmentation parameters has to be done to obtain the best results. However, identification of best combination of the parameters is challenging since each IIF image has a unique pattern with variation in textural information.

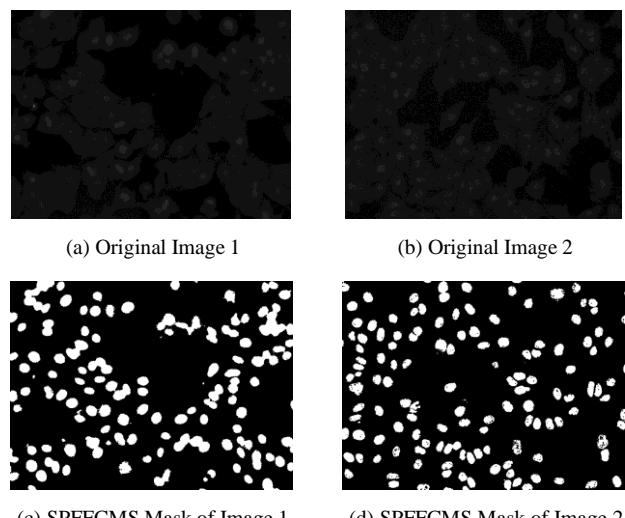


Fig. 5. SPFFCMS Segmentation Images

In SP-GCS segmentation the images are segmented using convex optimization algorithm where in the edges are identified. The extracted cell object boundaries and the final output is shown in the figure 6. From the images, it could be observed that the cell boundaries are extracted in a smooth manner when compared to the other two methods. Also, the overlay with the original images is done with minimal distortion indicating the method has better segmentation of the cell boundaries.

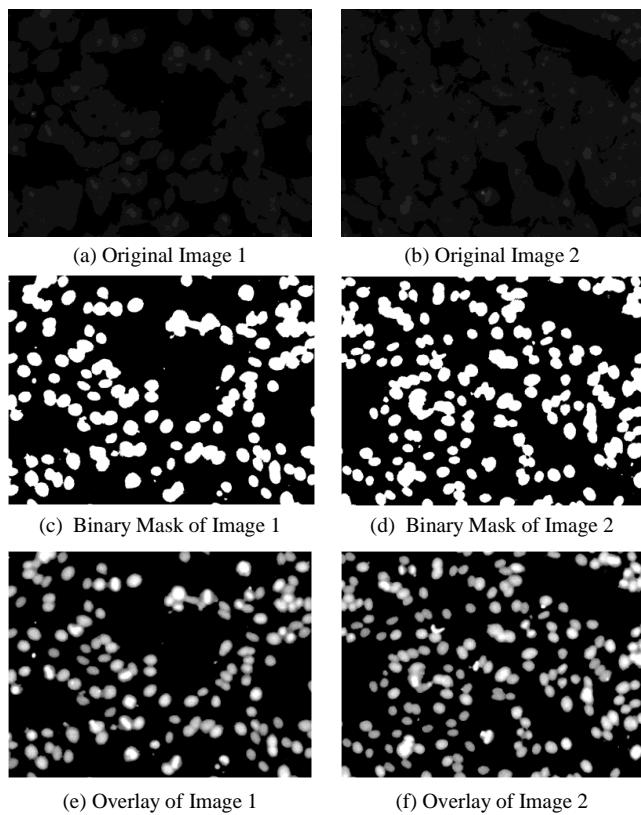


Fig. 6. SP-GCS Segmentation Images

The SP-GCS method validation is shown in figure 7. The number of cell objects in each image is counted and based on that the segmentation accuracy is validated. The cell count indicates that the segmentation method has extracted all the objects in the given image. The cell objects before and after segmentation is found to be equal for SP-GCS method when compared to the other two segmentation algorithm indicating better object extraction.

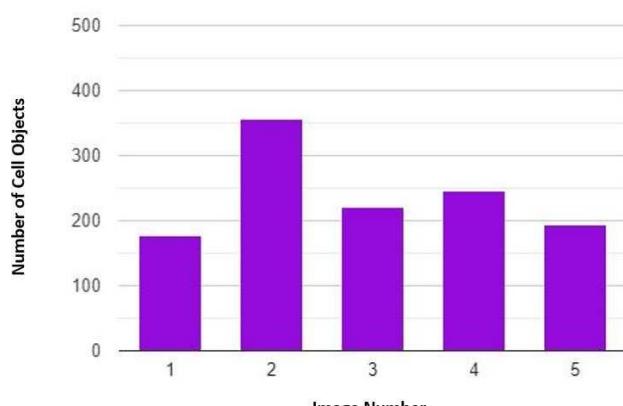


Fig. 7. Cell Count Validation

IV. CONCLUSION AND FUTURE SCOPE

In this study, the most commonly found homogenous ANA pattern was considered. The images are obtained from a public dataset and four different pre-processing methods were performed. The contrast enhancement process is validated with intensity profile analysis. Results indicate that IA method has better contrast enhancement when compared to the other three pre-processing techniques.

Segmentation was performed in the pre-processed images. Three different methods namely WS, SPFFCMS and SP-GCS were performed. The results are validated by counting the cell objects extracted. From the observations, it has been found that SP-GCS method has better segmentation results and was able to extract more cell objects than WS method. Hence for homogenous ANA pattern, IA+ SP-GCS combination has better contrast enhancement and segmentation.

The limitations of the current study are that only homogenous pattern is taken for analysis. The pre-processing based on LCE method was able to suppress the background information and extract the cell objects however the intensity was very low. IA method has improved the overall image intensity but the background was also enhanced along with the cell objects. Validation of segmentation was performed based on cell object extraction since the ground truth information was not available in the database.

In future, this study could be extended by taking all ANA patterns into consideration for pre-processing and segmentation. The validation of pre-processing using contrast enhancement measures could be performed. Segmentation can be done with boundary-based detection techniques. Classification of ANA patterns could be performed after segmentation for identifying various autoimmune diseases.

REFERENCES

- [1] D. A. Smith, D. R. Germole, "Introduction to immunology and autoimmunity", *Environ Health Perspect*, pp. 661-5, 1999.
- [2] J. M . Anaya, Y. Shoenfeld, A. Rojas-Villarraga, R. A. Levy, R. Cervera, "Autoimmunity: from bench to bedside", 2013.
- [3] S. Sahai, M. Adams, D. Kamat, "A Diagnostic Approach to Autoimmune Disorders: Clinical Manifestations: Part-1", *Pediatr Ann*, Vol. 45, No.6, pp.223-9, 2016.
- [4] A. Rigon, P. Soda, D. Zennaro, G. Iannello, A. Afeltra, "Indirect immunofluorescence in autoimmune diseases: assessment of digital images for diagnostic purpose", *Cytometry B Clin Cytom*, Vol.72, No.6, pp.472-7, 2007.
- [5] J. A. Al-Mughales, "Anti-Nuclear Antibodies Patterns in Patients with Systemic Lupus Erythematosus and their Correlation with other Diagnostic Immunological Parameters", *Front Immunol*, pp.13:850759, 2022.
- [6] T. Goldstein, X. Bresson, S. Osher, "Geometric applications of the split Bregman method: segmentation and surface reconstruction", *J Sci Comp*, Vol.45, No.1, pp.272-293, 2010.
- [7] N. Sharma, L. M. Aggarwal, "Automated medical image segmentation techniques", *J Med Phys*, Vol.35, No.1, pp.3-14, 2010.
- [8] A. Kornilov, L. Safonov, I. Yakimchuk, "A Review of Watershed Implementations for Segmentation of Volumetric Images", *J Imag*, Vol.8, No.5, pp.127, 2022.
- [9] M. Sonka, V. Hlavac, R. Boyle, "Image pre-processing. In: Image Processing", *Analy and Mach Vis*, pp.56-111,1993.
- [10] M. Stephen, E. Pizer, A. Philip, John D. Austin, C. Robert, A. Geselowitz, T. Greer, B. H. Romeny, John B. Zimmerman, Z. Karel, "Adaptive histogram equalization and its variations", *Com Vis, Graphics, and Imag Proc*, Vol. 39, No.3, pp.355-368, 1987.

- [11] A. Fava, M. Petri, "Systemic lupus erythematosus: Diagnosis and clinical management", *J Autoimmun*, Vol. 96, pp.1-13, 2019.
- [12] X. Du, and S. Dua, "Segmentation of fluorescence microscopy cell images using unsupervised mining", *Open Med Infor Journal*, Vol.4, pp.41-49, 2010.
- [13] I. Ersoy, F. Bunyak, J. Peng, K. Palaniappan, "HEp-2 cell classification in IIF images using shareboost", *Proc of 21st Inter Conf on Patt Recog*, pp. 3362-3365, 2012.
- [14] T. Goldstein, X. Bresson, S. Osher, "Geometric applications of the split Bregman method: segmentation and surface reconstruction", *J scienc comp*, 45(1), 272-293, 2010.
- [15] A. Lerner, P. Jeremias, T. Matthias, "The world incidence and prevalence of autoimmune diseases is increasing", *Inter J Celiac Dis*, Vol.3, No.4, pp.151-155, 2015.
- [16] K. M. Pollard, "Autoantibodies and Autoimmunity: Molecular Mechanisms in health and Disease", 2006.
- [17] S. Rahman, L. Wang, C. Sun, L. Zhou, "Deep Learning based HEp-2 Image Classification: A Comprehensive Review", *arXiv preprint*, arXiv:1911.08916, 2019.
- [18] A. Saleem, A. Beghdadi, B. Boashash, "Image fusion-based contrast enhancement", *EURASIP J Imag and Video Proc*, Vol.1, No.1, 2012.
- [19] S. Tonti, S. D. Cataldo, A. Bottino, Elisa Ficarra, "An automated approach to the segmentation of HEp-2 cells for the indirect immunofluorescence ANA test", *Comp Med Imag and Graphics*, Vol.40, pp.62-69, 2015.
- [20] B. C. Lovell, G. Percannella, A. Saggesse, M. Vento, A. Wiliem, "International contest on pattern recognition techniques for indirect immunofluorescence images analysis", *Proc of 23rd International Conf on Patt Recog*, pp.74-76, 2016.
- [21] K. Devanathan, D. Eeswar Samhithan. D, "Lesion Segmentation in Dermoscopic Images using Superpixel based Fast Fuzzy C-means Clustering", *Proc of IEEE Congreso Bienal de Argentina*, 2020.

TRACK 4

**Ninth International Conference in Biosignals, Images and Instrumentation,
SSNCE, Chennai, March 16 – 17, 2023**

Facial Expression Recognition using Convolutional Neural Network

Alben Richards MJ

*Department of Electronics Engineering
Madras Institute of Technology, Anna
University
Chennai, India
albenmarks@gmail.com*

Prakash P

*Department of Electronics Engineering
Madras Institute of Technology, Anna
University
Chennai, India
prakashp79@gmail.com*

Kaaviya Varshini E

*Department of Electronics Engineering
Madras Institute of Technology, Anna
University
Chennai, India
kaavyapriya6@gmail.com*

Diviya N

*Department of Electronics Engineering
Madras Institute of Technology, Anna
University
Chennai, India
popdivi25@gmail.com*

Kasthuri P

*Department of Electronics Engineering
Madras Institute of Technology, Anna
University
Chennai, India
kasthurip94@gmail.com*

Abstract—Facial expression is a way of non-verbal communication by using eyes, lips, nose and facial muscles. Smiling and rolling eyes are some examples. Facial expression recognition is the process of extracting facial features from a person. Facial expressions include anger, happy, disgust, sad, neutral, fear and surprise. By the use of machine learning, an expression recognition model is built using Convolutional Neural Network. The input data is fed to the system in order to give the expected results. The model is trained using Facial Expression Recognition (FER) dataset. The Convolutional Neural Network (CNN) gives good and accurate results. The Haar cascade classifier classifies the face and non-face regions in the input image which helps the convolutional network to classify the images. Good classification of images can be desirable by the use of classifiers. These classifiers can be implemented by using the OpenCV library.

Keywords—*Facial Expression Recognition, Convolutional Neural Network, Haar cascade classifier*

I. INTRODUCTION

The visible representation of an individual's affective state, cognitive activity, intention, personality, and psychopathology is their facial expression, which also serves as a means of communication in interpersonal interactions. It has undergone much study and made advances in recent decades. Despite significant advancements, it is still challenging to accurately identify facial expressions because of their complexity and variety. In general, nonverbal cues like involuntary movements, expressions on people's faces, and gestures can be used to communicate intents and feelings. This method has the potential to become an extremely effective nonverbal form of interpersonal communication. The main consideration is how well the system detects and extracts the facial emotion from the image.

The method is widely used in a variety of industries, including lie detection, medical assessment, and human computer interaction. Humans regularly distinguish between emotions by their distinctive facial features as part of an expression. For instance, a smile or a movement of the corners of the lips are certainly signs of satisfaction. Similar to how one emotion can be distinguished from another by additional deformations unique to that emotion. In studies on the automatic identification of facial expressions, the problems related to the representation and classification of static or dynamic features of these deformations of face pigmentation

are addressed expressions. The technique categorises a person's facial expressions based on their basic emotions, including happy, angry, disgust, fear, neutral, sad, and surprise. The fast advancement of Artificial Intelligence (AI) capabilities has led to an increase in interest in automatic Facial Expression Recognition (FER). They now serve a variety of purposes, and human contact with them is growing. Machines must be given the ability to comprehend the environment, especially human intentions, in order to enhance Human Computer Interaction (HCI) and make it more natural. Through cameras and other sensors, machines can record the state of their surroundings. Deep Learning (DL) algorithms have excelled at capturing environmental conditions in recent years. Since facial expressions reveal information about a person's inner state, expression detection is essential for machines to function more effectively. DL techniques can be used by a machine to analyse a series of facial photos and identify.

Several strategies and approaches are employed to resolve face emotion recognition challenges. Support Vector Machine (SVM) classification along with appearance and geometry characteristics was offered as a method for detecting face emotions [1]. The technique makes use of SVM and local region-specific features to identify facial reactions from a single picture frame. [2] suggested a technique utilising Faster R-CNN. The process includes normalising the image of the facial expression, extracting implicit features with a trainable convolution filter, and utilising maximum pooling to reduce the dimensionality of the feature. High-quality region suggestions were then produced by the Region Proposal Networks (RPNs) and put into the Faster R-CNN for detection. With channel selection and the categorization of EEG signals associated to good and negative emotions, emotion detection based on EEG features [3] in video clips is possible. EEG signals are categorised using a multi-layer perceptron neural network and the k-nearest neighbour technique. Initially, the classifier algorithms were applied for selecting the most effective EEG channels, resulting in the identification of the top five channels for each participant that demonstrated the best classification results. These systems involve complex processes to predict the expressions fed to the system. In this paper, a simple method involving CNN and Haar Cascade Classifier (Viola-Jones algorithm) for facial expression recognition is proposed.

II. METHODOLOGY

Convolutional Neural Networks (CNNs) are advanced deep learning techniques designed to process and analyze images by assigning weight and importance to various elements within the image and allowing the network to differentiate between images. This type of neural network comprises a combination of a pooling layer and a convolution layer. The pooling layer selects the most relevant information within a defined area, while the convolution layer condenses the data into a smaller region to extract unique features. Unlike other image classification algorithms, CNNs can learn various filters and characteristics without the need for extensive preprocessing.

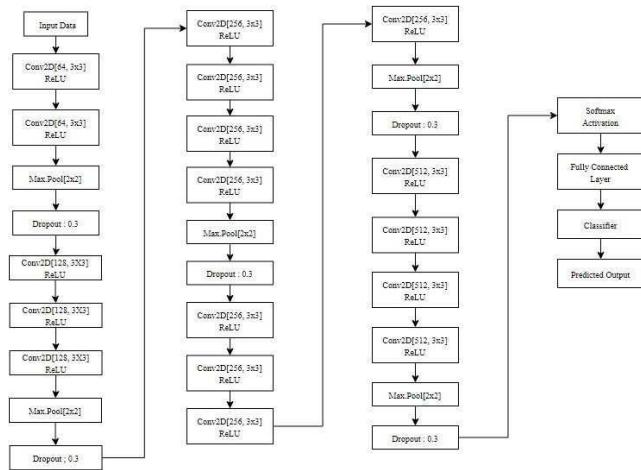


Figure 1. CNN Model Architecture

The specific model described in Figure 1 is comprised of 17 Conv2D convolution layers, activation layers, pooling layers, dropout layers, a fully connected layer, and a softmax activation function. All classifiers were implemented using Keras with Tensorflow as the backend and underwent training. The Haar Cascade Classifier is a method used for object identification, specifically for locating faces in still photos or moving videos. It operates by scanning the picture from the left corner to bottom corner and moving pixel by pixel. The Haar feature searches for specific features and characteristics in the image. The Haar Cascade Classifier is constructed by layering multiple weak classifiers, resulting in a more robust overall classifier. To effectively train the classifier, a sizable number of both positive pictures (those with faces) and negative pictures (those without faces) is required.

III. SYSTEM ARCHITECTURE AND DATASET OVERVIEW

Testing and Training comprises the two sections of the core algorithm. The CNN model must be trained to classify the expressions.



Figure 2. FER Sample Dataset

Few samples of FER Dataset are shown in Figure 2. The algorithm begins by determining if the trained data are present or not. If not, the system must be trained before doing testing for expression classification. The FER dataset is comprised of a set of labeled images that are used for training, development, and testing purposes. The training set includes 28,709 labeled images, the development set includes 3,589 labeled images, and the test set includes 3,500 images. Each image in the FER dataset is assigned one of seven possible facial expressions, including happy, sad, angry, fear, surprise, disgust, and neutral, with happy being the most commonly occurring expression.

The images within the FER dataset feature both posed and unposed headshots and are presented in grayscale format with a resolution of 48x48 pixels. The dataset has 4953 sample images for Anger, 547 sample images for Disgust, 5121 sample images for Fear, 8989 sample images for Happy, 6077 sample images for Sad, 4002 sample images for Surprise and 6198 sample images for Neutral.

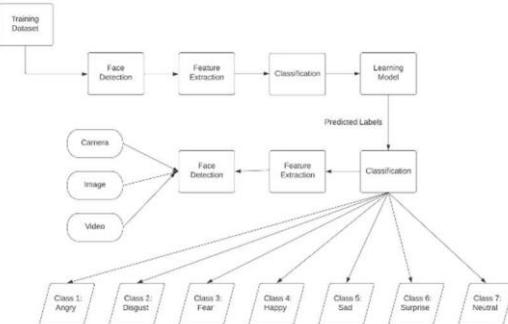


Figure 3. Pipeline of Facial Expression Recognition

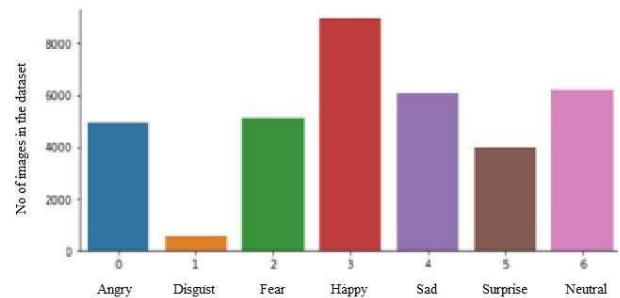


Figure 4. Graphical Representation of FER Dataset

The pipeline of the FER system is shown in Figure 3. The Graphical representation of the image count in FER dataset is given by Figure 4.

IV. IMPLEMENTATION OF THE CNN MODEL

A. Haar Cascade Classifier

Faces can be found in pictures or videos using the object detection method known as Haar Cascade Classifier. It employs Haar features, which move pixel by pixel across the image. A strong classifier is formed by cascading small classifiers to generate the classifier. Numerous positive (faces) and negative (no faces) images are needed for training. OpenCV offers models that have already been trained. It was suggested to use the Integral Image concept, which uses a single value for each feature, to speed up the

process. Each pixel in the original image is equal to the sum of all the pixels to its left and above when calculating the integral image. This lessens the complexity of time and accelerates the procedure.



Figure 5. Haar Features on Human Face

Figure 5 is a visual depiction of the Haar feature traversal as a whole in a human face.

B. Data Preprocessing

Pre-processing, which includes actions like face identification and alignment, lighting correction, position correction, and data augmentation, helps the FER system perform better. Haar Cascade is a classifier used for face detection. For a simpler background, the face has been trimmed. With pixel intensity values restricted to 0-1, images are downsized to 48x48 grayscale and normalized. The model is enhanced with data to make it more resistant to factors that are frequently encountered, such as rotating, flipping, and cropping photographs. As a result, performance is improved by addressing problems like overfitting and data shortages.

C. Feature Extraction and Classification

The detection, feature extraction, classification, and expression evaluation components make up the face expression evaluation system. The objective is to build a webcam-oriented real-time framework for analyzing and classifying human emotions based on facial expressions. Data preprocessing, feature extraction, and emotion classification make up the system's three processes. The model employs 17 convolutional layers, 0.3 dropout rate activation functions, and pooling layers. For expression classification, SoftMax activation is utilized at the end since it provides classification probabilities. Mathematically convolution is given by:

$$O(x, y) = \sum_{k=1}^n \sum_{l=1}^n I(x + k - 1, y + l - 1)K(k, l) \quad (1)$$

where x ranges from 1 to M - m + 1 and y ranges from 1 to N-n+1. The result of the convolution process is described by equation 1 and is referred to as the Feature Map, which provides details on the image's characteristics, that is edges and corners. This Feature Map is then utilized by other layers to uncover additional features of the input image. After the convolution layer, the convolution operation is performed on the input, it forwards the outcome to the subsequent layer.

The CNN model processes the picture data to extract features. Filters are used in convolution to produce a feature map. ReLU activation introduces nonlinearity, followed by Max Pooling. Dropout Layer sporadically deactivates some neurons. The result is delivered to a SoftMax activation function, which transforms real values into probabilities and

uses them to forecast the output emotion class, after passing through all convolutional layers.

0.043	→	0	Angry
0.035	→	0	Disgust
0.030	→	0	Fear
0.79	→	1	Happy
0.045	→	0	Sad
0.020	→	0	Surprise
0.037	→	0	Neutral

Figure 6. Softmax Activation Function Sample Output

From Figure 6, it is inferred that happy emotion has the highest probability value. Hence with respect to one hot encoding the classifier will classify the expression as happy.

D. Interfacing with webcam

A trained CNN model analyses a live image taken by a webcam of a subject to detect facial emotions such as happy, anger, neutral, surprise, disgust, fear and sad. The expression's label is generated by the model and displayed as the outcome. The output from the object file will then be compared to the person's taken image, after which the matching labels will be shown.

V. RESULTS AND DISCUSSION

A. Evaluation Parameters

The performance of a model is evaluated using accuracy, loss, and confusion matrix as metrics. Accuracy is a metric used to evaluate the ability of a model to identify patterns and relationships within a dataset based on the input or training data. True Negative (TN) and True Positive (TP) results make up correct predictions, while all predictions are made up of Negative (N) and Positive (P) examples. P includes both True Positive (TP) and False Positive (FP) results, and N includes both True Negative (TN) and False Negative (FN) results. A correct emotional prediction is represented by TP, an incorrect emotional prediction is represented by FP, a correct forecast of an incorrect expression is represented by TN, and an incorrect prediction of an incorrect expression is represented by FN. The cross-entropy loss function, also referred to as logarithmic loss, log loss, or logistic loss, is used to calculate a score or loss that disincentivizes the predicted class based on the probability of how it differs from the actual expected class. This loss is calculated when adjusting the model weights during training, with the goal being to reduce the loss. The smaller the loss, the better the model is considered to be. A perfect model has a cross-entropy loss of zero. Cross-entropy is given by equation 2.

$$L_{CE} = - \sum_{i=1}^n t_i \log(p_i) \quad (2)$$

where t_i is the truth label, n is the number of classes which is expressions and p_i are the Softmax probability for the i^{th} class. Adam is an optimization strategy that can be applied to iteratively modify network weights based on training instead of the traditional stochastic gradient descent method.

The alluring advantages of applying Adam includes efficient in terms of computing, and requiring little memory for problems with plenty of data or parameters, suitable for non-stationary goals, suitable for issues with extremely noisy or sparse gradients.

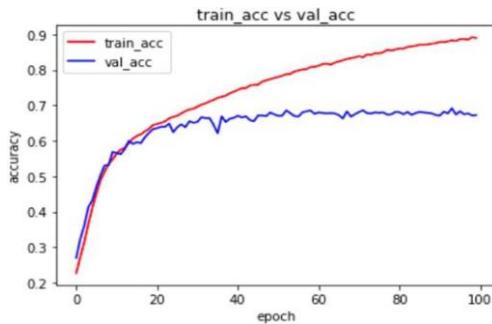


Figure 7. Validation and Training Accuracy for 100 Epochs

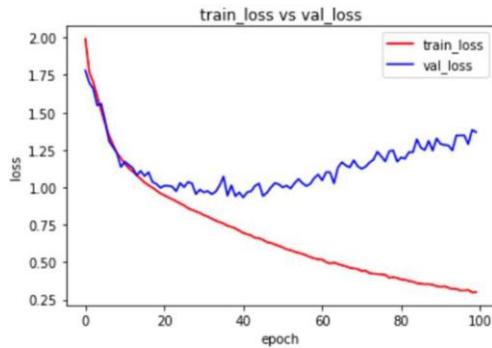


Figure 8. Validation and Training Loss for 100 Epochs

The model is trained for 100 epochs and the Figures 7 and 8 gives the Accuracy and Loss generated. Table 1 shows that the CNN model was able to predict emotions correctly with an average training accuracy, validation accuracy, training loss, and validation loss. The parameters are represented as percentages.

Table 1. Accuracy and losses for training and validation

MODEL	ACCURACY(%)	LOSS(%)
Training	89.46	0.30
Validation	70.45	1.26

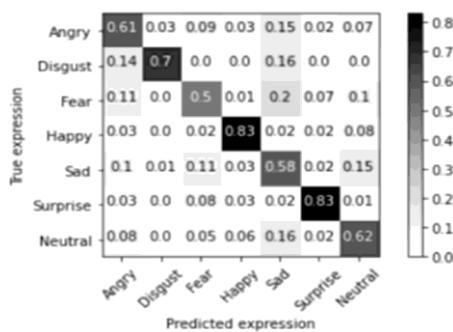


Figure 9. Confusion Matrix of the CNN model

The confusion matrix of the CNN model developed is shown in Figure 9.

B. Computational Results

Real time implementation of the Facial Expression Recognition (FER) CNN model for each of the seven expressions is shown in the below figures. Confidence is given by the maximum and rounded value of the output list from the Softmax Activation Function. The examples of the expressions detected are given below:



Figure 10. Expression Detected: Fear

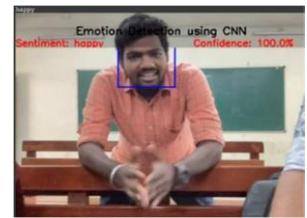


Figure 11. Expression Detected: Happy



Figure 12. Expression Detected: Sad



Figure 13. Expression Detected: Angry

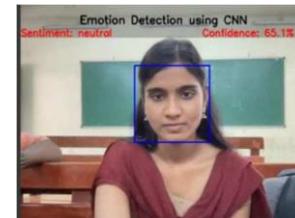


Figure 14. Expression Detected: Neutral



Figure 15. Expression Detected: Disgust



Figure 16. Expression Detected: Surprise

VI. CONCLUSION AND FUTURE SCOPE

The study aims to propose a CNN-based method for facial expression recognition that can automatically identify pattern features and overcome limitations from artificially designed features. Nearly 500 real time images were used for validation. The recognition of facial expressions is used in various fields, which includes video surveillance, robotic

vision, security, digital cameras and human-computer interaction. This work involves pre-processing of facial images, followed by feature extraction through a series of CNN layers and activation functions, and then the classification of emotion classes.

The performance of the proposed algorithm is analyzed by the metrics of accuracy, loss, and confusion matrix were utilized. The results obtained from the test data showed that, as illustrated in Figures 11 and 12, the model achieved 95% confidence in predicting the sad expression, 100% confidence in predicting the happy expression, 97% confidence in predicting the fear expression, and 86% confidence in predicting the disgust expression. The FER dataset was used for both training and testing, with a 8:2 split for training and testing samples, respectively.

There are 13,111,167 total parameters in the proposed CNN model, of which 13,103,431 are trainable. The best parameter values were determined using accuracy and loss metrics. The model was tested using a real-time web camera, and can also be trained to predict emotions using live images captured from a camera, such as at traffic junctions, hospitals, or sports stadiums.

REFERENCES

- [1] Ghimire, D., Jeong, S., Lee, J. et al. Facial expression recognition based on local region-specific features and support vector machines. *Multimed Tools Appl* **76**, 7803–7821 (2017).
- [2] Li, Jiaxing & Zhang, Dexiang & Zhang, Jingjing & Zhang, Jun & Li, Teng & Xia, Yi & Yan, Qing & Xun, Lina. (2017). Facial Expression Recognition with Faster R-CNN. *Procedia Computer Science*.
- [3] Özerdem, M.S., Polat, H. Emotion recognition based on EEG features in movie clips with channel selection. *Brain Inf.* **4**, 241–252 (2017)
- [4] L. Zahara, P. Musa, E. Prasetyo Wibowo, I. Karim and S. Bahri Musa, "The Facial Emotion Recognition (FER-2013) Dataset for Prediction System of Micro-Expressions Face Using the Convolutional Neural Network (CNN) Algorithm based Raspberry Pi," 2020 Fifth International Conference on Informatics and Computing (ICIC), Gorontalo, Indonesia, 2020
- [5] Zhang, R. Making convolutional networks shift-invariant again. In International Conference on Machine Learning (ICML), 2019.
- [6] Zhou, P., Han, X., Morariu, V. I., and Davis, L. S. Learning rich features for image manipulation detection. In IEEE International Conference on Computer Vision (ICCV), 2018.
- [7] S. Li and W. Deng, “Reliable crowdsourcing and deep locality-preserving learning for unconstrained facial expression recognition,” *IEEE Trans. Image Process.*, vol. 28, no. 1, pp. 356–370, Jan. 2019.
- [8] H. Ma and T. Celik, “FER-Net: Facial expression recognition using densely connected convolutional networks,” *Electron. Lett.*, vol. 55, no. 4, pp. 184–186, Feb. 2019.
- [9] S. H. Lee, K. N. K. Plataniotis, and Y. M. Ro, “Intra-class variation reduction using training expression images for sparse representation based facial expression recognition,” *IEEE Trans. Affect. Comput.*, vol. 5, no. 3, pp. 340–351, Jul./Sep. 2014.
- [10] D. K. Jain, P. Shamsolmoali, and P. Sehdev, “Extended deep neural network for facial emotion recognition,” *Pattern Recognition Letters*, 2019.
- [11] N. Christou and N. Kanjolya, “Human facial expression recognition with convolution neural networks,” in Third International Congress on Information and Communication Technology. Springer, 2019, pp. 539–545.
- [12] L. Tran, X. Yin, and X. Liu, “Representation learning by rotating your faces,” *IEEE transactions on pattern analysis and machine intelligence (TPAMI)*, vol. 41, pp. 3007–3021, 2018.
- [13] Singh, S., & Nasoz, F. (2020). Facial Expression Recognition with Convolutional Neural Networks. 2020 10th Annual Computing and Communication Workshop and Conference (CCWC)

Infrared Thermograms for Diagnosis of Dry Eye: A Review

Persiya. J
School of Electronics and Engineering
Vellore Institute of Technology, Chennai India
persiya.j2022@vitstudent.ac.in

Sasithradevi A
Centre for Advanced Data Science Vellore Institute of Technology, Chennai India
sasithradevi.a@vit.ac.in

S. Mohamed Mansoor Roomi
Department of Electronics and Communication Engineering, Thiagarajar College Of Engineering, Madurai India
smmrroomi@tce.edu

Abstract — A non-intrusive, contactless temperature measurement technique that provides real-time surface temperature distribution is infrared thermography. Ocular Surface Temperature (OST) can be measured using thermography without harming the subjects. It is used in wide variety of applications, especially in medical applications. Recently Infrared thermal images of eye are used to diagnose and detect many diseases and features of human eye. This paper examines the methods currently in use for diagnosing dry eye. The main focus is on thermal images. Thermographic technique is proved to be a highly sensitive, satisfying method and accurate for detecting eye disorders. Various machine learning and Deep learning algorithms are discussed. Finally, it is concluded as deep learning combined with thermography is more likely to be used to detect dry eye disease.

Keywords—Infrared Thermogram, Dry eye, Machine learning, Deep learning.

I. INTRODUCTION

Multifaceted condition of the ocular surface called dry eye, in which tears can harm the ocular surface and cause irritation, instability of the tear film, and vision problems [12]. Millions of individuals across the world suffer from dry eye disease (DED), which is the most common causes for patients to see an eye professional. Lacrimal functional unit (LFU) dysfunction, which affects the ocular surface (meibomian gland, conjunctiva, and cornea) and lids in addition to the lacrimal glands, is the primary contributing factor to DED. The changed tear film or tear film dysfunction that comes from these dysfunctions is a further consequence [4]. Urban life style that affects productivity at work include prolonged computer use, air conditioner use, exposure to stressful events, and sleep problems, exacerbate the issue [13]. Because individuals spend more time gazing at screens on mobile, Laptop and desktop devices, computer vision syndrome, which causes blurry vision, fatigue, and a decreased quality of life, is becoming more prevalent in today's society. There isn't a single standard test that can be used to diagnose DED as of yet. Figure 1 shows the healthy eye and eye having Dry eye syndrome.

DE symptoms are evaluated using a combination of diagnostic procedures, including the slit lamp test, tear breakup time, conjunctival and corneal staining using LG, clinical history, Schirmer's test and questionnaires [14, 19,

20]. Despite the fact that ophthalmologists and eye doctors regularly employ all of these tests to confirm the diagnosis, the lack of a standardised approach and the poor correlation with the patient's symptoms make them ineffective rendering the diagnostic evaluations as unreliable, inaccurate and incorrectly diagnosing DE [6, 16]. All of these examinations are excessively intrusive, and measurements can easily be influenced by mechanical, chemical, or other simulations [26].

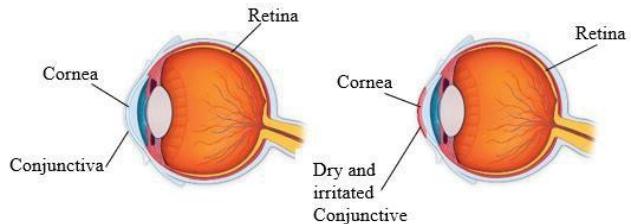


Fig. 1. Normal eye and Dry eye

The temperature of the ocular surface (OST) can now be measured non-invasively utilizing a novel diagnostic technique called infrared thermography [9]. There are several uses for infrared thermography in a variety of applications, including medical diagnosis, where it can be used to identify anomalies in the body by identifying temperature differences between healthy and diseased tissues. By monitoring the physiology of the eyes, thermologists can diagnose dry eye syndromes and other ocular conditions in the field of human ophthalmology [18]. The non-intrusive Infrared thermography (IRT) procedure finds aberrant dry eye temperature behaviors. Researchers use a non-invasive IR thermal imaging to evaluate and study the OST that is influenced by the stability of tear film [7]. Therefore, this paper's main objective is to study the various approaches used to diagnosis of dry eye disease using thermogram.

The sections below are organized as follows: Part II focusses on infrared thermograms. The various approaches utilized to analyze the IR images are briefed in Part III. The methodology is discussed in Section IV. Section V compared and evaluated the effectiveness of the literature review. Section V concludes the overall work by providing a summary.

II. INFRARED THERMOGRAMS

A visual representation of infrared radiation emitted by a scene or an object is called an infrared thermogram. As shown in figure 2 any object can either absorb or transmit or reflect the incident radiation. Infrared radiation, a kind of electromagnetic radiation, is emitted by all objects having a temperature higher than absolute zero.

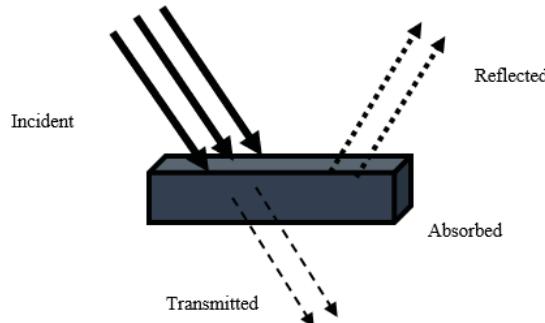


Fig. 2. Radiation Effects

The fundamental principles that control how objects emit and absorb thermal radiation are incorporated into the equations for infrared thermography. According to the Stefan-Boltzmann Law, there is a direct correlation between an object's temperature and the quantity of thermal radiation it emits. It is given in equation 1.

$$Q = T^4 \sigma \quad (1)$$

Where Q is the emissive power, T is absolute temperature and σ is the Stefan-Boltzmann constant. Planck's Law, illustrated in equation 2 describes the spectrum distribution of heat radiation emitted by an object.

$$P_{\lambda,T} = \frac{2hc^2}{\lambda^5 [e^{\frac{hc}{\lambda kT}} - 1]} \quad (2)$$

Where $P_{\lambda,T}$ denotes the spectral radiance, λ is the radiation's wavelength, T is the object's temperature, h denotes the Planck's constant, c denotes the speed of light, and k is the Boltzmann's constant. The Wien's Displacement Law, which is represented by equation 3, explains how the temperature of an item and its maximum emission wavelength are related.

$$\lambda_{max} T = w \quad (3)$$

Where λ_{max} is the wavelength of maximum emission, T is the temperature of the object, and w is the Wien's displacement constant.

In order to record the temperature distribution of an object or a scene, infrared thermography requires a specific camera that is sensitive to infrared light. Different colours or tints in the digital image created by the camera from the infrared radiation represent different temperatures. A typical infrared thermal image of normal eye captured by infrared thermal camera is depicted in figure 3.

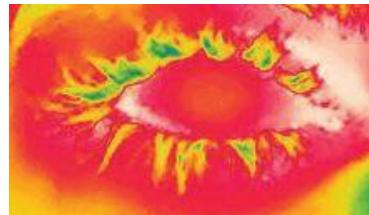


Fig. 3. A typical IR image of normal eye

These equations are used to analyse the thermal radiation emitted by objects and to interpret the information captured by infrared cameras based on the temperature profile.

III. APPROACHES USED

Advancements in the use of algorithms for machine learning and deep learning paved a way to be used in various automated medical diagnosis methods. There are various classification algorithms used in various medical infrared thermal images. Some of the commonly used effective classification algorithms are briefly explained below.

A. Machine Learning Approaches

The branch of artificial intelligence known as machine learning (ML) enables machines to think like people do and to make decisions on their own without the need for human supervision [8]. There are many different machine learning approaches. The list below includes in most of the reviewed papers.

1) *ANN*: A neural network made of artificial neurons has three levels: one input layer and output layer, and maybe two to more hidden layers (ANN). Each connection between layers has a weight attached to it. The neural network develops by modifying the weight. By updating the weight iteratively, network performance is enhanced. Error is decreased since ANN picks up on the training data.

2) *SVM*: SVM is the most popular classifier for thermal imaging in medicine. SVM represents the model by mapping points in space. The two classes are then divided by a hyperplane. The distribution of points is done according to the class to which they belong. The most effective hyperplane for minimising generalisation error is one that maximises the margin. A support vector machine's objective is to identify the ideal separation hyperplane that maximises the margin of the training data [25].

3) *Naive Bayes*: A probabilistic classifier that operates under the assumption of input source independence is the Naive Bayes classifier. The categorization process for NB uses a probabilistic technique. It operates with numerical data, and when the different categories are joined, the data creates a probabilistic model with the use of Bayes' theorem as its base. As it assumes the features are conditionally independent, it simplifies the computation of the probabilities. The naive bayes equation is:

$$P(w|d_1, d_2, \dots, d_n) = \frac{P(w) \times P(d_1|w) \times P(d_2|w) \times \dots \times P(d_n|w)}{P(d_1, d_2, \dots, d_n)} \quad (4)$$

In equation 4, $P(w|d_1, d_2, \dots, d_n)$ is the posterior probability of the class label w for the given features d_1, d_2, \dots, d_n ; $P(w)$ is the prior probability; $P(d_1|w), P(d_2|w), \dots, P(d_n|w)$ are the conditional probabilities of the features d_1, d_2, \dots, d_n for the given class label w , which are estimated from the training data; and $P(d_1, d_2, \dots, d_n)$ is the marginal probability of the features. Naive Bayes is a simple yet effective classification algorithm.

4) *AdaBoost*: AdaBoost is a popular and powerful algorithm that has been successfully applied to many classification problems. A simple or weak classifier must first be trained before making predictions on the training set. The revised weight is used to train the classifier, and the weight of the erroneously classified set is raised. It is made to predict once more on the training set in order to produce a competent classifier. Its simplicity and resistance to overfitting are its key advantages. However, if the weak classifiers are overly complicated, it may perform poorly and be vulnerable to noisy data and outliers.

5) *K-NN*: The k-NN method for classifying objects is one of the simpler machine learning algorithms. The training sample that is closest to the classification model is denoted by "k." Given that neighbours are given weight, the average of the model is increased more by the closest neighbour than by the farthest neighbours [10]. Infrared image analysis and interpretation can be done using the straightforward and understandable kNN method. The k nearest neighbours and the distance metric used to compare the similarity of the feature vectors have an impact on how well the kNN performs, therefore thorough customization and optimization is required.

6) *Decision Tree(DT)*: Both classification and regression use a supervised machine learning called the decision tree. It divides the data into categories and displays the results as a tree structure that resembles a flowchart. It categorises using a set of characteristics known as the root node at the top, which is then divided into leaf nodes that contain the class name [11]. Each leaf node of the tree corresponds to a class label or decision on a particular feature or attribute of the IR picture, whereas each internal node represents a test or judgement regarding that feature or attribute. Recursively dividing the feature space into subsets according to the values of the features, the algorithm builds the decision tree by choosing the feature that yields the greatest information. Overall, decision trees are an adaptable and comprehensible method that may be used to analyse and decipher IR images

7) *Random forest*: Random forest is the most well-liked and successful machine learning method. There are numerous decision trees in it (DT). Increasing the number of decision trees allows for the creation of a forest. The random forest method increases the randomness of tree development. Each decision tree is trained using a different set of training examples. The candidate with the most votes is chosen by the random forest.

B. Deep Learning Approaches

Deep learning networks have shown promising results in various medical imaging applications, including infrared (IR) imaging. Deep learning (DL) is already among the most preclusively used methods that lie within the ML umbrella. In

fact, ML, which itself refers to a larger Artificial Intelligence (AI) family, includes the technique known as DL. The DL method was created in order to outperform traditional ANN while using deep architectures. A deep ANN has a lot of hidden layers that determine how deep the network is, as opposed to a single hidden layer. Convolutional neural networks (CNNs) are among the various deep ANNs that have gained popularity in computer vision applications [17]. Convolutional neural networks, in particular, have helped deep learning perform very well in tasks involving the categorization and processing of images (CNN). Because of their accuracy, pre-trained CNN algorithms like ResNet18, ResNet50, Inception V3, AlexNet, VGG16, and VGG19 are frequently used in the literature to analyse medical images. Pre-trained models aid in evaluating and simplify its use. Convolution, pooling, and full connections are the three types of layers that make up the CNN structure. These layers are stacked in several layers and are arranged hierarchically. An input layer and an output layer are present in addition to these layers. Each layer serves the purpose of learning particular features of the image. The representation-learning algorithms known as deep neural networks made up of a hierarchy of processing layers and a limited number of nonlinear units (i.e., artificial neurons). Medical image processing also makes use of DNN.

C. Statistical Approaches

The use of statistical approaches is widespread. Ocular Surface Temperature(OSTs) or other related statistical parameters [15] such as sample size, mean, variance and standard deviation were compared using statistical tests of significance such as the p-test, f-test, chi-square test, or t-test. Patient with dry eye have low mean compared to normal eye. The probability threshold for statistical significance is mostly at low probability level. Since lower p-value suggests that something is more significant [24]. Validity of null hypothesis is tested. Using the Kolmogorov-Smirnov test, normality is first tested. The Independent Sample p-test is utilised for comparing normally distributed data [1]. To compare OST between normal eye and dry eye, the analysis of variance test (ANOVA) was used [27].

IV. METHODOLOGY

The methodology used for diagnosis of dry eye using machine learning algorithms and deep learning algorithms are depicted in figure 4 and figure 5 respectively. First step is IR image acquisition, which is followed by image pre-processing. In figure 4, in machine learning approach the crucial phase in the automatic diagnosis of dry eye is feature extraction. Statistical texture features are very much useful in the classification of images in the grey level spatial distribution. After feature extraction appropriate features must be selected to obtain high accuracy. Then machine learning algorithms are used to classify normal eye and dry eye.

In Deep learning approaches of diagnosis of dry eye as depicted in figure 5, after IR image acquisition image pre-processing is done. Deep learning algorithm is employed to classify normal eye and dry eye.

V. DISCUSSION

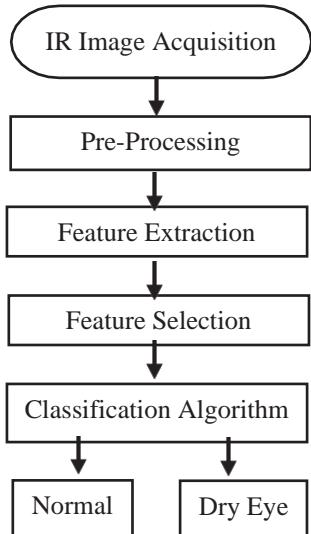


Fig.4. Methodology using Machine Learning approaches

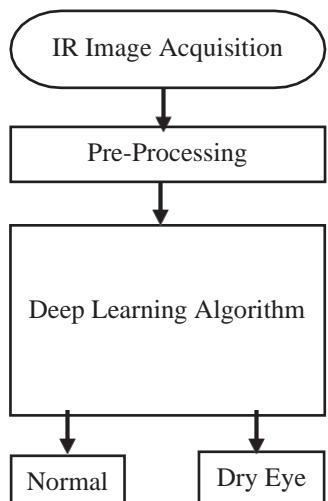


Fig. 5. Methodology using Deep Learning approaches

Statistical texture analysis employs a variety of techniques, including first-order statistics, second-order statistics, and higher-order statistics. First-order statistics, such as expectation, dispersion, asymmetry, entropy, excess, homogeneity, skewness, and kurtosis, are frequently referred to as such since they are derived from the probability of a specific pixel value. First-order statistical features are extracted using histogram texture features. Two-pixel values are used to calculate second-order statistics. The likelihood of witnessing a set of values for pixels next to the current pixel in the image is used to construct higher-order statistics. The image is used to build the gray-level co-occurrence matrix (GLCM), which is used to calculate higher-order statistics. Contrast, Correlation, Homogeneity, Energy are higher-order statistical parameters. Using these parameters different authors have used different methods to classify normal eye and dry eye. Acharya et al [3], extracted the features from Higher order spectra, a nonlinear technique. Then used different classifiers such as KNN, Naïve Bayes, SVM, Probabilistic neural network and DT. And they found KNN and SVM shows high accuracy of 99.8%. Tai Yuan Su et al, [24] used statistical method to differentiate normal image from dry eye image and obtained Sensitivity of 84%, specificity of 83%, receiver operating characteristic (ROC) area is 87%. V.K. Sudarshan et al, [23] added noise level of 0 and noise level of 10 and extracted bispectrum and cumulant features and used DT, k-NN, Probabilistic neural network (PNN), Linear discriminant analysis, SVM. They concluded SVM has high accuracy. Mostly used performance metrics to evaluate the performance are accuracy, sensitivity and specificity.

Table 1 summarizes the different methods used by different authors, number of normal and dry eye samples used in their method and different performance metrics used to evaluate their technique. It is found that number of samples used for analysis was less. The dataset used by different authors are private. No public datasets are available. The performance of different methods is not much efficient. Hence larger datasets are required to improve the efficiency and reliability.

TABLE I. AN OVERVIEW OF RELATED STUDIES

Authors	Year	Total Number of Images	Technique	Results/Findings	Dataset
Azharuddin, Mohammad et al, [5]	2014	42 diseased and 36 healthy	Statistical method	Temperature difference between normal eye and dry eye	From the outpatient's department, (RIO) Regional Institute of Ophthalmology, Calcutta Medical College
Acharya, U.R et al, [3]	2015	83 diseased and 21 healthy	KNN, Naïve Bayesian, DT, PNN, and SVM	Left eye: Accuracy, Sensitivity, and Specificity of 99.8%, 99.8% 99.8% respectively using KNN and PNN classifiers. Right eye: Accuracy, Sensitivity, and Specificity of 99.8%, 99.9%, 99.4% respectively using SVM classifier	From (SERI) Singapore eye research institute.

Su et al, [21]	2017	89 diseased and 65 healthy	Statistical method	Temperature difference between normal eye and dry eye	From Far Eastern Memorial Hospital, Banciao District, New Taipei City, Taiwan
K. Sudarshan et al, [22]	2017	83 diseased and 21 healthy	t-test SVM	Accuracy for different frames	From (SERI) Singapore Eye Research Institute.
V.K. Sudarshan et al, [23]	2017	42 diseased and 42 healthy	DT, k-NN, PNN, Linear discriminant analysis, SVM	Bispectrum Features: Accuracies, Specificities and Sensitivities of 86.90%, 88.10% and 85.71% , using 24 features with the noise level of 0, and 80.95%, 76.19% and 85.71%, using 15 features with the noise level of 10. Cumulant features: Accuracies, Specificities and Sensitivities of 90.48%, 90.48% and 90.48%, using 11 features with the noise level of 0, and 73.81%, 71.43% and 76.19%, using 14 features with the noise level of 10.	From (SERI) Singapore Eye Research Institute.
Acharya et al, [2]	2014	81 responded patients 40 subjects not responded and 41 healthy	DT, k-NN, NB, SVM	Sensitivity of 99.7%, Specificity of 100%, Accuracy of 99.8%	From (SERI) Singapore eye research institute.
Tai Yuan Su et al, [24]	2011	76 diseased, 47 normal	Statistical method (Compactness value (CV), Temperature difference value (TDV))	Sensitivity of 84%, specificity of 83%, receiver operating characteristic (ROC) area is 87%	From Far Eastern Memorial Hospital, Banciao District, New Taipei City, Taiwan.
Zhang et al, [27]	2021	138 diseased, 46 normal	Statistical method	Temperature difference between normal eye and dry eye	From the He Eye Specialist Hospital, Shenyang, China.
Abusharha, [1]	2021	25 diseased, 25 normal	Statistical method	Temperature difference between normal eye and dry eye	From optometry clinics at the College of Applied Medical Sciences, King Saud University.

VI. CONCLUSION

Infrared radiation is a non-ionising radiation and therefore it will not cause cancer. Hence infrared thermography is a non-destructive method. The growth in patients with eye problems necessitates automated pathology diagnosis and screening. This paper explains the effectiveness to detect dry eye using the non-invasive thermographic technique. Only fewer samples were used for the investigation in the previous papers, hence more samples would be required to be added in order to assess the results. For the classification of medical images, deep learning algorithms are also being developed, and they exhibit incredibly high accuracy. As a result, in the future, the accuracy of detecting dry eye using infrared thermal images can be evaluated and improved using machine learning and deep learning algorithms.

REFERENCES

- [1] ABUSHARHA, Ph D. "Analysis of Ocular Surface Temperature in Patients with Dry Eye." *The Medical Journal of Cairo University* 89. December (2021): 2549-2553.

- [2] U.R. Acharya, J.H. Tan, S. Vidya, S. Yeo, L.T. Cheah, W.J.E. Lim, C.K. Chua, L. Tong, Diagnosis of response and non-response to dry eye treatment using infrared thermography images, *Infrared Phys. Technol.* 67 (2014) 497–503.
- [3] Acharya, U.R., Tan, J.H., Koh, J.E., Sudarshan, V.K., Yeo, S., Too, C.L., Chua, C.K., Ng, E.Y.K. and Tong, L., 2015. Automated diagnosis of dry eye using infrared thermography images. *Infrared Physics & Technology*, 71, pp.263–271.
- [4] S.B. Ankita, G.B. Samir, M.J. Sunita, A review on recent advances in dry eye: pathogenesis and management, *Oman J. Ophthalmol.* 4 (2011) 50–56.
- [5] Azharuddin, M., Bera, S.K., Datta, H. and Dasgupta, A.K., 2014. Thermal fluctuation-based study of aqueous deficient dry eyes by non-invasive thermal imaging. *Experimental eye research*, 120, pp.97-102.
- [6] A. Behrens, J.J. Doyle, L. Stern, et al., Dysfunctional tear syndrome: a Delphi approach to treatment recommendations, *Cornea* 25 (2006) 900–907.
- [7] E. Hosaka, T. Kawamorita, Y. Ogasawara, N. Nakayama, H. Uozato, K. Shimizu, M. Dogru, K. Tsubota, E. Goto, Interferometry in the evaluation of precorneal tear film thickness in dry eye, *Am. J. Ophthalmol.* 151 (2010) 18–23.

- [8] Ibrahim, Ibrahim, and Adnan Abdulazeez. "The role of machine learning algorithms for diagnosing diseases." *Journal of Applied Science and Technology Trends* 2.01 (2021): 10-19.
- [9] H.T. Jen, E.Y.K. Ng, U.R. Acharya, C. Chee, Infrared thermography on ocular surface temperature: a review, *Infrared Phys. Technol.* 52 (2009) 97–108.
- [10] D.T. Larose, Discovering Knowledge in Data: An Introduction to Data Mining, vol. 90–106, Willey Interscience, New Jersey, USA, 2004 (Chapter 5: KNN).
- [11] D.T. Larose, Discovering Knowledge in Data: An Introduction to Data Mining, vol. 108–126, Willey Interscience, New Jersey, USA, 2004 (Chapter 6: Decision Trees).
- [12] Lemp, Michael A., and GARY N. Foulks. "The definition and classification of dry eye disease." *Ocul Surf* 5.2 (2007): 75-92.
- [13] Y. Masakazu, M. Yoshinobu, S. Chika, Impact of dry eye on work productivity, *ClinicoEconomics Out. Res.* 4 (2012) 307–312.
- [14] Methodologies to diagnose and monitor dry eye disease, Report of the diagnostic methodology subcommittee of the international dry eye workshop, *Ocul. Surf* 5 (2007) 108–152.
- [15] Morgan, Philip B., Andrew B. Tullo, and Nathan Efron. "Infrared thermography of the tear film in dry eye." *Eye* 9.5 (1995): 615-618.
- [16] K.K. Nichols, J.J. Nichols, G.L. Mitchell, The lack of association between signs and symptoms in patients with dry eye disease, *Cornea* 23 (2004) 762–770
- [17] Oliveira, Dario Augusto Borges, et al. "A review of deep learning algorithms for computer vision systems in livestock." *Livestock Science* 253 (2021): 104700.
- [18] Ring, E. F. J., and Kurt Ammer. "The technique of infrared imaging in medicine." *Infrared Imaging: A casebook in clinical medicine*. IoP Publishing, 2015.
- [19] K. Santosh, T. Alan, M. Angus, D. Charles, R. Kannu, Dry eye diagnosis, *Invest. Ophthalmol. Vis. Sci.* 49 (2008) 1407–1414.
- [20] K. Snjezana, T. Martina, S.R. Jasmina, N. Branko, Diagnostic procedures and management of dry eye, *Biomed. Res. Int.* (2013).
- [21] Su, Tai-Yuan, et al. "Infrared thermography in the evaluation of meibomian gland dysfunction." *Journal of the Formosan Medical Association* 116.7 (2017): 554-559.
- [22] K. Sudarshan, J.E.W. Koh, U.R. Acharya, J.H. Tan, M.R.K. Mookiah, C.K. Chua, L. Tong, Evaluation of evaporative dry eye disease using thermal images of ocular surface regions with DWT and gabor transform, in: E.Y.K. Ng, M. Etehadtavakol (Eds.), *Application of Infrared to Biomedical Sciences*, Springer Singapore, Singapore, 2017, pp. 359–375.
- [23] V.K. Sudarshan, J.E.W. Koh, J.H. Tan, Y. Hagiwara, C.K. Chua, E.Y.K. Ng, L. Tong, Performance evaluation of dry eye detection system using higher-order spectra features for different noise levels in IR thermal images, *J. Mech. Med. Biol.* 17 (07) (2017) 1740010
- [24] Tai Yuan Su, Huihua K. Chiang, Chen Kerh Hwa, Po Hsuan Liu, Ming Hong Wu, David O. Chang, Po Fang Su, Shu Wen Chang, "Noncontact detection of dry eye using a custom designed infrared thermal image system," *J. Biomed. Opt.* 16(4) 046009 (1 April 2011) <https://doi.org/10.1117/1.3562964>
- [25] Abdelsalam, Mohamed M., and M. A. Zahran. "A novel approach of diabetic retinopathy early detection based on multifractal geometry analysis for OCTA macular images using support vector machine." *IEEE Access* 9 (2021): 22844-22858.
- [26] N. Yokoi, A. Komuro, Non-invasive methods of assessing the tear film, *Exp. Eye Res.* 78 (2004) 399–407.
- [27] Zhang, Qing, et al. "Screening evaporative dry eyes severity using an infrared image." *Journal of Ophthalmology* 2021 (2021).

Design of an automated healthcare system using hand gestures for paraplegic people

Bethanney Janney J*

Dept. of Biomedical Engg

Sathyabama Institute of Science and

Technology

Chennai, India

bethanneyj@gmail.com

Aakesh U

Dept. of Biomedical Engg

Sathyabama Institute of Science and

Technology

Chennai, India

appu944304@gmail.com

Yaswant Rajasekaran

Dept. of Biomedical Engg

Sathyabama Institute of Science and

Technology

Chennai, India

yaswantrajasekaran@gmail.com

Renny Valentina

Dept. of Biomedical Engg

Sathyabama Institute of Science and

Technology

Chennai, India

rennytina0@gmail.com

S.Krishnakumar

Dept. of Biomedical Engg

Sathyabama Institute of Science and

Technology

Chennai, India

drkrishnakumar_phd@yahoo.com

T. Sudhakar

Dept. of Biomedical Engg

Sathyabama Institute of Science and

Technology

Chennai, India

drsudhakar35@gmail.com

Abstract—Paralyzed individuals often struggle to express their needs due to a lack of motor control in the brain, leading to difficulties in speech or sign language communication. The proposed system offers a solution for disabled individuals to communicate by displaying messages on a monitor through simple hand movements. The device uses an accelerometer to detect various hand tilt directions, allowing the user to convey different messages by tilting the device in different directions. The accelerometer measures the motion statistics and sends the information to the microcontroller for processing. Once processed, the appropriate message is displayed and a buzzer sounds. If the message is a call for assistance, the IFTTT app is activated and a message is sent to the doctor's phone. This system enables paralyzed patients to effectively communicate their basic needs.

Keywords—Paralysis, Electromechanical, hand glove, IFTTT (If This Then That) app, microcontroller, healthcare.

I. INTRODUCTION

According to the Global Burden of Diseases, stroke is a leading cause of death, taking the lives of around 5.8 million people annually. It is also the most common cause of paralysis, impacting about 33.7% of the population. Paralysis can also occur from severe spinal cord injury due to an accident. A recent survey by the World Health Organization estimated that 5.6 million people were predicted to be paralyzed, making up 1.9% of the population among the 50 studied. However, there is currently no reliable system in place to monitor their health and daily needs. Automation [1], sensor systems, machine vision [2], brain-computer interaction [3], and human motor interface [4] are some of the different approaches and systems that have been created to assist people with disabilities using various technologies. In today's fast-paced world, it is challenging to constantly care for those in need of assistance. To address these challenges, a device has been proposed that uses the ESP32 chip microcontroller and accelerometer to detect the patient's hand movements, allowing them to communicate their basic needs to their caregivers [5]. This system is aimed at individuals who have suffered paralysis as a result of a stroke or injury, and who struggle to communicate due to a lack of motor control in the brain. The device allows the disabled person to display a message on an LCD screen through body movements with motion capabilities. In case no assistance is available, the system sends a message via the ESP32 microcontroller and

SMS. To convey a message, the user only needs to tilt their hand at a specific angle, and different messages are conveyed by tilting the device in different directions. The accelerometer measures motion statistics, and the ESP32 processes the data and displays the appropriate message on the LCD screen. When it receives a motion signal, it also triggers a buzzer and displays a message. The proposed system automates the care process for paralysis patients, ensuring timely attention and optimal health [6].

II. RELATED WORKS

Kinjal Raykarmakar et al. (2022) created a device using microcontroller-based circuitry. It makes use of a receiver and transmitter circuit as well as a hand motion recognition circuit. The accelerometer and gyroscope are utilized in the hand motion circuit to detect hand movements, which are subsequently wirelessly transmitted to the receiver system through radiofrequency (RF). The receiver system is designed to accept and interpret these commands, show them on the LCD, and transfer the data online to a Gecko Server for the Internet of Things (IoT). The IoT Gecko Server subsequently presents this information online to obtain the desired result [7].

Sujin J. S (2021) and team created an IoT-based system to monitor and notify the medical needs of paralyzed outpatients. This helps the patients communicate their health concerns, diagnosis, and other important information to medical staff or guardians via the internet or by sending a message. If the patient's vital signs, such as blood pressure or oxygen levels, deviate from normal, the system will analyze the patient's health data and initiate automated nursing care, ensuring consistent attention [8].

Kate et al. (2022) described a tracking system that is employed to monitor the well-being of patients. The surveillance senses the patients' physiological data using wearable sensors like airway detectors, sphygmomanometers, and heart rate monitors. These parameters are directly supervised and conveyed to the care recipient utilizing GSM. A microcontroller can be useful in solving this problem (MSP430) [9].

In order to enhance interaction between the care recipient and the immobilized patient, Hira Beenish et al. (2021) proposed a system that would let the sufferer convey their

requirements by employing hand, finger, and foot movement. In order to track motion, the gyro MPU6050 is coupled to the Arduino UNO, which is affixed to the gloves. Flex sensors are used to track gestures, and gyroscope sensors are attached to the sufferer's hand to track any changes to that area of the body. Three main tasks are connected to the system. These sensors are used to transmit and comprehend the sufferer's needs via hand, finger, and foot movements [10].

Aziz et al. (2019) introduced an intelligent primary care system that aids the paralytic sufferer if they are notified by the signal from the person over a mobile network. To aid the health officer in meeting the demands of the paralyzed patient, many motion-signaling sensor instructions are offered in this work. Every time a patient instructs a basic hand movement, an SMS is sent to the healthcare provider, and a warning alerting them to the instruction is shown on a display panel. This enables the physician to contact the appropriate health officials to provide assistance [11].

III. MATERIALS AND METHODS

The proposed design is depicted in the block diagram as shown in Fig. 1. It features an ESP32 microcontroller that receives power from a 3.3V power source. The microcontroller is then connected to the microelectromechanical System, which delivers data about the patient's hand movement to the ESP32, which generates a customized message based on the patient's hand movement's specific direction and set angle. Every message generated by the system is shown on an LCD (16X2) screen attached to the microcontroller. When the Call Attendant message is generated, the IFTTT app (IFTTT app is connected to ESP32 via Wi-Fi) sends a message to the doctor's phone, instructing him to go aid the patient immediately. When the device sends a message, the LED linked to it will light up, indicating that a message has been sent and the buzzer also beeps providing an auditory alarm. As a result, this technology assists paralyzed patients in meeting their most fundamental demands at the appropriate time, allowing them to overcome their challenges. Only patients who are partially paralyzed are eligible to use this device. The device is designed for individuals who have experienced partial paralysis as a result of a tremor, and is not intended for use during an active tremor. This device communicates four fundamental requirements that address the primary assistance needs of partially paralyzed patients.

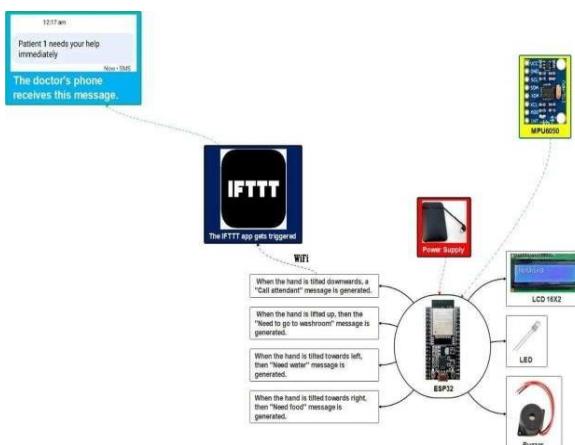


Fig. 1. Proposed block diagram.

A. ESP32 Microcontroller module

Espressif Systems, the company behind the famous ESP8266 System on Chip (SoC), has released the ESP32 SoC Microcontroller. This device includes a Tensilica 32-bit Xtensa LX6 Microprocessor with Wi-Fi and Bluetooth capabilities, making it a successor to the ESP8266 SoC. The ESP32 is available in single-core and dual-core configurations. Cadence's modular and extendable CPU cores provide a one-of-a-kind solution for System-on-Chip (SoC) designers. These cores, such as the Xtensa LX6 data plane [12].

B. MPU6050 MEMS

An accelerometer and three-axis gyroscope are features of the MPU6050 Micro Electromechanical System. It can measure a range of motion-related attributes, including acceleration, displacement, and others. The capabilities of the MPU6050's Digital Motion Processor are achieved by intricate computations. The MPU6050 has a 16-bit analog-to-digital converter and can concurrently record motion in three dimensions. The MPU6050 may be readily integrated with an Arduino or other well-known microcontrollers because of its functions that are well-known and useful [13].

C. LCD 16X2

Electronic screens known as liquid crystal displays (LCDs) have a variety of uses. A 16x2 LCD is a fundamental part that may be used in a variety of gadgets and circuits. Each character is represented by a 5x7 pixel matrix, and it has the ability to show 16 characters per line on each of its two lines. The display's 16 x 2 intelligent alphanumeric dot matrix can display 224 distinct letters and symbols. Command and Data registers make up the LCD [14].

D. Buzzer

A Piezoelectric Buzzer Alarm is a flexible tool with a variety of applications. It has two 35mm mounting holes and runs on a voltage range of 3 to 12 volts. It is simple to place on a flat surface. The alarm's dimensions are 30 mm in diameter and 10mm in height, and it is built of piezoelectric materials. It is black in color. Its rated voltage is 12 volts DC, and it produces 95 dB of sound pressure [15].

E. LED

When electricity passes through a light-emitting diode, light is produced (LED). When electrons in a semiconductor material mix with electron holes to form photons, energy is released. The hue of the light, which is comparable to the photon energy, is determined by the energy needed for electrons to pass across the semiconductor's bandgap. The semiconductor device may have many semiconductors or a phosphor layer to create white light. The first LED, which emits low-intensity infrared light and is utilized in remote control circuits for numerous consumer gadgets, was initially launched in 1962 [16].

F. Arduino IDE

A free, open-source tool for authoring, building, and uploading code to Arduino or ESP32 boards is the Arduino Integrated Development Environment (IDE). It is compatible with the programming languages C and C++ and is accessible on a variety of operating systems, including Windows, Mac, and Linux. Sketching is the term used to describe the act of creating code in the Arduino IDE. The Arduino board must be linked to the IDE in order to compile a sketch into a format

that can be uploaded over USB to the board using the IDE [17].

G. IFTTT

The programming expression "if this, then that" served as the inspiration for the moniker IFTTT. It provides a software platform that makes it possible to join programs, services, and gadgets from different sources in order to start automated activities. Applets, which function like macros linking several programs to carry out certain tasks, make this feasible. Through the IFTTT website, mobile app, or app widgets, users may manage applets. IFTTT's user-friendly and straightforward interface enables users to make their own applets or edit already-existing ones [18].

IV. RESULTS AND DISCUSSION

The gadget detects the motion and generates the exact message matching the particular direction of motion, allowing patients to communicate their basic demands simply by moving their hands. Visual and auditory alarms notify the attendants and doctors whenever a message is being conveyed through the device by the patient. If the patient is in an emergency scenario, he or she can use this device to send a message to the doctor's mobile phone by moving his or her hand in a specific direction. The prototype of an automated paralysis patient care system is shown in Fig. 2.



Fig. 2. Prototype.

When the prototype is powered ON, the LCD screen starts glowing. The LCD starts showing a message "**Nothing**" when the glove which will be worn by the patient shows no movement. In Fig. 3 the "**Nothing**" message is displayed on the screen when there is no movement in the hand. When the glove worn by the patient is tilted towards the right, the device pops up a message "**Need food**" as depicted in Fig. 4.



Fig. 3. When there is no movement.



Fig. 4. When the patient tilts his hand toward the right.



Fig. 5. When the patient tilts his hands toward the left.

When the patient's glove is turned to the left, a notification that reads "**Need water**" appears on the device. Fig. 5 shows the device with the message "**Need water**" and the patient's hand glove inclined to the left. When the patient tilts his hand downwards, the LCD screen displays the word "**Call Attendant**" as shown in Fig. 6, indicating that there is an emergency, the caregiver or the doctor should attend to the patient immediately.



Fig. 6. When the patient tilts his hand downwards.



Fig. 7. When the patient tilts his hand upwards.

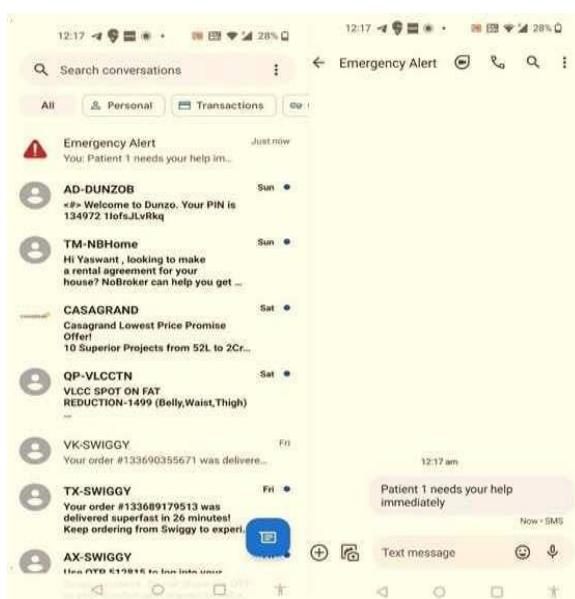


Fig. 8. Mobile message generated by the IFTTT app.

When the Call Attendant message is generated, the IFTTT app (The IFTTT app is connected to ESP32 via Wi-Fi) notifies the doctor's or attendant's mobile phone, as seen in Fig. 8, so they may assist the patient right away. When the device transmits a message, an LED connected to it will flash up to show that a notification is being sent, and a buzzer will also ring to provide an audible alarm.

The LCD panel displays the message "Need to go to the washroom" when the patient tilts his hand upwards, alerting the caregiver to prepare hygienic supplies. Fig. 7 demonstrates the device with the message "Need to go to the washroom" when the patient tilts his hand upwards.

The IFTTT (If This Then That) programme is incorporated into the suggested concept, setting it apart from others by enabling emergency SMS messages to be sent via Wi-Fi without a GSM module.

V. CONCLUSION

There are various medications and physiotherapy available for paralyzed individuals, but there is no specific system in place to meet their needs. To address these issues, a system has been implemented to communicate their basic needs just by the movement of their hand. This method aids paralyzed persons in meeting their most basic needs by communicating their needs to their caretakers. This device is easy to handle for the patient or other persons. In the future, an Artificial Intelligence-based patient idleness monitoring module along with a self-speech training module can be introduced in the patient's application to keep the patient aware of himself. Also, the patient can be directly linked with their doctor to track the activities of the patient to provide even more personal care depending on the wants and the status of the patient's health.

ACKNOWLEDGMENT

We would like to express our gratitude to the management of the Sathyabama Institute of Science and Technology, the School of Bio and Chemical Engineering, and the Department of Biomedical Engineering for providing us with the direction and assistance necessary to carry out the research successfully.

REFERENCES

- [1] Nikpour, M., Huang L., Al-Jumaily A.M, "Stability and Direction Control of a Two-Wheeled Robotic Wheelchair through a Movable Mechanism," IEEE Access, vol. 8, pp. 45221–45230, 2020.
- [2] Kim E.Y, "Vision-Based Wheelchair Navigation Using Geometric AdaBoost Learning," Electron. Lett., vol. 53, pp. 534–536, 2017.
- [3] Pinheiro O.R, Alves L.R.G, Souza J.R.D, "EEG Signals Classification: Motor Imagery for Driving an Intelligent Wheelchair," IEEE Lat. Am. Trans., vol. 16, pp. 254–259, 2018.
- [4] Hartman A, Nandikolla V.K, "Human-Machine Interface for a Smart Wheelchair," Journal of Robotics, 2019, vol. 2019, pp. 1-11, 2019.
- [5] Akash SD, Anusha Upadhy, Harshitha BR, Hemanth, Mr. Shivarudraiah B, "Smart wearable system for the elderly using IOT," International Research Journal of Modernization in Engineering Technology and Science, vol. 4(7), pp. 2799-2806, 2022.
- [6] Mohana M, S. Priyadarshini, N. Sowmiya, G. Pavithra Devi, "An IOT based automated communication system for paralyzed patients using simple hand gestures," European Journal of Molecular & Clinical Medicine-2020, vol. 7(4), pp. 2681-2686, 2020.
- [7] Kinjal Raykarmakar, Shruti Harrison, Anirban Das, "Paralyzed Patient Healthcare," Medical Internet of Things Techniques, Practices and Applications, pp. 161-173, 2021, DOI: 10.1201/9780429318078-11.
- [8] J. S. Sujin, S. Mukesh, M. J. Praveen Raj, M. Sashwanth and R. Ramesh Kumar, "IoT based Patient Monitoring System using TCP/IP Protocol," 2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS), 2021, pp. 1076-1080, doi: 10.1109/ICACCS51430.2021.9441910.
- [9] Kate, M.D., Wadhai, M.A., Vaidya, M.K., Kadak, M.S., Shrikhande, M.R., & Dhurve, M.P, "Paralysis Patients Monitoring System using GSM," International Journal of Advanced Research in Science, Communication and Technology, vol. 2(1), pp. 43-53, 2022.
- [10] Hira Beenish, Ahsan Sheikh, Fakheha Nasir, "Design and Implementation of a Monitoring System for Paralysis patient using IoT," KIET Journal of Computing & Information Sciences, vol. 4(2), pp. 49-58, 2021.
- [11] Aziz S. A. C, A.F., Kadmin, N. Rahim, W. H. W. Hassan, "Development of automatic healthcare instruction system via movement gesture sensor for paralysis patient," International Journal of Electrical and Computer Engineering, vol. 9(3), pp. 1676~1682, 2019, DOI: 10.11591/ijece.v9i3.
- [12] Deepasri T, Gokulpriya M, Arun Kumar G, Mohanraj P, M Shenbagapriya, "Automated Paralysis Patient Health Care Monitoring

- System," South Asian Journal of Engineering and Technology, vol.3(2), pp. 85–92, 2017.
- [13] Renee N, Sigrid Reddiyar, S Remina, S Sabrin, M Subhashini, "IOT Based paralysis patient health care monitoring system," Journal of the University of Shanghai for Science and Technology, vol. 23(6), pp. 431-437, 2021.
- [14] Vidya Sarode, Victor Thomas, Akash Dubey, Shashank Shukla, "Automated Paralysis Patient Health Care System," International Research Journal of Engineering and Technology, vol. 8(5), pp. 680-686, 2021.
- [15] Diptee Gaikar, Pradnya Porlekar, Divya Shetty, Akash Shitkar, Prof.Kalindi Kalebere, "Automated Paralysis Patient Healthcare System," International Journal of Creative Research Thoughts, vol. 9(8), pp. b251-b256, 2021.
- [16] Viancy V, Avinash Wilson, Anusha. P, "Paralyzed Patient Monitoring Equipment – IoT," International Journal of Advanced Research in Basic Engineering Sciences and Technology, vol. 6(10), pp. 25-28, 2020.
- [17] Rohit Malgaonkar, Saurabh Kamble, Manthan Jadhav, Satyam Parkale, "Survey on Automated Paralysis Patient Healthcare Monitoring System," International Journal for Scientific Research & Development, vol. 7(10), pp. 517 to 519, 2019.
- [18] Sayali A Bhurke, Prajakta A Jadhav, Amruta P Erudkar, Nayana S Palekar, Snehal B Jadhav and M. A. Pardesi, "An IOT based approach in paralysis patient healthcare system," International Journal of Applied Research, vol. 7(2), pp. 459-462, 2021.

Development of a Fiber-Optic Spectral Domain based Optical Coherence Tomography in Near Infrared Range for Biomedical Applications

1st Richa Parihar

Department of Engineering Design
Indian Institute of Technology
Madras
Chennai 600036
richalparihar@gmail.com

2nd Nilesh J. Vasa

Department of Engineering
Design
Indian Institute of Technology
Madras
Chennai 600036
njvasa@iitm.ac.in

3rd Joy Mammen

Department of Transfusion Medicine
and Immuno Haematology
Christian Medical College
Vellore
joymammen@cmcvellore.ac.in

4th Pauline John

Department of Bioengineering New
York University Abu Dhabi
Abu Dhabi ,UAE
paulinejhn@gmail.com

Abstract— In this paper, a fiber-optic low coherence interferometry technique, in the spectral domain, is used for depth detection by considering different spectrometer resolutions from 0.1 to 2 nm. Depth resolved result from the sample was obtained for 0.1 nm resolution of the spectrometer compared to 2 nm resolution. Two glass coverslips of thickness 150 μm were used as samples. A physical depth equal to the actual thickness of the sample was estimated using the fiber-optic spectral domain low coherence interferometry. Further, the preparation of phantom which mimics the real tissue has also been illustrated for further extending this non-invasive and non-contact study using optical coherence tomography, towards the detection of skin burns or abnormalities such as skin necrosis and skin melanoma.

Keywords— Low Coherence Interferometry, Optical Coherence Tomography, Phantom, Skin Necrosis

I. INTRODUCTION

Over the past few decades, the occurrence of skin diseases has increased all across the world. This not only has increased the pressure on the healthcare system but also made the researchers to make their efforts towards this direction to develop a non-invasive diagnostic technique. Although mortality is not very significant in skin diseases, according to [1] there is a chance that 4% of skin diseases mature into melanoma which is the most threatening form of the cancerous disease that can cause death. In many cases, skin and subcutaneous diseases can cause Years Lived with Disability (YLDs) in individuals. From 1990 to 2017, an increase of 53.7% in years lived with disability was reported for all the skin and subcutaneous diseases [2]. In India around 332.96 cases of cardiovascular diseases were reported, whereas it is 455.06 per 100,000 for years lived with disability were recorded in 2017. According to the survey conducted for European countries in 2020, 47.20% were affected by skin diseases. Based on this study, 201 million people aged 18 years or more were predicted to be affected with skin cancer [3]. Also, optical coherence tomography-based angiography gives the correlation between vessel depth and tissue injury depth. Therefore, early and accurate detection of all skin diseases especially, melanoma is necessary in the direction to get relevant management of the disease as well as promote survival rates. Human skin serves a large spectrum of

functions and that is why it is defined as the complex human organ. It basically includes three layers, i.e., epidermis, dermis and fat layer or hypodermis. It is of utmost importance to understand the interaction of light and skin tissue for various medical diagnostic and therapeutic techniques. It is essential to understand the light and tissue interaction based on optical properties such as absorption, scattering, attenuation and refractive index. The most prevailing imaging techniques used for examining skin tissue include confocal microscopy, multispectral (MS) imaging, 3D topography, optical coherence tomography (OCT), and polarized imaging (PI) amid other imaging modalities. In addition, the usage of machine learning for the automated detection of skin lesions has also become quite popular. Some of the popular imaging modalities are confocal microscopy in which the optical resolution and contrast of the micro-graph is imaged by means of a pinhole to bar the out-of-light in image formation. In multispectral imaging, various spectral bands of the electromagnetic spectrum are used for examining the reflected or transmitted light from the skin. Whereas, 3D topography technique, obtains the height maps of the surface of the lesions in order to acquire pointwise information from the surface of skin. In polarization imaging, polarization of light is being used. When the reflected light coming from the skin surface passes through the polarization state analyzer, it carries information about anisotropic properties of the tissue structure. The optical anisotropy or the birefringence which exhibits different refractive index depending upon plane of polarization. Skin tissues shows birefringent properties which can be analyzed by polarization imaging. In the direction of inspecting the internal microstructure of the living tissue a high resolution and a non-contact approach based on optical coherence tomography can be a valuable diagnostic tool. OCT technique is mainly based on low coherence interferometry (LCI) and uses Michelson interferometry configuration with a low coherent light source. It can detect amplitude and depth from the backscattered light from the sample and can obtain a cross-sectional image of the sample [4].

Frequency domain LCI techniques around 1300 nm and 1600 nm have also been used for applications such as glucose

detection from oral mucosa mimicking phantoms and eye models using a free-space configuration [5]. OCT technique is capable of providing real-time, in vivo, cross-sectional evaluation of the skin without cutting or excising the skin tissue [6]. It detects ballistic photons of light backscattered from the sample when there is a mismatch of refractive index, by using a low-coherence interferometry, a light source at the near infrared wavelength range of 1305 nm [6]. The corneal thickness of human eye can be calculated in vivo using scattering approximation technique [7]. The reflected signal from the sample at a particular depth is being compared with the reference signal and when both the path lengths are equal an interference occurs within the coherence length. Moreover, frequency domain differential absorption LCI technique had been developed in order for the detection of glucose for aqueous humor of eye model [8]. In this proposed technique, a fiber-optic spectral domain low coherence interferometry approach is used and preliminary studies are conducted to optimize the spectrometer resolution to obtain depth information from thin samples, for further conducting studies on skin necrosis.

II. METHODOLOGY

The schematic of fiber-optic spectral domain low coherence interferometry is shown in Fig.1. A superluminiscent light emitting diode (SLED) (Exalos) EXS210047-02 centered around 1310 nm was used as a light source. The reference mirror was mounted on a translating x-y stage. FC1310-70- 50-APC optical fiber-based coupler was employed for splitting the light in the ratio of 50:50. For reducing the output power from the reference arm, a neutral density filter was used. For acquiring spectral intensity pattern of reference, sample and the interference signals, a Yokogawa AQ6370D optical spectrum analyzer was used. Reference was kept fixed for acquiring interference. The interference signal intensity was measured at 1310 nm for spectrometer resolutions of

0.1 nm, 0.2 nm, 0.5 nm, 1 nm and 2 nm. The depth scan was obtained for the corresponding interference spectra at 0.1 nm, 0.2 nm, 0.5 nm, 1 nm and 2 nm of spectrometer resolution, using inverse Fourier transform using MATLAB 2021 software. The maximum scanning depth (Z_{\max}) can be estimated by the following equation

$$Z_{\max} = \frac{\lambda_0^2}{4n\delta\lambda} \quad (1)$$

Where λ_0 is the center wavelength of the source, $\delta\lambda$ is the wavelength sampling interval and n is the refractive index of the sample. Here the optical depth can be calculated with respect to different peaks which correspond to reflecting surface interfaces from the sample. Subsequently, physical depth was verified in relation with the refractive index of cover slip which is approximately 1.52. The sample shown in Fig.1 will be replaced with skin phantom/real skin. The skin in bound and unbound states constitutes water and protein structures called collagen. The collagen refractive index is approximately 1.55 at 1310 nm, remarkably larger than that of water, 1.32 at 1310 nm [9,10]. For mimicking biological tissue, generally the phantoms which mimic the optical

properties of real tissue are used. Figure 2 represents the schematic of phantom preparation process. Base material consists of a transparent material with low attenuation characteristics for the wavelengths of interest. Scattering agent consist of colloids or suspensions in order mimic the scattering property of the sample of interest Absorbing agent usually inks or dyes are added to this mixture in order to mimic the absorption property of the sample of interest. The India ink is used as the absorbing agent because of its optical properties in the near infrared region [11].

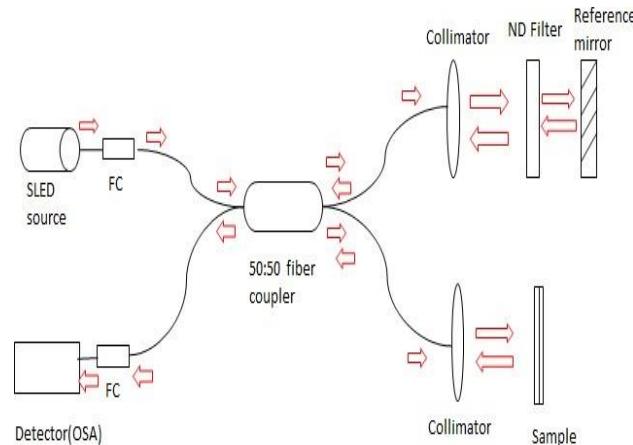


Fig. 1 Schematic of LCI setup with two cover slip as sample

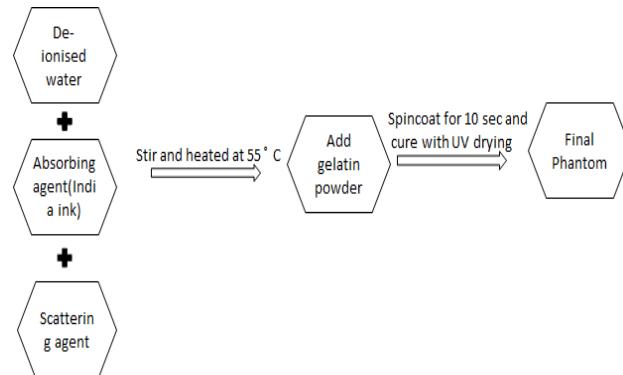


Fig. 2 Schematic representation of phantom preparation process.

III. RESULTS AND DISCUSSION

Figure 3 (a) shows the result obtained using MATLAB code, for 0.5 nm resolution showing different peaks which corresponds to reflecting interfaces in the sample. Center peak reciprocate the DC term of the OCT equation. Both side peaks resemble the mirror image on either side of DC term. P1 corresponds to air-glass interface of first coverslip, P2 attributes to the glass-air interface, P3 corresponds to the air-glass interface of second coverslip and P4 shows the glass-air interface of the sample. By measuring the difference between two peaks, the optical depth can be calculated. Finally taking the refractive index into account, the physical depth can be obtained which is the actual thickness of the sample

$$\text{Physical pathlength} = \frac{\text{Optical pathlength}}{\text{Refractive Index}} \quad (2)$$

Figure 3 (b) shows the result for 2 nm resolution. It is evident that even the four peaks are not clearly visible.

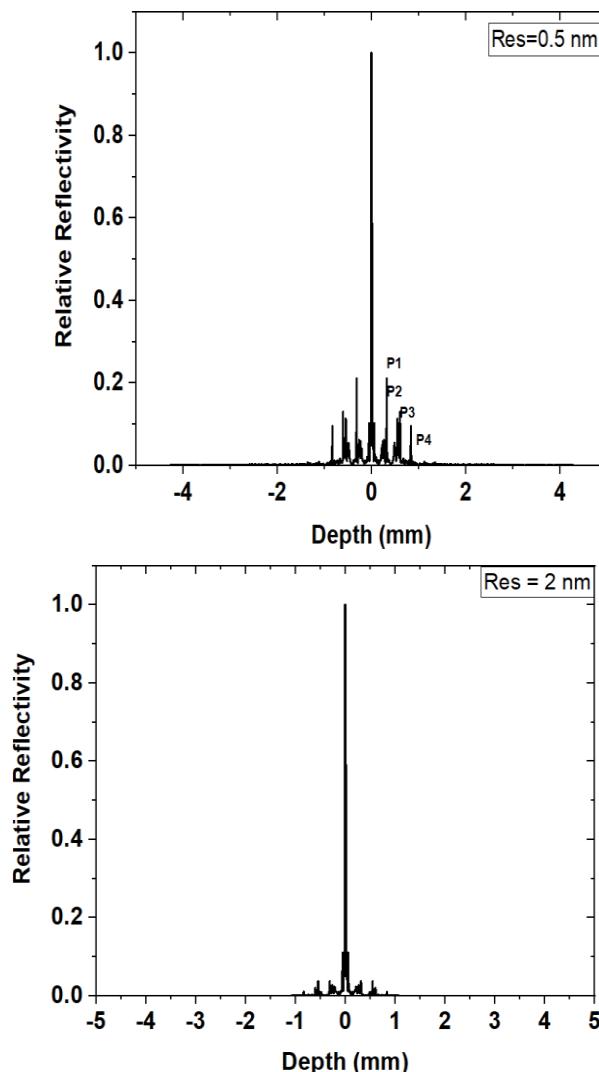


Fig 3 (a) Depth scan obtained by applying inverse Fourier transform showing peaks which correspond to reflecting surfaces of the sample for the resolution of 0.5 nm (b) Depth scan obtained by applying inverse Fourier transform showing peaks which correspond to reflecting surfaces of the sample for 2 nm resolution.

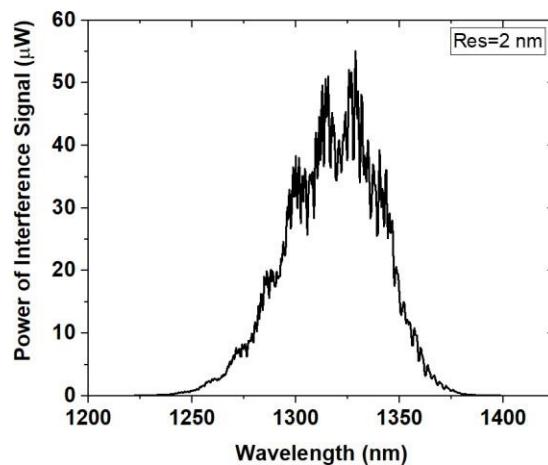
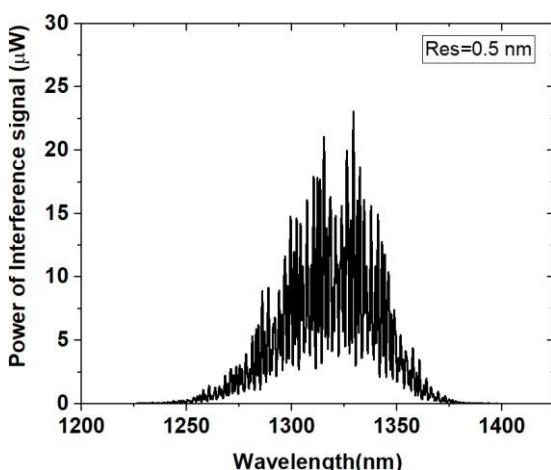


Fig 4 (a) Interference spectrum for 0.5 nm resolution (b) and for 2 nm resolution.

Figure 4 shows the interference spectrum obtained from two coverslips. For 0.5 nm resolution, the interference was prominent as shown in Fig. 4 (a), while for 2 nm resolution although the interference power of the signal was increased as shown in Fig. 4 (b), the interference spectrum was not easily seen.

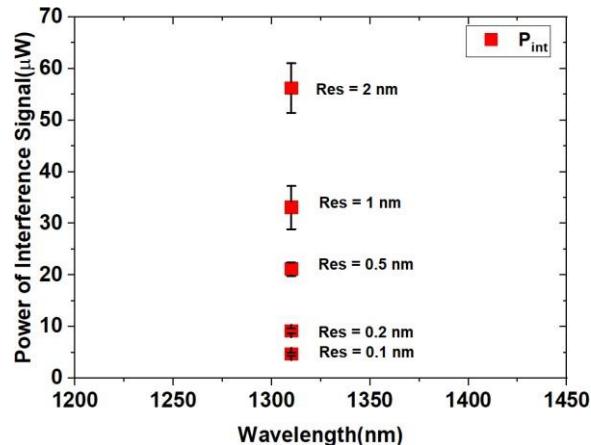


Fig 5 Interference Power (nW) plotted at 1310 nm wavelength for different spectrometer resolutions of 0.1 nm, 0.2 nm, 0.5 nm, 1 nm and 2 nm.

In Fig. 5, it is visible that with increase in resolution the interference power increases, as the interference power is highest for the case of 2 nm resolution.. Spectrometer resolution of 0.5 nm was found to be optimum for obtaining depth resolved information from the sample compared to 2 nm resolution.

TABLE I. RESOLUTION (nm) AND CORRESPONDING POWER OF INTERFERENCE SIGNAL (μW) FOR WAVELENGTH 1310 nm

S.No.	Resolution(nm)	Power of Interference signal(μW)
1	0.1	4.79
2	0.2	9.33
3	0.5	20.4
4	1	34.7
5	2	55

Table I indicates Power of the Interference signal for different resolution. It can be concluded that as the

value of resolution increases the interference power also increases. Also, for maximizing the imaging depth the spectral resolution needs to be minimized [12].

IV. CONCLUSION

A fiber optic low coherence interferometry system has been developed using SLED source center wavelength around 1310 nm. Studies to determine the optimum spectrometer resolution, for depth resolved detection from the sample has been demonstrated. The spectrometer resolution of 0.5 nm was found to be optimum for obtaining depth resolved information from the sample compared to other resolution which was verified with the actual thickness provided by the manufacturer. A physical depth of coverslip was estimated from the optical depth obtained and known refractive index of the sample. This study will be further advanced towards detecting skin abnormalities or burns, by using the high-resolution, non-invasive, non-contact and depth wise detection advantages of optical coherence tomography technique.

REFERENCES

- [1] Ajay Kumar, Sarika Hinge, Hemant Dixit, Rajesh Kanawade, and Gauri Kulkarni, "Skin Mimicking Solid Optical Tissue Phantom Fulfillment and its Characterization", Photonics in Dermatology and Plastic Surgery 2022, Proc. of SPIE Vol. 11934 119340.
- [2] Korde, V.R.; Bonnema, G.T.; Xu, W.; Krishnamurthy, C.; Ranger-Moore, J.; Saboda, K.; Slayton, L.D.; Salasche, S.J.; Warneke, J.A.; Alberts, D.S.; et al., "Using optical coherence tomography to evaluate skin sun damage and precancer", Lasers Surg. Med. 2007,39, 9.
- [3] M.A. Richard, C. Paul, T. Nijsten,P. Gisondi, C. Salavastru,C. Taieb,M. Trakatelli, L. Puig,A. Stratigos, " Prevalence of most common skin diseases in Europe: a population-based study", DOI: 10.1111/jdv.18050 JEADV 2022, 36, 1088–1096.
- [4] B. E. Bouma and G. J. Tearney, Handbook of optical coherence tomography (Marcel Dekker, Inc., 2002).
- [5] John, P., Manoj, M., Sujatha, N., Vasa, N.J. and Rao, S.R., 2015, June. Aqueous glucose measurement using differential absorption-based frequency domain optical coherence tomography at wavelengths of 1310 nm and 1625 nm. In European Conference on Biomedical Optics (p. 953711). Optica Publishing Group.
- [6] Themstrup, L.; Welzel, J.; Ciardo, S.; Kaestle, R.; Ulrich, M.; Holmes,J.; Whitehead, R.; Sattler, E.C.; Kindermann, N.; Pellacani, G.; et al."Validation of Dynamic optical coherence tomography for non-invasive, in vivo microcirculation imaging of the skin", Microvasc. Res. 2016, 107, 97–105.
- [7] A. F. Fercher, C. K. Hitzenberger, G. Kamp, and S. Y. El-Zaiat,"Measurement of intraocular distances by backscattering spectral interferometry," Optics Communications, vol. 117, pp. 43-48, 1995.
- [8] John, P., N., Vasa and Sujatha, " Fiber-Optic Differential Absorption Low Coherence Interferometry Technique for Sensing Glucose Selectively," 2020 Conference on Lasers and Electro-Optics Pacific Rim (CLEO-PR), Sydney,NSW,Australia,2020,pp.1-2,doi:10.1364/CLEOPR.2020.P3_11.
- [9] A. Delalleau, G. Josse, J.-M. Lagarde, H. Zahouani, and J.M. Bergheau, "A nonlinear elastic behavior to identify the mechanical parameters of human skin in vivo," Skin Research and Technology, vol. 14, no. 2, pp. 152–164, 2008.
- [10] Abhijit A. Gurjarpadhye, William C. Vogt, Yajing Liu, and Christopher G. Rylander, " Effect of Localized Mechanical Indentation on Skin Water Content Evaluated Using OCT", International Journal of Biomedical Imaging, Vol. 2011,8 pages, doi: :10.1155/2011/817250.
- [11] Paola Di Ninni, Fabrizio Martelli, and Giovanni Zaccanti, "The use of India ink in tissue-simulating phantoms," Optics Express, Vol. 18, issue 26, pp.26854-26856,2010
- [12] Jiraporn Saenjae, Jadsada Saetiew, Kunakorn Palawong, and Panomsak Meemon, " An Experimental Procedure for Quantitative Evaluation of Spectral Resolution of Spectrometer based Spectral Domain Optical Coherence Tomography", 978-1-5386-0882-1/17/\$31.00 ©2017 IEEE.

An analysis of the design and evolution of braces for lower back pain

Bethanney Janney J*

Dept. of Biomedical Engg

Sathyabama Institute of Science and Technology
Chennai, India
bethanneyj@gmail.com

Ponni S

Dept. of Biomedical Engg.

Sathyabama Institute of Science and Technology
Chennai
ponnibryavi@gmail.com

Kishore R

Dept. of Biomedical Engg.

Sathyabama Institute of Science and Technology
Chennai
kishorekishore31849@gmail.com

Keerthivasan V

Dept. of Biomedical Engg.

Sathyabama Institute of Science and Technology
Chennai
keerthivas2610@gmail.com

Meenatchi S

Dept. of Biomedical Engg.

Sathyabama Institute of Science and Technology
Chennai
meenu372003@gmail.com

Saraswathi B

Dept. of Biomedical Engg.

Sathyabama Institute of Science and Technology
Chennai
sarawathibalaji1810@gmail.com

Abstract— The ever-increasing sophistication of the world's technologies is drastically altering the way people live. The spread of industrialization and urbanization is directly responsible for the increase in sedentary lifestyles in the modern world. People who overuse technology often neglect their posture and slouch for long periods of time. According to statistics, back pain is the third leading cause of doctor visits, making proper posture even more important for a healthy lifestyle. But individuals often undervalue our capacity for upright movement, whether on purpose or accidentally. The suggested work seeks to develop an effective method for identifying and warning of posture position. The methodology centres on the evaluation of sensor placement in the brace to detect the wearer's spinal condition and design the circuit combining force sensors with an Arduino to evaluate the spine disfigurement and patient comfort.

Keywords— Brace, Detection, Posture, FSR sensor, Arduino, IMU MPU, Android.

I. INTRODUCTION

Every second of our lives is marked by a new set of body mechanics, and it is well known that certain postures if maintained over time, can have serious consequences for the health of the spine. Because of the increased demands of today's consumers and the competitive nature of the manufacturing sector, many people unknowingly spend long periods of time in an improper working posture without realising it. Major back issues can be traced back to prolonged incorrect posture [1].

In India, the combined incidence of lower back pain at the point, annually, and over the course of a lifespan was 66 percent. The aggregate occurrence percentages were greatest among females, people living in rural areas, and secondary laborers. The results of this research can serve as the foundation for developing policy for the prevention and treatment of lower back pain in a significant portion of the world's population. [2]. The state of our posture should be at the top of our list of health concerns, yet most of us are unable to make a intensive effort toward maintaining the correct posture as we go about our daily lives. Neglecting to maintain a posture that doesn't put our spine in danger, as failure to do so can result in shoulder and aching muscles, decreased respiratory performance, gastro symptoms, scoliosis, positional disorders, and a number of other disorders that can seriously affect our day-to-day life [3].

II. RELATED WORKS

C. C. Lim et al. (2014) designed and created a device that can identify incorrect spinal alignment by considering each component of the human back. The accelerometer in the device performs the function that was designed for it. Leaning forward or backward is one of the ways that incorrect posture can be identified, and after certain accelerometer readings have been reached, the user will be notified by the sound of an audible buzzer [3].

An intelligent wearable device was created by Ozgül et al. (2022) to identify and manage posture ailments. The suggested method is made to suggest the best workouts to prevent any physical inconveniences. Additionally, by gathering information about the individual wearing the vest through sensors, it seeks to identify hunched posture. Furthermore, it is advised to warn the person vocally when they are in a hunched position to rectify the posture ailment [4].

Tanlamai et al. (2022) demonstrated a device that uses three torso-mounted sensors to monitor body posture. The system was created, designed, and evaluated in two steps: First creating a posture monitoring device that considers both technological and user experience factors; and second determining the attitudes and opinions of senior citizens living in residential homes towards the tracking of bodily posture using a portable device fastened to the torso [5].

A computing strategy for the development of braces in case of adolescent scoliosis was suggested by Kardash et al. (2022). To maximize the transfer between active treatment and therapeutic response the suggested technique used a unique modelling of the patient's trunk to identify the brace shape. They obtained trunk models with individualized geometry for five patients with adolescent idiopathic scoliosis, and created Boston-style braces to assess the recommended methods. Under suitable comfort settings, the proposed braces enhance clinical indicators in a simulation context by an average of 45% [6].

Ferado et al. (2022) created the tools to facilitate research of poor posture using motion capture sensors. With the help of these resources, a benchmark sensor for the automatic identification of incorrect sitting postures was developed, and it was then tested to see how well Hjorth's parameters like activity, mobility, and complexity are applied to the categorization problem. They particularly generated Hjorth's time-domain parameters based on accelerometer data,

stacked them as feature vectors, and submitted them to a binary classifier. The suggested sitting position identification approach was experimentally evaluated in a configuration with standard and ergonomic keyboard type, and it observed an overall detection accuracy of up to 98.4% [7].

Pinero et al. (2021) used a specific hardware system that interprets real time video using convolutional neural network and devised and tested the model based on the postural detection of workers. In order to help the worker, avoid potential health issues brought on by bad posture, this system identifies the worker's neck, shoulders, and arm posture. According to the suggested system's results, the video computation may be executed in real time with little electricity consumption (less than 10 watts) and with a pattern recognition rate of more than 80% using powerful computers [8].

The performance of a squat was studied by Anne Schmitz et al. (2015) utilizing a markerless motion capture system made up of a single camera. They described the various squat positions that are appropriate for men and women. In their investigation, markerless motion capture demonstrated a technique with the potential to broaden the application of motion capture technology in therapeutic settings. However, there has been no in-vivo research done on the efficacy of using a single markerless camera system to assess clinically significant joint angles in the lower extremities. Because of this, they decided to incorporate an in-vivo comparison into their research. Joint angles during squatting were measured with a marker-based motion capture system and a Microsoft Kinect [9].

Using a force sensor, Yong-ren Huang and Xu-feng Ouyang (2012) describe an architecture that can record and analyse information relevant to sitting position. Installing force sensors into seat cushions and then constructing microcontroller circuits to collect data from the sensors are the two steps involved in the construction of the system [10].

A. Existing Techniques

By fighting against the pull of gravity, a person can keep their posture upright when they are standing, sitting, or walking. Image processing and sensor technologies are two different approaches that can be utilised to ascertain a person's posture. In the latter method, the author calculates the pressure distribution on various weight distribution surfaces. A sensor method to detect sitting posture is used to measure the angle using a goniometer and an electrogoniometer after placing two accelerometers on different parts of a person's spine. Despite its ease of use, effectiveness and ability to be worn, this design can only be put into practice when the user is seated. However, in a different method, an actuator is used as a biomechanical posture detector. The purpose of this method is to identify the user's movement phase or change in motion states by analyzing the user's position as he sits, stands up, sleeps. The actuator connects with an application, and this method also involves showing sensorial operation by a virtual [11].

On the other hand, a user's posture can be categorized and corrected with the use of an intelligent chair. To categorise the postures, neural networks are trained and introduced Wireless sensor node for a Motion Capture system with Accelerometers (WiMoCA), a custom-designed wireless body area-based sensor network for wireless and wearable posture identification. WiMoCA recognizes a person's

posture based on the position of the wearer's body. In this scenario, the sensors are portrayed as triaxial integrated MEMS (microelectromechanical system) accelerometers, and the sensor tip is customized to be wearable and operate with a minimum amount of power. To create an orientation detector foam, they adopt a technique that involves merging several straightforward sensors. Due to the need to alter the dimensions of the sensors and the associated costs, they have decided against pursuing this commercial alternative. Another approach involves the utilization of a belt that is fitted with a Bluetooth module, a gyroscope, and an accelerometer to monitor the user's movement and provide feedback on their posture. The gyroscope monitors the user's body rotation, the accelerometer determines the angle at which the user's body is tilted, and the Bluetooth module establishes a connection between the belt and the phone app [12].

III. METHODOLOGY

The setup concentrated on creating a circuit that can detect any discomfort and also use the values acquired to analyse the progress of the spine's ailment. As a result, the FSR sensor integrated into the Arduino successfully detected various pressure types concurrently and individually. The input readings for the accelerometer and gyroscope come from the IMU MPU sensor in the user's device and are transmitted to the user's android app via the bluetooth module. The final product will be a device that rests on the user's lower back. Data from the device's sensors, including whether the user is sitting or standing incorrectly, will be transmitted via Bluetooth to the user's Android app. The Android app will track the user's position and respond appropriately. Whenever the user is in a slouching position, whether seated or standing, the device will buzz or vibrate and immediately send data via bluetooth to the phone's app. Users will be able to see their posture in both a seated and a standing position through a new app for mobile devices.

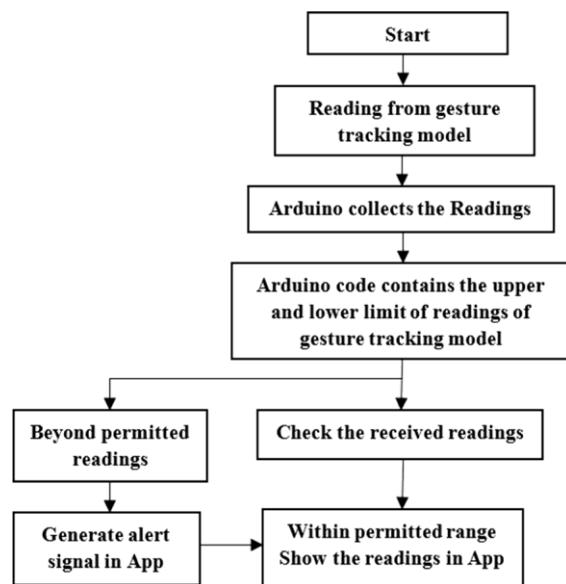


Fig. 1. Flowchart of a posture-monitoring and adjusting gadget

A. MPU-9250

It is a nine-axis gesture tracking gadget with a pin well-suited with the MPU-6515 (3x3x1mm module) that houses a 3-axis gyroscope, three - axis accelerometer, three - axis magnetometer, and a Digital Movement Processor TM (DMP). Consumer-grade motion performance, and its auxiliary I2C port is built to connect to a wide variety of non-inertial digital sensors like pressure sensors (Fig.2) [13].

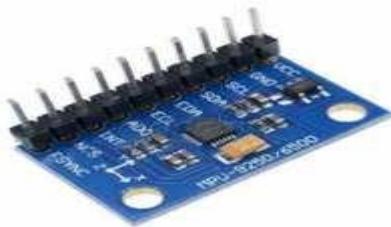


Fig. 2. Inertial Measurement Unit and Microprocessor

B. Arduino Nano

Depending on the ATmega328 (Arduino Nano 3.x) or ATmega168 (2.x), the Arduino Nano is a compact, full-featured board that is easy to use with a breadboard. It has nearly all the capabilities of the Arduino, with the exception of a DC power jack [14]. The Arduino UNO is on the right in Fig. 3.



Fig. 3. The Arduino Nano and Uno.

C. Buzzer

A buzzer, with its tiny size and 2-pin structure that allows it to be used on breadboards, Perf Boards, and even PCBs, is a useful component for adding audio to the system [15].

D. Bluetooth HC-05

HC05 module is a transparent Bluetooth SPP (Serial Port Protocol) module for setting up wireless serial connections. It can be used in either a Master or Slave configuration [10].

Solution for wireless communication, The CSR Blue core 04 Exterior sole chip Bluetooth model employs CMOS technology and AFH to provide a Bluetooth V2.0+EDR (Enhanced Data Rate) 3Mbps Signal processing serial interface Bluetooth module that is completely authorized (Adaptive Frequency Hopping Feature) [16, 17].

IV. RESULTS AND DISCUSSION

The hardware model Fig. 4 is put together to investigate the state of lower back pain in relation to backbone abnormalities.

Various analogue values were obtained in relation to the pressure sensor measured. It comprises of an Arduino board coupled to two Force Sensing Resistors (FSR) to detect the pressure of the body applies to the brace. At the location, one of the sensor's lead is fixed and resistors are positioned. Analog inputs are connected to arduino uno's a0 and a1 pins, and the power source is drawn from the Arduino board. The microprocessor gets data from the tri-axial accelerometer at regular intervals, normalizes the received data and determines a postural description of the user. The sensor device (Fig. 5) may additionally comprise memory, an actuator, and a power source.

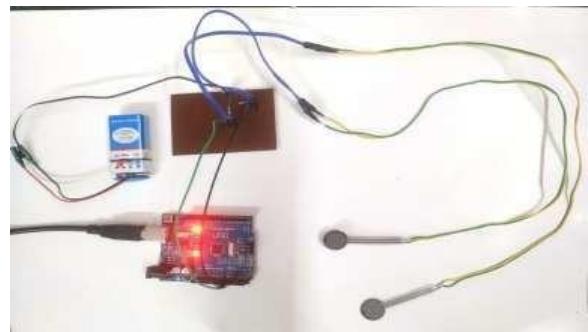


Fig. 4. Hardware model



Fig. 5. Prototype of brace for lower back pain.

The sensors were used to take readings based on the pressure that was experienced by the brace at different point of time. The output was obtained using the arduino IDE in the serial monitor shown in Fig. 6. The serial monitor showed different analog readings (FSR resistance in ohms) such as light touch, light squeeze, medium squeeze, big squeeze based on the amount of pressure the brace experienced and no pressure when the brace was left undisturbed. The FSR's lowest and maximum detectable pressures depend on its sensing range. FSR is essentially a resistor whose resistive value changes depending on the pressure given to it. The output voltage of the sensor ranges from 0V to 5V. This analogue voltage is transformed by the Arduino into a 10-bit integer with a range of 0 to 1000. Therefore, depending on how firmly you squeeze the sensor, the number shown on the serial monitor will range from 0 to 1000. Table I. shows the information about which command is executed based on the analog value of pressure indication.

REFERENCES

TABLE I. Command generated based on analog values

Analog Values (FSR Resistance in Ohms)	Command
<10	No Pressure
101-300	Light Touch
301-500	Light Squeeze
501-700	Medium squeeze
> 700	Big Squeeze

```

COM5
:
: Analog reading = 0 - No pressure
Analog reading = 0 - No pressure
:
: Analog reading = 0 - No pressure
Analog reading = 0 - No pressure
:
: Analog reading = 746 - Medium squeeze
Analog reading = 0 - No pressure
:
: Analog reading = 207 - Light squeeze
Analog reading = 0 - No pressure
:
: Analog reading = 321 - Light squeeze
Analog reading = 0 - No pressure
:
: Analog reading = 0 - No pressure
Analog reading = 0 - No pressure
:
: Analog reading = 654 - Medium squeeze
Analog reading = 47 - Light touch
:
: Analog reading = 642 - Medium squeeze
Analog reading = 288 - Light squeeze
:
: Analog reading = 737 - Medium squeeze
Analog reading = 309 - Light squeeze
:
: Analog reading = 754 - Medium squeeze
Analog reading = 287 - Light squeeze
:
: Analog reading = 0 - No pressure
Analog reading = 155 - Light touch
:
: Analog reading = 0 - No pressure
Analog reading = 0 - No pressure
:
: Analog reading = 176 - Light touch

```

Fig. 6. Serial monitor output showing different analog readings.

V. CONCLUSION

The proposed work led us to conclude that a wearable device is used to monitor and alert the poor back posture. Designers settled on an efficient approach based on the use of an arduino nano and an inertial measurement unit (IMU) as its main components, which would allow the device to be small enough to be attached to the user's back with minimal effort. Bluetooth allows for the development of an Android app that can track the user's spinal lordosis. This proposed model demonstrates that it is possible to create a sensor-integrated lower back brace using additive manufacturing, which is significantly more convenient, accurate, and time-efficient than traditional approaches. Additionally, it is a tried-and-true approach for non-surgically treating scoliosis. In future work the posture correction could be incorporated with the help of curvature analysis through the 3-matic research13.0 software.

ACKNOWLEDGMENT

We are grateful for the comfortable workspace provided by our college and the experts at Sathyabama Institute of Science and Technology, which allowed us to finish our work.

- [1] Q. W. Oung, M. Hariharan, H. L. Lee, S. N. Basah, M. Sarilee, and C. H. Lee, "Wearable multimodal sensors for evaluation of patients with Parkinson disease," Proc. - 5th IEEE Int. Conf. Control Syst. Comput. Eng. ICCSCE, pp. 269–274, 2016.
- [2] Shetty GM, Jain S, Thakur H, Khanna K. Prevalence of low back pain in India: A systematic review and meta-analysis. Work. 2022;73(2):429-452. doi: 10.3233/WOR-205300. PMID: 35964222.
- [3] A. B. Crane, S. K. Doppalapudi, J. O'Leary, P. Ozarek and C. T. Wagner, "Wearable posture detection system," 2014 40th Annual Northeast Bioengineering Conference (NEBEC), pp. 1-2, doi: 10.1109/NEBEC.2014.6972764, 2014.
- [4] Özgül, G. & Patlar Akbulut, F, "Wearable sensor device for posture monitoring and analysis during daily activities: A preliminary study," International Advanced Researches and Engineering Journal, vol. 6 (1), pp. 43-48, 2022.
- [5] Tanlamai U, Jaikengkit A-O, Jarutach T, Rajkulchais S, Ritbumroong T, "Use of daily posture and activity tracking to assess sedentary behavior, toss-and-turns, and sleep duration of independently living Thai seniors," Health Informatics Journal, vol. 28(1), 2022. doi:10.1177/14604582211070214.
- [6] Kardash Kateryna, Koutras Christos, Otaduy Miguel A, "Design of personalized scoliosis braces based on differentiable biomechanics—Synthetic study," Frontiers in Bioengineering and Biotechnology, vol. 10, pp. 1-13, 2022.
- [7] Feradov, Firgan, Valentina Markova, and Todor Ganchev, "Automated Detection of Improper Sitting Postures in Computer Users Based on Motion Capture Sensors," Computers, Vol. 11(7), pp. 116-125, 2022, <https://doi.org/10.3390/computers11070116>.
- [8] Pinero Fuentes, Enrique, Salvador Canas-Moreno, Antonio Rios-Navarro, Manuel Domínguez-Morales, José Luis Sevillano, and Alejandro Linares-Barranco, "A Deep-Learning Based Posture Detection System for Preventing Telework-Related Musculoskeletal Disorders," Sensors 21, vol. 15(5236), pp. 1-16, 2021.
- [9] A. Schmitz, M. Ye, G. Boggess, R. Shapiro, R. Yang, and B. Noehren, "44 measurement of in vivo joint angles during a squat using a single camera markerless motion capture system as compared to a marker-based system," Gait Posture, vol. 41(2), pp. 694– 698, 2015.
- [10] Y.R. Huang y X.-F. Ouyang, "Sitting Posture Detecting And Recognition Using Force Sensor," International Conference on BioMedical Engineering and Informatics, Chongqing, 2012.
- [11] J. Gálvez-Goicuría, J. Pagán, A. B. Gago-Veiga, J. M. Moya and J. L. Ayala, "Cluster-Then-Classify Methodology for the Identification of Pain Episodes in Chronic Diseases," IEEE Journal of Biomedical and Health Informatics, vol. 26(5), pp. 2339-2350, 2022.
- [12] Cajamarca G, Rodríguez I, Herskovic V, Campos M, Riofrío JC, "StraightenUp+: Monitoring of Posture during Daily Activities for Older Persons Using Wearable Sensors," Sensors, vol.18(10), pp. 3409, 2018. <https://doi.org/10.3390/s18103409>
- [13] R. Lim and M. Sikkandhar, "Pressure Sensor Substrate for Prolonged Sitting and Posture Monitoring," 2021 IEEE 23rd Electronics Packaging Technology Conference (EPTC), Singapore, Singapore, 2021, pp. 284-287, doi: 10.1109/EPTC53413.2021.9663903.
- [14] Chen, Duo and Zhang, Haihong and Kavitha, "Scalp EEG-Based Pain Detection Using Convolutional Neural Network," IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 30, pp. 274-285, 2022.
- [15] P. Werner, D. Lopez-Martinez, S. Walter, A. Al-Hamadi, S. Gruss and R. W. Picard, "Automatic Recognition Methods Supporting Pain Assessment: A Survey," IEEE Transactions on Affective Computing, vol. 13(1), pp. 530-552, 2022.
- [16] V. Srhoj-Egekher, M. Cifrek and S. Pešarec, "Feature Modeling for Interpretable Low Back Pain Classification Based on Surface EMG," in IEEE Access, vol. 10, pp. 73702-73727, 2022.
- [17] L. Cao, H. Zhang, N. Li, X. Wang, W. Ri and L. Feng, "Category-Aware Chronic Stress Detection on Microblogs," in IEEE Journal of Biomedical and Health Informatics, vol. 26(2), pp. 852-864, 2022.

Sitting posture Analysis using CNN and RCNN

Nishitha R

Biomedical Engg. Department
Sathyabama Institute of Science and
Technology
Chennai

N. R. Krishnamoorthy

Dept. of ECE
Sathyabama Institute of Science &
Technology, Chennai
moorthy26.82@gmail.com

T. Sudhakar*

Biomedical Engg. Department
Sathyabama Institute of Science and
Technology, Chennai
drsudhakar35@gmail.com

Bethanney Janney J

Biomedical Engg. Department
Sathyabama Institute of Science and
Technology, Chennai
bethanneyj@gmail.com

Vigneshwaran S

M.Tech-Medical Instrumentation
Sathyabama Institute of Science and
Technology, Chennai
vickyselvaraj20@gmail.com

Abstract—Based to statistics, about 80 percent of the population follows incorrect sitting posture. From the year 2020, due to the Covid-19 curfew, the traditional work office of many professionals turned into Work from home and remote office lifestyle. Many of the working professionals has no proper ergonomic workplace setup in their house leading to improper sitting position. Monitoring the sitting posture is needed to avoid chronic complication of MSD, neck, and spine related injuries. Many techniques were emerged for correcting and monitoring the posture. Neural Network based posture detection is being into the field of research as the come with more range of accuracy. This paper discusses about the various Neural Networks used and their accuracies.

Keywords—MSD-musculoskeletal disorder, ergonomic workplace, convolution neural networks (CNN) and region based convolutional neural network (RCNN).

I. INTRODUCTION

As many of the people adopt for the work from home lifestyle the number of people affected by the spine and neck disorder has been increased to greater extend. The primary cause for the complication of disorder is that many professionals are not supported by the ergonomic workspace. All employees cannot make ergonomic workplace which result in spinal complications. In order to prevent the chronicity of the disorder, preventive posture monitoring and sitting posture correcting tools been emerging in the market. Both wearable and non-contact tools have been under research to make a preventive tool for the position monitoring. Many sensors been implemented in the posture management like Force resist sensor which is placed on the cushion of sitting to calculate the force acting on the sensors, Pressure sensor to detect the evenly distributed pressure on the chair. Also, some research works on the non-contact supportive tools so that the persons are free from sensors. One of the predominant non-contact tools is the use of Neural Network to monitor the posture of the person. Many research work has been carried out to extract the accurate posture detection.

II. EXISTING METHODOLOGIES

Research for sitting posture detection and analysis have been increasing for past couple of years. Predominantly the posture detection is carried out either by wearable sensors or by the Machine learning methods. Many research work stated the use of sensors like Flex sensors [1], flexible array pressure sensor [12], FRS and Sitting pressure sensors (SPS) [10], [2] will help in measuring the pressure and/or force exerted over the cushion while sitting. Some articles state the combinational use of sensors and Neural network for obtaining more accuracies. Technologies like IoT [2], [4] are

used alone with sensors and machine learning for storing the data on cloud. By using sensors, the changes in pressure and/or force are monitored so that the persons are alerted when the values exceed the threshold in any of the sensors being used. In order to remove the noise and artifacts during sensor recording, they are supported by various filters like Kalman filter [2]. The monitoring and value analysis are made ease by creating a supporting android application [2]. Some research uses k-Nearest neighbor's, support vector machine, random forest, decision tree [10] for experimental analysis along with raspberry pi. Fully conventional neural network to implement the embeddable human posture recognition system [5] are used and proved that the accuracy is achieved at 98.99%. The use of Kinect V2 [17], [6] for classification of human body joints like using the Open Pose library. For performing the image classification either MATLAB [17] or python are used. In a research paper, integration of SVM, CNN and LSTM [19] are made to study about the posture of a person [9]. 3-D RCNN [20] with spatiotemporal process is used for multi-stage technique to recognize the position of a person. Neural network is emerging in this domain as they produce high range of accuracy with greater classification. Optical neural network [22] has random shift wavelet pattern for easy posture recognition.

III. MATERIALS AND METHODS

A. Neural Network

Neural Network is a part of Machine Learning which makes the device to learn itself with the existing inputs. As the name implies, Neural network resembles the structure of Brain Neurons having many networks of weight layers such that the accuracy and the learning rate of the device is being improvised.

B. CNN

The Convolutional Neural Network is a part of Neural Networks that performs the mathematical convolution to train the network. This project uses a three-layer CNN. The CNN has feature map which generate both Part Confidence Maps and Part Affinity Field.

C. RCNN

Region based convolutional neural network is more derived version of neural network that concentrate deep on the region- based classification of the object in the image. RCNN is used because it mainly concentrates on the image classification based on region detection and object segmentation. Dataset are supported by MS-COCO dataset. Key point RCNN is slightly difference from Mask-RCNN.

D. Python

For coding and training the Neural networks, python is being used as python is one of the general-purpose, user-friendly programming language and is easily adopted by many non-programmers. Python along with libraries like TensorFlow, Scikit-learn, Keras, Pytorch can be used for Machine learning tasks.

E. Keras

Python written deep learning API which runs on the top of machine learning platform TensorFlow. This API is simple, flexible, and powerful used for research as it supports fast experimental procedure. Layers and models are the core data structures of Keras.

F. Gaussian filter

The gaussian filter is a type of low pass filter. This filter is used for reducing the noise and blurring in an input image. Here multidimensional gaussian filter is used. It targets to retain the edges so that key point extraction is made at ease.

G. Open Pose

OpenPose is a multi-person detecting real-time library can be used for detecting the human body key points. OpenPose is written in C++ and Caffe. OpenPose with python can support to create different operating systems and languages.

H. TensorFlow

An open-source platform for machine learning where the pose estimation can be achieved by merely estimating the key points and key joints output. It is a core library and has comprehensive, ecosystem of tools, community resources that helps the programmer to have state of art in development of Neural network.

I. Torchvision

Torchvision is a model that contains pretrained key point detection model which are built on top of the ResNet-50 FPN. FPN is a process of fusing features maps at multiple scales for preserving the information at multiple levels.

IV. PROPOSED METHODOLOGY

The sitting posture analysis using CNN and RCNN is a comparative work to analyse the accuracy and latency of the CNN and RCNN in detecting the posture of the person. The output of the network indicated the keypoints extracted from the given input and the latency for detection of the system. The postures that we considered to detecting the position are: a. Hunchback, b. Reclined, c. Straight, d. hand folds, e. kneeling, f. cross legs. This software detects the sitting position by getting the input from the lateral view of the person as we aim to calculate the posture by calculating the angle and degree between the keypoint of reference. The key points are extracted from the image by using OpenPose model. In this work 18 keypoints in the body are extracted and analysed.

Steps carried to recognize the sitting posture are:

- Gather the keypoint coordinates for different body parts.
- Calculate the sitting angle of the person.
- Getting the ear position

- Getting the hip position
- Calculating the angle between ear and hip

If angle > 110 → Reclined position; Elif angle > 70 → Hunchback position; Else → Straight position

Initially the software gets the input image and extract the keypoints from the person. This is achieved by constructing three python files: one for containing the whole architecture of the software, second containing the parameters essential to predict the keypoints, third having the functions to calculate the co-ordinates of keypoints.

Once the keypoints are extracted the system starts detecting the angle between the keypoints to analyse the deviation of spine from straight posture. Here we considered Hip and Ear keypoint since both the keypoints lay on same line when they are in straight position.

The above work has four programs:

- Posture detection in an input image using CNN.
- Posture detection in an input image using RCNN.
- Posture detection in an input video and real-time video input using CNN.
- Posture detection in an input video and real-time video input using RCNN.

A. CNN Posture detection

Initially the image is passed through baseline CNN network to extract the feature maps of input. The feature map is processed in 3 stage CNN to obtain Part Confidence Maps and Part Affinity Field. Greedy bipartite matching algorithm are used to get Part Confidence Maps (PCM) and Part Affinity Field (PAF). Confidence Map is a 2D representation that a particular body part can be located in pixel. PAF is a set of 2D vector field which encodes location and orientation of limbs of people. It encodes pairwise connections between body parts.

Three steps carried out in CNN are:

- To predict the Part Affinity Fields from the Feature maps of base network
- Use the output PAF from previous layers to refine the prediction of Confidence maps.
- Confidence map and Part Affinity Field are passed into the greedy algorithm.
- The loss function is used to calculate the loss between PCM and PAF to the ground truth maps and fields.

Proposed CNN model consists of two convolutional layers, followed by pooling layer. CNN network accepts image with pixel size of 28x28x3, and finally results in the 14x14x3 size features signal after passing through the two set of convolution layer with Relu activation function and max pooling layer. The input posture image will be processed by two convolution layers with 16 filters; each has kernel size of 3 in the first set and has kernel size of 4 in the second set of convolution layer. Further the extracted features will be down sampled to 14x14x3 image size using two max pooling layers. The confusion matrix of CNN model is shown in the

figure 1. The accuracy of the model is evaluated for six different posture and overall accuracy of the CNN is obtained as 92.5%

Confusion Chart - CNN						
True Class	Hunchback	18	1	1		
	Reclined	1	19			
	Straight	1	1	18		
	cross legs			19		1
	hand folds			1	18	1
	kneeling			1	19	
	Hunchback	Reclined	Straight	cross legs	hand folds	kneeling
Predicted Class						

Fig.1. Confusion matrix of CNN Model

B. CNN Realtime posture detection

The CNN is coded for detecting the live input video streaming on the web camera attached to the laptop/PC and to process the recorded video input for detecting the posture of the person. The Torchvision is used for the dataset with gaussian filter to remove the noises in the input. If the posture is not able to detect or if the camera is not in the lateral position to the person, then the system responds with a notification of “unable to calculate the angle”.

C. RCNN Posture detection

Keypoint RCNN have same functionality as Mask-RCNN, they differ by output size and the way keypoints encoding in the keypoint mask. Keypoint RCNN has one-hot encoding a keypoint of the detected object instead of whole mask. Class-wise output feature map from Mask-RCNN: Here the feature map has two channels person and the background class. Pretrained keypoint detection model built on top of the ResNet– 50 FPN backbone. FPN is a fusion feature maps at different scales so that the information is preserved. The confusion matrix of RCNN model is shown in the figure 2. The accuracy of the model is evaluated for six different posture and overall accuracy of the RCNN is obtained as 95%

Faster RCNN has numerous layers where the Region-proposal- layer predicts the locations of N number of objects in feature maps. These regions are individually passed to ROI-Pooling layer which resizes the feature maps by quantising size grid to fixed size grid and getting the max-values to place in the fixed grid. A Fully Connected (FC) layer follows ROI-pooling layer. Mask-RCNN has some modification of layers, like ROI-Align layer is used instead of ROI-Pooling layer and the output of ROI-Align is passed to Mask-RCNN head. COCO dataset has 80 classes for detection and segmentation, the annotations are offered for person class. Torchvision is particularly trained to identify key points in a person.

Confusion Chart - RCNN						
True Class	Hunchback	19		1		
	Reclined	1	19			
	Straight		1	19		
	cross legs				19	
	hand folds				1	19
	kneeling				1	19
	Hunchback	Reclined	Straight	cross legs	hand folds	kneeling
Predicted Class						

Fig.2. Confusion matrix of RCNN Model

D. RCNN Realtime Posture detection

RCNN is coded for analysing the live streaming video and recorded video input. The real time input is taken and processed with dataset from Torchvision, and the noise are removed by using the gaussian filter. If the camera is in wrong position or if the lateral view of the subject is not properly detected, then the system comes with a alert message of “Unable to calculate the angle”, later which the position can be altered.

V. RESULTS AND DISCUSSION

The comparison of CNN and RCNN output time for same set of images are represented in Table I and Table II.

TABLE I. RESULTS FOR CNN

Input (Image)	Expected Result (Position)	Actual Result (Position)	Time in Seconds
Sample 1	Hunchback	Hunchback	56.30753
Sample 2	Hunchback	Hunchback	61.99616
Sample 3	Hunchback	Hunchback	72.53529
Sample 4	Reclined	Reclined	52.14701
Sample 5	Reclined	Reclined	70.73289
Sample 6	Reclined	Reclined	64.77853
Sample 7	Straight	Straight	82.54892
Sample 8	Straight	Straight	69.10063
Sample 9	Straight	Straight	69.49605

TABLE II. RESULTS FOR RCNN

Input (Image)	Expected Result (Position)	Actual Result (Position)	Time in Seconds
Sample 1	Hunchback	Hunchback	9.21453
Sample 2	Hunchback	Hunchback	9.80878
Sample 3	Hunchback	Hunchback	9.56103
Sample 4	Reclined	Reclined	8.6486
Sample 5	Reclined	Reclined	8.51954
Sample 6	Reclined	Reclined	8.89083
Sample 7	Straight	Straight	8.1843
Sample 8	Straight	Straight	8.41329
Sample 9	Straight	Straight	9.71633

Results for CNN, RCNN are shown in Table. III and Table IV. Based on the result, the CNN has more latency period in detecting the posture than the time taken by the RCNN network. The key point detection for the image and output of the posture detection are as follows:

TABLE III. OUTPUT FOR CNN

Position	Key point detected Output Image	Output of CNN
Hunchback		processing time1 is 58.31735 secs processing time2 is 153.17488 secs Person is in Hunchback position! Incorrect Spine Posture
Straight		To enable them in other operations, rebuild TensorFlow's <code>tf.flags</code> . processing time1 is 69.10063 secs processing time2 is 151.47656 secs Person is in Straight position! Correct Spine Posture
Reclined		processing time1 is 52.1701 secs processing time2 is 94.5633 secs Person is in Reclined position! Incorrect Spine Posture

TABLE IV. OUTPUT FOR RCNN

Position	Key point detected output image	Output of RCNN
Hunchback		start processing... C:\Users\91988\anaconda3\lib\site-packages\torch\func.py: pass the indexing argument. (Triggered internally at ... return _VF.meshgrid(tensors, **kwargs) # type: ignore[call-arg] Degree is: 50 Person is in Hunchback position! Incorrect Spine Posture
Straight		C:\Users\91988\anaconda3\lib\site-packages\torch\func.py: pass the indexing argument. (Triggered internally at ... return _VF.meshgrid(tensors, **kwargs) # type: ignore[call-arg] processing time2 is 8.41329 Angle deviated from correct posture: 1.00005/0.04549104 rad Degree is: 78 Person is in Straight position! Correct Spine Posture
Reclined		start processing... C:\Users\91988\anaconda3\lib\site-packages\torch\func.py: pass the indexing argument. (Triggered internally at ... return _VF.meshgrid(tensors, **kwargs) # type: ignore[call-arg] processing time2 is 8.64866 Degree is: 114 Person is in Reclined position! Incorrect Spine Posture

VI. CONCLUSION

The spinal posture of the person while working with computer have been analysed using the CNN and RCNN network. A comparative study has been plotted between the time taken by both the networks for detecting the posture. Based on the study for multiple images, it is clear that the RCNN have minimal latency for detecting the person position when compared to CNN (see Tab A and Tab B). This software supports to detect key points of many people, but posture analysis can be done only for one person at a time. This can be extended by iterating the detection part of the code for overall key points that are detected on each individual in the input. The research can further be extended by taking other type of Neural Networks and compare their accuracy and latency in detecting the posture.

ACKNOWLEDGMENT

I wish to acknowledge the help provided by Biomedical department and my college. I would like to show my gratitude to my guides for their guidance and support in the project phase.

REFERENCES

- [1] Q. Hu, X. Tang and W. Tang, "A Smart Chair Sitting Posture Recognition System Using Flex Sensors and FPGA Implemented Artificial Neural Network," in IEEE Sensors Journal, vol. 20(14), pp. 8007-8016, 2020, doi: 10.1109/JSEN.2020.2980207.
- [2] Arif Reza Anwary, Deniz Cetinkaya, Michael Vassallo, Hamid Bouchachia, "Smart-Cover: A real time sitting posture monitoring system," Sensors and Actuators A: Physical, vol. 317, pp.112451, 2021.
- [3] R. Hasib, K. N. Khan, M. Yu and M. S. Khan, "Vision-based Human Posture Classification and Fall Detection using Convolutional Neural Network," 2021 International Conference on Artificial Intelligence (ICAI), Islamabad, Pakistan, pp. 74-79, 2021. doi: 10.1109/ICAI52203.2021.9445263.
- [4] Gupta, Jahnvi; Gupta, Nitin; Kumar, Mukesh; Duggal, Ritwik, "An Introduction to IoT Based Posture Detection and Underlying Sensor Technology," TechRxiv. Preprint, pp.1-21,2020. <https://doi.org/10.36227/techrxiv.13466597.v1>.
- [5] Licciardo GD, Russo A, Naddeo A, Cappetti N, Di Benedetto L, Rubino A, Liguori R, "A Resource Constrained Neural Network for the Design of Embedded Human Posture Recognition Systems," Applied Sciences, vol. 11(11), 4752, 2021. <https://doi.org/10.3390/app11114752>.
- [6] Zhang, Y., Zhu, T., Ning, H. et al. "Classroom student posture recognition based on an improved high-resolution network," J Wireless Com Network, vol. 140, 2021. <https://doi.org/10.1186/s13638-021-02015-0>.
- [7] Piñero-Fuentes E, Canas-Moreno S, Rios-Navarro A, Domínguez-Morales M, Sevillano JL, Linares-Barranco A, "A Deep-Learning Based Posture Detection System for Preventing Telework-Related Musculoskeletal Disorders," Sensors, 21(15):5236, 2021, <https://doi.org/10.3390/s21155236>.
- [8] U Ran, Cong Wang, Yao Xiao, Xuliang Gao, Zhiyuan Zhu, Bin Chen, "A portable sitting posture monitoring system based on a pressure sensor array and machine learning," Sensors and Actuators A: Physical, vol. 331, 112900, 2021, <https://doi.org/10.1016/j.sna.2021.112900>.
- [9] S. Liaqat, K. Dashtipour, K. Arshad, K. Assaleh and N. Ramzan, "A Hybrid Posture Detection Framework: Integrating Machine Learning and Deep Neural Networks," in IEEE Sensors Journal, vol. 21, no. 7, pp. 9515-9522, 2021, doi: 10.1109/JSEN.2021.3055898.
- [10] Ahmad J, Sidén J, Andersson H, "A Proposal of Implementation of Sitting Posture Monitoring System for Wheelchair Utilizing Machine Learning Methods," Sensors, vol. 21(19), 6349, 2021, <https://doi.org/10.3390/s21196349>.
- [11] Silvia Caggiari, Peter R. Worsley, Sarah L. Fryer, Joseph Mace and Dan L. Bader, "Detection Of Posture and Mobility in Individuals at Risk of Developing Pressure Ulce," Rs Elsevier Medical Engineering and Physics. Vol. 91, pp: 39-47, 2021.
- [12] Cai W, Zhao D, Zhang M, Xu Y, Li Z, "Improved Self-Organizing Map-Based Unsupervised Learning Algorithm for Sitting Posture Recognition System," Sensors, vol. 21(18), pp. 6246, 2021, <https://doi.org/10.3390/s21186246>.
- [13] Weili Ding, Bo Hu1, Han Liu, Xinming Wang and Xiangsheng Huang, "Human Posture Recognition Based on Multiple Features and Rule Learning," International Journal of Machine Learning and Cybernetics. vol. 11(11), pp:2529–2540, 2020.
- [14] Luna-Perejón F, Montes-Sánchez JM, Durán-López L, Vazquez-Baeza A, Beasley-Bohrquez I, Sevillano-Ramos JL, "IoT Device for Sitting Posture Classification Using Artificial Neural Networks. Electronics," Sensors, vol. 10(15), 1825, 2021. <https://doi.org/10.3390/electronics10151825>.
- [15] Wan Q, Zhao H, Li J, Xu P, "Hip Positioning and Sitting Posture Recognition Based on Human Sitting Pressure Image," Sensors, vol. 21(2), 426, 2021. <https://doi.org/10.3390/s21020426>
- [16] T. -H. Tran, D. T. Nguyen and T. Phuong Nguyen, "Human Posture Classification from Multiple Viewpoints and Application for Fall Detection," 2020 IEEE Eighth International Conference on Communications and Electronics (ICCE), Phu Quoc Island, Vietnam, pp. 262-267, 2021. doi: 10.1109/ICCE48956.2021.9352140.
- [17] Klishkovskaia T, Aksenenov A, Sinitca A, Zamansky A, Markelov OA, Kaplun D, "Development of Classification Algorithms for the Detection of Postures Using Non-Marker-Based Motion Capture Systems," Applied Sciences, vol. 10(11), 4028, 2020. <https://doi.org/10.3390/app10114028>.
- [18] Wafaa M. Salih Abedi, Dr. Ibraheem Nadher and Dr. Ahmed T. Sadiq, "Modification Of Deep Learning Technique for Face Expressions and Body Postures Recognitions," International Journal of Advanced Science and Technology, vol. 29 (3), pp. 313-320, 2020.
- [19] R. Gupta, D. Saini and S. Mishra, "Posture detection using Deep Learning for Time Series Data," 2020 Third International Conference on Smart Systems and Inventive Technology (ICSSIT), Tirunelveli, India, 2020, pp. 740-744, doi: 10.1109/ICSSIT48917.2020.9214223.

- [20] Wafaa M. Salih Abedi, Dr. Ibraheem Nadher and Dr. Ahmed T. Sadiq, "Modified Deep Learning Method for Body Postures Recognition," International Journal of Advanced Science and Technology. vol. 29, No. 2, (2020), pp. 3830 – 3841, 2020.
- [21] Shumei Zhang And Victor Callaghan, "Real Time Human Posture Recognition Using an Adaptive Hybrid Classifier," International Journal of Machine Learning and Cybernetics, vol. 12, pp.489–499, 2020.
- [22] H. Zhang, S. Shui, Y. Wu, Q. Yang and Y. Cai, "Posture Recognition Based on the Improved Optical Diffractive Neural Network," 2020 IEEE 3rd International Conference on Electronic Information and Communication Technology (ICEICT), Shenzhen, China, pp. 783-785, 2020, doi: 10.1109/ICEICT51264.2020.9334358.

TRACK 5

**Ninth International Conference in Biosignals, Images and Instrumentation,
SSNCE, Chennai, March 16 – 17, 2023**

Parkinson's Disease Detection And Classification Of Stages From Drawing Patterns Using Deep Learning

Karthikeyan B,
Assistant Professor,
Department of Electronics and
Communication Engineering,
*Velammal College of
Engineering and Technology*,
Madurai, India
bkk@vcet.ac.in

Jerolin Futina J B,
UG Final year,
Department of Electronics and
Communication Engineering,
*Velammal College of
Engineering and Technology*,
Madurai, India
futinafredy0706@gmail.com

Pandia Abinaya P,
UG Final Year,
Department of Electronics and
Communication Engineering,
*Velammal College of
Engineering and Technology*,
Madurai, India
pandiaabinaya@gmail.com

Shanmuga Priya R,
UG Final Year,
Department of Electronics and
Communication Engineering,
*Velammal College of
Engineering and Technology*,
Madurai, India
priyashanmuga734@gmail.com

Abstract—Parkinson's disease (PD) is a common neurodegenerative disease by which millions of people are affected worldwide. Early detection of PD and accurate classification of the disease stage is a prerequisites for effective therapy and management. In this study, we propose a new approach to detect and classify PD phases using drawing models and the universal You Only Look Once (YOLO) object recognition framework. Data used in the study were collected from 70 patients with mild PD, 70 patients with moderate PD, 70 patients with severe PD, and 70 healthy controls, who were asked to draw certain shapes with pen and paper. The images are pre-processed and fed into the YOLO object recognition model, which has been trained to recognize and classify patterns into different stages of PD. The results of the study show that the proposed approach achieved high accuracy in PD detection and disease stage classification. These results show the potential of using drawing patterns and the YOLO object recognition framework for the early detection and classification of PD. The proposed approach could have a significant impact on the early diagnosis and treatment of PD and possibly improve the quality of life of people with this disease.

Keywords—Parkinson's, drawing patterns, deep learning, CNN, YOLO.

I. INTRODUCTION

Parkinson's disease (PD) is a neurodegenerative disease that affects more than 10 million people worldwide. It is characterized by the gradual loss of dopaminergic neurons in the brain, leading to motor symptoms such as tremors, rigidity, and bradykinesia. Early detection of PD is essential for timely treatment and better patient outcomes. However, the clinical diagnosis of PD is subjective and relies heavily on observations of physical symptoms. In recent years, there has been a growing interest in the use of technology in the early detection of PD[1]. One promising approach is to use machine learning algorithms to analyze data collected from drawing patterns. Drawing patterns is a simple, non-invasive, and inexpensive method to detect PD. Studies show that individuals with PD tend to be unique in their drawing patterns, characterized by smaller and more fragmented lines [2]. In this paper, we suggest a machine learning-based approach to detect and classify PD phases

using design patterns. Our approach involves collecting data from a large group of individuals with and without PD and using that data to train a classifier.[3] The classifier is then used to classify the new data as either PD or non-PD and classify individuals with PD into different stages based on their drawing patterns. We evaluated our method using data collected from a cohort of PD patients and healthy controls. Our results show that our approach is very accurate in detecting PD with more than 90% accuracy[4]. Our approach can be a valuable tool for the early detection and classification of PD. It is non-invasive, easy to administer, and can be used in a variety of settings, including primary care and remote monitoring[5]. There is enough potential for real-life applications of machine learning techniques in this regard.[6]

II. RELATED WORKS

A study was conducted in which individuals with Parkinson's disease as well as healthy individuals were asked to provide examples of their drawings drawn on paper and with a pen. It is then through the horizontal lines in their drawings that they can be distinguished from one another. As a result, in order to train all these features for autonomous feature generation, they used various classification methods, including Naive Bayes, AdaBoost, Logistic Regression, Support Vector Machines (SVM), J48, and Random Forest classifiers, in order to train all of these features. In terms of accuracy, they have a significant amount of accuracy, and the accuracy was 91%, which is quite good.[7]

A study that uses spiral drawings made by participants in their homes and telemetric touchscreen devices to differentiate between episodic and peak dosage dyskinesia. Many of the characteristics used to distinguish between people with Parkinson's disease and those who are healthy have been taken by these characteristics. These can be entered into algorithms for machine learning. Support Vector Machine (SVM), Logistic Regression, Random Forest classifiers, and Multilayer Perceptron are some of the techniques they have applied. All of these techniques have been applied to training and the development of automatic

systems. At the end, they came to the conclusion that the Multilayer Perceptron technique had achieved well, with an accuracy result of 84%. [8]

The method of separating individuals with Parkinson's disease from healthy individuals is presented using audio equipment. The most effective results have been achieved by a variety of methods, and those methods are now also employed in training. To choose features for the model, they used principal component analysis as a method of selecting features. In addition, they were able to compare data from two different groups, such as PCA-derived features and Random Forest classifications, by taking the data from two groups and comparing them. Using the Random Forest classifier, they were able to achieve a high accuracy of 96.83% [9]

A study proposes applying machine learning and computer vision techniques to distinguish between people with Parkinson's disease and healthy people based on handwriting. Their handwriting was used to create the dataset. Their handwriting can be used to detect Parkinson's disease. They employed proper techniques including OPF, SVM, and a Naive Bayes classifier (NB). With an accuracy of 78.9%, Naive Bayes provided the most accurate results of these techniques. There is a need to improve the accuracy of this method because it is currently low. [10]

In order to evaluate Parkinson's disease patients' spatial drawings and stability motions, Deep Echo State Networks (Deep ESNs) were employed. With the use of this method, the results were 89.3% accurate. In order to diagnose PD, a study used a convolutional neural network (CNN). They used fewer layers in this work, which resulted in limited datasets and biased results. Because of this, they have a relatively poor accuracy of 72.5%. A study also employed a condensed form of AlexNet. Rather than feeding raw data into the CNN, they fed spectral points into it in place of raw data. The researcher in this instance had a respectable accuracy of 96.5%. Parkinson's spiral drawings on a digital whiteboard dataset, which included 62 participants with PD, were utilized in all three papers. [11]

A study has suggested the use of convolutional neural networks in order to develop a multistage classification system based on spiral and wave sketches in order to identify Parkinson's disease. It can be used to distinguish between those who have Parkinson's disease and healthy individuals. In this work, the system was trained and validated using a total of 102 spiral sketches and 102 wave sketches. In the study, 93.5% of the results were found to be accurate based on the data they collected. [12]

III. METHODOLOGY

A. Data Collection

We collected datasets from Kaggle's data repository, Roboflow data repository as well as from patients with Parkinson's disease in order to conduct this study. All

subjects were asked to take a test, the Spiral Drawing Test. The design of the spiral is captured with A3 size paper. The spiral design is captured on A3 size paper [13]. Using an ink pen by the individual, the spiral design is drawn on the A3 paper that has been placed on the tablet. Data were collected from healthy subjects with mild, moderate, and severe PD.

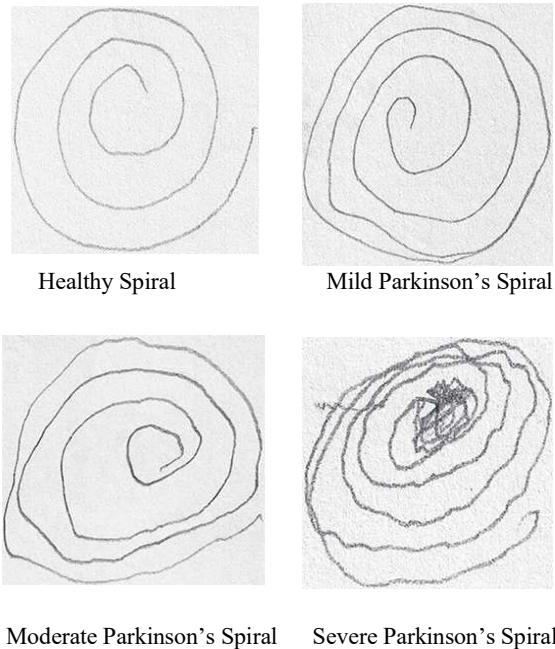


Fig.1. Images from Dataset

The folders are labeled and classified according to the stages that the images were collected. The images are divided into testing and training folders with 80% and 20% of the total images in each folder

B. Data Preprocessing

Image preprocessing is an important step in detecting Parkinson's disease and classifying its stages from drawing patterns. The goal of image preprocessing is to enhance the quality of the images and to remove any noise or artifacts that may interfere with the detection and classification process [14]. The image preprocessing steps used are

Auto-orient: Auto-orient in pre-processing typically refers to the automatic adjustment of the orientation of an image or video to a standard orientation, such as upright or landscape, based on its content or metadata. In image processing, auto-orient can be used to rotate images that were captured with a camera held in different orientations, such as portrait or landscape, so that they are all upright and have the same orientation. This can be particularly useful for large datasets or when dealing with images from multiple sources.

Resizing: The digital images of the drawing patterns may have different sizes and resolutions. Resizing the images to a standard size and resolution can reduce variability in the input data and improve the performance of the CNN [15]. All the images in the dataset are resized to 640x640 pixels.

Auto-Adjust Contrast: According to the images collected for the study, the images lack contrast, brightness, and clarity overall. All images, therefore receive contrast enhancement and adjustments using adaptive equalization.

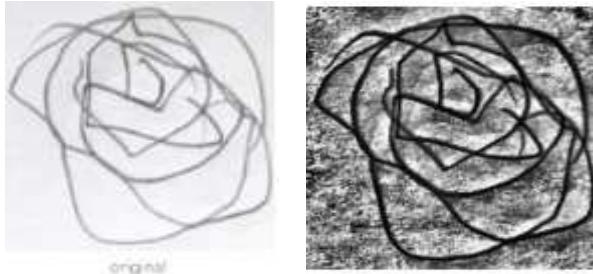


Fig.2. Auto Adjust Contrast

Grayscale: Grayscale in pre-processing typically refers to converting an image or video from a colour format to a grayscale format as a part of pre-processing steps. It can simplify processing and reduce computational complexity, as the resulting grayscale image or video contains only a single channel of brightness information instead of three channels of red, green, and blue color information. They can be used in various applications such as edge detection, feature extraction, and image compression[16]. By performing these image pre-processing steps, the quality of the input data can be improved, which can lead to better performance of the CNN in detecting and classifying Parkinson's disease stages from drawing patterns. Images may require different pre-processing steps depending on the specific research question and the input data characteristics.

C. Data Augmentation

Data augmentation techniques can help create new images that fill in this missing information. Since there are quite a few images in the dataset, the implementation of deep learning algorithms such as CNN becomes quite difficult. Thus, picture additions were finished in order to incorporate a few fabricated samples from the training dataset and increase the dataset's diversity[17]. The augmentation parameters applied to the dataset are

Flipping: Flip Flipping an image horizontally or vertically is a simple image augmentation technique that can help to increase the amount of data available for training a model. In horizontal flipping, the image is flipped along the vertical axis, while in vertical flipping, it is flipped along the horizontal axis.

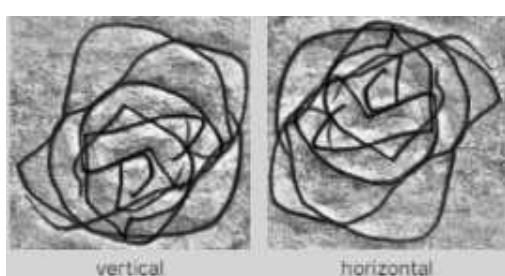


Fig.3.1. Flipping

90-degree Rotate: Rotating an image by 90 degrees can also help to increase the amount of data available for training. This can be done by rotating the image clockwise or counter clockwise or upside-down by 90 degrees.

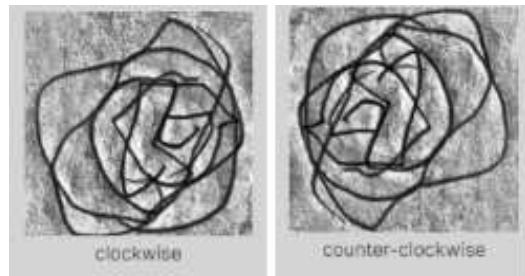


Fig.3.2. 90-degree Rotate

Rotation: Rotating an image by an arbitrary angle can help to simulate different viewpoints of an object in an image. This can be done by specifying the rotation angle and using affine transformations to apply the rotation[18]. Below the image, we have done rotation by 0 degrees, 15 degrees, and -15 degrees.

Shear: Shearing an image involves skewing it along either the x-axis or the y-axis. This can help to simulate the effect of camera angle or perspective.

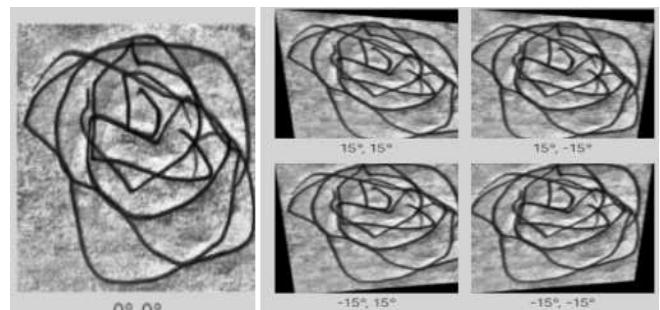


Fig.3.3 Shear

Exposure: Adjusting the exposure of an image can help to simulate different lighting conditions, such as a bright sunny day or a dimly lit room. This can be done by changing the overall brightness and contrast of the image.

Brightness: Adjusting the brightness of an image can help to simulate different lighting conditions. It can be done by increasing or decreasing the brightness of the image pixels. In our dataset, we have increased the brightness of the images up to 25%.

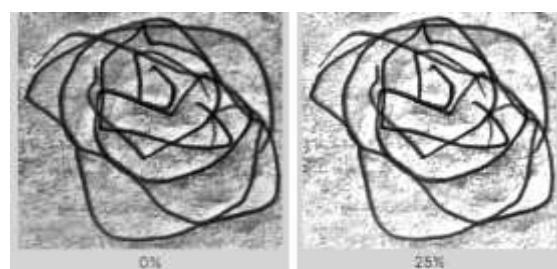


Fig.3.4. Brightness

Noise: Adding noise to an image can help to simulate the effect of random noise in the image sensor. This can be done by adding random values drawn from a Gaussian distribution to the pixel values of the image. 25% noise was added to the images in the dataset for the augmentation process.

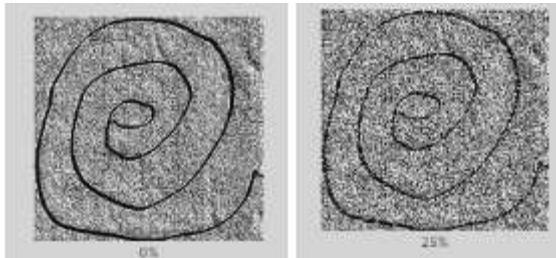


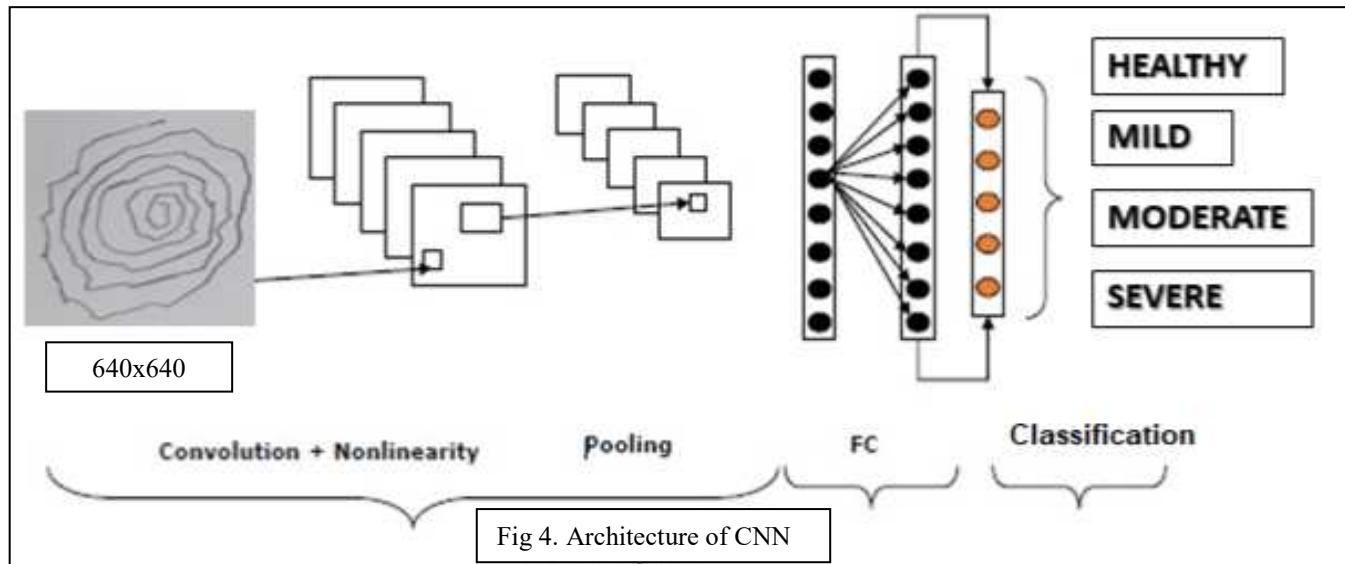
Fig.3.5. Noise

D. MODEL DEVELOPMENT

With the latest developments in the field of device intelligence, deep learning has plotted a paradigm shift in all sectors. Mainly, within the healthcare sector, the emergence of deep analysis and advanced computer detection systems required even more accurate and green analysis. As far as Parkinson's disorder is concerned, more than one biomarker needs to be analysed to determine a patient's medical condition. One of the impairments commonly visible in Parkinson's disease is hand tremors and impaired motor flexibility of the arms, leading to reduced skills in comic strips and writing. The ability to write and sketch can be

system to analyse spiral and wave patterns in patients with Parkinson's disease and to detect Parkinson's disease. The architecture of a convolutional neural network (CNN) for Parkinson's disease detection and classification of stages from drawing patterns typically consists of several layers, including convolutional layers, pooling layers, and fully connected layers[19]. Input Layer: CNN's input layer takes the drawing patterns as input. The size of the input layer is determined by the size and resolution of the images.

1. **Convolutional Layers:** The convolutional layers apply a set of learnable filters to the input image to extract features at different spatial scales. Each filter corresponds to a specific feature, such as edges or textures, and can be optimized during training to improve the performance of the CNN.
2. **Pooling Layers:** The pooling layers downsample the feature maps produced by the convolutional layers to reduce the computational complexity of the CNN. Common
3. **Fully Connected Layers:** The fully connected layers take the output of the last pooling layer as input and perform the final classification task. They typically consist of several dense layers with nonlinear activation functions, such as ReLU or sigmoid.
4. **Dropout Layer:** The dropout layer is a regularization technique used to prevent overfitting in the CNN. It randomly drops out a fraction of the neurons in the fully connected layers during training.
5. **Output Layer:** The output layer of the CNN produces the classification result, indicating



considered an essential biomarker for detecting and diagnosing Parkinson's disease. This study develops a

whether the input image belongs to a Parkinson's disease patient or a healthy control, as well as the stage of Parkinson's disease for each patient.

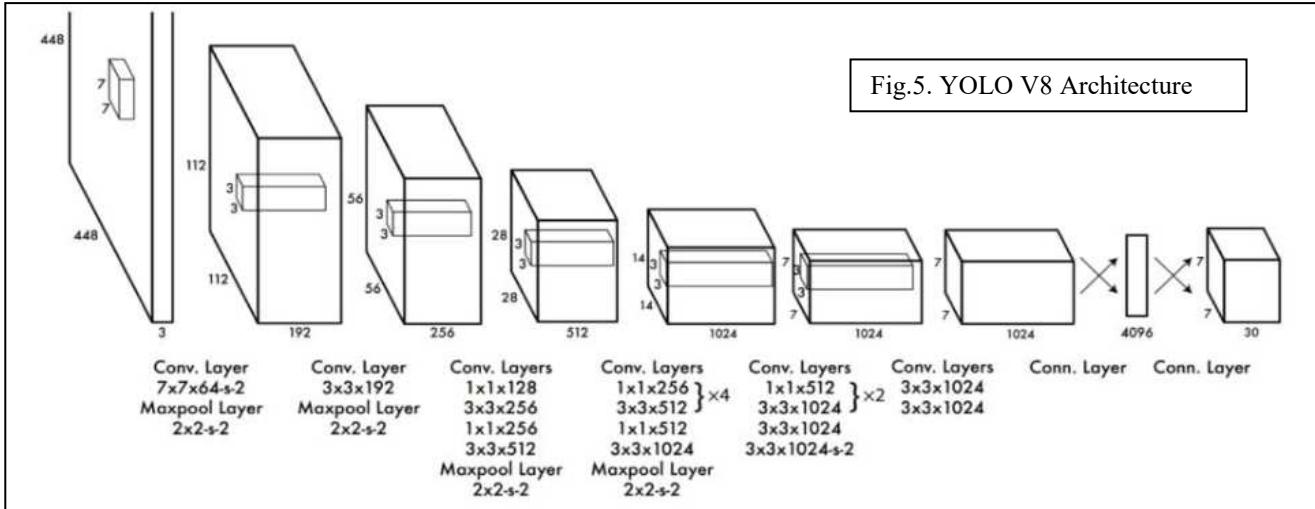


Fig.5. YOLO V8 Architecture

The output obtained by the CNN model exhibited low performance as well as low accuracy. We switched from the basic CNN model to the YOLO V8 classification model to increase accuracy. In the world of computer vision, You Only Look Once (YOLO) has become the most popular modeling architecture that has dominated the field of machine learning[20]. Its object detection algorithm is known for its fast image processing results and good accuracy.

There is a goal to the YOLO algorithm in that it aims to determine the class of an object in the input image, as well as its bounding box in which it is located in the input image. Built on PyTorch, it supports both CPU and GPU. YOLO v8 is superior to YOLO v5,6,7. The extensibility of YOLOv8 is a significant characteristic. The framework was created to work with all prior YOLO iterations, making it easy to replace them and examine their performance. Because of this, YOLOv8 is an ideal option for those who want to benefit from the most recent YOLO generation and at the same time wish to benefit from the most recent generation of YOLO. As a result, they can keep their YOLO models as practical as possible in their modern-day society. The detection network has 24 convolutional layers followed by 2 fully connected layers. Alternating 1 x 1 convolutional layers reduces the feature space from preceding layers. The convolutional layers are trained on the ImageNet classification task at half the resolution (224 x 224 input image) and then double the resolution for detection. Given the dataset without any augmentation, CNN had a 20% accuracy rate. After dataset augmentation, CNN achieved a 50% accuracy rate. When using the given dataset with augmentation, YOLOv8-x was more accurate than CNN models in terms of accuracy.

IV. RESULTS AND DISCUSSION

YOLOv8 is the latest family of YOLO-based object recognition models from Ultralytics, offering state-of-the-art performance. YOLOv8 provides already trained classification models, does the same image classification inference using the yolov8x-cls model. This is the largest

classification model offered by the archive. The activation function under the hood is ReLU except for the last layer which uses a linear activation function. Some additional techniques, such as bump normalization and removal, adjust the model and prevent it from overfitting. Reasons to choose the YOLO framework are Speed, Detection accuracy, good generalization, and Open source. The models were trained on different algorithms with different augmentation parameters and the accuracies of the model were recorded. Among the algorithms tested, we recorded a high accuracy when the dataset was trained with the YOLOv8-x algorithm.

Accuracy of the best model = 94%

Epoch	GPU mem	loss	Instances	Size
195/200	2.53G	0.04475	13	224: 100% 23/23 [00:03<00:00, 7.41it/s]
classes	top1_acc	top5_acc: 100%	1/1 [00:00<00:00, 8.92it/s]	
all	0.94		1	

Fig.6. Accuracy

TABLE I. COMPARISON TABLE

Models	Accuracy
CNN (Without Augmentation)	20%
CNN (With Augmentation)	50%
Transfer Learning (With Augmentation)	82%
YOLOv8-n (With Augmentation)	64%
YOLOv8-x (With Augmentation)	94%

When tested the generated YOLOv8x model with the sample test images using the API and the results obtained were

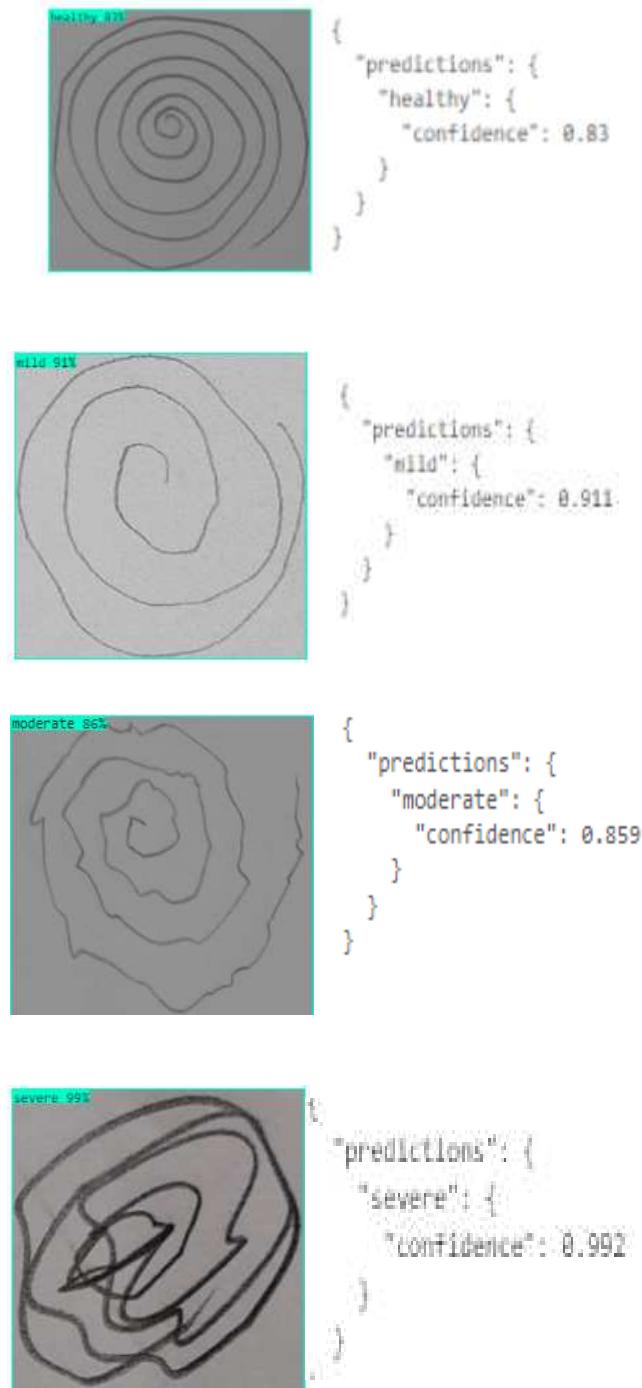


Fig.6. Prediction Results

V. CONCLUSION

As a result of using the YOLO v8 architecture, a convolutional neural network that is capable of both training and testing end-to-end and is capable of detecting and recognizing multiple targets in images, we were able to construct in this paper a multistage classification system that

successfully detects and recognizes numerous targets in images. Following the application of the augmentation parameters to the given drawing pattern image dataset, the model plotted an accuracy of 94%. In spite of the fact that the classifier's results at the moment seem to be acceptable, both the paper's technique and the classifier's performance can be improved in the future. After several iterations, the existing process can be improved by increasing the number of data samples, using different kinds of drawings instead of spirals, and selecting the second section as state-of-the-art architecture. The current system however gave us a great deal of confidence that this kind of technology could be used in real-life situations and in reliable production environments based on its performance in the current system.

VI. REFERENCES

- [1] Joseph Jankovic.,2015.Book Parkinson's disease and movement disorder .
- [2] Jankovic, J. Parkinson's disease: Clinical features and diagnosis. *J Neurol. Neurosurg. Psychiatry* 2008, 79, 368–376.
- [3] Contreras-Vidal, J.L.; Stelmach, G.E. Effects of parkinsonism on motor control. *Life Sci.* 1995, 58, 165–176
- [4] Poluhá, P., Teulings, H.L. and Brookshire, R., 1998. Handwriting and speech changes across the levodopa cycle in Parkinson's disease. *Acta psychological*, 100(1-2), pp.71-84.
- [5] Saunders-Pullman, R., Derby, C., Stanley, K., Floyd, A., Bressman, S., Lipton, R.B., Deligtisch, A., Severt, L., Yu, Q., Kurtis, M. and Pullman, S.L., 2008. Validity of spiral analysis in early Parkinson's disease. *Movement disorders: official journal of the Movement Disorder Society*, 23(4), pp.531-537
- [6] Aich, S., Pradhan, P., Park, J., Sethi, N., Vaths, V. and Kim, H.C., 2018. A validation study of freezing of gait (FoG) detection and machine learning-based FoG prediction using estimated gait characteristics with a wearable accelerometer. *Sensors*, 18(10), p.3287
- [7] Kotsavasiloglou, C., Kostikis, N., Hristu-Varsakelis, D. and Arnaoutoglou, M., 2017. Machine learning-based classification of simple drawing movements in Parkinson's disease. *Biomedical Signal Processing and Control*, 31, pp.174-180.
- [8] Memedi, M., Sadikov, A., Groznik, V., Žabkar, J., Možina, M., Bergquist, F., Johansson, A., Haubenberger, D. and Nyholm, D., 2015. Automatic spiral analysis for objective assessment of motor symptoms in Parkinson's disease. *Sensors*, 15(9), pp.23727-23744
- [9] Aich, S., Sain, M., Park, J., Choi, K.W. and Kim, H.C., 2017, November. A mixed classification approach for the prediction of Parkinson's disease
- [10] C. R. Pereira et al., "A step towards the automated diagnosis of Parkinson's disease: Analyzing handwriting movements," *IEEE 28th Int. Symp. Comput.-Based Med. Syst.*, 2015, pp. 171–176.
- [11] Gallicchio, C.; Micheli, A.; Pedrelli, L. Deep Echo State Networks for Diagnosis of Parkinson's Disease. *arXiv*2018, arXiv:1802.06708.
- [12] Sabyasachi Chakraborty.,2020. Parkinson's Disease Detection from Spiral and Wave Drawings using Convolutional Neural Networks: A Multistage Classifier Approach.
- [13] Impedovo, D.; Pirlo, G. Chapter 7: Online Handwriting Analysis for the Assessment of Alzheimer's Disease and Parkinson's Disease: Overview and Experimental Investigation. In Series on Language Processing, Pattern Recognition, and Intelligent Systems. *Pattern Recognition and Artificial Intelligence*; World Scientific Publishing: Singapore, 2019; pp. 113–128.
- [14] Gallicchio, C.; Micheli, A.; Pedrelli, L. Deep Echo State Networks for Diagnosis of Parkinson's Disease. *arXiv* 2018, arXiv:1802.06708
- [15] O'Neil King, R. Biometrics in Healthcare; Biometrics Research Group, Inc.: Toronto, ON, Canada, 2017;
- [16] M. A. E. Van Stiphout, J. Marinus, J. J. Van Hilten, F. Lobbezoo, and C. De Baat, "Oral health of Parkinson's disease patients: a case-

- control study," Parkinson's Disease, vol. 2018, Article ID 9315285, 8 pages, 2018.
- [17] B. M. Bot, C. Suver, E. C. Neto, M. Kellen, A. Klein, C. Bare, M. Doerr, A. Pratap, J. Wilbanks, E. R. Dorsey et al., "The empower study, parkinson disease mobile data collected using researchkit," Scientific data, vol. 3, p. 160011, 2016.
- [18] Zham, P., Kumar, D.K., Dabnichki, P., Raghav, S. and Keloth, S.M., 2016. Dynamic handwriting analysis for assessing movement disorder. In 13th International Conference of Applied Computing.
- [19] Ravina B, Putt M, Siderowf A et al. (2005) Donepezil for dementia in Parkinson's disease: a randomised, double blind, placebo controlled, crossover study. *J. Neurol Neurosurg Psychiatry*, 76, 934–939.
- [20] Saman Sarraf, Danielle D. DeSouza, John Anderson, Ghassem Tofighi, DeepAD: Alzheimer's Disease Classification via Deep Convolutional Neural Networks using MRI and fMRI, Cold Spring Harbor Laboratory Press.

Convolutional Neural Networks for Image Emotion Recognition by fusing Differential and Supplementary Information

Shashank Rapolu

Department of Mechanical Engineering
Indian Institute of Technology Kanpur
Kanpur, India
ssrapolu20@iitk.ac.in

Aman Singh

Department of Electrical Engineering
Indian Institute of Technology Kanpur
Kanpur, India
asingh20@iitk.ac.in

Ayush Dhingra

Department of Mechanical Engineering
Indian Institute of Technology Kanpur
Kanpur, India
ayushd20@iitk.ac.in

Ajaz Hussain

Department of Mechanical Engineering
Indian Institute of Technology Kanpur
Kanpur, India
ajazh20@iitk.ac.in

Tushar Sandhan

Department of Electrical Engineering
Indian Institute of Technology Kanpur
Kanpur, India
sandhan@iitk.ac.in

Abstract—Emotions arise out of complex phenomena, which are believed to have a biological basis. Neuroscience research has demonstrated that emotions are related to distinct patterns of brain activity and the release of certain chemicals, such as hormones and neurotransmitters, into the bloodstream. Emotion recognition is used in many applications, such as advertising, social networking and cinema. In this paper, we propose a deep convolutional neural network (CNN) fusion technique made up of two parts: a differential-CNN system to extract emotional features from an image and combine them using convex-combination, and a supplementary-CNN to extract central object details of an image and combine it with differential-CNN features. We amplify the difference between minute latent representations of an image via differential-CNN features. The type of emotion an image elicits is highly influenced by its central object, hence supplementary-CNN is helpful in supplementing the differential-CNN features to improve emotion recognition. In comparison to the contemporary state-of-the-art methods, our proposed method showed an absolute gain of 6% on the primary dataset and outperformed them on the extra secondary datasets social media web scrapping process.

Index Terms—Deep learning, image emotion, latent information, convolutional neural networks

I. INTRODUCTION

Emotions are strongly related to the activation of particular brain regions and the release of specific hormones and neurotransmitters into the bloodstream. For instance, the amygdala, a little almond-shaped structure in the brain, initiates the release of adrenaline and other stress chemicals that prepare the body to respond to a perceived threat when we experience fear. Research on emotion analysis has grown significantly in the past two decades. Previous attempts to solve this challenge used handcrafted methods based on psychological ideas [1] and image perception. Color, edges, lines, texture, composition and picture descriptors like Histogram of Oriented Gradients

(HOG) were handcrafted. Balance, emphasis and harmony were linked to the emotions elicited from an image. Lu et al. [2] attempted to retrieve shape features and classify the emotions of an image. Yanulevskaya et al. [3] retrieved Gabor and Wicest surface texture features from images and mapped them to emotions. The type of line influences the emotional response as well. In the image, oblique, horizontal and vertical lines had distinct effects on the emotional response. Chang et al. [4] extracted edges, ridges and lines to classify images. Zhao et al. [5] extracted features using image descriptors (such as Histogram of Oriented Gradients) and combined them with other handcrafted features. Both Mehrabian and Valdez [6] attempted to map color primitives to emotions and classify images.

Following the introduction of convolutional neural networks (CNNs) [7] and their rising prevalence in ubiquitous applications, the emphasis switched from handcrafted methods to neural networks. Based on the emotion categories inspired by Machadjik and Hanbury [8], You et al. [9] created a large-scale dataset for visual emotion identification. The benchmark results were produced with AlexNet [10]. Xu et al. [11] suggested a CNN-based framework for visual sentiment prediction. Considering the emotional subjectivity of an image, Yang et al. [12] provided a framework for retrieving and classifying affective images. Campos et al. and Jou et al. [13] employed CNNs to predict visual sentiment by visualizing the image's local patterns. Chen et al. [14] presented a method for classifying images based on visual sentiment concepts using deep CNN architectures.

In our work, we employed a differential-CNN system to extract emotional elements from images and combine them with supplementary information to predict emotions. We also developed a novel idea of convex combination of features,

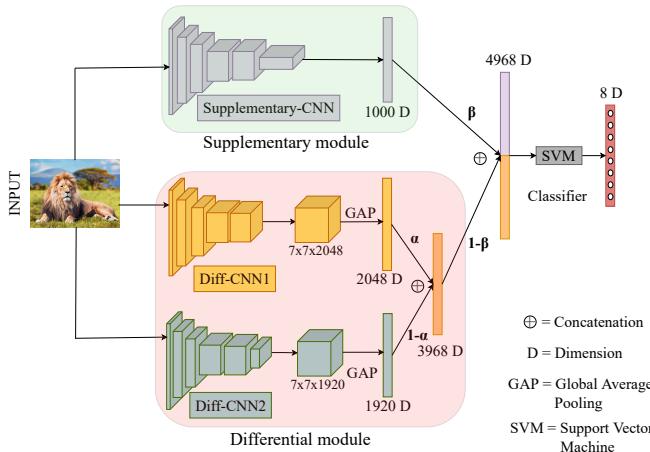


Fig. 1: Our final proposed network consists of a differential module containing a dual-CNN network for emotional-feature extraction and a supplementary module for image's central object details. Parameters α and β are differential and amplification factors, respectively. Convex combination of differential features and supplementary-feature amplification are performed separately and are further fine-tuned using SVM classifier with the CNN layers frozen.

which uses emotional data extracted from the differential-CNN system and predicts using the combined features. For training and evaluation, we used a large-scale dataset classified into eight emotions as well as additional small-scale datasets only for evaluation. As previously demonstrated [15], the emotion evoked by an image depends on the central object present in that image. As a result, we had provided the central object information using supplementary-CNN along with emotional aspects (differential-CNN features) for emotion prediction. Not only does the central object has an influence, but our experiments demonstrate that amplification of it abruptly alters emotion detection from an image.

II. OUR METHOD

Different CNNs extract and highlight different features from the same image. Our model consists of a differential-CNN module containing two deep CNNs (dual-CNN network) for emotional-feature extraction and a supplementary-CNN module containing a single CNN for primary object details present in the image. The intuition behind employing a dual CNN network in differential-CNN structure for feature extraction is that by merging the features extracted from the two CNNs, the emotional elements from an image ignored or suppressed by one CNN shall be recognized by the other and thus enhance overall emotion recognition. As finding the right combination is challenging, we have tested with several already available CNN models. It is observed from the experiments conducted by us that deep CNNs are superior at emotion recognition than shallow CNNs and hence we have used deep CNNs for differential feature extraction and for final combination.

TABLE I: Number of images present in each emotion category of various datasets evaluated in this paper. The abbreviations IE, E-ROI, and AP stand for Image Emotion, EmotionROI and AbstractPaintings, respectively and ‘-’ denotes absence of that specific emotion category.

Emotion	IE [9]	E-ROI [12]	Artpphoto [8]	AP	Ours
Amusement	4724	-	101	33	-
Anger	1176	330	77	9	-
Awe	2883	330	102	28	-
Contentment	5131	-	70	72	407
Disgust	1591	330	70	29	471
Excitement	2727	-	105	36	430
Fear	969	330	115	41	418
Sadness	2635	330	166	32	420
Total	21836	1650	806	280	2146

A. Differential CNNs for amplifying incremental latent information

First, we train Diff-CNN1 and Diff-CNN2 separately on the Image Emotion dataset. In this process, the feature matrix extracted from Diff-CNN1 is passed through GlobalAveragePooling (GAP) layer obtaining a vector representing the corresponding emotional features (differential-CNN features). These features extracted from the GAP layer are then passed to a dense layer containing eight nodes based on eight different emotions. The output from this last layer is transformed into a probability distribution using the softmax function. Similar training is implemented for Diff-CNN2, where a different feature matrix is obtained, eliciting a different vector after passing through the GAP layer. These two differential feature vectors are concatenated before passing them through the final 8-node classifier.

Not only does the combination of features extracted from differential-CNN module has an effect, but the proportion of this combination also significantly affects the performance. With this idea, we have introduced a concept of convex-combination of features extracted from the differential module, governed by a parameter α called differential factor:

$$X = \alpha X_1 \oplus (1 - \alpha) X_2 \quad (1)$$

where X_1 and X_2 are feature vectors of Diff-CNN1 and Diff-CNN2 respectively, X is a concatenated feature vector with a differential factor α and \oplus represents concatenation.

$0 \leq \alpha < 0.5$	X_1 is suppressed, X_2 is highlighted
$\alpha = 0.5$	X_1 and X_2 are equally considered
$0.5 < \alpha \leq 1$	X_1 is highlighted, X_2 is suppressed

B. Supplementary feature fusion and amplification

Furthermore, the differential-CNN system is supplemented with additional information containing central object details. As elucidated in Section I, a 1000-D vector is concatenated with the differential-CNN feature vector, acting as supplementary information. Another pre-trained CNN initialized with imagenet weights producing a probability distribution of 1000 different objects is used, called as Supplementary-CNN. Similar to convex-combination, amplifying the image's central ob-

TABLE II: Performance of our methods (before fine-tuning) on all the datasets used for evaluation. The parameter $\beta=1$ is irrelevant for the supplementary-feature amplification model since emotional qualities are ignored, and ‘-’ reflects it. Abbreviations IE, E-ROI, and AP stand for Image Emotion, EmotionROI, and AbstractPaintings and the best results obtained are represented in **bold**. It can be observed that accuracy increases with increase in parameter α till 0.75 and then decreases and considerably varies with the parameter β on the IE (primary) dataset.

Parameter	Image emotion recognition accuracy (%)									
	Differential convex combination model (α)					Supplementary-feature amplification model (β)				
	IE	AP	ArtPhoto	E-ROI	Ours	IE	AP	ArtPhoto	E-ROI	Ours
0	64.85	19	44	21	64	66.71	26	45	23	71
0.25	65.76	35	14	21	68	65.27	26	44	21	71
0.5	66.62	32	43	62	67	66.65	26	45	21	71
0.75	66.80	34	44	61	70	65.76	26	45	22	70
1	66.28	33	43	60	66	-	-	-	-	-

ject information might affect how emotions are perceived. The supplementary-feature amplification, representing the central object w.r.t the combined emotional aspects (from differential module), is controlled by an amplification factor, denoted by the parameter β . It should be noted that supplementary-feature amplification is performed for direct concatenation of differential features i.e. without their convex combination. A higher value of β denotes higher prominence of central object present in that image, whereas lower β denotes its lesser prominence.

C. Fine-tuning with SVM classifier

To further enhance emotion recognition of our models, fine-tuning with a Support Vector Machine (SVM) classifier for different values of regularization parameter (C) is implemented. The features extracted from the differential-CNN and supplementary-CNN modules are contained in the concatenated vector, which is extracted and passed through the SVM classifier with rbf kernel. Only the SVM classifier is trained at this stage after all preceding CNN layers are frozen.

III. EXPERIMENTS

Experiments were conducted on various publicly available datasets, implemented using specific training parameters and compared with the baselines as mentioned below.

A. Datasets

In our work, both large-scale and small-scale datasets are taken into account. The primary dataset is the Image Emotion dataset [9], which consists of 21,836 images distributed disproportionately into 8 emotion categories, details of which are mentioned in Table I. Like prior studies, we randomly and uniformly split the entire dataset into 85% for training and 15% for testing. Abstract Paintings, ArtPhoto [8] and EmotionROI [12] are small-scale datasets and are only used for evaluation. We also constructed a dataset from scratch to further test our approach on a medium-scaled dataset. Initially, each of us collected images from the internet totaling up to 2500 images. The gathered images were then examined together and classified into five emotion categories. In order to scrape images from the web, we employed several keywords. As an illustration, we used cold beverages, games, and

TABLE III: Comparison of our two fine-tuned CNN variants (in **bold**) with various state-of-the-art methods.

Type	Methods	Accuracy (%)
Handcrafted	SIFT [16]	37.56
	HOG [12]	44.67
	SentiBank [14]	49.09
Deep Learning	DeepSentiBank [12]	54.10
	AlexNet (fine-tuned) [9]	56.55
	Binary Assisted [17]	61.31
	LiteEmote [18]	61.67
	Supplementary-feature amplification ($\beta=0$)	67.41
	Differential convex-combination ($\alpha=0.75$)	67.75

romance to create excitement and trash, filth, and diseases to create disgust. However, we did not include the emotion anger because it is hard to locate images that elicit anger. We omitted awe and amusement as image visuals overlap with contentment and excitement. To test our final model, we gathered 2146 images based on the remaining categories in this manner. Details of all the datasets are given in Table I.

B. Implementation Details

The primary dataset used in our work is split into 85% for training and 15% for testing while preserving this ratio for each emotion class present in the dataset. ResNet101 and DenseNet201 are chosen for Diff-CNN1 and Diff-CNN2, respectively and ResNet101 is used as supplementary-CNN. A 224x224x3 pixel image is passed through the convolutional layers. Images were rescaled by a factor of 1/255 and a batch size of 64 was used for training and evaluation. Before passing the input data to the convolutional layers, we preprocessed it with CNN-specific preprocessors. The optimizer used was SGD with a learning rate of 0.001 and categorical cross-entropy as a loss function. All models were trained until the accuracy of the test dataset increased while test loss fell drastically and then increased marginally towards the end.

C. Baselines

We contrasted our model against both handcrafted and deep learning approaches. The handcrafted methods (mentioned in Table III) extracted low-level features like HOG and mid-level features like SentiBank from images and performed predictions on emotions. In handcrafted methods we compared with SIFT [16], HOG [12] and SentiBank [14]. For deep

	Amusement	Anger	Awe	Contentment	Disgust	Excitement	Fear	Sadness	Amusement	Anger	Awe	Contentment	Disgust	Excitement	Fear	Sadness	Amusement	Anger	Awe	Contentment	Disgust	Excitement	Fear	Sadness	Amusement	Anger	Awe	Contentment	Disgust	Excitement	Fear	Sadness
Actual	0.8	0.01	0.03	0.07	0	0.08	0	0.01	0.79	0.01	0.03	0.06	0.01	0.09	0	0.01	0.81	0.01	0.03	0.05	0	0.08	0.01	0.01	0.8	0.01	0.03	0.05	0.01	0.09	0	0.01
	Amusement	Anger	Awe	Contentment	Disgust	Excitement	Fear	Sadness	Amusement	Anger	Awe	Contentment	Disgust	Excitement	Fear	Sadness	Amusement	Anger	Awe	Contentment	Disgust	Excitement	Fear	Sadness	Amusement	Anger	Awe	Contentment	Disgust	Excitement	Fear	Sadness

Fig. 2: Confusion matrices to evaluate performance on each emotion category of the Image Emotion dataset. The two confusion matrices (from left) are for two Diff-CNN models trained individually. The third confusion matrix is for the best differential convex-combination model ($\alpha=0.75$) and the final is for the best ($\beta=0$) supplementary-feature amplification model. It can be observed that emotion prediction is enhanced by the final models compared to differential CNNs for the exact same dataset.

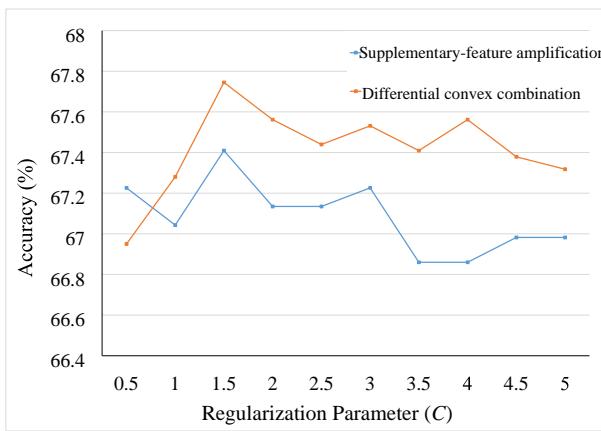


Fig. 3: Accuracy (vs) regularization parameter C of our two best-performing final models on the Image Emotion dataset.

learning methods, we compared with DeepSentiBank [12], which classifies emotions using adjective-noun pairs, a fine-tuned Alexnet model [9] trained on the Image Emotion dataset, the Binary Assisted Classification Network [17], which recognizes emotions using a single CNN and binary classification of emotions, and the LiteEmote model [18], which combines object- and category-specific features with emotional features.

D. Results on the Image Emotion (IE) dataset

Experiments were carried out on the test dataset comprised of 15% of the images from the Image Emotion dataset. Diff-CNN1 and Diff-CNN2, trained individually on Image Emotion dataset, scored 64.14% and 65.06%, respectively. Our fine-tuned differential convex-combination model ($\alpha=0.75$) achieves the best accuracy of 67.75%. The differential convex combination in Table II demonstrates how performance is significantly impacted by the differential factor (α). Accuracy increased from 65.06% to 66.28% for $\alpha = 1$, which only includes Diff-CNN1 with the supplementary-CNN. Accuracy

increased from 64.14% to 64.85% when $\alpha = 0$, which only applies to Diff-CNN2 with the supplementary data. Furthermore, accuracy increased to 66.62% when $\alpha = 0.5$, denoting that features from the differential-CNN system are given equal importance. This proves that combining two different CNN characteristics will mask any CNN features that are missing from one CNN. The supplementary-feature amplification model's performance increased by 1.38% when β value increased by 25% from $\beta=0.25$ to $\beta=0.5$, demonstrating that amplification of the image's central object has a substantial impact on the model's performance. Performance was further improved by fine-tuning the two aforementioned models with the SVM classifier. Accuracy increased from 66.71% to 67.41% for supplementary-feature amplification model with $\beta=0$ and from 66.80% to 67.75% for the differential convex-combination model with $\alpha=0.75$.

E. Effect of regularization parameter (C)

We also tested the effect of SVM classifier's regularization parameter C . Fig. 3 demonstrates how changes in the parameter have a sudden impact on the fine-tuning of our final two models. The two models, convex-combination with $\alpha=0.75$ and supplementary-feature amplification with $\beta=0$, respectively had the best results for $C=1.5$.

F. Results on small-scale datasets

We also evaluated our models on publicly available small-scale datasets. With α values of 0.25 and 0.5, our differential convex-combination model performed best on AbstractPaintings and EmotionROI datasets, respectively. Surprisingly, the supplementary-feature amplification model had the best accuracy on Artphoto and our proposed dataset, compared to the aforementioned model. On the Artphoto dataset, our model outperformed LiteEmote model [18] by 2%. Despite having lesser images, the model's performance on our proposed dataset, obtaining 71%, is unexpectedly better than on the Image Emotion dataset.

IV. CONCLUSION

In this study, we suggested a unique method for recognizing emotions from images. Based on the idea that a single CNN would not be able to extract all the important features from an image that are required for emotion recognition, a dual-CNN network is implemented in a differential module. Experiments conducted by us proved our hypothesis. We also proposed the convex combination of differential features and supplementary feature amplification using differential and amplification factors for observing the effect of relative importance between the differential features. Adding more CNNs to the differential module without sacrificing computing intensity might result in even better outcomes. Other emotion-dependent elements, such as low-level or background features, can be investigated further to enhance the detection of an image's emotion.

REFERENCES

- [1] D. Joshi, R. Datta, E. Fedorovskaya, Q.-T. Luong, J. Z. Wang, J. Li, and J. Luo, "Aesthetics and emotions in images," *IEEE Signal Processing Magazine*, vol. 28, no. 5, pp. 94–115, 2011.
- [2] X. Lu, P. Suryanarayanan, R. B. Adams Jr, J. Li, M. G. Newman, and J. Z. Wang, "On shape and the computability of emotions," in *Proceedings of the 20th ACM international conference on Multimedia*, pp. 229–238, 2012.
- [3] V. Yanulevskaya, J. C. van Gemert, K. Roth, A.-K. Herbold, N. Sebe, and J.-M. Geusebroek, "Emotional valence categorization using holistic image features," in *2008 15th IEEE international conference on Image Processing*, pp. 101–104, IEEE, 2008.
- [4] L. Chang, Y. Chen, F. Li, M. Sun, and C. Yang, "Affective image classification using multi-scale emotion factorization features," in *2016 International Conference on Virtual Reality and Visualization (ICVRV)*, pp. 170–174, IEEE, 2016.
- [5] S. Zhao, Y. Gao, X. Jiang, H. Yao, T.-S. Chua, and X. Sun, "Exploring principles-of-art features for image emotion recognition," in *Proceedings of the 22nd ACM international conference on Multimedia*, pp. 47–56, 2014.
- [6] P. Valdez and A. Mehrabian, "Effects of color on emotions.,," *Journal of experimental psychology: General*, vol. 123, no. 4, p. 394, 1994.
- [7] 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, 1998.
- [8] J. Machajdik and A. Hanbury, "Affective image classification using features inspired by psychology and art theory," in *Proceedings of the 18th ACM international conference on Multimedia*, pp. 83–92, 2010.
- [9] Q. You, J. Luo, H. Jin, and J. Yang, "Building a large scale dataset for image emotion recognition: The fine print and the benchmark," in *Proceedings of the AAAI conference on artificial intelligence*, vol. 30, 2016.
- [10] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Communications of the ACM*, vol. 60, no. 6, pp. 84–90, 2017.
- [11] C. Xu, S. Cetintas, K. Lee, and L. Li, "Visual sentiment prediction with deep convolutional neural networks (2014)," *arXiv preprint arXiv:1411.5731*.
- [12] J. Yang, D. She, Y.-K. Lai, and M.-H. Yang, "Retrieving and classifying affective images via deep metric learning," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 32, 2018.
- [13] V. Campos, B. Jou, and X. Giro-i Nieto, "From pixels to sentiment: Fine-tuning cnns for visual sentiment prediction," *Image and Vision Computing*, vol. 65, pp. 15–22, 2017.
- [14] T. Chen, D. Borth, T. Darrell, and S.-F. Chang, "Deepsentibank: Visual sentiment concept classification with deep convolutional neural networks," *arXiv preprint arXiv:1410.8586*, 2014.
- [15] H.-R. Kim, Y.-S. Kim, S. J. Kim, and I.-K. Lee, "Building emotional machines: Recognizing image emotions through deep neural networks," *IEEE Transactions on Multimedia*, vol. 20, no. 11, pp. 2980–2992, 2018.
- [16] T. Rao, M. Xu, H. Liu, J. Wang, and I. Burnett, "Multi-scale blocks based image emotion classification using multiple instance learning," in *2016 IEEE International Conference on Image Processing (ICIP)*, pp. 634–638, IEEE, 2016.
- [17] X. He and W. Zhang, "Emotion recognition by assisted learning with convolutional neural networks," *Neurocomputing*, vol. 291, pp. 187–194, 2018.
- [18] Y.-H. Chew, L.-K. Wong, J. See, H.-Q. Khor, and B. Abivishaq, "Liteemo: lightweight deep neural networks for image emotion recognition," in *2019 IEEE 21st International Workshop on Multimedia Signal Processing (MMSP)*, pp. 1–6, IEEE, 2019.

A novel design for automatic measurement of reaction time for audiovisual and muscular stimulus

1st Keerthika.N

*School of Electronics Engineering
Vellore Institute of Technology
Chennai, India
keerthika.n2021a@vitstudent.ac.in*

2nd E.Sathish

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

3rd V. Kiruthika

*School of Electronics Engineering
Vellore Institute of Technology
Chennai, India
kiruthika.v@vit.ac.in*

4th Santhakumar.M

*School of Electronics Engineering
Vellore Institute of Technology
Chennai, India
santhakumar.2020@vitstudent.ac.in*

Abstract—Reaction Time (RT) is crucial for detecting cognitive abilities in sports and clinical applications. RT Measurements can be used to evaluate the performance and sensory-motor integration of individuals. It determines a person's attentiveness because RT indicates how rapidly an individual reacts toward a stimulus. A novel experimental setup called Automatic Reaction Time Tester (ARTT) system is proposed in this study to measure the RT using Audio Stimulus (AS), Visual Stimulus (VS), and Muscular Reaction Time (MRT). The ARTT system helps in reducing human intervention and time consumption. It improves accuracy and makes it easier to test the RT in terms of AS, VS, and MRT in a single system. In the sports field, coaches are able to analyze the current condition of the players and modified their training sessions accordingly. Moreover, the individual player can also check their performance through self-diagnosis methods for improving their performance. In the medical field, it assists clinicians in determining a patient's response to medication and facilitates a speedy recovery through this RT test.

Index Terms—Reaction Time, Audio Stimulus (AS), Visual Stimulus (VS), Muscular Reaction Time (MRT), Sports

I. INTRODUCTION

The human body reacts to many environmental actions and there is a need to react to the actions immediately. So, Reactions are essential to human life and the necessity to measure the instantaneous response is referred to as Reaction Time (RT). It is defined as the time elapsed between stimulus and response. The various types of reaction times are Simple Reaction Time (SRT), Recognition Reaction Time (RRT), and Choice Reaction Time (CRT). SRT is the minimal time needed to respond to a stimulus. RRT is the time duration of a specific stimulus and response including the presence of multiple false stimuli.CRT means a person requires time to distinguish between two or more stimuli and there may be a requirement to select one or more responses. RT can be tested in a non-invasive way. It is used to evaluate the functionality of both peripheral and central nervous system structures. The two categories of RT are Visual Stimulus (VS) and Auditory Stimulus (AS). RT is one of the methods to find how efficiently the central nervous system is able to process

information. RT is applied for analyzing color blindness, the coordination between body and mind for paralyzed patients, performing fitness tests for sports individuals' and hearing tests for analyzing the intensity of deafness. [1]

II. RELATED STUDY

Humans are able to perceive the conditions in the outside environment and react accordingly with the help of their sensory nervous systems and infer sensory information. RT is reliable, it indicates the frequencies at which sensory information is processed and converted into motor actions. [2] Fig.1 shows the biological representation of RT.

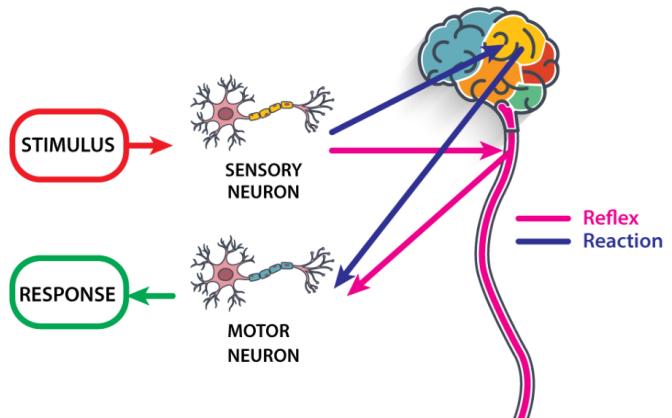


Fig. 1. Biological Representation of RT [3]

A broad literature survey was carried out to analyze the significance of RT in sports. RT is crucial in the physical fitness profile for evaluating the AS and VS of the individuals. This would be helpful for individuals to improve their performance and evaluate their current condition in their sports careers. [4] RT is an important aspect in indoor and outdoor games, namely basketball, volleyball, football, cricket,

athletes, table tennis, and gymnastics. Another study was carried out to compare Visual Reaction Time (VRT) and Visual Anticipation Time (VAT) between athletes and non-athletes ($n= 228$). VAT is measured using a Bassin anticipation timer and VRT is measured using a Lafayette reaction timer. Non-athletes made more errors and were less consistent in both VAT and VRT, whereas athletes made fewer errors and were more consistent in VAT. Moreover, the survey indicated that gender was not a barrier to participating in sports. [5] A study on the Anticipation and Reaction Time (A&RT) test was carried out for 11 volleyball players (VB) and sprinters (athletes). The A&RT test consisted of a sensory and cognitive ability test that measured the visual and auditory reaction time in eye-open and eye-closed conditions for the players. The result showed that time was identical between the groups. It was predicted that athletes' timing is better than VB players in AS, but the opposite was true for A&RT. However, the authors concluded that athletes do more sensory and cognitive actions with their eyes open and closed in their respective sports events. [6] In sports, field anticipation plays a vital role. The author analyzed the VB players' anticipation and RT towards the opponent in various situations during the entire match. If the players have good anticipation and RT, it would lead to victory for the team. Moreover, the study inferred that VB players' anticipatory prowess can be forecasted by looking back on past games and spotting a shifting pattern in the game situation easily. [7]

An experimental analysis was done for Football players (adults), who underwent 22-weeks training session. The individuals were tested by the pre- and post-training sessions with a few parameters such as endurance, flexibility, coordination, agility, and speed. In addition to that, VS and AS tests were also taken by the experimental setup. In this test, once the player completed a 22-week training session they were tested for their present condition to analyze the important skill abilities. The experimental setup contained lights and tones such as blue, red, and yellow lights and 500, 1000, and 3000 Hz tones respectively. In VS when any one color comes into the screen the person will immediately press the respective color button. Similarly, for AS when any number is heard from the surroundings, the concerned person will press the number, and finally, the RT was calculated. [8], [9]

An extensive survey was carried out to analyze the importance of RT in individuals with health abnormalities. RT is a combination of AS and VS recognition. In clinical uses, when the same stimulus was given to the same person over and again, it was useful to predict and learn about their current condition. The analysis was made to measure the RT in a simple and non-invasive way for autism patients (Adults). The Visual Attention (VA) and VS test were conducted manually on the adults. The final result showed that autistic adults had weaker RT compared with healthy adults but, there was an improvement in their behaviors as a response to the given medication.[16].The analysis was carried out to measure the RT for the adults, who were affected by Autism Spectrum Disease (ASD). The RT test was conducted in terms of Simple

Reaction Time (SRT) or Choice Reaction Time (CRT). From this analysis, the authors inferred that the ASD-affected adults had a lot of cognitive and information-processing skills, but SRT or CRT was not affected to a greater extent. [10] Several

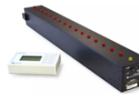
Instrument Name & Model	Image	Instrument Name & Model	Image
Bassin Anticipation Timer - Model 35575 [12]		MOART Reaction and Movement Time Panel with Psymcon Control - Model 35600 [14]	
Modified Bassin Anticipation Timer- Model 35580 [13]		Foot Switch for MOART- Model 35603 [15]	

Fig. 2. Measuring Instruments for RT [11], [12], [13], [14]

researchers have examined RT in patients with Parkinson's disease. The hunt for distinct patterns of slower RT in Parkinson's disease has yielded conflicting results. [15] The author performed a study to measure RT for head-injured victims using Ruler Drop Test (RDT) experimental setup. This analysis proved that repeated RDT tests for the patients showed improvements in their RT. [16] The detailed survey carried out indicates that RT is predominantly utilized in the medical and sports fields. Medical applications of RT include evaluating the patient's progress and determining if the patient will respond appropriately to the given treatment. In the sports field, RT is a great tool for coaches to evaluate the players' performance. The player and the coach can get an awareness of their present condition through the RT study. In the future, it will also help the trainer to modify the workouts based on this evaluation.

III. METHODS

The Ruler Drop Test (RDT) is one of the non-invasive methods to measure RT. The RDT can be performed manually using a timer or by applying the assessment variables in an equation.

A. Manual Measurement of RDT

Everyone requires a small amount of time to react to every incident in their life. For example, when watching a cricket match in a stadium, if a player strikes the ball and it rolls towards the spectators' gallery, a spectator attempts to catch it. The time elapsed between the strike and catch of the ball is referred to as RT. There are several ways to determine the RT. The Ruler Drop Test (RDT) is a standard and reliable method for testing. This can be completely accomplished manually by hand. A time period can be calculated manually in two different ways: by counting seconds directly or by applying an equation. Fig.3 shows the manual measurement of RT.

B. Manual Measurement of RDT using equations

The terms and variables in the equations are defined which include Maximum Value - K, Minimum Value - R, (Initial)

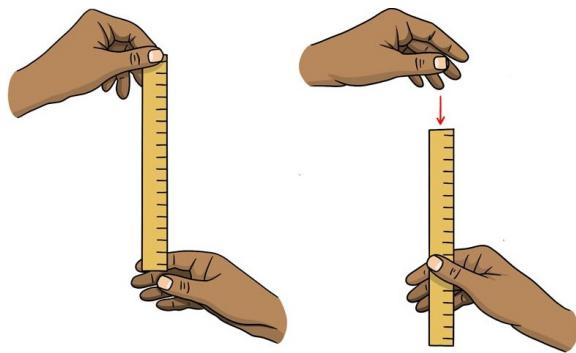


Fig. 3. Manual Measurement of RT [17]

Velocity - V1, (End) Velocity - V2, Displacement - ω , Time - t, Acceleration Average – S, Velocity Average $- \alpha$. [18]

$$\alpha = \frac{1}{2}(v_1 + v_2) \quad (1)$$

$$\omega = \frac{1}{2}(v_1 + v_2 + st)t \quad (2)$$

$$\omega = v_1 t + \frac{1}{2}st^2 \quad (3)$$

$$t = \sqrt{2} \frac{\omega}{s} \quad (4)$$

$$\text{Calculate the error} = \frac{1}{2}(K - R)^2 \quad (5)$$

(i.e) K is max value and R is min value.

Both RDT procedures discussed above were performed manually. There were few limitations while implementing the method. If the test was performed repeatedly for the same person, it gave different values for every test, resulting in a lack of accuracy with respect to time. Both methods need manual intervention to carry out the task. To overcome these limitations, a novel method called Automatic Reaction Time Tester (ARTT) system is proposed in this study to measure RT. Fig.4 shows the block diagram of ARTT system. This ARTT system will be used to measure the AS, VS, and RDT tests using a single system.

C. Automatic Reaction Time Tester system. (ARTT)

The proposed design is an automatic system for measuring the RT. The block diagram comprises of the embedded board with a timer, light for visual stimulus, and speaker for auditory stimulus. Initially, the electromagnet is coupled with a steel ruler. The instance at which the electromagnet opens up, the ruler falls and the timer will be started automatically in the embedded board. When the person catches the ruler, the time will be stopped automatically. The time elapsed between the fall and catch is obtained.

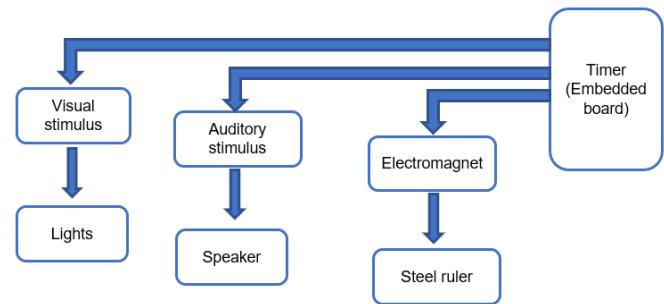


Fig. 4. ARTT system

D. Circuit Diagram of ARTT system

The AS, VS and MRT can be performed in a single ARTT system as shown in figure 5. The initial condition is needed, to fix the respective pins for input and output. The electromagnet (12V, size 20 x 15mm) is coupled with a steel ruler and the led bulbs, switches, speaker, force sensor (thickness 0.203mm, length-56.9mm, diameter-25.4mm), 5V Relay module (operating time 10 msec), and LCD display (16X2) are fixed in AT Mega board (2560). This circuit will perform AS, VS and MRT tests using a single system. When VS test is performed in the system, the circuit is in open condition initially. Once the switch is on in the circuit, a particular color LED will glow and the user will press the respective color switch immediately and RT time will be displayed on the LCD.

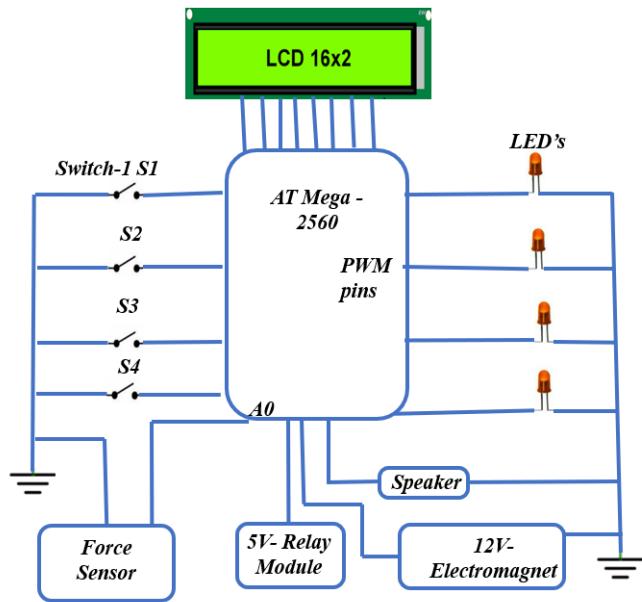


Fig. 5. ARTT circuit diagram

In VS, when the LED color is seen, the user will press the respective switch immediately. In AS, when the name of the color sound is heard, the user will press the respective color switch and the RT is displayed on the LCD. In addition to this,

the ARTT system will perform the MRT for the individuals to measure the RT. In this case, the circuit is in off condition the system holds the ruler using an electromagnet. Once the circuit is switched on, the electromagnet will drop the ruler within 3 to 8 milliseconds. When the ruler is caught by the sensor glove the RT will be displayed on the LCD immediately. Moreover, if the force is not applied properly to the sensor, the loop will run infinitely inside the system.

IV. RESULTS AND DISCUSSION

The simulation was done through the Porteus software. Fig.5 shows the result of ARTT simulation. It consists of the

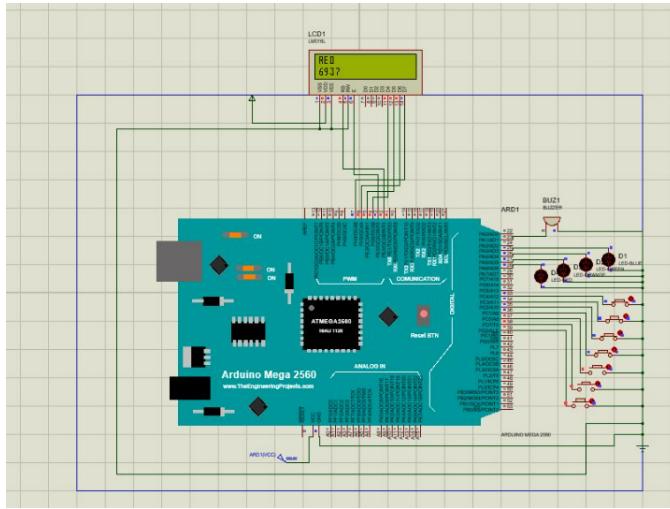


Fig. 6. Simulation of ARTT system

relay module of 5V, Arduino ATmega 2560, electromagnet DC 12V, five push buttons, speaker, four LEDs, and a 10K resistor in the design in order to construct the entire system and compute the RT in terms of VS and AS using a single system.

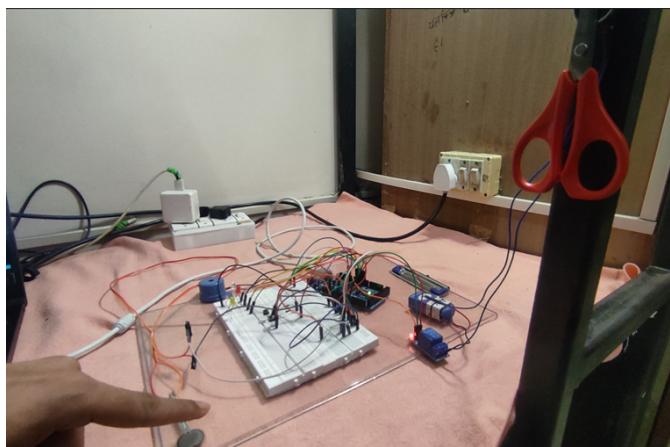


Fig. 7. Measurement of MRT in ARTT system

Fig.6 shows the experimental setup to perform the MRT. In this system, initially the circuit is in closed condition and the

electromagnet will hold the object. When the circuit is on, the object will be dropped and it will be caught by the person with force sensor glove. RT will be immediately displayed on the LCD.

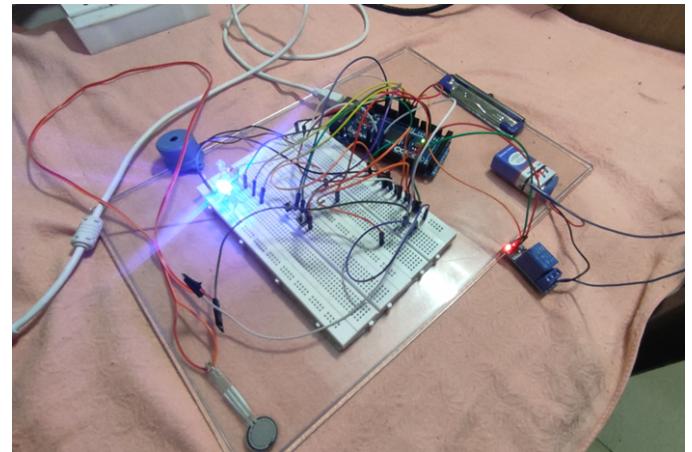


Fig. 8. Measurement of VS and AS in ARTT system

Fig 7, shows the experimental setup to perform AS and VS. In Fig 8, a battery is used for supplying the input circuitry. All the switches and LEDs are connected to Arduino AT Mega 2560 board. When the specific stimulus is received, the individual responds to the stimulus and RT can be measured and displayed on the LCD.

RT- Reaction Time	AS- Audio Stimulus
SRT- Simple Reaction Time	VRT- Visual Reaction Time
CRT- Choice Reaction Time	VAT- Visual Anticipation Time
VS- Visual Stimulus	A&RT- Anticipation & Reaction Time
RT- Reaction Time	RDT- Ruler Drop Test.
RRT- Recognition Reaction Time	ARDT- Automatic Ruler Drop Test.

Abbreviations

Table.2 shows the outcome of the RT analysis for various fields. Manual and instrument-based measurements are used for testing RT in terms of AS and VS. [4], [5], [19], [15] But the present work performed by the authors will measure the RT in terms of AS and VS using a single system. The accuracy and time taken for performing the experiment has also significantly improved. The previous works have measured only the AS, VS, and MRT by manual setup for sports and clinical applications. [8], [16] All three tests were not performed in a single system. The proposed work by the authors is to measure the AS, VS, and MRT using a single system.

Comparative Analysis of Study on RT								
The symbols represent the **- Not implement and √ - implement								
Reference	Field of Study	Measuring Instrument			RT			Outcome
		Device	Experimental Setup	Manual	VS	AS	MRT	
[2]	Sports	**	**	√	√	√	**	VS and AS for the different sports people was examined and players' pre- and post-training performance was assessed.
[7]	Sports	√	**	**	√	**	**	The VS for athletes and non-athletes was measured by using Lafayette reaction timer. Finally, VS was compared for the groups to find the RT.
[9]	Sports	**	√	**	√	√	**	The AS and VS was measured for the football players, who underwent 22 weeks training session. The individuals were tested for the RT in both pre- and post-training sessions
[16]	Medical	**	**	√	√	**	**	The VS for the autism patients and healthy subjects was measured and the RT between the groups was evaluated to analyse the response of medication in autism people.
[17]	Medical	**	**	√	√	√	**	The SRT or CRT for ASD patients was evaluated. It was inferred that ASD patients had more cognitive abilities.
[18]	Medical	**	**	√	√	√	**	RT was examined for Parkinson's patients and their important due to medication was assessed.
[20]	Medical	**	√	**	**	**	√	The RT was measured using RDT for head injured victims. Repeated RDT test showed improvement in RT.
Proposed ARTT system	Sports/Medical	**	√	**	√	√	√	For sports and clinical applications, the ARTT system measures the RT in terms of AS, VS and MRT

Table. 1. Comparative Analysis for RT

V. CONCLUSION

RT plays a vital role to sense the cognitive abilities. Various methods are attempted for calculating the RT manually. In this study an experimental setup is used to measure AS, VS and MRT to analyze RT. The novel setup reduces the human intervention, and time consumption. It provides improved accuracy and facilitates the testing of Visual Stimulus and Auditory Stimulus response in addition to Muscular Reaction Time in a single system. This system will be helpful for sports coaches to analyze the current condition of the players and give appropriate training. It will enable players to perform self-diagnosis and modify their training in order to improve their performance in tournaments. It supports physicians to understand the response of medication for the patients undergoing treatments and enables quick recovery.

REFERENCES

- [1] J. Y. Fang and T. L. Davis, "Reaction time in parkinson's disease," in *Reference Module in Neuroscience and Biobehavioural Psychology*. Elsevier, 2017.
- [2] W. Wang, L. Hu, H. Cui, X. Xie, and Y. Hu, "Spatio-temporal measures of electrophysiological correlates for behavioral multisensory enhancement during visual, auditory and somatosensory stimulation: A behavioral and erp study," *Neuroscience Bulletin*, vol. 29, no. 6, pp. 715–724, 2013.
- [3] "microgate," accessed:2020. [Online]. Available: <https://training.microgate.it/en/solutions/reaction-time>
- [4] H. B. Temur and R. Baytar, "Comparison of the reaction time period of individuals in sport, fine arts, and classroom education," *Asian Journal of Education and Training*, vol. 5, no. 3, pp. 495–500, 2019.
- [5] Y. M. Kuan, N. A. Zuhairi, F. A. Manan, V. F. Knight, and R. Omar, "Visual reaction time and visual anticipation time between athletes and non-athletes," *Malaysian Journal of Public Health Medicine*, vol. 2018, pp. 135–141, 2018.
- [6] L. Nuri, A. Shadmehr, N. Ghotbi, and B. Attarbashi Moghadam, "Reaction time and anticipatory skill of athletes in open and closed skill-dominated sport," *Eur J Sport Sci*, vol. 13, no. 5, pp. 431–436, 2013.
- [7] A. Piras, R. Lobietti, and S. Squatrito, "Response time, visual search strategy, and anticipatory skills in volleyball players," *Journal of Ophthalmology*, vol. 2014, p. 189268, 2014.
- [8] L. Ricotti, J. Rigosa, A. Niosi, and A. Menciassi, "Analysis of balance, rapidity, force and reaction times of soccer players at different levels of competition," *PloS one*, vol. 8, no. 10, p. e77264, 2013.
- [9] Y. H. Song, S.-M. Ha, J.-S. Yook, and M.-S. Ha, "Interactive improvements of visual and auditory function for enhancing performance in youth soccer players," *International Journal of Environmental Research and Public Health*, vol. 16, no. 24, p. 4909, 2019.
- [10] F. R. Ferraro, "No evidence of reaction time slowing in autism spectrum disorder," *Autism*, vol. 20, no. 1, pp. 116–122, 2016.
- [11] "Lafayette instrument," accessed:2019. [Online]. Available: <https://lafayetteevaluation.com/products/bassin-anticipation-timer>
- [12] "Lafayette instrument," accessed:2019. [Online]. Available: <https://lafayetteevaluation.com/products/modified-bassin-timer>
- [13] "Lafayette instrument," accessed:2017. [Online]. Available: <https://lafayetteevaluation.com/products/moart-panel>
- [14] "Lafayette instrument," accessed:2018. [Online]. Available: <https://lafayetteevaluation.com/products/moart-foot-switch>
- [15] E. Evarts, H. Teräväinen, and D. Calne, "Reaction time in parkinson's disease," *Brain*, vol. 104, no. 1, pp. 167–186, 1981.
- [16] G. Del Rossi, A. Malaguti, and S. Del Rossi, "Practice effects associated with repeated assessment of a clinical test of reaction time," *J Athl Train*, vol. 49, no. 3, pp. 356–359, 2014.
- [17] "maatauranga," accessed:2020. [Online]. Available: <https://maatauranga.co.nz/index9t.html>
- [18] "Ukessays.com," accessed:2020. [Online]. Available: <https://us.ukessays.com/essays/engineering/ruler-dropping-test-measure-reaction-8206.php>
- [19] P. s. Bhakare and A. Vinchurkar, "Study of auditory reaction time in severity of autism," *Indian journal of applied research*, vol. 8, 2018.

Active Feedback using Riemannian Features for Motor Imagery Classification

Vinay Gupta

*Department of Electrical Engineering
Indian Institute of Technology Kanpur
Kanpur, India
guptak@iitk.ac.in*

Rishi Mittal

*Department of Engineering Physics
Indian Institute of Technology Mandi
Mandi, India
B21261@students.iitmandi.ac.in*

Durgesh Ameta

*IKSMHA Center
Indian Institute of Technology Mandi
Mandi, India
D22117@students.iitmandi.ac.in*

Laxmidhar Behera

*Director, Senior Member, IEEE
Indian Institute of Technology Mandi
Mandi, India
director@iitmandi.ac.in*

Tushar Sandhan

*Department of Electrical Engineering
Indian Institute of Technology Kanpur
Kanpur, India
sandhan@iitk.ac.in*

Abstract—Brain-Computer Interface (BCI) technologies employing electroencephalography (EEG) signals heavily depend on effective and accurate signal classification strategies. Researchers have extensively developed various machine learning (ML) algorithms. However, very little has been done to improve the user's ability to elicit better brain patterns easily distinguishable across different stimuli. This paper proposes four feedback mechanisms through which the user trains himself for a motor imagery (MI) task via active feedback. The feedback strategies involve t-Distributed Stochastic Neighbor Embedding (t-SNE)-transformed Riemannian covariance matrices, mean correlations between C3 and C4 channels, tangent space transformed mean correlations between C3 and C4 channels and the power spectral density difference between C3 and C4 channels. Using a standard SVM classifier, the subjects showed significant improvement in MI accuracy post the training session. An increase in accuracy of more than 9% is achieved for two feedback mechanisms on an in-house dataset of 24 subjects indicating the effectiveness of proper feedback for eliciting better MI ability from users.

Index Terms—Electroencephalogram (EEG), Brain-Computer Interfaces (BCI), Motor Imagery (MI), user feedback, Riemannian manifold

I. INTRODUCTION

Brain-computer interfaces (BCIs), also known as neural interfaces, allow for direct brain-to-external device connection [1]–[3]. Electroencephalography (EEG) based non-invasive BCIs [4], [5] have shown promise for a variety of uses [3], including rehabilitation [6]. Assistive technologies (such as communication or smart wheelchair control) [7]–[9]. Mental rotation, mental calculation [10], or motor imagery [11]–[13] are some of the tasks performed in order to control BCIs. We shall refer to BCIs that are motor imagery (MI) based in this article. In MI-based BCIs, users are encouraged to visualize the limb movements rather than performing them [14].

Although promising, MI-BCIs need to be more trustworthy to be employed outside laboratories in practical applications. Significant error rates are associated with decoding users' mental commands [15], [16]. Furthermore, it is predicted that

10 to 30 percent of BCI users would be unable to use current MI-BCI apps (based on studies mostly carried out on healthy naive subjects using the currently accepted training protocols) [16]–[19]. The initial literature frequently referred to these unsatisfactory interactions as BCI insufficiency or BCI inefficacy [16], [20]–[22] however doing so implies incorrectly that users are to blame for the issue [23]. As a result, decoder/classifier techniques [24]–[26] rely on user data gathered beforehand and during the training for co-adaptive MI-BCIs [27]–[30] have found widespread use. The overwhelming research for better classification is almost saturated at a certain decoder accuracy.

The significance of user training, which aims to aid users in developing or enhancing their control over BCI, may have been dominated by immense studies devoted to decoding BCI instructions from EEG data. The effectiveness of MI-BCIs intrinsically relies on the users' capacity to effectively convert mental commands into neural impulses. In other words, it depends on the users' ability to create individual EEG patterns for each type of mental command and stable patterns for the same orders every time they want to give them. If the user is unable to provide different EEG signals, no machine learning algorithm will be able to do so. While the need for entirely stable patterns may be replaced by future decrypting algorithms, stability appears necessary for existing systems to effectively decode mental commands. It is necessary to understand the neurophysiological mechanisms behind this capacity to effectively encode mental directives using MIs and create consistent, recognizable EEG patterns, which is necessary for improving user training and, in turn, MT-BCI dependability.

Research is crucial for areas involving MI-BCI user training to comprehend the amount to which MI-BCI

users may get to know when and how to modify their MI approach in order to improve the system's ability to recognize their mental commands.

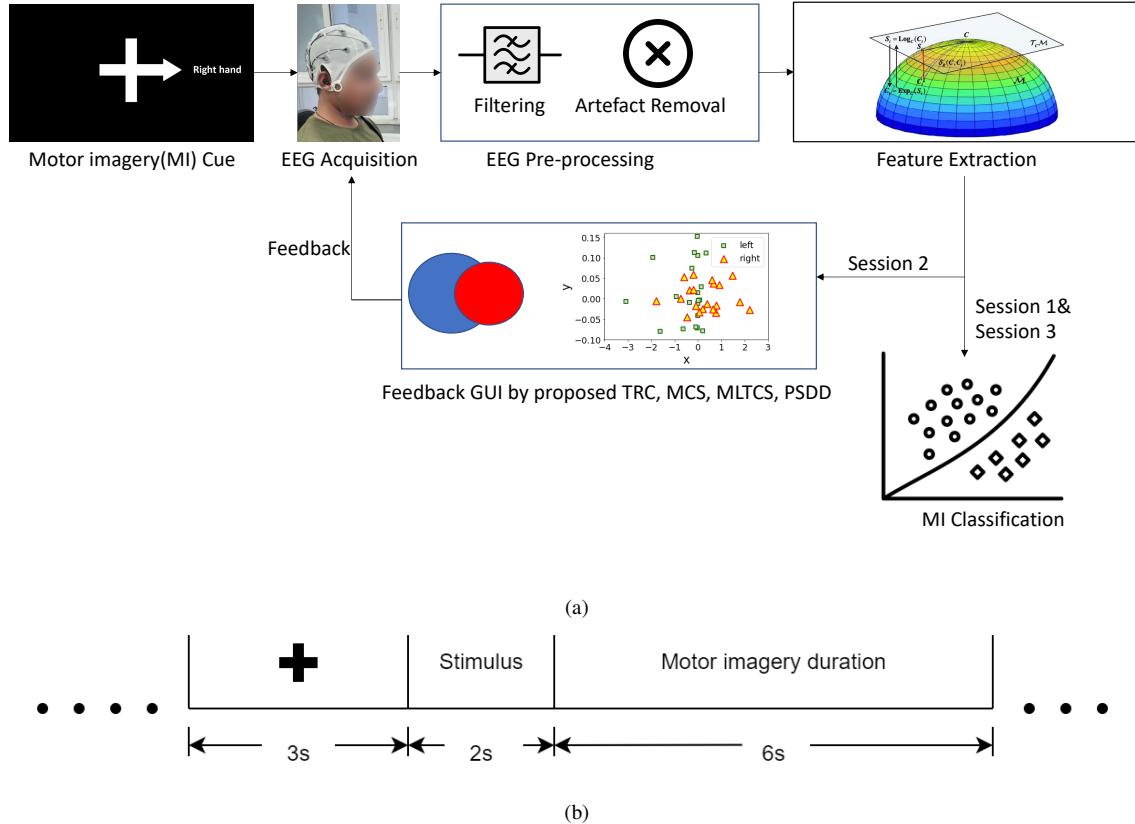


Figure 1: 1(a): Overall pipeline of the proposed method. 1(b): Timeline of single motor imagery (MI) trial. Here, the stimulus is the instruction to imagine the left/right hand.

In the user-training approaches, initially, researchers focused on acclimatizing users to a fixed system [5], [20], [31], [32] of trial-and-error operand conditioning strategy [33]. However, they might not be enough to make MI-BCIs useful in real-world applications. Users can be informed about the MT that the BCI has identified by either continuous visual feedback or discrete tactile input [34], [35]. A review of feedback mechanisms is given in Table I. One of the standard MI-BCI feedback methods comprises a swiftly increasing bar or moving cursor that demonstrates how confidently and precisely the system recognizes the task performed by the learner. Feedback provided to MI-BCI users during training is normally based on the output of the BCI classifier, which typically measures the odds of the current trial corresponding to a particular MI class [36], [37]. However, the concept of classifier output is unfamiliar to the majority of MT-BCI users. For the learner who must comprehend and evaluate feedback, it does not stand for anything tangible. Users may therefore find it challenging to comprehend the feedback and apply it in order to make required behavioral changes. In this paper, we handle the limitations of classifier reliability by proposing four novel MI-feedback approaches based on Riemannian Covariance features, their tangent space projections, and power

Table I: A review of various feedback approaches for EEG signal capture in motor imagery settings.

Reference	Feedback Mechanism
C. Zich et al. [38]	Two-dimensional feedback that included data on contra vs. ipsilateral activity and contra lateral activity during rest vs. during mental visualization
F. Cincotti et al. [34]	Vibro-tactile feedback (such as vibration on hand)
John et al. [33]	trial-and-error operand conditioning strategy
Pillette L et al. [39]	Users were given emotional feedback in between experiments via treatments that included both spoken statements and visible facial expressions.
Wolpow J R et al. [5]	Visual feedback in the form of a rapidly increasing bar or moving cursor that shows how accurately and confidently the system identifies the task carried out by the learner

spectrum density-based features. The overall pipeline is shown in Fig. 1. The paper is organized as follows: Section II gives a description of the proposed methodology, Section III discusses the results, with Section IV concludes the paper.

II. METHODOLOGY

In this section, we describe the proposed feedback methodologies. EEG data is recorded while a subject performs left

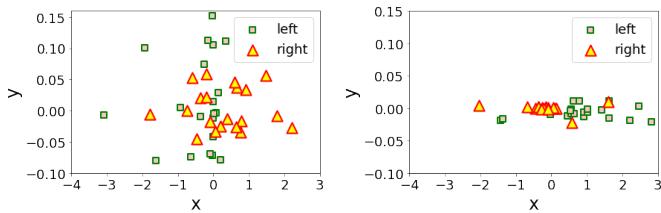


Figure 2: The Reimannian covariance matrices of MI trials, which are flattened and transformed to two dimensions using t-SNE (TRCs), are plotted in the above figures. Left Figure: Trials for S1 (Pre-training) and left and right MI are randomly distributed. Right Figure: Trials for S3(Post-training), left and right MI are clustered distinctly.

and right MI tasks as per the instruction of the experimenter. An EEG signal is a time-series signal, which is represented by a matrix C of size $(n_channels \times n_timesamples)$. Each MI task took a duration of 6 seconds.

A. Feedback approaches

1) *T-SNE transformed Riemannian-covariance matrices (TRC)*: In this method, the user was given visual feedback through a point plot where each point is a 2-D T-SNE transformed Riemannian covariance matrix [40]. Each EEG trial of 6 seconds is split into six one-second epochs. Epochs corresponding to right and left imaginations were marked differently in the plot, as shown in Fig. 2.

The normalised Riemannian covariance matrix $M \in \mathbb{R}^{c \times c}$ is defined as the following:

$$M = \frac{C * C^T}{\text{trace}(C * C^T)} \quad (1)$$

Where $C \in \mathbb{R}^{c \times T}$ is the EEG epoch matrix, where c is the number of EEG channels and T is the number of time samples. The details of the EEG system and channels used are provided in the next section. These matrices fall within the smooth Riemannian manifold on account of being Symmetric Positive Definite (SPD) matrices. The manifold dimension will be $c(c + 1)/2$. These matrices are flattened to column vectors and then transformed into two dimensions using t-SNE. The 2-dimensional points are plotted and shown on the display screen with different markers corresponding to different imagery to the user. The user is instructed to devise his own mental strategies in a way that the left and right markers get clustered separately.

2) *Mean Covariance Space (MCS)*: In this method, the user was given visual feedback through a Venn diagram, as shown in Fig. 3, consisting of two circles, one for left imagery and the other for right imagery. Each circle for the corresponding imagery has a moving average correlation of channels C3 and C4 as the center and the moving standard deviation of these channels as the radius.

$$\mu = \frac{1}{n} \sum_{i=1}^N x_i, \sigma = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (x_i - \mu)^2} \quad (2)$$

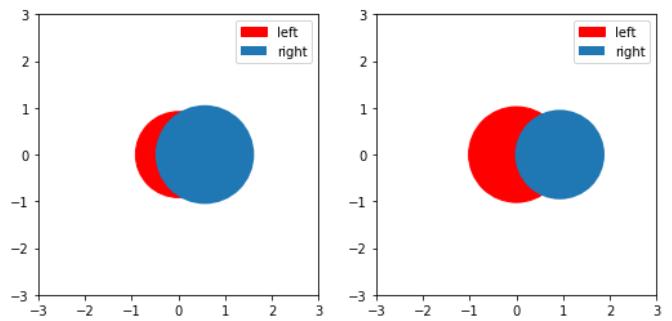
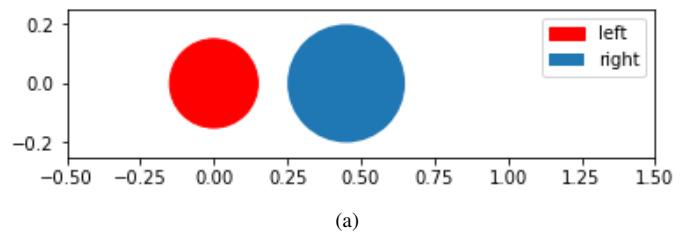
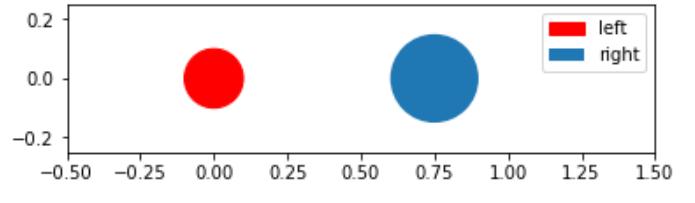


Figure 3: Venn-Diagrams representing the mean of correlation of channels as the center and standard deviation of correlation of channels as the radius (MCS) for left and right MI. Left Figure: Trials for S1 (Pre-training), left and right MI circles are almost overlapping; right Figure: Trials for S3 (Post-training), left and right MI circles are more distinct.



(a)



(b)

Figure 4: Venn-Diagrams representing the mean of LDA filtered Tangent space transformed Covariance Space correlations of channels (MLTCS) as the center and standard deviation as radius for left and right MI. Fig (a): Trials for S1 (Pre-training), left and right MI circles are less distinct. Fig (b): Trials for S3 (Post-training), left and right MI circles are more distinct.

Where μ is the mean of correlation, σ is the standard deviation of correlation, N is the number of trials, i is the correlation which is nothing but the lower triangular value of normalized Riemannian covariance matrix $M \in \mathbb{R}^{2 \times 2}$ which is defined as in equation 1.

3) *Mean LDA filtered Tangent space transformed Covariance Space (MLTCS)*: The Riemannian manifold is not a space where traditional machine learning algorithms can be applied. Through the use of linearization, the Riemannian manifold is mapped to the tangent space. For P_n , any point P in the tangent space is homogenous to the set of $n \times n$ symmetric matrices P_n . The mathematical details can be seen in [4]. For Euclidean tangent space, we may classify data using

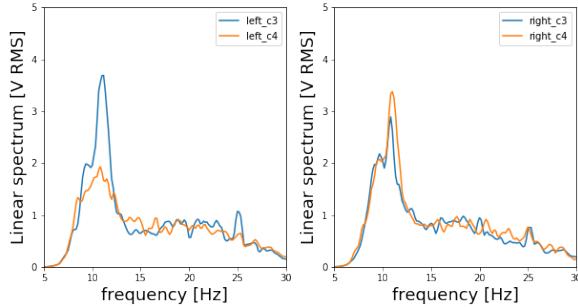


Figure 5: Power spectrum density graphs for left and right MI over C3 and C4 channels for a subject in Group 4.

any traditional ML method. Using the Linear Discriminant Analysis (LDA) criterion [41], we compute the various directional vectors after projecting data into the tangent space. This is an eigenvector decomposition maximizing the ratio of the between-class scatter matrix to the within-class scatter matrix. These LDA-weighted tangent space transformed features will provide a more distinguishing moving average and moving standard deviation through which a better venn diagram (given as visual feedback) will be obtained as shown in Fig.4, in turn reducing the user training load.

4) Power spectral density difference (PSDD): In this paradigm, users are given visual feedback in the form of line plots for both left and right imagery. Initially, we plotted PSDD [42] for channels C3 (corresponding to right imagery) and C4 (corresponding to left imagery) and marked their plots with blue and orange colors, respectively. Then for each corresponding imagery, we subtract the average PSDD plot obtained after n number of trials of the respective channel from that of the other. If the line plot obtained for the left imagery has a major segment as orange, then the user's performance is good and vice versa, while for the right imagery major segment will be blue for good user performance.

The motor imagery activity is prominent in the frequency range of 8-13 Hz, so considering a margin, we have plotted PSDD in the range of 5-15 Hz as shown in Fig. 5

III. RESULTS AND DISCUSSION

In this section, we will discuss the results achieved by the proposed feedback mechanisms to the MI user.

A. Participants

Twenty-four healthy subjects aged 19–30 (20 male, four female) were divided into four groups, G1, G2, G3, and G4, which consisted of 6 members containing, including five males and one female each. These four groups would receive four different feedback mechanisms. None of them has experience using MI-based BCIs. Before the trial started, each participant read and signed a consent form.

B. Experimental Design

Participants were made to sit comfortably on a chair during the MI tasks for both groups. The screen displays a single cross

in the middle as an attention-grabbing symbol for 3s, following which an arrow appears to indicate the hand (left/right) to which the subjects had to imagine clenching their wrist. Subjects were asked to visualize a hand gesture for 6 seconds while seeing a blank screen. The Timeline of each trial is shown in Fig. 1b.

There were three sessions, each consisting of 20 left and 20 right trials. In session 1, we recorded baseline MI data without letting the subjects practice. In session 2, each group was made to practice with their respective feedback approach. In session 3, the participants performed the MI task without feedback.

Throughout the MI sessions, EEG recordings were collected with an OpenBCI Electrode Cap (10–20 electrode position system), which allows up to 19 channels of EEG data to be obtained using wet electrodes. The Electrode Cap was paired with the OpenBCI CytonDaisy 16-channel Biosensing Board. The sampling rate was 250 Hz with contact impedance below five kΩ and the extracephalic reference electrode placed on the mastoid. The MI epoch was selected at time 0 from the on-screen appearance of a black screen after the stimulus arrow was shown. The epoch lasted for 6 seconds. The signals were band-pass filtered between 8 to 30 Hz. Artifact rejection was done. Baseline subtraction was performed.

The experimental sessions were carried out in a sound-proof room. The dataset consisted of 3 sessions(S1-S3) for 24 participants, which were grouped into four groups. Each session consisted of 20 left and 20 right trials. Each trial lasted for 6 seconds, which resulted in 1500 samples per trial. When considering two channels C3 and C4, the resultant data matrix is of size (40,2,1500), i.e. (trials, channels, samples) per session. The Riemannian covariance matrices thus shape to be (40,2,2). Only the lower triangular element is considered a feature that denotes the correlation of C3 and C4 channels. Therefore, we have 40 data points per session. Support vector machines(SVMs) are used to classify the trials of every session for every subject.

The classification accuracies for the pre-training session (S1) and post-training session (S3) are reported in Table II. A definite improvement was seen in all the subjects in all four groups from pre to post-training. Even if one may argue that the accuracies must be improved on account of practice in the first two sessions, the effect of different feedback mechanisms can be justified by observing the fact that there is different improvement in accuracies using them. G1, which used the TRC feedback mechanism, showed a mean improvement of 2.3 % in classification accuracy, while G2, which used MCS, showed 5.1 %, G3, which used MLTCS, showed 9.1 %, and G4, which used PSDD, showed 9.2 % mean improvement. The minor improvements in G1 and G2 can be attributed to the fact that TRC and MCS mechanisms employed raw eeg signals without any feature extraction and classification, whereas MLTCS employed the LDA classification technique, which provided better feedback leading to a major improvement in accuracy. PSDD employed power spectrum feature extraction and showed considerable improvement. On performing the t-

- [18] C. Guger, G. Edlinger, W. Harkam, I. Niedermayer, and G. Pfurtscheller, "How many people are able to operate an eeg-based brain-computer interface (bci)?" *IEEE transactions on neural systems and rehabilitation engineering*, vol. 11, no. 2, pp. 145–147, 2003.
- [19] M. Ahn and S. C. Jun, "Performance variation in motor imagery brain-computer interface: a brief review," *Journal of neuroscience methods*, vol. 243, pp. 103–110, 2015.
- [20] A. Kübler, B. Kotchoubey, T. Hinterberger, N. Ghanayim, J. Perelmouter, M. Schauer, C. Fritsch, E. Taub, and N. Birbaumer, "The thought translation device: a neurophysiological approach to communication in total motor paralysis," *Experimental brain research*, vol. 124, no. 2, pp. 223–232, 1999.
- [21] C. Vidaurre, C. Sannelli, K.-R. Müller, and B. Blankertz, "Machine-learning-based coadaptive calibration for brain-computer interfaces," *Neural computation*, vol. 23, no. 3, pp. 791–816, 2011.
- [22] F. Lotte, F. Larue, and C. Mühl, "Flaws in current human training protocols for spontaneous brain-computer interfaces: lessons learned from instructional design," *Frontiers in human neuroscience*, vol. 7, p. 568, 2013.
- [23] M. C. Thompson, "Critiquing the concept of bci illiteracy," *Science and engineering ethics*, vol. 25, no. 4, pp. 1217–1233, 2019.
- [24] K.-R. Müller, M. Tangermann, G. Dornhege, M. Krauledat, G. Curio, and B. Blankertz, "Machine learning for real-time single-trial eeg-analysis: from brain-computer interfacing to mental state monitoring," *Journal of neuroscience methods*, vol. 167, no. 1, pp. 82–90, 2008.
- [25] B. Blankertz, G. Dornhege, M. Krauledat, K.-R. Muller, V. Kunzmann, F. Losch, and G. Curio, "The berlin brain-computer interface: Eeg-based communication without subject training," *IEEE transactions on neural systems and rehabilitation engineering*, vol. 14, no. 2, pp. 147–152, 2006.
- [26] F. Lotte, L. Bougrain, A. Cichocki, M. Clerc, M. Congedo, A. Rakotomamonjy, and F. Yger, "A review of classification algorithms for eeg-based brain-computer interfaces: a 10 year update," *Journal of neural engineering*, vol. 15, no. 3, p. 031005, 2018.
- [27] C. Vidaurre, A. Schlogl, R. Cabeza, R. Scherer, and G. Pfurtscheller, "A fully on-line adaptive bci," *IEEE Transactions on Biomedical Engineering*, vol. 53, no. 6, pp. 1214–1219, 2006.
- [28] P. Shenoy, M. Krauledat, B. Blankertz, R. P. Rao, and K.-R. Müller, "Towards adaptive classification for bci," *Journal of neural engineering*, vol. 3, no. 1, p. R13, 2006.
- [29] M. Sugiyama, M. Krauledat, and K.-R. Müller, "Covariate shift adaptation by importance weighted cross validation," *Journal of Machine Learning Research*, vol. 8, no. 5, 2007.
- [30] S. Lu, C. Guan, and H. Zhang, "Unsupervised brain computer interface based on intersubject information and online adaptation," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 17, no. 2, pp. 135–145, 2009.
- [31] N. Birbaumer, T. Elbert, A. G. Canavan, and B. Rockstroh, "Slow potentials of the cerebral cortex and behavior," *Physiological reviews*, vol. 70, no. 1, pp. 1–41, 1990.
- [32] N. Birbaumer, A. Kubler, N. Ghanayim, T. Hinterberger, J. Perelmouter, J. Kaiser, I. Iversen, B. Kotchoubey, N. Neumann, and H. Flor, "The thought translation device (ttd) for completely paralyzed patients," *IEEE Transactions on rehabilitation Engineering*, vol. 8, no. 2, pp. 190–193, 2000.
- [33] B. W. El legado de John, "Skinner, bf (1945). the operational analysis of psychological terms. psycho."
- [34] F. Cincotti, L. Kauhanen, F. Aloise, T. Palomäki, N. Caporaso, P. Jylänki, D. Mattia, F. Babiloni, G. Vanacker, M. Nuttin *et al.*, "Vibrotactile feedback for brain-computer interface operation," *Computational intelligence and neuroscience*, vol. 2007, 2007.
- [35] C. Jeunet, C. Vi, D. Spelmezan, B. N'Kaoua, F. Lotte, and S. Subramanian, "Continuous tactile feedback for motor-imagery based brain-computer interaction in a multitasking context," in *IFIP Conference on Human-Computer Interaction*. Springer, 2015, pp. 488–505.
- [36] C. Jeunet, F. Lotte, and B. n'Kaoua, "Human learning for brain-computer interfaces," *Brain-Computer Interfaces 1: Foundations and Methods*, pp. 233–250, 2016.
- [37] F. Lotte and C. Jeunet, "Defining and quantifying users' mental imagery-based bci skills: a first step," *Journal of neural engineering*, vol. 15, no. 4, p. 046030, 2018.
- [38] C. Zich, S. Debener, M. De Vos, S. Frerichs, S. Maurer, and C. Kranczioch, "Lateralization patterns of covert but not overt move-
- ments change with age: an eeg neurofeedback study," *Neuroimage*, vol. 116, pp. 80–91, 2015.
- [39] L. Pillette, C. Jeunet, B. Mansencal, R. N'kambou, B. N'Kaoua, and F. Lotte, "A physical learning companion for mental-imagery bci user training," *International Journal of Human-Computer Studies*, vol. 136, p. 102380, 2020.
- [40] L. Van der Maaten and G. Hinton, "Visualizing data using t-sne," *Journal of machine learning research*, vol. 9, no. 11, 2008.
- [41] A. Barachant, S. Bonnet, M. Congedo, and C. Jutten, "Riemannian geometry applied to bci classification," *LvaIca*, vol. 10, pp. 629–636, 2010.
- [42] J. Dempster, *The laboratory computer: a practical guide for physiologists and neuroscientists*. Academic Press, 2001.
- [43] R. Fazel-Rezai, S. Amiri, A. Rabbi, and L. Azinfar, "A review of p300, ssvep, and hybrid p300/ssvep brain-computer interface systems," 2013.

Statistical and Deep Convolutional Feature Fusion for Emotion Detection from Audio Signal

Durgesh Ameta
IKSMHA Center

Indian Institute of Technology Mandi
Mandi, India
D22117@students.iitmandi.ac.in

Vinay Gupta

Department of Electrical Engineering
Indian Institute of Technology Kanpur
Kanpur, India
gupta@iitk.ac.in

Rohit Pilakkottil Sathian

Department of Software Engineering
Delhi Technological University,
Delhi, India
rohitpilakkottilsathian@gmail.com

Laxmidhar Behera
Director, Senior Member, IEEE
Indian Institute of Technology Mandi
Mandi, India
director@iitmandi.ac.in

Tushar Sandhan
Department of Electrical Engineering
Indian Institute of Technology Kanpur
Kanpur, India
sandhan@iitk.ac.in

Abstract—Speech serves as a crucial mode of expression for individuals to articulate their thoughts and can offer valuable insight into their emotional state. Various research has been conducted to identify metrics that can be used to determine the emotional sentiment hidden in an audio signal. This paper presents an exploratory analysis of various audio features, including Chroma features, MFCCs, Spectral features, and flattened spectrogram features (obtained using VGG-19 convolutional neural network) for sentiment analysis in the audio signals. This study evaluates the effectiveness of combining various audio features in determining emotional states expressed in a speech using the Ryerson Audio-Visual Database of Emotional Speech and Song (RAVDESS). Baseline techniques such as Random Forest, Multi-Layer Perceptron (MLP), Logistic Regression, XgBoost, and Support Vector Machine (SVM) are used to compare the performance of the features. The results obtained from the study provide insight into the potential of utilizing these audio features to determine emotional states expressed in speech.

Index Terms—Audio Emotion Classification (AEP), Multi-Layer Perceptron (MLP), Mel Frequency Cepstral Coefficient (MFCC), Visual Geometry Group19 (VGG19), Convolutional Neural Network(CNN)

I. INTRODUCTION

Emotions are key components of human communication and are expressed in a variety of ways across several modalities, including speech, biosignals, and text. From a specific field to a crucial element of Human-Computer Interaction (HCI), significant developments have been made in recognizing emotions expressed in speech [1]. As an alternative to using conventional devices as a means of comprehending spoken information and enable human listeners to react, it aims to provide natural connection with machines through direct voice interaction [2]. Spoken language dialogue systems, such as those utilized in call centre discussions, onboard car driving systems, and use of vocal emotion patterns in medical applications are a few examples of typical applications [3]

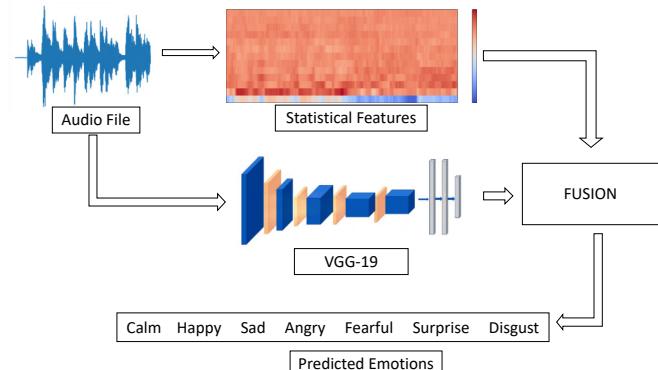


Fig. 1. Flow chart describing the overview of the framework. Features are extracted with two different approaches and then combined to produce an enhanced representation for analyzing emotional state hidden in the speech.

It is a unique assignment to ascertain a person's emotional state, and it can serve as a benchmark for any emotion identification software [4]. Among the many models used, a discrete emotional approach is considered to be one of the fundamental models used for classifying various emotions. It makes use of a variety of emotions, including grief, boredom, rage, surprise, contempt, and disgust [5]. Another significant model utilized was a continuous three-dimensional space, which incorporated characteristics for arousal, valence, and potency [6].

Researchers have established a few efficient Audio Emotion Prediction (AEP) techniques using digital audio speech signals [5]. These techniques typically involve two stages: feature extraction and feature classification.

In the initial stage, either several features are extracted from the spectrogram [7] using deep learning approaches (like CNN), or the hand-crafted features (like frequency bands [8], chroma soft [9], MFCCs [10], spectral centroid [11]) are

extracted from the Librosa library.

The second stage involves the classification of features using linear or non-linear classifiers. The most commonly employed linear classifiers for emotion recognition are Multi-Layer Perceptron (MLP), Bayesian Networks (BN) and Support Vector Machine (SVM). The speech signal is typically regarded as non-stationary. From previous researches conducted, it is also believed that non-linear classifiers, such as the Hidden Markov Model (HMM) and Gaussian Mixture Model (GMM), perform well for AEP [12].

To the author's best knowledge, none of the studies have combined features from the spectrogram and librosa for AEP. In this study, we concatenated both sets of features and observed that simple concatenation of multi-domain features lead to better accuracy.

The paper's layout is as follows: After the introduction, Section II lists related works. Section III covers the Methodology, Section IV covers the Experimental Setup, Section V covers the Results, and Section VI covers the Conclusion.

II. RELATED WORK

AEP tasks are handled by various approaches in the literature, which can be classified broadly into two categories namely,

1) *Hand-Crafted Feature-Based Audio Emotion Prediction (AEP)*: In this approach, before recognition of emotion, spectral and prosodic features from raw audio data are extracted. Few of these important features are pitch, formants, and energy, as well as the linear prediction cepstrum coefficient (LPCC), Mel energy spectrum dynamic coefficient (MEDC), Mel frequency cepstrum coefficients (MFCCs), and perceptual linear Prediction cepstrum coefficient (PLP) [10], [16]. For the extracted features, the proposed classifiers include Hidden Markov Models (HMM), Bayesian Networks (BN), K-nearest Neighbors (KNN), SVM, GMM or Artificial Neural Networks (ANN) [13]- [15]. According to Dave et al [16], among several speech emotion features, the performance of AEP trained with MFCCs is superior to that of other low-level features like formant, linear productivity code (LPC), loudness, etc. On the Berlin Database of Emotional Speech [18], Yixiong Pan utilised SVM for three classes of emotion classification and reported 95.1 percent accuracy [17].

2) *Convolutional Neural Network (CNN)-Based AEP*: In contrast to these conventional methods, numerous novel papers have lately been published, reporting promising outcomes using Deep Neural Networks in their experiments. George et al., classified the REEmote COLlaborative and Affective (RECOLA) natural emotion database, by providing a method for spontaneous AEP based on CNN and LSTM [20]. In order to determine the speaker's emotions, the author utilised two CNN layers to extract the discriminative feature from the entire utterance. They then fed this information to two LSTM layers for sequence learning. On the CASIA Chinese dataset, Lian et al. [21] employed the DBN model to learn features

and extract hidden information from speech, and utilized the SVM classifier to predict emotions. In the literature, for extracting discriminative features from audio signals, several techniques have been described using the CNN model for AEP with various forms of input [5]. Deep learning AEP methods enhanced the recognition ratio for real-time spontaneous AEP utilising a variety of voice datasets, such as RAVDESS [22], IEMOCAP [27], SAVEE [28] etc.

III. METHODOLOGY

This section describes the proposed methodology for AEP. We explain the feature extraction and classification techniques employed. Statistical or the spectral features and spectrogram or CNN features are explained. The proposed concatenation of these features is also detailed here.

A. Feature Extraction

1) *Statistical features*: An audio signal is denoted as y and it is sampled at the frequency of sampling rate ($sr=22050\text{ Hz}$). Fig. 2 shows a waveform plot of a three-second audio clip of happy sound.

The feature set included the mean and variance computed for the following list of spectral features.

- Chroma STFT (*short-time Fourier transform*): This feature mainly represents a variation of STFT which is used for the extraction of harmonic information from audio signals [5]. $S(n)=$ Speech Signal, segmented into M samples using a hop size of L samples where w represents discrete-time window of length N . The resulting STFT represent:

$$Y_{k,l} = \sum_{n=1}^{N-1} S(lL + n)w(n)e^{-j\Omega_k n} \quad (1)$$

with segment index l , frequency index k , and the normalized angular frequencies $\Omega_k = \frac{2\pi k}{N}$.

- Chroma CQT (*constant-Q transform*): This feature set is used for the representation of the audio waveform so that it encodes the pitch content of the signal. Pitch content contains important information about genders, speakers, speaking styles, and content of the audio [28]. Chroma CQT is extracted by taking a logarithmically spaced set of frequency band [8]. Fig. 3 shows graphical representation of Chroma CQT extracted from the audio signal y .

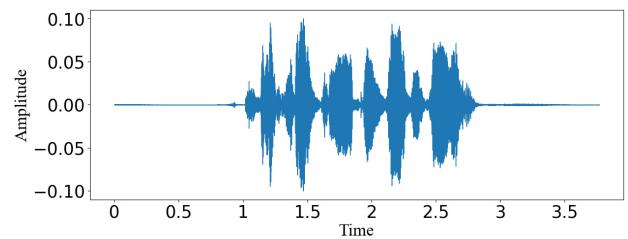


Fig. 2. An image depicting the waveform of a three-second audio clip featuring the emotion of happiness, denoted as y .

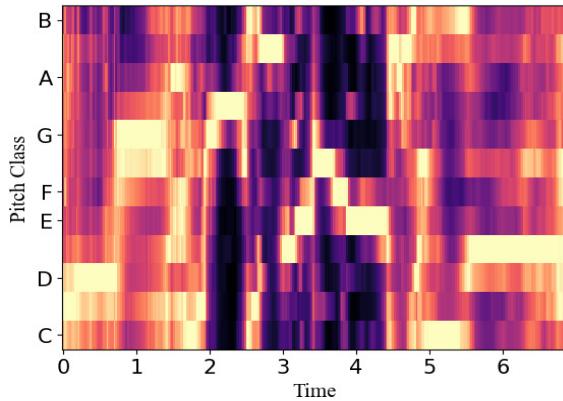


Fig. 3. Graphical Representation of Chroma CQT which was taken from the audio signal y .

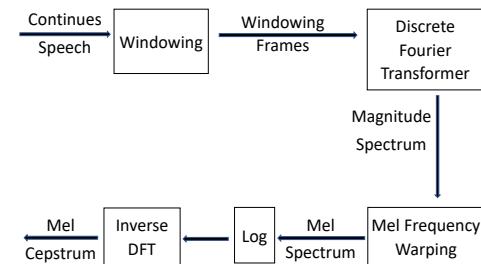


Fig. 4. Complete pipeline for obtaining MFCCs [10]

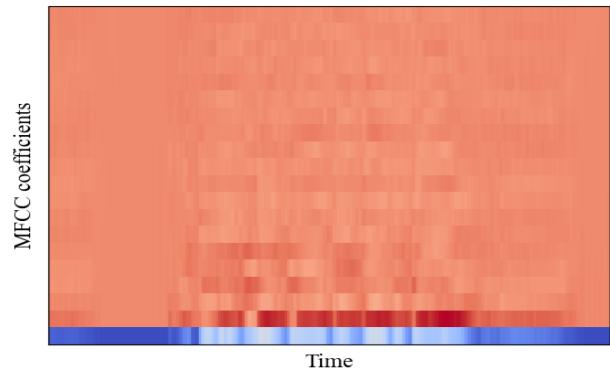


Fig. 5. MFCCs plot of audio signal y of happy sound.

- Chroma CENS (*Chroma Energy Normalization Statistics*): In this implementation of the Chroma feature space, the energy in each dimension is normalised by dividing the same with the sum of the energies [9]. The short-term energy is computed according to the equation

$$E(i) = \sum_{n=1}^{W_L} |x_i(n)|^2 \quad (2)$$

where $x_i(n), n = 1, \dots, W_L$ represents the sequence of audio samples of the i^{th} frame and W_L represents the length of the frame.

The normalized energy is given by:

$$E_N(i) = \frac{1}{W_L} E(i) \quad (3)$$

- MFCCs (*Mel Frequency Cepstral Coefficients*): This is a set of coefficients which is derived by first performing Fourier transform on the audio signal, given by:

$$S_i(k) = \sum_{n=1}^N S_i(n) h(n) e^{-\frac{j2\pi kn}{N}} \quad (4)$$

where $S(n)$ denotes time domain signal, n and i is the range over the no. of samples and frames respectively, $h(n)$ is hamming window and N is the length of DFT.

The power spectral estimate for the audio frame $s_i(n)$ based on the periodogram is given by:

$$P_i(k) = \frac{1}{N} |S_i(k)|^2 \quad (5)$$

The resultant spectrum obtained is then mapped to the “mel-scale” (a nonlinear frequency scale that is based on the human’s auditory system’s response to sound), then the logarithm of the power is determined and finally to decorrelate the coefficients a discrete cosine transform is applied. A total of 13 MFCCs were used in this experiment [10]. Fig. 4 shows complete pipeline for obtaining MFCCs. Fig. 5 shows MFCCs plot of audio signal y .

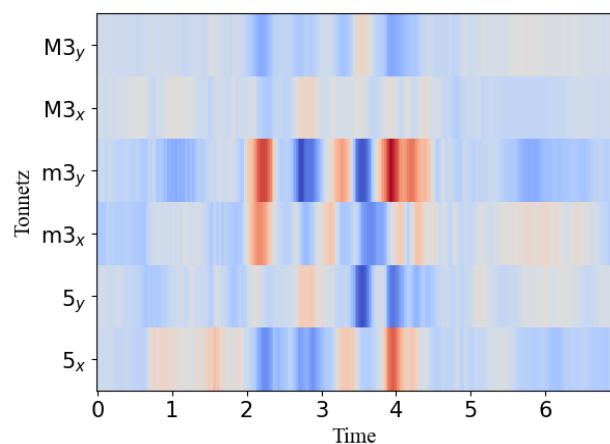


Fig. 6. Graphical representation of Tonnetz of audio signal y

- ZCR (zero crossing rate): a count of the instances where a signal's sign flips from positive to negative or vice versa [5]. Fig.7 shows the number of zero crossings of waveform
- RMS: This feature is used to measure the power/ magnitude of a signal [5].

$$RMS = \sqrt{\sum_{n=1}^N x^2(n)} \quad (6)$$

where $x(n)$ represents value of n^{th} frame.

- Spectral Centroid: This feature indicates the frequency at which the spectral energy is concentrated [8]. Spectral centroid, C_i of the i^{th} audio frame is calculated as:

$$C_i = \frac{\sum_{k=1}^{Wf_L} k X_i(k)}{\sum_{k=1}^{Wf_L} X_i(k)} \quad (7)$$

where Wf_L represents the number of coefficients that are used in the computations and $X_i(k)$ represents the magnitude of the DFT coefficients of the i^{th} audio frame, which is calculated as:

$$X_i(k) = x(i) \exp(-j \frac{2\pi}{N} ki) \quad (8)$$

where $x(i)$ is the i^{th} discrete-time signal sample and $j = \sqrt{-1}$.

- Spectral Bandwidth: This feature measures the width of the frequency distribution of sound [8].
- Spectral Contrast: this feature represents the relative strength of different frequency components in the audio

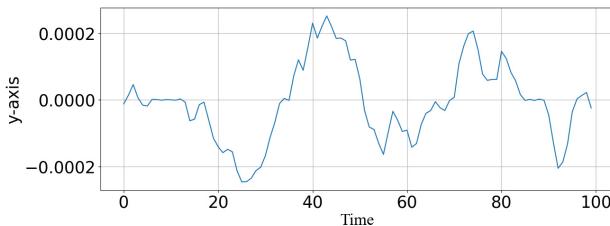


Fig. 7. Graph depicting Zero Crossing Rate

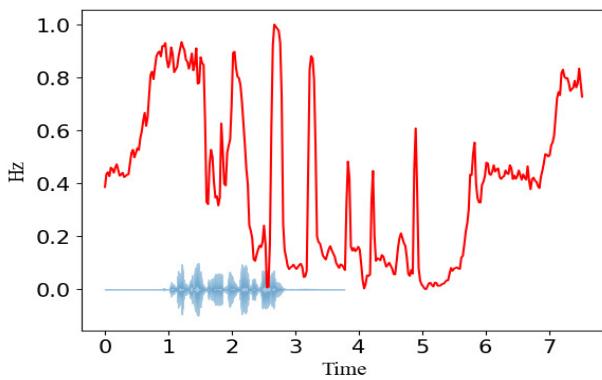


Fig. 8. Energy of the spectrum shown by Spectral Centroid

signal, it's useful to identify various types of patterns that occur in different sound [11].

- Spectral Flatness (SF): this features quantifies the level of tonal and noisy components present in an audio file [24].

$$SF = \frac{\sqrt{\prod_{n=0}^{N-1} x(n)}}{\sum_{n=0}^{N-1} x(n)} = \frac{\exp(\frac{1}{N} \sum_{n=0}^{N-1} \ln x(n))}{\frac{1}{N} \sum_{n=0}^{N-1} x(n)} \quad (9)$$

- Spectral Roll-off: the frequency below which a particular proportion of the total spectral energy is present [24]. It satisfies the following condition:

$$\sum_{k=1}^m X_i(k) = C \sum_{k=1}^{Wf_L} X_i(k), \quad (10)$$

(where C = adopted percentage.)

- Onset Detection: Identification of the beginning point of an audio signal [26].
- Onset Strength:It is a measure of prominence of onset in an audio signal [26].
- In addition to these features, beats and tempo were also taken into consideration.

2) *Spectrogram Features using VGG19*: In addition to the features obtained above, the audio files were then passed through the Mel spectrogram (in such cases, overlapping triangle filters are used to scale the frequency axis to the Mel scale) to obtain a spectrogram which is then resized using inter nearest interpolation technique to finally obtain a $224 \times 224 \times 3$ image. The proposed methodology makes use of a pre-trained VGG-19 network (on ImageNet) to obtain the flattened features of the spectrogram image.

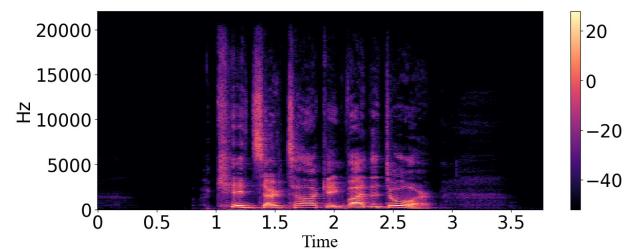


Fig. 9. This is a visual representation of the audio signal's Mel spectrogram with 128 Mel bands.

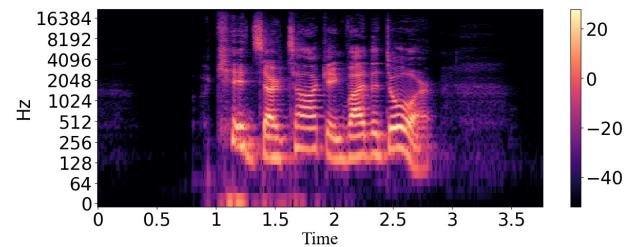


Fig. 10. Plot of Log Spectrogram

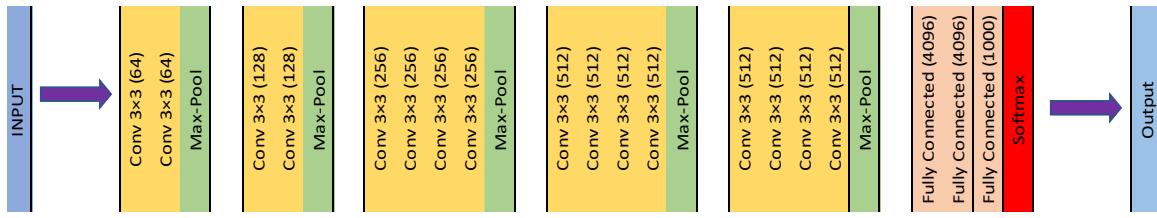


Fig. 11. Architecture of VGG19 Model [30]

VGG19 is a convolutional neural network that comprises of 19 layers, including 16 convolutional layers, 3 fully connected layers, 5 max pooling layers, and 1 SoftMax layer [30]. The network is trained on the ImageNet database, made of over a million images from 1000 categories. As a result, it possesses the capability to categorize images into one of the 1000 object categories. It has 3×3 filters in each convolutional layer. Max Pooling is a down-sampling strategy in Convolutional Neural Networks. The expression for the down-sampling layer is:

$$\chi_{pj}^n = f(\tau_j^n \text{down}(\chi_j^{(n-1)}) + b_j^n) \quad (11)$$

$\text{down } \chi_j^{(n-1)}$ is the maximum pooling sampling function, τ_j^n is the coefficient corresponding to the j^{th} feature map of the n^{th} layer, and $f(\tau_j^n \text{down}(\chi_j^{(n-1)}) + b_j^n)$ is the ReLU activation function. Fig. 11 describes architecture of VGG19 Model. The flattened VGG-19 features are reduced in dimension using Principal Component Analysis (PCA) [31].

PCA transforms n -vectors (v_1, v_2, \dots, v_n) of Z -dimensional space into Z' space with principles (v_1, v_2, \dots, v_n) where Z and Z' are positive integers with size $Z' \leq Z$.

The output features of PCA can be represented as:

$$f_{n'} = \sum_{k=1}^{U'} C_{k,i} G_k \quad (12)$$

where G_k = eigenvectors and $C_{k,i}$ = principal components. In this study, PCA was applied to the flattened features of the spectrogram converted it into a 50-dimensional vector.

B. Feature Classification

To analyze the data, several machine learning models were employed. The models used in this study include Random Forest, Logistic Regression, XgBoost, MLP and SVM. The performance of these models was evaluated and compared in order to determine the most effective model. Different combinations of the features sets were taken into consideration in order to know how each feature influenced one another. Statistical features alone; PCA-reduced flattened spectrogram features alone; a combination of both were considered. Fig. 12 depicts the entire pipeline of feature classification. To analyze the data, several machine learning models were employed. The models used in this study include Random Forest, Logistic Regression, XgBoost, MLP and SVM.

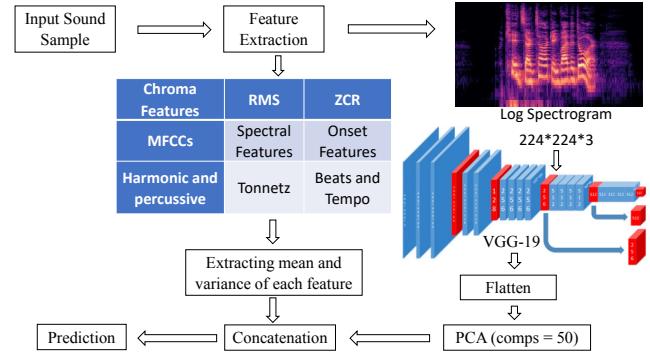


Fig. 12. Emotion Classification after concatenating statistical features and flattened spectrogram features.

The evaluation and comparison of these models' performance were conducted to establish the model that delivers the highest level of effectiveness. Different combinations of the features sets were taken into consideration in order to know how each feature influenced one another. Statistical features alone; PCA-reduced flattened spectrogram features alone; a combination of both were considered.

IV. EXPERIMENTAL SETUP

A. Dataset

The RAVDESS [22] dataset was chosen for the present study. It includes speech and song records, 247 untrained Americans categorized these records into: Calm, Happy, Sad, Surprise, Disgust, Fear, Neutral, and Anger base for each performer. Following is a brief description of the dataset:

- A total of 24 trained performers—12 men and 12 women—make up the data collection.
- Each of the audio recordings has the same statements with an American accent, and they were all recorded in a controlled environment. There are two further related file formats.
- The audio file has 1440 audio files total—60 trials for each actor multiplied by 24 actors.
- The raw audio file format uses a sampling rate of 48 kHz and a bitrate of 16 bits for both kinds of audio files. No information from the original recordings of the audio files in the dataset is lost or changed because both recordings are lossless, uncommented audio files.

B. Libraries

We used the python ‘librosa’ library to extract all the statistical features of the audio files [32]. Keras and tensorflow libraries are used to train the VGG-19 classifier [33].

V. RESULTS AND DISCUSSIONS

In this section, we have evaluated our proposed set of features on RAVEDESS [22] using Random Forest, Logistic Regression, XgBoost, SVM, and MLP. Several experiments were conducted by changing the hyperparameters to fine-tune the model and obtain the best results. The baseline models were tested on three feature sets i.e. statistical features [A], flattened spectrogram features obtained after passing through pre-trained VGG-19 layers on which PCA was applied to reduce the vector dimensions to 50 [B], and [C] obtained after concatenation of (A) and (B).

TABLE I

CLASSIFICATION PERFORMANCE FOR FEATURE SET [A], CONSISTING OF ONLY STATISTICAL FEATURES ON RAVEDESS DATASET.

Models	Accuracy (%)	Training Time (s)
Random Forest	61.80	0.26
Logistic Regression	62.26	0.08
XgBoost	64.35	0.61
MLP	71.06	8.16
SVM	60.18	0.07

TABLE II

CLASSIFICATION PERFORMANCE FOR FEATURE SET [B], FLATTENED SPECTROGRAM FEATURES OBTAINED FROM PRE-TRAINED VGG-19.

Models	Accuracy (%)	Training Time (s)
Random Forest	55.55	0.27
Logistic Regression	56.71	0.03
XgBoost	56.48	0.61
MLP	55.79	9.79
SVM	53.47	0.09

TABLE III

CLASSIFICATION PERFORMANCE FOR FEATURE SET [C], OBTAINED AFTER CONCATENATION OF SET [A] AND SET [B].

Models (Our fusion)	Accuracy (%)	Training Time (s)
Random Forest	63.88	0.34
Logistic Regression	67.59	0.09
XgBoost	64.12	0.82
MLP	73.61	19.96
SVM	64.58	0.05

As indicated by the results presented in Tables I, II, and III, the Multi-Layer Perceptron (MLP) classifier achieved the top accuracy with a score of 71.06% for the feature set [A]. In addition, SVM took the shortest training time of 0.07 seconds. For the feature set [B], the Logistic Regression classifier obtained the highest accuracy of 56.71% and the least training

duration of 0.03 seconds. When using feature set [C], the MLP classifier again obtained the best performance with an accuracy of 73.61%. The accuracies for different classifiers for all three sets of features are plotted in Fig. 13.

From Table I and Table II, it is observed that the feature set [A] gave good results as compared to feature set [B], this might be because statistical features, such as the mean and variance of spectral features, MFCCs, chroma features, etc., provide appropriate information about the distribution of the audio signal over time. Since these features have a variety of sub-features associated with them, they are able to differentiate between different audio classes.

There could be several reasons for the low accuracy obtained for feature set [B]. One reason may be that we used the flattened features directly without fine-tuning the layers of VGG-19, as fine-tuning also requires a significant amount of training data. Additionally, the dataset only had 192 entries per emotion, which might have resulted in insufficient data for the CNN model to learn from. To improve the accuracy of feature set [B] in future implementations of the study, we can consider several options: firstly, we can fine-tune the model by unfreezing several layers or entire blocks of VGG-19; secondly, we could increase the size of the dataset by combining multiple datasets or by gathering new dataset.

While attempting to train the models using feature set [B], we noticed that we achieved better accuracies without applying PCA. In fact, for some models, the increase in accuracy was around 5%-7%. However, there was a downside: the models took a lot of time to converge, and some models did not converge at all when using the VGG-19 feature set without PCA-based dimensionality reduction. Additionally, when we combined large dimensional CNN features with statistical features, we observed that the accuracy was similar to the accuracy without these feature fusion. This could be due to the fact that the large number of flattened spectrogram features (without PCA) were inherently given more priority than the statistical features. Therefore, we utilized PCA-based low dimensional CNN features, even though they produced

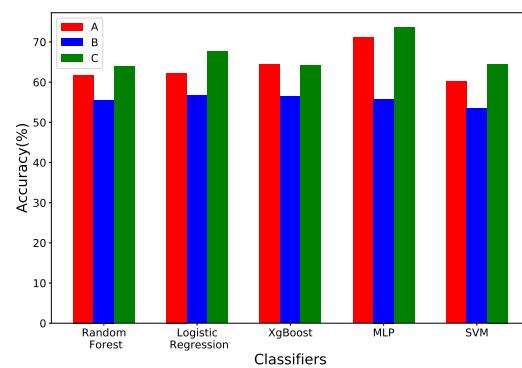


Fig. 13. Bar Graph comparing the accuracies obtained by different classifiers using the proposed set of features.

lower accuracy when used in isolation as reported in Table II. However, the fusion of these features with statistical features helped to increase the overall accuracy of the model, as demonstrated in Table III.

VI. CONCLUSION

In this work, the audio emotion recognition system has been designed using various supervised machine learning techniques on the RAVDESS dataset using our proposed set of features. The study's results have shown that the combination of statistical features and flattened spectrogram features produced better results than using these features separately. Among the baseline models, MLP performed best, achieving an accuracy of 73.61%, while SVM had the shortest training time at 0.05 seconds.

The models we have utilized in this study were traditional machine learning classifiers, and in future, we would explore end-to-end learning CNN architectures.

REFERENCES

- [1] R. Cowie, E. Douglas-Cowie, N. Tsapatsoulis, G. Votsis, S. Kollias, W. Fellenz, and J. G. Taylor, "Emotion recognition in human-computer interaction," *IEEE Signal Process. Mag.*, vol. 18, no. 1, pp. 32–80, Jan. 2001.
- [2] J. G. Rázuri, D. Sundgren, R. Rahmani, A. Moran, I. Bonet, and A. Larsson, "Speech emotion recognition in emotional feedback for human-robot interaction," *Int. J. Adv. Res. Artif. Intell.*, vol. 4, no. 2, pp. 20–27, 2015
- [3] B. Schuller, G. Rigoll and M. Lang, "Speech emotion recognition combining acoustic features and linguistic information in a hybrid support vector machine-belief network architecture," 2004 IEEE International Conference on Acoustics, Speech, and Signal Processing, Montreal, QC, Canada, 2004, pp. I-577, doi: 10.1109/ICASSP.2004.1326051.
- [4] M. S. Hossain and G. Muhammad, "Emotion recognition using deep learning approach from audio-visual emotional big data," *Inf. Fusion*, vol. 49, pp. 69–78, Sep. 2019
- [5] McFee, B. et al. (2015) "Librosa: Audio and Music Signal Analysis in python," Proceedings of the 14th Python in Science Conference [Preprint]. Available at: <https://doi.org/10.25080/majora-7b98e3ed-003>.
- [6] Goudbeek, M., Scherer, K. (2010). Beyond arousal: valence and potency/control cues in the vocal expression of emotion. *The Journal of the Acoustical Society of America*, 128(3), 1322–1336. <https://doi.org/10.1121/1.3466853>
- [7] Tachibana, H., Ono, N., and Sagayama, S. (2014) "Singing voice enhancement in monaural music signals based on two-stage harmonic/percussive sound separation on multiple resolution spectrograms," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 22(1), pp. 228–237. Available at: <https://doi.org/10.1109/taslp.2013.2287052>.
- [8] Brown, J.C. (1991) "Calculation of a constant q spectral transform," *The Journal of the Acoustical Society of America*, 89(1), pp. 425–434. Available at: <https://doi.org/10.1121/1.400476>.
- [9] Meinard Müller and Sebastian Ewert "Chroma Toolbox: MATLAB implementations for extracting variants of chroma-based audio features" In Proceedings of the International Conference on Music Information Retrieval (ISMIR), 2011.
- [10] Hossan, M.A., Memon, S. and Gregory, M.A. (2010) "A novel approach for MFCC feature extraction," 2010 4th International Conference on Signal Processing and Communication Systems [Preprint]. Available at: <https://doi.org/10.1109/icspcs.2010.5709752>.
- [11] Dan-Ning Jiang et al. (no date) "Music type classification by spectral contrast feature," Proceedings. IEEE International Conference on Multimedia and Expo [Preprint]. Available at: <https://doi.org/10.1109/icme.2002.1035731>.
- [12] A. D. Dileep and C. C. Sekhar, "GMM-based intermediate matching kernel for classification of varying length patterns of long duration speech using support vector machines," *IEEE Trans. neural Netw. Learn. Syst.*, vol. 25, no. 8, pp. 1421–1432, Aug. 2014.
- [13] R. Chakraborty and S.K. Kopparapu, "Improved speech emotion recognition using error correcting codes", *Multimedia Expo Workshops (ICMEW) 2016 IEEE International Conference on*, pp. 1-6, 2016, July.
- [14] W. Fei, X. Ye, Z. Sun, Y. Huang, X. Zhang and S. Shang, "Research on speech emotion recognition based on deep auto-encoder", *Cyber Technology in Automation Control and Intelligent Systems (CYBER) 2016 IEEE International Conference on*, pp. 308-312, 2016, June.
- [15] G. Vyas, M.K. Dutta, K. Riha and J. Prinosil, "An automatic emotion recognizer using MFCCs and Hidden Markov Models", *Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT) 2015 7th International Congress on*, pp. 320-324, 2015, October.
- [16] Dave, N. Feature extraction methods LPC, PLP and MFCC in speech recognition. *Int. J. Adv. Res. Eng. Technol.* 2013, 1, 1–4.
- [17] Pan, Y., Shen, P. and Shen, L., 2012. Speech emotion recognition using a support vector machine. *International Journal of Smart Home*, 6(2), pp.101-108.
- [18] Burkhardt, F., Paeschke, A., Rolfes, M., Sendlmeier, W.F. and Weiss,B., 2005, September. A database of German emotional speech. In *Interspeech* (Vol. 5, pp. 1517-1520).
- [19] Erol, B.; Seyfoglou, M.S.; Gurbuz, S.Z.; Amin, M. Data-driven cepstral and neural learning of features for robust micro-Doppler classification. In *Proceedings of the Radar Sensor Technology XXII*, Orlando, FL, USA, 16–18 April 2018; p. 106330J
- [20] Trigeorgis, G.; Ringeval, F.; Brueckner, R.; Marchi, E.; Nicolaou, M.A.; Schuller, B.; Zafeiriou, S. Adieu features? end-to-end speech emotion recognition using a deep convolutional recurrent network. In *Proceedings of the 2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Shanghai, China, 20–25 March 2016; pp. 5200–5204.
- [21] Zhu, L.; Chen, L.; Zhao, D.; Zhou, J.; Zhang, W. Emotion recognition from Chinese speech for smart effective services using a combination of SVM and DBN. *Sensors* 2017, 17, 1694.
- [22] 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), Article e0196391. <https://doi.org/10.1371/journal.pone.0196391>
- [23] "Signal Processing Methods for music transcription" (2006). Available at: <https://doi.org/10.1007/0-387-32845-9>.
- [24] Dubnov, S. (2004) "Generalization of spectral flatness measure for Non-Gaussian Linear Processes," *IEEE Signal Processing Letters*, 11(8), pp. 698–701. Available at: <https://doi.org/10.1109/lsp.2004.831663>.
- [25] Harte, C., Sandler, M., Gasser, M. (2006). "Detecting Harmonic Change in Musical Audio." In *Proceedings of the 1st ACM Workshop on Audio and Music Computing Multimedia* (pp. 21-26). Santa Barbara, CA, USA: ACM Press. doi:10.1145/1178723.1178727.
- [26] Böck, Sebastian, and Gerhard Widmer. "Maximum filter vibrato suppression for onset detection." 16th International Conference on Digital Audio Effects, Maynooth, Ireland. 2013.
- [27] C. Busso, M. Bulut, C. Lee, A. Kazemzadeh, E. Mower, S. Kim, J. Chang, S. Lee, and S. Narayanan, "IEMOCAP: Interactive emotional dyadic motion capture database," *Journal of Language Resources and Evaluation*, vol. 42, no. 4, pp. 335–359, December 2008.
- [28] Jackson, Philip and ul haq, Sana. (2011). Surrey Audio-Visual Expressed Emotion (SAVEE) database.
- [29] Ayadie ME, Kamelb MS, Karayb F: Survey on speech emotion recognition: features, classification schemes, and databases. *Pattern Recognit* 2011, 44(3):572-587.
- [30] Núñez-Marcos, Adrián Azkune, Gorka Arganda-Carreras, Ignacio. (2017). Vision-Based Fall Detection with Convolutional Neural Networks. *Wireless Communications and Mobile Computing*. 2017. 1-16. 10.1155/2017/9474806.
- [31] Karl Pearson F.R.S. , 1901. LIII. On lines and planes of closest fit to systems of points in space. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 2(11), pp.559–572.
- [32] McFee, B., Raffel, C., Liang, D., Ellis, D. P., McVicar, M., Battenberg, E., Nieto, O. (2015). librosa: Audio and music signal analysis in python. In *Proceedings of the 14th python in science conference* (Vol. 8).
- [33] Chollet, F., others. (2015). Keras. GitHub. Retrieved from <https://github.com/fchollet/keras> 1

A Review on Preprocessing of EEG Signal

Suveetha Dhanaselvam P,
Associate Professor,

Department of Electronics and Communication Engineering,
Velammal College of Engineering and Technology,
Madurai, India
suveethaj@gmail.com

C.Nadia Chellam
Research Scholar

Department of Information and Communication Engineering
Velammal College of Engineering and Technology,
Madurai, India
nadiachellamc@gmail.com

Abstract—Electroencephalogram (EEG) is the documentation of brain's electrical activity tapped from the scalp. The signals picked from the scalp do not express an accurate representation of the brain signals. These bio-signals need to be processed in order to be used for the desired application. To unravel this problem, there is a necessity to define a strong and repeatable EEG pre-processing method. EEG data pre-processing specifies a procedure of remodeling the raw EEG data into a clean EEG data by removing the undesirable noise and artifacts thereby converting it into suitable format for further analysis and interpretable by the user. This paper tends to review various EEG preprocessing techniques that has been described within the published literatures so as to focus on the acceptable preprocessing modality for a specific application.

Keywords—electroencephalogram, signal preprocessing, artifacts.

I. INTRODUCTION

With the emanation of non-invasive techniques, numerous research has been conducted on brain activity and related disorders. EEG tends to be a vital, noninvasive and cost effective technique accustomed monitor the state of the brain. The International Federation of Clinical Neurophysiology defines EEG as that (i) the science relating to the electrical activity of the brain, and (ii) the technique of recording electroencephalograms.

The EEG signal described as a nonlinear and non-stationary random weak signal[1]. On contrast with techniques as fMRI or PET, EEG has a high temporal resolution. The electroencephalogram records the cerebral electrical potentials by means of electrodes positioned over the scalp. The EEG has an amplitude of some 5 to 200 μ V and is exposed to artifacts, from both bioelectric and physical sources[2]. The signal reflects the brain's functional state in addition to mental condition and this vital information is required in order to track patient's health [3].

II. EEG RECORDING AND PREPROCESSING

The brain signal has characteristic information in many regions at any given time. The EEG records voltage fluctuations by means of scalp electrodes, due to the flow of electrical charge for each excitation of the synapse in the neurons of the brain [4]. Here The noninvasiveness and inexpensive nature made EEG the most popular modality. The number of electrode varies from 1 to 256 for different EEG headsets. The EEG signals can be arranged in accordance with their frequency bands each of which has specific biological significance.

TABLE I. EEG FREQUENCY BANDS

Band	Frequency	Amplitude (μ V)	Location	Activity
Delta	0.5 – 4 Hz	100 – 200	Frontal and during deep sleep	Deep sleep
Theta	4 – 8 Hz	5 – 10	Hippocampus region	Drowsiness, Light sleep
Alpha	8 – 13 Hz	20 – 80	Occipital head region	Relaxed
Beta	13 – 30 Hz	1 – 5	Symmetrical distribution and most evident in the frontal and central head regions	Active thinking, alert
Gamma	>30 Hz	0.5 - 2	Widely over the cerebral cortex	Hyperactivity

The EEG electrodes appear as small metal disc made up of stainless steel or tin or silver alongwith silver chloride coating. These electrodes are placed in special positions on the scalp by means of internationally recognised system - International 10/20 system. EEG electrodes can be Wet or Dry electrodes. Traditionally wet electrodes made of Ag/AgCl were used. While recording EEG, the electrodes not only pick up the clean brain signal rather it will be

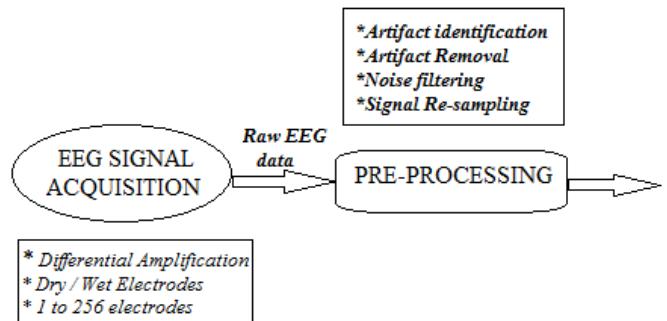


Fig 1. Pre-processing methodology

adulterated with a variety of noise and artifacts. Following the recording step, preprocessing of the EEG signal is needed. Here, preprocessing refers to mainly removing noise , to urge closer to true neural signals and converting the information into a more useful format that may be useful for further analysis and research work. The principal pre-processing steps are the artifacts identification and removal, noise filtering, and re-sampling the signal to adjust to detector input specifications.

III. EXISTING METHODOLOGIES

A. EEG Signal Filtering and Re-sampling

Artifacts – are signals recorded by EEG that is not of cerebral origin. The origin of artifacts may be physiological or extra-physiological. Physiological artifacts are generated within the patient (e.g., ocular movements [6], cardiac, eye blinks and muscular activity [7] and extra-physiological artifacts are from outside the body or from external environment (e.g., EMI, 50/60 Hz artifact, cable movements, electrode paste-related). Ocular artifacts (OA) and myogenic artifacts(MA) contaminate EEG signals in which OA is visible as relatively large pulses in the frontal region [8], and MA appears in the temporal and occipital regions with wide frequency spectrum [7].

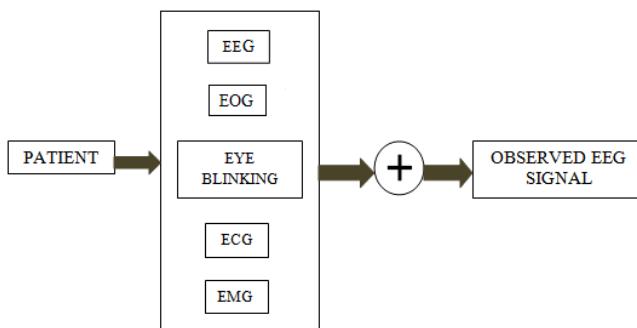


Fig 2. Bio-potential artifacts

Removal of artifacts needs careful consideration so as to retain the useful information of the EEG signal. Differential measurements prevent artifacts induced by Electromagnetic Interference (EMI), mainly by power lines [5]. Good electrode placement avoids errors and eliminates artifacts [2]. Wu Wen et al. [9] proposed a technique with which the EEG signals are used to detect the sleep quality.

The electrophysiological signal - EEG has low amplitude and is highly prone to noise intervention. Hence, in the preprocessing stage, the signal is filtered to scale high-frequency noise, then it is split into 30s component so as to eliminate baseline drift and artifact interference which is illustrated by means of a flowchart.

Miao Shi et al. [10] proposed a method on EEG signals classification by pattern recognition method and the Support Vector Machine (SVM) And Optimization by means of Improved Squirrel Search Algorithm (ISSA). In this method, the preprocessing is done by (i)extract the time-domain features as feature vectors, (ii) feature vectors

directed to SVM, and (iii) Classification and identification of the required data. To filter the acquired data, Miao Shi et al. uses an elliptical filter to extract the required information from the μ wave and β rhythm of the EEG signal which has the energy band within 8–30 Hz frequency. Here the elliptical filter has the pass band frequency that matches with the acquired waves' energy band during when the pass band ripples under 1 db and signal attenuation occur at a range of 5 Hz on each side of the pass band which is 40 db.

A multichannel Weighted Weiner filter been proposed for attenuating eye blink artifacts[11] here Hierarchical Fully Connected Topology (HFCT) and Ad-hoc Nearest-Neighbor Topology (ANNT) are employed. This provides 5% better results for artifact attenuation in comparison with the existing approaches like Principal Component Analysis (PCA) and Independent Component Analysis (ICA). However, the proposed approach so far not been used in real medical devices.

Hao Chao et al.[12] in his study proposed an EEG emotion recognition framework combining multiband feature matrix and a capsule network (CapsNet). The Database for Emotion Analysis using Physiological Signals (DEAP) dataset was employed to validate the the proposed emotion recognition framework. From DEAP, the EEG dataset of signal frequency 512 Hz recorded with 32 electrodes were acquired. In the preprocessing stage, (i) the acquired EEG signals were down-sampled from 512 Hz to 128 Hz, (ii) the Electrooculogram (EOG) effects removed and, (iii) band-pass filtering implemented with cut-off frequencies set at 4.0 and 45.0 Hz.

For the attention dataset, 2-channel EOG signals below and right of the right eye were acquired and the ocular artifacts removed using least-mean square (LMS) based adaptive filtering. The EEG signals in the valence and relaxation datasets were not severely contaminated by the eye saccadic artifacts and hence LMS filter not applied. For removing the ocular artifacts, the MSDW algorithm been applied as that doesn't require any additional EOG signal.

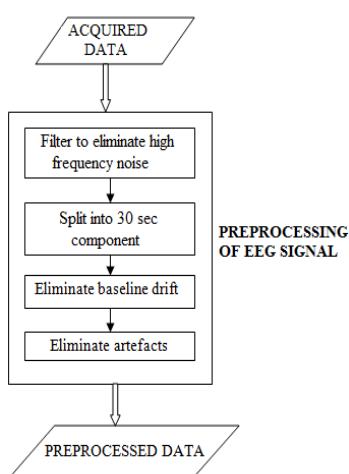


Fig 3. Flow diagram showing Signal preprocessing

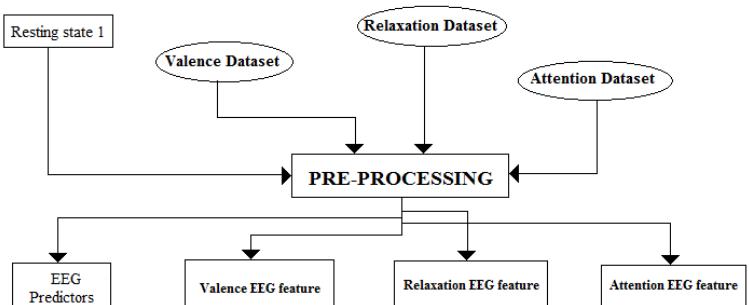


Fig 4. Flowchart showing the pre-processing

Narusci S Bastos et al. [14] proposed the EEG analysis based on two different band frequencies in EEG (full band and Beta band) through data mining. In this work, Narusci et al. focused in analyzing the Beta frequency band as this band is particularly associated with attention, visual precision, and coordination state of human. Here in the preprocessing, software is used and the steps are - (1) Conversion of raw EEG data (recorded without using filters)

in graph data format (GDF) to comma-separated value (CSV), (2) Data labeling, (3) unwanted data cleaning and (4) Transformation of CSV data into Attribute- Relation File Format (ARFF), a machine learning format. Again, in another case study Narusci et al. [14] used a filter to pick out the Beta frequency band ie, between 13 Hz to 30 Hz.

Ho- Seung Cha et al.[16] proposed a method to anticipate individual user's dynamic range of the EEG features and this helps in determining individuals who need additional calibration of the features. Initially, four EEG datasets are acquired - Resting state EEG (RS-EEG) dataset, the valence dataset, the relaxation dataset and the attention dataset. Here in the pre-processing stage all the datasets excluding attention dataset were down-sampled to 256 Hz, and the Multi-window Summation of Derivatives within a Window (MSDW) algorithm is applied in which (a) the down-sampled EEG data band pass-filtered at 0.5–30 Hz, (b) segmented into a series of short segments (c) a sliding window introduced with 1s length and an overlap of 50%. (d) by means of visual examination the eye blink artifacts removed.

Alexander Craik et al.[30]in his study specified that the study of various artifact removal did not address any specific artifact removal process. The foremost frequent artifact-removal algorithms used were independent component analysis (ICA) and discrete wavelet transformation (DWT).

TABLE II. SIGNAL FILTERING METHODS PROPOSED IN THE LITERATURE

Filter	Type of Artifact removed	Advantages	Drawbacks
Low pass filter	Noise	Eliminates high frequency noise	Significant frequency components may be lost.
FIR Band pass filter	Noise	Allows only the selected band of frequency to pass through	
Elliptical filter	Particular band of frequencies	To extract information from μ wave and β rhythm of EEG Signal	Results in accuracy problems
Multichannel Weighted Weiner filter	Eye blink artifacts	5% better results compared to PCA and ICA	Not employed in real medical devices so far.
LMS based adaptive filter and MSDW algorithm	Eye saccadic artifacts and ocular artifacts	LMS – for removing eye saccadic artifacts and MSDW – for removing ocular artifacts	Convergence rate is low so that cannot be used in real time systems.

B. EEG Signal Decomposition

Gen Li et al.[13] idealized a maximum marginal approach on EEG signal pre-processing. In the preprocessing stage , the wavelet transform is used to

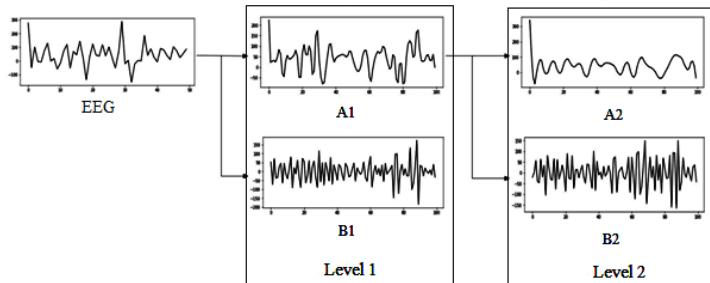


Fig 5. Decomposition of EEG at two levels

break up the EEG data. At level 1, the EEG signal decomposed into the approximate component and detail component. The detail component at level 2 been calculated by applying wavelet transform on the approximate component at the level 1 component.

Min Kang et al.[15] proposes a deep-asymmetry method to convert brain asymmetry feature into a matrix-type image which is then fed to a convolution neural network. The study uses the EEG data set, an open access data set at the Hospital Universiti Sains Malaysia (HUSM). Elimination of power line noise done by 50 Hz Notch filter and filtration of data performed in the range of 0.5–70.0 Hz. The EEG data were collected with a sample rate of 256 samples per second. In the data preprocessing stage, (i) each channel were normalized using the min-max normalization method. (ii) The Independent Component Analysis (ICA) was used to remove other EEG noise. (iii) The data segmentation process by dividing the 5 min data set into epochs of 4s (1024 samples) in order to meet the ML problem needs.

Xiao zhong et al.[20] proposed an EEG signal processing algorithm using ICA-EMD. This concentrates in removing noisy artifacts from EEG signals. (a)By means of ICA, the multichannel EEG recording is decomposed into statistically independent components., and (b) by applying EMD filtering the elimination of physiological artifacts in single-component EEG were achieved. The experiment shows that the ICA-EMD algorithm effectively eliminates the artifacts thereby retaining the useful neural information.

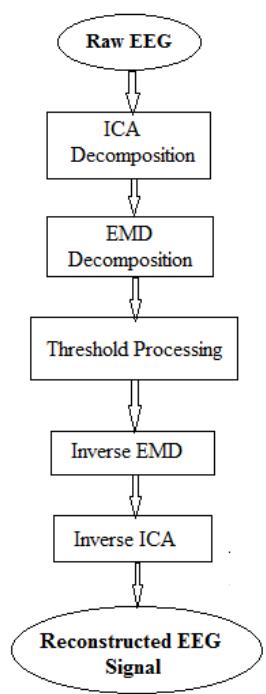


Fig 6. Process flow showing EEG signal preprocessing using ICA-EMD decomposition

C. EEG Signal Processing with Graphic User Interface (GUI)

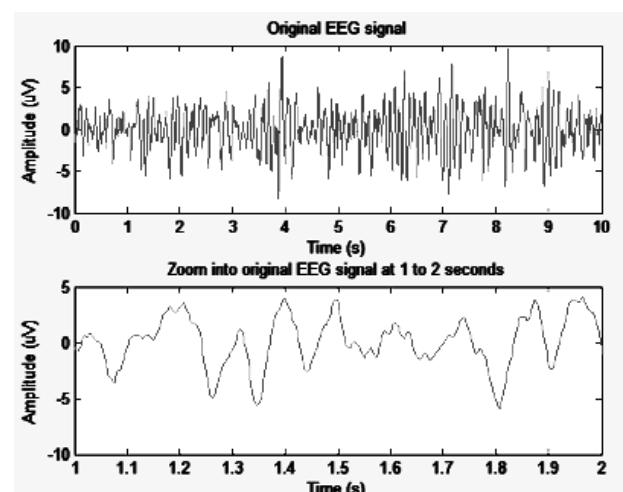


Fig 7. Original and processed EEG signal using MATLAB

Alejandra L. Callara et al.[17] presented a framework to study the brain activity in the Cheyne- Stokes

Respiration patients with the help of EEG signal.. Initially, The acquired data including EEG were sampled at a frequency of 512 Hz and filtered with a two pole anti-aliasing low- pass filter with 3dB at 105 Hz on all channels. With the help of an interactive GUI toolbox, EEGLAB, the EEG data were high-pass filtered above 1 Hz with a wavelet-based filter to improve stationarity. To remove line noise, an adaptive method is utilized which omits the line-noise-related sinusoidal artifacts. After filtering, the identification of bad channels done with the help of a Kurtosis-based approach (Z-score threshold: ± 10) Visual inspection is performed to remove the corrupted data segments which in turn allows proper ICA decomposition.

Haoran Liu et al.,[18] in this paper described the steps involved in the emotion recognition algorithm using EEG while reviewing the existing EEG- based emotion recognition methods along with assessment of their classification effect. The preprocessing of EEG signal is done using EEGLAB which performs the tasks such as channel identification, filtering, baseline correction and independent principal component analysis. The dataset acquired from open datasets of affective computing. After importing EEG data to EEGLAB, filtering the noise in the signal performed by means of suppressing the signal. To remove the electromagnetic interference, the Butterworth band-pass filters are used. EOG artifacts been removed with the help of either regression method or adaptive filtering method.

Ala Hag et al.[19] proposed a novel methodology that identifies the mental stress by comparing the statistical difference between stress and rest conditions. The EEG

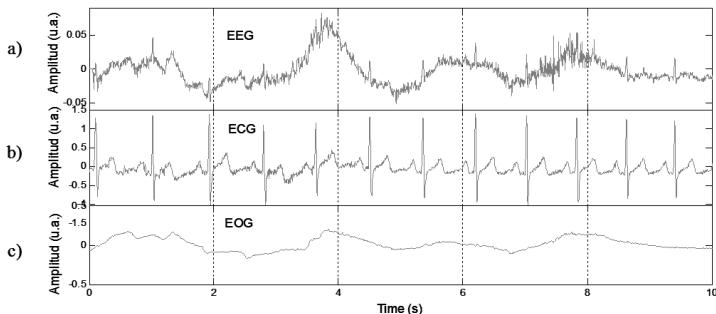


Fig 8. Separation of ECG and EOG signal from the corrupted EEG signal

signal data is acquired and is sampled at 256 Hz frequency. Here in the Preprocessing stage, Python and an external MNE package is used. (i) The unprocessed EEG signals filtered using a band-pass finite impulse response (FIR) filter between 1 Hz and 35 Hz bandwidth. (ii) Power-line noise of 50/60 Hz excluded. Furthermore, (iii) Fast-ICA performed to eliminate the noise ocular artifacts (EOG) under 4 Hz, muscle artifacts (EMG) beyond 30 Hz, and cardiac artifacts (ECG). In frequencies less than 16 Hz, Fast-ICA exhibits an unique ability of de-noising ocular artifacts (OAs).

D. Signal Cutting - EEG

Giuseppe Varone et al.[21] in his paper proposed a data based machine learning (ML) pipeline consisting of a semi-automatic signal processing technique along with supervised ML classifier to aid in the diagnosis of Psychogenic Non-Epileptic Seizures(PNES) by means of EEG data. The

acquired EEG signal is down sampled at 256 Hz and segmented into 20 minute records. While pre-processing the signal, to avoid imbalances the 20 min long records changed to 15 min records. Then the EEG data were manually reviewed to label and reject the artifactual time series such as (i) Eye blinking/ EOG, (ii) Muscular movement /EMG, (iii) Heart rate / ECG, and (iv) electrode artifact. Then the manually reviewed signal were band pass filtered between 1 and 70 Hz with the 3rd order Butterworth band pass filter of 50 Hz to obtain the important EEG rhythms. At the end of the 20 minute recording, clean EEG time series were obtained and segmented into non-overlapping EEG epochs.. Each EEG epoch are of size, $L = 5$ s and therefore 1280 samples obtained. With the help of handwritten Matlab 2018a algorithms, each one of the EEG epochs are preprocessed individually and later stored as .mat files.

E. EEG Processing in BCI system

Mamunur Rashid et al.[22] in his article provided a concise review of EEG-based BCI systems, the pre-processing strategies involved, suggested feature extraction methods, existing classification algorithms, evaluation metrics utilized and this helps the reader to select the appropriate method for a specific BCI system. Mamunur Rashid et al. explains that pre-processing is a non-trivial process, as the process removes any unnecessary components lodged within the EEG signal. Efficient preprocessing enhances signal quality, a better feature separability and thereby classification performance. The primary and effective method to attenuate artifacts when there is no signal overlap in EEG are- low, high, and band pass filters. [13]. In case of spectral overlap, artifact removal techniques such as Adaptive filtering, Wiener filtering, Bayes filtering [23], Common Average Referencing (CAR) [24], and blind source separation (BSS) [25] are practiced. The algorithm SuBAR (Surrogate-based Artifact Removal) removes muscular and ocular artifacts effectively from EEG [26]. When there is limited number of channels, the combination of EEMD-IVA demonstrated and that outperforms other existing methods in a situation. [27].

The joint BSS approach and quadrature regression IVA (q-IVA) removes artifacts effectively in both the time and frequency domains [28]. By combining BSS and Regression (REG) technique, can be used for BCI applications and thus useful for real-time purpose.[29].

IV. RESULTS AND DISCUSSION

It has to be noted that Pre-processing of EEG signal refers to transforming the raw EEG data into a suitable format that can be used for further analysis. And so, various researchers used different pipelines based on their application. In general, EEG tends to record the electrical activity of the brain and is prone to artifacts which is either physiological or non-physiological in nature. Physiological artifacts include EOG, EMG and ECG whereas the main non- physiological artifacts are EMI, 50/60 Hz artifact, cable movements, electrode paste-related and so on.

From the above study, it was stated that the ocular artifacts can be removed effectively by means of band-pass filtering with cut-off frequencies of 4.0 and 45.0 Hz.[12], LMS-based adaptive filtering with 2-channel EOG measurements [16], Fast-ICA for ocular artifacts in low

frequencies less than 16 Hz.[19], SuBAR to remove muscular and ocular artifacts [26], Regression method[18] and Adaptive filtering with stability and fast convergence [18].

In order to remove muscular and cardiac artifacts, Fast ICA with frequency beyond 30 Hz been proposed[19] and the particular time series will be rejected[21] or band pass filtered between 1 and 70 Hz [21]. And the non-physiological artifacts such as EMI by power lines can be removed by implementing differential measurement techniques while recording EEG [4], use of Butterworth band-pass filters [18], filtering in the range of 0.5 to 70.0 Hz with 50 Hz notch filter [15]. For removing line noise, an adaptive method to subtract line noise associated sinusoidal artifacts been deployed[17].The noise caused by 50/60 Hz of line power can also be omitted [19].

For attenuating eye blink artifacts, a multichannel Weighted Weiner filter been presented [11]. Miao Shi et al. uses an elliptical filter to acquire the information from the μ wave and β rhythm of the EEG signal (8-30 Hz)[9].Xiao zhong et al.[20] proposed an ICA-EMD based EEG signal processing algorithm to remove noisy artifacts from EEG signals. In case of eye saccadic artifact attenuation, the MSDW algorithm is proposed [16].

In few other studies, the researchers utilized different methods in addition to artifact removal step in the EEG preprocessing phase. Gen Li et al. utilizes wavelet transform to break down the EEG signal for calculating the frequency components [13].

Naruscic et al. [14] in the preprocessing phase converts EEG data (recorded without using filters) in graph data format (GDF) into ARFF, a machine learning format which is suitable for further data mining software. Min Kang et al. [15] normalized the EEG data using the min-max normalization method. The ICA was used to remove EEG noise.

Ho- Seung Cha et al.[16] by using MSDW algorithm detected time periods with eye blink artifacts and are data segmented. Data segments with eye blink artifacts were excluded. The EEG data were high-pass filtered above 1 Hz with a wavelet-based filter and this improves stationarity. After filtering, the presence of bad channels identified by means of a kurtosis-based approach (Z-score threshold: ± 10), The corrupted data segments were removed by visually inspecting the data. [17].

The preprocessing of EEG signal includes channel location, filtering, baseline correction and principal component analysis using EEGLAB [18].

The experts reviews and reject the artifactual time series such as (i) Eye blinking / EOG, (ii) Muscular movement / EMG, (iii) Heart rate / ECG, and (iv) Electrode artifact [21]. ICA is an effective method for removing artifacts though it disregards the temporal or spatial relations within sources and results in the loss of relevant information. And again due to the less time consumption, CCA can be utilized for real-time applications [30]. As per the principle of BSS algorithms, more number of channels are required in order to be more accurate. For which the wavelet transform and EMD based methods require a single channel but can be decomposed into multiple components. After which the number of channels can be reduced which eventually

increases the computational complexity and this becomes a major limitation for BCI application [30].

From our review, it is clear that ICA-based algorithms deal with all kinds of artifact occurred in EEG recordings. ICA and CCA rather than being alone their combination with other methods seems to be an interesting choice for removal of muscle artifacts. For applications using few channels, EMD, IVA, and its fusion methods with BSS or WT becomes a proper choice. The necessity of reference signal limits adaptive filter or regression methods, to be utilized for the removal of artifacts. Artifacts when overlapped with spectral properties, the wavelet transform is not suitable. EMD suffers from the drawback of mode-mixing. Accordingly, it is strenuous to find a single method that is both efficient and accurate enough to assure all the conditions.

In addition, the automatic methods are not advised for artifact removal, as there are multiple types of artifacts that exist in the recordings. The manual rejection of segment directly neglects the epochs contaminated by artifacts but there is a possibility of losing information from the EEG signal [9]ood electrode placement avoid errors and eliminates artifacts [2].

TABLE III. SUMMARY OF THE EEG PREPROCESSING METHODS

Methods discussed	Purpose
Band pass filtering with cut-off frequencies at 4.0 and 45.0 Hz [12]	Ocular artifacts can be removed
LMS based adaptive filtering [16]	EOG removed
Fast ICA with freq less than 16 Hz[19]	Ocular artifacts in low frequencies less than 16 Hz.
Surrogate based artifact Removal (SuBAR) / Regression method [26]/ Adaptive filtering [18]	To remove muscular and ocular artifacts
Fast ICA with frequency beyond 30 Hz [19] / Band pass filtering between 1 and 70 Hz [21]	To remove muscular and cardiac artifacts
Multichannel Weighted Weiner filter [11]	To attenuate eye blink artifacts
Elliptical filter [9]	To acquire information from μ wave and β rhythm of EEG
ICA – EMD [20]	To remove noisy artifacts from EEG
Multi-window Summation of Derivatives within a Window [16]	To remove eye saccadic artifacts
Wavelet transform [13]	EEG signal decomposed to calculate the frequency components
Data mining software [14]	EEG signal converted to machine learning format – ARFF.
Min-max normalization method [15]	Each EEG channel been normalized followed by ICA to remove noise.
Wavelet filter (High pass filtered above 1 Hz) [16]	To improve stationarity of the Signal.
Kurtosis – based approach [17]	Determines the presence of bad channels.
GUI – Matlab [18]	Performs tasks such as channel location, filtering, baseline correction and ICA.
Wavelet transform and EMD based methods [30]	For Blind Source Separation algorithms involving more number of channels.
Elimination of time series [21]	Experts label and rejects the artifactual time series by review.

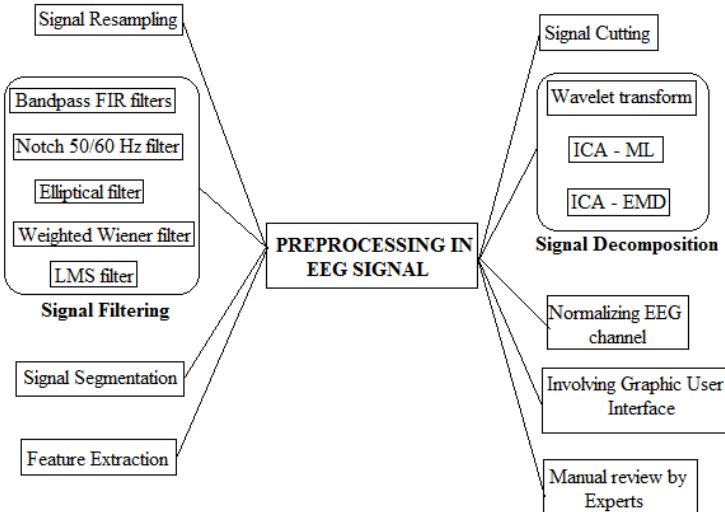


Fig 9. Pre-processing techniques suggested for acquiring useful information from the EEG Signal

V. CONCLUSION

From the above discussion we conclude that it is complicated to find a unique method that fits to satisfy all the conditions perfectly. Among the many artifact removal methods/ algorithms available in the literature, some are used to remove only particular artifacts such as EOG, ECG, and EMG. In some cases, the methodology requires reference channel to enhance the accuracy. Again when it comes to online applications the methods which are highly complex is not suitable. Hence, from the above discussion, we conclude that there is no exceptional choice to remove all types of artifacts from an EEG signal. Such published techniques were proposed by progressing existing algorithms or by combining various methods or making removal process a real-time process by means of machine learning. As EEG preprocessing is still an ongoing area of research study, there is no comprehensively approved EEG preprocessing pipeline, and that researchers have some privilege in choosing how to transform the raw data. Good preprocessing increases the signal quality, and thereby results in enhanced feature separability and better classification performance [22]. Apart from this, as a precaution the artifact avoidance can be tried to avoid unnecessary motion.

REFERENCES

- [1] Radüntz T, Scouten J, Hochmuth O, Meffert B (2017), "Automated EEG artifact elimination by applying machine learning algorithms to ICA-based features", Journal of neural engineering, 14(4), 046004. <https://doi.org/10.1088/1741-2552/aa69d1>.
- [2] C'D Binnie, P F Prior (1994) "Electroencephalography", Journal of Neurology, Neurosurgery and Psychiatry; 57(11): 1308-1319, doi:10.1136/jnnp.57.11.1308.
- [3] Van de Steen F, Faes L, Karahan E (2019), "Critical comments on EEG sensor space dynamical connectivity analysis", BrainTopogr32, 643–654, <https://doi.org/10.1007/s10548-016-0538-7>.
- [4] Riiita Salmelin, Sylvain Baillet (2009), "Electromagnetic Brain Imaging", Human brain mapping 30(6), 1753-7, doi: 10.1007/hbm.20795.
- [5] M A Lopez-Gordo, D Sanchez Morillo, F Pelayo Valle (2014), "Dry EEG Electrodes", MDPI Sensors, 14,12847-12870, doi:10.3390/s140712847.
- [6] Croft R J, Barry R J (2000), "Removal of ocular artifact from the EEG: a review", Neurophysiologie clinique = Clinical neurophysiology, 30(1), 5–19. [https://doi.org/10.1016/S0987-7053\(00\)00055-1](https://doi.org/10.1016/S0987-7053(00)00055-1).
- [7] Pion-Tonachini, Luca (2019), "ICLabel: An automated electroencephalographic independent component classifier, dataset, and website", NeuroImage vol. 198: 181-197, doi:10.1016/j.neuroimage.2019.05.026.
- [8] Chaumon M, Bishop D V, Busch N A (2015), "A practical guide to the selection of independent components of the electroencephalogram for artifact correction", Journal of neuroscience methods, 250, 47–63. <https://doi.org/10.1016/j.jneumeth.2015.02.025>.
- [9] Wu Wen (2021), "Sleep Quality Detection Based on EEG Signals Using Transfer Support Vector Machine Algorithm", Frontiers in Neuroscience, Vol.15, pp.370, doi:10.3389/fnins.2021.670745.
- [10] Miao Shi, Chao Wang, Xian-Zhe Li, Ming-Qiang Li, Lu Wang, Neng-Gang Xie(2021),"EEG signal classification based on SVM with improved squirrel search algorithm", Biomedical Engineering / BiomedizinischeTechnik, Vol. 66(2), pp.137-152. <https://doi.org/10.1515/bmt-2020-0038>.
- [11] Manoj Prabu Manoharan and V.R.Sarma Dhulipala (2020), "Improved energy efficient design in software defined wireless electroencephalography sensor networks (WESN) using distributed architecture to remove artifact", Computer Communications,152, 266 - 271, doi:10.1016/j.comcom.2019.12.056.
- [12] Hao Chao, Liang Dong, Yongli Liu and Bao-Yun Lu (2019), "Emotion Recognition from Multiband EEG Signals Using CapsNet", Sensors (Basel, Switzerland), Vol.19(9), doi:10.3390/s19092212.
- [13] Gen Li and Jason J Jung (2020), "Maximum Marginal Approach on EEG Signal Preprocessing for Emotion Detection", Applied Sciences 10, no. 21: 7677, <https://doi.org/10.3390/app10217677>.
- [14] Narusci S Bastos, Bianca P Marques, Diana F Adamatti, Cleo Z Billa (2020), "Analyzing EEG Signals Using Decision Trees: A Study of Modulation of Amplitude", Computational Intelligence and Neuroscience, vol. 2020, Article ID 3598416, 11 pages, <https://doi.org/10.1155/2020/3598416>.
- [15] Min Kang, Hyunjin Kwon, Jinhyeok Park, Seokhwan Kang and Youngho Lee (2020), "Deep-Asymmetry: Asymmetry Matrix Image for Deep Learning Method in Pre-Screening Depression", Sensors (Basel, Switzerland), Vol.20, doi:10.3390/s20226526.
- [16] Ho-Seung Cha & Chang-Hee Han and Chang-Hwan Im (2020), "Prediction of Individual User's Dynamic Ranges of EEG Features from Resting-State EEG Data for Evaluating Their Suitability for Passive Brain-Computer Interface Applications", Sensors (Basel, Switzerland), Vol 20, doi:10.3390/s20040988.
- [17] Alejandro L. Callara, Maria Sole Morelli, Valentina Hartwig, Luigi Landini, Alberto Giannoni, Claudio Passino, Michele Emdin and Nicola Vanello (2020), "Ld-EEG Effective Brain Connectivity in Patients With Cheyne-Stokes Respiration", IEEE Transactions on Neural Systems and Rehabilitation Engineering, Vol.28 (5), 1216-1225, doi:10.1109/TNSRE.2020.2981991.
- [18] Haoran Liu, Ying Zhang, Yujun Li and Xiangyi Kong (2021), "Review on Emotion Recognition Based on Electroencephalography", Frontiers in Computational Neuroscience. Vol.15, doi:10.3389/fncom.2021.758212.
- [19] Ala Hag, Dini Handayani, Thulasyamall Pillai, Teddy Mantoro, MunHou Kit and Fares Al-Shargie (2021), "EEG Mental Stress Assessment Using Hybrid Multi-Domain Feature Sets of Functional Connectivity Network and Time-Frequency Features", Sensors (Basel, Switzerland), Vol.21(18), <https://doi.org/10.3390/s21186300>.
- [20] G Xiao zhong, L Dezhi and Z Xintong (2020), "Research on EEG signal preprocessing based on ICA fusion EMD," 2020 International Conference on Virtual Reality and Intelligent Systems (ICVRIS), pp. 450-453, doi: 10.1109/ICVRIS51417.2020.00113.
- [21] Giuseppe Varone, Sara Gasparini, Edoardo Ferlazzo, Michele Ascoli, Giovanbattista Gaspare Tripodi, Chiara Zucco, Barbara Calabrese, Mario Cannataro and Umberto Aguglia (2020), "A Comprehensive Machine-Learning-Based Software Pipeline to Classify EEG Signals: A Case Study on PNES vs. Control Subjects", Sensors (Basel, Switzerland), Vol.20(4), <https://doi.org/10.3390/s20041235>.

- [22] Mamanur Rashid, Norizam Sulaiman, P P Abdul Majeed Anwar, Musa RabiuMuazu, Ab Nasir Ahmad Fakhri, Bari BiftaSama, KhatunSabira (2020), "Current Status, Challenges, and Possible Solutions of EEG-Based Brain-Computer Interface: A Comprehensive Review", *Frontiers in Neurorobotics* Vol. 14(25), doi:10.3389/fnbot.2020.00025.
- [23] K T Sweeney, T E Ward and S F McLoone (2012), "Artifact Removal in Physiological Signals—Practices and Possibilities," *IEEE Transactions on Information Technology in Biomedicine*, vol. 16, no. 3, pp. 488-500, doi: 10.1109/TITB.2012.2188536.
- [24] Zaizullyas M, PutehSaad and Imran Ahmad M (2015), "A survey of analysis and classification of EEG signals for brain-computer interfaces," in 2nd International Conference on Biomedical Engineering (ICoBE) (Penang), 1–6, doi: 10.1109/ICoBE.2015.7235129.
- [25] Naoya Oosugi, Keiichi Kitajo, Naomi Hasegawa, Yasuo Nagasaka, Kazuo Okanoya, and Naotaka Fujii (2017), "A new method for quantifying the performance of EEG blind source separation algorithms by referencing a simultaneously recorded ECoG signal", *Neural Networks*, 93, 1–6, <https://doi.org/10.1016/j.neunet.2017.01.005>.
- [26] M Chavez, F Grosselin, A Bussalb, F De VicoFallani and X Navarro-Sune (2018), "Surrogate-Based Artifact Removal From Single-Channel EEG," in *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, Vol. 26(3), pp. 540-550, doi: 10.1109/TNSRE.2018.2794184.
- [27] Xueyuan Xu, Xun Chen and Yu Zhang (2018), "Removal of Muscle Artifacts from Few-Channel EEG Recordings Based on Multivariate Empirical Mode Decomposition and Independent Vector Analysis", *Electronics Letters*, Vol 54, doi:10.1049/el.2018.0191.
- [28] S Lee, M J McKeown, Z J Wang and X Chen (2019), "Removal of High-Voltage Brain Stimulation Artifacts From Simultaneous EEG Recordings," in *IEEE Transactions on Biomedical Engineering*, Vol. 66(1), pp. 50-60, doi: 10.1109/TBME.2018.2828808.
- [29] Guarnieri R, Marino M, Barban F, Ganzetti M and Mantini D (2018), "Online EEG artifact removal for BCI applications by adaptive spatial filtering", *Journal of neural engineering*, 15(5), 056009, <https://doi.org/10.1088/1741-2552/aacfdf>.
- [30] Alexander Craik, Yongtian He and Jose Contreras-Vidal (2019), "Deep learning for Electroencephalogram (EEG) classification tasks: A review", *Journal of Neural Engineering*, Vol.16, doi:10.1088/1741-2552/ab0ab5.
- [31] Shailaja Kotte and J R K Kumar Dabbakuti (2020), "Methods for removal of artifacts from EEG signal: A review", *Journal of Physics.: Conference Series*, 1706, <https://doi.org/10.1088/1742-6596/1706/1/012093>

TRACK 6

**Ninth International Conference in Biosignals, Images and Instrumentation,
SSNCE, Chennai, March 16 – 17, 2023**

Wearable Device for Early Detection of Muscle Cramp with Heating Pad

Prasanna K M

*Department of Biomedical Engineering
Dr.N.G.P.Institute of Technology
Coimbatore, India
prasannakm0@gmail.com*

Raagini R

*Department of Biomedical Engineering
Dr.N.G.P.Institute of Technology
Coimbatore, India
raagini0710@gmail.com*

Sadhana K

*Department of Biomedical Engineering
Dr.N.G.P.Institute of Technology
Coimbatore, India
sadanakanajaraj@gmail.com*

Revathi J

*Department of Biomedical Engineering
Dr.N.G.P.Institute of Technology
Coimbatore, India
revathi.J@drngpit.ac.in*

Abstract— Muscle cramps are usually characterised by a continuous, painful, and localised involuntary contraction of group of muscles or very particular muscle fibres. Although, the causation is undetermined in several situation. Cramp can occur due to dehydration, overusing of muscle, straining, or even remaining still for extended amount of time. Despite the fact that most muscle spasms are innocuous, some could be caused by a medical condition, such as inadequate blood flow or nerve compression. Muscle cramps must be treated early as they happen during or after exercise, long-distance travel, or both. If not, it will persist for a long time and will mostly impede people's ability to travel, sleep and engage in sports. The proposed project is a proactive step that aims to provide first aid for muscle cramps that may occur during transportation, exercise, and other conditions. Novel wearable systems make it possible to measure these extremely complicated physiological event, expanding their capacity and sustaining their non-invasiveness with the use of moisture sensor, MAX30100, and DHT11 sensors. If the analyzed parameters point to a state to a muscular cramp, the Peltier module will be enabled to deliver a therapeutic dose of heating to the affected muscle.

Keywords— Muscle cramp, Peltier, Involuntary contraction, First Aid, Nerve Compression

I. INTRODUCTION

A painful, involuntary muscle contraction is known as a muscle cramp. It often lasts a few seconds to a few minutes and can harm healthy people as well as people who are already ill. When a cramp strike, the muscles contract excessively and without conscious brain control. Pain receptors near the muscles alert the brain that something is amiss. A muscle cramp, which can happen in a variety of circumstances, despite its brief length, it causes skeletal muscles to suddenly contract involuntarily and with serious distress that is shown in Fig.1.



Fig. 1: Muscle at rest and Muscle at cramp

Cramps have the ability to knock even the best-trained athletes to their knees and forced them out of the game. Due to the stress on the muscles and the resulting dehydration from sweating, cramping is a potential [10]. The loss of electrolytes like sodium [1], potassium, magnesium and calcium while sweating causes muscle cramping [2].

The occurrence of cramps is high; it is estimated that 60% of 65 years or older suffer from cramps frequently [3]. Night time cramps are more severe for women than for males. Hypothyroidism [4], myopathies, Alcoholism, renal failure with dialysis, and excessive perspiration are some of the conditions linked to cramping. Approximately 60% of people with insulin dependent diabetes mellitus experience cramps, compared to 80% of people with adult-onset diabetes. About 50% of pregnant women experience muscle cramps, which are most common in the final three months and at night [5]. More areas of the body that can be affected due to cramp is shown in Fig.2. A change in myoneural junction, obesity, compression of the peripheral nerves, deficient blood to the muscles, or greater use of the muscles in the lower limbs could all be contributing factors.

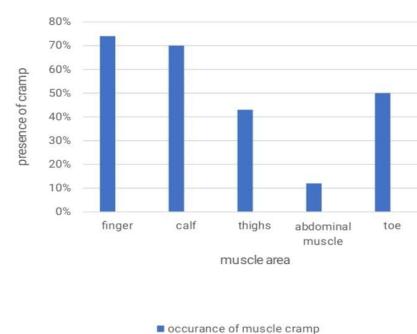


Fig. 2: Presence of Cramps vs Muscle area

Although there is evidence that some occurrences of cramps may be related to irregular monosynaptic activity as a result of the exertion of the affected muscles, quick treatment must be given for cramps that arise in the course of exercise or after activity [5]. If not, it will persist for a long time and will mostly impede people's ability to travel, sleep and engage in sports. The proposed system is a

sensor-based muscle cramp monitoring system packaged as a wearable device with heating pad for treatment.

II. LITERATURE SURVEY

Wing Yin Lau et al. (2021) studied [6] the hypothesis in a cross-over design, ten men are requested to drink ORS and mineral water while ran downward-sloping for 1hour to lose body weight. After DHR, changes in dead weight were measured at specific intervals. The loss in body weight in each interval, the participants ended up drinking either spring water or ORS. The calf muscle cramp liability is determined using a TF of an electrical stimulation. Potassium, magnesium, serum sodium and serum osmolarity and haemato-globulin (Hb) were determined from blood draw in certain time interval. ANOVA method used to compare variation in the above determined variables over time. Thus results in TF increased for ORS condition but decreased from start value soon after DHR for mineral water condition. Soon after-DHR Osmolarity, Blood sodium and chloride concentrations reduced in the mineral water condition. Results were indicated that ORS consumption during exercise reduced the risk of muscle cramps.

Neeru Jayanthi et.al (2020) studied [7] the pathophysiology, risk factors, and potential treatments for EAMC in tennis players. The two main postulated hypotheses for the origin of EAMC are impaired neuromuscular theory, fluid and electrolyte imbalance theory. According to impaired neuromuscular theory, persistent irregular spinal reflex activation accompanied by muscle cramp. Explains how the stretch receptor's afferent activity increased and the Golgi organ's afferent activity dropped, inducing muscle cramp by increasing the muscle's activation and loss of control. Fluid and Electrolyte Imbalance Theory proposed that an electrolyte deficit provokes interstitial fluid migration into the intravascular space, causing neuromuscular junctions become excitable due to malformation and higher sensitive to excitant ECF and chemical transmitter, resulting in immediate ejection. Massage therapy, has proved to be the fast and fruitful method for cramp relief when symptoms of acute EAMC appear.

Han Li, Jun Zhang* et.al (2020) [8], proposed a technique for sensing and controlling muscle temperature with a wearable device It is made of custom radiation fins, platinum resistors, Peltier stack and thermal disperse film. Multiple temperature sensors are placed in the middle of thermal disperse film and thermal control device to detect the temperature at the superficial skin. About 25°C is utilized to control the room temperature, and 10°C is set as the target temperature. The experiment was then carried out by presence and absence of super cooling, and then accumulated the input from sensors. Then the mean was calculated. With the help of PID controller, muscle temperature can be maintained via voltage supply. Current direction through the Peltier array can be changed according to the muscle stimulation for heat and cold therapy. As a result, the response time slows down. Compare to cold stimulation (0°C and 15°C) the system has excellent performance in thermal stimulation because it requires temperatures between 45° and 60°C.

Ajay R, et al. [9] provide a method for the detection of muscle cramp along with the primary therapy in the course of transportation. It is done by the detection of Electromyography and Photoplethysmography then it allows the hospital authorities to exploit these signals. PPG and EMG sensors served an ideal function in this system. The patient must then use vibration to relieve the body's muscle cramps.

III. METHODOLOGY

In existing system, stretching [9], massaging, and applying heat to the injured muscle are the mainstays of the primary treatment for muscular cramps. Other therapies, such as rehydration, calcium supplementation, hormone therapy, electrolyte replenishment etc., target the underlying cause of the muscle cramps. The lack of appropriate primary aid in the transportation industry is one of the key problems with the muscular cramp segment. It is essential to offer the primary treatment for muscle cramp during transportation because they can happen to long-distance travelers. The existing method provide an initiative to detect EMG and PPG signals and to provide primary therapy for muscle cramps during travelling time.

Exercise-related cramps may sometimes be caused by excessive spinal reflex activity as a result of the stress on the affected muscles; hence, immediate treatment for these cramps is necessary. If not, it will remain for a long period and will mostly make it difficult for individuals to travel, sleep, and engage in sports. In the proposed system, to detect muscle cramp feature of lower extremity in sports men, a sweat sensor, heart rate and SpO₂ sensor is placed in wearable device for quickly detecting and assessing. Due to the stress on the muscles and the resulting dehydration from sweating, it is potential for muscle cramping. As a result of a muscular cramp, the heart rate will increase comparatively while the blood oxygenation amount decreases. For the detection of muscle cramp MAX30100, DHT11 and moisture sensors are used for detect the heart rate, temperature and sweat rate respectively as shown in Fig.3. The acquired parameters are then sent as an analog signal to the Node MCU. Then these values are converted into digital values by the Node MCU. If the assessed parameters indicate muscle cramp condition Node MCU activates the heating pad to provide therapeutic dose of 40-45°C heating to the affected muscle. The detected values are stored automatically in the cloud for viewing the values of previous cramp using a web application.

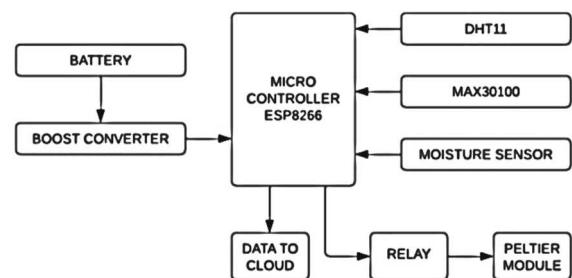


Fig. 3: Block diagram of Hardware Module

A. Hardware Description

1) Node MCU

As a source code platform, modification, redraft and assembling is very convenient in Node MCU. A Node MCU device is similar to an Arduino. ESP8266 is its key component. The pins are programmable. Built-in WIFI is available. It is inexpensive. It can be programmed in a variety of computer languages [10]. The Node MCU helps in editing, modifying and building the hardware component for sensing and it also helps in conversion of analog to digital signal.

The General Purpose Input/Output is accessible via Node MCU. The node MCU receives and processes the sensor-detected values.

In this project, DHT11 sensor, Moisture sensor, MAX30100 sensor, LCD and Relay is connected to the D5, A0, D1, D2, D4 pin of Node MCU respectively.

2) DHT11 Sensor

DHT11 includes an NTC/Termistor for temperature measurement. Fig.4 depicts the internal circuit diagram of the DHT11 sensor. This sensor operates on the thermistor principle, that is resistance changes in response to temperature. Fig.5 represents the Negative Temperature Coefficient, which indicates that resistance decreases as temperature increase. The IC helps in measuring and processing the analog signal by calibration coefficients that can be stored and convert it into digital signal with the temperature [11]. At the time of muscle cramp, temperature of the body will be increased.so it is mainly used for detection of temperature in the body

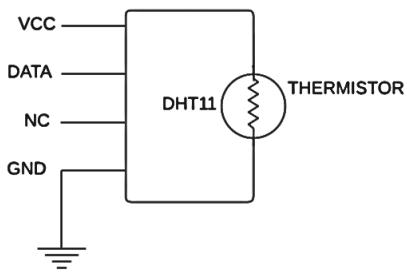


Fig. 4: Circuit Diagram of DHT11 Sensor

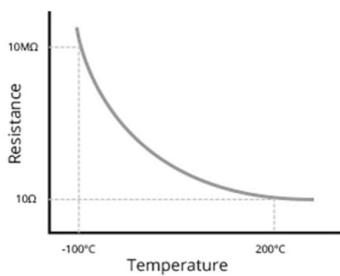


Fig. 5: Thermistor characteristics NTC curve.

3) MAX30100 Sensor

The heart rate sensor is used to monitor pulse rate and oxygenation. This MAX30100 sensor consist of two LEDs, photodiode, improved optics and low noise frequency modulation is shown in fig.6. Heart rate and blood oxygenation in the body are determined with the help of an infrared LED and a red LED [21].

Using light-emitting diodes, the oxygen content is measured. Infrared light is mostly absorbed by the oxyhemoglobin blood where red light passed through the oxyhemoglobin blood. Red light is mostly absorbed by deoxyhemoglobin blood where infrared light passed through the deoxyhemoglobin blood. MAX30100 sensor's photodiode is used to measure changes in light intensity. The photodiode measures the amount of light passing through the blood, sends the signal to an analogue to digital converter. When the muscle cramp strike, pulse rate and O₂ saturation will increased simultaneously.so this sensor helps in detection of heart rate and SpO₂ in the body.

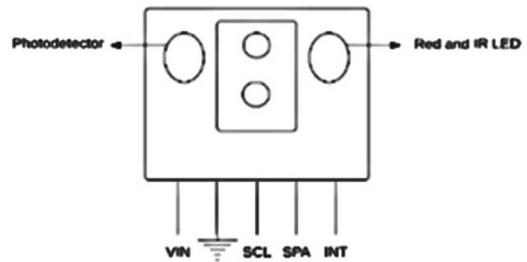


Fig. 6: Circuit Diagram of MAX30100 sensor

4) Moisture Sensor

The sensor which is used for measuring the moisture content in the skin is moisture sensor. The forficate probe like moisture sensor contains two sensitive conductor that helps to measure the moisture level in the skin. When exposed to moisture it functions as potentiometer, changing resistance.

There are four pins present known as AO, DO, VCC and GND. As the AO (Analog Output) provides analogue output voltage proportionate to the level of moisture, a high moisture content causes a more voltage whereas a less moisture content causes a less voltage. A digital output (DO) shows whether the soil moisture level is within the acceptable range. D0 changes from HIGH to LOW depending on whether the moisture level is above the threshold value (determined by the potentiometer). VCC supplies 5V power to the sensor. GND is the ground pin. Dehydration causes muscle cramp. This sensor helps in detection of moisture content that is perspiration rate in the body.

5) Peltier Module

It is a thermal control device that has both heat and cold effects. The surface temperature can be adjusted and maintained at desired temperature by running an electric through the module. Based on the Peltier effect, Peltier modules work. Two ceramic plates, one of which distributes heat and the other of which absorbs it, are separated by semiconductor pallet. By transferring heat between two junctions, the Peltier effect raises the temperature difference between them. This Peltier module is used for heating or maintaining a temperature as electrons transition from a high- to a low-energy state, they carry heat with them and release it on the opposing ("hot") side. This module helps in producing and maintaining heat about 40°C for the treatment of muscle cramp.

IV. RESULTS AND DISCUSSION

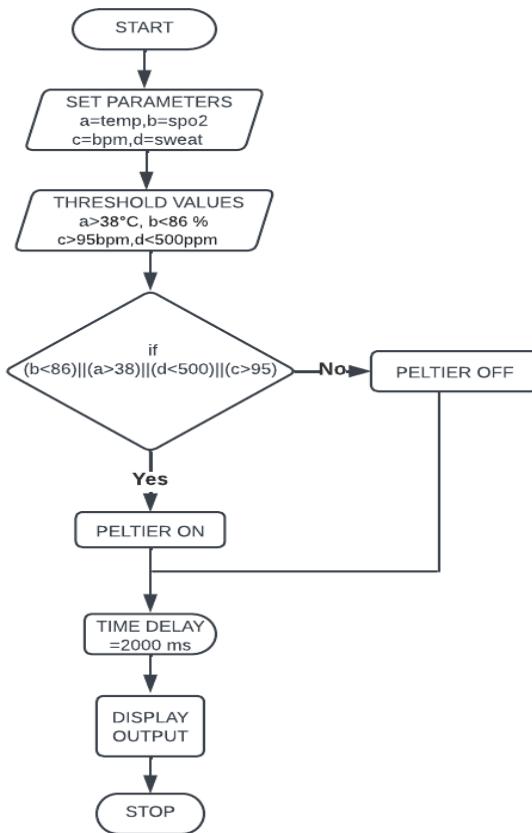


Fig. 7: Flowchart of Proposed System

Fig.7 represents the flow diagram of the proposed technology. The first process of this work is initialization i.e., assigning a variable for every parameter. The next process is about setting the threshold value for each and every parameter and also allotted a time delay, which is used to separate the occurrence of two events. Then the system checks for the condition. If condition is true, Peltier will be ON, if condition is false, Peltier will not turn ON. Finally, these parameter values will be displayed.

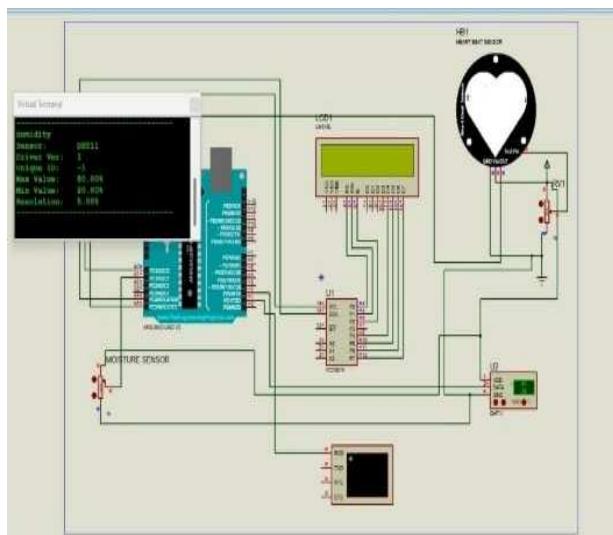


Fig. 8: Simulation Result

Fig.8 shows the wearable muscle cramp monitor's final design, that was simulated to predict its performance and verify that it followed to the stated design goals. The Node MCU is connected to the DHT11, MAX30100 and Moisture sensor. Temperature, heart rate, and moisture content are detected using the above mention sensors. The Node MCU processes these input data and compares them to the threshold level. The I2C LCD screen displays the digitized output. A virtual terminal is added to the controller which represents the ESP8266 that will send the value to the cloud to perform the machine for the later process. The values are displayed in the LCD display.

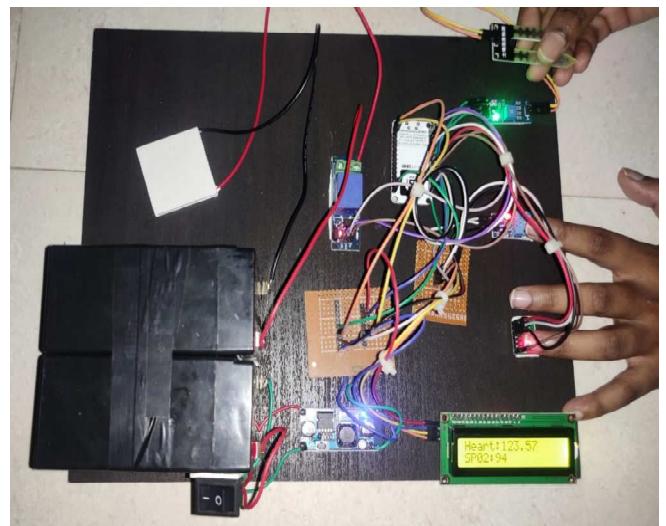


Fig. 9: Hardware output of the proposed system

The hardware output of the proposed system is shown in fig. 9. The various characteristics of human - Heart rate, sweat rate, temperature and spo2 predict the occurrence of cramp. Muscle cramp detected patients treated with heat therapy. Temperature at $40^{\circ}\text{C} - 45^{\circ}\text{C}$ with frequency of $100 - 1000$ Hz gives optimum relief of muscle pain.

TABLE I. NORMAL READINGS

Parameters	Normal	Abnormal
Temperature	$36.1^{\circ}\text{C} - 37.2^{\circ}\text{C}$	$>38^{\circ}\text{C}$
Heart Rate	60 – 100 bpm	>100 bpm
SPO2	96% or more	$< 96\%$

The threshold values of moisture sensor is less than 500 ppm, Temperature is greater than 38°C , Heartrate is greater than 95 bpm, SPO2 less than 86%. The values obtained are shown according to the threshold values.

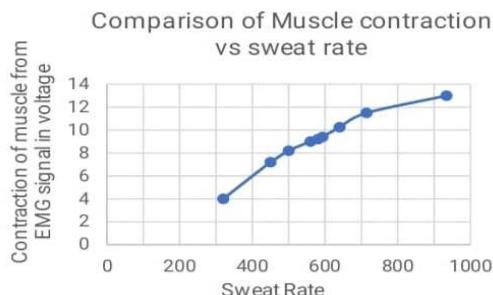


Fig. 10: Comparison of Muscle Contraction vs Sweat rate

With the use of vital parameters, we can determine whether or not a muscle cramp has occurred. Muscle cramps cause an increase in heart rate, drop in oxygen saturation, rise in sweating, and an increase in body temperature. Figure 10 depicts the correlation between sweat rate and EMG signal in a muscle cramp condition.

TABLE II. OBSERVED READINGS

S. no	Heart rate in bpm	Spo2 in %	Sweat rate in mg/l	Temperature in °c	Peltier status
1	89	94%	637.000	37.2	Peltier off
2	87	80%	403.000	37.4	Peltier on
3	96	79%	540.000	37.8	Peltier on
4	75	86%	650.000	36	Peltier off
5	94	76%	480.00	38.2	Peltier on
6	85	90%	640.00	37.7	Peltier off
7	88	99%	590.00	36.2	Peltier off
8	104	88%	490.00	38.4	Peltier on

The observation reveals that subject 5 has low blood oxygen saturation level, high sweat rate and high body temperature. Similar to subject 5, subject 8 likewise notices increase in sweating and an increase in body temperature accompanied by an elevated heart rate. As a result, the Peltier module provide heat before any potential muscular cramps so they can be avoided.

From the above readings it's clear that, when bpm, SPO2, Temperature, Moisture Values are in normal rate, the Node MCU doesn't trigger the relay of Peltier module for treatment. If the parameters exceeds/falls behind its Preset threshold values the node MCU trigger the relay of Peltier module for therapy.

V. CONCLUSION

The implementation of this invention in fields of medicine promotes the real time health care monitoring. This simplest technique reduces the risk of painful, involuntary contraction of muscle. The proposed work is to detect the muscle cramp in an individual with or without any disease and to provide treatment using heating pad while travelling, sleep as well as during physical activities.

REFERENCES

- [1] Miller, K.C., McDermott, B.P. and Yeargin, S.W., 2020. Sweat characteristics of cramp-prone and cramp-resistant athletes. *International journal of sport nutrition and exercise metabolism*, 30(3), pp.218-228.
- [2] Lau WY, Kato H, Nosaka K 2019, Water intake after dehydration makes muscles more susceptible to cramp but electrolytes reverse that effect *BMJ Open Sport & Exercise Medicine*;5: e000478.
- [3] Maughan, R.J. and Shirreffs, S.M., 2019. Muscle cramping during exercise: causes, solutions, and questions remaining. *Sports Medicine*, 49(2), pp.115-124.
- [4] Swash, M., Czesnik, D. and De Carvalho, M., 2019. Muscular cramp: causes and management. *European Journal of Neurology*, 26(2), pp.214-221.
- [5] Bordoni, B., Sugumar, K. and Varacallo, M., 2018. Muscle Cramps.
- [6] Lau, W. Y., Kato, H., & Nosaka, K. (2021). Effect of oral rehydration solution versus spring water intake during exercise in the heat on muscle cramp susceptibility of young men. *Journal of the International Society of Sports Nutrition*, 18(1).
- [7] Troyer, W., Render, A. and Jayanthi, N., 2020. Exercise- associated muscle cramps in the tennis player. *Current Reviews in Musculoskeletal Medicine*, 13(5), pp.612-621.
- [8] Li, H., Zhang, J., Jiang, C., Liu, Q., Zhang, M. and Zhou, J., 2020, October. Muscle Temperature Sensing and Control with a Wearable Device for Hand Rehabilitation of People After Stroke.
- [9] Ajay. R, Akshay P Rajan, Joe Jophy, Christo Varghese, 2018, Relief for Muscle Cramp ness and Detector for EMG and PPG, INTERNATIONAL JOURNAL OF ENGINEERING RESEARCH & TECHNOLOGY (IJERT) ICONNECT – 2018 (Volume 6 – Issue 07).
- [10] Singh, P.Y., 2019. Internet of Things and Nodemcu A review of use of Nodemcu ESP8266 in IoT products. *Journal of Emerging Technologies and Innovative Research (JETIR)*, 1085, p.8.
- [11] Novelan, M.S. and Amin, M., 2020. Monitoring System for Temperature and Humidity Measurements with DHT11 Sensor Using Node MCU. *International Journal of Innovative Science and Research Technology*, 5(10), pp.123-128.
- [12] Evangeline, C.S. and Lenin, A., 2018. Human health monitoring using wearable sensor. *Sensor Review*.

Segmentation and Severity Classification of Dementia in Magnetic Resonance Imaging using Deep Learning Networks

Sarath Vignesh A

Dept. of Electronics Engineering

Madras Institute of Technology, Anna University

Chrompet, Chennai, India

sarathvignesh333@gmail.com

Denicke Solomon H

Dept. of Electronics Engineering

Madras Institute of Technology, Anna University

Chrompet, Chennai, India

denickesolomon2462@gmail.com

Dheepan P

Dept. of Electronics Engineering

Madras Institute of Technology, Anna University

Chrompet, Chennai, India

pdheepan2002@gmail.com

Kavitha G

Dept. of Electronics Engineering

Madras Institute of Technology, Anna university

Chrompet, Chennai, India

kavithag_mit@annauniv.edu

Abstract- Magnetic resonance imaging is the accepted standard for analyzing any deformation in brain. There are many biomarkers which can be considered for analyzing the effect of Alzheimer's disease in brain. One such biomarker is the ventricle which expands during the progression of Alzheimer's disease. Ventricle segmentation plays a vital role in the diagnosis. Automated segmentation approaches are preferred since manual segmentation takes a longer time. In this work, the magnetic resonance images are skull stripped using a combination of Fuzzy C-means clustering and the Chan-Vese contouring technique. Segmentation of ventricle is performed by deep learning architectures, U-Net and Seg-Unet on 1164 transverse MR images acquired from ADNI (Alzheimer's DiseaseNeuroimaging Initiative) database which is an open-source database for carrying researches on Dementia. The features are extracted from the segmented images using ResNet-101 and they are classified using a classifier merger approach which consists of 3 classifiers. The final class label is obtained by majority voting on the individual classifier predictions. The results were compared and analyzed.

Keywords- Alzheimer, FCM, Chan-Vese Contouring, Unet, Seg-Unet, Classifier Merger

I. INTRODUCTION

Dementia [1], is a term that relates to loss in cognitive function which affects the daily activities of an individual. The major cause of dementia in adults of age 65 and older is reported as Alzheimer's disease (AD), which accounts for almost two-thirds of all dementia cases. Alzheimer's disease is a neuro- degenerative disorder that results in gradual degradation of behavioral and cognitive abilities. The clinical stages of AD can be classified into Control Normal (CN), Early Mild Cognitive Impairment (EMCI), Late Mild Cognitive Impairment (LMCI), Alzheimer's Disease (AD) [2]. During the progression of AD, the brain starts to shrink starting in the hippocampus region. The

effect of Alzheimer can be analyzed using many biomarkers in the brain. One such biomarker is ventricle which contains Cerebrospinal Fluid (CSF). In the due course of AD, ventricle enlarges due to the tissue degeneration in the surrounding area [3].

Sinaga et al [4] compared and concluded that proposed K-Means algorithm is fast and generic. In a similar work, Christ and Parvathy [5] concluded Fuzzy C-Means (FCM) clustering technique can be used for tissue segmentation in medical images. This algorithm is derived from fuzzy set theory. Chan and Vese [6] proposed an active contour model to detect the objects in a given image based on curve evolutions and level sets. The algorithm works towards minimization of energy in the neighboringregions by grouping the pixels. Ronneberger et al [7] proposed a convolutional architecture called Unet for segmentation with a limited number of training set images. The model depends more on data augmentation. Badrinarayanan et al [8] introduced a similar model called Seg-Net with a novelty that the decoder upsamples the lower resolution input feature map. Dinthirang et al [9] combined the architectures of Unet and Seg-Net and implemented a hybrid deep learning architecture called Seg-Unet for tumour segmentation in brain MRI. Veluppal et al [10] extracted texture features using kernel density functions whereas Liu et al [11] implemented a hierarchical network to extract node based and edge-based texture features. Cinar et al [12] compared commonly used deep learning models and concluded that ResNet-101 classifies brain tumors efficiently. Shaikh and Ali [13] proposed a novel classifier merger strategy to increase classification efficiency. The ensemble of classifiers provided a very low mean rate and an increased accuracy.

The manuscript is organized as follows, Section II briefs the methodology of the work, the results and the corresponding discussions are provided in Section III and Section IV presents the conclusion.

II. METHODOLOGY

MRI volumes considered in this work for studying the extent of AD were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) Dataset. The acquired dataset is composed of 388 subjects as follows: 97 patients with severity level of normal, 97 patients with severity level of EMCI, 97 patients with severity level of LMCI, 97 patients with severity level of AD. The flow of the work is depicted in Fig.1.

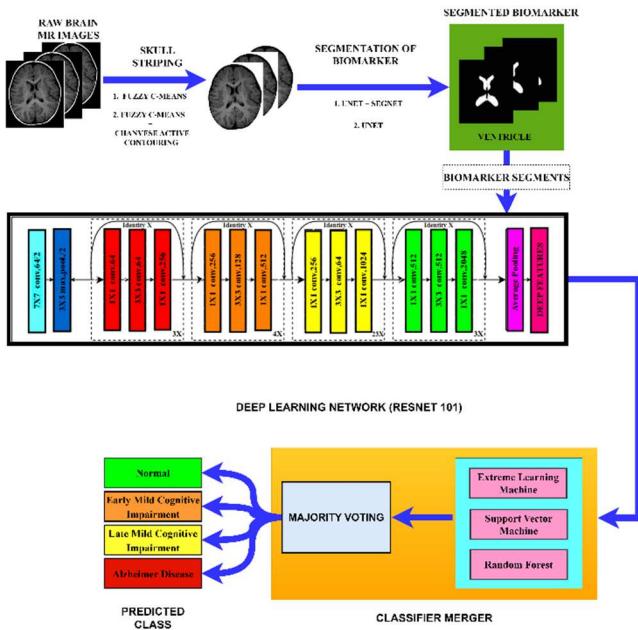


Fig. 1 Flow Diagram

The first process done is skull stripping which is to remove the undesirable skull part from the MRI. Next process is segmentation. Since, biomarker is considered as ventricle, the ventricles are segmented using deep neural network from the skull stripped images. Deep features are extracted from the segmented ventricle images using deep learning network and are classified using classifier merger which contains conventional and deep learning network, with the merging rule as majority voting.

A. Skull Stripping

The removal of the skull from Magnetic Resonance (MR) images is known as skull stripping. From an MRI of the brain, it entails distinguishing brain tissue from various non-brain tissues.

i. Fuzzy C-means:

FCM algorithm works by assigning membership values to each of the data points with respect to each cluster centre. The membership value is assigned on the basis of distance between the cluster centre and the data point.

Smaller the distance between the data point to the cluster centre, it is more likely that the data belongs to that cluster centre. The multiplicative lower frequency bias field estimation is used to modify fuzzy c-means clustering. The assigning of membership values allows a pixel to belong to multiple clusters. FCM operates by minimizing the object function:

$$J = \sum_{i=1}^n \sum_{j=1}^m u_{ij}^q d(x_i, c_j) \quad (1)$$

where the fuzzification of clustering is controlled by 'q', the fuzzy membership of data x_i to the cluster with the centroid c_j is u , and d is the distance between the data point and the centroid of cluster j .

ii. Chan-Vese Active Contour without edges:

This algorithm is a method for segmenting regions. It utilizes curve evolutions as well as the level sets. The base for the Chan-Vese algorithm is the Mumford-Shah model, which uses energy function for segmentation. The energy function formulated for the intensity of the u_0 image in the point (x, y) separated by C contour into two regions in the Chan-Vese active contour algorithm is given in equation (2)

$$\begin{aligned} F(c_1, c_2, C) = & \mu \cdot \text{Length}(C) + \nu \cdot \text{Area}(\text{inside}(C)) + \\ & \lambda_1 \int_{\text{inside}(C)} |u_0(x, y) - c_1|^2 dx dy + \\ & \lambda_2 \int_{\text{inside}(C)} |u_0(x, y) - c_2|^2 dx dy \end{aligned} \quad (2)$$

where $\lambda_1 = \lambda_2 = 1$, $\nu = 0$ are fixed parameters and μ must be greater than or equal to 0.

B. Segmentation

Semantic segmentation is popularly used to group individual pixels in an image belonging to the same class together. It is basically a pixel-level prediction since each pixel in an image is classified according to a class.

i. Unet:

Unet [7] is a fully linked Conventional Neural Network that successfully classifies images into semantic groups. The autoencoder network, which converts inputs to outputs, is the basis of the U-Net architecture. U-Net has two paths, an encoder path that contracts and a symmetric expanding path that expands, both of which are similar to an autoencoder network (decoder). The context of the input image is captured by U-Net encoder's path. This path is nothing more than a convolutional and pooling layer pipeline. Transposed convolutions are used in the decoder path to enable accurate localization.

ii. Seg-Unet:

The Seg-Unet architecture is a hybrid structure that combines features from the SegNet15 and U-Net architectures, both of which are widely used for image segmentation. In comparison with Unet,

each convolution block in SegNet has one more convolutional layer. It is included to match the third-dimension feature map size. In the first convolution blocks of SegNet15, a skip connection inspired by U-Net is added. The skip connection allows features that were already captured in the corresponding encoder to be concatenated in the decoder, allowing for image reconstruction. This also eliminates the chance of loss of local information and promotes consolidation of finer contextual information at the up-sampling layer of the encoder. Each max-pooling layer loses some information during down-sampling. The depth concatenation layer is used to implement the skip connection. The use of batch normalization between the convolutional layers in the Unet model constitutes the Seg-Unet model.

C. Feature Extraction

The key reason to go for feature extraction instead of direct classification is the reduction in the time taken in the whole process. Though the time taken is reduced same level of accuracy can be achieved. Transfer learning (TL) is the idea of training a model with fewer data. The target model does not have to be trained in TL from scratch. Using the idea of transfer learning, Fig.2 illustrates the model learning process. The pretrained ResNet-101 model is retrained on the segmented ventricle images and the skull stripped whole brain images. The deep features are extracted from the layer before the fully connected layer which is in the form of a vector of dimension $N \times 2048$.

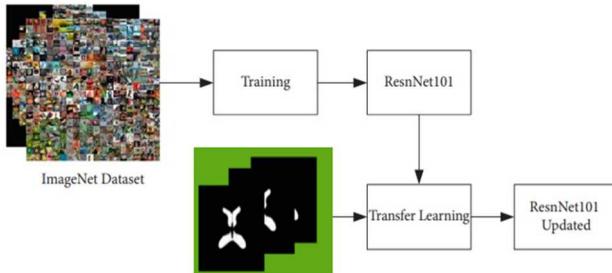


Fig.2 Transfer Learning applied to segmented ventricle images

D. Classification

Classification is the process of predicting the label of a particular image or a feature. This work utilizes a classifier merger approach for classification. The inputs to the classifiers are the deep features extracted using ResNet-101.

Classifier merger is a fusion of three individual classifiers namely Support vector machine (SVM), Random Forest (RF) and Extreme Learning Machine (ELM). The output classes from individual classifiers are merged using majority voting and the final class label for the input image is the class which is predicted by majority of the classifiers.

III. RESULTS AND DISCUSSION

Validating the results at each stage is necessary to conclude on the efficiency of the algorithms utilized in the work. The metric that is widely adopted to check the level of similarity between two images is the Dice Similarity Coefficient. It is given by equation (3)

$$\text{Dice Similarity Coeffecient} = \frac{2 * \text{Area of overlap of white pixels}}{\text{Total number of white pixels}} \quad (3)$$

A. Skull Stripping

Brain extraction tool is capable of skull stripping the images to a certain extent. It can be used as a ground truth image to be compared with the results of the skull stripping algorithms used in this work. The output images from the Fuzzy C-Means (FCM) and FCM with active Chan-Vese contouring are compared with the ground truth images. The dice score obtained for the algorithms are 78% and 86% correspondingly. The FCM clustering algorithm fails due to the homogeneity in the images obtained from the ADNI dataset. On the other hand FCM with Chan-Vese contouring algorithm works perfectly on the dataset. The algorithm was able to strip all sized skulls. For the purpose of comparison between the two algorithms, a sample set of 100 images were considered.

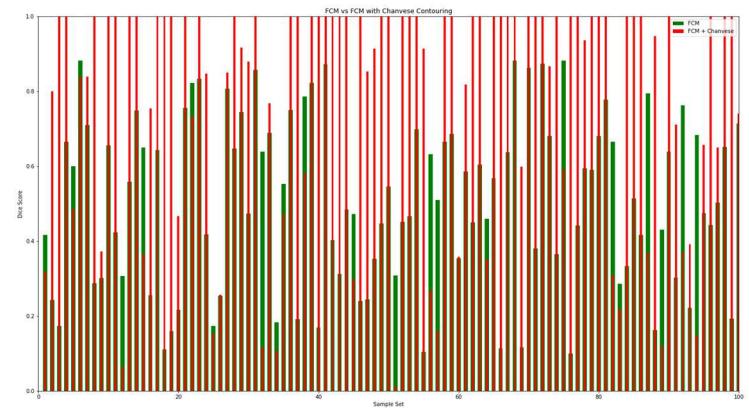
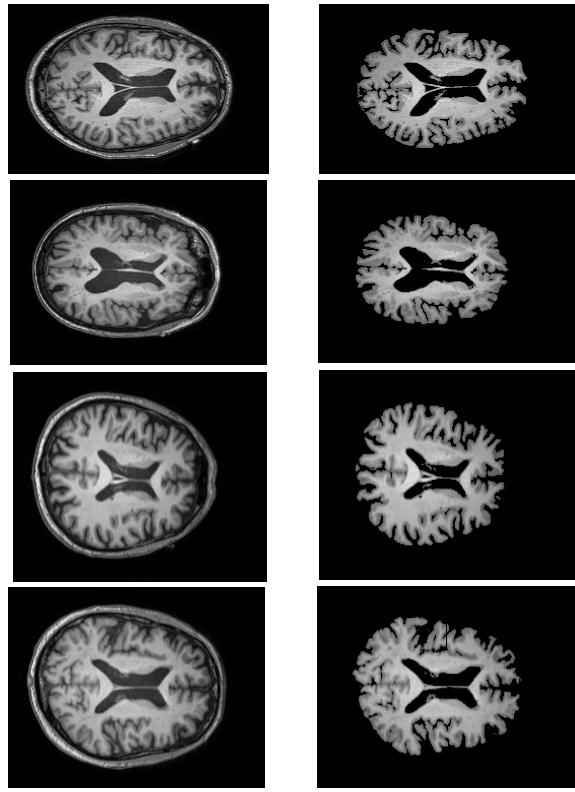
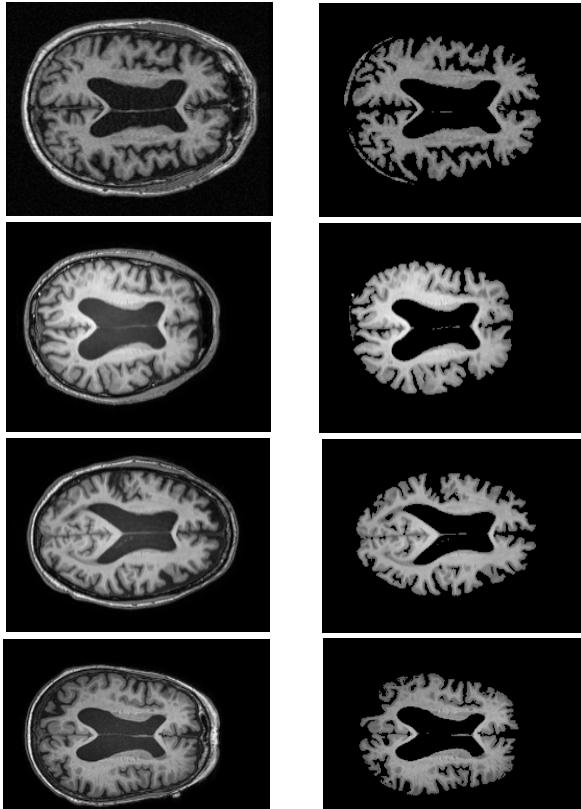


Fig.3 FCM vs FCM-Chan-Vese Contouring algorithm

Fig.3 is a plot between the sample image set on the x-axis and corresponding dice score on the y-axis for FCM and FCM with Chan-Vese contouring algorithms. The Dice score obtained in the sample set by FCM and FCM with Chan-Vese contouring algorithms are 56% and 85.3% respectively. Fig.4 and Fig.5 shows few skull-stripped images from the FCM with Chan-Vese Contouring algorithm for Normal and Abnormal Categories.



a)
Fig. 4 FCM-Chan-Vese contoured Normal brain
b) Original image b) Skull Stripped image



a)
Fig. 5 FCM-Chan-Vese contoured Abnormal brain
b) Original image b) Skull Stripped image

B. Segmentation

The segmentation of biomarker plays a vital role in analyzing the extent of Alzheimer's disease in a subject. The ventricle segmented using Unet and Seg-Unet deep learning models are validated with the ground truth images which is manually extracted with the MIPAV open-source tool. Both the models were accurate in segmenting the ventricle with Seg-Unet marginally outperforming the Unet model. Similar to the case of skull stripping, a sample set of images are considered for comparison purpose. The Dice score obtained for the sample set is 89.3% and 91.3% by Unet and Seg-Unet models respectively.

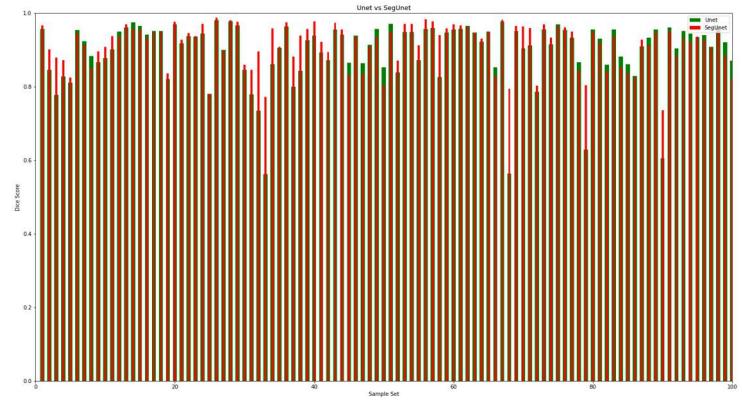


Fig.6 Unet vs Seg-Unet segmentation results

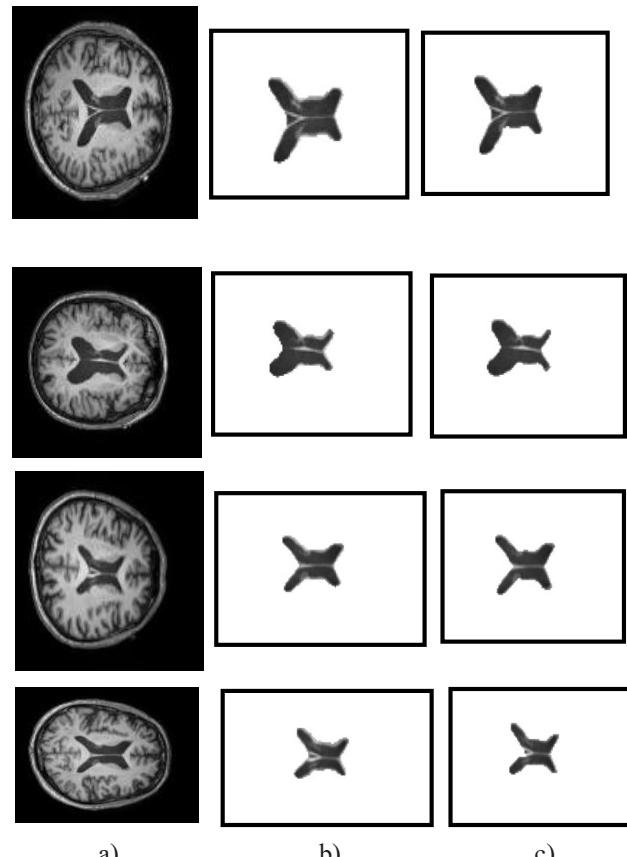


Fig. 7 Segmentation results of Unet and Seg-Unet for normal brain
a) Original image b) Unet result c) Seg-Unet result

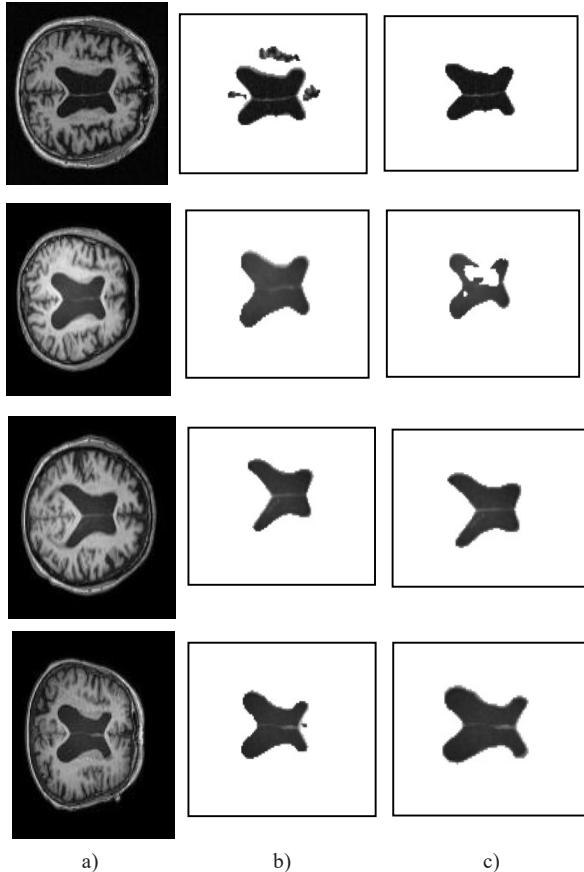


Fig.8 Segmentation results of Unet and Seg-Unet for Abnormal brain
a) Original image b) Unet result c) Seg-Unet result

Fig.6 shows the plot between sample image set on x-axis and dice score on y-axis for Unet and Seg-Unet models. It is quite easy to understand the close margin of the considered models. The Dice scores obtained for the whole dataset is 86.35% and 87.85% by Unet and Seg-Unet models respectively. Fig.7 and Fig.8 shows few images segmented by Unet in comparison with that of Seg-Unet for normal and abnormal classes.

C. Feature Extraction

The deep features are extracted from the Resnet-101 algorithm. The features are extracted before the final fully connected layer. A total of 4096 features from skull stripped whole brain images and its corresponding segmented ventricle images are obtained for a single subject. Fig. 9 portrays the heatmap of sample deep features obtained.

D. Classification

The deep features are fed as input to all three classifiers (SVM, RF and ELM) individually. The final output is obtained by majority voting on the individual classifier results, which is the output class label prediction from the classifier merger.

The Tables I-III shows the confusion matrix obtained from the individual classifiers and table IV shows the confusion matrix of the classifier merger.

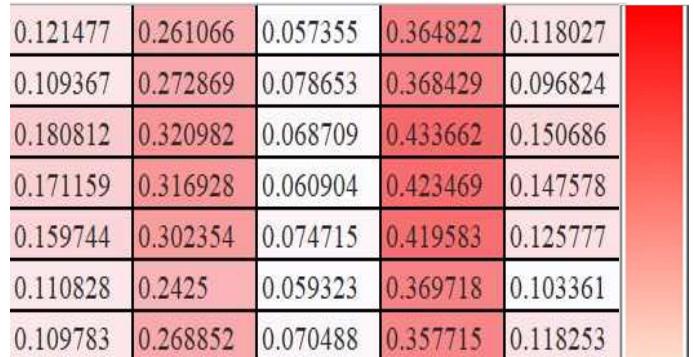


Fig. 9 Heatmap of sample deep features

Table I
Confusion Matrix of SVM Classifier

SVM		Predicted			
		Normal	EMCI	LMCI	AD
Actual	Normal	58	0	0	22
	EMCI	0	73	0	14
	LMCI	0	0	69	26
	AD	0	0	0	91

Table II
Confusion Matrix of RF Classifier

RF		Predicted			
		Normal	EMCI	LMCI	AD
Actual	Normal	88	0	0	0
	EMCI	0	43	0	38
	LMCI	0	0	87	0
	AD	0	0	0	97

Table III
Confusion Matrix of ELM Classifier

ELM		Predicted			
		Normal	EMCI	LMCI	AD
Actual	Normal	88	0	0	0
	EMCI	0	43	0	0
	LMCI	0	27	72	0
	AD	0	0	26	97

Table IV
Confusion Matrix of Classifier Merger

CLASSIFIER MERGER		Predicted			
		Normal	EMCI	LMCI	AD
Actual	Normal	88	0	0	0
	EMCI	0	47	10	0
	LMCI	0	0	85	26
	AD	0	0	0	97

Table V lists the corresponding classifiers and its accuracies. It can be easily understood that classifier merger which works by majority voting on the predictions of individual classifiers is able to provide a better accuracy than the individual classifiers.

Table V
Accuracy comparison between classifiers

CLASSIFIER	ACCURACY
SVM	82.15%
RF	86.95%
ELM	83.22%
MERGER	89.55%

IV. CONCLUSION

MR brain images were obtained from the open source ADNI dataset and a total of 1176 images (97 subjects from each of the 4 classes) were used. Before the biomarker segmentation, skull stripping is performed to remove the skull which is of no significance in the Alzheimer's disease prediction. The skull stripping process is carried out using FCM- Chan-Vese Contouring. The images are then subjected to segmentation by both U-Net and Seg- Unet models. Seg-Unet performed exceptionally in segmenting the biomarker. The segmented biomarker images are then passed through ResNet-101 for deep feature extraction. The concept of transfer learning is applied and the initial weights were taken as the imangenet weights. A total of 2048 features are extracted from a single image. The deep features are provided to three classifiers namely SVM, RF and ELM classifiers. The classifiers are trained using 823 samples and 353 samples is kept for testing. The accuracies obtained with individual classifiers are 82.15%, 86.95% and 83.22% respectively. The accuracy obtained from classifier merger is 89.55%. This increase in accuracy is attributed to the majority voting merging technique. It gives the output class label as the class which is predicted by atleast two of the classifiers.

V. REFERENCES

- [1] Kumar A, Sidhu J, Goyal A, Tsao J.W, "Alzheimer Disease" in Treasure Island (FL): StatPearls Publishing, 2022.
- [2] Breijeh, Zeinab, and Rafik Karaman, "Comprehensive review on Alzheimer's disease: Causes and treatment" in Molecules, vol.25, issue.24, 2020.
- [3] Nestor SM et al., "Ventricular enlargement as a possible measure of Alzheimer's disease progression validated using the Alzheimer's disease neuroimaging initiative database" in Brain, vol 131, 2008.
- [4] K. P. Sinaga and M. S. Yang, "Unsupervised K-means clustering algorithm," in *IEEE Access*, vol. 8, pp. 80716-80727, 2020.
- [5] M. C. J. Christ and R. M. S. Parvathi, "Fuzzy C-means algorithm for medical image segmentation" in 3rd International Conference on Electronics Computer Technology, vol. 2, pp. 108-115, Apr 2011.
- [6] T. Chan and L. Vese, "An Active Contour Model Without Edges" in Scale-Space Theories in Computer Vision, vol. 10, pp.266-277, Sept 1999.
- [7] O. Ronneberger, P. Fischer and T. Brox, "Convolutional networks for biomedical image segmentation" in Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015, vol. 9351, pp.234-241, Nov 2015.
- [8] V. Badrinarayanan, A. Kendall and R. Cipolla, "SegNet: A deep convolutional encoder-decoder architecture for image segmentation", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 39, issue. 12, pp. 2481-2495, Dec 2017.
- [9] D. Daimary, M. B. Bora, K. Amitab and D. Kandar, "Brain tumor segmentation from MRI images using hybrid convolutional neural networks" in Procedia Computer Science, vol. 167, pp. 2419- 2428, June 2020.
- [10] A. Veluppal, D. Sadhukhan, V. Gopinath and R. Swaminathan, "Detection of mild cognitive impairment using kernel density estimation-based texture analysis of the corpus callosum in brain MR images" in Innovation and research in Biomedical engineering, vol. 43, issue. 5, pp. 340-348, 2022.
- [11] Liu J, Wang J, Hu B, Wu FX and Pan Y, "Alzheimer's disease classification based on individual hierarchical networks constructed with 3-d texture features", IEEE Transactions on NanoBioscience, vol.16, issue.6, pp. 428-437, 2017.
- [12] N. Çınar, B. Kaya and M. Kaya, "Comparison of deep learning models for brain tumor classification using MRI images" in 2022 International Conference on Decision Aid Sciences and Applications (DASA), pp. 1382-1385, 2022.
- [13] T.A. Shaikh and R.Ali, "Enhanced computerised diagnosis of Alzheimer's disease from brain MRI images using a classifier merger strategy", International Journal of information technology, issue.14, pp. 1791–1803, Jan 2021.

EXTRACTION, PROCESSING AND ANALYSIS OF SURFACE ELECTROMYOGRAM SIGNAL AND DETECTION OF MUSCLE FATIGUE USING MACHINE LEARNING METHODS

D. V. Pravin

Department of Electronics Engineering
MIT campus, Anna University
Chennai, India
vpdvpravin963@gmail.com

S. Abinesh

Department of Electronics Engineering
MIT campus, Anna University
Chennai, India
abinesh.saravana29@gmail.com

A. J. Ragavkumar

Department of Electronics Engineering
MIT campus, Anna University
Chennai, India
ragavaj03@gmail.com

G. Kavitha

Department of Electronics Engineering
MIT campus, Anna University
Chennai, India
kavithag_mit@annauniv.edu

Abstract— Muscle fatigue is a condition where a muscle or group of muscles lose their ability to contract and generate force. This can happen due to a variety of factors, including prolonged physical activity, lack of oxygen, and depletion of energy stores in the muscle. The raw sEMG signal is extracted by means of gel electrode attached to biceps of right arm. The preprocessing method used in the work involves different order of filters to process the raw signal. Further, the filtered signal is also amplified using instrumentation amplifier. The designed hardware extracts the signal at a frequency range between 56 Hz and 170 Hz. Six statistical features are extracted from the filtered signal in the time domain. The extracted features are given to various trained machine learning models using different algorithms such as Random Forest (RF), Support Vector Machine (SVM) and Logistic Regression (LR). The highest accuracy of about 87.5 % is achieved using random forest algorithm with the precision of 90%. The results that are obtained proves that machine learning methods can be used effectively to detect muscle fatigue from sEMG signals. The proposed method shows the propitious results in terms of accuracy and decisiveness. It can be used in areas such as sports training, rehabilitation, and ergonomics. This complete circuit is easy to produce and implement which could be used in the development of wearable and portable devices.

Keywords— Muscle fatigue, sEMG, machine learning

I. INTRODUCTION

Muscular system generates the electric signal, namely Electromyogram signal, which is essential to analyze the muscular activities, detection of muscular dysfunction or problems with signal transmission between nerve and muscle [1]. Muscle fatigue, which is common muscle

tiredness or is a neuromuscular condition where muscle fails to initiate an activity with the required force. It is difficult to point the reason for muscle fatigue. Neuromuscular fatigue can be grouped under three headings. i) central fatigue ii) fatigue of the neuromuscular junction and iii) muscle fatigue [2]. Central fatigue arises due to prolonged exercise and also due to neurochemical changes in brain [3]. Common diagnosis processes include CT scans, MRI, nerve tests and blood tests. These are done to test the conditions of nerves and to check the signs of infections or other disorders. In this research, sEMG is chosen, which is non-invasive method of data acquisition from subjects [4]. sEMG signal, a bio signal that is generated in muscles during its contraction and relaxation representing neuromuscular activities [5]. Determining fatigue can be accomplished by measuring the force or power produced during a maximum voluntary contraction (MVC) test, which is a test of maximum effort intensity [6]. To understand the activity of muscles under normal and abnormal conditions sEMG is analysed. The acquired signal through the electrodes would be corrupted with noises such as noise from electrical equipment, ambient noise, ECG noise, motion artifacts, etc., [7]. The phenomena of shift of sEMG mean frequency and median frequencies towards lower frequency under isometric conditions to find the fatigue level of muscle [8]. The noise generated must be removed for further analysis. The typical range of sEMG signals is between 20 Hz and 500 Hz. The amplitude of the sEMG signal is concentrated between 0.01 to 5 millivolts (mV) [9]. The acquired signal has to be amplified to the level as the signal lies within the noise level. The circuit should be designed such that it should have high input impedance, CMRR and accuracy. Also, the output impedance and noise factors should be low [10]. The placement of electrodes plays a important role in the exact acquisition of the signal. Once the signal is acquired, feature selection plays significant role in the classification of normal and abnormal signals.

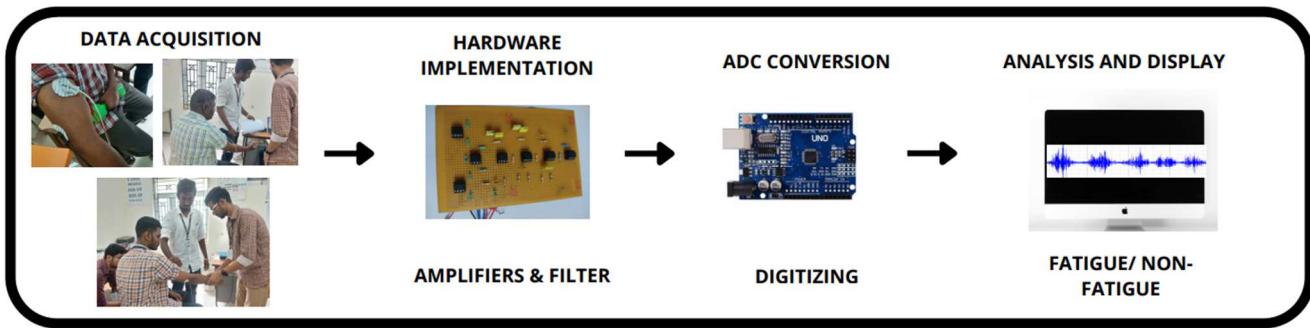


Figure 1 Work flow of the proposed work

II. METHODOLOGY

Figure 1 depicts work flow of overall proposed work. The sEMG is extracted from the subject with electrode attached to cleaned surface of biceps using circuit designed. The signal is digitized using Arduino UNO, and then recorded using computer. It is then fed to trained machine learning model. The machine learning model takes six time-domain features into account and classifies the signal into two conditions namely fatigue or non-fatigue.

A. DESIGN OF HARDWARE

1) INSTRUMENTATION AMPLIFIER

AD620, an instrumentation amplifier which yields high accurate results. It is an 8 pin DIP IC. Its gain can be adjusted up to ten thousands. These amplifiers find its application in precise data acquisition. It presents low noise and input bias current. Due to its compactness, efficiency and low power usage, it is used in a variety of medical devices and applications. It is used in designing differential amplifier, where the extracted sEMG signal from two electrodes are given as an input and the amplified output is the difference between them [11].

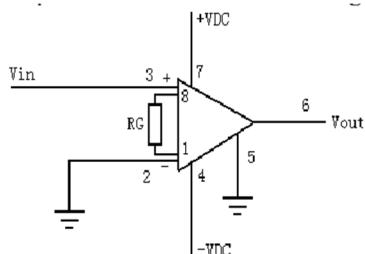


Figure 2 Instrumentation amplifier

Instrumentation amplifier is designed using output voltage gain. Thus, the gain equation of instrumentation amplifier is given by equations (1) and (2).

$$G = \frac{49.4 \text{ k}\Omega}{R_G} + 1 \quad (1)$$

For any arbitrary gain, R_G is

$$R_G = \frac{49.4 \text{ k}\Omega}{G - 1} \quad (2)$$

2) BAND PASS FILTER

OP-Amp OP-07 is used for the purpose of filtering and as a final amplifier [12], which is low noise op-amp and has CMRR of 110 dB. The OP07 is available in 8-pin DIP. The designed filter here is 6th order bandpass filter. The filter gain is set to 10 to amplify the sEMG signal, which can vary from -10 to 10 mV, and utilize the full range of the amplifier without reaching saturation. The design of this 6th order filter is achieved by implementing three 2nd order filters simultaneously. The filter designed here is done using Butterworth.

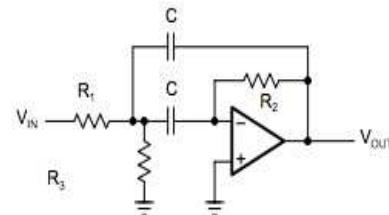


Figure 3 Multiple feed-back band pass filter

Figure 3 depicts the circuit diagram of multiple feedback band pass filter. This filter is designed using equations (3) to (7) by choosing appropriate C_1 :

$$R_1 = \frac{Q_{BP}}{2\pi f_0 A C_1} \quad (3)$$

$$R_2 = \frac{1}{4\pi f_0 Q_{BP} C_1} \quad (4)$$

$$R_3 = R_4 = \frac{\left(\frac{Q_{BP}^2}{Q_{BP}^2 + A} + 1\right) Q_{BP}}{2\pi f_0 C_1} \quad (5)$$

$$C_2 = \left(\frac{A}{Q_{BP}^2} + 1\right) C_1 \quad (6)$$

$$Q_{BP} = \frac{f_0}{f_h - f_l} \quad (7)$$

Here, A is the gain of the considered filter, f_0 is its central frequency. f_l and f_h denotes the low and high frequencies respectively. The quality factor is denoted as Q_{BP} . Increasing the order of filter will result in better noise removal, because of sharper response. Three 2nd order MFB band pass filter stages are used to construct 6th order band pass filter. It is mandatory to find the above mentioned

parameters for the filter design. For a band pass filter of 2nd order, the equations (8) to (14) are used.

$$Q_{BPx} = Q_{BPy} = \sqrt{\frac{\left(k_{LP}^2 + 4Q_{BP}^2 + \sqrt{(k_{LP}^2 + 4Q_{BP}^2)^2 - \frac{4k_{LP}^2 Q_{BP}^2}{Q_{LP}^2}}\right)}{\frac{2k_{LP}^2}{Q_{LP}^2}}} \quad (8)$$

$$f_{0x} = \frac{f_0}{\sqrt{\frac{k_{LP}Q_{BPx}}{2Q_{LP}Q_{BP}} + \sqrt{\left(\frac{k_{LP}Q_{BPx}}{2Q_{LP}Q_{BP}}\right)^2 - 1}}} \quad (9)$$

$$f_{0y} = f_0 \left(\frac{k_{LP}Q_{BPy}}{2Q_{LP}Q_{BP}} + \sqrt{\left(\frac{k_{LP}Q_{BPy}}{2Q_{LP}Q_{BP}}\right)^2 - 1} \right) \quad (10)$$

$$A_x = A_y = \sqrt[n]{A^2} \sqrt{1 + Q_{BPx}^2 \left(\frac{f_0}{f_{0x}} - \frac{f_0}{f_0} \right)^2} = \sqrt[n]{A^2} \sqrt{1 + Q_{BPy}^2 \left(\frac{f_{0y}}{f_0} - \frac{f_0}{f_{0y}} \right)^2} \quad (11)$$

$$Q_{BPZ} = \frac{Q_{BP}}{k_{LP}} \quad (12)$$

$$f_{0z} = f_0 \quad (13)$$

$$A_z = \sqrt[n]{A^2} \quad (14)$$

Here, n is order of the filter, Q_{LP} and k_{LP} is factors acquired from the table of approximation of Butterworth low pass filters of higher order. 3rd order low pass and high pass filter combinedly used to design the 6th order band pass filter. Considering the filter parameters as lower and higher cut-off frequencies are 50 and 150 Hz respectively and gain equal to 10, the results are obtained as follows:

$$Q_{BP1} = Q_{BP2} = 2.1889 \quad (15)$$

$$f_{01} = 64.9679 \text{ Hz} \quad (16)$$

$$f_{02} = 153.9222 \text{ Hz} \quad (17)$$

$$A_1 = A_2 = 4.7158 \quad (18)$$

$$Q_{BP3} = 1; f_{03} = 100 \text{ Hz}; A_3 = 2.1544 \quad (19)$$

Components were chosen carefully such that it produces the minimum difference between the calculated values of components (resistance and capacitance) and the conventional ones by carefully choosing capacitor C₁ value in each filter stage.

3) FINAL AMPLIFIER

The sEMG signal ranges from 0.01 to 5 millivolts (mV). This amplifier has been used to amplify output of filter, as it is difficult to adjust total gain of designed filters. Inverting amplifier is used and gain is defined by equation (Figure 4). The gain is adjusted here in this work by varying resistance values of R_i and R_f.

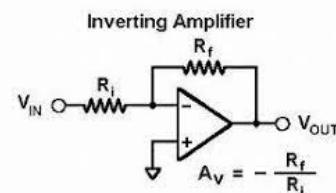


Figure 4 Inverting amplifier circuit

The possible values of resistors designed to amplify to the range of interest are R_f = 30kΩ and R_i = 240kΩ.

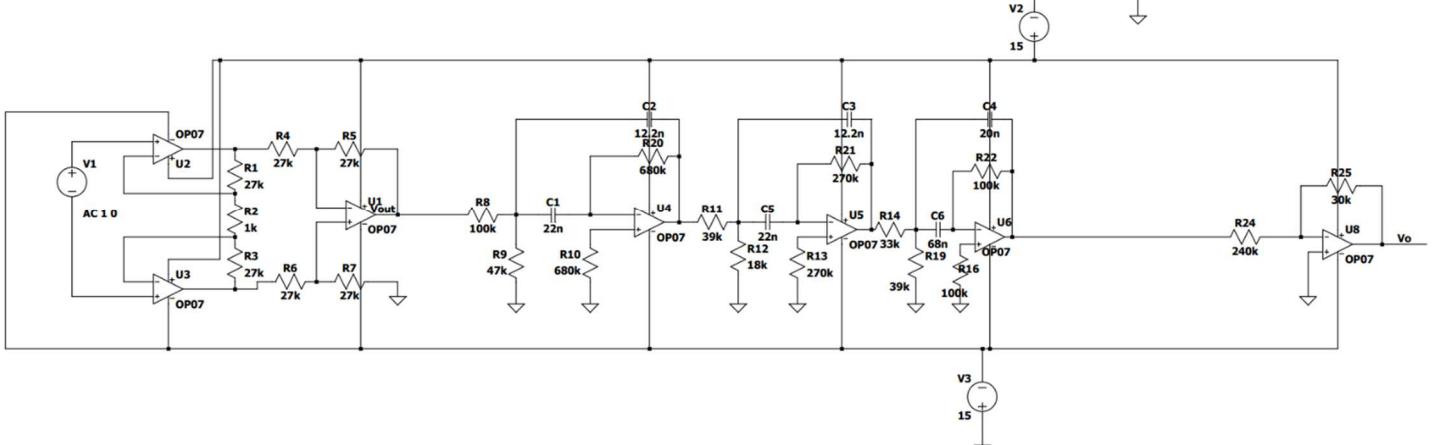


Figure 5 Circuit simulation using LTSpice

Figure 5 shows complete circuit designed for extracting, amplifying and filtering sEMG signal which is simulated in LTSpice XVII.

4) ADC CONVERSION

The final amplified and filtered signal is analog in nature. It has to be digitized for further analysis. So, the signal is given to Arduino UNO for Analog to Digital conversion. It has 10-bit ADC, which converts it into digital form at a baud rate of 9600. It is then stored in the form of '.txt' or '.csv' format. It is then reconstructed using Python programming for feature extraction. The features are extracted from the reconstructed signal and then are stored in '.csv' format for further classification using machine learning classifier models.

B. DATASET COLLECTION

sEMG dataset for training machine learning model is acquired from various subjects using data acquisition device. The participant's demographic data were collected before recording of the individual signals. sEMG signals were recorded using BIOPAC MP36 data acquisition device. In data collection, SS2LB electrode lead set with sensor connector and non-invasive BPL-04 disposable and adhesive Ag/AgCl electrodes (5.00 x 3.5 cm size) were used. The relative location of forearm muscles were determined by physician. The BPL-04 solid gel electrodes were placed and sensors were calibrated. The subjects were asked to contract and relax their biceps with and without dumbbells. To minimize noises, all the data were collected in the same environment. Figure 6 shows placement of electrodes on bicep muscle. Figure 7 shows real time dataset acquisition.



Figure 6 Placement of electrodes

The sEMG data was acquired from the right hand biceps of the subjects. There are two phases in recording process. Firstly, the subjects were asked to contract and relax the biceps, which corresponds to normal condition. Secondly, subjects were given a dumbbell of 3 Kg and were asked to lift it until their arm gets tired, and signal was recorded.



Figure 7 Data acquisition using electrodes

sEMG data was collected and recorded at a sampling frequency of 2 kHz in filtered form. Filtering is done using the sixth-order Butterworth band pass filter. The range of band pass filter is between 5 and 500 Hz with a second order notch filter at 50 Hz in order to eliminate various noises. The noises include motion artifacts, ECG artifacts and high frequency noises, which could be removed using software BSL 4.0 filtering options. The sEMG signals were recorded for duration of 8120 seconds. The amplitude of the sEMG signals was in the range of -5 to 5 mV. Figure 8 shows typical sEMG signal recorded using BIOPAC MP36 DAQ system. From the figure 7, it can be observed that sEMG signal is random bioelectric signal corrupted with noises.

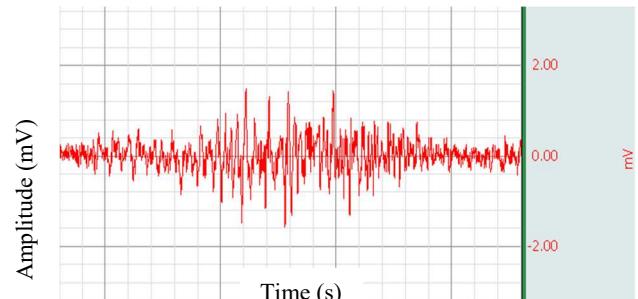


Figure 8 Typical sEMG signal

C. DATA PREPROCESSING

The sEMG signals recorded were filtered with FIR bandpass filter with lower and upper cutoff frequency of 50 and 150 Hz respectively using the filtering option in BIOPAC Software BSL 4.0. This filter applied is to remove noises such as ECG noise, Electrical equipment noise, motion artifacts, and cross-talk. The signals are then used for feature extraction. This signal classification approaches utilize time domain analysis.

E. FEATURE EXTRACTION

Generally, EMG signals are classified using the different statistical features [4]. The features which were extracted are Peak-to-peak amplitude, Mean, Median, Skew, Kurtosis and Standard deviation. The above-mentioned features were recorded in .CSV file for a window of 5 seconds.

The various features are given as follows:

- **MEAN**

$$\bar{Y} = \frac{1}{N} \sum_{i=0}^N Y_i \quad (20)$$

- **SKEW**

$$Skew = \frac{\sum_{i=1}^N (Y_i - \bar{Y})^3 / N}{\sigma^3} \quad (21)$$

- **KURTOSIS**

$$K = \frac{\sum_{i=1}^N (Y_i - \bar{Y})^4 / N}{\sigma^4} \quad (22)$$

- **STANDARD DEVIATION**

$$\sigma = \sqrt{\frac{\sum_{i=1}^N (Y_i - \bar{Y})^2}{N - 1}} \quad (23)$$

- **MEDIAN**

$$M = Y \left[\frac{N+1}{2} \right], \text{ if } N \text{ is odd} \quad (24)$$

$$M = \frac{Y \left[\frac{N}{2} \right] + Y \left[\frac{N}{2} + 1 \right]}{2}, \text{ if } N \text{ is even} \quad (25)$$

III. RESULTS AND DISCUSSION

A. RESULTS OF HARDWARE DESIGNED

Circuit for sEMG extraction is designed using design equations (equations (1) to (14)) and was simulated in LTSpice XVII. Transient response and frequency response for circuit and filter is given below. Figure 9 shows input sEMG signal, filtered sEMG signal, and output signal of the circuit, which is amplified and filtered. Figure 9(a) shows extracted raw sEMG signal from biceps which is corrupted with noises. 9(b) and 9(c) shows output of filter

which is amplified and filtered and output of final inverting amplifier stage which is to be fed to ADC convertor.

Transient response of circuit simulated in LTSpice produces voltages in the range of +10V to -10V after instrumentation amplifier stage, ranges of +14V to -14V after bandpass filter stage and final amplifier produces a maximum of 4.5V only.

The simulated circuit produces frequency response with central frequency of 79.2 Hz and gain of 35.86 dB with higher cut-off frequency of 161 Hz and gain of 32.89 dB and lower cut-off frequency of 59.8 Hz and gain of 32.86 dB.

Transient response of circuit tested produces voltages in the range of +17.8 Vpp after instrumentation amplifier stage, ranges of +27.6 Vpp after bandpass filter stage and final amplifier produces a maximum of 5V only. The higher cutoff frequency was found to be 170 Hz. The lower cutoff frequency is found to be 56 Hz.

Figure 10 shows the frequency response of circuit. It shows that the sharpness of filter increases with increase in order and at six stage band pass filter produced cutoff frequency approximately equal to desired frequencies.

Figure 11 shows the output of designed circuit. The sEMG signal from electrodes is given to circuit and the circuit produced an amplified and filtered sEMG signal which is suitable to classification. The output signal sEMG is free from noises.

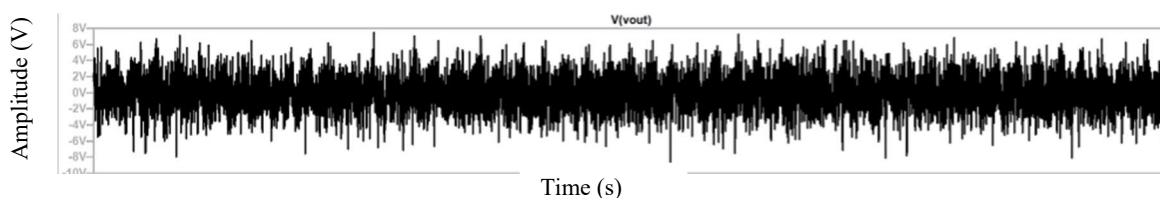


Figure 9 (a) Extracted raw sEMG signal

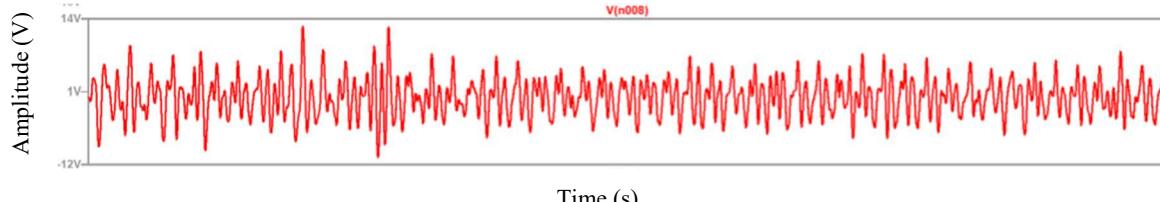


Figure 9 (b) Output of filter

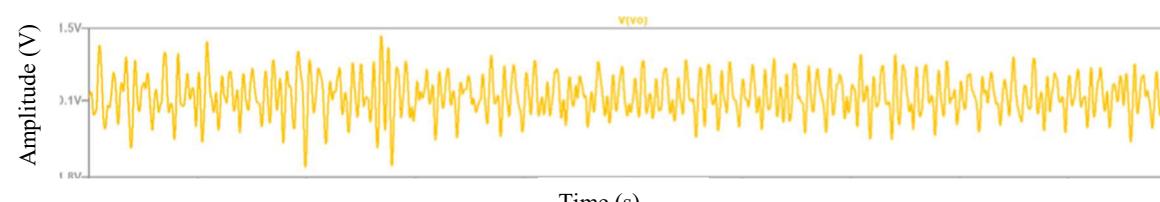


Figure 9 (c) Output of final inverting amplifier

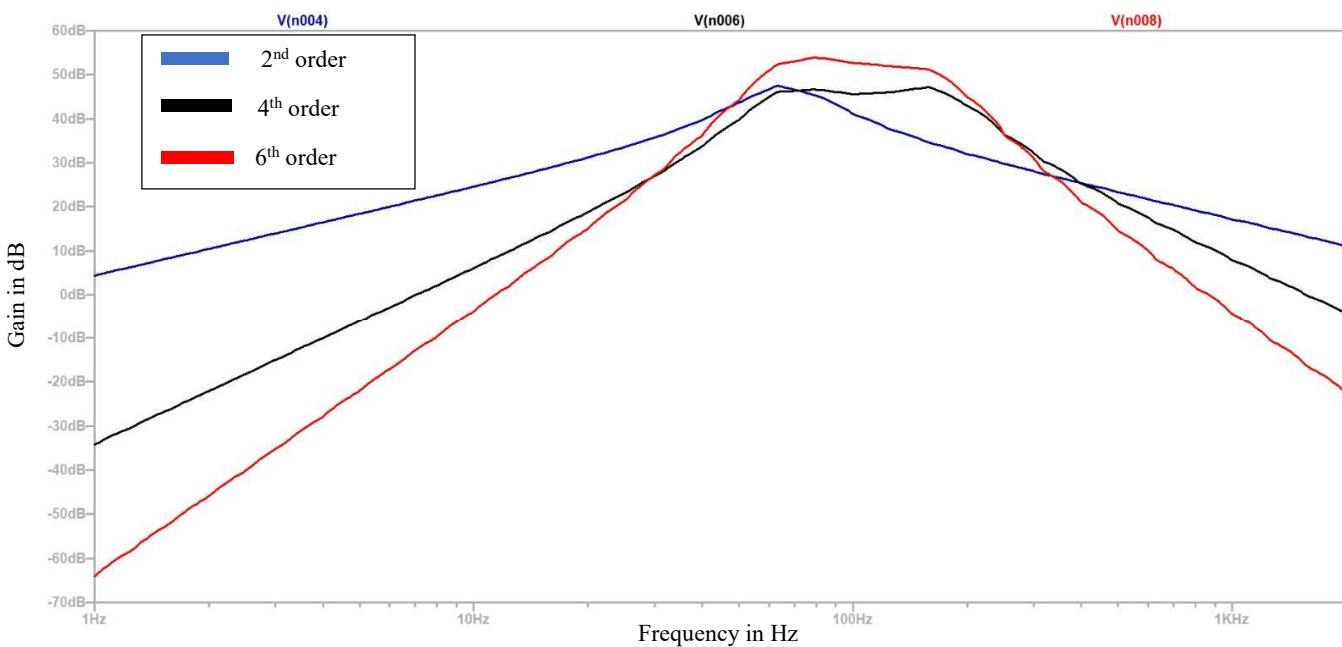


Figure 10 Frequency response of circuit for different order of filters

This signal will be given to ADC convertor and digital signal is received and converted to required type to be given to trained model.



Figure 11 Amplified and filtered sEMG signal output of circuit

B. RESULTS OF DATA PRE-PROCESSING METHOD

A Sample of 50 normal condition signals and 50 fatigue conditioned signals were taken and data pre-processing techniques were applied to filter noise such as ECG noise, Electrical equipments, motion artifacts, cross talk. The required sEMG signals were filtered with FIR band pass filter with lower cutoff frequency and upper cutoff frequency of 50 Hz and 150 Hz respectively using filtering option in software BSL 4.0. Figure 12 shows a typical sEMG signal with noises before applying band pass filter. Figure 13 shows sEMG after removing noises using band pass filter.

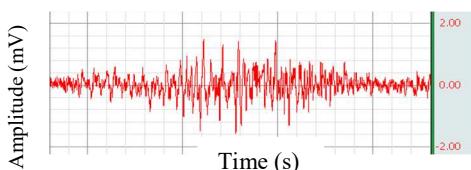


Figure 12 sEMG signal without band pass filter

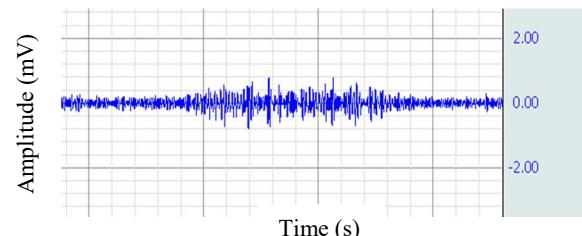


Figure 13 sEMG signal after using band pass filter

C. RESULT OF CLASSIFICATION MODELS

The models were given with test data and performance of various machine learning models on dataset is listed in Table 3.1. Classifier model trained using dataset include Logistic regression, RF and SVM classifiers. Random forest classifier model outperforms all other models in all metrics. Overall accuracy of model is high in Random Forest. In SVM, even though accuracy is low, it has good precision score. Recall score is also high in Random Forest. Thus, Random Forest has good accuracy, high precision and recall score, so it produces less misclassification.

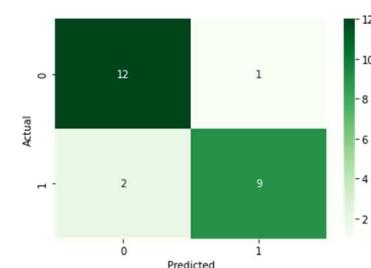


Figure 14 Confusion matrix of Random Forest

Table 1 Comparison of different models and its performance metrics

CLASSIFICATION MODEL	OVERALL ACCURACY (%)	PRECISION (%)	SENSITIVITY (%)	SPECIFICITY (%)
RANDOM FOREST	87.5	90	81.8	92
LOGISTIC REGRESSION	79.16	72.7	80	78.57
SUPPORT VECTOR MACHINE	83.3	87.3	80	90.9

Trained classifier models were given test data. Figures 14, 15 and 16 shows the confusion matrix of RF classifier, Logistic regression and SVM classifier respectively.

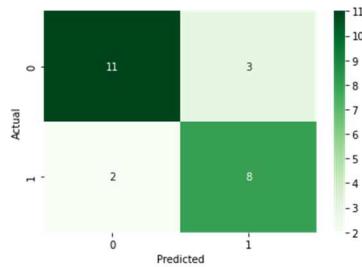


Figure 15 Confusion matrix of Logistic regression

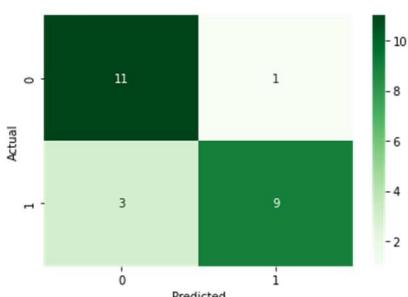


Figure 16 Confusion matrix of SVM

IV. CONCLUSION

In this work, the extraction of sEMG is done using circuit designed, and the detection of muscle fatigue from sEMG signal is done using Logistic regression, RF and SVM classifiers. Original signals were obtained by placing gel or patch electrodes on cleaned surface of biceps with reference electrode was placed on bony part of hand. The circuit designed produced sEMG output which was an amplified and filtered version of raw sEMG signal and it was given to computer for processing. The trained machine learning model uses six-time domain features for classification. The algorithms used were Logistic regression, RF and SVM classifiers. These classifiers performed well for the prepared dataset. Random Forest classifier model with 87.5% overall accuracy, outperforms other models. Support Vector Machine classifier equally

performed well with overall accuracy of 83.3%. Thus, it can be concluded that Random Forest classifier could be used to differentiate sEMG signal for fatigue and non-fatigue conditions.

REFERENCES

- [1] P. A. Karthick and S. Ramakrishnan, "Surface electromyography based muscle fatigue progression analysis using modified B distribution time-frequency features," *Biomedical Signal Processing and Control*, vol. 26, pp. 42–51, 2016.
- [2] M. Cifrek, V. Medved, S. Tonković, and S. Ostojić, "Surface EMG based muscle fatigue evaluation in Biomechanics," *Clinical Biomechanics*, vol. 24, no. 4, pp. 327–340, 2009.
- [3] M. Gruet, J. Temesi, T. Rupp, P. Levy, G. Y. Millet, and S. Verges, "Stimulation of the motor cortex and corticospinal tract to assess human muscle fatigue," *Neuroscience*, vol. 231, pp. 384–399, 2013.
- [4] P. A. Karthick, D. M. Ghosh, and S. Ramakrishnan, "Surface electromyography based muscle fatigue detection using high-resolution time frequency method and machine learning algorithms," *Computer Methods and Programs in Biomedicine*, vol. 154, pp. 45–56, 2018.
- [5] G. Venugopal, M. Navaneethakrishna, and S. Ramakrishnan, "Extraction and analysis of multiple time window features associated with muscle fatigue conditions using SEMG signals," *Expert Systems with Applications*, vol. 41, no. 6, pp. 2652–2659, 2014.
- [6] H. A. Yousif, A. Zakaria, N. A. Rahim, A. F. Salleh, M. Mahmood, K. A. Alfarhan, L. M. Kamarudin, S. M. Mamduh, A. M. Hasan, and M. K. Hussain, "Assessment of muscles fatigue based on surface EMG signals using machine learning and statistical approaches: A Review," *IOP Conference Series: Materials Science and Engineering*, vol. 705, no. 1, pp. 2-5, 2019.
- [7] E. F. Shair, S. A. Ahmad, M. H. Marhaban, S. B. Mohd Tamrin, and A. R. Abdullah, "EMG processing based

measures of fatigue assessment during manual lifting,”
BioMed Research International, vol. 2017, pp. 1–12, 2017.

[8] Ü. Kaljumäe, O. Hänninen, and O. Airaksinen, “Knee extensor fatigability and strength after bicycle ergometer training,” Archives of Physical Medicine and Rehabilitation, vol. 75, no. 5, pp. 564–567, 1994.

[9] V. Simic De Torres, “Instrumentation amplifiers for measurement of biosignals,” Treball Final de Grau, UPC, Escola d'Enginyeria de Barcelona Est, Departament d'Enginyeria de Sistemes, Automàtica i Informàtica Industrial, 2020.

[10] R. Merletti, “Electromyography: Physiology, engineering and non-invasive applications,” John Wiley & Sons, 2004, pp. 133–168.

[11] S. F. Ahmed, A. Ali, M. K. Joyo, M. Rehan, F. A. Siddiqui, J. A. Bhatti, A. Liaquat, and M. M. Dezfouli, “Mobility Assistance Robot for disabled persons using electromyography (emg) sensor,” 2018 IEEE International Conference on Innovative Research and Development (ICIRD) Bangkok, Thailand, 2018, pp. 1-5.

[12] K. Sharmila, T. V. Sarath, and K. I. Ramachandran, “EMG Controlled Low Cost Prosthetic Arm,” 2016 IEEE Distributed Computing, VLSI, Electrical Circuits and Robotics (DISCOVER), Mangalore, India, 2016, pp.169-172.

Resampling-free fast particle filtering with application to tracking rhythmic biomedical signals

Mohammed Ashik¹, Ramesh Patnaik¹

¹Dept. of Instrument Technology,

Andhra University

ashikmd909@gmail.com, ramesh_patnaik@yahoo.com

Praveen B. Choppala^{2,*}

²Dept. of E.C.E.

WISTM, Andhra University

praveen@wistm.edu.in

Abstract—The particle filter is known to be a powerful tool for the estimation of time varying latent states guided by nonlinear dynamics and sensor measurements. Particle filter's traditional resampling step is essential because it avoids degeneracy of particles by stochastically eliminating the small weight particles that do not contribute to the posterior probability density function and replacing them by copies of those having large weights. Nevertheless, resampling is computationally costly since it requires extensive and sequential communication among the particles. This work proposes a novel method of particle filtering that eliminates the need for resampling and prevents degeneracy by substituting low-weight particles with a simple cutoff decision strategy based on the cumulative sum of weights. The proposed scheme limits replacement over only a few important particles and hence substantially accelerates the filtering process. We show the merits of the proposed method via simulations using a nonlinear example and also apply the method for tracking harmonics of real biomedical signals.

Index Terms—Particle filtering, resampling, resampling-free particle filter, root mean square error, computational time

I. INTRODUCTION

Many scientific problems require sequential estimation (or tracking) of the latent state of a dynamic system using noisy sensor observations. The appropriate solution for problems as this is Bayesian state estimation [1]. The particle filter (PF) [2], [3] is known to be a powerful nonlinear nonGaussian Bayesian state estimator. The filter operates on the principle of approximating the posterior pdf by a set of weighted samples, which in this context are known as particles and the higher the number of particles the accurate the representation of the pdf. The PF comprises of two steps. The first, sequential importance sampling (SIS) specifies the process of propagating new particles at each time sample and updating their weights, and the second, the resampling, that replaces (or duplicates) the lower weight particles by those having larger weights so that the PF is a better representation of the posterior [4], [5]. The interest of this paper is in resampling.

The stochastic resamplers are the fundamental class of resamplers most conventionally used in PFs. One of the first resampling methods, the multinomial selection [2], operates by first evaluating the cumulative sum of the normalised particle weights and then finding a value of the sum greater than a random sample drawn from $\mathcal{U}(0, 1)$. Sampling from the full

interval (0, 1) outcomes in large Monte Carlo error variance. This problem was overcome in the stratified [6] and systematic [3] resampling that divide the interval (0, 1) into strata and draw one sample per particle from each stratum. The residual resampler [7] improved upon these by duplicating the particles stochastically using principle of proportional allocation. The stochastic resamplers are known to be theoretically accurate as they display the posterior distribution in an unbiased manner pdf. That being said, using resampling to improve upon previous random sampling involves sequential search and extensive communication overhead within all the particles. This leads to enormous computational complexity and hence impedes the use of more particles for accurate estimation and leads to difficulty in hardware realisation [8].

The Gaussian particle filter (GPF) [9], on the other hand, is a resampling-free PF approach that operates by assuming the densities are Gaussian. The method is fast but suffers from estimation bias. Among other classes of resamplers, the partially deterministic resamplers [10] use thresholding techniques to reliably reproduce large-weight particles and randomly sample the smaller-weight ones. For the treatment of the inefficient replication of large-weight particles, similar methods were presented by Fearnhead [11], Barembach et al [12] and Choppala et al [13]. A review of recent advances in resampling can be found in [8], [14].

This paper proposes a novel and simple PF approach that does not involve traditional resampling over all the particles but replacement across a few low weight particles. The key idea here is to determine the large weight particles using a cutoff of the cumulative sum of the weights and retaining them, and then replace the low weight particles using those having large weights. This replacement scheme avoids the need to resample all the particles and thus minimises the communication overhead to over only a minimum number of particles. The key benefit of this approach is the ability to achieve accelerated filtering when compared to the standard PF with stochastic resampling.

The rest of the paper is organised as follows. Section II describes the PF approach and the need for conventional resampling. Section III proposes the proposed resampling free PF. Sections IV and V present a evaluation study and application to tracking harmonics of rhythmic biomedical signals. We finally conclude in section VI.

* Corresponding author.

II. PARTICLE FILTERING

Using all available data, Bayesian state estimators attempt to iteratively estimate a posterior pdf for the hidden target state is given by $p(\mathbf{x}_t|\mathbf{y}_{1:t}), t = 1, \dots, T$ where the real-valued hidden target state is $\mathbf{x}_t \in \mathbb{R}^{d_x}$ at time instant $t \in \mathbb{N}$ and the sequence $\mathbf{x}_{1:t} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_t\}$ is a hidden Markov process with initial distribution $p(\mathbf{x}_0)$ and the Markov state transition pdf $\mathbf{x}_t|\mathbf{x}_{t-1} \sim p(\mathbf{x}_t|\mathbf{x}_{t-1}, \theta)$ with d_x denoting the state dimensionality, and the sensor observation $\mathbf{y}_t \in \mathbb{R}^{d_y}$ is conditionally independent of previous observations $\mathbf{y}_{1:t-1} = \{\mathbf{y}_1, \dots, \mathbf{y}_{t-1}\}$ given the state at time t and follows Density of observations, denoted by $p(\mathbf{y}_t|\mathbf{x}_t, \theta)$. Observation dimensionality is denoted by d_y . For simplicity, we will ignore from now on that the model also has a vector of known parameters denoted by the θ .

As part of the Bayesian recursion, the pdfs are represented by the PF as an approximation utilising a collection of weighted particles. Take the posterior pdf of the target state $\{\mathbf{x}_{t-1}^i, w_{t-1}^i\}_{i=1}^N$ to be a collection of N weighted particles $p(\mathbf{x}_{t-1}|\mathbf{y}_{1:t-1})$ at time $t-1$. In order to advance to time t , the PF creates a new ensemble of particles from the current ensemble, using a proposal distribution $q(\cdot)$ as its seed $\bar{\mathbf{x}}_t^i \sim q(\mathbf{x}_t|\mathbf{x}_{t-1}^i, \mathbf{y}_t)$ and then weighs them using the new observation according to

$$\bar{w}_t^i = w_{t-1}^i \frac{p(\mathbf{y}_t|\bar{\mathbf{x}}_t^i)p(\bar{\mathbf{x}}_t^i|\mathbf{x}_{t-1}^i)}{q(\bar{\mathbf{x}}_t^i|\mathbf{x}_{t-1}^i, \mathbf{y}_t)}, i = 1, \dots, N \quad (1)$$

and normalises the weights. But nevertheless, the disparity in weights grows after a few rounds, resulting in decadence. Decadence causes the PF to completely diverge/bias from the true target state. The solution to this is resampling. Conventional resampling methods that accomplish this obtains a new weighted set of N particles $\{\bar{\mathbf{x}}_t^i, \bar{w}_t^i\}_{i=1}^N \rightarrow \{\mathbf{x}_t^i, w_t^i\}_{i=1}^N$ as follows [4], [5]; for $i = 1, \dots, N$, we sample an index $j(i)$ distributed according to the probability $P(j(i) = m) = \bar{w}_t^m, m = 1, \dots, N$ and assign $\mathbf{x}_t^i = \bar{\mathbf{x}}_t^{j(i)}$ and set $w_t^i = 1/N$. The weights are then normalised and this weighted particle set is representative of the posterior at t as

$$p(\mathbf{x}_t|\mathbf{y}_{1:t}) \approx \sum_{i=1}^N \bar{w}_t^i \delta(\mathbf{x}_t - \bar{\mathbf{x}}_t^i) = \sum_{i=1}^N \frac{1}{N} \delta(\mathbf{x}_t - \mathbf{x}_t^i) \quad (2)$$

where $\delta(\cdot)$ is the Dirac-delta function.

Inferring that this resampling is consecutive and comprehensive across all particles is assumed. As a result, computing the PF is a significant expense. One approach to mitigate the computational complexity is to resample only when the estimated effective sample size (ESS), defined as,

$$\text{ESS} = 1 / \sum_{i=1}^N (\bar{w}_t^i)^2 \quad (3)$$

falls below a certain threshold. That being said, highly nonlinear models or highly manoeuvring targets and/or low noise scenarios cause high degeneracy and require frequent resampling. In the subsequent section, we propose a novel

resampling-free approach the PF will run much faster as a result.

III. PROPOSED PARTICLE FILTER

In this section, we propose the resampling-free PF approach. Let the sampled particles and the normalised weights $\{\bar{\mathbf{x}}_t^i, \bar{w}_t^i\}_{i=1}^N$ be obtained from (II) and (1) respectively. We then rearrange them in accordance to the ascending order of the weights as

$$\{\bar{\mathbf{x}}_t^i, \bar{w}_t^i\}_{i=1}^N : \bar{w}_t^1 \leq \bar{w}_t^2 \leq \dots \leq \bar{w}_t^N$$

The weights are then divided by the weight of the N th particle as

$$\mathbf{c} = \{c_t^i = \bar{w}_t^i / \bar{w}_t^N, i = 1, \dots, N\} \quad (4)$$

so that the maximum value c_t^N is limited to one. It can be understood that the function $c_t^{1:N}$ is a monotonically increasing function and mimics the cumulative sum. We then determine the first index \hat{i} at which the increasing scaled weights $\{c_t^1, c_t^2, \dots, c_t^{\hat{i}}, \dots, c_t^N\}$ is greater than a chosen threshold γ as

$$c_t^{\forall i \geq \hat{i}} \geq \gamma \quad (5)$$

We now determine the set of dominant set of indices D that will be retained as

$$D = \{\hat{i}, \hat{i} + 1, \dots, N\} \quad (6)$$

Figure 1 shows the increasingly scaled weights for weights that are moderately peaked (moderately informative) in blue to being very peaked (or highly informative) in red. It can

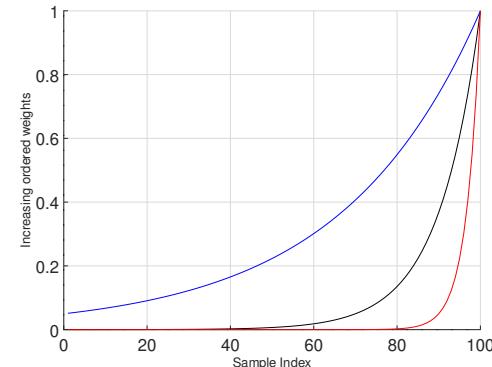


Fig. 1. The function \mathbf{c} for different cases from not-so-peaky to very peaky weights.

be observed that bigger value of γ and peaky weights cause very few very high weights to be gathered in the dominant set. For moderate weights and γ value, more particles would be gathered into the dominant set. We aim to retain the set of particles with indices contained in D and use them to replicate others, hence the value of γ that determines the set should be chosen carefully. In this paper, we set $\gamma = 0.8$.

The non-dominant set of $\hat{i}-1$ particles, i.e., those having a

weight less than γ are eliminated and a new set of samples $\mathbf{s}_t^{i=1, \dots, \hat{i}-1}$ are drawn instead as follows, sample a set of $\hat{i} - 1$ samples uniformly at random, with replacement, from the set of dominant particles $\tilde{\mathbf{x}}_t^{i \in D}$. The final set of particles will contain the new sampled particles $\{\mathbf{s}_t^i\}_{i=1}^{\hat{i}-1}$ appended with the particles corresponding to the dominant set $\{\mathbf{x}_t^i\}_{i=\hat{i}}^N$. It should, however, be noted that the new set contains replicates of particles corresponding to the dominant set, therefore for very peaky weights (very low noise condition), we may run into the problem of very few particles replicating the entire new particle set causing degeneracy. To avoid this, we perturb the new samples by a small regularising factor as

$$\mathbf{s}_t^i = \mathbf{s}_t^i + u, u \sim \mathcal{N}(0, \epsilon \lll), i = 1, \dots, \hat{i} - 1 \quad (7)$$

Then the final set of particles are arranged and their weights are then computed according to

$$\mathbf{x}_t^{1:N} = \{\mathbf{s}_t^{i=1, \dots, \hat{i}-1}, \mathbf{x}_t^{\hat{i}}, \mathbf{x}_t^{\hat{i}+1}, \dots, \mathbf{x}_t^N\} \quad (8)$$

$$w_t^i = p(\mathbf{y}_t | \mathbf{x}_t^i), i = 1, \dots, N \quad (9)$$

Figure 2 illustrates the proposed approach at a single time

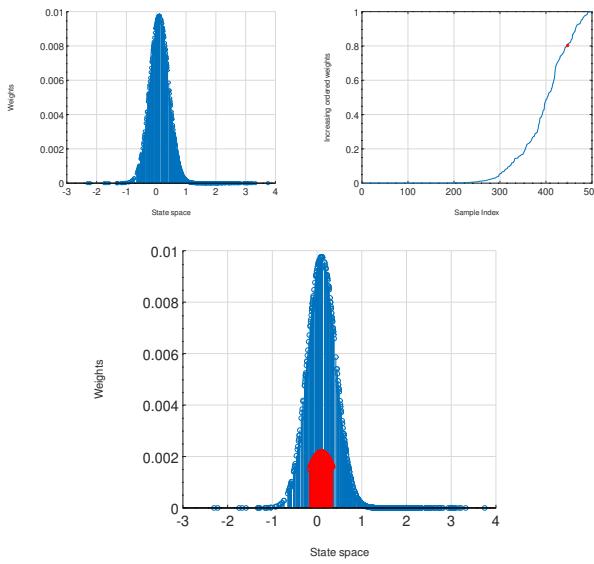


Fig. 2. An illustration of the proposed PF approach at one time instant. The left top panel shows the stem plot of the particles and weights obtained from (II) and (1). The right top panel corresponds to the increasing ordered function c from (4); $\hat{i} = 447$ is shown using a red dot. The bottom panel shows the original weighted samples in blue and the final resampled particles and weights obtained from (8) and (9) in red.

stamp. The original particles, with $N = 500$, are sampled according to $p(\mathbf{x}) = \mathcal{N}(1, 1)$ and weighed as $p(\mathbf{y} = 0.1 | \mathbf{x}) = \mathcal{N}(\mathbf{x}, 0.1)$. For $\gamma = 0.8$, we have $\hat{i} = 447$ (shown as a red dot on the monotonically increasing scaled weight function c shown in the right top panel), meaning that samples from index 447 to 500 are retained and 1 to 446 samples are drawn from the retained ones. The final resampled set shown in the bottom

panel and it can be observed that the low weight particles are effectively eliminated and new samples are drawn from regions of high importance (weight).

The key merits of the proposed method is that resampling of particles is now minimised to over only those having high weights while the conventional method resamples across all the particles. This proposition substantially accelerates the filtering process. The proposed method, however, does not contribute to the estimation accuracy anymore than does the standard particle filter. In the subsequent section, we evaluate the performance of the proposed method.

IV. EVALUATION

In this section, we present the evaluation of the proposed PF compared with the standard PF using the popularly used systematic resampling for the popularly used nonlinear growth model. Consider the nonlinear state space model

$$\mathbf{x}_t = \frac{\mathbf{x}_{t-1}}{2} + \frac{25\mathbf{x}_t}{1 + \mathbf{x}_t^2} + 8\cos(1.2t) + \mathbf{a}_t \quad (10)$$

$$\mathbf{y}_t = \tan^{-1}(\mathbf{x}_t) + \mathbf{e}_t \quad (11)$$

for $t = 1, \dots, T = 100$ where the process noise random variable is $\mathbf{a}_t \sim \mathcal{N}(0, \tau^2 = 5)$ and the observation noise random variable is $\mathbf{e}_t \sim \mathcal{N}(0, \sigma^2 = 0.1)$. We evaluate 4 cases: (a) proposed fast PF CASE 1 - with no additional resampling, (b) proposed fast PF CASE 2 - with resampling if $\text{ESS}/N \leq 0.3$, (c) standard PF CASE 1 - with systematic resampling at all time instants, and (d) standard PF CASE 2 - with systematic resampling only when $\text{ESS}/N \leq 0.3$. All the simulations are averaged over 200 Monte Carlo runs.

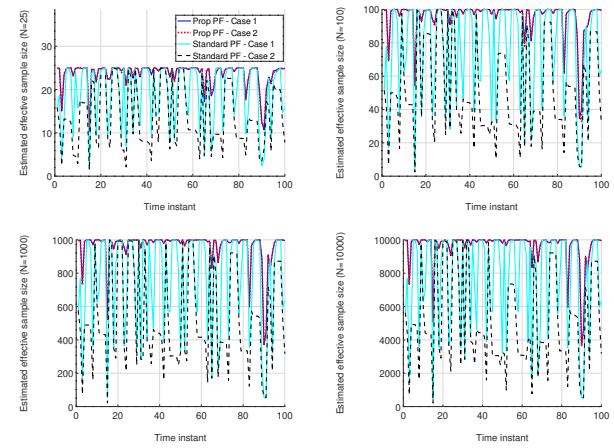


Fig. 3. The estimated ESS versus the time instant t for $N = 25, 100, 1000, 10000$.

Firstly, Figure 3 shows the degeneracy effect in the proposed filter for different values of N . The most important benefit of the suggested approach is that it avoids degeneracy by eliminating particles having low weights and sampling new ones instead using particles having large weights. Accordingly, The suggested method's ESS has been shown to stays

high throughout the simulation thus overcoming degeneracy whereas the standard PF with continuous resampling suffers. It can also be observed that the sample size N does not effect the impact of degeneracy on either filters. This is due of the highly nonlinear nature of the target dynamics that cause the filter to diverge easily. Figure 4 shows the CASE 2 of the proposed and the standard PF. It is clear that the proposed PF's ESS remains over the resampling threshold since it samples from strategically significant places hence does not require additional resampling, whereas the standard PF requires resampling at least 20% of the time.

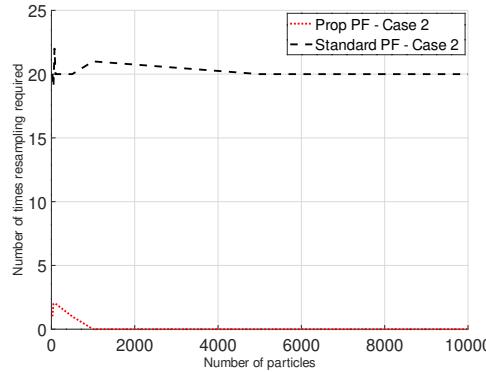


Fig. 4. The estimated ESS versus the time instant t for $N = 25, 100, 1000, 10000$.

Now that we have established that the proposed filter does not suffer from degeneracy, we focus on the computational time and tracking accuracy. In the Figure 5 the left panel shows the RMSE versus the computational time (in seconds) T_c for the values of $N \in (100, 10000)$. The initial value at $N = 100$ is depicted using a solid dot. It can be understood that the standard PF shows higher tracking accuracy but requires proportionately higher computational time than the proposed filter. The gain in terms of the computational time in the proposed filter is by virtue of minimising communication overhead within particles by resampling across only a few particles. An accurate measure of the tradeoff between computational time and tracking accuracy is the $T \times \text{RMSE}$ defined as

$$T \times \text{RMSE} = T_c \times \text{RMSE} \quad (12)$$

A low value of $T \times \text{RMSE}$ is desired as both the computational time and the tracking error is desired to be small. The right panel of Figure 5 plots the number of particles against the $T \times \text{RMSE}$. It can be observed that when the computational time and the tracking accuracy are taken cojointly, the proposed filter exhibits superior performance which increases with the sample size N . For $N = 100, 500, 1000, 5000, 10000$, the standard PF CASE 1 (resampling all the time) exhibits 3.3806, 10.3886, 15.0777, 22.8699 and 24.3066 times higher $T \times \text{RMSE}$ than the proposed PF CASE 1 respectively. Also the standard PF CASE 2 (conditionally resampling) exhibits 0.88150, 2.28424, 3.35400, 4.78495 and 5.06672 times more $T \times \text{RMSE}$ than the proposed PF respectively. This establishes

that the proposed method exhibits superior performance over the standard PF in terms of the combined measure of accuracy and computational time.

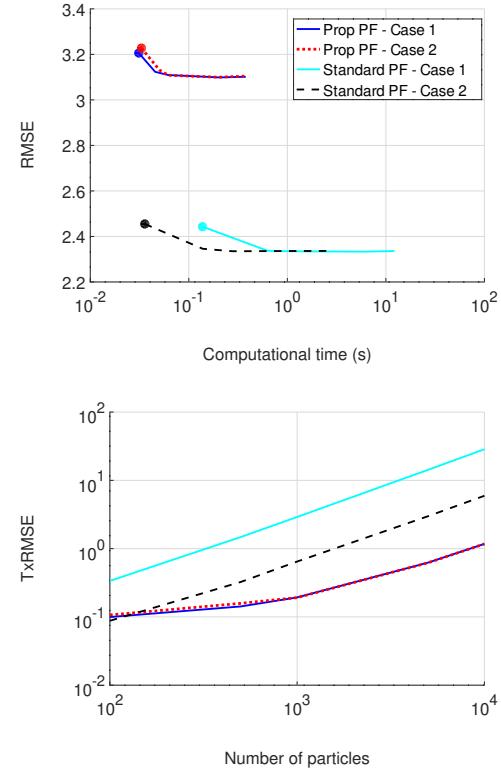


Fig. 5. The top panel shows the RMSE versus computational time (in seconds). The bottom panel shows the $T \times \text{RMSE}$ versus the number of particles.

Finally, we show how the proposed filter scales with noise. Table I shows the RMSE, the computational time and the $T \times \text{RMSE}$ for a low noise ($\sigma^2 = 0.1$) and a high noise ($\sigma^2 = 2$) with $N = 1000$. During low noise, the weights are very peaky (informative). This peaky nature leads to particles near the tails to have small weights. Several particles having small weights lead to degeneracy, and we need to resample often to avoid it. Resampling often in the standard PF leads to high computation complexity, whereas the proposed filter does not suffer from this problem as it always ensures that the sampled set of particles are drawn from regions of importance/information. It can be observed from the table that standard PF exhibits very high $T \times \text{RMSE}$ even when resampled fully or intermittently. During high noise, the weights are moderately peaked and this does not require frequent resampling. Hence it can be observed from the table that the % increase in $T \times \text{RMSE}$ of the standard PF over the proposed is not as much as the case is for low noise.

V. APPLICATION TO BIOMEDICAL SIGNAL TRACKING

We finally employ a challenging problem of tracking multiple harmonics in periodic rhythmical signals [15], [16]. These

Measure	σ^2	Stan PF CASE 1	Stan PF CASE 2	Prop PF
RMSE	0.1	2.3346	2.3351	3.1092
RMSE	2.0	4.5982	4.5949	5.8693
T_c	0.1	1.2455	0.1848	0.0620
T_c	2.0	1.2807	0.1848	0.0659
$T \times RMSE$	0.1	2.9080	0.6468	0.1928
$T \times RMSE$	2.0	0.2661	0.0795	0.0967
$\uparrow T \times RMSE$	0.1	1407.7	175.1	-
$\uparrow T \times RMSE$	2.0	235.40	-17.82	-

TABLE I

THE TABLE SHOWS THE RMSE, THE COMPUTATIONAL TIME (IN SECONDS) T_c , THE $T \times RMSE$ AND THE % INCREASE IN $T \times RMSE$ FOR THE STANDARD PF CASE 1 (RESAMPLING ALL THE TIME), STANDARD PF CASE 2 (CONDITIONAL RESAMPLING) AND THE PROPOSED FAST PF (WITH NO RESAMPLING). THE LAST TWO ROWS SHOW THE % INCREASE IN THE $T \times RMSE$ OF THE STANDARD FILTERS OVER THE PROPOSED FILTER.

signals are a model of the biomedical signals including the electrocardiogram (ECG) signals. The measurement model for this example is

$$y(t) = \sum_{k=1}^{N_h} \left(a_{k,t} \cos(k\theta_t) + b_{k,t} \sin(k\theta_t) \right) + \bar{y}_t + e_t \quad (13)$$

where N_h is the known number of harmonics, θ_t is the instantaneous phase of the fundamental frequency, \bar{y}_t is the signal mean (assumed known), $a_{k,t}$ and $b_{k,t}$ are the sinusoidal coefficients and $e_t \sim \mathcal{N}(0, \sigma^2)$ is the additive white measurement noise variable. The state variables that transit over time are defined as

$$\theta_t = \theta_{t-1} + 2\pi t_s f_t \quad (14)$$

$$\bar{f}_t = g[\bar{f}_{t-1} + u_{f,t}] \quad (15)$$

$$f_t = \bar{f}_{t-1} + \alpha(f_{t-1} - \bar{f}_{t-1}) + u_{f,t} \quad (16)$$

$$a_{k,t} \sim \mathcal{N}(a_{k,t-1}, 10^{-2}), k = 1, \dots, N_h \quad (17)$$

$$b_{k,t} \sim \mathcal{N}(b_{k,t-1}, 10^{-2}), k = 1, \dots, N_h \quad (18)$$

for $t = 1, \dots, T$ where f_t is the fundamental frequency, \bar{f}_t is the mean fundamental frequency, $t_s = 1/f_s$ is the sampling interval, $\alpha = 0.99$ is the auto-regressive coefficient assumed to be known and the Markov state transition noise variables are $u_{f,t} \sim \mathcal{N}(0, 10^{-4})$ and $u_{\bar{f},t} \sim \mathcal{N}(0, 10^{-6})$. The mean frequency is assumed to follow a nonlinear reflecting function

$$g[f] = \begin{cases} f_{\max} - (f - f_{\max}), & \text{if } f_{\max} \leq f, \\ f, & \text{if } f_{\min} \leq f \leq f_{\max}, \\ f_{\min} + (f_{\min} - f), & \text{if } f \leq f_{\min} \end{cases} \quad (19)$$

The state space is now $2N_h + 3$ dimensional and the state vector is $\mathbf{x}_t = (\theta_t, f_t, \bar{f}_t, a_{1,t}, a_{2,t}, \dots, a_{N_h,t}, a_{1,t}, a_{2,t}, \dots, a_{N_h,t})^\top$. The reader is referred to [15] for more detailed description of the model.

The real data is obtained from [17] which were originally taken from the PhysioNet service (<http://www.physionet.org>) from the MIT-BIH Arrhythmia database. Three ECG signals corresponding to the normal sinus rhythm class are taken. Each ECG signal is sampled at $f_s = 360$ Hertz and we use a 3.3 second signal containing $T = 1200$ samples. We set $N_h = 5$, $f_s = 360$ Hz, $f_{\max} = 1$, $f_{\min} = 5$ Hertz, $\bar{y}_{\forall t} = 0$, $\sigma^2 = E(y_{1:T})/10$. The initial state values are set to $\theta_{t=0} = 0$ and

$f_{t=0} = 2$, $\bar{f}_{t=0} = 1.5$ Hertz. Since the original model (ground truth) is unavailable for real data, we show the normalised mean square error (NMSE) measure defined as

$$NMSE = \frac{\sum_{t=1}^T (y_t - \hat{y}_t)^2}{(y_t - \bar{y}_t)^2} \quad (20)$$

where $\hat{\cdot}$ is the estimated value. The NMSE hence captures the similarity between the received noisy measurements $y_{1:T}$ and the estimate of the measurement $\hat{y}_{1:T}$ obtained from the estimated target state $\hat{\mathbf{x}}_{1:T}$, and a value of less than one indicates good harmonic tracking [15]. We show the estimated measurements in Figure 6 and the NMSE values, averaged over 50 Monte iterations are shown Table II.

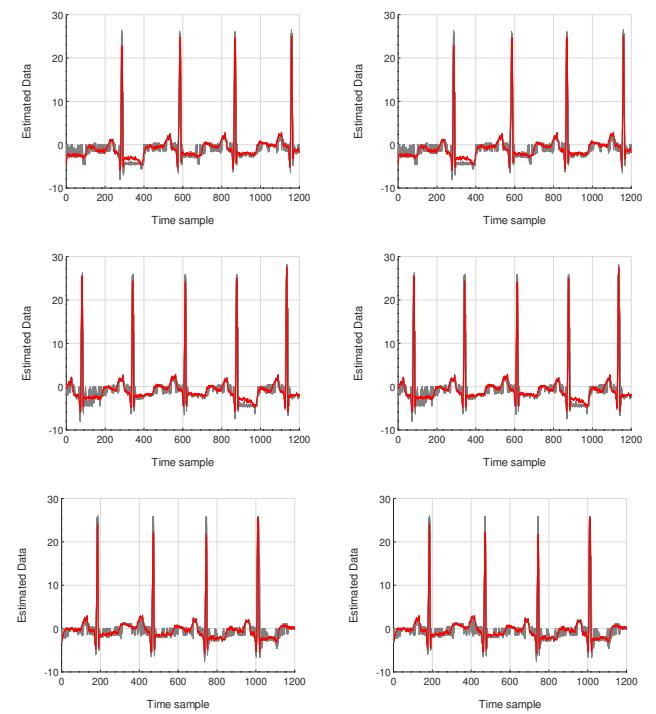


Fig. 6. The left panels correspond to the proposed PF harmonic tracker for the 3 ECG signals and the right panels to the standard PF. The measured signal is shown in red and the 50 Monte Carlo simulations of the estimated signal are shown in gray.

It can be observed that the proposed method shows an improvement of 83.87%, 82.43% and 82.43% (% reduction in

		Data1	Data2	Data3
$E(NMSE)$	Proposed	0.10976	0.13175	0.12995
$E(NMSE)$	PF	0.62687	0.74569	0.69508
$Var(NMSE)$	Proposed	0.000036660	0.0011593	0.00034820
$Var(NMSE)$	PF	0.026454962	0.0041143	0.05708340
T (s)	Proposed	12.891	15.286	32.131
T (s)	PF	32.317	44.880	49.303

TABLE II

THE MEAN AND VARIANCE OF THE NMSE VALUES OVER 50 INDEPENDENT SIMULATIONS FOR THE REAL ECG DATASETS.

NMSE) over the standard PF in terms of NMSE. It can also be observed that the proposed method exhibits an improvement of 60.37%, 65.89% 34.88% (% reduction in computational time) in computational time (taken in seconds). This evaluation study has shown that the proposed resampling-free fast PF approach substantially accelerates the filtering operation as it involves replacing only a few large weight particles. Also the proposed method adds in terms of tracking accuracy by leveraging the proposed replacement strategy on the high weight particles. The MATLAB code developed for this model is available at bit.ly/3IPLOoH. The study of the proposal for high noise conditions will be explored in the future.

VI. CONCLUSION

To get over degeneracy and focus the particles on the most important areas, the resampling step in the particle filter is essential. This aids in accurate exploration of the state space by the particles. However the conventional resampling methods involve a computationally expensive sequential operation over all the particles. This paper proposed a resampling free PF approach to overcome the computational overhead suffered by the traditional PF. The key idea here is to form a monotonically increasing function of the weights (scaled to a maximum value of one) and then determine a set of important indices with large weights using cutoff thresholding and replicate the others using the particles corresponding to the important indices. The key benefit of this proposition, as evidenced in the evaluation results is that we avoid degeneracy and filter is substantially accelerated. The proposed method has been successfully applied to track harmonic frequencies in real biomedical signals. In the future, we aim to adapt the value of the threshold in order to capture the tails of the posterior density accurately.

REFERENCES

- [1] Bar-Shalom, Yaakov, X. Rong Li, and Thiagalingam Kirubarajan, "Estimation with applications to tracking and navigation: theory algorithms and software," John Wiley & Sons, 2004.
- [2] N. Gordon, David J. Salmond, and Adrian FM Smith, "Novel approach to nonlinear/non-Gaussian Bayesian state estimation," In IEE proceedings F (radar and signal processing), vol. 140, no. 2, pp. 107–113. 1993.
- [3] M.S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking," IEEE Trans. Signal Proc., vol 50, no. 2, pp. 174–188, 2002.
- [4] R. Douc, and O Cappe, "Comparison of resampling schemes for particle filtering," In Proc. IEEE Symp. on Image and Signal Processing and Analysis, pp. 64–69, 2005.
- [5] J. D. Hol, Thomas B. Schon, and F. Gustafsson, "On resampling algorithms for particle filters," Proc. 2006 IEEE Workshop on Nonlinear Statistical Signal Proc., pp. 79–82. 2006.
- [6] G. Kitagawa, "Monte Carlo filter and smoother for non-Gaussian nonlinear state space models," J. of Computational and Graphical Statistics, pp. 1–25, 1996.
- [7] J. Liu, and R. Chen, "Sequential Monte Carlo methods for dynamic systems," J. American Statistical Association, Vol. 93, No. 443, pp. 1032–1044, 1998.
- [8] P. B. Choppala, P. D. Teal, and M. R. Frean, "Resampling and Network Theory," IEEE Trans. Signal and Information Processing over Networks, vol 08, pp. 106–119, 2022.
- [9] J.H. Kotecha, and P. M. Djuric, "Gaussian sum particle filtering," IEEE Trans. Signal Processing, vol 51, No. 10, pp. 2602–2612, 2003.
- [10] M. Bolic, P.M. Djuric, and S. Hong, "Resampling algorithms and architectures for distributed particle filters," IEEE Transactions on Signal Processing, Vol. 53, No. 7, pp. 2442–2450, 2005.
- [11] P. Fearnhead, "Particle filters for mixture models with an unknown number of components," J. Statistics and Computing, vol. 14, no. 1, pp. 11–21, 2004.
- [12] S. Baremburh, A. Garivier, and Eric Moulines, "On approximate maximum-likelihood methods for blind identification: How to cope with the curse of dimensionality," IEEE Trans. Signal Processing, vol 57, no. 11 pp. 4247–4259, 2009.
- [13] P. B. Choppala, P. D. Teal, and M. R. Frean, "Soft systematic resampling for accurate posterior approximation and increased information retention in particle filtering," In Proc. IEEE Workshop on Statistical Signal Processing, pp. 260–263, 2014.
- [14] T. Li, M. Bolic, and Petar M. Djuric, "Resampling methods for particle filtering: classification, implementation and strategies" in IEEE Signal processing magazine, Vol. 32, No. 3, pp. 70–86, 2015.
- [15] Kim Sunghan, Lars Andreas Holmstrom, and James McNames. "Tracking of rhythmical biomedical signals using the maximum a posteriori adaptive marginalized particle filter," J. British Journal of Health Informatics and Monitoring, Vol. 2, No. 1, pp. 1–23, 2015.
- [16] Kim Sunghan, Fouzia Noor, Mateo Aboy, and James McNames, "A novel particle filtering method for estimation of pulse pressure variation during spontaneous breathing," J. Biomedical engineering Online, Vol. 15, No. 1, pp. 1–18, 2016.
- [17] Paweł Plawiak, "ECG signals (1000 fragments)," Mendeley Data, V3, 2017. DOI: 10.17632/7dybx7wyfn.3

Design of an IP core for motion blur detection in fundus images using an FPGA-based accelerator

Rohit Jacob George

Department of Electronics Engineering
Madras Institute of Technology
Chennai, India
rohit20011002@gmail.com

Charaan S

Department of Electronics Engineering
Madras Institute of Technology
Chennai, India
kumarcharaan27@gmail.com

Swathi R

Department of Electronics Engineering
Madras Institute of Technology
Chennai, India
swathirk4053@gmail.com

S. P. Joy Vasantha Rani

Department of Electronics Engineering
Madras Institute of Technology
Chennai, India
joy_mit@annauniv.edu

Abstract— This paper focuses on applying an algorithm for real-time blur detection in fundus images via hardware acceleration. Blur in fundus images is caused due to many factors, but most of the time, with a reasonable degree of accuracy, they could be classified as motion blur. A motion blur could be modelled as an image convolved with a blur transfer function. Blur metrics are identified via techniques such as Haar DWT as it gives reasonable accuracy for most types of linear blur. First, a hardware architecture using Verilog HDL is created that computes the edge maps of images. This architecture is based on a novel algorithm that encompasses a series of Haar DWT Units. The simplicity and flexibility in this proposed architecture allow any kind of software or hardware platform to integrate the proposed model with very little to no modification, onto them. Subsequently, the IP core for the proposed architecture is developed, which can be further extended into an SoC, which can then be programmed onto a suitable FPGA system, which could then be uploaded with images that get classified as blurred and clear images. The on-chip processing system of the FPGA-SoC reads the image data and sends it to the Blur Detector IP via the DMA IP in the SoC. The whole process uses a double-buffered design in order to reduce IP stall time and increase efficiency.

Keywords—Fundus Images, Blur Detection, Haar DWT, IP Core, Xilinx Vivado, FPGA

I. INTRODUCTION

In today's era, digital images are ubiquitous. They have found their usage in almost every single modern application. One such important usage of digital images is in the pathology assessment of fundus in several specimens. Fundus refers to the interior portion of human eye, and comprises of the retina. Fundus images are captured using various cameras at various locations. As a result, there are numerous images of varying quality, some of which have pathologies that are either artificially created or cannot be readily seen. An ophthalmologist should review these poor-quality images and obtain new ones if necessary [1]. In order to perform digital image processing techniques such as compression, enhancement, restoration, and so on successfully, it is required that the image is devoid of motion blur [2]. Motion blur is defined as the de-focusing of parts or the entirety of the image due to the relative motion between the image background and the image capturing device, typically a camera [3]. As a result of motion blur, loss of clarity arises which makes the processing of images to be difficult. For example, object detection in images requires the images to have great clarity. Thus, motion blur tends to pose a serious issue for image

processing techniques, and early detection of motion blur must be done in order to find out whether the image is ready to be pre-processed for several applications.

Image blur detection is one of the most fundamental applications in the field of image processing, and it is often performed at the start of any image restoration method. There have been numerous numbers of reference image blur detection approaches that are proposed. A wide variety of blur detection algorithms have been presented till date. The blur metrics for these techniques are derived from a variety of different methodologies , including the Haar DWT, Sobel edge detection, DCT, and cosine transformations [4]. The Haar discrete wavelet transform is the most basic of these techniques, and it may be implemented on hardware without requiring a large number of resources.

Hardware–Software Codesign techniques are geared toward the design of embedded system cores and System-on-Chip (SoC) cores, both of which entail the integration of memory and ASIC blocks, DSP and microcontrollers, and interface peripherals [5]. Traditionally, a system is divided between hardware and software components, which are designed independently of one another with the exception of some common standards that are required owing to compatibility difficulties. Because of the increasing size of systems and the growing significance of power consumption, a new trend emerged to take into account the entirety of the system design process. This trend attempted to reduce the amount of design challenges by splitting the different system jobs early in the system design process between the system's hardware and software. The goal of these techniques is to determine, for a particular application, the ideal method of work splitting and assignment between software that is running on hardware such as FPGA and ASIC cores. The system tasks are divided between software parts that run on the core of a general microprocessor and sections of hardware that run on ASIC cores in this technique. The goal of this strategy is to reduce the overall amount of power that is being wasted by the system. [6]

In order to detect the presence of linear blur in static images, a novel hardware architecture needs to be developed, that is capable of determining the intensity of blur in images. This architecture must contain a blur detector unit that follows a certain algorithm which gives out the intensity of blur by computing the edge-map of images. After developing the architecture, the IP for a fully optimized System-on-Chip must

be generated using Xilinx Vivado Design Suite which incorporates the designed core blur detection unit.

The work presented in this paper is described as follows: Section 2 presents the related work undergone and section 3 illustrates the algorithm used for blur detection in hardware. Appropriate mathematical expressions and tabulations are added in this section which provide more information about the Discrete Wavelet Transform Function that is used in the blur detection algorithm. Section 4 contains the proposed design methodology that was followed during the course of the project. It follows a top-down approach, where the overall block diagram for the entire project is given; followed by separate sub-sections explaining the hardware designs undergone for this project. Section 5 discusses the results obtained at various stages of the hardware accelerator. This includes the results obtained during the IP Core generation. It also presents the utilization summary, along with power and timing summary. Section 6 presents the project conclusion and further work that is intended to be done.

II. RELATED WORK

Significant research has been made in the niche area of FPGA based accelerators for image processing applications – particularly on image blur detection. Rezk A. A., et. al. [6], have proposed an FPGA based IP core, that is capable of performing blur detection, compression and encryption via parallel processing. Their work presents a novel architecture that can detect blur and at the same time perform image compression and chaotic compression on-the-fly. Askari Javaran, et. al. [7], have designed a blur metric that estimates linear motion blur parameters. Their work presents the most important parameters that need to be considered for estimating blur percentage in static images. Additionally, their paper also explains how to produce the Point Spread Function once we compute the blur parameters. Li Yang [8], has proposed an algorithm that detects motion blur from photographs and later performs image restoration. It introduces the fractional derivative approach to estimate motion blur parameters. Koik B. T., et. al [9] have developed a set of blur detection algorithms for image segmentation. They have used methods such as non-reference (NR) block, wavelet-based histograms and blind image de-convolution for developing the blur detection algorithm. Marais, et. al. [10] have developed a defocus blur metric that may be used to assess the quality of blind images. They present a spectral subtraction technique variant based on a power spectrum surface of revolution that outperforms already existing direct deconvolution methods for defocus detection. This strategy has been proven to be effective for discriminating between blurred and clear images.

III. THE BLUR DETECTION ALGORITHM

In the field of mathematics, a Haar wavelet is defined as a succession of rescaled functions that have a "square-shaped" shape and, when combined, constitute a wavelet family or basis [11]. Wavelet and Fourier analysis may both represent a target function across an interval as an orthonormal basis. Fourier analysis is a kind of wavelet analysis. It is now generally accepted that the Haar sequence was the first known wavelet basis, and it is widely used as an example in educational settings. In this section of work, we carry out the traditional approach to computing the Haar transform, which

involves the use of convolution. In addition, rather than being constructed as a series of combinational circuits, the entire architecture is designed in a pipelined fashion. This ensures that the architecture is unaffected by the dimensions of the image. The Haar-based blur detection technique is capable of providing a high level of accuracy when identifying photographs that have been blurred. Furthermore, this approach may be integrated with neural networks to conduct an overall analysis of the image's quality. The result of running this method was the identification of the feature that had the greatest potential for performance when assessing the image's quality. In addition, this algorithm can be used in a range of applications, such as the lifelogging wearable cameras, to filter out fuzzy photographs so that only clear ones are displayed [12].

A. 1-Level 2D DWT Unit

From the original image, every four adjacent pixels are changed into A, B, C and D, where the terms "average and vertical coefficients" pertain to A and B, whereas "horizontal and diagonal coefficients" refer to D and C respectively [13]. The average coefficients will be found in the LL (Low-Low) band as opposed to the detail coefficients in the other bands. By using the conventional convolutional method, the 1-level 2D DWT Unit is obtained. This technique uses down samplers after passing the coefficients via a High and Low Pass Filter [14].

B. 3-Level 2D DWT Unit

The 3-level 2d DWT is accomplished by cascading units that are only one level deep. Controlling the down-sampling of the three levels is accomplished with the help of the enable signals, which range from E1 to E6. In contrast to the more common method of scanning rows one at a time, the image needs to be scanned in a particular order in order to make the most of these cascaded units.

C. Image Blur Detection Algorithm

The Dirac, the A-step, the G-step, or the Roof edge types are used in the blur detection approach to categorize the image's edges. [10]. The fluctuation in the intensity of the pixel is used as the basis for the categorization. Images that have been blurred usually are devoid of A-step or Dirac edges, and on the contrary, a significant amount of Roof and G-Step edges are found. As a result, the algorithm is able to recognize blurred pictures by distinguishing between the various types of edges present in the image. The 3-level 2d Haar wavelet transform is the first thing that the algorithm does once it has been initialized, followed by constructing the edge map for every level.

The third step is to use partitioning windows to create sub-edge maps. The fourth step is to determine which partitioning window has the highest value. This value is assigned to the term ' $\text{Edge}_{b,a}$ ', where 'b' represents the partition level index and 'a' represents the level of decomposition [6]. For three levels – The partition with index b will have an edge point if any of the aforementioned values is larger than the preset threshold (Th). Utilizing the guidelines, one may determine the kind of edge that is present. The total number of edges is divided by the sum of the Dirac and A-step edges in

the last step. In the case that this ratio falls below a very low threshold, the image will be deemed blurry. By dividing the total number of unsharp G-step and Roof edges by the total number of G-step and Roof edges, the blur extent is calculated [6]. Fig. 1 summarizes the above steps in the form of a workflow diagram to calculate the presence of blur along with the blur extent.

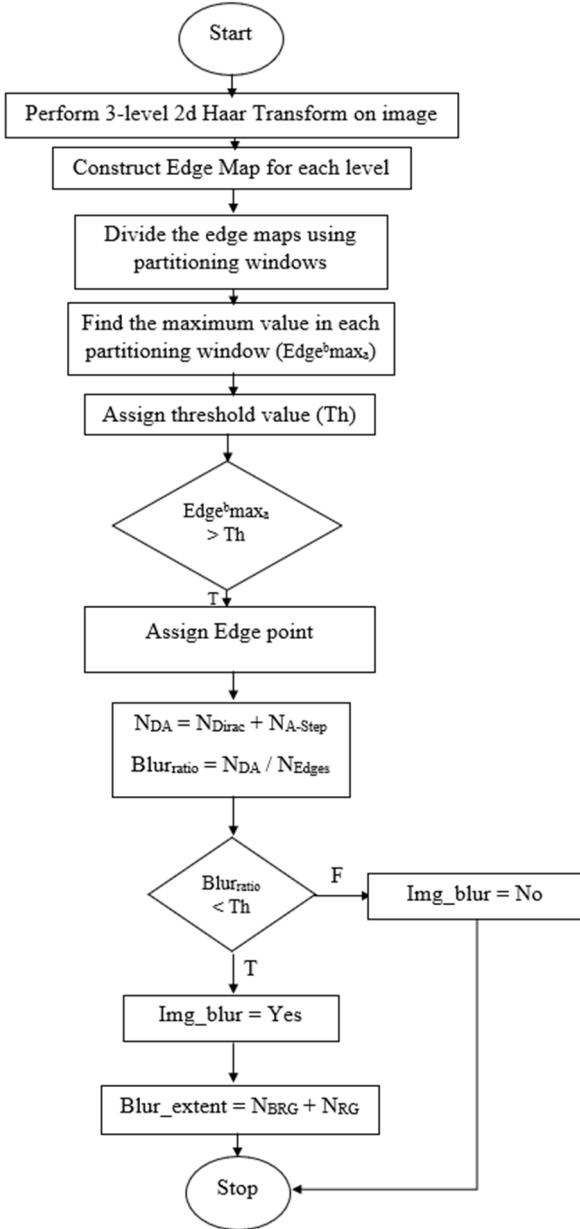


Fig. 1. Workflow Diagram of the Blur Detector Algorithm

IV. PROPOSED DESIGN METHODOLOGY

The overall workflow of the hardware-software co-designed System-on-Chip unit for motion blur detection is given in Fig. 2, in a top-down approach. On the hardware front, first, the core blur detection unit is designed using Verilog HDL. The detector encompasses a series of 3-D Haar DWT units along with High and Low Pass Filters, which execute the blur detection algorithm, by computing the edge maps of the image. After that, the IP for the Blur unit is generated from the Verilog code written. This IP can be interfaced with proprietary IPs of the Direct Memory Access (DMA), AXI, and AXILite present in the Vivado Design Suite.

Corresponding specifications are included and the SoC that houses the core blur detection unit can be designed.

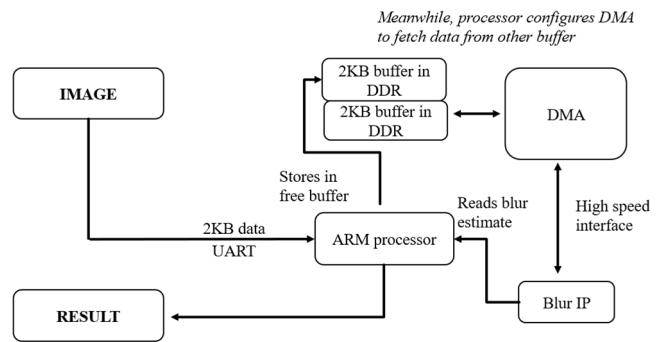


Fig. 2. Proposed overall workflow diagram of the Blur Detector SoC

A. Hardware Design

The hardware design front is initiated by the design of the core blur detection unit, using Verilog HDL. Once the design is realized, an IP is created from the Blur Detector Unit using the Vivado IP Manager tool. This IP carries the functionality of the Core Blur Detection Unit and is interfaced with a couple of other standalone IPs that are available in the Vivado IP manager library.

B. Core Blur Detection Unit

This unit is extensively designed using Verilog HDL and consists of an arrangement of comparators, registers, and an edge-type detector that produces the values of the four types of edges present in the image. The inputs given to the core blur detector unit, constitute the values obtained from the 3-D Haar DWT units which are classified into four types of bands: High-High (HH), High-Low (HL), Low-High (LH) and Low-Low (LL) [6]. The enable signals for the 3-level 2-D Haar Transform units are given by a sampler circuit. It finally produces the blur estimate of the image. In order for one batch (2KB) of data to get processed, the design takes 2048 clock cycles. By assuming a 100 MHz clock frequency, the processing takes ~20us to complete. The overall block diagram for the Blur Detector is given in Fig 3.

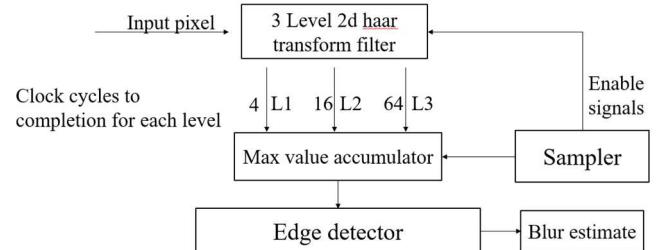


Fig. 3. Block Diagram of Blur Detection Unit

C. IP Core Generation

The modules for the core blur detection unit are written in Verilog and are as follows: Sampler, blur_detection_unit, edge_detector, 3_level_dwt. The code was then packaged into an IP core using the Vivado IP manager tool that comes under the Xilinx Vivado Design Suite. It reduces the amount of time needed for the design process by giving users access to highly characterized Intellectual Property (IP) core for Xilinx FPGAs. The Xilinx Vivado Design Suite gives access to an IP-centric design flow, which enables the inclusion of IP

modules derived from a variety of design sources into several designs. With the help of CORE Generator, a list of intellectual property that is particular to an architecture, a domain, or a market is made available. From routinely used functions like memory and FIFOs to system-level building blocks like filters and transformations, the complexity of these user-customizable IP functions spans the range [15]. All of these functions can be utilized in a variety of applications. These IP blocks might save anything from a few days to many months of design effort. The highly optimized IP allows FPGA designers to focus their efforts on swiftly creating designs, which accelerates the process of bringing products to market. Some intellectual property from the CORE Generator IP portfolio may be included into the user's chosen technique. Fig. 4 contains the generated IP core design for this blur detection algorithm.

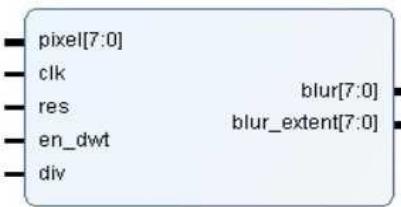


Fig. 4. IP generated for the designed Blur Detector using Vivado IP manager tool

D. Vivado Synthesis

The synthesis process in Xilinx Vivado is memory-usage and performance-optimized and at the same time, is timing-driven by virtue. The IDE features an implementation and synthesis environment that allows for a simple flow between synthesis and implementation processes. The utility automatically handles the run data, allowing for several run attempts with different Register Transfer Level (RTL) source versions. VHDL, Verilog, and System Verilog are all supported by Vivado. As a result, the Blur IP unit can now be configured with the AXI4 Full and AXI4 Stream Interface of FPGA Processing Systems [16]. The overall intended block diagram of the full IP Core of the blur detector design is given in Fig 5.

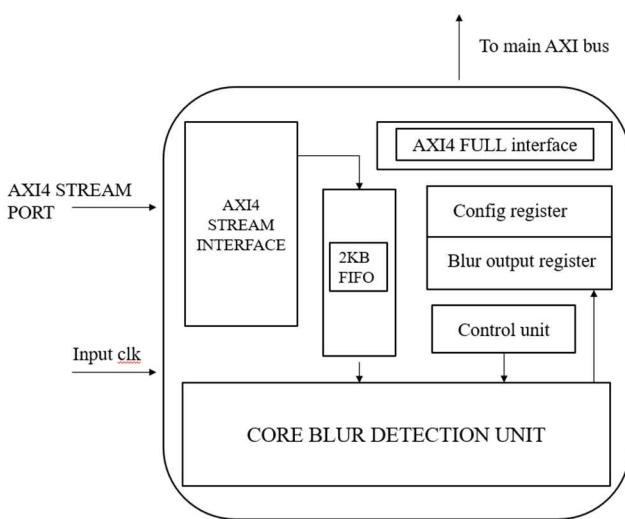


Fig. 5. Block Diagram of Full IP Core

V. RESULTS AND DISCUSSION

A. Simulation of Blur Detection Unit

Before moving further with the actual hardware implementation of the Haar DWT method, as is standard practice for any system, a simulation was run to evaluate its effectiveness first. In order to carry out this task, a standalone Vivado project known as testAlgo was developed. The only Verilog files that were used for this project were those for the core blur unit. In order to replicate the top-level file (core.blur.unit), a testbench was developed. This testbench's sole purpose was to read the contents of the image file from preprocessed memory files and then put those contents in a register. After that, the data included in the image was sent on to the implemented blur unit so that the process of the DMA engine moving the data could be replicated. Because of this, an extremely accurate and precisely regulated simulation of the blur unit that was designed was possible.

The simulation was run in Vivado by using the typical technique for running simulations in that program, and the results obtained, revealed that the IP core running the blur detection algorithm was efficient for the majority of the images that were used. It was discovered that the algorithm was able to identify a variety of blurs, including motion, out-of-focus, linear, and others like them.

The simulation results along with the RTL-level view of the blur detector unit are included in Fig. 6 and Fig. 7.

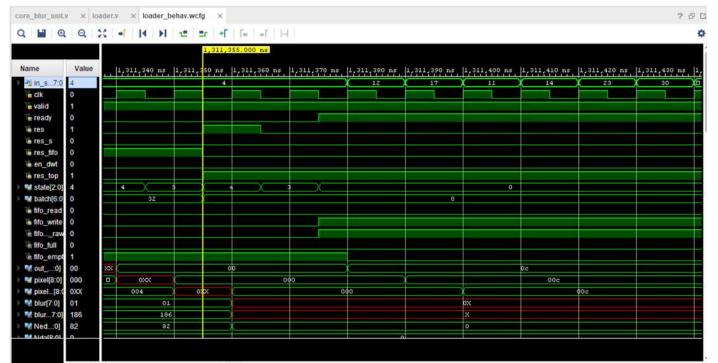


Fig. 6. Simulation results of Core Blur Detector Unit in Xilinx Vivado

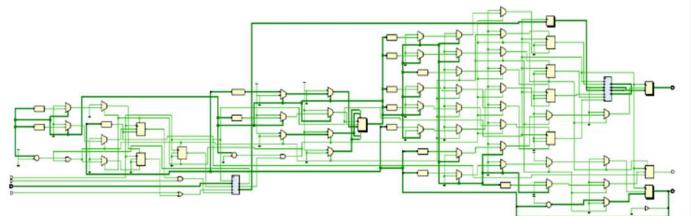


Fig. 7. RTL View of Blur Detector Unit with FIFO and Control Logic

B. IP Design

The final IP was then created by adding an AXI LITE interface and an AXI stream interface. These interfaces were created using the Vivado interface wizard and modified according to our needs. The input image data is sent through the AXI stream interface and the output data is read through the AXI LITE interface. The obtained IP Core is illustrated in Fig. 8.

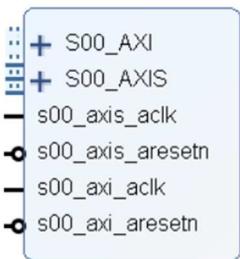


Fig. 8. IP Core with AXI and AXILite

C. Power Summary

The power analysis is derived directly from the netlist as a result of the implementation. The total power spent on the chip was assessed to be 1.22W; and the obtained power value was determined to be a small reasonable amount, which satisfied the essential requirements of the critical embedded system constraint. The targeted device had a static power consumption of 10% and a dynamic power consumption of 90%, as shown in Fig. 9.

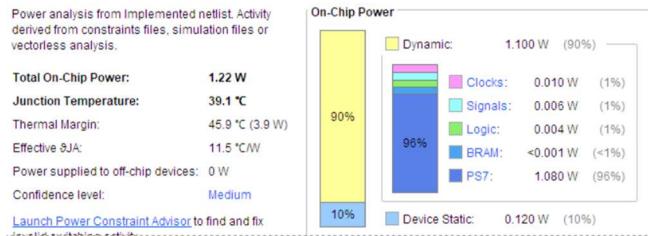


Fig. 9. Power Analysis Summary

D. Timing Summary

As a consequence of this, the timing summary of the design is also acquired from the netlist, and from this, it can be deduced that all of the user-specified time requirements have been satisfied. This demonstrates that there were no failures caused by significant critical path delays that was associated with critical time constraints. The timing summary obtained is illustrated in Fig. 10 below.

Design Timing Summary			
Setup	Hold	Pulse Width	
Worst Negative Slack (WNS):	0.159 ns	Worst Hold Slack (WHS):	0.059 ns
Total Negative Slack (TNS):	0.000 ns	Total Hold Slack (THS):	0.000 ns
Number of Falling Endpoints:	0	Number of Falling Endpoints:	0
Total Number of Endpoints:	56729	Total Number of Endpoints:	56729
All user specified timing constraints are met.			

Fig. 10. Timing Analysis Summary

E. Utilization Summary

The output of the design and implementation processes is the generation of the real resource occupation in the hardware system. It is possible to properly configure the architecture given that the quantity of resources actually required is lower than the number of resources occupied by the architecture as it is implemented. For this analysis, the Xilinx Zynq 7000 FPGA is considered whose processing system contains 17600 LUTs, 8806 of which can be used as memory or shift registers. This represents 50.034% of the total LUTs in the system, and hence can be effectively used for off-loading the designed IP Core. In a Zynq processing

system, there are 35200 slice registers available, of which 18927 of them have been used thus far (53.76%). In a similar fashion, the utilization of other resources is tabulated below in Table 1. Table 2 contains the performance report summary.

TABLE I. FPGA HARDWARE RESOURCE OCCUPATION SUMMARY

Resource	Utilized	Available	Utilization Percentage (%)
Slice LUTs	7066	17600	40.14
Slice Registers	16753	35200	47.5
LUT as Memory	201	6000	3.35
Block RAM Tile	1	60	1.67
Bonded IOPADs	130	130	100
BUFGCTRL	1	32	3.125
F7 Muxes	2184	8800	24.818
F8 Muxes	1088	4400	24.727

TABLE II. PERFORMANCE SUMMARY

Parameter	Value
Latency (Cycles)	65536
Latency (ns)	1.3×10^6

VI. CONCLUSION AND FURTHER WORK

In this paper, an algorithm that estimates the intensity of motion blur in fundus images was implemented. The algorithm encompasses a series of 3-D Haar Discrete Wavelet Transform units that compute the Edge Map of images and predict if the Roof and G-Step edge count value is greater than Dirac and A-Step edges. Later on, an IP of the Blur Detector Unit containing a Control unit and FIFO is created, and the same can be off-loaded onto an FPGA Board for accelerated computation in real-time. The AXI and AXILite communication protocols are being used for interfacing purposes, and the DMA IP is instantiated with a double buffer and the Core Blur Detector Unit. As a result, the presence of blur along with the blur intensity is returned for a series of test images. The obtained IP core is found to function as a hardware accelerator and speeds up the execution of the blur detection process manifold times, in contrast to the conventional way of detecting blur through software methodologies.

Furthermore, the detection of motion blur in real-time can be made possible by configuring the ARM Cortex Processors present in Xilinx FPGA systems using Xilinx Vivado SDK. Moreover, a python script can be written which handles the pre-processing of input images. The uploader program for pre-processing is written as a simple python script which is invoked on the PC after the FPGA is programmed. The program reads each image in the dataset and sends it to the device one by one. The result for each image can then be displayed as either blurred or not blurred on the console. Once all images are sent, a summary of the results is given which contains the total number of transmitted images to the FPGA, the number of correctly predicted images, and the total time taken. Additionally, the de-blurring of images in real-time can also be done using a suitable architecture and algorithm.

REFERENCES

- [1] A. Corrêa Silva, "Meta-Learning Applications in Digital Image Processing," *International Conference on Systems, Signals and Image Processing (IWSSIP)*, 2020, pp. 19-20.
- [2] A. A. Chernomorets and A. S. Krylov, "Blur detection in fundus images," *5th International Conference on BioMedical Engineering and Informatics*, 2012, pp. 243-246.
- [3] Y. Shen, J. Dang, T. Lei and W. Luo, "Motion blur parameters estimation based on frequency and spatial domain analysis," *2nd International Conference on Consumer Electronics, Communications and Networks (CECNet)*, 2012, pp. 387-390.
- [4] A. N. Almustafa, Y. Nugraha, A. Sulasakiin, I. D. Bhaswara and J. I. Kanggrawan, "Exploration of Image Blur Detection Methods on Globally Blur Images," *10th International Conference on Information and Communication Technology (ICoICT)*, 2022, pp. 275-280.
- [5] X. S. Hu, R. C. Murphy, S. Dosanjh, K. Olukotun and S. Poole, "Hardware/software co-design for high performance computing: Challenges and opportunities," *IEEE/ACM/IFIP International Conference on Hardware/Software Codesign and System Synthesis (CODES+ISSS)*, 2010, pp. 63-64.
- [6] A. A. Rezk, A. H. Madian, A. G. Radwan and A. M. Soliman, "On-the-Fly Parallel Processing IP-Core for Image Blur Detection, Compression, and Chaotic Encryption Based on FPGA," *IEEE Access*, vol. 9, 2021, pp. 82726-82746.
- [7] Taiebeh Askari Javaran, Hamid Hassanpour, "Using a Blur Metric to Estimate Linear Motion Blur Parameters", *Computational and Mathematical Methods in Medicine*, vol. 2021, 2021, pp. 1-8.
- [8] L. Yang, "Image Restoration from a Single Blurred Photograph," *3rd International Conference on Information Science and Control Engineering (ICISCE)*, 2016, pp. 405-409.
- [9] B. T. Koik and H. Ibrahim, "A Literature Survey on Blur Detection Algorithms for Digital Imaging," *1st International Conference on Artificial Intelligence, Modelling and Simulation*, 2013, pp. 272-277.
- [10] Marais I. V. Z. and Steyn W. H., "Robust defocus blur identification in the context of blind image quality assessment," *Signal Processing Image Communication*, 2007, vol. 22, no. 10, pp. 833-844.
- [11] Hanghang Tong, Mingjing Li, Hongjiang Zhang and Changshui Zhang, "Blur detection for digital images using wavelet transform," *IEEE International Conference on Multimedia and Expo (ICME*, 2004, pp. 17-20 Vol.1.
- [12] S. Chowdhury, M. S. Ferdous and J. M. Jose, "A user-study examining visualization of lifelogs," *14th International Workshop on Content-Based Multimedia Indexing (CBMI)*, 2016, pp. 1-6.
- [13] L.-T. Ko, J.-E. Chen, H.-C. Hsin, Y.-S. Shieh, and T.-Y. Sung, "Haar-Wavelet-Based Just Noticeable Distortion Model for Transparent Watermark," *Mathematical Problems in Engineering*, 2012, vol. 2012, pp. 1-14.
- [14] Angelopoulou, M.E., Masselos, K., Cheung, P.Y.K. et al., "Implementation and Comparison of the 5/3 Lifting 2D Discrete Wavelet Transform Computation Schedules on FPGAs," *Journal of Signal Processing Systems*, 2008, pp. 3-21.
- [15] V. M. Khvatov and D. A. Zheleznikov, "Development of an IP-cores Libraries as Part of the Design Flow of Integrated Circuits on FPGA," *IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (ElConRus)*, 2021, pp. 2686-2691.
- [16] R. Chandaluri and U. Nelakuditi, "Design and Implementation of AXI4-lite Interface in Zynq SoC," *IEEE Delhi Section Conference (DELCON)*, 2022, pp. 1-4.

Design of DCT Hardware Accelerator using FPGA for Medical Image Authentication

Swetha V

*Department of Electronics
Engineering
MIT Campus, Anna University
Chennai, India
swetha290502@gmail.com*

Evangeline Divya Sagayee G

*Department of Electronics
Engineering
MIT Campus, Anna University
Chennai, India
evangelinedivya26@gmail.com*

Brintha J S

*Department of Electronics
Engineering
MIT Campus, Anna University
Chennai, India
brinthajs123@gmail.com*

Karthika K

*Department of Electronics
Engineering
MIT Campus, Anna University
Chennai, India
karthikakaliapan@gmail.com*

Joy Vasantha Rani S P

*Department of Electronics
Engineering
MIT Campus, Anna University
Chennai, India
Joy_mit@annauniv.edu*

Abstract— The biomedical sector has undergone significant changes due to the continuous advancements in technology and computer networks. Multimedia data is now being utilized to transmit crucial biomedical information, including patient reports, diagnoses, and hospital information. To convert multimedia files to transform domain from spatial domain for encryption with a watermark, a Discrete Cosine Transform (DCT) is utilized. This involves incorporating the DCT into the process of transforming the file, allowing it to be modified and encrypted while retaining its original content. To achieve this a hardware accelerator that performs DCT technique is designed and implemented using Vivado HLS. The result shows that the designed architecture possesses high level of resilience, and the image quality is preserved after watermarking.

Keywords—Discrete Cosine Transform, Medical images, Field-programmable gate array, Telemedicine, Image watermarking, Hardware accelerator, Intellectual Property core, High Level Synthesis

I. INTRODUCTION

Digital image watermarking is a strategy that embeds hidden information, such as a logo, text, or an identifier, into digital images to protect the image's ownership and authenticity. In the biomedical industry, digital image watermarking is used in several ways. In Medical Image Authentication, biomedical images, ultrasound images, are critical in diagnosing and treating various diseases. Digital image watermarking can be used to authenticate these images, ensuring that they are original and have not been altered or manipulated. In Medical Image Ownership Protection, Digital image watermarking can be used to protect the copyrights to biomedical images. Radiologists and surgeons, often take images as part of their diagnosis and treatment. Watermarking these images ensures that the ownership is not lost or transferred without proper authorization.

Cryptography, steganography, and watermarking are the three basic methods of secure communication [1]. Digital watermarking involves the technique of inserting data into a digital multimedia content, which can be retrieved or detected at a later stage.

The purpose of this is to prevent copying and control the content. It is classified based on carrier and representation. When classified by the carrier, digital image watermarking can be image, video, audio or text watermarking [2].

Digital image watermarking may be divided into two types based on representation: visible and invisible image watermarking. Although digital watermarking can be used in any medium, it is most embedded in images to protect their copyright [3].

FPGA based digital image watermarking is discussed in [4, 5]. Digital image watermarking can be accomplished by performing DCT on the Discrete Wavelet Transform coefficients of the middle frequency LL sub bands from the original image one. But because of the disadvantages of DWT like, Shift sensitivity and poor directionality, DCT is preferred over it.

DCT is an abbreviation for Discrete Cosine Transform. DCT is a rapid computing method for Fourier transform that transforms real signals into frequency domain values. As many real-world signals do not have any complex components, DCT only operates on the real part of the complex signal. The discrete cosine transform is a method that represents an image as a group of sinusoids with different frequencies and magnitudes. The DCT has a unique property where, for a typical image, only a few DCT coefficients contain a significant amount of visually important information [6]. Optimization of the algorithm and data precision to fulfill the future requirements of watermarking batch processing is discussed in [7] which represents the designed hardware accelerator unit on the Field programmable logic array to accelerate the process. If this process is done by a CPU, the image watermarking technique will be consuming a lot of time. The issue of low rates is highly evident while performing these operations on many numbers of images. The CPU-only implementation of the digital watermarking algorithm may not fulfill the real time processing requirements. The integration of DWT and SVD techniques is used for ensuring the robustness of image watermarking [8]. The paper [9] implements three types of integer DCTs (DCT-II, DCT-V and DCT-VIII), and two types of and two types of integer DSTs (DST-I and

DST-VII) to reduce the numbers of addition and multiplication operations. To prevent separation between the data and pictures, we apply a hidden watermarking technique in the frequency domain of the image. To verify the origin of the medical image, we employ a watermarking technique that is visible but hard to detect in the spatial domain of the image [10]. Results show that the SVD and scrambling can extract the watermark information completely, and it has better anti-attack performance [11]. In [12], a method of authentication and self-recovery of tampered information in digital images is proposed.

This article is organized as follows: Section II illustrates the various forms of image watermarking techniques available and DCT compression algorithm. Section III describes the characteristics of FPGA and details about HLS. Section IV describes the architecture of the hardware accelerator designed for DCT compression. Section V gives the modelling, synthesis, and implementation results of the architecture involved. Section VI describes the conclusion and the applications that can be incorporated as a result of this work.

II. IMAGE WATERMARKING AND DCT

A. Image watermarking

Image watermarking involves the technique of inserting data into the multimedia product, which can then be retrieved or detected in the watermarked. The primary objective of image watermarking is to make the data imperceptible to the human eye, while making it difficult for unauthorized individuals/candidates to remove the product.

Various algorithms have been employed to implement the technique in the frequency and spatial domains, each with its own set of benefits and drawbacks. Visible watermarks are secondary images that are translucent and overlaid on of the primary image. An imperceptible watermark is embedded in a way it makes changes to pixel values imperceptible to the human eye. Retrieval of the watermark is only possible with a suitable decryption mechanism is embedded in such a way that changes to pixel values are perceptually undetectable, and the watermark can only be retrieved using an appropriate decryption mechanism.

Digital image watermarking is a technique that is used for concealing a hidden information that serves to secure a multimedia content's copyright or to provide evidence that the data is accurate, complete, and has not been tampered with. It provides protective digital information encoding and decoding, and authentication of received data by the recipient, and offers robust and reliable marks indicating copyright and legal ownership. However, if the private key is lost, pirates can remove watermarks from all images owned by a particular owner, leading to system instability. Pirates may employ non-traditional image-processing techniques to deceive monitoring software or desynchronize the detector.

B. Discrete Cosine Transform

With the advancement of digital techniques and the internet, digital watermarking techniques have become increasingly important in protecting digital product intellectual property rights. The paper proposes a digital watermarking algorithm based on the DCT technique. When a digital watermark is

embedded after DCT, it has an elevated level of robustness. Experiment results show that this algorithm is resistant to typical image analysis and malicious attacks. The DCT is a numerical approach for converting a spatial (or time) domain signal or information into orthogonal domain elementary frequency components (or frequency domain). The discrete cosine transform can be described as a linear, invertible function $F: RN \rightarrow RN$ (where R represents the set of real numbers) or as an $N \times N$ square matrix.

In the case of a two-dimensional DCT, it is applied to an M-by-N matrix and is explained as follows.

$$= \frac{2}{\sqrt{MN}} C(m)C(n) \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x,y) \cos \frac{(2x+1)mx}{2M} \cdot \cos \frac{(2y+1)ny}{2N}$$

Where $C(m) = C(n) = \frac{1}{\sqrt{2}}$ for $m, n = 0$ and $C(m), C(n) = 1$ otherwise.

The inverse of the two dimensional DCT is given as

$$f(x,y) = \frac{2}{\sqrt{MN}} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} C(m)C(n)F(m,n) \cos \frac{(2x+1)mx}{2M} \cdot \cos \frac{(2y+1)ny}{2N}$$

By Transformation matrix approach:

The forward and inverse transform of DCT in matrix expression is defined below

$$DCT: F = C.f.C^T$$

$$IDCT: f = C^T.f.C$$

When all pixels are transformed, they have positioned from most significant to least significant. When the least significant pixels are quantized to 0, pixel loss occurs. The Discrete Cosine Transform (DCT) helps to break down an image into different spectral components with varying levels of significance, which can impact the overall visual quality of the image. In most images, the low-order DCT coefficients contain a significant portion of the signal energy after transformation. This makes it possible to quantize these coefficients with greater precision compared to the higher-order coefficients.

The one-dimensional DCT is well-suited for processing signals that are one-dimensional. However the analysis of two-dimensional signals needs a two dimensional version of the DCT. The 2D DCT for an $N \times N$ matrix is quantified in a concise manner. The 1D DCT is implemented to every row of the result, accompanied by each column.

$$y_u = \alpha(u) \sum_{j=0}^{N-1} x_j \cos \frac{(2j+1)\pi u}{2N}$$

$$\text{Where } \alpha(u) = \begin{cases} \sqrt{\frac{1}{N}} & \text{for } u = 0 \\ \sqrt{\frac{1}{N}} & \text{otherwise} \end{cases}$$

While the practical application of this formula would require $O(N^2)$ operations, it is possible to achieve the same outcome with an $O(N \log N)$ complexity by breaking down the computation in a similar fashion to the Fast Fourier Transform (FFT).

III. FIELD PROGRAMMABLE GATE ARRAY AND HIGH LEVEL SYNTHESIS

A. Field Programmable Gate Array

An FPGA (Field-Programmable Gate Array) is a type of logic chip that contains a two-dimensional grid of cells and programmable switches. In essence, they are integrated circuits (ICs) that consist of a grid of identical processing units and configurable connections between them. Application Specific Integrated Circuit (ASIC) have the advantage of being faster. However, FPGAs can operate similarly to ASICs, which is a benefit. In the incident of a glitch, FPGAs can also be reprogrammed. ASICs are incompetent of doing so.

The special kind of FPGA utilized in this work is Zynq edge processing system. Zynq is an APSoC, which means that in addition to integrating most, if not all, of a computer's components into a single chip, developers can also use the FPGA technology found within it. As a result, when referring to Zynq, the term "system" refers to the system of dual dedicated processors (Dual-core ARM Cortex-A9 Processors) and FPGA technology. Developers can take advantage of the best of both worlds by having access to both processor and FPGA functions.

Zynq design flow comprises the following stages. The initial stage of developing a system involves determining its specifications and requirements. During the system design phase, the different tasks or functions are allocated for implementation either in PL (Programmable Logic) or PS (Processing System), which is referred to as task partitioning. This step is crucial since the overall system performance depends on which tasks are assigned to be implemented using the most appropriate technology, whether hardware or software. The next steps involve developing and testing the hardware and software components. For the PL, the focus is on identifying the functional blocks necessary to achieve the design objectives and integrating them as IPs while ensuring proper connectivity. Similarly, the software development activity involves writing code to run on the PS. Finally, system integration and testing are conducted to complete the design process.

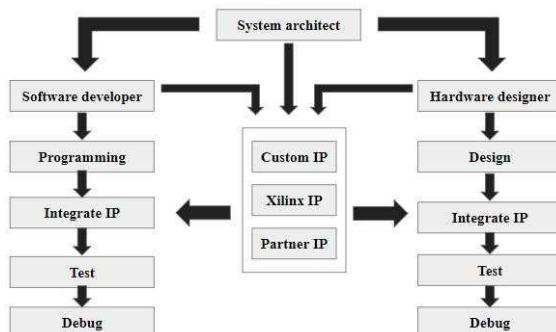


Fig. 1. Hardware Software co-design flow

B. High Level Synthesis

To simplify, synthesis involves the conversion of an abstract design into an accurately implemented chip using logic gates. The basic RTL design is transformed into a gate-level netlist that adheres to all the designer's specified constraints. On the other hand, HLS (High-Level Synthesis) is an automated design process that converts a design's behavioral or descriptive model into a digital hardware implementation.

The aim of HLS is to extract parallel processing from the input description and build a microarchitecture that is faster and less expensive than simply running the input description as a program on a microprocessor.

In this work, the control algorithms are programmed using the high-level language C. The automated tool utilized provides the Register Transfer Level (RTL) hardware description of the design. HLS (High-Level Synthesis) expedites the design implementation process by allowing C specifications to be targeted directly onto Xilinx devices, eliminating the need for manual RTL creation.

The goal of HLS is to extract parallelism from the input description and build a microarchitecture that is faster and less expensive than simply running the input description as a programme on a microprocessor.

IV. HARDWARE ACCELERATOR FOR DCT COMPRESSION ARCHITECTURE

A. Hardware acceleration

Hardware acceleration involves transferring specific computational tasks of an application to specialized hardware components within the system. This process results in greater efficiency compared to running software exclusively on a general-purpose CPU. Implementing applications higher up in the digital computing system hierarchy combines the versatility of CPUs, which are general-purpose processors, with the efficiency of specialized hardware like GPUs and ASICs, resulting in a significant increase in efficiency by several orders of magnitude.

B. DCT General Architecture

In order to embed watermark information, DCT is utilized to convert a particular image from frequency domain from frequency domain. To accomplish this, a hardware accelerator using the DCT technique is designed and implemented in Xilinx Vivado. To optimize space, we leverage the separability property of the 2D DCT formula. This means that the formula can be split into two 1D DCT operations - one along the row vector and the other along the column vector of the previous row vector's outcomes.

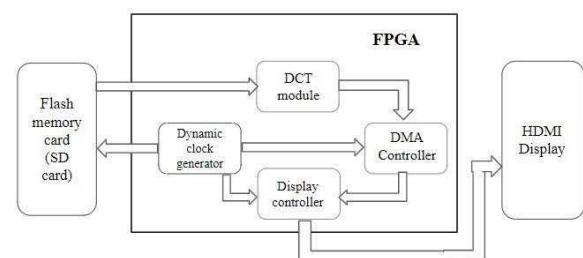


Fig. 2. Block diagram of DCT Hardware Accelerator

V. MODELLING SYNTHESIS AND IMPLEMENTATION RESULTS

A. IP core generation using vivado HLS

In this work, Vivado HLS is used to design a user-defined IP core for the DCT algorithm. The high level language (C language) used in IP core design is simulated. Hence the purpose of pre-synthesis validation is to ensure that the C program correctly implements the required functionality. The synthesis process then converts the source code into RTL implementation and generates a synthesis report. A post-synthesis verification process named C/RTL co-simulation is performed wherein verilog files are created as a result. Finally, the user-defined IP core is imported as a block into the Vivado IP catalog for hardware design. The IP core can be exported as a reusable block for other design applications that require DCT operation.

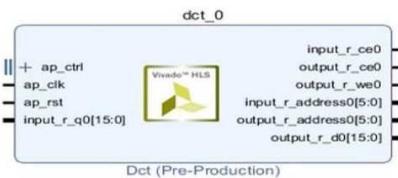


Fig. 3. User defined IP core for DCT by HLS

B. Overall Hardware Block Design

The overall functionality of the hardware accelerator is implemented by creating a block design with Vivado IP integrator. The Vivado user IP repository receives the user-defined IP core. The Zynq 7 processing system is included in the block design because it is compatible with the functionality of the hardware accelerator. The remaining functional cores required for the block design are derived from the default IP repository. The default block and connection automations are generated first, followed by manual interconnections generated in accordance with the architecture's functionality. The block design is then validated before synthesis and implementation. Finally, the generated bit stream is used to program the FPGA board to perform the desired functionality.

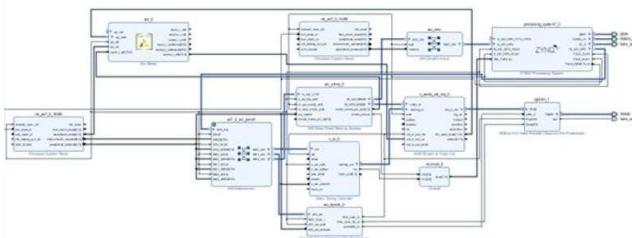


Fig. 4 Block Design

C. Utilization Summary

As a result of the design and implementation, the actual hardware system resource occupation is generated. The architecture can be successfully configured because the resources occupied by the implemented architecture are less than the actual quantity of resources. A Xilinx Zynq 7 processing system contains 17600 LUTs, 3937 of which can be used as memories or shift registers (22.37%). A Zynq processing system has 35200 flipflops available, of which

6938 (or 19.71%) have been used. The utilization of resources is also tabulated below.

TABLE I. FPGA CHIP HARDWARE RESOURCE OCCUPATION

Resource	Utilization	Availability	Utilization %
LUT	3937	17600	22.37
LUTRAM	243	6000	4.05
FF	6938	35200	19.71
BRAM	10.50	60	17.50
DSP	1	80	1.25
IO	11	100	11.00
BUFG	2	32	6.25
MMCM	1	2	50.00

D. Power summary

As a result of the implementation, the power analysis from the netlist is obtained. With a total on-chip power of 1.508 W, the implementation's power consumption is deemed to be reasonably low, satisfying the essential embedded system constraint. The targeted device has a static power consumption of 8% and a dynamic power consumption of 92%.

TABLE II. POWER ANALYSIS

Resource	Power consumption(W)	Percentage (%)
Clocks	0.033	2
Signals	0.013	1
Logic	0.008	1
BRAM	0.007	1
DSP	0.001	<1
MMCM	0.107	8
I/O	0.140	10
PS7	1.078	76

E. Experimental Setup

The Hardware design created by the Vivado IP integrator has been worked upon by the Software Development Kit (SDK), thus combining the hardware and software. The original image to be compressed is processed by the FPGA board. The board is programmed with the bit stream generated by the Vivado IP integrator. The image is DCT compressed and the quality of the original image is preserved.



Fig. 5. Interfacing with FPGA board

The experimental analysis of the hardware accelerator is performed for various images. For instance, the original image of size 5.9 MB is been compressed by the designed accelerator to an image of size 1.4 MB, thereby performing compression of ratio 4:1

TABLE III. RESULTS

Original image size(MB)	Compressed image size(MB)	Compression ratio
5.9	1.4	4:1
3.5	1.4	2.5:1

VI. CONCLUSION

To enhance computational efficiency, The DCT algorithm takes advantage of the frequency domain's energy compactness and matrix sparsity properties. The algorithm performs a two dimensional DCT decomposition to one set of one dimensional DCTs for a given spatial domain image block size, after which the calculation can be executed. The algorithm takes advantage of frequency coefficient redundancy to simplify the implementation of a 2D DCT. The design of the DCT algorithm is more effective when using a 2D DCT data matrix, as it contains minimal non-zero elements in the low-frequency range for a spatial domain data matrix.

In conclusion, DCT-based digital watermarking application has a confident impact on the achievement of the watermarking system. By employing a variety of robust features and appropriate embedding techniques, it is possible to increase the precision of watermarking methods. The use of image watermarking has become increasingly important in the medical field for safeguarding the authenticity and confidentiality of medical images. With the integration of a watermark into an image, medical experts can ensure the image's integrity by detecting any unauthorized modifications, thereby enhancing the accuracy of diagnoses and treatment plans and providing an added layer of confidence. By embedding a watermark into an image, medical professionals can ensure that the image has not been altered or tampered with, providing confidence in the accuracy of diagnoses and treatment plans. This technology also allows for secure sharing of medical images among healthcare providers, ensuring patient privacy and reducing the risk of data breaches. Additionally, image watermarking can help in identifying the ownership of the medical images, ensuring legal compliance, and protecting the intellectual property of medical institutions. In the medical industry, it is crucial to ensure the security and integrity of biomedical images, and image watermarking plays a significant role in achieving this objective. Therefore, image watermarking is considered a vital application in the medical industry.

REFERENCES

- [1] N. S. Kulkarni, I. Gupta, and S. N. Kulkarni, "A Robust Image Encryption Technique based on Random Vector," International conference on Emerging trends in engineering and technology (ICETET), pp. 15-19, 2018.
- [2] Lamri Laouamer and Omar Tayan, "A Semi-Blind Robust DCT Watermarking Approach for Sensitive Text Images," Arab J Sci Eng, Vol.40, pp.1097-1109, 2015.
- [3] Rahate Kunal B, A. S. Bhalchandra, S. S. Agrawal "VLSI Implementation of Digital Image Watermarking," International Journal of Engineering Research & Technology (IJERT), Vol. 2 Issue 6, 2016.
- [4] Majdi Farag Mohammed El Birek, M.F.L. Abdullah, Ali Abdrhman M. Ukasha and Ali A. Elrowayati (2016) "Digital Image Watermarking Based On Joint (DCT-DWT) and Arnold Transform" International Journal of Security and Its Applications, Vol.10, no. 5, pp.107-118
- [5] Majdi Farag Mohammed El Birek, M.F.L. Abdullah, Ali Abdrhman M. Ukasha and Ali A. Elrowayati (2016) "Digital Image Watermarking Based On Joint (DCT-DWT) and Arnold Transform" International Journal of Security and Its Applications, Vol.10, no. 5, pp.107-118
- [6] Raunak Phade, Rucha Mahajan, Sampurna Mankar, Pratik Patil (2013) "Digital Image Watermarking using DCT and ZIP Compression Technique" Global Journal of Computer Science and Technology Graphics & Vision, Vol. 13 2013, Issue 3 Version 1.0
- [7] Yanpeng Cao, Feng Yu, Yongming Tang (2020) "A Digital watermarking encryption technique based on FPGA Cloud accelerator" IEEE Access, Vol. 8
- [8] Narima Zermi, Amine Khaldi, Med Redouane Kafi, Fares Kahlessenane, Salah Eusch (2021) "A lossless DWT-SVD domain watermarking for medical information security" Multimedia Tools and Applications, Springer Nature 2021
- [9] Woonsung Park, Bumshik Lee, and Munchurl Kim (2019), "Fast Computation of Integer DCT-V, DCT-VIII and DST-VII for Video Coding" IEEE Transactions on Image Processing
- [10] Nunez-Ramirez, D, Cedillo-Hernandez, M., Nakano-Miyatake, M., & Perez-Meana, H (2020), "Efficient Management of Ultrasound Images using Digital Watermarking. IEEE Latin America Transactions"
- [11] Lei Pei and Naeem Jan (2022), "Research on Digital Image Watermarking Algorithm Based on Scrambling and Singular Value Decomposition", Hindawi , Journal of Mathematics, Volume 2022.
- [12] J. Molina, V. Ponomaryov, R. Reyes, S. Sadovnychiy, and C. Cruz (2020), "Watermarking Framework for Authentication and Self-recovery of Tampered Colour Images", IEEE Latin America Transactions, Vol. 18

**Ninth International Conference in Biosignals, Images and Instrumentation,
SSNCE, Chennai, March 16 – 17, 2023**

THANK YOU
- ICBSII 2023

PROCEEDINGS OF THE
NINTH INTERNATIONAL CONFERENCE ON
BIOSIGNALS, IMAGES, AND INSTRUMENTATION
March 16 & 17, 2023

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 Ninth International Conference on Biosignals, Images and Instrumentation (ICBSII-2023) is conceived with the thought of bringing together scientists, engineers and researchers from various domains all over the world. It is a platform where some of the greatest minds of the country and abroad could interact, exchange ideas and work together towards a common goal. Research papers were received from diverse areas such as Physiological Modelling, Medical Imaging, Medical Robotics, Biomechanics, Biomedical Instrumentation and Nanomaterials. After a rigorous review process by the expert review committee, papers that displayed quality in ideas and work were selected for the final presentation at the conference.

This conference is the fruit of a vision of the Management, faculty, and students of the Department of Biomedical Engineering, SSN College of Engineering in association with the Centre for Healthcare Technologies (CHT), a multi-disciplinary R&D centre, which works unanimously towards materializing the advancements in healthcare innovations. The Department of Biomedical Engineering, since its inception in 2005, has been a pioneer in the field of biomedical technology, instrumentation, and imaging. The department has excellent infrastructure, experienced faculty members and motivated students. Department also has foreign collaborations with Birmingham City University UK, Drexel University Philadelphia, University of Bologna, Italy, and several industries and hospitals such as Kauvery hospitals, Sri Ramachandra Medical College, Perfint Healthcare, NIEPMD etc to name a few.

To add feather to the crown, the department has conducted eight international conferences (ICBSII) in 2013, 2015, 2017, 2018, 2019, 2020, 2021, 2022 and two national conferences (NCABES) in 2014, 2016 so far.