

Sukanta Kumar Sabut
Arun Kumar Ray
Bibudhendu Pati
U Rajendra Acharya *Editors*

Proceedings of International Conference on Communication, Circuits, and Systems

IC3S 2020

Lecture Notes in Electrical Engineering

Volume 728

Series Editors

Leopoldo Angrisani, Department of Electrical and Information Technologies Engineering, University of Napoli Federico II, Naples, Italy

Marco Arteaga, Departament de Control y Robótica, Universidad Nacional Autónoma de México, Coyoacán, Mexico

Bijaya Ketan Panigrahi, Electrical Engineering, Indian Institute of Technology Delhi, New Delhi, Delhi, India
Samarjit Chakraborty, Fakultät für Elektrotechnik und Informationstechnik, TU München, Munich, Germany

Jiming Chen, Zhejiang University, Hangzhou, Zhejiang, China

Shanben Chen, Materials Science and Engineering, Shanghai Jiao Tong University, Shanghai, China

Tan Kay Chen, Department of Electrical and Computer Engineering, National University of Singapore, Singapore, Singapore

Rüdiger Dillmann, Humanoids and Intelligent Systems Laboratory, Karlsruhe Institute for Technology, Karlsruhe, Germany

Haibin Duan, Beijing University of Aeronautics and Astronautics, Beijing, China

Gianluigi Ferrari, Università di Parma, Parma, Italy

Manuel Ferre, Centre for Automation and Robotics CAR (UPM-CSIC), Universidad Politécnica de Madrid, Madrid, Spain

Sandra Hirche, Department of Electrical Engineering and Information Science, Technische Universität München, Munich, Germany

Faryar Jabbari, Department of Mechanical and Aerospace Engineering, University of California, Irvine, CA, USA

Limin Jia, State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing, China
Janusz Kacprzyk, Systems Research Institute, Polish Academy of Sciences, Warsaw, Poland

Alaa Khamis, German University in Egypt El Tagamoa El Khames, New Cairo City, Egypt

Torsten Kroeger, Stanford University, Stanford, CA, USA

Qilian Liang, Department of Electrical Engineering, University of Texas at Arlington, Arlington, TX, USA

Ferran Martín, Departament d'Enginyeria Electrònica, Universitat Autònoma de Barcelona, Bellaterra, Barcelona, Spain

Tan Cher Ming, College of Engineering, Nanyang Technological University, Singapore, Singapore

Wolfgang Minker, Institute of Information Technology, University of Ulm, Ulm, Germany

Pradeep Misra, Department of Electrical Engineering, Wright State University, Dayton, OH, USA

Sebastian Möller, Quality and Usability Laboratory, TU Berlin, Berlin, Germany

Subhas Mukhopadhyay, School of Engineering & Advanced Technology, Massey University, Palmerston North, Manawatu-Wanganui, New Zealand

Cun-Zheng Ning, Electrical Engineering, Arizona State University, Tempe, AZ, USA

Toyoaki Nishida, Graduate School of Informatics, Kyoto University, Kyoto, Japan

Federica Pascucci, Dipartimento di Ingegneria, Università degli Studi "Roma Tre", Rome, Italy

Yong Qin, State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing, China

Gan Woon Seng, School of Electrical & Electronic Engineering, Nanyang Technological University, Singapore, Singapore

Joachim Speidel, Institute of Telecommunications, Universität Stuttgart, Stuttgart, Germany

Germano Veiga, Campus da FEUP, INESC Porto, Porto, Portugal

Haitao Wu, Academy of Opto-electronics, Chinese Academy of Sciences, Beijing, China

Junjie James Zhang, Charlotte, NC, USA

The book series *Lecture Notes in Electrical Engineering* (LNEE) publishes the latest developments in Electrical Engineering - quickly, informally and in high quality. While original research reported in proceedings and monographs has traditionally formed the core of LNEE, we also encourage authors to submit books devoted to supporting student education and professional training in the various fields and applications areas of electrical engineering. The series cover classical and emerging topics concerning:

- Communication Engineering, Information Theory and Networks
- Electronics Engineering and Microelectronics
- Signal, Image and Speech Processing
- Wireless and Mobile Communication
- Circuits and Systems
- Energy Systems, Power Electronics and Electrical Machines
- Electro-optical Engineering
- Instrumentation Engineering
- Avionics Engineering
- Control Systems
- Internet-of-Things and Cybersecurity
- Biomedical Devices, MEMS and NEMS

For general information about this book series, comments or suggestions, please contact leontina.dicecco@springer.com.

To submit a proposal or request further information, please contact the Publishing Editor in your country:

China

Jasmine Dou, Editor (jasmine.dou@springer.com)

India, Japan, Rest of Asia

Swati Meherishi, Editorial Director (Swati.Meherishi@springer.com)

Southeast Asia, Australia, New Zealand

Ramesh Nath Premnath, Editor (ramesh.premnath@springernature.com)

USA, Canada:

Michael Luby, Senior Editor (michael.luby@springer.com)

All other Countries:

Leontina Di Cecco, Senior Editor (leontina.dicecco@springer.com)

**** This series is indexed by EI Compendex and Scopus databases. ****

More information about this series at <http://www.springer.com/series/7818>

Sukanta Kumar Sabut · Arun Kumar Ray ·
Bibudhendu Pati · U Rajendra Acharya
Editors

Proceedings of International Conference on Communication, Circuits, and Systems

IC3S 2020



Springer

Editors

Sukanta Kumar Sabut
School of Electronics Engineering
Kalinga Institute of Industrial Technology
Bhubaneswar, Odisha, India

Bibudhendu Pati
Department of Computer Science
Rama Devi Women's University
Bhubaneswar, Odisha, India

Arun Kumar Ray
School of Electronics Engineering
Kalinga Institute of Industrial Technology
Bhubaneswar, Odisha, India

U Rajendra Acharya
Ngee Ann Polytechnic
Singapore, Singapore

ISSN 1876-1100

Lecture Notes in Electrical Engineering

ISBN 978-981-33-4865-3

<https://doi.org/10.1007/978-981-33-4866-0>

ISSN 1876-1119 (electronic)

ISBN 978-981-33-4866-0 (eBook)

© The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2021

This work is subject to copyright. All rights are solely and exclusively licensed by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721,
Singapore

Preface

This book is a compilation of selected high-quality peer-reviewed papers from the International Conference on Communication, Circuits and Systems (ic³s2020) held during October 16–18, 2020, at the Kalinga Institute of Industrial Technology, Deemed to be University, Bhubaneswar, India. The book presents the latest development in the field of intelligent computing, advances in communication, signal processing, control systems, VLSI and embedded systems, artificial intelligence, and IoT. The book offers a timely and thorough survey of the latest research in the field of electronics and computing. It consists of papers that share the latest breakthroughs in and promising solutions to the most important issues facing today's society. Written by scientists, academicians, and research scholars from leading institutions and universities, it will be beneficial for readers from both academia and industry.

Bhubaneswar, India
Bhubaneswar, India
Bhubaneswar, India
Singapore

Sukanta Kumar Sabut
Arun Kumar Ray
Bibudhendu Pati
U Rajendra Acharya

Contents

Academic Students Attendance System: A Case Study of Alexa Skill Development	1
K. P. Swain, Soumya Ranjan Samal, I. S. Amiri, M. N. Mohanty, and G. Palai	
Implementation of Low-Cost and Low-Power-Based Temperature and Air Quality Monitoring System for a Local Area in Odisha	7
Rudra Swarup Sahoo, Umakanta Samantasinghar, Adyasa Priyadarsini Acharya, and Priyabrata Biswal	
Comparison Between LSTM and RNN Algorithm for Speech-to-Speech Translator	15
Bageshree Pathak, Shipra Mittal, Komal Shinde, and Pranjali Pawar	
Interactive and Non-interactive Control-Based Lossless Grounded Negative Inductance Simulator Using Current Differencing Buffered Amplifier	23
Shekhar Suman Borah and Mourina Ghosh	
Exploring Novel Techniques to Detect Aberration from Metal Surfaces in Automobile Industries	31
Debaniranjan Mohapatra, Amit Chakraborty, and Ankit Kumar Shaw	
Detecting Diseased Leaves Using Deep Learning	41
Sourodiip Ghosh, Anwita Chakraborty, Ahana Bandyopadhyay, Ishita Kundu, and Sukanta Sabut	
Energy Storage System Analysis and Optimization for Photovoltaic-Based Pico Hydro System in Remote Village	47
Tapas Chhual Singh, G. R. K. D. Satyaprasad, Kali Charan Rath, and P. Srinath Rajesh	
Coplanar Wave Guide Fed Dual Band Antenna Loaded with Metamaterial Split Ring Resonator	53
Mekala Harinath Reddy, D. Sheela, J. Premalatha, and Abhay Sharma	

Design of Low Power Multipliers Using Approximate Compressors	65
Vishal Gundavarapu, M. Balaji, and P. Sasipriya	
Block-Based Local Binary Pattern for Recognition of Handwritten Odia Numerals	73
Suchismita Behera and Niva Das	
Performance Evaluation of Different Machine Learning Techniques for Detection of Non-technical Loss	81
Adyasha Banajyoti and C. N. Bhende	
An Overview with Current Advances in Industrial Internet of Things (IIoT)	89
T. Primya, G. Kanagaraj, and G. Subashini	
Video Indexing Through Human Face	99
Sanjoy Ghatak and Debotosh Bhattacharjee	
Acquisition and Analysis of Skin Impedance in Parkinson's Disease	109
Revati Shriram, Akshata Shinde, Radhika Nibhande, Anchal Guleria, and Rashmi Atre	
Automation of Soil Nutrient Measurement System and Irrigation Control	117
U. B. Mahadevaswamy, R. Pavan Nayak, M. N. Darshan, Tallam Vineeth Kumar, and S. Gautham Gopi	
Performance Evaluation of Perovskite Solar Cells at Elevated Temperatures	127
Sumanshu Agarwal, Archana Kumari Munda, Vedika Pandey, and Kundan Kumar	
Implementation of Transfer Learning Technique for the Detection of COVID-19	135
Amit Prakash Sen and Nirmal Kumar Rout	
Blockchain Enabled and Changeable Threshold-Based Group Specific Multiple Keys' Negotiation	141
Anindya Kumar Biswas and Mou Dasgupta	
Speech to Sign Language Conversion Using Neuro Fuzzy Classifier	149
Pooja Narshetty and H. Y. Vani	
Design and Development of an Ultra-wideband Millimetre-Wave Antenna for Short-Range High-Speed Communication	157
Debraj Dhang, Satyadeep Das, and Sudhakar Sahu	
A Comparative Analysis of BTC Variants	163
D. Nayak, K. B. Ray, and T. Kar	

Crosstalk Noise Reduction in Long Wire Interconnects Using MTCMOS Inverters	171
Jayashree Mallidu and Saroja V. Siddamal	
GPS and GSM Enabled Smart Blind Stick	179
Sourodip Ghosh, Moinak Bose, and Ankit Kudeshia	
Prediction of Speed for Smart Insulin Pump Utilizing Adaptive Neuro-fuzzy Inference System and ANN	187
J. V. Alamelu and A. Mythili	
Curvelet Transform and ISODATA Thresholding for Retinal Vessel Extraction	195
Sakambhari Mahapatra, U. R. Jena, and Sonali Dash	
Design a T-Shape Cantilever Beam Using by Scilab and COMSOL	205
Vasagiri Suresh, Burra Rajesh Kumar, and Vankara Jyothi	
Design and Analysis of PZT-Based Piezoelectric Speakers	215
Vasudha Hegde, Christina David, E. Shirley Jesseca, Shreya Nadgouda, and S. M. Vrinda	
Drowsiness and Yawn Detection System Using Python	225
Surabhi Kumari, Kumari Akanksha, Sasmita Pahadsingh, Swati, and Supriya Singh	
Intelligent Analysis of X-Ray Images for Detecting Bone Abnormality in Upper Extremities	233
Puja Kalivarapu, Rufina Flora George Rajan, and B. Niranjana Krupa	
Swarm Optimization of Multiple UAV's for Resource Allocation in Humanitarian Aid and Disaster Relief Operations	241
Anant Shukla, Rishav Choudhary, and Malavika Prabhuram	
Hardware Design of a Turbo Product Code Decoder	249
Gana C. Nair, B. Yamuna, Karthi Balasubramanian, and Deepak Mishra	
Multiple Object Detection and Tracking Using Deep Learning	257
Shreyas Burde and Suneeta V. Budihal	
Blood Donation Management System Using Android Application	265
Ahana Bandyopadhyay, Ishita Kundu, Anwita Chakraborty, Rajat Kumar, Ayush Kumar, and Sukanta Sabut	
Reinforcement Learning for Improving Coherence of Multi-turn Responses in Deep Learning-Based Chatbots	273
D. G. Suhaas Kiran, Swapneel, Safal Deepak Pansare, and B. N. Krupa	

Intensification of Bandwidth, Return Loss and Gain of Ultra-Wideband Microstrip Antenna with Single-Band-Notch (U-Slot) Characteristics	281
Manisha Mohanty and Bikram Choudhury	
Design of Multi-utility Bottleneck Horn Antenna for Ku-, K-, and Ka-Band Applications	289
Yagnik Mehta, Nisarg Patel, Ronak Vashi, and Jagdish Rathod	
Cyberbully Detection Using 1D-CNN and LSTM	295
Sourodipt Ghosh, Aunkit Chaki, and Ankit Kudesia	
Comparative Study of Skin Lesion Segmentation and Feature Extraction in Different Color Spaces	303
A. Veeramuthu, A. Anne Frank Joe, B. S. Sathish, L. M. I. Leo Joseph, P. Ganesan, and V. Elamaran	
Recognition of Handwritten Digit in Free Space	313
Manas Chandan Behera, B. Shivalal Patro, and Ritik Gupta	
An Area-Efficient 1.23 V Current-Mode Bandgap Reference with Start-Up Circuit	321
Aditya Kumar Hota and Kabiraj Sethi	
Design of AlGaAs/InGaAs/GaAs-Based PHEMT for High Frequency Application	329
Geeta Pattnaik and Meryleen Mohapatra	
Signal Propagation Modelling on the Inner Surface of Human Arm	339
Deepak Kumar Rout, Asha Rani Mohapatra, Rishu Bhardwaj, Raunak Kumar, and Vedant Raj Singh	
Detection of Brain Tumor from MR Images Using BWT and SOM-SVM with Authentication	347
Nilesh Bhaskar Rao Bahadure, Sidheswar Routray, and Arun Kumar Ray	
Wavelet Transform for Signal Compression in Sparse Algorithms	355
Rosalin, Nirmal Kumar Rout, and Debi Prasad Das	
Performance Evaluation Based on Classification of Web Log Data: A Machine Learning Approach	363
Varun Malik, Ruchi Mittal, Jaiteg Singh, and Pawan Kumar Chand	
Energy and Direction Aware Routing Protocol for Flying Ad Hoc Networks	371
Sudesh Kumar, Ram Shringar Raw, and Abhishek Bansal	
Performance Comparison of Tree-Based Machine Learning Classifiers for Web Usage Mining	379
Ruchi Mittal, Varun Malik, Vikas Rattan, and Deepika Jhamb	

Classification and Analysis of Water Quality Using Machine Learning Algorithms	389
Amandeep Kaur, Meenu Khurana, Preetinder Kaur, and Manpreet Kaur	
Extracellular Conductivity and Nerve Signal Propagation: An Analytical Study	399
Satyabrat Malla Bujar, Biswajit Das, and Soumik Roy	
EBN-Net: A Thermodynamical Approach to Power Estimation Using Energy-Based Multi-layer Perceptron Networks	407
Koustav Dutta, Rajarshi Pal, and Rajendra Prasad	
Deep Convolutional Neural Network-Based Knee Injury Classification Using Magnetic Resonance Imaging	419
R. Pandey, S. Mitra, A. Prajapati, A. K. Singh, and V. K. Shrivastava	
Performance Comparison of LSTM Models for SER	427
Tanushree Swain, Utkarsh Anand, Yashaswi Aryan, Soumya Khanra, Abhishek Raj, and Suprava Patnaik	
Harmony Search Optimization-Based Direct Estimation of Harmonic Components	435
Y. Ramya Sree, Bh. Sudharani, K. Sravan Kumar, D. J. V. Prasad, and Ch. Durga Prasad	
Di-zeta Converter-Based Solar-Wind Generating System with a Novel SMES Battery Hybrid Energy Management	443
K. Aseem and S. Selva Kumar	
Design and Implementation of Mixed Signal ASIC-Event Logger in TSMC 0.18 μm CMOS Technology	451
Saroja V. Siddamal, Suhas B. Shirol, Sujata S. Kotabagi, Shraddha B. Hiremath, and Nalini C. Iyer	
Invasive Weed Optimized Drift Parameter Involved Differential Protection Scheme	459
D. J. Abhishek, P. Jagadeesh, N. Srinivasu, and Ch. Durga Prasad	
Breast Cancer Detection and Classification Employing Convolutional Neural Networks and Support Vector Machine for Histopathological Images	467
K. S. Babu and N. R. Sharvani	
Lobe Correction of a Miniaturized Microstrip Antenna Using SIW Cavity	475
Seshadri Binaya Behera, Sraddhanjali Mohapatra, and Subhrakanta Behera	
Implementation of Custom DRC in TSMC 0.18 μm for Optimized Layout	481
B. H. Shraddha, Nagaratna Shanbhag, Saroja V. Siddamal, and Nalini C. Iyer	

Design of Current Mode MOS Logic for Low-Power Digital Applications	493
Amit Bakshi, S. N. Mishra, and Sandeep Kumar Dash	
Augmented Reality as a Supported Educational Method for Embedded Devices and Technology	501
Sambit Prasad Kar, Maitrish Ghosh, and Nirmal Kumar Rout	
Reweighted Zero-Attracting Modified Variable Step-Size Continuous Mixed p-Norm Algorithm for Identification of Sparse System Against Impulsive Noise	509
Ansuman Patnaik and Sarita Nanda	
A Semantic-Based Information Retrieval System	517
Alka Ranjan and Soumya Priyadarshini Panda	
Multiband Ultrathin Terahertz Metamaterial Absorber for Sensing Application	525
Shruti and Sasmita Pahadsingh	
Frequency Estimation Using Adaptive Algorithm	533
Sananda Kumar	
An Improved Ridge Regression-Based Extreme Learning Machine for the Prediction of Diabetes	541
Priya Das and Sarita Nanda	
MIMO-OFDM Outdoor Channel Estimation Using Sparse Momentum Fractional Adaptive Filter	549
Swetaleena Sahoo, Sarita Nanda, and Harish Kumar Sahoo	
Real-Time Electromyographic Hand Gesture Signal Classification Using Machine Learning Algorithm Based on Bispectrum Feature	559
Deepa Nair, R. Rajapriya, and K. Rajeswari	
Design of Microwave Filters Using Swarm-Based Optimization Techniques	565
Piali Chakraborty, Arindam Deb, and Jyoti Ranjan Panda	
Author Index	573

About the Editors

Prof. Sukanta Kumar Sabut received the B.E. and M.Tech. degrees in Electronics Engineering from the VT University of Karnataka in 2005 and the Ph.D. degree in Medical Science and Technology from IIT Kharagpur, India, in 2011. Presently, he is working as Associate Professor in the School of Electronics Engineering, KIIT Deemed to be University, Bhubaneswar, Odisha, India. Dr. Sabut is a member of IEEE, IET, IFESS, Rehabilitation Council of India, and the Institution of Engineers (India). He is the author of 75 articles in journals and conferences having Google scholar citations of 872. He has over 20 years of experience in both teaching and research. His research interests include biomedical signal and image analysis, machine learning and deep learning for health care, and neural and rehabilitation engineering.

Prof. Arun Kumar Ray is currently working as Professor and Director, School of Electronics Engineering and KIIT Deemed to be University. He is now involved both in teaching and administration. He has completed his Ph.D. from IIT Kharagpur. He has been with KIIT since the last 21 years. He has published many research papers in peer-reviewed international/national journals and conferences. He is also having active memberships of IEEE Signal Processing and IET and lifetime member of ISTE. During his teaching professional, he had taught subjects like digital signal processing, digital image processing, and computer networking. His research interest includes image processing and computer networks.

Prof. Bibudhendu Pati is Head in the Department of Computer Science, Rama Devi Women's University, Bhubaneswar, India (Only Government Women's University in the state of Odisha, India). He received his Ph.D. degree from Indian Institute of Technology Kharagpur, India. He has around 22 years of experience in teaching and research. His areas of research interests include wireless communication, cloud computing, big data, Internet of things, and advanced network technologies. He has got several papers published in reputed journals, conference proceedings, and books of international repute. He was General Chair of ICACIE 2016, IEEE ANTS 2017, ICACIE 2018, and ICACIE 2019. He is the life member of Indian Society for Technical Education, Computer Society of India, and a senior member of IEEE.

Prof. U Rajendra Acharya is a senior faculty member at Ngee Ann Polytechnic, Singapore. He is also (i) Adjunct Professor at the University of Malaya, Malaysia, (ii) Adjunct Faculty at Singapore Institute of Technology, University of Glasgow, Singapore, and (iii) Associate Faculty at Singapore University of Social Sciences, Singapore. He received his Ph.D. from National Institute of Technology Karnataka (Surathkal, India) and D.Eng. from Chiba University (Japan). He has published more than 630 papers, in refereed international SCI-IF journals (345), international conference proceedings (42), books (17) with 26,681 citations in Google Scholar (with h-index of 76), and ResearchGate RG Score of 47.76. Presently, he is Editor of *Knowledge-Based Systems Journal*, *BME Online Journal*, *International Journal of Neural Systems* etc. His major academic interests are in biomedical signal and imaging processing, data mining, and biophysics for better healthcare systems. He is ranked in the top 1% of the Highly Cited Researchers (2016 and 2017) in Computer Science according to the Essential Science Indicators of Thomson.

Academic Students Attendance System: A Case Study of Alexa Skill Development



**K. P. Swain, Soumya Ranjan Samal, I. S. Amiri, M. N. Mohanty,
and G. Palai**

Abstract ‘Alexa’ which is developed by Amazon Lab 126 is an intelligent virtual assistant used to improve the user’s comfort. Its expanded technology and modern range of facilities empowers the users to connect the different things wirelessly. Due to its compatibility with Android and iOS platform and ever growing applications, peoples are using Alexa as a house assistant. It is able to respond numbers of voice commands like maintaining to-do-list, alarm setting, controlling electrical appliances, providing world information. It can also be used as a gateway device which accepts the voice commands coming from different IoT enabled devices. Generally, upon requesting, the task performed by Alexa is called ‘Alexa Skills’ which is essentially a voice-driven Alexa app. In this work, the basics of Alexa skill development is discussed by building a simple custom skill called ‘Academic Students Attendance System’ which can be implemented in any educational institution where the number of students present in the class can be obtained from the voice command.

Keywords Alexa · Skill development · Hotword · Amazon · Lambda server

K. P. Swain (✉) · G. Palai

Department of Electronics and Communication Engineering, GITA, Bhubaneswar, India
e-mail: kaleep.swain@gmail.com

S. R. Samal

Faculty of Telecommunications, Technical University of Sofia, Sofia, Bulgaria

I. S. Amiri

Computational Optics Research Group, Advanced Institute of Materials Science, Ton Duc Thang University, Ho Chi Minh City, Vietnam

Faculty of Applied Sciences, Ton Duc Thang University, Ho Chi Minh City, Vietnam

M. N. Mohanty

Department of ECE, I.T.E.R., SoA University, Bhubaneswar, India

1 Introduction

Voice technology has travelled a long distance from performing a simple task to controlling different things, and most importantly, it is easier than a mobile app. So in the coming future, it will play a vital role to become a next big thing. Every year the selling of Alexa from Amazon is exponentially growing. Once it is set up (first time) by the help of a mobile app from an Google app or Apple app store, it can be operated flawlessly through a Wi-Fi network. It has many advantages over a mobile app like; fastest way of operation, hands free and multitasking experience, achieved task by keeping away from the mobile screen, optimal choice for blind and hand disabled persons and appropriate for old or uneducated people who finds tough to work with the smartphone. Some of the cotemporary voice control devices are Cortana (Microsoft), Siri (Apple) and Google Assistant (Google). Alexa is an artificial intelligence personal assistance made by Amazon. It powers Alexa echo and Alexa Echo dot smart speakers. The Alexa smart speaker is a hardware device with speakers and microphone built into it. It has audio processing unit which converts analogue audio signals into digital audio. The audio is then transferred to Amazon (AWS) servers to further process it. The digital audio signals are first converted into text using speech to text software, and then, the meanings are extracted using natural language processing software. Alexa starts listening to user commands by detecting hot word. For example, ‘Alexa’ is a hotword following which commands can be spoken for actions to be performed. Some of the literatures which are involved in some application related to Alexa skill are discussed in [1–6].

In [1], an intelligent virtual assistant Alexa’s ecosystem using a wireless smart speaker is discussed for forensics application. In the system, both mobile and Web application are developed to support the digital examination where wireless smart speaker acts as gateway for Alexa. An attendance system is developed in [2] which is worked on real-time basis using Alexa where inputs are taken from the students after Alexa is initialized by teacher. A local server (Raspberry Pi) is also used which acts as a real-time database to update the students attendance. The entire processing along with the authentication is taken care by Alexa skill sets. A review is conducted in [3] to obtain the customer reaction (emotion analysis) about the voice interactive devices which are used for home application. More than one lakhs of samples are collected regarding customer emotion by considering Amazon Echo and DNight MagicBox, and it is found Amazon Echo is very popular among the people where it is treated as a family member. A study is carried out in [4], whether the data collected by Alexa or similar type of device can be placed before the law enforcement body or not and it is revealed from the forensic analysis that the footprint of these type of devices can be used as a proof in the court of law. The alertness for patients whom are suffering from FITS seizure is performed by using Alexa along in companion with IoT devices [5]. Here, with the help of sensors equipped in IoT, wearable gadget senses the patient condition and accordingly transmits the voice signal to the patient, and if the sensors do not receive any return message from the patient, then the gadget raises the alert to different concern body. An extension of home automation with

more security features is discussed in [6] where the front door of a home is operated by Alexa. These facilities can be performed remotely by a smartphone, and also, the status of a door (open or close) can be read through the phone with the help of a local server like Raspberry Pi.

2 Basic Working Principle

Figure 1 demonstrates the basic working principle of Alexa where automatic speech processing along with natural language processing is used as in-build technology in Alexa to know the actual function of the voice command. Aside this, a cloud server is also used which executes the code with the help of lambda function.

Alexa skills are essentially programs with converted speech data to text data as input parameters by speech and language processing. There are a number of functions that can be programmed to perform specific tasks for specific commands. The function which performs an action on receiving command is known as ‘intent’. A skill can have multiple intents which maps to multiple speech commands. Intents can be defined in Alexa skill console through Internet, and intent can have multiple variants of a speech command which can be defined in the console. The intent is then passed to the programming environment where the programming environment is basically an Amazon lambda function. Amazon lambda functions are a part of AWS (Amazon Web server) service which is employed as a server less application. A lambda function takes in input as ‘http request’ and returns ‘http response’ to the client in case of a Web application or calls the other application program interfaces (APIs) and returns a speech response in the form of JavaScript Object Notation (JSON) object in case of an Alexa skill. Amazon provides skill development in Python and JavaScript at the time of coding.

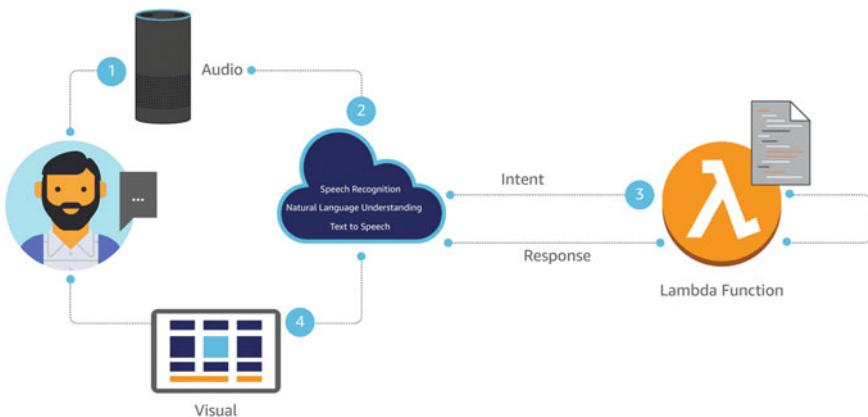


Fig. 1 Demonstration of Alexa working principle

3 Implementation

In this case, REST-representational state transfer (RESTful) API is extensively used for the development of entire application. Here, the storage of data is carried out in EpsumThings server, and business logic is executed in AWS lambda. Both of the systems are kept in synchronization through RESTful API. REST API is essentially a request and response model from client and server, where client (lambda function in this case) sends request to the server, and response is sent back from sever after the operation is successful.

The implementation of the attendance system is done as per the flow diagram shown in Fig. 2. At the first step, Alexa skill is created through Alexa skill console where developer can choose whether to use the backend at their own server or use Amazon lambda function. In this case, the backend is implemented on AWS lambda. After selection of skill development backend, intent definition is presented where speech intents can be defined. It can be implemented either by graphical user interface or bulk intents by using JSON objects. Intent can have multiple hotwords which can be defined in the console to perform a same action for a set of keywords or sentence. Once the intents are defined, they are connected with backend program using lambda function. Here, the backend program is using Alexa skill API to receive parameters parsed from the voice intent and passed to the backend functions. In this case, Python is used as the backend language which takes the parsed parameters as function arguments and string is returned as voice response to the user.

Figure 3 indicates the data flow between Alexa and a local server through API. In this case, the data about the attendance is stored in the EpsumThings platform which is updated using a simple API call. Here, EpsumThings is the name of online cloud server which is used to store the daily attendance data. The data stored in the platform is then pulled using REST API provided by EpsumThings platform through a simple ‘GET’ request. The attendance is represented by a number which shows number of students present in the class. When a professor asks Alexa using hotword ‘What is today attendance’ to know about the present attendance, the Alexa sends the sound signals to the Amazon skill server which in turn parses the audio signal into text. The text is then sent to the AWS lambda for processing. The lambda function gets the parameters as payload and calls API to get the attendance data from EpsumThings server and responds the user as speech about the attendance of the class. As the

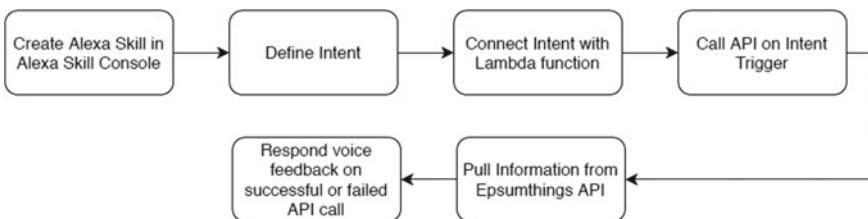


Fig. 2 Flow diagram for Alexa skill implementation

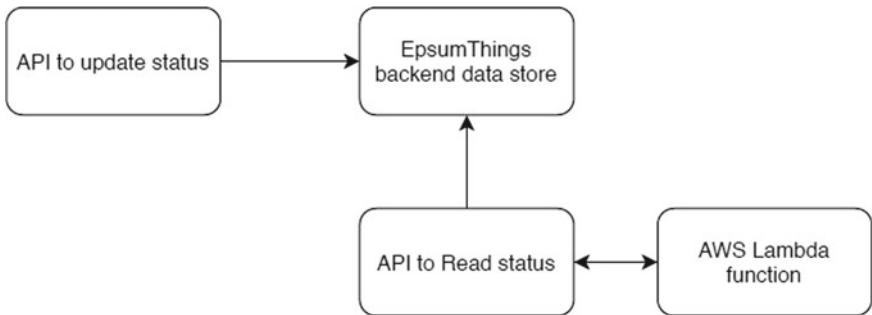


Fig. 3 Flow diagram for data flow between Alexa and EpsumThings through API

above procedure cannot be represented in text, so the above block diagram (Figs. 2 and 3) is used to represent the entire process. In this process, one has to count the number of students present in the class and update the database in regular basis. Also, sometimes the process can be failed due to any server or network error.

4 Conclusion

A basic attendance system is successfully developed using Alexa skill development in this work which can send the voice answer about the daily student attendance after being asked by any person. Here, EpsumThings (online server) is used to store the daily attendance using API at each day. Upon using the hotwords ‘what is today attendance’, the intent is passed to the AWS lambda function by Alexa for processing. Then, the attendance data which is stored in the ‘EpsumThings’ server is fetched to Alexa, and at last, Alexa is responded the answer through the voice message.

References

1. H. Chung, J. Park, S. Lee, Digital forensic approaches for Amazon Alexa ecosystem. *Digit. Invest.* **22**, S15–S25 (2017)
2. A.S. Byalpi, Anush, Alexa based real-time attendance system, in *Proceedings of the International Conference on Communication and Electronics Systems (ICCES)* (2018), pp. 121–124
3. Y. Gao, Z. Pan, H. Wang, G. Chen, Alexa, my love: analyzing reviews of Amazon echo”, in *IEEE Smart World, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (Smart World/SCALCOM/UIC/ATC/CBDCom/IOP/SCI)* (IEEE, 2018), pp. 372–380
4. D.A. Orr, L. Sanchez, Alexa, did you get that? Determining the evidentiary value of data stored by the Amazon® echo. *Digit. Invest.*, 72–78 (2018)

5. M.S. Shaik, R. Tamilkodi, V. Subramanyam, Detection of FITS seizure by Alexa using IoT, in *International Conference on Vision Towards Emerging Trends in Communication and Networking* (2019)
6. S. Srikanth, S.K. Saddamhussain, P. Siva Prasad, Home anti-theft powered by Alexa, in *International Conference on Vision Towards Emerging Trends in Communication and Networking* (2019)

Implementation of Low-Cost and Low-Power-Based Temperature and Air Quality Monitoring System for a Local Area in Odisha



Rudra Swarup Sahoo, Umakanta Samantasinghar,
Adyasa Priyadarsini Acharya, and Priyabrata Biswal

Abstract Environmental pollution is a major issue for public health and safety. Air contamination is one of the inherent causes of environmental issues. The reported monitoring and tracking systems have low precision, low sensitivity, and demonstrated laboratory scaled devices which cannot be accessed by the public. Therefore, improved observatory systems are highly essential. To overcome the limitations of reported systems, we propose a three-step air pollution monitoring system, where we can evaluate the data in three stages, that is a public display, website, and mobile application. An IoT kit was prepared using the MQ-135 Gas sensor, DHT11, LCD, and NodeMCU. We have deployed our developed system in our institution to provide local awareness among the general public and academic individuals as well as a platform for research and development under the smart city application. We can also monitor the data such as temperature, humidity, dew point, air quality index on a website so that individuals will be able to acquire relevant data from the database. A normal user or a person suffering from respiratory disorder is also capable of monitoring the data and will be able to take the required actions to prevent the upcoming unwanted situations. The power consumption and cost of the proposed system are the two important features of any modern devices. Hence, the authors here also give a sight of the power consumption as well as prospective of powering IoT devices through energy harvesting, which is not discussed in many reported literatures.

Keywords Air quality index · Sensors and IOT · Energy harvesting · Low-power experimental set-up · Power consumption

R. S. Sahoo (✉) · U. Samantasinghar
B.J.B. (Autonomous) College, Bhubaneswar 751014, India
e-mail: rudraswarup99@gmail.com

A. P. Acharya
Capital Engineering College, Bhubaneswar 752055, India

P. Biswal
National Institute of Technology Rourkela, Rourkela 769008, India

1 Introduction

The sensors and actuators are actively participating in the field of structural health monitoring, air quality monitoring in smart city applications as well as human activity monitoring. Air pollution is one of the largest problems which has a drastic impact on the environment. Air is getting contaminated because of the release of toxic gases by different organization, increased vehicles, and exponential rise in concentration of harmful gases and particulate matters [1], which in affecting human as well as animal health. There are some factors present in the air like oxygen, ozone, carbon dioxide, carbon monoxide, methane, ammonia, particulate matter, etc. [2]. Air pollution causes harmful diseases like asthma, which motivates researchers to monitor the air quality and analyse of real-time data. The current work demonstrates a real-time air quality observatory system which is IoT enabled. This technology is playing a very useful pillar in our air quality monitoring system.

As shown in Table 1, there are many authors with their models that show different types of parameters of the air to calculate the air quality value. The parameters that are listed in the table can be easier to understand for the experts, but it is not that easy for a simple citizen. The reported monitoring system consumes high power and expensive. Some need costly sensors and some costly software to monitor this big amount of data. These models consuming more power also which is also a disadvantage. Many of these models cover a larger area like a smart city [3], and some also cover out of the city data [4]. Some mobile models can move from one location to another using Drone [5] and GPS [6]. To measure air quality, there is an air quality index (AQI) which is different according to the country. As per Central Pollution Control Board

Table 1 Comparison among different monitoring systems reported in literatures

Ref.	Parameters analysed	Power consumption (mW)	Total cost (INR) (k)	Application area
1	CO, CO ₂ , PM, NO ₂ , SO ₂ , O ₃	1712	495	Specific areas
2	PM, CO, CO ₂ , Temperature, humidity, pressure	16,370	5	Outdoor
3	Temperature, humidity, dust CO, O ₃	2620.45	4.5	Wide area
4	PM	15,595	6 k	Academic
5	O ₂ , O ₃ , CO, CO ₂ , NO ₂ , NH ₃ , CH ₄	3010	20.2	Large area
6	SO ₂ , NO _x , O ₃ , CO, PM, VOC	1205	3.2	Outdoor
7	CO ₂ , Temperature, VOC	750	3.03	Indoor

of India, if the data is between 0–50, PPM then it is good, if it is between 51 and 100, then satisfactory, 101–200 is moderate, and 201–300 is poor. More than 300 is very much unhealthy which is mainly found in coal mines [1]. The paper is organized as follows: Section 2 presents overall system describing different modules as well as how the data accessibility is efficiently improved. Section 3 discussed experimental results as well as power consumption of different modules and prospective of energy harvesting for the developed system. Finally, the Section 4 summarizes the current work.

2 Overview of the System

The current work focuses on low-cost cum low-power air quality observatory and accessing system that consumes a very less amount of power as 1.43 W and shows real-time data. Not only experts but also everyone can monitor the data. If we send the data to the cloud only, then only connected persons can monitor the data and the public cannot see this and they will not aware of the problem. So an LCD is connected to display the real-time sensor data in public purpose. This model is mainly designed for educational institutions. The students need a healthy environment so that they will be healthy physically and mentally. If an air quality monitor is installed on the campus, then it will keep providing the real-time data to the authorities in the database and the students on the LCD. This data can be stored and monitored in the database so that authorities can take the required steps. The overall block diagram of the air quality monitoring system is given in Fig. 1.

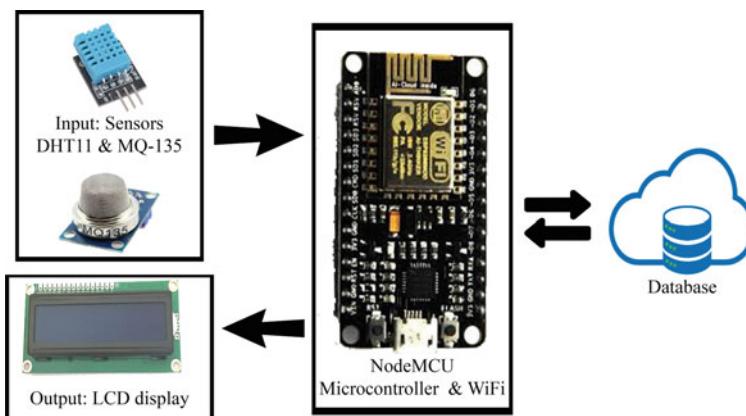


Fig. 1 Block diagram of overall system

2.1 NodeMCU Unit

Here, we used the NodeMCU board because it is a low-cost open-source IoT platform. It is included with a WiFi chip which is mounted with the Microcontroller unit. We are utilized this because of its low-cost, low-power consumption capability, and small compact size.

2.2 DHT11 Sensor and MQ135 Sensor

DHT11 senses two factors, temperature and humidity, using the substrate used in the sensor. We have calculated the dew point value using the relation between temperature and relative humidity as in (1) [7].

$$T_d = T - ((100 - RH)/5) \quad (1)$$

where T_d is dew temperature ($^{\circ}\text{C}$), T is the temperature ($^{\circ}\text{C}$), and RH is relative humidity (%). MQ-135 is the sensor that can give the air quality data with proper calibration and air quality is measured in parts per million (PPM). We need to convert this to logarithmic value to get the data in PPM and is given in (2) [1], where X_{PPM} is the air quality value in PPM, y is the ratio of resistances, m is the slope, b is the y-intercept.

$$X_{\text{PPM}} = 10^{(\log(y)-b)/m} \quad (2)$$

2.3 Data Storage and Accessibility

The data has been monitored through smart modules and collected at the database through the IoT devices. We can see the real-time data of the sensor value live on the 20×4 LCD screen as well as the data will be stored in the Cayenne database.

3 Experimental Set-Up and Discussion

3.1 Experimental Procedure

As shown in Fig. 1, the individual modules are initially tested and then interfaced with each other using Arduino. The power supply to the individual systems is given through battery. However, the devices can be powered through different energy harvesting techniques. The experimental set-up along with prototype, which has been

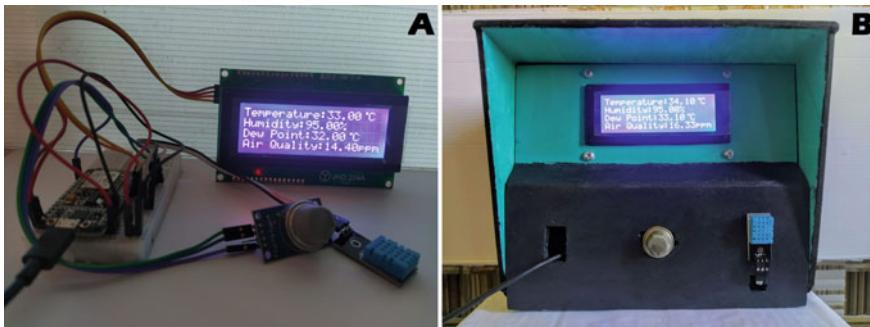


Fig. 2 Experimental Set-up and prototype **a** Interconnection of Modules, **b** Final packaged module

Table 2 Power consumption of different modules

S. No.	Component used	Characteristics/feature	Min. Operating conditions		Power consumption (W)
			Voltage (V)	Current (mA)	
1	NodeMCU	Microcontroller and WiFi chip	5	160	0.8
2	LCD	Output display	5	50	0.25
3	DHT11	Temperature and humidity	3.3	40	0.132
4	MQ-135	Air quality	5	50	0.25

commissioned in the premises of B.J.B College(academic), Bhubaneswar, is illustrated in Fig. 2. The power consumption of different modules has been experimentally observed and is given in Table 2.

3.2 Results and Discussion

The data from sensor module can be continuously analysed and monitored for 24×7 h. For manifesting the characterization of different parameters such as temperature, humidity, dew point, and air quality index, a group of data has been considered for a particular day starting from 10 a.m. for 24 h. It can be seen from the Fig. 3a–c the temperature and humidity were changing according to time and weather and dew point according to temperature and humidity. We noticed that the humidity was constant after 10:00 p.m. This condition affected the dew point graph. It can be obvious from Fig. 3a, c, the graph looks very much similar after 4:00 a.m., but there is a difference of 2–3 °C found between temperature and humidity. We can see the data on air quality in PPM in Fig. 3d, where the value stays under 20 PPM throughout

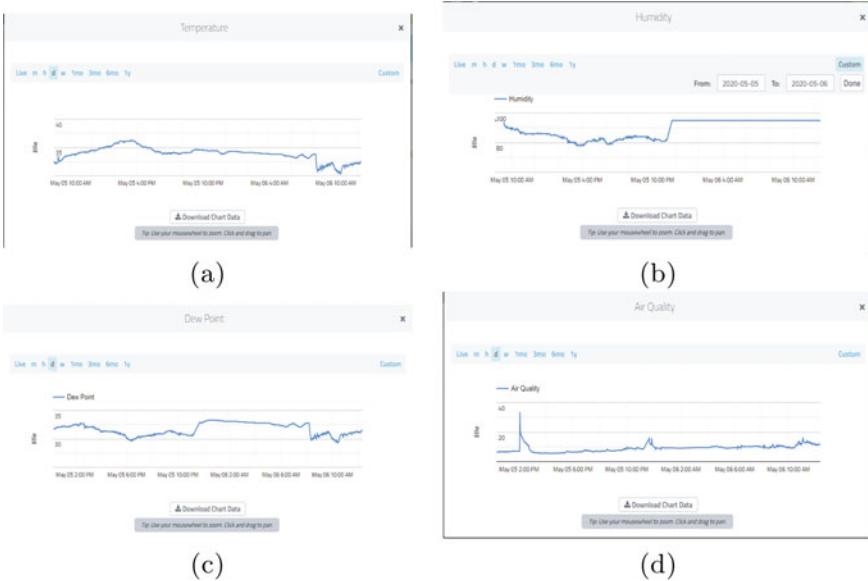


Fig. 3 Sensor data after testing **a** Temperature. w.r.t time, **b** Humidity w.r.t time, **c** Dew point w.r.t time, **d** Air quality w.r.t time

the day, and at 2:00 p.m., it was higher for some reason. It can be dust or any kind of gas exposed to it. As per the AQI, if the value is under 50, then it is good for everyone. The data we got is under 40, so we can say that the air is not polluted. The accuracy of temperature, humidity, dew point, and air quality index in terms of standard deviation of measurement results are found to be ± 0.286 , ± 2.77 , ± 0.563 , and ± 0.436 , respectively. In our developed system, the modules get powered from the battery. However, it is proposed to develop some energy harvesting modules based on piezoelectric and or photovoltaic to harvest the required amount of power such as minimum of 1.432 W. There are widely explored literatures on energy harvesting prospective which can be explored for our work [8, 9].

4 Conclusion

In this work, we have demonstrated a low-cost air quality monitoring system, whose power consumption is also very low and is accessible by general public use and or researchers. The power consumption of individual units as well as from energy harvesting prospective as a future work has also been presented, which has not been discussed in many literatures. This work gives a roadmap for researchers to work on low-power applications as well as to integrate the energy harvesting modules to any activity monitoring systems.

References

1. K.B.K. Sai, S.R. Subbareddy, A.K. Luhach, IOT based air quality monitoring system using MQ135 and MQ7 with machine learning analysis. *Scalable Comput. Pract. Experience* **20**(4), 599–606 (2019)
2. J. Botero-Valencia, L. Castano-Londono, D. Marquez-Viloria, M. Rico-Garcia, Data reduction in a low-cost environmental monitoring system based on LoRa for WSN. *IEEE Internet Things J.* **6**(2), 3024–3030 (2018)
3. S. Kumar, A. Jasuja, Air quality monitoring system based on IoT using Raspberry Pi, in *2017 International Conference on Computing, Communication and Automation (ICCCA) (IEEE)*, pp. 1341–1346 (2017)
4. S. Dhingra, R.B. Madda, A.H. Gandomi, R. Patan, M. Daneshmand, Internet of things mobile-air pollution monitoring system (IoT-Mobair). *IEEE Internet Things J.* **6**(3), 5577–5584 (2019)
5. S. Zhi, Y. Wei, Z. Cao, C. Hou, Intelligent controlling of indoor air quality based on remote monitoring platform by considering building environment, in *2017 4th International Conference on Systems and Informatics (ICSAI)(IEEE)*, pp. 627–631 (2017)
6. Z. Hu, Z. Bai, K. Bian, T. Wang, L. Song, Real-time fine-grained air quality sensing networks in smart city: design, implementation, and optimization. *IEEE Internet Things J.* **6**(5), 7526–7542 (2019)
7. M.G. Lawrence, The relationship between relative humidity and the dewpoint temperature in moist air: a simple conversion and applications. *Bullet. Am. Meteorol. Soc.* **86**(2), 225–234 (2005)
8. P. Biswal, N. Verma, S.K. Kar, B. Mukherjee, Development and performance analysis of a low cost experimental set up for piezoelectric based energy harvester using loudspeaker, in *2018 15th IEEE India Council International Conference (INDICON)(IEEE)*, pp 1–5 (2018)
9. S.M. Antony, S. Indu, R. Pandey, An efficient solar energy harvesting system for wireless sensor network nodes. *J. Inf. Optim. Sci.* **41**(1), 39–50 (2020)

Comparison Between LSTM and RNN Algorithm for Speech-to-Speech Translator



Bageshree Pathak, Shipra Mittal, Komal Shinde, and Pranjali Pawar

Abstract This paper presents the implementation of a speech-to-speech translator using python that can overcome the barrier of different languages. The user can speak in Marathi which will be taken as the input and output will be the translated speech in English. The proposed methodology may be used to bridge the language barrier between a doctor and patient in a rural scenario. The machine learning model used here is sequence-to-sequence model. Keras layers are used which includes encoding, dense, RNN.

Keywords RNN · Speech-to-text · Text-to-text · Text-to-speech

1 Introduction

India is a diverse country having about 22 constitutionally approved languages. Around 1650 dialects are spoken by different communities. As per 2011 Census, the average Indian literacy rate is 69.30%. Fewer than 5% people can read/write English. Two languages that are English and Hindi are regarded as the official languages of India. These languages were made official, so that there can be some common languages which can be used to communicate all over India in common. But, being a democratic country, the languages were not imposed on everyone. So, there are still places all over the country where people only know their local language, and when there is a need, they are not able to communicate with the people outside their region and even the digital media is mostly present in English. All these things make it necessary to have a language translator.

B. Pathak · S. Mittal · K. Shinde (✉) · P. Pawar

Department of Electronics and Telecommunication, MKSSS's Cummins College of Engineering for Women, Pune, India

e-mail: komalshinde1811@gmail.com

S. Mittal

e-mail: shipramttl@gmail.com

So, the purpose of this translator is to get over the language barrier and also the technology barrier. To be specific, we are sticking to a rural scenario where a doctor and patient need to talk with each other and both are unaware of each other's language. In this way, the project is contributing to the society as well.

This translator system can further be extended to include more languages, so that the translator is more usable.

2 Literature Survey

There are four modules in [1], namely voice recognition, translation, speech synthesis, and image translation. The translated language is converted to audio and is given as the output. This application can also accept written text and can convert it into the required language. This application is also enabled to recognize any text present in the image which can be stored in the system or can be captured by making use of camera and translate the text into the desired language. The translation can be displayed on the screen.

Debnath et al. [2] is used for conversion of English text to speech of multiple languages. Input is given by the user as English text to GUI or images of text, and the output is given as translated speech in different language with the help of OCR technique. This methodology can be used to assist people lacking speech or the non-native speakers.

In [3], translation of text is done by a computer without human intervention. Machine translation is a part of NLP that aims to change text/speech of one natural language to another with the help of software. An automatic translation system is built for conversion of text from English to Hindi.

In [4], translation is done by using automated computing. This reviewed various important machine translation system and present comparison of the core technology.

Hermanto et al. [5] presents a comparison between neural-based network that uses RNN and statistical-based network for Indonesian to English and vice versa machine translation. The perplexity value evaluation of both the models is done. The result shows that the use of RNN obtains a much better result. Also, BLEU and RIBES values were found to be increased by 1.1 and 1.6 higher than those obtained using statistical based.

3 Methodology

The implementation is done in three phases:

1. Speech recognition
2. Text-to-text translation
3. Text-to-speech conversion (Fig. 1).



Fig. 1 Block diagram

3.1 *Speech Recognition*

This block is responsible for speech recognition and then converting the speech to text. Firstly, the mic is configured, and then, the chunk size is set which is nothing but the bytes of data that we want to read at once. Then, the sampling rate is decided, and this tells us how often the values are recorded for processing. Since the surrounding noise is variant, the program is allowed some time to adjust its energy threshold of recording, so that it is adjusted in accordance to the external noise level. Finally, the speech is converted to text. All these functions are accomplished with the help of Google Speech-to-Text API.

3.2 *Text-to-Text Translation*

This block is responsible for converting the text output from the previous block (i.e., Marathi) to the other language (i.e., English). It makes use of the neural machine learning model created with the help of Keras library. The steps followed are shown in the flowchart (Fig. 2).

Dataset There is essentially a dataset present in the form of a text file. The file contains pair of Marathi and their translated counterparts in English. This is the dataset that is used for training (80%) and testing (20%). The file is read and is followed by text cleaning, i.e., punctuations removed, uppercases converted to lowercase. The size of total dataset is 38,008 sentences.

Text-to-Sequence Conversion Now, to convert the text to sequence, first the maximum length of sentences in each language is determined by plotting histogram. Then, the text (training and testing) data is vectorized using Keras Tokenizer() class individually for each language set.

Encoder, Decoder Layers The first layer is embedding which is used to convert positive into dense vectors of fixed size. This is the form that deep neural network can understand. Dense representation helps in capturing words relation. In this representation, words with similar meanings are kept closer to each other in the vector space. There are two RNN layers present, the first one encodes. Encoding is done to convert the sequence of symbols into a fixed length vector representation. The second one decodes. Decoding means converting the representation into another sequence of symbols. It applies the activation function on the output.

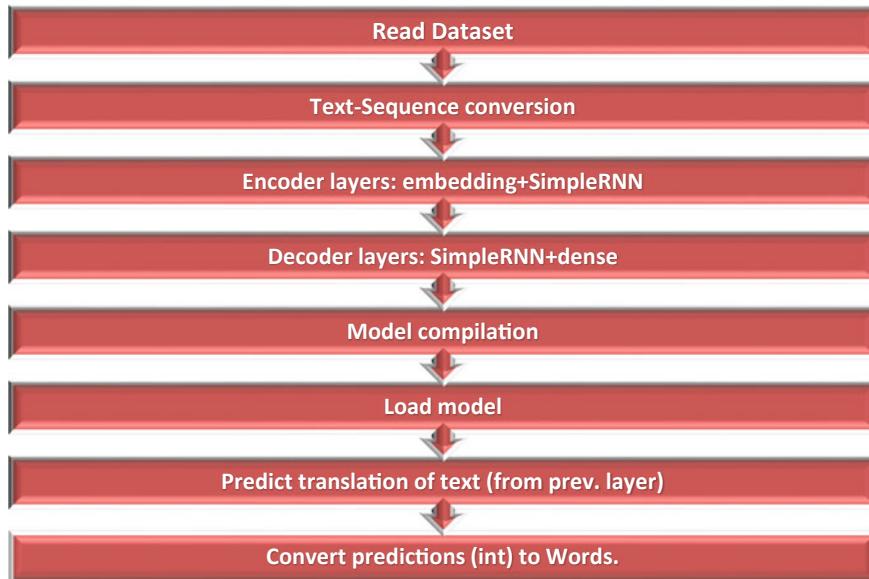


Fig. 2 System architecture

Model Compilation, Saving, and Loading In this section, the model is compiled, followed by saving and loading.

Prediction Using the above model, prediction has been made, and the predictions are converted back to words using Tokenizer() class.

3.3 Text-to-Speech Translation

This block is responsible for converting the translated text to speech. Firstly, the language and pace are chosen. Then, the text is converted to speech and saved in desired format in the machine. This has been implemented using Google Speech-To-Text API, i.e., gTTS API.

4 Result

This result is shown for training and validation for LSTM model where x-axis represents the number of epochs and y-axis represents the loss (Tables 1 and 2; Figs. 3 and 4).

Table 1 LSTM result comparison

S. no.	LSTM actual	LSTM predicted
1	I'll meet you tomorrow	I'll meet you tomorrow
2	Besides teaching English, he writes novels	Besides teaching English, he writes
3	She shot him	She shot him
4	Did you really like it?	Did you really like?
5	My wife is Canadian	My girlfriend is Canadian
6	I'm doing this to help you	I'm doing to help you
7	Why do you need a new television?	Why do you need a new?
8	I'm coming home tom	I'm coming to home
9	I can help you out	I can help you
10	Give me your belt	Give me your cane
11	Whose son is tom	Toms is toms
12	What vegetables do you like?	What games do you like?
13	Do you have a lot of pens?	Do you have books than?
14	I don't get it	It doesn't not
15	They seem American	He burned hard

Table 2 RNN result comparison

S. no.	RNN actual	RNN predicted
1	I'll meet you tomorrow	I'll meet you tomorrow
2	Besides teaching English, he writes novels	Besides teaching English, he writes novels
3	Give me your belt	Give me your belt
4	They left the room one by one	They left the room one
5	Do you like earl grey tea?	Do you like earl tea?
6	She shot him	She shot him
7	Stop yelling	Stop off
8	Your hair is too long	Your are too long
9	Are those your kids?	Are they your children?
10	I don't get it	I don't understand
11	They seem American	He is American
12	What vegetables do you like?	What vegetables do you?
13	My wife is Canadian	My wife is Canadian
14	I'm doing this to help you	I'm trying to help you
15	I left home	I left home

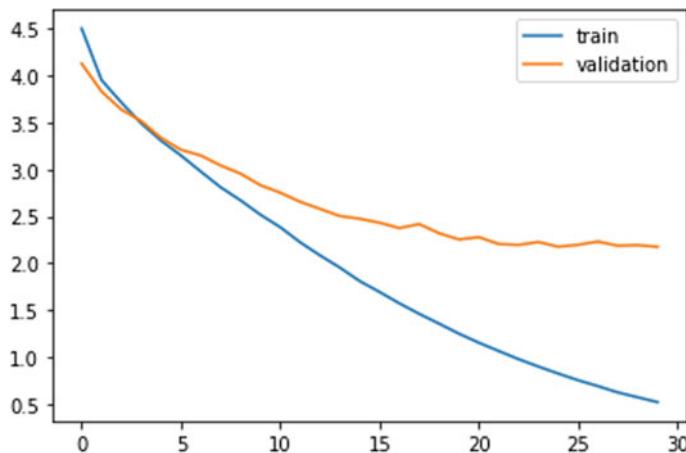


Fig. 3 Plot for loss versus no. of epoch (LSTM model)

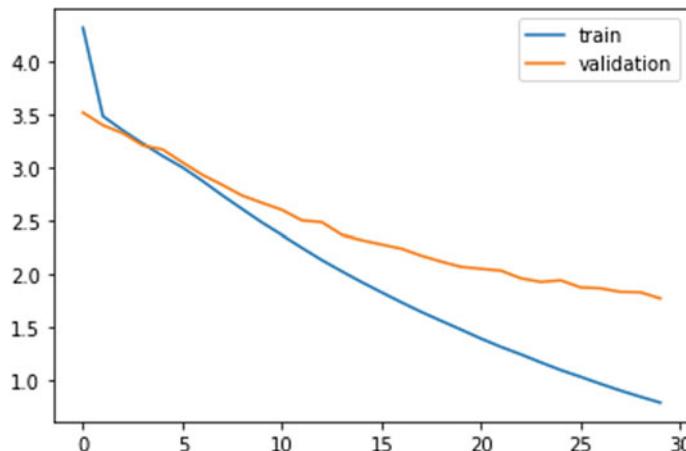


Fig. 4 Plot for loss versus no. of epoch (RNN model)

This result is shown for training and validation for RNN model where x-axis represents the number of epochs and y-axis represents the loss (Table 3).

Table 3 Table for calculation of efficiency of both the models

	Predicted correct	Predicted incorrect	Total sentences
Actual/LSTM	9	6	15
Actual/RNN	11	4	15

5 Conclusion

We have developed a system for speech-to-speech translation, where the Marathi speech is converted to English and vice versa for conversation between a doctor and a patient in rural areas. We have used RNN and LSTM algorithms for the same. It has been seen that RNN performed better than LSTM. The efficiency for LSTM came out to be 60%, whereas for RNN, it came out to be 73.33%.

References

1. A. Lahane, S. Patil, S. Prajapati, M. Phonde, S. Rane, Multilingual speech and text recognition and translation using image multilingual speech and text recognition and translation using image. *Int. J. Eng. Res. Technol. (IJERT)* **5**(04) (2016). ISSN:2278-0181
2. R. Debnath, V. Hanumante, D. Bhattacharjee, D. Tripathi, S. Roy, Multilingual speech translator using MATLAB, in *International Conference on Electrical, Electronics, Signals, Communication and Optimization (EESCO)* (2015)
3. P. Kumar, S. Srivastava, M. Joshi, Syntax directed translator for English to Hindi language, in International Conference on Research in Computational Intelligence and Communication Networks (ICRCICN) (2015)
4. S. Saini, V. Sahula, A survey of machine translation and system for Indian languages, in *IEEE International Conference on Computational Intelligence & Communication Technology* (2015)
5. A. Hermanto, T.B. Adji, N.A. Setiawan, Recurrent neural network language model for English-Indonesian machine translation: experimental study, in *2015 International Conference on Science in Information Technology (ICSITech)*, 27–28 Oct 2015

Interactive and Non-interactive Control-Based Lossless Grounded Negative Inductance Simulator Using Current Differencing Buffered Amplifier



Shekhar Suman Borah and Mourina Ghosh

Abstract Two novel topologies of lossless grounded negative inductance simulator configuration using only one Current Differencing Buffered Amplifier (CDBA) and five passive elements are introduced in this paper. The value of the simulated inductances can be independently controlled without disturbing the condition of realization of inductance, that provides the non-interactive control. The properties of the proposed circuits are validated using PSPICE simulation. The value of the inductance simulator has also been achieved experimentally using commercially accessible Current Feedback Operational Amplifier (CFOA)-based IC AD844AN and found to agree well with the theoretical hypotheses.

Keywords CDBA · Grounded lossless inductor · Negative inductance simulation

1 Introduction

Inductors find broad applications in various areas such as in wireless tuning circuitry, sensors, relays, power storage devices, transformers, oscillators, filters, ferrite beads, phase shifters and parasitic element cancellation. However, a monolithic printed spiral inductor suffers from plenty of shortcomings inherent to the spiral arrangement of the inductors which need a prohibitively extensive chip area that makes them too expensive [1]. Negative inductance is equally important as that of positive inductance. At a fixed frequency, a negative inductance can be observed as a capacitor which acts the equivalent impedance as an inductor but with -90-degrees of phase shift. It may be used for various applications such as cancellation of parasitic inductance, impedance matching in microwave circuits, and generation of chaotic oscillations. [2].

S. S. Borah · M. Ghosh

Department of Electronics and Communication Engineering, IIIT Guwahati, Guwahati, Assam, India

e-mail: borahsumanshekhar90@gmail.com

CDBA, a unique multi-terminal active circuit analog building block, was introduced by Acar and Ozoguz in 1999 [3]. As CDBA can function in both current mode (CM) and voltage mode (VM), a vogue of progression has been observed and adopted in the last few decades which can provide inherently extended bandwidth and linearity, larger slew rate, wide effective range and simple circuitry [4]. A comparative summary of the proposed configurations with earlier reported CDBA-based lossless/lossy inductance simulators is shown in Table 1. It can be perceived that, topologies presented in [4–6] are lossless/lossy floating inductance, but all have positive value. Also, the supply voltage is quite high as compared to the proposed one. Also, the power dissipation is quite high for the configuration proposed in [6]. The topology in [7] implements two circuits; one of them is lossless grounded negative inductance, but it uses two CDBA as an active component.

In this paper, an attempt is made to propose two new lossless grounded negative inductance simulators employing single CDBA as an active element and some passive components. In the first proposed negative inductance simulator, the value of inductance and condition of realization cannot be controlled independently. But the second CDBA-based proposed configuration, negative inductance can be independently varied without disturbing the condition of realization. Both the circuits are simulated in PSPICE using 0.5 μm MIETEC CMOS-based CDBA to verify

Table 1 Comparison with the previously reported papers

References	[4] Figure 2a	[5] Figure 3	[6] Figure 3	[7] Figure 2	Proposed work Fig. 2
No. of CDBA	1	3	3	2	1
R and C counts	2R + 1C	1C	1C	1C	3R + 2C
Technology used	AD844	0.5 μm CMOS	Bipolar process parameters	0.35 μm CMOS	0.5 μm MIETEC CMOS
Transistors used	36	60 + 8 (MRC)	84	78	20
Type of simulated inductor	Lossy positive L	Synthetic floating L	Floating positive L	Floating and grounded negative L	Lossless grounded negative L
Passive component matching requirement	No	Yes	Yes	No	Yes
Non-interactive tuning of L_{eq}	No	Yes	Yes	Yes	Yes
Supply voltage	± 12 V	± 2.5 V	± 3 V	± 1.5 V	± 1.5 V
Hardware verification	Yes	No	No	No	Yes
Power consumption	–	–	7.37 mW	–	0.821 mW

the different properties of negative inductance. An experimental analysis is also performed for the proposed circuits using commercially available IC AD844AN.

2 Circuit Description

CDBA is a two input, two output terminal device. Figure 1 shows the schematic circuit of CDBA [8]. The port relationship of CDBA is expressed by the following equation as:

$$I_z = I_p - I_n, \quad V_w = V_z, \quad V_p = V_n = 0 \quad (1)$$

where V_p , V_n and I_p , I_n are the input voltages and currents at the terminal p and n respectively, terminal z and w are the current and the voltage output terminals. CDBA can be designed by employing MOSFETs with suitable technology or by employing commercially available IC of CFOA, i.e. AD844AN as defined in the literature [8].

2.1 Proposed Lossless Negative Inductance Simulator

The first and second proposed lossless negative inductance Simulator I and Simulator II with interactive and non-interactive control of inductance is shown in Fig. 2. It consists of one CDBA and five passive components.

From the routine analysis, the equation of input admittance for the Simulator I can be obtained as,

Fig. 1 Schematic diagram of CDBA

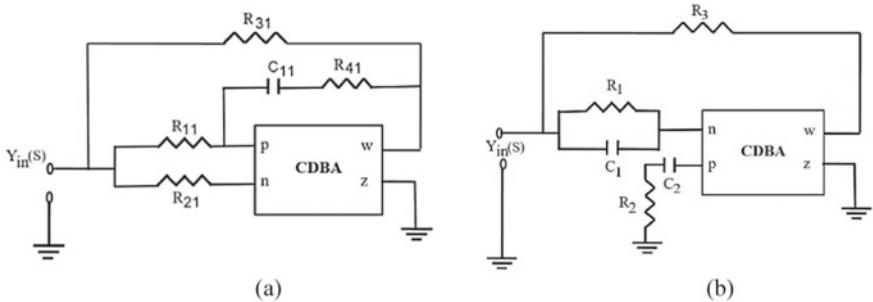


Fig. 2 Proposed lossless negative inductor, **a** simulator I **b** simulator II

$$\begin{aligned} Y_{\text{in}}(s) &= \frac{I_{\text{in}}(s)}{V_{\text{in}}(s)} = \frac{1}{R_{11}} + \frac{1}{R_{21}} + \frac{1}{R_{31}} - \frac{R_{41}}{R_{21}R_{31}} - \frac{1}{R_{21}R_{31}sC_{11}} \\ &\quad + \frac{R_{41}}{R_{11}R_{31}} - \frac{1}{R_{11}R_{31}sC_{11}} \end{aligned} \quad (2)$$

If the condition,

$$\frac{1}{R_{11}} + \frac{1}{R_{21}} + \frac{1}{R_{31}} = \frac{R_{41}}{R_{31}} \left(\frac{1}{R_{21}} - \frac{1}{R_{11}} \right) \quad (3)$$

Then $Y_{\text{in}}(s)$ is purely inductive and if $\frac{1}{R_{11}} = 0$, the condition for lossless negative inductance can be expressed as,

$$L_{\text{eq}} = -C_{11}R_{21}R_{31} \quad (4)$$

Thus, the proposed topology uses one active component and five passive components to realize the negative inductor with R_{11} or without R_{11} . It can be noticed that the inductance value cannot be controlled independently by the condition of realization; i.e. it is interactively controlled.

Similarly, the conventional analysis of the proposed Simulator II appears in an interpretation of input admittance as:

$$Y_{\text{in}}(s) = \frac{I_{\text{in}}(s)}{V_{\text{in}}(s)} = \frac{sC_2(R_3 - R_2)R_3 + sR_1(C_2 - C_1) + s^2C_1C_2R_1(R_3 - R_2) - 1}{sC_2R_1R_3} \quad (5)$$

If $R_2 = R_3 = R$ and $C_1 = C_2 = C$, then $Y_{\text{in}}(s)$ is purely inductive and inductance value can be represented as:

$$L_{\text{eq}} = -CR_1R \quad (6)$$

The sensitivity of L_{eq} for passive components is expressed in (7) and (8) for both the simulators, respectively. It reveals the low sensitivity of the circuits.

$$S_{C_{11}}^{L_{\text{eq}}} = S_{R_{11}}^{L_{\text{eq}}} = S_{R_{21}}^{L_{\text{eq}}} = S_{R_{31}}^{L_{\text{eq}}} = S_{R_{41}}^{L_{\text{eq}}} = 1 \quad (7)$$

$$S_C^{L_{\text{eq}}} = S_{R_1}^{L_{\text{eq}}} = S_R^{L_{\text{eq}}} = 1 \quad (8)$$

3 Simulation Results

The PSPICE simulations are executed by employing the MIETEC 0.5 μm CMOS technology parameters [8]. Supply voltages are, $V_{DD} = V_{SS} = \pm 1.5$ V and bias currents are considered as 25 μA and 30 μA , respectively. Here, only the results obtained from Fig. 2b are displayed to conserve space. To signify the performance of the proposed Simulator II, an inductance of -2 mH is designed using passive element values as $C_1 = C_2 = C = 2$ nF, $R_2 = R_3 = R = 1$ k Ω and $R_1 = 1$ Ω . The theoretical and simulated impedances w.r.t frequency responses are displayed in Fig. 3a. The inset represents the magnified appearance of impedance responses. The inductance value remains within $\pm 10\%$ in the frequency range from 80 Hz to 393 kHz. However, it is found that for lesser values of inductance, the frequency range increases. Inductance and phase response of -2 mH is presented in Fig. 3b, c.

Figure 4 exhibits the voltage and current waveforms for the proposed negative inductance simulator II. It can be observed that the current leads the voltage by 90°. The theoretical peak value of the resulted voltage obtained is 1.256 mV which is very near to the simulated value of 1.271 mV. Therefore, it reveals that the performance of the proposed lossless grounded negative inductance simulator is considerably good.

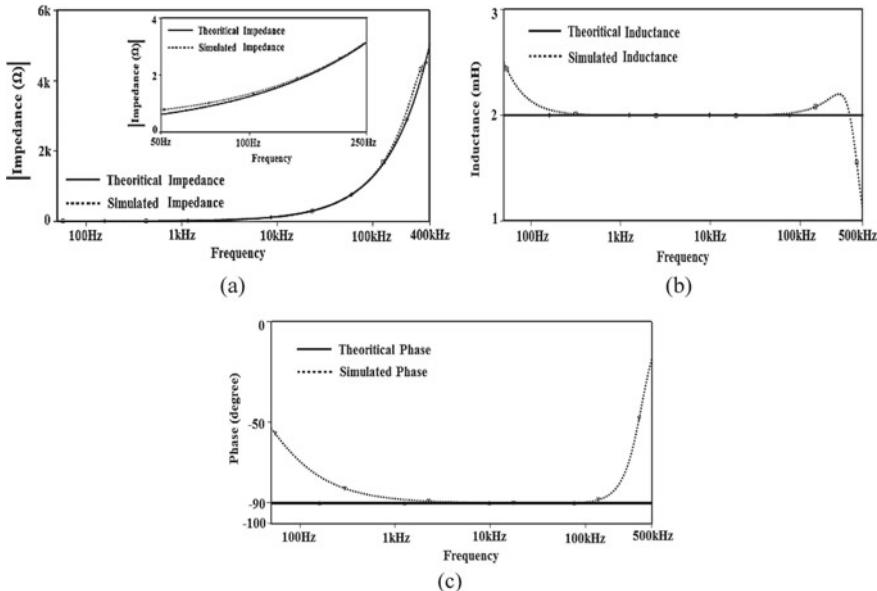


Fig. 3 Proposed simulator II **a** impedance magnitude response, **b** relative change (ΔL) response, **c** phase response of input impedance of $L_{eq} = -2$ mH for CMOS-based CDBA

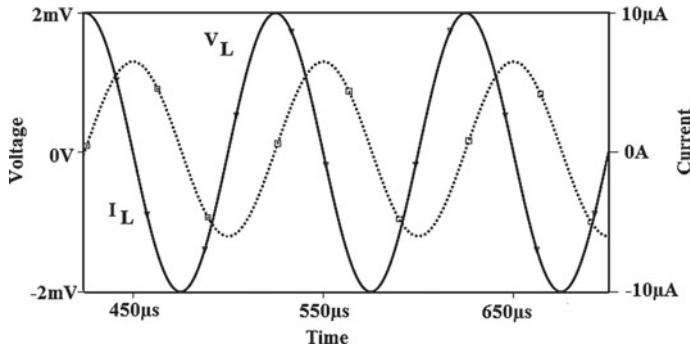


Fig. 4 Voltage and current waveforms ($L_{eq} = -2$ mH) of the proposed simulator II

4 Experimental Verification

The theoretical and simulated propositions of the proposed configurations are also tested experimentally by applying commercially obtainable IC AD844AN [3]. The experimental setup of CDBA using AD844 is displayed in Fig. 5a with a DC power supply of ± 5 V. To design an inductance simulator of $L_{eq} = -1$ mH, the required components are taken as $C_1 = 50$ nF, $C_2 = 50$ nF, $R_1 = 20 \Omega$, $R_2 = 1$ k Ω and $R_3 = 1$ k Ω . A 9 V_{P-P} triangular voltage (V_{in}) is applied through $R_{eq} = 1$ k Ω resistor (to represent a triangular current source of 9 mA_{P-P} with $R_{in} = 10$ k Ω inner resistance). Figure 5b shows the waveform observed on the oscilloscope. From Fig. 5b, the value of rate of change of current is obtained as $\frac{di}{dt} = \frac{9 \text{ mA}}{0.120 \text{ ms}} = 75 \text{ A/s}$ and that of peak value of the square waves as $0.4 \times 200 \text{ mV} = 80 \text{ mV}$. The magnitude of experimental L_{eq} is obtained using the relationship $V = |L_{eq}| \frac{di}{dt}$ as $\frac{80 \times 10^{-3}}{75} = 1.06 \text{ mH}$. It is noticed that the experimental inductance value is little deviated from the theoretical value. This may be due to the tolerance of $\pm 10\%$ of the components as used in the experiment.

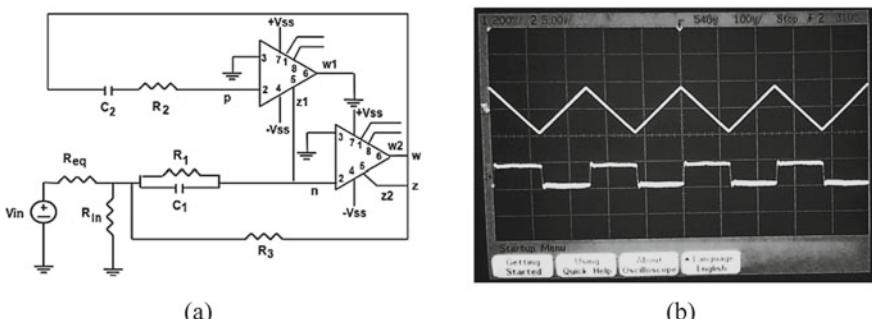


Fig. 5 Proposed simulator II using AD844AN **a** experimental setup, **b** input triangular current and square output voltage waveforms (time/div = 100 μ s)

5 Conclusion

This paper presents two lossless grounded negative inductance simulator using CDBA as an active element. Both the circuit offers interactive and non-interactive control of condition of realization of inductance. The advantageous characteristics offered by the proposed configurations are as follows: (i) Uses of single CDBA, (ii) The capacitors are virtually grounded, which is ease the VLSI implementation, (iii) It has wide frequency range of operation for low inductance and (iv) Low power consumption.

References

1. G. Thanachayanont, A. Payne, CMOS floating active inductor and its application to bandpass filter and oscillator designs. *IEEE Proc. Circ. Dev. Syst.* **147**(1), 42–48 (2000)
2. S. Kilinc, K.N. Salama, U. Cam, realization of fully controllable negative inductance with single operational transresistance amplifier. *Circ. Syst. Sig. Process.* **25**(1), 47–57 (2005)
3. C. Acar, S. Ozoguz, A new versatile building block: a current differencing buffered amplifier suitable for analog signal processing filters. *Microelectron. J.* **30**(2), 157–160 (1999)
4. J.K. Pathak, A.K. Singh, R. Senani, New Canonic Lossy Inductor using a Single CDBA and its Application. *Int. J. Electron.* **103**(1), 1–13 (2016)
5. A.U. Keskin, E. Hancioğlu, CDBA-based synthetic floating inductance circuits with electronic tuning properties. *ETRI J.* **27**(2), 239–242 (2005)
6. W. Tangsrirat, W. Surakampontorn, Electronically tunable floating inductance simulation based on current-controlled current differencing buffered amplifiers. *Thammasat Int. J. Sci. Technol.* **11**(1), 60–65 (2006)
7. Jaikla, W., Siripruchyanun, M., Current controlled CDBAs based novel floating and grounded negative inductance simulators, in *The 21st International Technical Conference on Circuits/Systems, Computers and Communications* (2006), pp. 701–704
8. A.U. Keskin, Multifunction biquad using Single CDBA. *Electr. Eng.* **88**, 353–356 (2006)

Exploring Novel Techniques to Detect Aberration from Metal Surfaces in Automobile Industries



Debaniranjan Mohapatra, Amit Chakraborty, and Ankit Kumar Shaw

Abstract Industry 4.0 and its adaptation is no longer a good to have enabler for all manufacturing industry now. Usage of artificial intelligence in perspective of Industry 4.0 is penetrative and the horizon vast. In today's world, developing an automated system for identification of fracture and its location is an emerging research area in automobile industries. A proper identification of fractures at the time of manufacturing of vehicles leads to both time and economical reclaims with excellent accuracy. In this paper, an automated system is implemented which is able to identify the features with its location with high accuracy. For this, 'OpenCV' resource is used for development of algorithm and de-noising process to extract the features. A machine learning technique, supported by region-based convolutional neural network (R-CNN) is implemented, and a transfer learning mechanism is used by which the system learns itself from raw data and classifies it with a good accuracy.

Keywords R-CNN · AlexNet · OpenCV · SIFT/SURF/ORB · Median filter · Keypoint

D. Mohapatra
S 'O' A University, Bhubaneswar, India
e-mail: mohapatradebaniranjan@gmail.com

A. Chakraborty · A. K. Shaw (✉)
University of Calcutta, Kolkata, India
e-mail: aankitshaw@gmail.com

A. Chakraborty
e-mail: halfblood.percyy@gmail.com

1 Introduction

Digital transformation is the most essential and widely used platform in various business processes. It aims to reduce cost, improve quality, and stay in the competitive age of ever changing business. Nowadays, large and small manufacturing automobile companies have faced problems in various material surface related problems in context of material manufacturing. Automobile industry faces many problems in material surface manufacturing, such as adhesion in metals, lose bonding, non-uniform coating, non-adhesion of paints, etc.

This paper introduces two ways to identify aberration in metal surface. In the first approach, a computer vision and OpenCV library is used for different types of metal image pre-processing and then detecting aberration in metal surface. The second approach elaborates deep learning with both region-based convolutional neural network and AlexNet to classify weather the metal surface is having aberration or not.

2 Literature Review

Dr. D. Selvathi, I. Hannah Nithilla and N. Akshaya in their paper [1] are able to detect the scratch in main surface and subsurface efficiently. They have collected the images from pulsed thermography video and denoised it with median filter. They have collected the information from images by principal component thermography. For segmentation purpose, they have used canny edge detector. After that they have taken some threshold value by convolution with Laplacian filter and also by background subtraction from the image and identified the scratch.

Familiana et al. [2] have carried out the research work on characterising aluminium surface using image processing and ANN. In their work, they have considered some parameters like area, colour intensity, entropy, RGB and HIS colour index, texture, RMSE, etc.

Li et al. [3] have worked on metal surface detection which was mainly focused on metal surface crack detection in spoons used in restaurants. Methodologies that have used in this work is median filtering, first order differential operation, second order differential operator, sobel operator, log operator and the parameters are grey interval of image and image back ground colour.

Mittle and Kerber [4] have worked in metal crack detection using transfer learning process. This transfer learning algorithm is supported by GoogLeNet and AlexNet. They have used supervised learning mechanism and trained the CNN architecture with stochastic gradient descent (SGD). They have got the data from different sources and have got the accuracy above 99.8%, and $F1$ score was 0.836.

3 Proposed Methodologies

In our first approach, an image (img) is taken and converted it into a grey scale image (img1) format. Then, a duplicate of img1 is created and applied de-noising process to it. The output image after de-noising is let say img2. After that we have applied feature extraction techniques to both the images (img1 and img2) and calculated the keypoints. Let, keypoint of img1 = a and keypoint of img2 = b. Now, the concept is if the input image is having any aberration, then the difference between the keypoint of img1 and img2, i.e. " $a - b$ " will be a large value. Mathematically;

$$\% \text{ of } (a - b) = \frac{b}{a} * 100 \begin{cases} \leq 30\%, \text{ Image has scratch} \\ > 30\%, \text{ No scratch} \end{cases}$$

Here, threshold value is set to 30%.

3.1 Solution Approach

Here, real-time computer vision using OpenCV library is used to detect aberration in metal surface. There are no training images present, hence the training images has to be generated from the public image sources. Due to the very nature of the generated images and compute capacity, images had to be scaled-down to 32 * 32 pixels in RGB channel to fit into initial process. The solution is built primarily using Fig. 1.

3.1.1 OpenCV

OpenCV is a library in python, which is designed to solve many real-time computer vision problems [5]. It supports very wide range of programming languages like C++, java, python, etc., in our work, images are imported by installing OpenCV. NumPy library, also called as numeric python is also installed for creating multi-dimensional arrays and matrices.

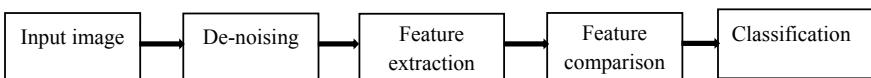
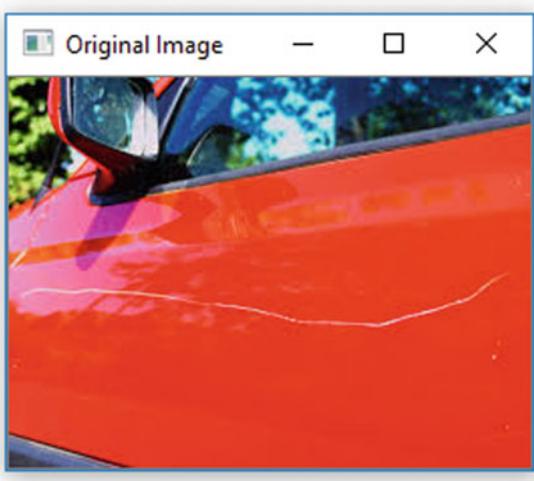


Fig. 1 Block diagram of the entire network

Fig. 2 Original image

3.1.2 De-noising Using Median Filter

For a proper feature extraction, image de-noising should be done; hence, median filter is used in our work for de-noising the input images. Median filter is mainly used to reduce salt and pepper noises. For a grey levelled image, the pixel value will range between 0 and 255 where 0 corresponds to black colour and 255 represents white. Hence, shades of grey colour in between 0 and 255 are obtained. The function cv2.medianBlur calculates the median value of all the pixels and places it in the central pixel. This central pixel value may not be presented inside image itself but, however, it reduces the noises (salt and pepper) quite effectively since the central pixel value continuously replaces by some pixel value of an image (Figs. 2, 3 and 4).

3.2 Feature Extraction

After de-noising feature extraction is done of both the images using STIF/SURF/ORB. STIF stands for scale-invariant feature transform. The advantage of STIF is it is rotation and scale invariant. That means, it will extract the feature from the image whether it is straight or rotate, scaled-up or scaled-down.

3.3 Keypoint and Descriptor

Keypoint is nothing but the point of interest or region of interest of an image. It is simply a point (x, y) on the image which can be used as its characteristics and it is

Fig. 3 Grey levelled image**Fig. 4** De-noised image

scale invariant. Descriptor describes the keypoint by recording the neighbourhood points of keypoint [6]. Detector detects the keypoint by analysing the magnitude and direction of intensity changes in the neighbourhood pixels. Descriptor is basically a binary string that keeps the information about magnitude, orientation, pattern, intensity gradient direction, etc. To get the keypoints, we need to find local extrema of an image. Extrema is a point where the minimum and maximum value of a function is obtained at some time interval.

Let the function be $f(x)$. To find 'Extrema', $f'(x) = 0$ is performed. Where $f'(x)$ is the first order derivative of the function $f(x)$. Now, different points of ' x '

is obtained and it is compared with other 8 neighbourhood pixels and also 9 pixels towards right and left of it. If it is the local extrema, it is considered as the keypoint.

3.4 Feature Comparison

Here, the keypoints of img1 and img2 are compared. Let keypoint of img1 = α and keypoint of img2 = β so now,

$$\text{Percentage of difference}(D) = \frac{\beta}{\alpha} * 100,$$

Finally, the classification stage is followed as,

$$D = \begin{cases} \leq 30\%, \text{ image has scratch} \\ > 30\%, \text{ no scratch} \end{cases}$$

3.5 Output

In this output (Fig. 5), it is visible that we have taken the difference between the keypoints of both the images.

Here (Fig. 6), it is observed that the input image is classified as the scratched image as it is having the differences ≤ 30 (Fig. 7).

Here (Fig. 8), it is observed that the input image is classified as the un-scratched image as it is having the differences > 30 .



Fig. 5 Differences between keypoints of grey scaled and de-noised images

Fig. 6 Identification of scratched image

```
img1_keypoint:274
img2_keypoint:44
%Difference:16.05839416058394
image has scratch
```



Fig. 7 Comparison is done between the un-scratched images

Fig. 8 Identification of un-scratched image

```
img1_keypoint:2503
img2_keypoint:1106
%Difference:44.186975629244905
image has no scratch
```

In our second approach, a deep learning mechanism, supported by R-CNN and AlexNet is implemented. Here, R-CNN stands for regions with CNN. In R-CNN [7] instead of running classification on huge number of regions, the images are passed through selective search region. In this way, instead of classifying huge number of regions, we need to just classify first 2000 regions. This makes this algorithm fast compared to previous techniques of object detection. Here, some selective regions with high probability scratched regions were taken and fed as input to AlexNet architecture. AlexNet is a transferred learning mechanism based on CNN algorithm with two numbers of activation function ReLU and softmax [8]. There are no training images present, hence the training images has generated from public image sources. Due to the very nature of the generated images and compute capacity problem as described the images had to be scaled-down to 32 * 32 pixels in RGB channel to fit into input layer while training (Fig. 9).

Here, MP stands for max-pooling, which is used as an activation function. It is used just after each convolutional layer (CV) so as to minimise the dimension without hampering the image characteristics. Finally, all the obtained neurons are supplied to fully connected layers (FC) and are the scratches are being classified by binary classifier.

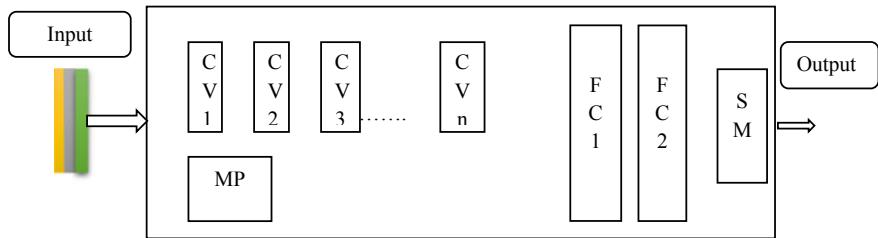


Fig. 9 AlexNet architecture

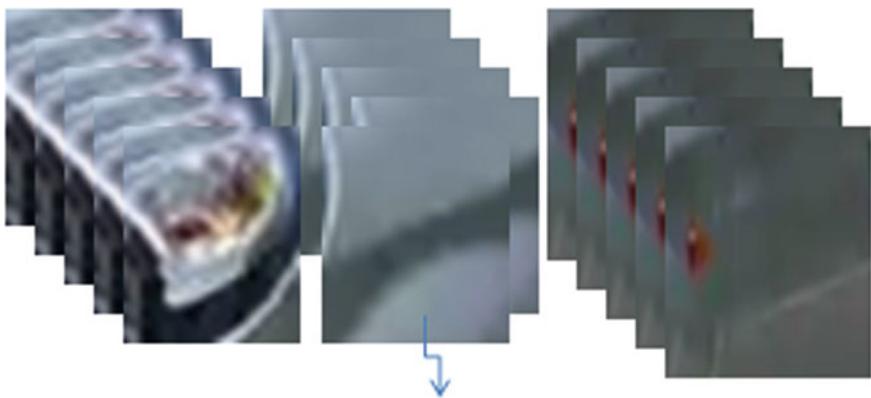


Fig. 10 Region-based images from the input image for scratch classification

3.6 Output

A vehicle image is taken to identify the cracks if present. After taking this image, it was processed for the selection of regions.

Here (Fig. 10), we have taken numbers of selected region from the input image and each region is considered as input to the AlexNet architecture where there images are convolved with numbers of kernel filters. After convolution operation, max-pooling operation works where it extracts the important characteristics from the image with reduction of size.

Finally, the cracks from the images are identified (Fig. 11) based on binary classification technique.

3.7 Comparison Table

A comparison Table 1 is given to explain the degree of performance between the two methods. For OpenCV, a total of 134 data were taken, and for R-CNN 212, data



Fig. 11 Identification of cracks from the input car image

Table 1 Comparison between OpenCV and R-CNN

Techniques	Multiple scratches	Library dependency	Training architecture	Execution time	Accuracy
OpenCV	Less efficient	More dependent	No need of training	Less execution time is required	Good but performance lacks in case of multiple scratches
R-CNN	More efficient	Less dependent	Training model is required	Execution time is more (16 h per 50 epochs)	Excellent

were chosen. The overall accuracy of 97.76 and 99.05% were found in OpenCV and R-CNN methods, respectively.

4 Conclusion

In this work, two techniques of metal crack detection in automobile industries have been proposed. A remarkable accuracy and high percentage of classifications accuracy have been achieved using OpenCV and R-CNN, respectively. The first method is very accurate in case of single crack detection and it is also image size independent, whereas R-CNN works significantly with accuracy above 96% in case of multiple cracks identification. The execution time is comparatively less in R-CNN but it can be better if the entire architecture is run in cloud. This techniques can work only on outer surface cracks detection but the cracks may occur in inner surfaces as well. Hence, in future, we will work in this sector and will try to introduce some sensor-based system which will classify and identify the entire cracked regions from the whole metal body.

References

1. D. Selvathi, I.H. Nithilla, N. Akshaya, Image processing techniques for defect detection in metals using thermal images, in *2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI)*, Tirunelveli, India (2019), pp. 939–944
2. H. Familiana, I. Maulana, A. Karyadi, I.S. Cebro, A. Sitorus, *2017 International Conference on Computing, Engineering, and Design (ICCED)* (2017), pp. 1–6
3. Z. Li, J. Zhang, T. Zhuang, Q. Wang, *IEEE 3rd Advanced Information Technology, Electronic and Automation Control Conference (IAEAC)* (2018), pp. 2365–2371
4. D. Mittel, F. Kerber, *24th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA)* (2019), pp. 544–551
5. D. Zhou, Y. Tian, X. Li, J. Wu, *Chinese Automation Congress (CAC)* (2019), pp. 4695–4699
6. F. Fang, L. Li, M. Rice, J.-H. Lim, *IEEE International Conference on Image Processing (ICIP)* (2019), pp. 2976–2980
7. L. Yu, Z. Yu, Y. Gong, *Int. J. Sig. Process. Image Process. Pattern Recogn.* **8**(5), 117–126 (2015)
8. S. Gibb, H.M. La, S. Louis, *IEEE Congress on Evolutionary Computation (CEC)* (2018), pp. 1–8

Detecting Diseased Leaves Using Deep Learning



Sourodip Ghosh , Anwita Chakraborty , Ahana Bandyopadhyay ,
Ishita Kundu , and Sukanta Sabut

Abstract Plants are an integral part of our ecosystem, and automation in detecting their diseases has intrigued researchers all around the world. In the proposed context, we illustrate a comparison analysis, to detect diseased leaf images of bell pepper and tomato plants. In this research, we have developed a custom-designed CNN architecture and a deep neural network, DenseNet121. The model was evaluated using standard parameters like precision, sensitivity, specificity, F-measure, FPR, and FNR, which guarantees the outperforming ability of pre-trained classifier with respect to the custom CNN. The balanced accuracy (BAC) of CNN and DenseNet121 was 96.5% and 98.7%, respectively, thus outperforming all other works on this particular dataset. The train data size was 80%, and the test data size was 12% with validation as 8%.

Keywords Leaf disease · CNN · DenseNet121 · Image processing

1 Introduction

Infected plants show degree of manifestations like, hid spots, streaks that affect leaves, stems, and seeds of the plant. Diseases can be primarily classified into three varieties, like viral, bacterial, and fungal. By utilizing image processing for illness discovery has become well known. However, detecting a plant's healthiness through a picture is an exceptionally troublesome errand. With intense research in this field, various optimized methodologies were applied for more customized disclosure of plant infections. These disclosure of plant ailments are tremendously investigated fields of research as it naturally perceives the contamination from the signs that appear on the plant leaves. Atabay et al. [2] used PlantVillage dataset that includes healthy and diseased tomato plant leaves, and CNN was applied to categorize the disease affecting tomato leaves. Arivazhagan et al. [1] used 1200 images of mango leaves

S. Ghosh (✉) · A. Chakraborty · A. Bandyopadhyay · I. Kundu · S. Sabut
KIIT University, Bhubaneshwar, India
e-mail: sourodip.ghosh02@gmail.com

consisting of five different diseases and applied the VGG19 model and obtained an overall accuracy of 96.67%. Shrivastava et al. [9] proposed a classification analysis using AlexNet and SVM across 619 rice plant diseased images, with the best accuracy of 91.37% using training and testing ratio of 80 : 20. In recent years, various experiments performed used the abilities of deep learning models and frameworks to effectively accomplish different precision levels on laboratory or field pictures for example in [3] by Konstantinos P. Ferentinos. The CNN framework was applied to detect disease of different plants, and an accuracy of 99.53% was achieved using VGG convolutional neural network on an open database of 87, 848 images including both healthy and diseased leaf images. In [8], authors used 13,262 images from PlantVillage dataset that were provided as input to AlexNet and VGG16 frameworks, and accuracy of 95.65% and 97.49% was obtained using AlexNet and VGG-16, respectively. Fuentes et al. [4] used several frameworks like Faster R-CNN, R-FCN, and SSD where the mean AP of the whole system was more than 80%. In [10], Wang G et al. considered four fine-tuned frameworks or models that were VGG19, VGG16, ResNet50, and Inception-v3 as architectures to classify the apple black rot images in the PlantVillage dataset. Among all the fine-tuned models, VGG16 model achieved an accuracy of 90.4%. As mentioned in [11], Zhang X et al. used 9 sorts of maize leaves that were distinguished, the two improved deep convolutional neural system models using CIFAR10 and GoogLeNet, can accomplish high recognizable proof precision, 98.9%, and 98.8%, separately. Mohanty et al. [7] applied CNN to identify 14 crop species and 26 diseases where accuracy achieved was 99.35%.

2 Proposed Methods

The authors provide a suggestive method by incorporating a deep convoluted neural network, DenseNet121 [5], and a custom-designed CNN architecture. The workflow is shown in Fig. 1. DenseNet is considered for this research with respect to other pre-trained DNNs, due to its superior performance in terms of feature extraction for this particular dataset. DenseNet was initially avoided by the previous researchers as they did not like the idea of experimenting with more than 50 layers as it directly enhances computing time. But, the method is feasible as the densely connected layers are incorporated, making the final output layer gain direct information from all the layers. A CNN architecture was developed to compare with the pre-trained model and to evaluate performance on several assessment parameters.

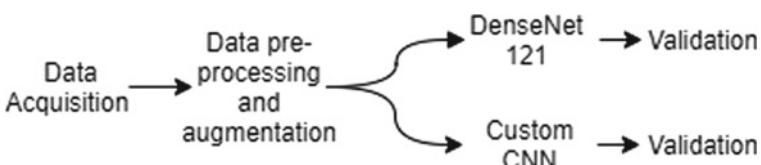


Fig. 1 Flowchart for the proposed model

2.1 Data and Preprocessing

The dataset [7] contains folders of different classes of labeled leaves. We have considered bell pepper and tomato leaves as the scope of this research, as these leaves are highly correlated and are often confused. The dataset contains 18,114 images from all the classes, described as follows.

Pepper bell—bacterial spot: 997 images, pepper bell—healthy: 1478 images, tomato—septoria leaf spot: 1771 images, tomato—two-spotted spider mite: 1676 images, tomato—yellow leaf curl virus: 3209 images, tomato—healthy: 1591 images, tomato—target spot: 1404 images, tomato—leaf mold: 952 images, tomato—bacterial spot: 2127 images, tomato—early blight: 1000 images.

The image size was reduced to 64×64 and was split into 80% training set, 12% test set and 8% validation set to validate the model. The train and test split was chosen as mentioned by monitoring best performance in the validation directory. To ensure that the overall model contains balanced classes, data augmentation was performed. The number of images for the particular classes was enhanced by setting and altering the shape, rotation range, width, and recreating images from the existing directory. The number of images after data augmentation was as follows: training set: 18,838 images, test set: 2826 images, and validation set: 1124 images.

2.2 Model Architecture

The model was created using the following architectures where initially the CNN was custom designed to predict output labels and then the network was compared to DenseNet121. Figure 2 represents the proposed convolutional network architecture. This network allows the layers to pass information about the features, layer by layer. The model architectures have been developed using Keras framework, API for tensorflow, and the system has been boosted with NVIDIA Tesla K80, a graphics processing unit (GPU). This speeds up the model training time, thereby allowing us to test and experiment more on the data. The model was trained for 10 epochs.

Convolutional Neural Network This is a type of feed-forward ANN, popularly used for image analysis like pattern recognition. It consists of convolutional layers, ReLU layers, pooling layers, and FC layers. Here, the convolutional networks receive

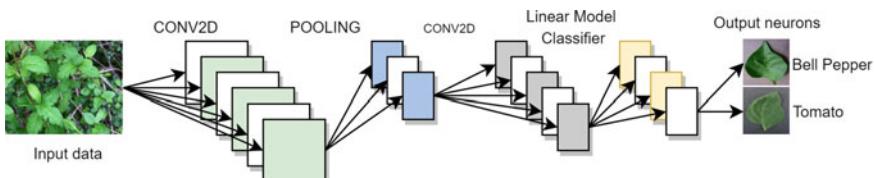


Fig. 2 Convolutional neural network (CNN) architecture

the input and transforms it to the hidden layer where the bias is added and finally output is produced. The model consists of 4 dropout layers (0.25, 0.25, 0.25, and 0.5), following which are FC layers; the final output layer corresponds to the number of classes (11 for this system). Softmax activation function was used as the final activation layer. Adam optimizer [6] was employed for the CNN models as shown in Eqs. (1–4). The total number of epochs was set to 10, and the initial learning rate of the model was 0.001 with a batch size of 32.

DenseNet The CNN architecture pass images through several padding, pooling, and convolutional layers to extract high-end features. DenseNet121 [5] allows training of the model with a high number of layers making much information accessible for the next layer and solves the vanishing gradient problem which is a major issue in any custom CNN architecture. It also decreases the total number of trainable parameters. The dropout values are limited to 0.5. The learning rate was set to 0.001, beta decay factors as 0.9 and 0.999, respectively, and decay value is set to null. All the values are set by fine-tuning the hyperparameters. The weights for the DenseNet121 architecture are taken from predefined “imagenet” which provides bounding-box annotations, thereby solving the problem with object localization. The optimizer used is Adam. The underlying equations for effective convergence and weight updates using Adam optimizer are explained in Eqs. (1–4).

Initial Weights:

$$\rho_m \leftarrow 1, \rho_v \leftarrow 1, \mathbf{m} \leftarrow \mathbf{0}, \mathbf{v} \leftarrow \mathbf{0} \quad (1)$$

Adam Optimiser Update equations:

$$\mathbf{m} \leftarrow \beta_m \mathbf{m} + (1 - \beta_m) \nabla_w J \quad (2)$$

$$\mathbf{v} \leftarrow \beta_v \mathbf{v} + (1 - \beta_v) (\nabla_w J \odot \nabla_w J) \quad (3)$$

$$\mathbf{w} \leftarrow \mathbf{w} - \alpha \left(\frac{\mathbf{m}}{\sqrt{\mathbf{v}} + \epsilon} \frac{\sqrt{1 - \rho_v}}{1 - \rho_m} \right) \quad (4)$$

Here, \odot refers to element multiplication, and in Eq. (4), the operations under the root are also handled element wise.

2.3 Result Analysis

The CNN model predicts 2727 test images correctly, while the DenseNet model correctly predicts 2788 test images out of 2826 test images. The correlation matrix for the actual and predicted test images has been provided in Fig. 3. From the figure, the classes are represented as:

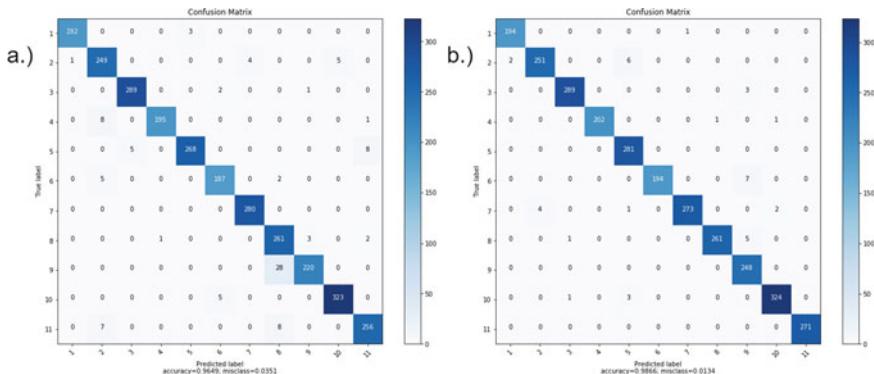


Fig. 3 Confusion matrix for **a** CNN **b** DenseNet121

Table 1 Comparison analysis of evaluation parameters

Evaluation parameter	Custom CNN	DenseNet121
Precision	0.965	0.986
Sensitivity	0.965	0.987
Specificity	0.965	0.943
F-measure	0.965	0.986
False positive rate	0.035	0.057
False negative rate	0.035	0.013
Balanced accuracy (BAC)	0.965	0.987

Class 1: Pepper bell—Bacterial spot, Class 2: Pepper bell—Healthy, Class 3: Tomato—Bacterial spot, Class 4: Tomato—Early blight, Class 5: Tomato—Late blight, Class 6: Tomato—Leaf mold, Class 7: Tomato—Septoria leaf spot, Class 8: Tomato—Two-spotted spider mite, Class 9: Tomato—Target spot, Class 10: Tomato—Yellow leaf curl virus, Class 11: Tomato—Healthy.

The class Tomato—Healthy has lowest error rate. Tomato—Target spot has the least sensitivity of all. The table for comparison analysis for evaluating necessary parameters for model evaluation is shown in Table 1.

DenseNet121, as shown in Table 1, performs better than the custom CNN architecture, designed by the authors. The pre-trained model had lesser chances of misclassification. The DenseNet model had lower specificity score than custom CNN. The feature mapping technique helps the DenseNet architecture to achieve better results in test accuracy. This also accounts for a high training accuracy in very low number of epochs.

3 Conclusion

We provide a comparison analysis using DenseNet121 and a custom CNN, which successfully detects healthy tomato and bell pepper leaves from diseased leaves. Due to the limited number of research works on the proposed dataset, our comparison with prior works is limited. These models can also predict the class of disease affecting the leaf, with a highly optimized performance. The disease classification for other types of plants with more data is aimed as a future scope to this research.

References

1. S. Arivazhagan, S.V. Ligi, Mango leaf diseases identification using convolutional neural network. *Int. J. Pure Appl. Math.* **120**(6), 11067–11079 (2018)
2. H.A. Atabay, Deep residual learning for tomato plant leaf disease identification. *J. Theoret. Appl. Inf. Technol.* **95**(24), (2017)
3. K.P. Ferentinos, Deep learning models for plant disease detection and diagnosis. *Comput. Electron. Agric.* **145**, 311–318 (2018)
4. A. Fuentes, S. Yoon, S.C. Kim, D.S. Park, A robust deep-learning-based detector for real-time tomato plant diseases and pests recognition. *Sensors* **17**(9), 2022 (2017)
5. G. Huang, Z. Liu, K.Q. Weinberger, Densely connected convolutional networks. *CoRR* (2016). <http://arxiv.org/abs/1608.06993>
6. D.P. Kingma, J. Ba, Adam: a method for stochastic optimization. *arXiv preprint arXiv:1412.6980* (2014)
7. S.P. Mohanty, D.P. Hughes, M. Salathé, Using deep learning for image-based plant disease detection. *Frontiers Plant Sci.* **7**, 1419 (2016)
8. A.K. Rangarajan, R. Purushothaman, A. Ramesh, Tomato crop disease classification using pre-trained deep learning algorithm. *Proced. Comput. Sci.* **133**, 1040–1047 (2018)
9. V.K. Shrivastava, M.K. Pradhan, S. Minz, M.P. Thakur, Rice plant disease classification using transfer learning of deep convolution neural network. *Int. Arch. Photogrammetry, Remote Sens. Spat. Inf. Sci.* (2019)
10. G. Wang, Y. Sun, J. Wang, Automatic image-based plant disease severity estimation using deep learning. *Comput. Intell. Neurosci.* **2017** (2017)
11. X. Zhang, Y. Qiao, F. Meng, C. Fan, M. Zhang, Identification of maize leaf diseases using improved deep convolutional neural networks. *IEEE Access* **6**, 30370–30377 (2018)

Energy Storage System Analysis and Optimization for Photovoltaic-Based Pico Hydro System in Remote Village



Tapas Chhual Singh, G. R. K. D. Satyaprasad, Kali Charan Rath, and P. Srinath Rajesh

Abstract This paper suggests a robust planning and management of various sources with the energy storage system (ESS) to overcome the energy demand of the remote village. Here in a proposed site, pico hydro plant is installed, and the continuous power generation deficit is avoided by the solar PV system with ESS to strengthen the distributed generation (DG) system. PV system is used to run the pico hydro turbine in pump mode to lift the downstream water into the reservoir, for which a suitable solar PV system is designed by considering the various parameters like irradiances, load demand, loss, etc. During peak load demand, ESS is used to store and balance the net energy generated by the system. A suitable ESS system is designed and optimized by using the software tool. Mathematical calculation is made based on the seasonal load demand and availability of power for the ESS. This proposed idea customized the hybrid model for internal rate of return (IRR). Simulation results obtained and compared with both the practical PV system and battery storage system and obtained the optimized result which decreases the upfront cost and operational cost.

Keywords Pico hydro · Solar PV · ESS

1 Introduction

Distributed generation system adopted by the many countries where their method of generating electricity is changed [1] and highly implementing the DG system by using new technologies [2]. This inclusion supports the use and implementation

T. C. Singh (✉) · G. R. K. D. Satyaprasad

Department of Electrical Engineering, GIET University, Gunupur, India

e-mail: tapasingh89@gmail.com

K. C. Rath

Department of Mechanical Engineering, GIET University, Gunupur, India

P. S. Rajesh

Department of Electrical Engineering, AKIT, Bhadrachalam, India

of various irregular renewable energy sources [3] that will reduce the no of remote villages that are far away from the centralized distribution system [4]. Many countries now adopting such generation system and developing new model in the assortment of upcoming power industry [5].

Solar PV system is supplementing the power during peak load time as well as used to run pump to store water in the upper reservoir (UR). PV system is integrated with battery energy storage system which is used to store enough energy during off peak time while PV system through inverter utilizes for the pump to lift the water. Here, a suitable storage system is designed based on the days of autonomy and monthly energy need for the residential. Storage system basically utilizes during off peak time, so that the cost of ESS can be reduced, and at the same time payback period of the whole system can be reduced.

2 Pico Hydro Plant Description

In the proposed model, pico hydro turbine generator set is used along the river flow at a suitable place. Cross-flow turbine is having more efficiency for the small-scale application which is coupled with the generator set to operate during peak time. It is having a suitable upper reservoir of sufficient capacity kept at a height of 15 m [6]. PV system is used to store energy and to pump the water during off peak time

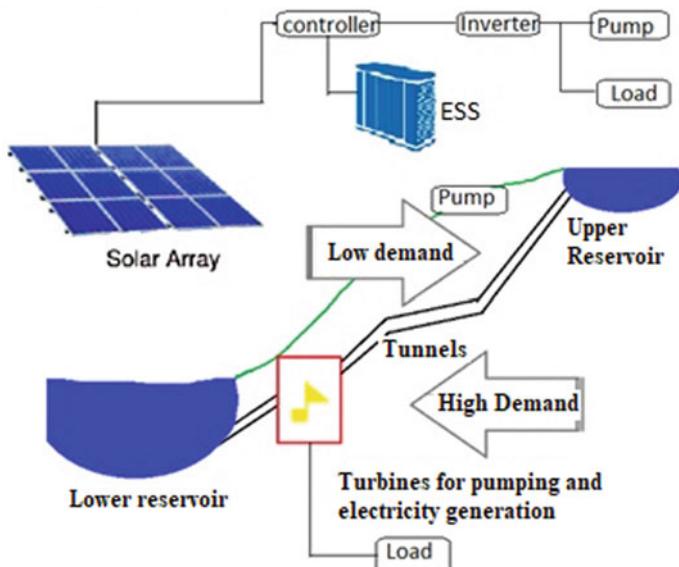


Fig. 1 Pico hydro plant layout with solar PV and battery storage system

Table 1 Pico hydro power generation with different heads and flow rate

Generator output in kW	Heads in (ft)				
	10	20	50	100	200
	Flow rate (cubic/ft/s)				
0.5	1.2	0.5	0.25	0.12	0.05
1.00	2.5	1.2	0.5	0.12	0.06
2.00	4.8	2.5	1	0.5	0.25
5.00	12	6	2.5	1.5	0.6
10.00	23.5	12	4.7	2.5	1.2

as shown in Fig. 1. The output power for variable water flow and height is as given in Table 1 [7].

3 Solar PV System for the Pico Hydro Plant

Photovoltaic system performance and predictions are calculated by the National Renewable Energy Laboratory (NREL's) PVWatts software by considering many inherent assumptions and uncertainties for the proposed site. More current generation is the main target achieved by connecting the modules in parallel and series for higher voltage for maximum power generation [8]. Photovoltaic system specification for the proposed model considered and simulated to obtain the yearly energy performance is given in Table 2.

Table 2 Annual solar radiation and energy performance

Month	Solar radiation (kWh/m ² /day)	AC energy (kWh)	Value (\$)
Jan	5.98	526	418
Feb	6.48	510	405
March	6.58	558	443
April	6.55	542	431
May	6.18	535	425
June	4.89	429	341
July	4.33	399	317
Aug	4.47	411	326
Sept	5.21	455	362
Oct	5.62	505	401
Nov	5.81	509	404
Dec	5.82	524	416
Annual	5.66	5903	\$4689

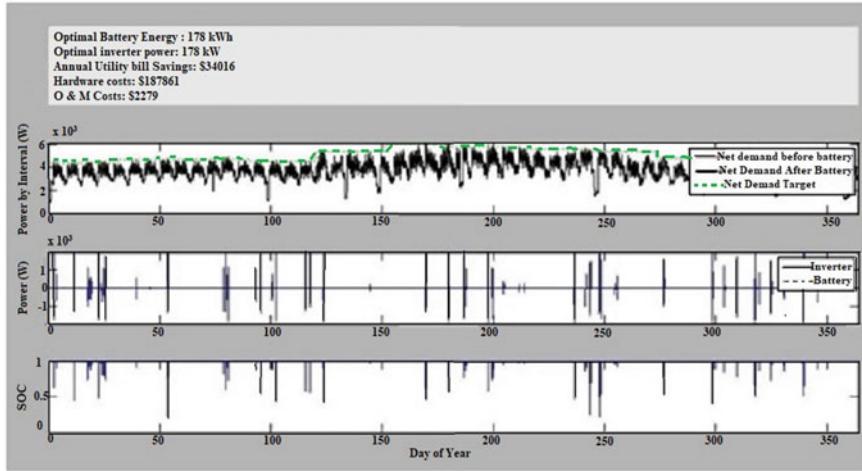


Fig. 2 Optimal battery configuration results output

4 Energy Storage System (ESS) and Optimization

ESS is designed based on the availability of PV power and load demand, to maximize the obtainable efficiency with minimum cost. For the designing of suitable battery bank, a battery size optimizer is added. Battery lifetime analysis and simulation tool (BLAST) is used for the above-said purpose. NREL's PVWatt is interlinked with BLAST to get most accurate and optimized result for the storage system design and analysis, which is as shown in Fig. 2. Internal rate of return (IRR) is for the cost analysis of the storage system. The total annual electricity bill with storage system is subtracted from the without storage system to obtained annual saving. From this annual savings, IRR for 20 year is obtained with respect to the power-to-energy ratio. It shows the calculated IRR (0.57) for all the simulated ESS and their data points, whereas the amount of available energy is obtained as 177.9 and 355.7 kWh. Figure 2 shows that battery power and its state of charge (SOC) during net demand before and after of battery internal power, and hence, the net demand target is obtained from the data.

5 Optimization of PV and ESS

Economical viability of the proposed PV and ESS is studied through the “RE opt” tool, and system performance for one year is obtained as shown in Fig. 3, for the financial analysis, system capital investment cost (\$/kW), operation and maintenance cost (\$/kW/year), etc. For the storage system, various parameters like power capacity cost (\$/kW), rectifier and inverter efficiency (%), energy capacity cost (\$/kWh), etc.,

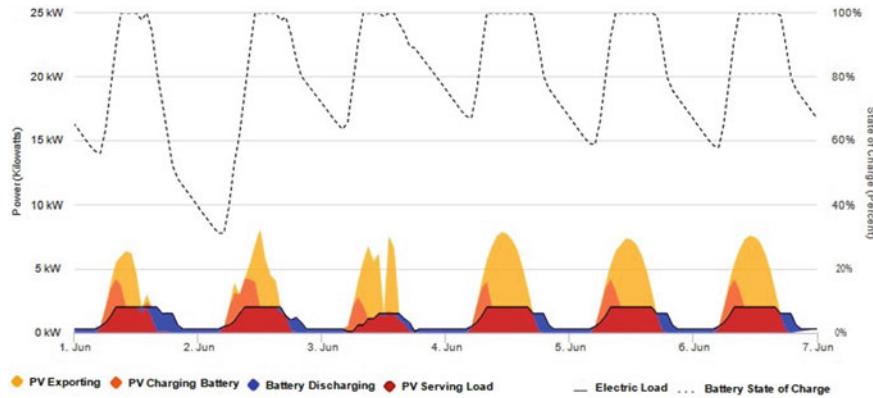


Fig. 3 Optimized dispatch strategy (for PV and ESS)

are considered for design and optimization of ESS as given in Table 3. The system minimizes the life cycle cost of energy, storage system capacity, and power optimized for the good economical routine.

In this case, we considered the above parameters for the dispatch strategy optimization. Financial model of photovoltaic structure is obtained by considering 11 kW PV system, for which annualized PV energy production is nearly 17,000 kWh, and replacement and operation cost is nearly \$18,600. Then, the total life cycle cost is \$1,062,802, and difference of \$1,008,108 is obtained after optimization. Above simulation is carried out by using NREL s “RE opt” tool.

Table 3 Simulation parameter

Inputs	Parameters	Figure
Site and utility	Annual energy consumption (kWh) Annual energy charge (\$/kWh) Annual demand charge (\$/kW/month)	10,000 \$11.03 \$20.00
PV	Initial principal cost (\$/kW) Operation and maintenance cost (\$/kW per year)	\$1600 \$16
Financial	Study period (years) Nominal discount rate (%) Effective tax rate (%) Electricity price growth rate, supposed (%) O&M charge growth rate (%)	25 8.3% 26% 2.3% 2.5%
Battery	Cost of energy capacity (\$/kWh) Cost of power capacity (\$/kW) Energy capacity replacement cost (\$/kWh) Efficiency of inverter (%) Minimum SOC (%) Initial SOC (%)	\$420 \$840 \$200 96% 20% 50%

Graph shows the dispatch scheme optimized by NREL's REopt Lite tool for specified outage time in addition to the rest of the year as shown in Fig. 3. The battery charging rate is more during the period of maximum PV generation while the PV serving load is average of the total generation. Battery discharge is noticeable when PV power is less and load is connected to the main supply.

Monthly data on solar PV power generation from solar irradiance is obtained as shown in Fig. 3, from which the maximum power per day in kW is obtained. REopt Lite's hourly dispatch chart is optimized for maximum savings and allows the user to see when PV is producing and when the battery is charged and discharged.

6 Conclusion

For the remote location where abundant water potential is available, pico hydro plant is a low cost and reliable option for the electricity generation. Problem with electricity transmission from long distance, power loss, less reliability, and less security could be avoided by constructing a decentralized energy generation system through pico hydro, solar PV, battery storage hybrid system for small-scale application. Minimum power requirement for light load, small electrical appliances could fulfill the power demand in a community. Also it reduces environmental pollutions, CO₂ emission, atmospheric temperature, etc., although initial upfront cost is more but the running cost is comparatively very less in long run. The future scope of pico hydro plant could strengthen the distributed generation system for sustainable development.

References

1. R. Ferreira, P.H. Corredor, H. Rudnick, X. Cifuentes, L. Barroso, Electrical expansion in South America: centralized or distributed generation for Brazil and Colombia. *IEEE Power Energy Mag.* **17**(2), 50–60 (2019)
2. OECD/IEA, Technology Roadmap, Hydropower, Paris (2012)
3. A. López-González, L. Ferrer-Martí, B. Domenech, Sustainable rural electrification planning in developing countries: a proposal for electrification of isolated communities of Venezuela. *Energy Policy* **129**, 327–338 (2019)
4. A. Berrada, K. Loudiyi, I. Zorkani, Valuation of energy storage in energy and regulation markets. *Energy* **115**, 1109–1118 (2016)
5. H.F. Mansour, N. Ghadimi, Optimal preventive maintenance policy for electric power distribution systems based on the fuzzy AHP methods. *Complexity* **21**(6), 70–88 (2016)
6. <https://greenbugenergy.com/get-educated-skills/measuring-power>
7. A.A. Nimje, G. Dhanjode, Pico-hydro-plant for small scale power generation in remote villages. *IOSR J. Environ. Sci. Toxicol. Food Technol. (IOSR-JESTFT)* 59–67 (2015)
8. T.C. Singh, G.R.K.D. Satya Prasad, Ch. Saibabu, K.Ch. Rath, Assessment and feasibility analysis of off-grid solar PV-Pico hydro hybrid system. *Int. J. Eng. Adv. Technol. (IJEAT)* **8**(6), 2461–2464 (2019)
9. V. Leite, T. Figueiredo, T. Pinheiro, Â. Ferreira, J. Batista (2012) Dealing with the very small: first steps of a picohydro demonstration project in an university campus. *Ren. Energy Power Quality* 1305–1310

Coplanar Wave Guide Fed Dual Band Antenna Loaded with Metamaterial Split Ring Resonator



Mekala Harinath Reddy, D. Sheela, J. Premalatha, and Abhay Sharma

Abstract Recently, antenna design based on metamaterial has become more common due to their improved bandwidth and radiation. In addition, for inherent impedance matching coplanar wave guide (CPW) antennas are used. In this papers, a high gain CPW fed dual bands antenna employed with metamaterial is demonstrated. Two circular rings are used as a radiating layers in the design, and the split rings resonator (SRR) element is loaded on the ground level to improve antenna output. Measured-results achieved resonance at 4.2 GHz (2.2–6.2 GHz) and 8.2 GHz (7–10.5 GHz) with a gain of above 2 dBi throughout the operational bands and with a high gain of 7 dBi at 10 GHz.

Keywords Coplanar wave guide · Impedance matching · Metamaterial · Split ring resonator

M. H. Reddy (✉)

University of Leicester, Leicester, UK

e-mail: harinath.m1995@gmail.com

D. Sheela

Saveetha School of Engineering, Saveetha University, Chennai, India

e-mail: dsheelatagorand@gmail.com

J. Premalatha

Tagore Engineering College, Chennai, India

e-mail: prema.joe@gmail.com

A. Sharma

Multimedia University, Cyberjaya, Malaysia

e-mail: sharma2abhay1@gmail.com

1 Introduction

Planar metamaterial is commonly used in the application of radio frequencies; for example, small resonators are one of these applications [1]. Metamaterials are either negative permeable or permittive or both and so they enhance the performance of antenna especially coplanar wave guide [2]. Metamaterial is mostly used in antennas to enhance bandwidth and radiation properties [3]. Also, various meta-material structures are being deployed to diminish the mutual coupling in MIMO antennas. Antennas might be impractical to use when they have narrow bandwidth or negative gain and in order to overcome these limitations metamaterial finds better choice [4]. Metamaterial antenna is one of the solutions for rising demand in the wireless communications, and it has brought a revolutionary progress in various aspects [5].

With development in today wireless communications, antennas fed by CPW are gaining attention since they can be built in MMIC circuits [5]. CPW antennas have an advantage of less dispersion and loss of radiation [6]. CPW feed lines can be used to connect active and passive components both serially and parallelly on one side of the substrate so that via hole connections can be eliminated [7]. For wide band applications, CPW fed antennas are more adequate. To enhance bandwidth, a broadband CPW fed slots antenna is presented in [8]. In [9], a dual band CPW slotted antenna with rectangular patch is discussed that is validated to perform at 2.2 and 5.8 GHz with a compact size. A T-shaped CPW antenna with a triangular slots on ground level is validated experimentally in [10] to work at 3.1–11.1 GHz. In [11], a circular disc antenna coupled with CPW and SRR is presented to cover WLAN and WIMAX. A compacts CPW antenna with two parasitic patches is demonstrated in [12] to broaden the bandwidth (2.3–7.4 GHz).

In this paper, a circular split ring antenna is demonstrated coupled with SRR and CPW. Parametric study is performed for various ground iterations. The theme of the paper is to investigate the performance of antenna when coupled with both SRR and CPW.

2 Proposed Antenna Geometry

The proposed design is produced on a standards 1.6 mm thick (h) FR44 substrate ($\epsilon_r = 4.4$). The structures of ground and radiating layers are illustrated below.

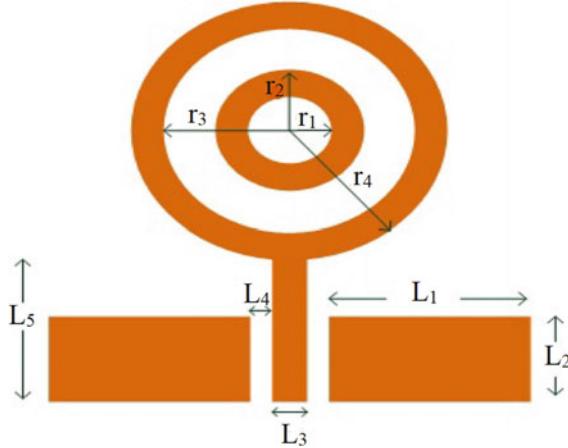


Fig. 1 Radiating layer configuration

2.1 Top Layer Configuration

Using Eqs. 1 and 2, the diameter of circular patches is determined [13]

$$R = \frac{F}{\left\{1 + \frac{2h}{\pi \epsilon_r F} \left[\ln \frac{\pi F}{2h} + 1.7726 \right]\right\}^{\frac{1}{2}}} \quad (1)$$

$$F = \frac{8.791 \times 10^9}{f_r \sqrt{\epsilon_r}} \quad (2)$$

The radiating layer consist of two circular split rings fed by CPW with radius r_1 and r_2 for inner ring and r_3 and r_4 for outer ring. Figure 3 reveals structure of antennas top layer (Fig. 1).

For the outer and inner rings, the average diameter is kept at $\lambda/2$ and $\lambda/8$, respectively. The inner and outer ring dimensions in the design are $r_1 = 6$, $r_2 = 9$, $r_3 = 10$, $r_4 = 13$ (units in mm). For the CPW feed, $L_1 = 20$, $L_2 = 13$, $L_3 = 3.5$, $L_4 = 1$ and $L_5 = 15$ (units in mm) are the dimensions defined. Over all antenna dimension is 38×38 mm².

2.2 Ground Layer Configuration

Ground layer has been analysed using three different configurations. Iteration-I and Iteration-II use full ground layer and grounds mid-open as shown in Fig. 2. Iteration-II dimensions are given by $G_1 = G_2 = 38$, $G_3 = G_4 = G_5 = G_6 = 2.5$ and $G_7 =$

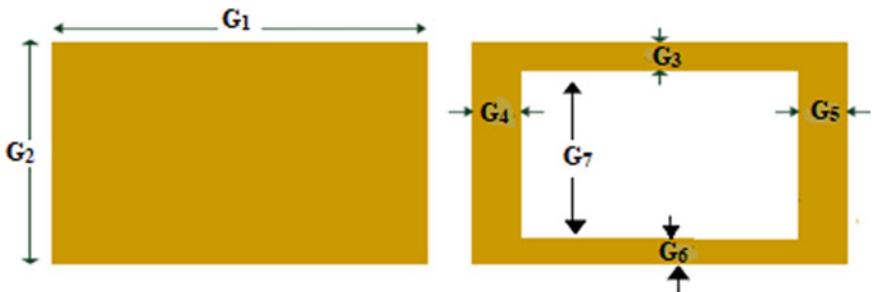


Fig. 2 Full ground (Iteration-I) and ground mid-open configuration (Iteration-II)

30 mm. As shown in Fig. 3, SRR is the 3rd iterative structure that is deployed on the ground layer and the superimposed view of whole structure is shown in Fig. 4. When compared to Iterations I and II, Iteration-III configuration provided improved

Fig. 3 Metamaterial-based ground configuration (Iteration 3)

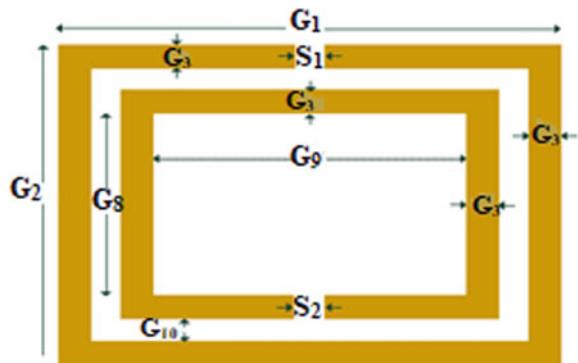
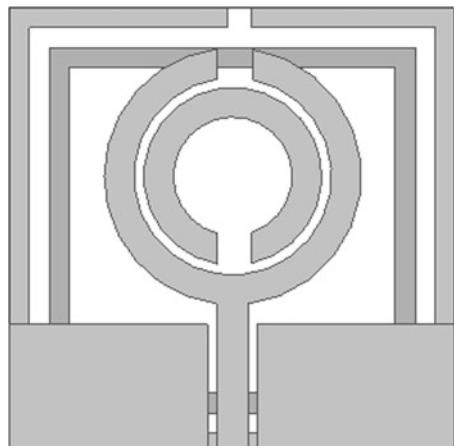


Fig. 4 Antenna coupled with SRR and CPW



bandwidth, high gain and low loss. Iteration-III ground structure dimensions are given by $S_1 = S_2 = 2.5$, $G_8 = 32$, $G_9 = 32$ and $G_{10} = 2$ mm, respectively. Magnetic resonant frequency of rectangular SRR can be calculated using Eqs. (3) and (4) [13].

$$f_{\text{SRR}_m} = \frac{C}{2\pi^2} \sqrt{\frac{3w}{\varepsilon_r(r - 2t - w)^3}} \quad (3)$$

$$\lambda_{\text{SRR}_m} = 2\pi^2 \sqrt{\frac{(r - 2t - w)^3}{3w}} \quad (4)$$

where f_{SRR_m} is magnetics resonant frequency, λ_{SRR_m} is the guided wavelength, ε_r is the substrate's dielectric constants, 'w' is the gap between inner and outer rings of SRR and 't' metal line width.

3 Results and Discussions

The design simulation is performed in HFSS, and the results are obtained for the various ground iterative configurations. Return loss plot in Fig. 5 shows that antenna with CPW and SRR exhibited a return loss of -43.31 and -57.75 dB at 4.2 and 8 GHz resonance frequencies, respectively. Full Ground configuration (Iteration-I) yielded peak resonance frequency at 8.45 GHz (7.9–9.4 GHz) and Iteration-II configuration exhibited resonance frequency at 8.5 GHz (7.7–10 GHz) as shown in Fig. 6. There was a drastic enhancement observed when SRR is loaded in to the antenna.

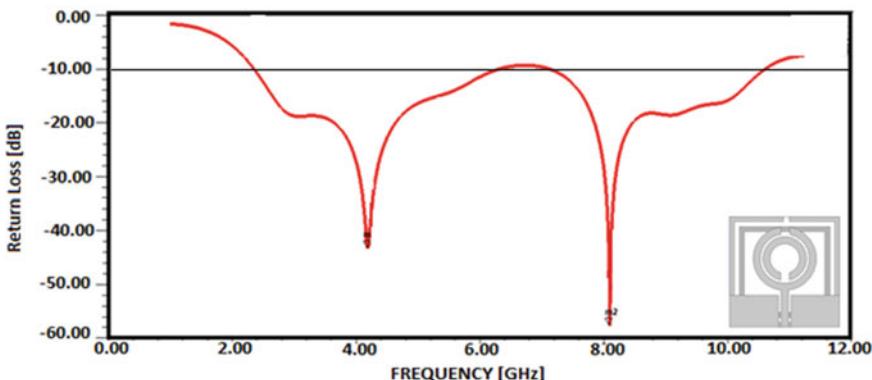


Fig. 5 Frequency versus return loss plot of proposed antennas

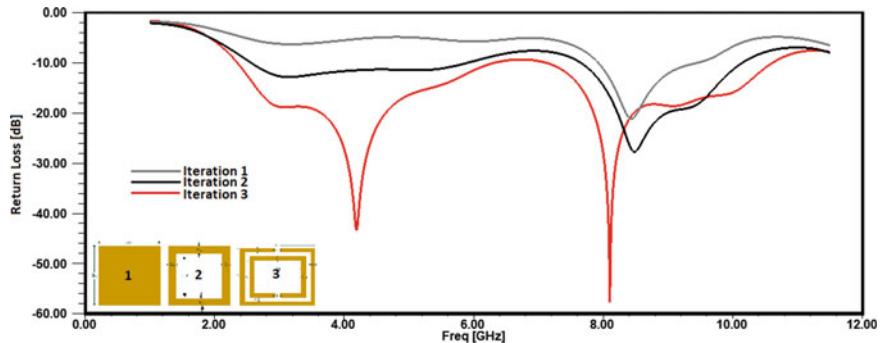


Fig. 6 Frequency versus return loss plot of various ground iterative structures with CPW

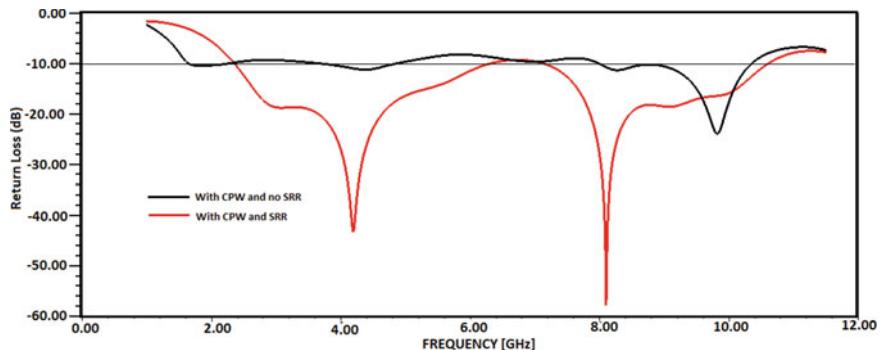


Fig. 7 Frequency versus return loss plot of antenna with CPW and SRR and with CPW and no SRR

Figure 7 illustrates the return loss for two different configurations of antenna, one antenna with CPW and no SRR and the second one with CPW and SRR. When CPW is used with no SRR, the antenna resonated at 4.2 and 10 GHz with a return loss of -10.5 and 21 dB. When both CPW and SRR are superimposed in to one antenna, the resonance was obtained at 4.2 GHz (2.2–6.2 GHz) and 8.2 GHz (7–10.5 GHz). This shows the benefit of having antennas with SRR.

4 Measured Result

The planned antennas are fabricated on FR4 material and tested using vector network analyser. Fabricated prototype is shown in Fig. 9.

Measured result in Fig. 8 shows that simulated and experimented results are in wise agreement operating at dual bands 4.2 (2.2–6.2 GHz) and 8.2 GHz (7–10.5 GHz) with improved bandwidth at operating frequencies (Fig. 9).

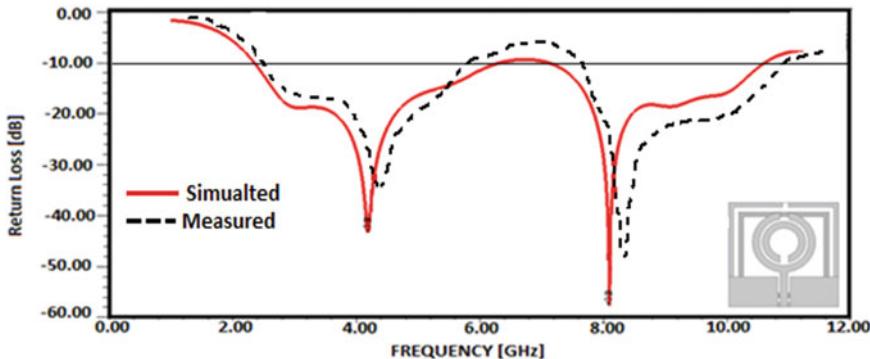


Fig. 8 Measured result of antenna

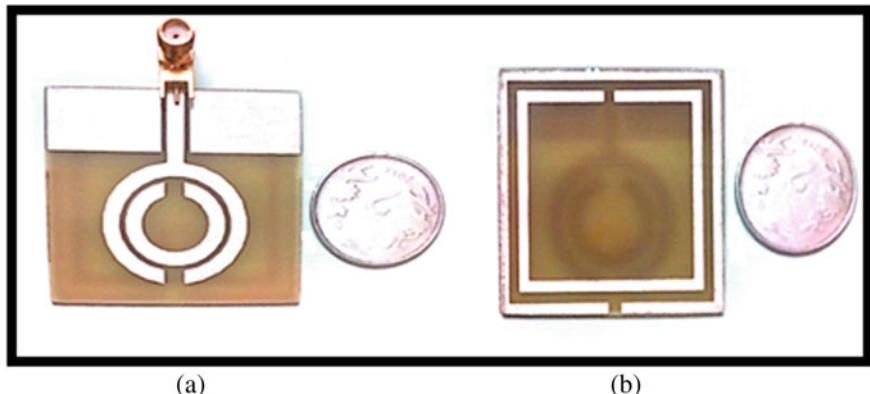


Fig. 9 Top layer (a) and bottom layer (b) of the fabricated antenna

Figure 10 illustrates the frequency versus gain plot of the antenna. Positive gain is observed throughout the operational bands with a maximum gains of 7 dBi at 10 GHz. It was investigated that antenna with iteration-III ground configuration obtained high gain when compared to iterations I and II. The fact is that the surface waves get suppressed due to the negative permeability characteristics of metamaterial, and hence, the gain is improved. At 4.2 and 8.2 GHz, the gain obtained is 5.5 and 6.5 dBi.

Figure 11 displays the surface current distributions of the antenna at 4.2 and 8.2 GHz. At 4.2 GHz current is equally distributed on the radiation layer and SRR. Whereas, at 8.2 GHz high current density is observed in SRR structure. Also, the distribution is stronger in higher band resonant frequencies compared to the lower frequencies. This current distribution validates that the gain is higher at higher frequencies. The final iterative structure's, simulated and measured (H- and E-Plane) patterns of radiations at 4.2 and 8 GHz are illustrated in Fig. 12 which shows eight-

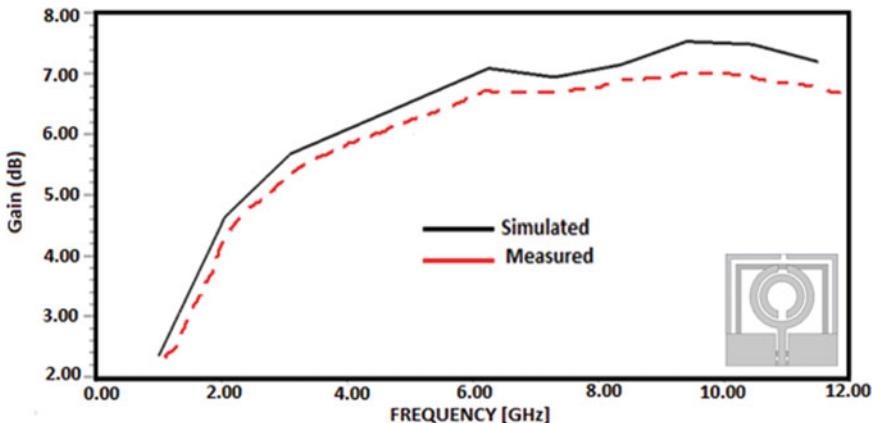


Fig. 10 Gain of antenna with CPW and SRR

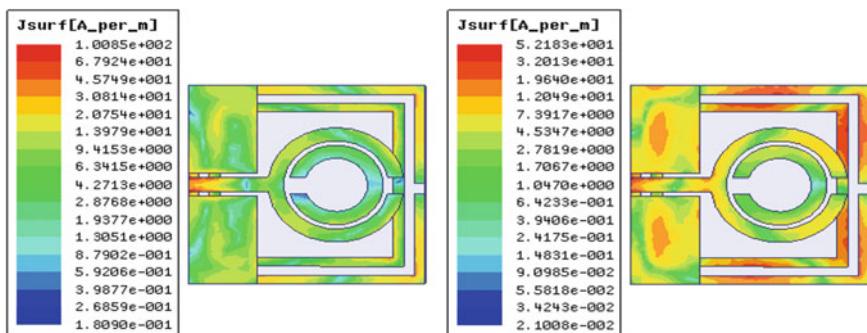


Fig. 11 Surface currents distributions at 4.2 and 8.2 GHz

shaped patterns in E-plane and omnidirectional patterns in H plane. Slight variations were observed in the radiation patterns for various ground iterations (Table 1).

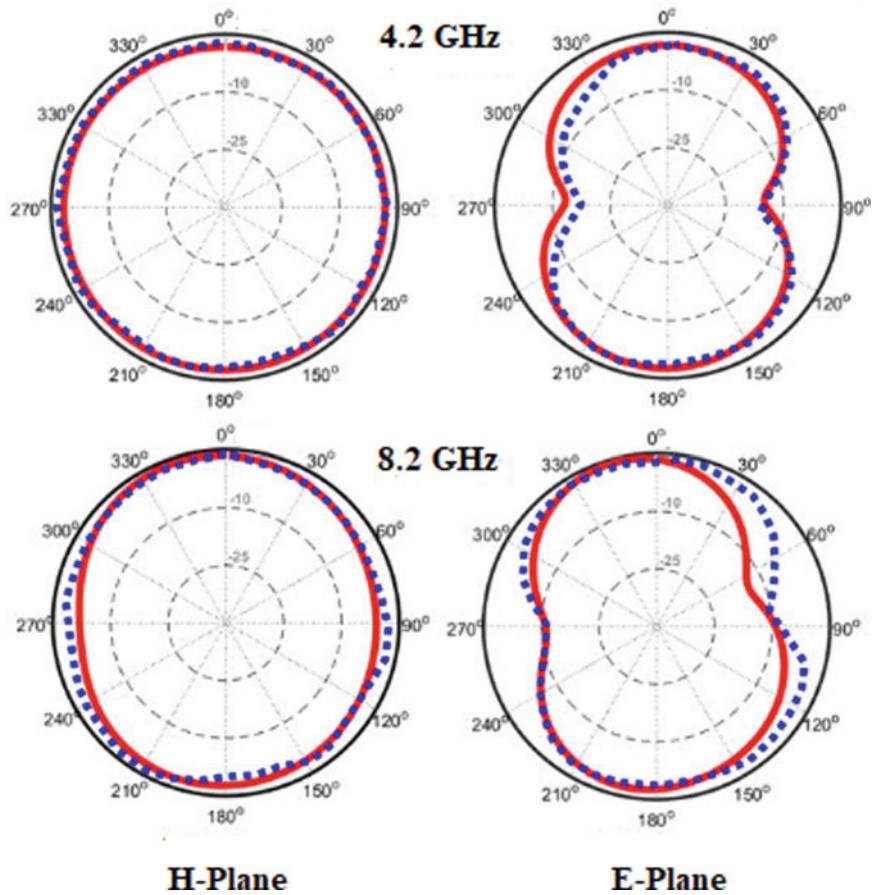


Fig. 12 Simulated (solid line) and measured (dotted line) radiation patterns

Table 1 Proposed work comparison with existing antennas

References	Size (mm ²)	f_{11} (GHz)	f_{22} (GHz)	Bandwidths (MHz)	Gains (dB)		
[14]	70×70	1.7	2.1	200	400	4.5	3.8
[15]	50×20	2.40	5.20	61	28	2.83	2.70
[16]	70×70	1.5	3.5	1600	200	3.5	4
[17]	70×70	1.6	2.2	200	400	3.8	4.1
[18]	63×75	2.5	5.6	1000	500	2.5	6
Proposed	38×38	4.2	8.2	4000	3500	5.5	6.5

5 Conclusion

A compact CPW fed antenna loaded with split ring resonator is proposed for 2.4 ISM band, C-band (3.6–4.2 GHz) and x-band applications (7–10 GHz) applications with a maximum gain of 7 dBi at 10 GHz. Parametric study has been done for various ground iterations in terms of reflection coefficient. Proposed design is fabricated and experimented, where both the results are in sink and show favourable characteristics for practical applications. Experimented results shows resonating frequencies at 4.2 GHz (2.2–6.2 GHz) and 8.2 GHz (7–10.5 GHz) with bidirectional pattern of radiation in H-planes and omnidirectional pattern in E-planes at centre frequencies.

References

1. M. Anand, Applications of metamaterial in antenna engineering. *Int. J. Tech. Res. Appl.* **2**(1), 49–52 (2014)
2. D. Deng, H. Zheng, A design of an ultra-wideband coplanar waveguide antenna. *Int. J. Eng. Res. Technol.* **4**(07), 1160–2116 (2015)
3. K. Sato, S.H. Yonak, T. Nomura, S.I. Matsuzawa, H. Iizuka, Metamaterials for automotive applications, in *IEEE Antennas and Propagation Society International Symposium*, Honolulu, HI (2007), pp. 1144–1147
4. Y.H. Xie, C. Zhu, L. Li, C.H. Liang, A novel dual-band metamaterial antenna based on complementary split ring resonators. *Microw. Opt. Technol. Lett.* **54**, 1007–1009 (2012)
5. D.R. Smith, S. Schultz, Determination of effective permittivity and permeability of metamaterials from reflection and transmission coefficients. *Phys. Rev. B* **65**, 195104 (2002)
6. G.V. Eleftheriades, A. Grbic, M. Antoniades, Negative-refractive-index transmission-line metamaterials and enabling electromagnetic applications, in *Proceedings of IEEE Antennas and Propagation Society International Symposium. Digest* (June 2004), pp. 1399–1402
7. J. Zhu, M.A. Antoniades, G.V. Eleftheriades, A compact tri-band monopole antenna with single-cell metamaterial loading. *IEEE Antennas Wirel. Propag. Lett.* **58** (2010)
8. J.-W. Niu, S.-S. Zhong, A CPW-fed broadband slot antenna with linear taper. *Microw. Opt. Technol. Lett.* **41**(3), 218–221 (2004)
9. W.-C. Liu, F.-M. Yeh, Compact dual-and wide-band CPW-fed slot antenna for wireless applications. *Microw. Opt. Technol. Lett.* **50**(3), 574–575 (2008)
10. V.A. Shameena, S. Mridula, A. Pradeep, S. Jacob, A.O. Lindo, P. Mohanan, A compact CPW fed slot antenna for ultra wide band applications. *AEU-Int. J. Electron. Commun.* **66**(3), 189–194 (2012)
11. G. Shehata, M. Mohanna, M.L. Rabeh, Tri-band small monopole antenna based on SRR units. *NRIAG J. Astron. Geophys.* **4**(2), 185–191 (2015)
12. S.N. Ather, P.K. Singhal, Broadband CPW-fed Rectangular Antenna with parasitic patches, in *2014 International Conference on Computational Intelligence and Communication Networks* (IEEE, 2014), pp. 26–29
13. M.A. Antoniades, G.V. Eleftheriades, A folded-monopole model for electrically small NRI-TL metamaterial antennas. *IEEE Antennas Wirel. Propag. Lett.* **7**, 425–428 (2008)
14. C.H. Chen, E.K.N. Yung, Dual-band circularly-polarized CPW-fed slot antenna with a small frequency ratio and wide bandwidths. *IEEE Trans. Antennas Propag.* **59**(4), 1379–1384 (2011)
15. M.E. Atrash, K. Bassem, M.A. Abdalla, A compact dual-band flexible CPW-fed antenna for wearable applications, in *2017 IEEE International Symposium on Antennas and Propagation & USNC/URSI National Radio Science Meeting*, San Diego, CA (2017), pp. 2463–2464

16. Y.Y. Chen, Y.C. Jiao, G. Zhao, F. Zhang, Z.L. Liao, Y. Tian, Dual-band dual-sense circularly polarized slot antenna with a C-shaped grounded strip. *IEEE Antennas Wirel. Propag. Lett.* **10**, 915–918 (2011)
17. C. Chen, E.K.N. Yung, Dual-band dual-sense circularly-polarized CPW-fed slot antenna with two spiral slots loaded. *IEEE Trans. Antennas Propag.* **57**(6), 1829–1833 (2009)
18. R.K. Saini, S. Dwari, CPW fed dual-band dual-sense circularly polarized monopole antenna. *IEEE Antennas Wirel. Propag. Lett.* **16**, 2497–2500 (2017)

Design of Low Power Multipliers Using Approximate Compressors



Vishal Gundavarapu, M. Balaji, and P. Sasipriya

Abstract Approximate computing is one of the most promising innovative methodologies to achieve low power consumption and high performance. In this paper, two designs of approximate 4:2 compressors have been proposed to reduce the latency and minimize the power consumption. DADDA multiplier has been designed using the proposed compressors, and it has been compared with the existing approximate multipliers. Hybrid multipliers consisting of two types of proposed 4:2 compressors have also been designed and included for comparison. Furthermore, an exact DADDA multiplier is also implemented to enable the calculation of power savings for the proposed multipliers. The multipliers are then employed in discrete cosine transform (DCT) for image compression, and the error characteristics such as PSNR and NED are analyzed. Simulation results prove that the proposed approximate multipliers have reduced power consumption when compared to the exact multiplier with the minimum error which makes them suitable for image processing applications. All the simulations are carried out using Cadence Virtuoso® design tools.

Keywords 4:2 compressor · Approximate compressor · Approximate multiplier · DADDA multiplier · Low power

1 Introduction

Multiplication is a fundamental operation widely used in signal processing and image processing applications. However, these applications do not benefit from a high level of accuracy. With the introduction of technology such as IoT, there is a growing need for image processors and signal processors to be more power efficient. We

V. Gundavarapu · M. Balaji · P. Sasipriya (✉)
SELECT, VIT, Chennai, India
e-mail: sasipriyap@yahoo.com

V. Gundavarapu
e-mail: g.vishal1998@gmail.com

can achieve this by compromising accuracy. Due to the approximate nature of these circuits, the hardware complexity can be reduced and thus offer reduced delay.

A multiplier consists of two stages, namely, partial product generation stage, and accumulation of the partial products stage. There are various designs of approximate multipliers that have been introduced in literature. Momeni et al. have proposed two 4:2 compressor circuits in [1]. The authors have designed four multipliers viz., multiplier 1 and multiplier 2 employ the proposed two approximate compressors, and multiplier 3, and multiplier 4 employ hybrid architecture by combining both the exact compressors and approximate compressors to reduce the error rate. Venkatachalam and Ko [2] introduces a novel method of generating partial products using probability statistics. An inaccurate counter is proposed in [3] with the aim of reducing the delay. This design is then used in a 4×4 Wallace tree multiplier. An error detection and correction circuit has also been proposed in [3] to ensure improved accuracy. Ha and Lee [4] uses truncation and approximate compressors for accumulation of bits of lower significance and the main focus is on using approximate 4:2 compressors to add these partial products. Three designs of approximate compressors have been proposed by having a low error rate as a constraining factor in [5]. The logic design of these approximate compressors has been realized by introducing error in the truth table of an exact compressor in such a way that the proposed approximate compressors have reduced hardware complexity which in turn reduces the number of gates. He et al. [6] proposes a Booth multiplier that uses a probabilistic prediction formula to generate some of the partial products. Venkatachalam et al. [7] and Liu et al. [8] propose approximate Booth multipliers where error is introduced into the Booth encoders.

In this paper, two approximate compressors have been proposed and five multipliers using the proposed compressors have been designed. To prove the accuracy of the multipliers, an image compression algorithm that uses the direct cosine transform (DCT) has been implemented in this paper. The remainder of the paper has been organized as follows. Section 2 presents the design of two novel approximate 4:2 compressors. Section 3 presents the implementation of the multipliers using the proposed compressors. Section 4 presents the simulation results and expands on the results from the image processing applications. Finally, the conclusion is presented in Sect. 5.

2 Approximate 4:2 Compressors

Traditionally, in order to add four bits along with the input carry, a combination of two full adders were used. However, 4:2 compressors have been designed in recent years to reduce the stage latency [9]. The design of novel approximate compressors proposed in this paper is discussed in this section. The ACCI2 compressor used in the design of a DADDA multiplier had the highest PSNR in comparison [10] and therefore has been used as a base for the novel designs proposed in this paper.

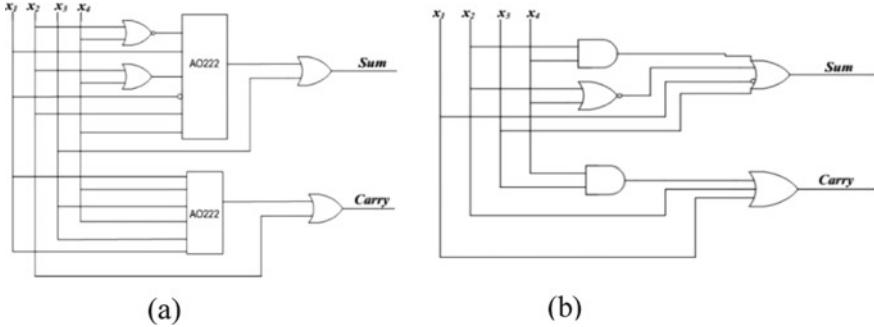


Fig. 1 Gate level implementation of **a** design 1 and **b** design 2

2.1 Design 1

The approximate compressors are proposed by considering the design constraint of reduced complexity in the logic design. The gate level realization of the proposed compressor is shown in Fig. 1a. A total of three error conditions were introduced for the sum output, and one error was introduced into the carry of this design. The Boolean expressions for Sum and Carry of the proposed 4:2 compressor is given by:

$$\text{Sum} = (X_1 X_2' X_4') + X_1' (X_2 + X_4) + X_2 X_4 + X_2 \quad (1)$$

$$\text{Carry} = X_2 + X_2 X_4 + X_1 X_4 + X_1 X_2 \quad (2)$$

The number of stages to generate carry signal in this compressor is in the form of one AO222 stage and an OR gate. In this design, the third output C_{out} is ignored and therefore introduces an error when all the inputs are equal to logic ‘1.’ Six out of the sixteen outputs are incorrect. Therefore, the error rate, i.e., percentage of wrong outputs to total number of outputs is 37.5% which is tolerable for image processing applications.

2.2 Design 2

The compressor design 1 is further modified by introducing ‘don’t care’ conditions which reduces the complexity of the circuit compared to Design 1. However, this increases the error generated by the circuit. The number of stages involved in generating the carry signal here consists of an AND gate and an OR gate. Thus, the number of delay stages is 2Δ which is lesser than the critical path delay in an exact 4:2 compressor. However, the number of wrong outputs is eight and the error rate becomes 50%. Figure 1b shows the gate level implementation of the proposed

compressor. The Sum and Carry expressions are given by:

$$\text{Sum} = X1' + X2X4 + X2 + X2'X4 \quad (3)$$

$$\text{Carry} = X1 + X2 + X3X4 \quad (4)$$

3 Design of Multipliers

In general, the architecture of multiplier consists of three stages. The partial product generation, accumulation of partial products (array/carry save or tree structure) and finally an adder (RCA/CLA or tree adder) to add the outputs generated by the previous stage.

In this paper, a DADDA tree multiplier is used. Figure 2 shows the architecture of the DADDA multiplier [9] realized using 4:2 compressors. An exact compressor has a propagation delay of 4Δ . Hence, the critical path delay of the exact multiplier to produce the final outputs is found to be $4\Delta + 4\Delta + 4\Delta = 12\Delta$. However, the proposed multiplier which uses the novel approximate compressors, which has only two outputs, has a critical path delay of $2\Delta + 2\Delta + 2\Delta = 6\Delta$. A Brent-Kung tree adder introduced in [11] is used to perform the final stage addition to improve the performance.

A total of five multipliers have been proposed in this paper. In the first multiplier, design 1 compressors are used to reduce the partial products and propagate the outputs to the final addition stage. Similarly, in the second multiplier, design 2 compressors have been used. The hybrid multipliers (multipliers 3, 4 and 5) are designed by

Fig. 2 Architecture of DADDA tree multiplier

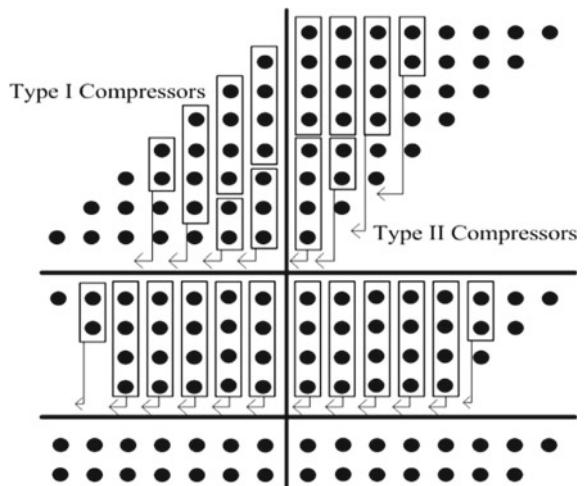


Table 1 Error parameters for the multipliers

Multiplier	ER	NED
Multiplier 1	0.9668	0.0525
Multiplier 2	0.9999	0.1243
Multiplier 3	0.9978	0.0554
Multiplier 4	0.8650	0.0082
Multiplier 5	0.9987	0.0113

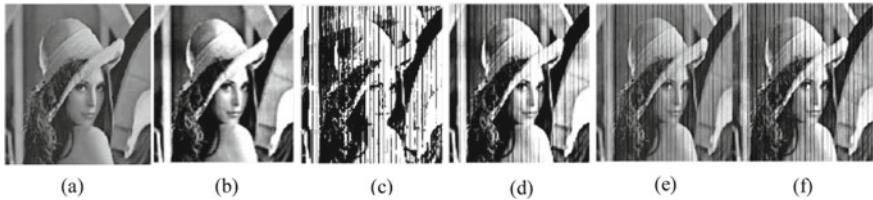
combining both the design 1 and design 2 compressors. The hybrid multipliers are designed as follows. One half side of the multiplier uses the compressors termed ‘Type I compressors,’ and the other side uses a different design of compressors termed ‘Type II compressors’ as shown in Fig. 2. The multiplier 3 is composed of design 1 and design 2 compressors proposed in this paper. Since design 1 has a lesser error percentage, it is used in the accumulation of bits with higher significance, i.e., it is used in the part of the reduction circuitry represented as ‘Type I compressors’ and design 2 is used as ‘Type II compressors.’ The multiplier 4 is composed of ACCI2 as ‘Type I compressor’ and design 1 compressors as ‘Type II compressor.’ Finally, multiplier 6 has ACCI2 as ‘Type I compressor’ and design 2 as ‘Type II compressor.’ The accuracy of the proposed approximate multipliers are measured using the error parameters error rate (ER) and normalized error distance (NED) [12]. NED is the normalization of the mean of the error distance which is a stable metric independent of the size. Error distance is the distance between a wrong output generated and the correct output for a given input. The error rates for the proposed multipliers are shown in Table 1.

4 Simulation Results

The 8-bit DADDA multiplier circuits have been realized using Cadence® Virtuoso. All the multiplier components used in the simulation are realized using the 45 nm CMOS process technology. The multipliers were operated at a switching frequency of 100 MHz, and it is observed that the multipliers implemented using approximate circuitry consume lesser power when compared to the exact counterpart. The power consumed and power savings when compared to the exact multiplier is shown in Table 2. It can also be noted that the multipliers using the approximate compressors have used lesser number of transistors when compared to the exact multiplier. From Table 2, it is observed that multiplier 2 has the lowest power consumption and also has the lowest transistor count amongst all the multiplier designs. Due to the higher complexity of ACCI2 and design 1, the power consumed by multipliers 1 and 4 are on the higher end. However, the power consumed by them is still comparatively lesser than the power consumed by the exact multiplier.

Table 2 Power consumption comparison and transistor count

Multiplier	Power consumed (μW)	Power savings (%)	Number of transistors
Exact	14.29		2296
Multiplier 1	9.012	36.93	2190
Multiplier 2	5.933	55.82	1646
Multiplier 3	7.730	47.10	1902
Multiplier 4	9.073	36.50	2126
Multiplier 5	7.748	48.47	1838

**Fig. 3** Image compression results for **a** exact multiplier **b** multiplier 1, PSNR = 14.882 **c** multiplier 2, PSNR = 8.3581 **d** multiplier 3, PSNR = 12.881 **e** multiplier 4, PSNR = 20.1896 and **f** multiplier 5, PSNR = 15.3463

To study the impact of approximation in real-time applications, the proposed multipliers were used to perform the matrix multiplications required to implement the DCT algorithm for image compression. For evaluation purposes, the proposed multiplier designs were realized using MATLAB. Performing DCT results in floating-point numbers. In order to use these numbers with the 8-bit multipliers proposed here, these floating-point numbers were rounded to their closest whole number. This would certainly add another layer of error into this process, and therefore, the errors obtained here would be higher than expected.

Peak signal to noise ratio (PSNR) is used to evaluate the quality of the resulting images shown in Fig. 3. The results show that the PSNR for the multiplier 2 is the lowest. The hybrid multipliers have effectively reduced the error rates as can be seen from the increase in PSNR values in the hybrid multipliers. The hybrid multipliers have been successful in keeping the power consumption low while also having an acceptable amount of clarity in the images as can be seen from Fig. 3.

5 Conclusion

Error tolerant approximate computing has been employed in this paper to achieve low power operation. This paper has presented two novel designs for approximate 4:2 compressors with reduced latency compared to the exact compressor. These compressors were used in the realization of five DADDA multipliers. Multiplier 2

has the lowest power consumption amongst all the proposed multipliers. It is found to be 55.82% lesser than an exact multiplier when operated at 100 MHz. However, multiplier 2 has the higher error rate compared to the other proposed multipliers. The hybrid multipliers are also designed in this paper and the results show that the schemes were efficient in terms of power consumption while also being able to maintain an acceptable error. Image processing application has been implemented using the proposed multipliers and the PSNR values are measured for all the proposed multipliers. The results prove that the proposed multipliers are suitable for real-life applications.

References

1. A. Momeni, J. Han, P. Montuschi, F. Lombardi, Design and analysis of approximate compressors for multiplication. *IEEE Trans. Comput.* **64**(4) (2015)
2. S. Venkatachalam, S. Ko, Design of power and area efficient approximate multipliers. *IEEE Trans. VLSI Syst.* **25**(5), 1782–1786 (2017)
3. C. Lin, I. Lin, High accuracy approximate multiplier with error correction, in *2013 IEEE 31st International Conference on Computer Design (ICCD)*, Asheville, NC (2013), pp. 33–38
4. M. Ha, S. Lee, Multipliers with approximate 4–2 compressors and error recovery modules. *IEEE Embed. Syst. Lett.* **10**(1), 6–9 (2018)
5. Z. Yang, J. Han, F. Lombardi, Approximate compressors for error-resilient multiplier design, in *2015 IEEE International Symposium on Defect and Fault Tolerance in VLSI and Nanotechnology Systems (DFTS)*, Amherst, MA (2015), pp. 183–186
6. Y. He, X. Yi, B. Ma, Z. Zhang, B. Zhang, A probabilistic prediction based fixed-width booth multiplier, in *2018 IEEE Asia Pacific Conference on Circuits and Systems* (2018)
7. S. Venkatachalam, E. Adams, H.J. Lee, S. Ko, Design and analysis of area and power efficient approximate booth multipliers. *IEEE Trans. Comput.* **68**(11), 1697–1703 (2019)
8. W. Liu, L. Qian, C. Wang, H. Jiang, J. Han, F. Lombardi, Design of approximate Radix-4 booth multipliers for error-tolerant computing. *IEEE Trans. Comput.* **66**(8), 1435–1441 (2017)
9. C.-H. Chang, Gu. Jiangmin, M. Zhang, Ultra low-voltage low-power CMOS 4–2 and 5–2 compressors for fast arithmetic circuits. *IEEE Trans. Circ. Syst. I Regul. Pap.* **51**(10), 1985–1997 (2004)
10. M. Balaji, V. Gundavarapu, P. Sasipriya, K. Bhaaskaran, Low power multiplier using approximate compressor for error tolerant applications. *Int. J. Eng. Adv. Technol.* **9**(3), 319–324 (2019)
11. P.B. Reddy, V.K. Bhaaskaran, Design of adiabatic tree adder structures for low power, in *International Conference on Embedded Systems (ICES 2010)* organized by CIT, Coimbatore and Oklahoma State University (2010), pp. 14–16
12. J. Liang, J. Han, F. Lombardi, New metrics for the reliability of approximate and probabilistic adders. *IEEE Trans. Comput.* **62**(9) (2013)

Block-Based Local Binary Pattern for Recognition of Handwritten Odia Numerals



Suchismita Behera and Niva Das

Abstract In this paper, we exploit the texture feature of local binary pattern (LBP) for handwritten Odia numeral recognition. There are several challenges in the handwritten recognition due to the different writing style of the individuals. The histogram of the block-based LBP is computed considering the block sizes of (4×4) , (8×8) and the whole image and these features are used to classify the numerals. For classification, SVM and single decision tree classifiers have been employed. In this experiment, we evaluated the performance of the block-based LBP on the ISI, Kolkata Odia numeral digital database. The performance of (4×4) block-based LBP shows high level of accuracy (97%) for both SVM and single tree classifiers. Single decision tree classifier shows promising results in terms of computational time and hence can be used for real-time applications.

Keywords Numeral recognition · LBP · SVM · Single decision tree classifier

1 Introduction

An optical character recognition system (OCR) recognises characters from scanned pages of the document images and facilitates text processing. Handwritten character recognition has been a challenging task in the field of pattern recognition due to the large variations in the writing style of the same character by the different individuals with difference in age, gender and education. However, offline handwriting recognition has gained prime importance in the creation of digital libraries, archiving historical monuments, bank cheque processing and postal pin code verification to mention a few examples. The progress in this field is due to the advancement in

S. Behera · N. Das (✉)

Department of Electronics and Communication Engineering, Siksha ‘O’ Anusandhan (Deemed to be University), Bhubaneswar, Odisha, India

e-mail: nivadas@soa.ac.in

S. Behera

e-mail: suchismitabehlera@soa.ac.in

automatic processing of large volumes of documents; one of the reason being global digitization of several handwritten scriptures and books.

Achieving high recognition accuracy on the handwritten character dataset is a challenging problem for which many solutions have been proposed. High accuracies have been obtained for the standard MNIST dataset, but there are many datasets of different languages where the work have been hindered due to less number of examples available in the database. Many research works have been done for various types of Indian scripts such as Bangla, Devanagari, Telugu, Tamil and Gujrati. Odia handwritten script recognition is a major challenging task for researchers due to its complex shape of characters and large variations in the writing styles of the individuals and also due to the insufficiency publicly available database.

In this paper, a variant of LBP known as block-based LBP have been considered for feature extraction and along with SVM and single decision tree classifier is applied for the recognition of handwritten Odia numeric characters.

The paper has been organised as follows: the literature review of the work done on handwritten numerals is provided in Sects. 2, 3 deals with a brief overview of handwritten Odia numerals, Sect. 4 highlights the feature extraction process, Sect. 5 analyses the experimental results and the conclusion in Sect. 6.

2 Literature

According to the reports published, there have been a number of research studies for Odia character recognition. Roy et al. used contour-based features for offline Oriya isolated numeral recognition in 2005 [1]. In this, they have made use of the bounding box technique and chain code histogram. The bounding box of each numeral is decided and then divided into blocks. For each block, the chain code histogram of the contour points is computed. Finally, quadratic and neural network (NN) classifiers have been used for recognition, thus achieving a recognition rate of 94.81% for quadratic and 90.38% for NN. In 2007, curvature features have been exploited by Pal et al. in [2] for identification of the Odia characters with accuracy. Desai [3] in 2010 performed Gujarati handwritten digit recognition with a success rate of approximately 82%. Jindal et al. [4] had classified eight major Indian printed scripts and have achieved 96.45–99.79%. Banashree et al. [5] in 2007 achieved an accuracy level of 98% by implementing half toning algorithm for Hindi numeral recognition using 16-segment concept for feature extraction and NN classifier. In 2009, Shanthi and Duraiswamy [6] applied support vector machine (SVM) for the recognition of handwritten Tamil character and attained an accuracy of 82.04%. Mishra et al. [7], in 2013 used discrete cosine transform (DCT) and discrete wavelet transform (DWT)-based features for Odia handwritten numeral recognition, thus achieving a recognition accuracy of 92% and 87.50%. In [8], Pushpalata and Babita implemented the Odia numeral recognition by taking the curvature and the gradient features using support vector machine. They obtained a recognition rate of 90.5% and 95.5%.

Fig. 1 Sample handwritten numeral Odia images from ISI Kolkata Database



Tasnuva and Haider in [9] have exploited three different variants of LBP for the recognition of handwritten Bangla numerals. They have used K nearest neighbours (KNN) classifier and achieved a recognition accuracy of 96.7%.

3 Odia Handwritten Numerals

In India, there are 22 major languages, which are written in different scripts. Odia is regarded as one of the finest and structured languages of India. At the same time, it is the official language of the state of Odisha. There are 10 numerals form 0 to 9 in the Odia script. Figure 1 shows different sets of Odia handwritten numerals form 0 to 9. From the figure, it can be seen that the numeral six, seven and nine appear in different shapes. Because of the variations in shape and structure of different numerals written by different individuals, the identification task becomes more challenging.

For implementation, the Odia numeral database provided by ISI, Kolkata have been considered. The database contains near about 500 images from each class.

4 Local Binary Pattern

The original LBP proposed by Ojala et al. in 1996 emerged as an effective texture descriptor which has proven to be highly discriminative [10]. Because of its benefits such as luminance invariance, rotation invariance, easy implementation and low computational cost, it is suitable for real-time applications. The LBP computation is done for a 3×3 window where the neighbouring pixels are thresholded with respect to the central pixel. The result is a binary string of 0 s and 1 s which are multiplied by powers of 2 and summed to obtain the label for the central pixel. The LBP operation is explained according to the following mathematical expressions:

$$s(x) = \begin{cases} 0, & x < 0 \\ 1, & x \geq 0 \end{cases}$$

$$\text{LBP}(x, y) = \sum_{p=0}^{P-1} s(i_p - i_c) 2^p$$

where i_p and i_c are the gray intensity values of the neighbouring pixels and centre pixel, respectively. A 2^p -bin histogram is computed which is a high dimensionality feature vector having a good discriminative property. Some pattern occurs more frequently than others and the number of bitwise transition for them is not more than two. Such patterns are called uniform patterns, and the rest patterns are non-uniform. For example, 00110000 is a uniform pattern where the number of transitions is 2 and 10101011 is considered to be a non-uniform pattern as it consists of 6 transitions. Each uniform pattern is allotted a bin, and all the non-uniform patterns are collected and summed up into a single bin, thus reducing the bin numbers. For $P = 8$, the number of bins is reduced from 256 to 59.

5 Feature Extraction and Classification

5.1 Feature Extraction

Uniform LBP has been employed for the recognition of Odia numerals as it provides a compact and discriminative feature vector. Three different blocks are considered in this experiment. First, the LBP converted image is considered as a block for which the histogram of size (59×1) is obtained which serves as the feature set. Second, the LBP image is split into 4×4 and 8×8 block size. Histograms are calculated for each block and then concatenated to obtain a feature set. The method is illustrated in Fig. 2. For an image size of (24×24) , a feature set of size (2124×1) is obtained for block size of 4×4 and for (8×8) block, the feature set size is (531×1) .

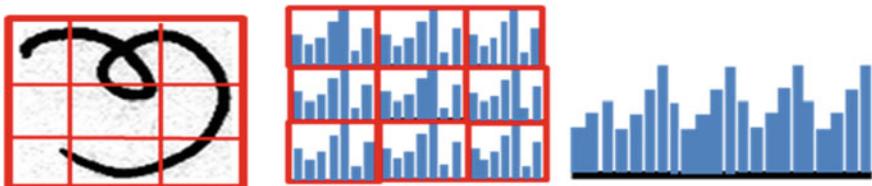


Fig. 2 Illustration of LBP feature extraction 8×8 zoning, corresponding local LBP histograms followed by concatenated histogram

6 Results and Discussion

We have successfully implemented the block-based uniform LBP for the recognition of Odia numerals. The performance of the proposed approach was evaluated based on its recognition ability of the various numerals. The Odia numeral database provided by ISI Kolkata was taken into consideration for this purpose. The data set consisting of near about 500 images for each of the 10 classes from 0 to 9 was divided into training set and the testing set. 400 randomly selected sample images from each numeral class were chosen as the training set where the rest 100 images formed the test set. The results for different schemes were compared using the recognition accuracy obtained from the confusion matrices. A sample confusion matrix is shown in Table 1 for SVM classifier in Table 1 using 4×4 block-based uniform LBP features. The number of correctly identified numerals is reflected along the diagonal of the confusion matrix for each numeral. The overall recognition accuracy of each block-based LBP with different classification schemes is shown in Table 2.

The LBP features with three different blocks were considered for performance evaluation. First, the whole image was considered as a block and upon applying uniform LBP on it, a histogram of size (59×1) was obtained. For this block size, SVM classifier resulted in 45.3% accuracy while single decision tree yielded an accuracy of 82.3%. The highest recognition accuracy of 97.4% was obtained with

Table 1 Confusion matrix of 4×4 block using SVM classifier

		Predicted class									
		0	1	2	3	4	5	6	7	8	9
Actual class	0	100	0	0	0	0	0	0	0	0	0
	1	1	99	0	0	0	0	0	0	0	0
	2	0	0	96	0	0	0	0	3	0	1
	3	0	0	0	98	1	0	0	0	0	1
	4	0	0	0	0	97	1	0	0	1	1
	5	0	0	0	0	0	100	0	0	0	0
	6	0	0	0	0	0	0	95	5	0	0
	7	0	0	5	0	0	0	2	93	0	0
	8	0	0	0	0	1	2	0	0	97	0
	9	0	0	0	0	1	0	0	0	0	99

Table 2 Overall accuracy of Odia numeral recognition

Database	Block size	Recognition accuracies	
		SVM	Single tree
ISI, Kolkata	1×1	45.3	82.3
	4×4	97.4	94.8
	8×8	92.3	94

Table 3 Computational time for different recognition techniques

Block size	Computational time in sec for classifier schemes	
	SVM	Single tree
1 × 1 (whole image)	14	9
8 × 8	24	21
4 × 4	26	23

SVM classifier for a block size of (4 × 4). For the same block size, single decision tree produced an accuracy of 94.8%. It was observed that with reduction in block size from (24 × 24) to (4 × 4), the recognition accuracy in both the classifier schemes improved. It was also noticed that single decision tree maintained the accuracy between 82.3 and 94.8%, and the rate of improvement with decrease in block size is small as compared to SVM. SVM showed rapid rate of improvement with reduction in block size, i.e. increased number of features. Decision tree produced fair results with reasonable accuracy even with larger block size which was not seen in SVM.

The computational time for each block-based scheme with the two classifiers is presented in Table 3. Single decision tree certainly outperformed SVM in terms of speed of computation.

7 Conclusion

This paper presents a new feature extraction technique for the Odia digit recognition. The performances are evaluated for uniform LBP using three different types of block and two classifier techniques (SVM and single tree). Looking at the outcomes, it is suggested that SVM is a better choice where accuracy is of utmost importance while decision tree is a better choice where faster computation is essential along with reasonable accuracy. The results are encouraging and can be further investigated for improvement in accuracy by combining LBP features with other conventional features and designing more robust classifier.

References

1. K. Roy, T. Pal, U. Pal, F. Kimura, Oriya handwritten numeral recognition system, in *Proceedings of Eighth International Conference on Document Analysis and Recognition*, vol 2 (2005), pp. 770–774.
2. U. Pal, T. Wakabayashi, F. Kimura, A system for off-line Oriya handwritten character recognition using curvature feature, in *10th International Conference on Information Technology* (2007), pp. 227–229
3. A. Desai, Gujarati handwritten numeral optical character recognition through neural network. *Pattern Recogn.* **33**, 2582–2589 (2010)

4. M.K. Jindal, R.K. Sharma, G.S. Lehal, Segmentation of horizontally overlapping lines in printed Indian scripts. *Int. J. Comput. Intell. Res.* **3**(4), 277–328 (2007)
5. N.P. Banashree, D. Andhre, R. Vasanta, P.S. Satyanarayana, OCR for script identification of Hindi (Devanagari) numerals using error diffusion Halftoning Algorithm with neural classifier. *Proc. World Acad. Sci. Eng. Technol.* **20**, 46–50 (2007)
6. N. Shanthi, K. Duraiswamy, A novel SVM-based handwritten Tamil character recognition. *Pattern Anal. Appl.* (2009)
7. T.K. Mishra, B. Majhi, S. Panda, A comparative analysis of image transformation for handwritten Odia numeral recognition, in *International Conference on Advances in Computing, Communications and Informatics (ICACCI)* (2013)
8. P. Pushpalata, M. Babita, Recognition of offline handwritten Odia numerals using support vector machine, in *International Conference on Computational Intelligence and Networks* (2015)
9. H. Tasnuva, K.A. Haider, Handwritten Bangla numeral recognition using local binary pattern, in *2nd International Conference on Innovations in Science, Engineering and Technology (ICISET)* (2015)
10. T. Ojala, M. Pietikainen, T. Maenpaa, Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *Pattern Anal. Mach. Intell. IEEE Trans.* **24**(7), 971–987 (2002)

Performance Evaluation of Different Machine Learning Techniques for Detection of Non-technical Loss



Adyasha Banajyoti and C. N. Bhende

Abstract The percentage of losses in India in transmission and distribution sector of electricity has been fairly high. In distribution system, a considerable amount of energy is dissipated which can be categorized into technical and non-technical losses. It is possible to control and compute technical losses, provided the load quantities are known for the given power system, whereas the non-technical losses do not have any recorded information as it is difficult to track energy theft due to the act of meter tampering or bypassing the measurement system. Generally, sudden or surprise checking is done in localities, where the electricity theft is suspected by the distribution companies. However, these operations alone are not enough to identify the miscreants or to reduce the energy losses. Moreover, manual inspection is quite tedious and can be very costly. Thus, certain advance technologies like machine learning techniques need be used to counter the electrical theft more effectively. In this paper, various machine learning techniques are discussed and their performances are compared for the detection of power theft in power system.

Keywords Machine learning · Power loss · Theft detection · Long short term memory · Boosting

1 Introduction

The economy of the country gets affected because of the losses in the power system. All energy that is generated and sent to the distribution and transmission system fails to arrive at the end consumer. We are losing a considerable amount of energy in our distribution system, which can be split off into the technical and non-technical losses. The distribution system is responsible for highest non-technical and technical losses.

The electrical components of the power system are responsible for the technical losses in power system. They occur naturally and it is mostly due to the power lost in

A. Banajyoti (✉) · C. N. Bhende
Indian Institute of Technology Bhubaneswar, Bhubaneswar, India
e-mail: Aab45@iitbbs.ac.in

measurement equipments, power transformers, most importantly transmission lines, etc. It is possible to compute and manage technical losses when we know load quantities and the system parameters.

However, non-technical losses are mostly because of the electrical theft, billing errors, defective energy meters, etc. Hence, these losses are more difficult to measure and there is no recorded information regarding the same.

Installation of smart meters can help till some extent but people have found advanced methods of the energy theft like attacking the operating system of the metering device.

The service providers usually perform unexpected visits to inspect in some suspected areas of increased electrical theft now a days. However, to capture the miscreants or to minimize the losses, these steps alone are not sufficient. Also, the manual inspection is a quite tedious and can be costly. Hence, to tackle the electrical theft more successfully, utility companies need to look into some of the latest technologies available in market. Using smart meters, we can very easily collect the electricity data of customers and we can use advanced data mining techniques to study vast data to find the electrical theft and hence the study of various machine learning techniques for the detection of energy theft. With the developments in the domain of machine learning, many strategies are put forward to identify and lessen the theft of energy, which are researched more in the literature survey.

In this paper, the performance of various machine learning schemes is analyzed for the detection of power theft. Section 1 contains the brief introduction on the power theft scenario in India. Section 2 consists of the researches on different type of energy theft detection techniques. There is a detailed working flow of the theft detection models and their results in the Sect. 4. Section 5 concludes the paper along with the comparison of all the models for the evaluation of their performance.

2 Various Machine Learning Techniques

Before adding details to the theft detection models, we need to be well acquainted with these machine learning techniques. The briefings of various machine learning schemes are given below.

2.1 Decision Tree Coupled Support Vector Machine (SVM)

SVM is used for classification problem may that be binary classification or multi-class classification. Many approaches based on SVM used the load profile of customers to expose irregularities or anomalous behavior which is highly related to NTL activities [1]. But in this model, the decision tree (DT) and SVM have been combined to detect non-technical loss [2]. At first, all the input features such as number of persons, temperature, electricity consumption, time of day are given to the regression DT,

which predicts the expected electricity consumption of the costumers. Then, along with the previously defined input features, this newly generated expected electricity consumption data (output of DT) is supplied as input to the SVM. Finally, the SVM helps in detecting the energy theft in power system.

2.2 *Probabilistic Neural Network (PNN)*

Along with the input and output layer, the PNN has two more layers namely pattern layer and summation layer. The output layer generates a vector of probabilities for each input. In the final layer, the highest of these probabilities are picked and are labeled as 1 (positive, i.e., done theft) and others are labeled as 0 (negative, i.e., No theft). This inspiration has been taken from a paper, where a probabilistic neural network (PNN) classifier along with S transform is used for the identification of different power quality problems [3].

2.3 *Stacked Long-Short Term Memory (LSTM)*

Long-short term memory network is an advancement of recurrent neural network (RNN), which has the capability to remember very old information. It is usually done by explicitly introducing a memory unit into the network. Our electricity consumption data is time series data, hence the use of LSTM. Different papers have used LSTM or gated recurrent unit (GRU), which is another advancement of RNN, for predicting the power consumption and theft identification [4, 5].

2.4 *Convolutional Neural Network with Long-Short Term Memory (CNN-LSTM)*

Like other deep neural networks CNN also has input and output layer along with many hidden layers. In the CNN, the convolutional and pooling hidden layers are subsequently accompanied by the dense layer for dimension alteration of output of hidden layer. In this model, the CNN and LSTM are combined for detecting non-technical loss. The CNN has the capability to extract feature from a given data set, which made it useful for our model [6].

2.5 Gradient Boosting Algorithms

Boosting algorithms combine multiple weak models to create a strong model. Over the last few years, boosting algorithms have become immensely popular in the machine learning field because they are less affected by the over-fitting problem. Gradient boosting machine (GBM) uses multiple decision trees to generate the final prediction. In GBM, the errors done by the previous trees are considered in the each new decision tree generated. In quite a number of papers, recently developed gradient boosting classifiers were used for the detection of NTLs [7]. GBM has two variants, one is extreme gradient boosting and other is light gradient boosting which are given below.

Extreme Gradient Boosting (XGBoost) It is an advancement of the GBM algorithm. XGBoost follows the same working procedure as GBM. In this algorithm, the tree is split level-wise or depth-wise. XGBoost implements parallel preprocessing, which in turn makes it faster than GBM. The introduction of variety of regularization techniques in XGBoost helps in reducing over-fitting and finally improves the overall performance. Also, the XGBoost model is capable of handling the missing value cases in the dataset on its own.

Light Gradient Boosting (LightBoost) LightBoost algorithm is also an alteration of the GBM. The speed of execution and efficiency of the LightBoost model is the reason of its popularity. It can easily handle bulk data, but it does not perform well with a small dataset. Instead of a level-wise growth, the trees in LightBoost grow in a leaf-wise manner. LightBoost uses binning technique for continuous data, which results in lower memory usage and better efficiency. The leaf-wise split enables the LightBoost to perform good with bulk datasets with a notable reduction in execution time as compared to XGBoost.

3 Data Processing and Model Evaluation Metrics

3.1 Data Preprocessing

To be able to relate more with the real-time scenarios, we have selected various features for our dataset. These features are number of persons, temperature, time of day, season, electricity consumption. To validate proposed schemes, we have taken the hourly energy consumption data of six months for the residential area of IIT Bombay [8]. From various other sources, we have collected other features of our dataset and later they are linked together. From the historical meteorological dataset, we have extracted the details about season and temperature with respect to time and date. As we couldn't find the number of persons in a specific house at a specific time, an average number of persons are reckoned in every house at that specific time.

3.2 *The Unbalanced Data Problem*

When the ratio of observations in each class are not in proportion that is called as an imbalanced dataset, which is a familiar issue in machine learning classification. Same is the case for our dataset, where the percentage of theft cases is merely a 10% of the whole dataset. So, the accuracy can be misleading in this case. For boosting methods, this problem was quite evident, hence data balancing technique called as synthetic minority over-sampling technique is used for boosting methods [9].

Synthetic Minority Over-Sampling Technique (SMOTE) Removing random observations from the majority class is called under-sampling, and adding duplicates of minority instances to the dataset is called over-sampling. SMOTE is one type of over-sampling method. While randomly removing data in random under-sampling, it may discard some important information and in random over-sampling duplicates may cause over-fitting. So, in SMOTE instead of taking duplicates, a subgroup of the dataset is considered from the minority class and then new synthetic similar instances are created, which solves the problem of over-fitting and useful information are not lost like under-sampling case.

3.3 *Model Evaluation Metrics*

In the field of machine learning, a confusion matrix depicts the results of the prediction by a classification problem. The number of right and wrong predictions are computed for each class and entered in the confusion matrix. From the confusion matrix, following performance measures can be found out,

Accuracy The number of cases correctly predicted to be doing theft or not doing theft divided by the total number of cases gives the accuracy. But sometimes accuracy can be misleading because it gives equal weightage to the errors of both the classes (theft and non-theft). Accuracy can be high even if our model is not detecting theft cases and only predicting the non-theft cases. That is why need other measures along with accuracy.

F-Measure As we cannot fully be dependent on accuracy, we compute F-measure. The total number of positive cases, who are also correctly predicted, divided by the total number of positive cases gives recall. The ratio of total number of positive cases to the cases who are predicted to be positive is called precision. It is the harmonic mean of both the measures. The *F*-measure takes into account the worst case between precision or recall.

Table 1 Comparison of different machine learning models

Model no.	Model name	Accuracy (%)	F-measure (%)
1	DT coupled SVM	87.91	63.93
2	Probabilistic neural network	92	70.7
3	Stacked LSTM	90.78	79.85
4	CNN with LSTM	89.8	72.8
5	XGBoost with SMOTE	92.18	82.44
6	LightBoost with SMOTE	95.124	88.33

4 Results and Discussion

Different machine learning models are analyzed; the same consumer data set is provided to them and each model is trained and evaluated. Confusion matrix for each model is generated and accuracy, recall, precision, and F-measure is determined.

To get a clear picture of the performances of the models, a comparison study is performed. The comparison of percentage of accuracy and F-measure between all the models have been given in the Table 1. From Table 1, it can be observed that although the SVM alone and DT coupled SVM have accuracy more than 85%, their F-measure is not up to the mark. The neural networks performed better in detecting non-theft but the detection rate of theft (evident from F-measure) is still not acceptable. The stacked LSTM performed quite well with acceptable accuracy and F-measure. The boosting methods outperformed all other methods, especially the LightBoost with SMOTE model which has more than 95% accuracy and nearly 90% F-measure.

5 Conclusion

In this paper, the performance of different machine learning models are analyzed. From the results, it is concluded that the LightBoost method outperformed all the other presented methods with the accuracy of more than 95% which makes the method suitable for practical use.

References

1. J. Nagi, A.M. Mohammad, K.S. Yap, S.K. Tiong, S.K. Ahmed, Non-technical loss analysis for detection of electricity theft using support vector machines, in *IEEE 2nd International Power and Energy Conference* (2008), pp. 907–912
2. A. Jindal, A. Dua, K. Kaur, M. Singh, N. Kumar, S. Mishra, Decision tree and SVM-based data analytics for theft detection in smart grid. *IEEE Trans. Industr. Inf.* **12**(3), 1005–1016 (2016)
3. S. Mishra, C.N. Bhende, B.K. Panigrahi, Detection and classification of power quality disturbances using s-transform and probabilistic neural network. *IEEE Trans. Power Deliv.* **23**(1), 280–287 (2008)
4. J. Goh, S. Adepu, M. Tan, Z.S. Lee, Anomaly detection in cyber physical systems using recurrent neural networks, in *IEEE 18th International Symposium on High Assurance Systems Engineering (HASE'2017)* (2017), pp. 140–145
5. X. Wang, T. Zhao, H. Liu, R. He, Power consumption predicting and anomaly detection based on long short-term memory neural network, in *IEEE 4th International Conference on Cloud Computing and Big Data Analysis (ICCCBDA'2019)* (2019), pp. 487–491
6. R.R. Bhat, R.D. Trevizan, R. Sengupta, X. Li, A. Bretas, Identifying non-technical power loss via spatial and temporal deep learning, in *15th IEEE International Conference on Machine Learning and Applications (ICMLA'2016)* (2016), pp. 272–279
7. R. Punmiya, S. Choe, Energy theft detection using gradient boosting theft detector with feature engineering-based preprocessing. *IEEE Trans. Smart Grid* **10**(2), 2326–2329 (2019)
8. Open energy Info. <https://seil.cse.iitb.ac.in/residential-dataset>. Last accessed 13 Nov 2019
9. H. Lee, S. Jung, M. Kim, S. Kim, Synthetic minority over-sampling technique based on fuzzy c-means clustering for imbalanced data, in *International Conference on Fuzzy Theory and Its Applications (iFUZZY'2017)*, Pingtung (2017), pp. 1–6

An Overview with Current Advances in Industrial Internet of Things (IIoT)



T. Primya, G. Kanagaraj, and G. Subashini

Abstract Industrial IoT is a gifted technology which combines many industrial systems with Internet connectivity. Perception techniques have been concerned to considerably get better product efficiency and diminish production cost by cooperating with intellect devices, in which highly developed computing, big data analysis and intelligence. IIoT is frequently involved in industries related to manufacturing sector linked to industrial IoT subset. Working of Internet of things developments are industrial processes, supply chains, products and services. This paper meticulously surveys the what is meant by IIoT, current advances of the industrial Internet with new budding technologies as IoT, big data analytics, cloud computing, cyber-physical-social systems, artificial intelligence.

Keywords Artificial intelligence (AI) · Cyber-physical-social systems (CPSS) · Cyber-space · Industrial Internet of things (IIoT)

1 Introduction

IIoT is nothing but interrelation between sensors, tools, various devices combined altogether with computers application of industries, involves developmental and energy supervision. This combination helps for collection of data, transfer, and overview together involves development in productivity and usefulness and also much economic increment. IIoT is growth of a distributed control system (DCS) helps to improve procedure control which includes major amount of computerization with cloud computing. IIoT facilitates by technology like edge computing, cloud

T. Primya (✉)
Dr. N.G.P. Institute of Technology, Coimbatore, India
e-mail: primyacse@gmail.com

G. Kanagaraj
Kumaraguru College of Technology, Coimbatore, India

G. Subashini
PSG College of Technology, Coimbatore, India

computing, mobile technologies, 3D printing, advanced robotics, cyber-security, big data, Internet of things, RFID technology, machine to machine, and cognitive computing [1].

IoT is combination of many applications as consumer and industrial. In consumer markets, it is classically worn to depict coupled applications like wearable, temperature control, systems for providing security in home, shopping, applications for planning travel and further. Speedy enlargement is consequence of push-pull of market forces, consumer command for expediency and smart, connected applications services is coordinated by corporate attention in collecting, leveraging that similar information into novel growth opportunity [2]. While reviewing through IIoT, found many extensive activities when linked with below terms.

2 Internet of Things (IoT) and Industrial Internet of Things (IIoT)

IoT is system with unique identifiers (UIDs) of consistent mechanical and digital machines, computing devices, and capability to reassign network data devoid of involves human-to-human or human-to-computer interaction. Due to junction of several technologies, real-time analytics, ML, commodity sensors, and embedded systems (ES), description of IoT have emerged. Fields which are contributing to enable IoT are control systems, WSN, ES. In market place of user, products are compared with IoT technology using notion of “smart home”, devices pertaining to casing, one or more ecosystem holding devices and inhibited via associated devices involved in that ecosystem, as smart speakers and smart mobiles [3].

IIoT differs from IoT in its architectural behaviours, system performance with more power, and various applications in medical field also are shown in Fig. 1. Internet has altered how people converse, what they perform, and how they effort mutually. At present, concentration has shifted to undertaking identical for machines. For precedent a small number of durations, transferring operations into noteworthy efficiency environments by system developers includes sensor interconnection, finding edged in nodes, put together elegant systems is IIoT [3].

Fourth industrial revolt brings everyone troublesome in history of industrial automation, disturbing energy industries to healthcare, and industrialized to transportation. Swiftness of revolutionize with technical leaps also accelerates. IIoT runs various industrial applications that involve campaign and sensors, consistent, and use of computerized machines. With a strapping focal point on machine learning and big data, IIoT includes industry and venture to supplement operation dependability and competence with dense involvement on communication between human and machine [4].

IIoT focuses mainly on various industrial applications like energy sector, manufacturing industries, agricultural field, and many more. It has diverse technical needs like interoperability, complexity-level increment, and aspects related to security. Device

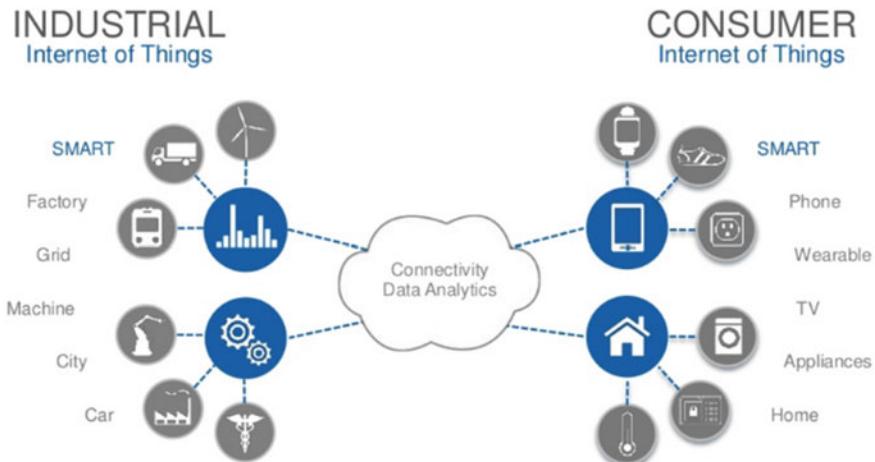


Fig. 1 Industrial Internet of things with Internet of things

which monitors personal fitness uses same technology, but it is different from recent industrial applications like remote robotic surgery. IoT versus IIoT faces similar challenges in technical wise but autonomous complexities and risks are very high. IIoT got good stalk with private-public partnership. To work IIoT in reliable, safe, and secure manner, many organizations unitedly work on innovations in technical sides [5].

3 Cyber-Physical Systems (CPS) and Cyber-Physical-Social Systems (CPSS)

CPS is important enabler to reunion physical machines that are formerly disjointed and also an essential equipment podium for IoT and IIoT. CPS gives abstractions, modelling, analysis, and design techniques which helps to integrate dynamics of substantial procedure with those of communication and software. Computer system inhibited by computer-based algorithms. Here, physical and software components are intensely tangled, work with reveal multiple and separate behavioural modalities, dissimilar temporal and spatial scales and help in modifying situation with traditions in former.

CPS examples are autonomous automobile systems, automatic pilot avionics, medical monitoring, and control systems in industries. In Lieu of standalone devices, CPS deliberates as network of interconnection elements with physical output/input. This extends its perspective in many aspects like surgery using robots, avoidance in collision, controlling air traffic, monitoring in healthcare appliances, and many more [6].

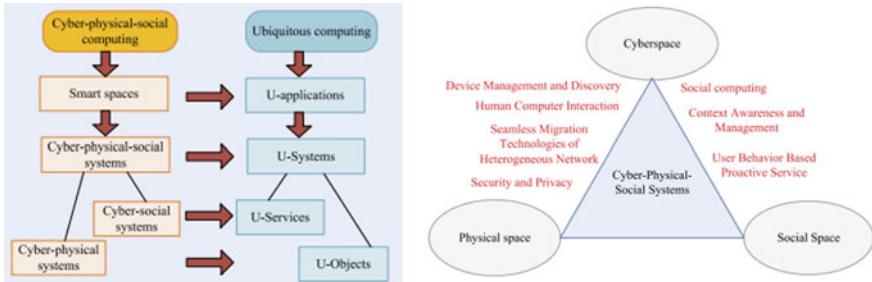


Fig. 2 Comparison of CPSS with CPS

Disclosure of CPSS has revolt in relationship with computer, human/physical environment. CPS with social domain, feature human contribution is CPSS. In this intelligence, communications with CPSS take in social description/associations. Interaction of socio-human exerts cyber-space on primary control. Figure 2 explains how the CPSS with CPS differs. It particularly establishes an effectual relationship among human and machine intelligence. Further method of CPSS over CPS is characteristic in social network. Main target is combining sensing applications of community which helps people in bonding with sensor objects [7].

Down with CPS, CSS enhanced commonly linked, ultimately incorporated into CPSS allow interfaces between physical and cyber-space. CPSS includes all functionalities of CPS and CSS. CPSS varies in need of obliging a range of submission. CPSS includes human centric services in computation with many capabilities. It helps to find peoples difficulty in social interaction with physical world. Hence, it develops bulky errands in both social/physical world for smooth work progress [8].

4 Cloud Computing Role in IoT Revolution

Cloud computing and IoT together play major role in everyday life. While considering future embracing Internet, and utilization of Internet are considered as very persistent and important component in case of IoT. Combination of cloud and IoT is predicted as innovative and many application scenarios are considered as narrative concept.

IoT is usually categorized by real-world diminutive belongings, extensively dispersed, with incomplete storage space and dispensation ability, which engage concerns concerning reliability, performance, security, and privacy. Processing power and storage are considered as very less capabilities in cloud computing so that problems in IoT are not able to solve completely. Hence, cloud and IoT are joined together upset as new paradigm called CloudIoT [9].

Generation of huge volume of big data by IoT makes Internet infrastructure to become more burdens. Because of this, companies are compelling to resolve their faults and also to find answer for their faults to reduce tension. Giving scalability in

relief of endeavour applications and SaaS, cloud computing is considered as important role in information technology. Cloud is providing services to companies also so all are migrating their data to cloud platform. Convention Internet connection or purposeful direct link is used by cloud providers to transfer data. The data is discontent and Internet does not cross traffic limit is ensured by cloud is direct link benefit and QoS can be embarrassed. In presence of IT services of cloud computing, Internet is used to upload and also retrieve resources from server's direct connection. Cloud-based storage systems are preferred to store files instead of local storage devices. Cloud computing and IoT together are interrelated and help to improve effectiveness in day-to-day tasks [10].

Huge amount of data is generated by IoT and cloud computing helps to maintain this data in effective way. If we use some resource, it is enough to pay for usage of those resources only not more than is way of charging provided by various cloud providers called as pay per use model. Cloud providers help small size IoT start up companies and also entire amount of IoT companies is scale of economy.

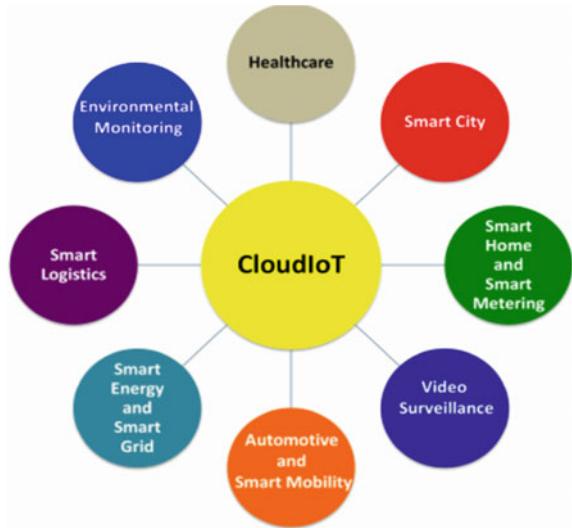
Main thing for today's developers is to enable better collaboration between IoT and cloud computing. It helps developers by accessing and storing data remotely and also finds way to work with no delay on project. Recourse allocation in various areas enables IoT companies to change their data as their need after cloud storage. Cloud plays even a major in this world even after emergence of big data technologies now. This way finds helpful to many companies feasible to access all big data via cloud technologies [11].

5 Challenges of Deploying Real-Time Workflows in Cloud

It is a method to make more efficient association of traffic from IoT devices and realize real-time local data analysis. With help of edge computing, data shaped by Internet of things devices be processed where it is shaped in its place of taking gone to the routes to data centres [10].

Before receiving data at data centre or cloud, at edge of network data are analysed by IoT devices by edge computing. Edge computing allows data from IoT devices to be analysed at edge of network before being sent to a data centre or cloud. Data storage of computer is brought closer to location; its needs are paradigm of distributed computing. Decentralization of processing data at networks end is edge computing compared with cloud computing. Industrial Internet needs both edge and cloud architecture then completely centralized cloud; it helps to switch over services, products, and productivity in world of industrial Internet. Hertzog [12] all explained in Fig. 3. Hustle matters in today's hyperconnected world. Approachable connectivity is way of finding, maybe it is of real-time contact or IoT device handling. Conventional architecture of cloud concentrates handling out data centres in cloud-based applications, or between data source and user. Even though computation of complex work is done by cloud, for real-time workflows, data *centres* are very slow

Fig. 3 Scenarios driven by the CloudIoT



in case of dispensation. Data is processed and analysed by reduction of remoteness among communications being inspected [17].

6 Roles of Big Data in IoT

In real time, many devices will collect, analyse, transfer data. IoT devices attain attention in worldwide because of capabilities and functionalities hold by big data. Huge quantity of processing data instantaneously is big data role in IoT [13]. In Fig. 4, it stores data processing solution by various technologies of storage. While comparing with big data systems, massive amount of generation of not structured information is given by IoT devices. These kinds of big data generated by IoT devices based on 3 V factors as volume, velocity, and variety. Massive amount of data storage done in big data files where shared and distributed database is big data system. Generation of analysed data report is done using analytic tools Hadoop MapReduce or Spark which helps in examining stored IoT big data. It has enormously collision between each other. IoT grows business grows quotes here related to data. Both have interdependency [14].

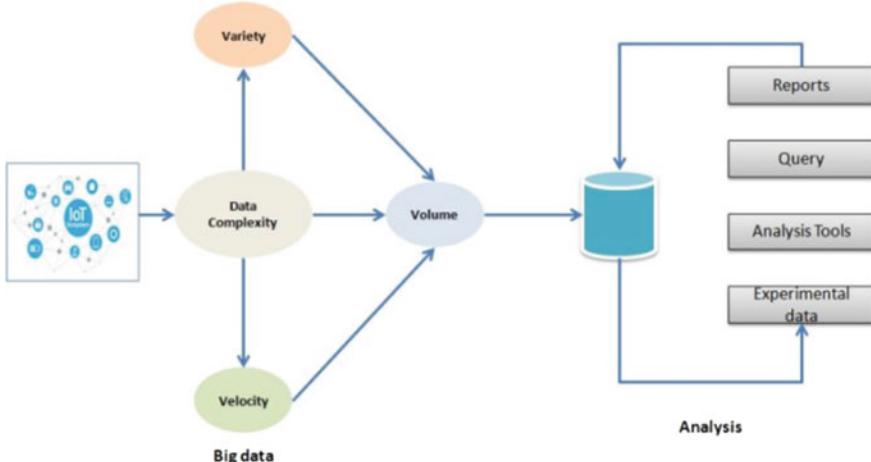


Fig. 4 IoT big data processing

7 Bringing Power to AI to IoT

Intelligence machines which are created to work like and react as humans are meadow in computer science called as artificial intelligence (AI). With no explicit programming, software which finds more accurate outcomes is core part of AI is machine learning. Internet of things is always smarter. Industries incorporate AI in particular, ML into their IoT applications. Finding insights in data is solution. With gesture of speculation, propel of new products, increasing surge of deployment in enterprise, AI is making splatter in IoT. Industries explore role for AI with existing IoT consumption for evaluating new IoT projects, which gives better efficiency in functioning.

ML can identify operating conditions which are not visible to human eye. Math went beside shipping industry instincts are Hulls kept smooth through frequent maintenance improve fuel effectiveness an adequate amount to enormously compensate improved maintenance costs [16]. New products and services can make by power of combining AI with IoT. NLP becomes powerful because it does not require human to operate, leasing persons speaks with machines instead. Nowadays, robot, drones are doing extraordinary performance which cannot done by humans such as scrutiny, monitoring [14].

8 Industry 4.0 and IIoT

Industry 4.0 was started by German government a part of “High-Tech Strategy 2020” in 2010. Connection of value chains is Industry 4.0 [15]. To make CPS, industrial industries can automatically connect and integrate things. Eventual objective is waste

reduction by using new technologies and manufacturing environmental value increment. While considering both, it have many similarities based on approaches. To increase productivity, both are linked together with machine or devices [12].

9 Conclusion

Industrial IoT is an intense incorporation of advanced computing. This paper surveys research works in industrial Internet with its included technologies. This paper with a great effort analyses that what is intended by current advances of the Industrial Internet of things with a variety of emerging technologies like IoT, big data analytics, cloud computing, CPSS and artificial intelligence. The cyber-security and security issues and challenges to be included in the future.

References

1. R. Baheti, H. Gill, Cyber-physical systems, in *The Impact of Control Technology*, ed. by T. Samad, A.M. Annaswamy, vol 2 (IEEE Control Systems Society, New York, 2011), pp. 161–166. Available: <https://ieeecss.org/main/IoCT-report>
2. B. Dorsemaine et al., Internet of things: a definition and taxonomy, in *Proceedings—NGMAST 2015 9th International Conference on Next Generation Mobile Applications, Services and Technologies* (2015), pp. 72–77. <https://doi.org/10.1109/NGMAST.2015.71>
3. The Internet of Things—High-Tech Strategy. <https://www.gtai.de/GTAI/Navigation/EN/Invest/Industries/Smarterbusiness/Smart-systems/internet-of-things.html>
4. M. Broy, Challenges in modeling cyber-physical systems, in *Proceedings of the 12th International Conference on Information Processing in Sensor Networks* (ACM, 2013)
5. D. Bandyopadhyay, J. Sen, Internet of things: applications and challenges in technology and standardization. *Wireless Pers. Commun.* **58**(1), 49–69 (2011)
6. I. Lee, An exploratory study of the impact of the internet of things IoT on business model innovation: building smart enterprises at fortune 500 companies. *Int. J. Inf. Syst. Soc. Change* **7**(3), 1–15 (2016) [Online]. Available: <https://doi.org/10.4018/IJISSC.2016070101>
7. S.G. Pease, R. Trueman, C. Davies, J. Grosberg, K.H. Yau, N. Kaur, P. Conway, A. West, An intelligent real-time cyber-physical toolset for energy and process prediction and optimisation in the future industrial internet of things. *Future Gener. Comput. Syst.* **79**, 815–829 (2018). [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0167739X1630382X>
8. L. Da Xu, W. He, S. Li, Internet of things in industries: a survey. *IEEE Trans. Industr. Inf.* **10**(4), 2233–2243 (2014)
9. H. Flatt, S. Schriegel, J. Jasperneite, H. Trsek, H. Adamczyk, Analysis of the cyber-security of industry 4.0 technologies based on rami 4.0 and identification of requirements, in *IEEE 21st International Conference on Emerging Technologies and Factory Automation* (2016), pp. 1–4
10. M. Hermann, T. Pentek, B. Otto, Design Principles for Industrie 4.0 scenarios: a literature review. Technische Universität Dortmund (2015), p. 11. Working paper (Accessed 12 Sept 2017)
11. L. Püschel, M. Roeglinder, H. Schlott, What's in a smart thing? Development of a multi-layer taxonomy, in *Proceedings of the International Conference on Information Systems—Digital Innovation at the Crossroads, ICIS 2016*, ed. by P. Ågerfalk, N. Levina, S.S. Kien (Association for Information Systems, Dublin, Ireland, 11–14 Dec 2016)

12. C. Hertzog, Smart Grid Trends to Watch: ICT Innovations and New Entrants (2012). Available: <https://www.smartgridlibrary.com/tag/ictot-convergence/>
13. S. Jeschke, C. Brecher, T. Meisen, D. Özdemir, T. Eschert, Industrial internet of things and cyber manufacturing systems, in *Industrial Internet of Things* (Springer, Cham, 2018), pp. 3–19. https://doi.org/10.1007/978-3-319-42559-7_1
14. B. Dickson, What Is the Difference Between Greenfield and Brownfield IoT Development? (2016) Available: <https://bdtechtalks.com/2016/09/22/>
15. Industrial Internet Consortium, What Is the Industrial Internet? (2018) [Online]. Available: <https://www.iiconsortium.org/about-industrial-internet.htm>
16. N. Shah, IT and OT Convergence –The Inevitable Evolution of Industry, Iotforall (2017) Available: <https://www.iotforall.com/it-and-ot-convergence/>
17. G. Kanagaraj, T. Primya, K. Rekha, C. Vinothini, P. Anitha, IoT-enabled water quality monitoring system (2020). https://doi.org/10.1007/978-981-15-0146-3_26

Video Indexing Through Human Face



Sanjoy Ghatak and Debotosh Bhattacharjee

Abstract There are so many techniques for video indexing using face recognition, but all the methods are unable to reduce space and time complexity. This paper proposes a novel way to deal with separate data from video through the human face and ordering and recover it at whatever point vital. This work is centered around the accompanying advances: key frame extraction from the info video, face recognition from key frame, face distinguishing proof utilizing standardized identification, and video ordering through the barcode. First, we recognize a video as information in this work, and from it, each of the edges is obtained. Using the color histogram difference, the main frames are taken from all the frames present in the video. A color histogram is used to find the difference between two images and return the color histogram. Using the Viola–Jones algorithm, faces are identified from the extracted main frames. Using sliding window technique, the image gradient is created from face image and identifies the human face extracted from video using individual EAN-8 barcode for video indexing purposes. The main objective of this work is to identify human face from video and to represent them as EAN-8 linear video indexing barcode. This approach is useful for security purposes, video surveillance network, video description of communication channel, etc. It is useful for purposes of video indexing and retrieval. This reduces the bandwidth of the correspondence, storage space, and time complexity. The effect of this approach is to index the video as a linear EAN-8 barcode based on the human face present in the video after describing the individual human face of the video.

Keywords Indexing · Facial image · EAN-8 barcode · Viola–Jones algorithm · Image gradient

S. Ghatak (✉)

Department of CSE, Sikkim Manipal University, Manipal, Sikkim, India
e-mail: sanjoy1cs@yahoo.co.in

D. Bhattacharjee

Department of CSE, Jadavpur University, Kolkata, West Bengal, India
e-mail: debotoshb@hotmail.com

1 Introduction

Most image indexing methods use low-level characteristics such as texture and color. It is difficult to recognize people who use low-level image and video indexing tools, and it is not possible to index people based on them. A person in the images and video scene is one of the important entities. An approach to combining the detection and recognition of people in images and video sequences is discussed in paper [1]. In the paper [1], they index the video according to a person's face, which is a process that takes more time and space. If we send the human face as information through a communication channel, then it will take more bandwidth, more time, and space. Paper [2] and [1] could not discuss this entire problem. Recognition based on a single image, however, is difficult due to well-known issues such as shifting lighting, posing variability, facial expression, and occlusion. Differences in the face image caused by these factors often exceed those caused by changes in identity. Video-based recognition is widely believed to be capable of resolving the inherent ambiguities of image-based recognition, such as low-resolution sensitivity, pose variations, and partial occlusion, by accurately capturing additional information, resulting in more accurate and reliable face recognition. In addition, video inputs allow facial dynamics to be recorded that are useful for facial identification. So this paper explores an indexing technique to solve the same problems discussed above using EAN-8 linear barcode representation of a human face.

Writing a plan for this paper is as follows: Sect. 2 includes a literature review (or related work) of the human face indexing image. Section 3 explains the current work (proposed method). Section 4 presents experimental results and analysis, and eventually, the paper is done with the conclusion in Sect. 5.

2 Related Work

Nowadays, face is considered an important object for video indexing. According to paper [3], in a video report for video indexing, three data sources or modalities are considered. These three modalities are visual modality, auditory modality, and textual modality. The main focus of this paper was a video search by semantic analysis. In this paper, it is mentioned that a video has three features: visual, audio, and textual and also discussed the details of different types of low-level video indexing. Video indexing through face recognition is not discussed in this paper. Paper [4] addressed the clustering of video system indexing based on the hidden Markov model's face recognition hybrid model and supporting vector machine. In this paper, the human face is divided into five parts: forehead, eye, nose, mouth, and chin, and then, the SVM is designed to classify those devices and check for the independent features of those devices. In paper [2], narrating structure of the video is consequently deteriorated into shorter lucid sections by methods for AI, with the goal that clients can without much of a stretch get a look of the substance and recover a particular piece of

video. In [5], the author proposed a method of indexing and retrieving video based on a sparse representation of bag of faces. Mingtao Pei [6] developed a face video retrieval system focused on profound learning of the representation of binary hash. The researcher developed a deep convolution neural network (deep CNN) in this paper to learn from face-to-face video retrieval discriminative and compact binary representations. This paper addressed the question of paper [7] and [8] (large intra-class facial variations and the strong demands of saving time and space) and solved these problems. An approach is addressed in paper [1] to overcome the problem of low-level video indexing related apps. The researcher used face detection and face recognition based on a neural network, a pseudo-two-dimensional HMM and a k-means clustering algorithm for face detection purposes. This method is unable to track the faces. Speech variation is a challenge in audio modality video indexing technique. The writer addressed a method of addressing this variability in paper [9] by choosing key frames from videos based on the lip motion's temporal analysis. Paper [10] discussed a technique for automatic detecting human face in a generic video sequence. The author used an iterative algorithm to give a confidence measure for the absence or presence of faces within the video shots. The author discussed software regions in paper [11], may video ordering and recovery difficulties, and some possible video ordering headings, video recovery of the system of the executive. After survey and analysis of all this paper, it is concluded that video indexing through low-level features is complicated for person recognition. All these methods are time and space consuming from storage space and an indexing point of view. The facial expression, pose, emotion, illumination change, and occlusions of the face image of the video are also the important factors of video indexing through the human face. It also observed that the inherent ambiguities such as sensitivity to low resolution, pose variations, and partial occlusion of cavities in the face are affected by video-based recognition. To solve these issues with the current implementation of face detection, there needs to be a system which should be able to fulfill the followings:

- (i) With minor variations in lighting conditions, facial expressions, and minor variations in face posture, the device will allow sufficiently accurate performance.
- (ii) The method should give sufficiently accurate results without discriminating against age and color of the person.
- (iii) Grant these results in quasi-real time.

To overcome these issues mentioned above, a new approach (video indexing through the human face) is proposed in this paper. In this technique, instead of the face image, linear EAN-8 barcode of face images are used for indexing human face present in many types of video. For storing the EAN-8 linear barcode, fewer amounts of storage space and time are required than original face image. Nevertheless, sending this barcode through the communication channel requires less bandwidth than the original face picture. But the drawback of facial image barcode representation is stability. In this work, the sliding window technique is used to create a stable barcode, which are mentioned in paper [12] and [13].

3 Proposed Method

The following are different stages of the method which is proposed in this document and represent in Fig. 1 with a block diagram. (a) The first stage is the extraction of frame from the input video; (b) key frame are extracted from the frames obtained from input video using color histogram difference; (c) use the Viola–Jones algorithm to detect faces from the main picture, (d) all the face images are converted into the grayscale face image; (e) from grayscale, face image gradient is calculated using a sliding window technique; and (f) from image gradients, an EAN-8 barcode is generated using the EAN-8 sequence table.

3.1 Frame Extraction

Dynamic video is a combination of the scene, shot, and frame. Therefore, the first step is to extract the still images, which are represented from input videos as a scene, shot and picture. The scene is a collection of shot, and shot is a set of frames. Traditional video has an abundance of content and includes 20–30 frames per second. The frame is a still image present in a video and contains redundant information.

3.2 Key Frame Extraction

The frame which represents vital information of each shot is known as key frame. In this paper, human face with different expressions, pose, and lighting conditions is considered as key frames. Different types of approaches for key frame extraction are discussed in paper [9, 14]. Some of the approaches are a pair-by-pair comparison, probability ratio, curve saliency motion capture data, and many more. But the color histogram technique is used in this process to retrieve the key frame from all of a given video's frames. Key frames are obtained from the frames using the difference in the color histogram. If the obtained difference is greater than the threshold value,

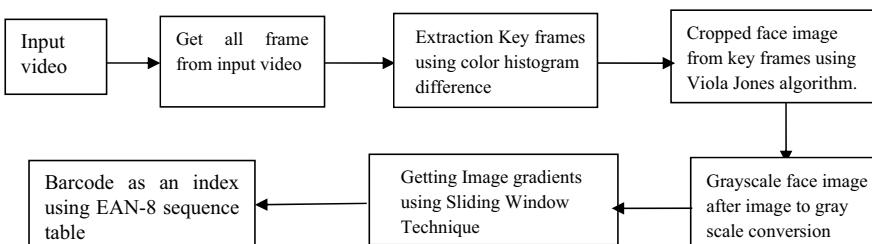


Fig. 1 Steps for the proposed method (with block diagram)

the frame will be selected as the next key frame. The threshold is the amount of variation in the color histogram.

3.2.1 Color Histogram Method

A color histogram used for key frame extraction is a color-based frame difference technique. The idea behind this approach is that two frames with uniform background and uniform (although moving) objects will have little difference in their corresponding histograms. Threshold selection is an essential part of this method. The formula for calculating the color histogram difference between the two consecutive frames is as follows:

$$D(F_I, F_{I+1}) = \sum_{J=1}^n \frac{|h_I(J) - h_{I+1}(J)|^2}{h_{I+1}(J)} \quad (1)$$

where h_I and h_{I+1} indicate the histogram of the two consecutive frame F_I and F_{I+1} , respectively. When $D(F_I, F_{I+1})$ is higher than the given threshold, then a shot transition occurs. “ n ” is the number of the frame.

3.3 Cropped Face Images from the Key Frames Using the Viola–Jones Algorithm

Using the Viola–Jones object detection process, faces are identified from the extracted key frames. Details of this method are discussed in paper [15]. Advantage of using Viola–Jones algorithm is that detection is fast, but training is slow.

3.4 Grayscale Face Image After Image to Grayscale Conversion

After getting the faces from key frames, face image to grayscale conversion is required to calculate the gradient of the face image.

3.5 Image Gradient Using Sliding Window Technique

From this grayscale image (faces), image gradients are calculated using the sliding window technique. A detail of this method was discussed in paper [12] and [13].



Fig. 2 Barcode of hollywood (victory bad scene) face video dataset

3.6 Barcode as an Index Using EAN-8 Sequence Table

In this section, the method for EAN-8 barcode generation from the obtained gradient value will be discussed. In our method, the barcode is a linear representation of the face image detected from the video. With the help of these barcodes, human faces detected from any types of videos can be indexed because these barcodes are a linear representation of face images. Advantage of indexing with barcode from the human face is that those barcodes will take less amounts of storage space and minimum time for indexing. Details of the barcode generation method are discussed in paper [12, 16], and [13], and an algorithm for the barcode generation process is also described in this paper. Figure 2 shows the examples of some barcodes, and those are detected from different Hollywood (victory bad scene) face video dataset.

3.7 Accuracy of Barcode Determination

The accuracy of the barcode of human face is measured based on 8-digit number of EAN-8 barcode of two same or different face images. Here, checksum digit is not considered because it is an error correction digit. Remaining 7 digits are considered. Accuracy is calculated after comparing the digit present in the two EAN barcode of the same location of two human faces of the same image or different images. Comparison is made between two different barcode numbers in the same place. If the same position number of two different barcodes is the same, then the different face image is considered the same face. For measuring the accuracy of particular barcode, compare all the barcode generated from the face image of the input video. After comparing the two barcodes if the accuracy is greater than 80%, then two barcodes are considered as same faces, and if it is less than 60%, then the barcode of the image is considered as different faces. However, if the accuracy is less than 60% and the barcode of the face images is the same faces, then these types of the barcode of the faces are not considered as indexing. But for record maintaining purpose, it can be indexed. In this case, it can create ambiguity. The formula for measure the accuracy of the barcode is as follows:

$$\text{Accuracy} = ((F - I)/F) * 100 \quad (2)$$

where F is the size of the barcode and I is the initial value, which is considered as zero.

4 Experimental Result and Discussions

In our research, the key frame is extracted from the human face-based video dataset, and then, barcode is generated from these cropped faces. After that, indexing is done based on the barcode of the human face present in the video. However, some video dataset directly contains key frame of the face image, so from this dataset directly, we are generating the barcode without extracting the key frame. (YouTube face video dataset). We use the three different video data sets to test our process. At first, Hollywood video data set and then we present a realistic TV series video dataset [15] after this You Tube video data set. Table 1 shows the barcode generated from some of the Hollywood Movies. All the videos of this data set are in AVI format. After observing the result of Table 1, we see that a number of faces detected from datasets *Butterfly Effect—00696*, *Forrest Gump—01277*, and *Gandhi—02262* are 15, 35 and 5, respectively, but a number of valid barcodes of individual face are generated from these data sets are 3, 16, and 4, respectively. The reason behind this is when we measure the accuracy of the barcode, then we observed that accuracy is less than 60% and barcode of the face images is the same faces. So, all of these face image barcodes are not considered to be indexed.

Table 1 Statistics of the no. of key frames generated, no. of faces detected, and no. of barcode generated from some of the Hollywood movies

Video name	Size (KB)	No.of key frames	No. of faces detected	No. of valid barcodes
American Beauty—00170	1420	2	1	1
As Good As It Gets—01766	6150	18	5	5
Big Fish—00664	1640	6	1	1
Butterfly Effect, The—00696	1940	15	15	3
Casablanca—00250	628	7	0	0
Crying Game, The—01482	3210	18	1	1
Forrest Gump—01277	22,300	424	35	16
Gandhi—02262	18,000	18	5	4

Table 2 No. of generated key frames, no. of identified faces, and no. of valid barcode generated from some TV series video dataset episodes

Video name	Size (KB)	No. of key frames	No. of faces detected	No. of the valid barcode
24_ep2	848,000	4471	1970	1970
Breaking_bad_ep1	1,130,000	2473	567	567
How_I_Met_Your_Mother_ep1	435,000	553	532	532
Mad_men_ep1	980,000	3092	1983	1983
Modern_Family_ep1	465,000	2356	1415	1415
Sons_of_Anarchy_ep1	1,110,000	3964	1648	1648

Table 3 Key frame no. statistics generated, faces no. detected, and barcode no. generated from some TV series video database video files

Video name	Size (KB)	No. of key frame detected	No. of faces detected	No. of valid barcode generated
Aligned_video_0	556	84	84	84
Aligned_video_5	819	166	166	166
Aligned_video_3	1171	173	173	173
Aligned_video_1	1942	307	307	307
Aligned_video_2	3158	468	468	468
Aligned_video_4	664	119	119	119

Result of barcode generated from the TV series video dataset is shown in Table 2. All the videos of these dataset are in MP4 format. Here, no of barcode considered for each face detected from key frames. If considered the no of the barcode of same face images, then it will be reduced.

Table 3 represents the result of our work for YouTube face video face dataset.

5 Conclusions

The paper introduces a novel approach that uses straight EAN-8 scanner tags to index video through face recognition. An EAN-8 linear barcode for face representation detected from the videos is used for this suggested method of video indexing through face recognition. In this technique, color histogram method is used for key frame detection; Viola-Jones object detector is used as face detector. Sliding window technique is used for image gradient calculation, and EAN-8 barcode is used for indexing the face as a barcode. In this method, facial expressions, little change of face direction, illumination variance, and occlusion are also considered during EAN-8 barcode generation from human faces of videos. In the case of illumination, invariance face images, and significant changes of poses of face images, the problem will occur

during stable barcode generation. It is a space and time efficient method. In this technique, the main problem is making linear barcode from illumination invariant face image and angular face images. The next examination will endeavor to deal with these issues. For an account, this video indexing system can be used to identify verification, affirmation, and person quest.

References

1. S. Eickeler, F. Wallhoff, U. Iurgel, G. Rigoll, Content-based indexing of images and video using face detection and recognition methods, in *ICASSP 2001*, IEEE Xplore
2. L. Baraldi, C. Grana, R. Cucchiara, Neural story: an interactive multimedia system for Video indexing and re-use, in *Proceedings of CBIM*, Florence, Italy, June 19–21 (2017)
3. C.G.M. Snoek, M. Worring, Multimodal video indexing: a review of state of the art. *Multimedia Tools Appl.* **25**, 5–35 (2005)
4. Y. Wan, S. Ji, Y. Xie, X. Zhang, P. Xie, Video program clustering indexing based on faced recognition hybrid model of Hidden Markov model and support vector machine, in *IWCIA 2004*, LNCS 3322, pp. 739–749 (2004)
5. B.-C. Chen, Y.Y. Chen, Y.-H. Kuo, T.D. Ngo, D.-D. Le, S.I. Satoh, W.H. Hsu, Scalable face track retrieval in video archives using bag-of-faces sparse Representation. *IEEE Trans. Circ. Syst. Video Technol.* (2015)
6. Z. Dong, S. Jia, T. Wu, M. Pei, Face video retrieval via deep learning of binary hash representations, in *Proceeding of the Thirtieth AAAI Conference on Artificial Intelligence (AAAI-16)*
7. Y. Li, R. Wang, Z. Huang, S. Shan, X. Chen, Face video retrieval with image query via hashing across Euclidean space and Riemannian manifold, in *CVPR* (IEEE, 2015), pp. 4758–4767
8. Y.C. Chen, V.M. Patel, S. Shekhar, R. Chellappa, P.J. Phillips, Video-based face recognition via sparse joint representation, in *FG* (IEEE, 2013), pp. 1–8
9. U. Saeed, J.-L. Dugely, Temporally consistent keyframe selection from video for face recognition, in *18th European Signal Processing Conference*, 23–27 Aug 2010, IEEE Xplore, 30 April 2015
10. C. Czirjek, N. O'Connor, S. Marlow, N. Murphy, Face detection and clustering for video indexing applications, in *ACIVS 2003—Advanced Concepts for Intelligent Vision Systems*, 2–5 Sept 2003
11. M. Ravinder, T. Venu Gopal, T. Venkat Narayana Rao, Video indexing and retrieval applications and challenges. *Orient. J. Comput. Sci. Technol.* **3**(1), 125–137 (2010)
12. Y. Matveev, G. Kukharev, N. Shchegoleva, A simple method for generating facial barcodes, in *WSCG2014 Conference on Computer Graphics, Visualization and Computer Vision in Co-operation with EUROGRAPHICS Association Exchange Anisotropy* (Academic, Czech Republic, 2014), pp. 213–220.
13. S. Ghatak, Facial representation using linear barcode, in *Advanced Computational and Communication Paradigms*, vol. 2, pp. 791–801 (2018)
14. R. Zheng, C. Yao, H. Jin, L. Zhu, Q. Zhang, W. Deng, Parallel key frame extraction for surveillance video service in a Smart City. *PloS One* **10**(8), e0135694. <https://doi.org/10.1371/Journal.pone.0135694>
15. R. De Geest, E. Gavves, A. Ghodrati, Z. Li, C. Snoek, T. Tuytelaars, Online Action Detection. arXiv: 1604.06506v2[cs.CV] 30 Aug 2016
16. Barcode generation algorithm for EAN-8and EAN13. https://www.cherry-notes.spb.ru/barcode_ean8.htm. 15 March 2014

Acquisition and Analysis of Skin Impedance in Parkinson's Disease



Revati Shriram, Akshata Shinde, Radhika Nibhande, Anchal Guleria, and Rashmi Atre

Abstract Progressive loss of neuron from substantia nigra region of brain results into a neurodegenerative disease named Parkinson's disease. Parkinson's disease (PD) has lots of symptoms which are mainly divided into two main groups like: motor symptoms and non-motor symptoms. Symptoms of PD and its severity vary with each patient. But the mainly observed symptoms are limb/neck tremors and rigidity and are usually asymmetric; it affects one side of the body more than the other. Due to deterioration of the nigral dopamine neurons present in the brain that control muscles, progression in the PD symptoms is observed over the years (moving from stage 1 to stage 5). By the time motor symptoms are observed in PD patients, more than 70% dopamine generating neurons are lost from the substantia nigra brain region. There is no cure for this condition as once the neurons are lost, they cannot be regenerated again, with treatment focusing on medications, or a surgical intervention can reduce the symptoms of PD to certain extent. The work was carried out by the authors on non-motor Parkinson's disease symptom to study the changes in skin properties in PD patients by measuring skin impedance, which can be used for early detection of PD. The methodology was applied to sixteen normal and six PD real subjects of both the genders, between the age group of 15–86 years. It was observed that skin impedance in subjects with Parkinson's disease was much lower than the skin impedance observed in normal subjects.

Keywords Skin impedance · Magnitude · Phase · Gold electrodes · Silver electrodes · Parkinson's disease

1 Introduction

Our nervous system consists of basic unit named nerve cell or a neuron. Neurons can be of different types, such as motor neuron, which is responsible for the transmission of the signal from the brain to muscles through spinal cord to carry out

R. Shriram (✉) · A. Shinde · R. Nibhande · A. Guleria · R. Atre
Cummins College of Engineering for Women, Pune, India
e-mail: revatishriram@yahoo.com

contraction or relaxation. Progressive loss of neuron from substantia nigra region of brain results into a neurodegenerative disease named Parkinson's disease [1, 2]. Neurons from substantia nigra (mid-brain region) produce a chemical named dopamine, which is responsible for various motor functions. Balance between the amount of dopamine and acetylcholine (another chemical neurotransmitter produced in the brain) is required for the smooth/normal motor functioning. Due to loss of neurons from mid-brain (substantia nigra), the dopamine production also reduces, which leads to increase in the amount of acetylcholine produced. Due to imbalance between these two neurotransmitters, improper muscle contraction occurs which leads to the patient experiencing jerks or rigidity in the muscles. Parkinson's disease is a progressive and chronic disease which has affected around 10 million people worldwide. It is also called as movement disorder [3]. Various motor and non-motor symptoms are seen in PD patients. Onset of non-motor symptoms are generally during the early stages of PD, whereas motor symptoms occur during the later stages of PD.

2 Methods

It is observed that the skin tends to be excessively oily and sweaty in a subject, which changes the skin impedance of the patient [4–6]. The method proposed utilizes an electrode excited by a frequency signal [7] which is placed on the skin [8, 9]. The impedance measured by the electrode is further processed by an impedance analyzer system. The system directly gives out a value of the impedance measured in kilo ohm. This value is further used to evaluate the parameters of the PD patients and a healthy person. Analysis of acquired data was carried out using MATLAB, MINITAB and Excel Software.

Figure 1 shows the system block diagram for measurement and analysis of magnitude reading, phase reading and skin Impedance reading. Figure 2 shows the silver and gold electrodes (sensors) of 2.5 cm and 5 cm distance, fabricated by the authors for the acquisition of skin impedance.

2.1 Impedance Analyzer

The method proposed utilizes an electrode excited by a frequency signal which is placed on the skin [9]. The impedance measured by the electrode is further processed



Fig. 1 System block diagram of skin impedance measurement and analysis

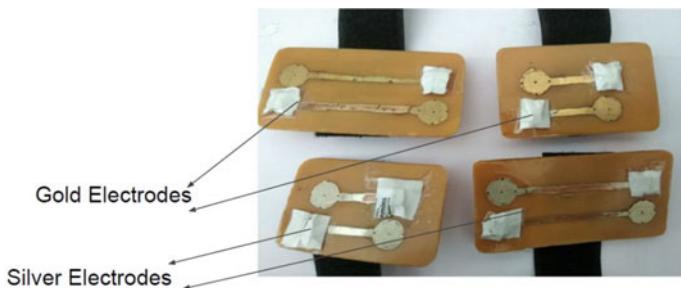


Fig. 2 Silver and gold electrodes designed for skin impedance measurement. Gold electrodes: 2.5 cm and 5 cm; silver electrodes: 2.5 cm and 5 cm

by an impedance analyzer circuit. Bioimpedance is measured by injecting a low-level sinusoidal current into the skin. The magnitude and phase are measured using the real and imaginary values obtained from the impedance analyzer board. Pmod IA is the impedance analyzer used for measurement of impedance by using silver and gold electrodes designed by the authors. Measurement range of analyzer is 100Ω to $10 M\Omega$. Calibration must be carried out for this analyzer before acquisition of skin impedance.

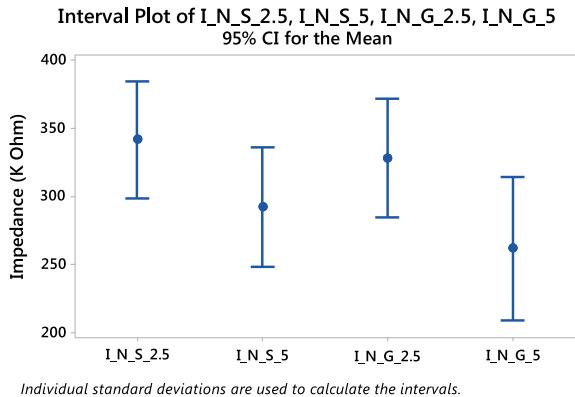
3 Database Collection Details

Data collection for normal subjects was carried out at Cummins College of Engineering for Women, Pune. Data collection of Parkinson's subjects was carried out with the help of members of Parkinson's Mitra Mandal (Parkinson's Support Group, Pune). We are incredibly grateful to the members of Parkinson's Mitra Mandal support group. It would not have been possible without the help of them. Database consists of total 22 real subjects (16 normal subjects and six PD subjects) of both the genders aged between 15 and 86 years, with the mean age of around 45 years. Verbal consent was given by all the normal and PD subjects for data acquisition and its use in the research. Modified Unified Parkinson's Disease Rating Scale (UPDRS) form was developed by the authors, and PD subjects/caretakers were requested to fill up that form for the future records, symptom variation analysis and study.

4 Results

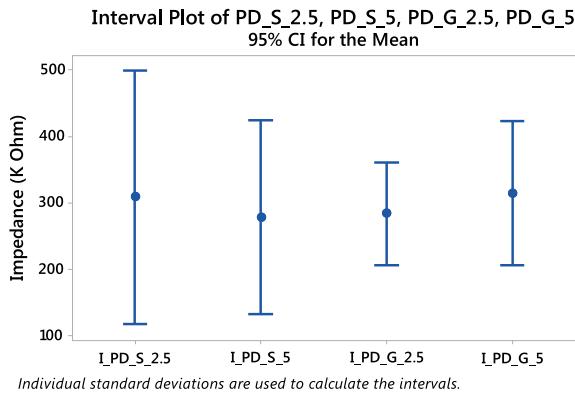
Skin impedance was recorded using total four electrodes (sensors), silver 2.5 cm, silver 5 cm, gold 2.5 cm and gold 5 cm. Study was carried out for the measurement of skin impedance and to check suitability of the electrode material and to study effect of distance between the two electrodes. Interval plot of skin impedance for

Fig. 3 Interval plot of impedance for normal subjects using silver and gold electrodes



Individual standard deviations are used to calculate the intervals.

Fig. 4 Interval plot of impedance for Parkinson's disease subjects using silver and gold electrodes



Individual standard deviations are used to calculate the intervals.

normal and Parkinson's subject is as shown in Fig. 3 and 4. Nomenclature use in the interval plot is: I—impedance, N—normal subjects, S—silver electrodes, 2.5 or 5—distance between the electrodes, G—gold electrodes and PD—subjects with Parkinson's disease. MINITAB 17 was used to obtain the internal plot. Figure 5 shows the box plot of skin impedance measured using two types of silver and gold electrodes each for the normal subjects. Figure 6 shows the box plot of skin impedance measured using two types of silver and gold electrodes each for the PD subjects.

5 Discussion

Lot of research related to Parkinson's disease is going on by the researchers from medical, engineering, biochemistry and pharmacology field. But, still the main cause of Parkinson's disease is not yet clear. Due to this, there is no one specific test to decide or predict PD. Generally, patients' history, visual examination (tremors, pin rolling movement, stoned face) and any other symptoms felt are discussed by the trained

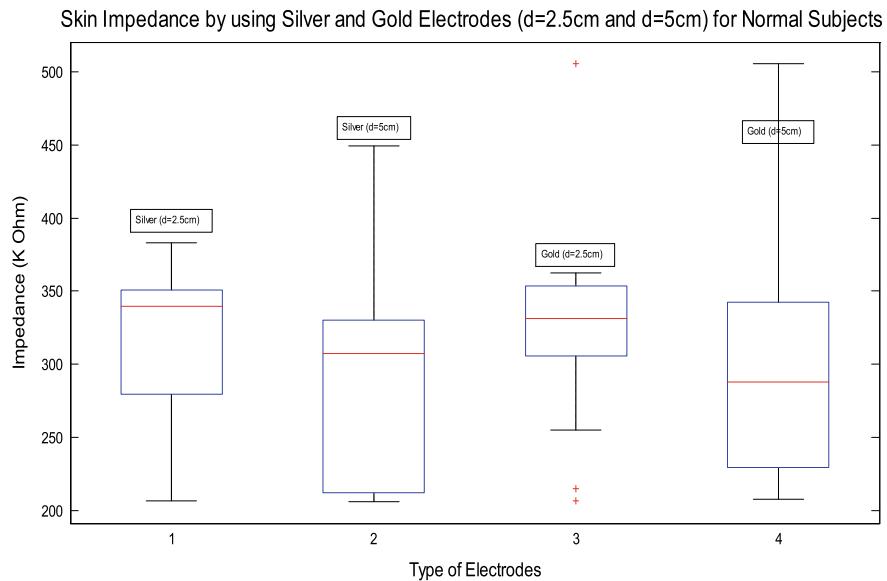


Fig. 5 Skin impedance measured using silver and gold electrodes for normal subjects

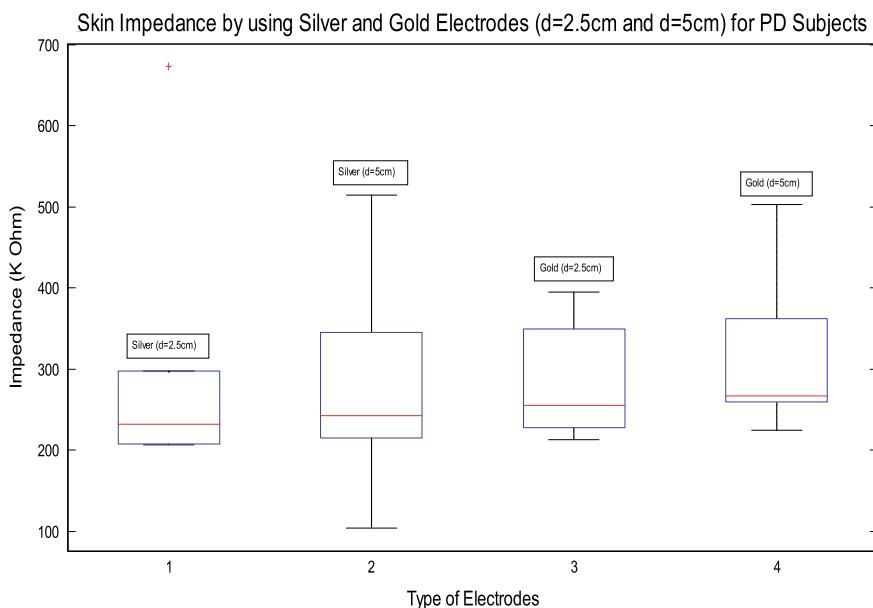


Fig. 6 Skin impedance measured using silver and gold electrodes for Parkinson's disease subjects

neuro-physicians before coming to the conclusion of having PD. Unified Parkinson's Disease Rating Scale (UPDRS) is another questionnaire-based diagnostics test which helps in deciding the PD and its stages. Blood testing or some imaging methods like MRI/fMRI can also be carried out to rule out other diseases with similar symptoms [10]. All these are the invasive or specialized techniques which are bit expensive and require specialized setup. Retinal changes can also be studied for the early prediction of PD. Another testing method is by prescribing carbidopa-levodopa in specific dosage, and feedback is taken from a patient on the severity of symptoms after the medication for dosage adjustment. Parkinson's disease is confirmed if an incredibly significant improvement is observed in the patient. The degeneration of the neurons from mid-brain cannot be stopped, but the rate at which the degeneration is occurring can be controlled. The main strategy is to increase the dopamine signaling in the brain. The chemical dopamine cannot cross the blood–brain barrier, but its precursor levodopa can. Amount of levodopa can be increased through medication. Another treatment available for PD is deep brain stimulation which involves implanting of a device which sends electrical signals directly to the brain [11].

6 Conclusion

The PCB had copper layer on which the design is decided. The type of metal in contact with skin had to be effective enough to transmit the proper impedance values to the designed system. The metals that were to be in contact with the skin had to be chemically inert. The metals under consideration were silver and gold. Gold and silver are good conductors of the electric signal and are non-reactive. Thus, the metal structure in the PCB was coated with gold and silver. The area of contact had to be considerate enough to cover the skin area. After analyzing the results for different surface areas, it was found that circular area with diameter of 1 cm was good enough to satisfy the desired criteria and gives a good signal. It was observed that the material (silver or gold) of electrode does not affect much on the values of the skin impedance recorded. At the same time, distance between the two electrodes also does not affect the impedance reading substantially. So, silver electrode of 2.5 cm can be used as a best-suited electrode for further skin impedance acquisition. It was observed that the impedance acquired for normal subject was between the range of 290–340 Ω , whereas impedance recorded for the Parkinson's disease subjects was between the range of 210–220 Ω which shows that skin impedance for PD subjects is lesser than skin impedance for the normal subjects. This finding of authors is in line with the research available till date. So, skin impedance measurement can also be used as one of the features in diagnosis of PD, though we need to mark that this is not the only feature. This feature can be used along with other 4–5 features in the diagnosis of PD.

References

1. M. Bhat, S. Inamdar, D. Kulkarni, G. Kulkarni, R. Shriram, Parkinson's disease prediction based on hand tremor analysis, in *IEEE International Conference on Communication and Signal Processing ICCSP 2017*, 6–8 April 2017, Chennai, Tamil Nadu, India
2. A. Shinde, R. Atre, A.S. Guleria, R. Nibandhe, R. Shriram, Facial features based prediction of Parkinson's disease, in *IEEE International Conference for Convergence of Technology (I2CT)*, 7–8 April 2018, Pune
3. S. Kulkarni, N. George, J. James, S. Parsewar, R. Shriram, Detection of Parkinson's disease through smell signatures, in *IEEE International Conference on Communication and Signal Processing ICCSP'20*, 9–11 April 2020, Chennai, Tamil Nadu, India
4. Y. Harth, D. Lischinsky, A novel method for real-time skin impedance measurement during radiofrequency skin tightening treatments. *J. Cosmetic Dermatol.* **10**, 24–29 (2010)
5. S.W. Baik, Y.J. Kim, J.H. Kim, W.Y. Jang, S.S. Kim, G.C. Park, J.M. Son, G.R. Jeon, Implementation of bioelectric impedance measurement system using multi-frequency applying method and two-electrode method, in *International Conference on Chemistry, Biomedical and Environment Engineering (ICCBEE'14)* 7–8 Oct 2014, Antalya, Turkey
6. S. Bjorklund, T. Ruzgas, A. Nowacka, I. Dahi, D. Topgaard, E. Sparr, J. Engblom, Skin membrane electrical impedance properties under the influence of a varying water gradient. *Biophys. J.* **104**, 2639–2650 (2013)
7. M. Frączek, T. Kręcicki, Z. Moroń, A. Krzywańska, J. Ociepka, Z. Rucki, Z. Szczepanik, Measurements of electrical impedance of biomedical objects. *Acta Bioeng. Biomedch.* **18**(1). <https://doi.org/10.5277/Abb-00294-2015-03>
8. F. Lu, C. Wang, R. Zhao, L. Du, Z. Fang, X. Guo, Z. Zhao, Review of stratum corneum impedance measurement in non-invasive penetration application. *Biosensors* **8**(31) (2018)
9. B.J. Nordbotten, C. Tronstad, Ø.G. Martinsen, S. Grimnes, Estimation of skin conductance at low frequencies using measurements at higher frequencies for EDA applications. *Physiol. Meas.* **35**, 1011–1018 (2014). <https://doi.org/10.1088/0967-3334/35/6/1011>
10. N. George, S. Kulkarni, J. James, S. Parsewar, R. Shriram, Detection of Parkinson's disease through speech and smell signatures, in *2nd International Conference on Advances in Electrical and Computer Technologies (ICAECT 2020)*, 24–25 April 2020, Coimbatore, India
11. J. James, N. George, S. Kulkarni, S. Pasewar, R. Shriram, M. Bhat, Detection of Parkinson's disease through speech signatures. Springer Book Chapter in AISC Series. ISBN No 978-981-15-1479-1

Automation of Soil Nutrient Measurement System and Irrigation Control



**U. B. Mahadevaswamy, R. Pavan Nayak, M. N. Darshan,
Tallam Vineeth Kumar, and S. Gautham Gopi**

Abstract The soil nutrients nitrogen, phosphorous and potassium (NPK) are primary and vital components for an agricultural land to be fertile. The excess or under usage of fertilizers will hamper the characteristics of the soil. The improper selection of manure would reduce the yield of the crop, resulting in loss of the yield. The proposed work aims at developing a sensor using the principle of optical transducer and hence determining NPK nutrients in the solution using Beer–Lamberts law. Also, under or over irrigation would lead to loss of the crop. The proposed method also aims at providing irrigation monitoring using a soil moisture sensor which determines water content in the soil using principle of conductivity. The system is developed as a three-tier application having sensors as data acquisition layer. The wireless communication is using Bluetooth as data link/data communication layer. Finally, Android mobile application is the presentation layer. The work aims at providing best-suited fertilizer along with optimal amount of fertilizer that gives high yield. In comparison with existing methods in same domain, the proposed work is economical and time efficient.

Keywords NPK · Beer–Lamberts law · Three-tier application · Optical transducer · Bluetooth

1 Introduction

Agriculture is the primary source of income for any country. Major part in economy of the country is derived from it. According to recent survey, it was found that one of the main reasons for failure of the crop was improper selection of fertilizer. Also, sometimes excess or inadequate usage of fertilizer would result in crop failure. Crops can grow effectively if its major nutrients are in adequate quantity. Primary nutrients of the soil are nitrogen (N), phosphorus (P) and potassium (K). Various crops need

U. B. Mahadevaswamy · R. Pavan Nayak (✉) · M. N. Darshan · T. V. Kumar · S. Gautham Gopi
Department of Electronics and Communication, JSS Science and Technology University, Mysore,
India
e-mail: pavnayak22@gmail.com

different amounts and different compounds consisting of these elements for effective growth of crop.

Soil nutrients NPK are calculated using chemical methods such as titration. These chemical processes involve dividing samples into parts, purifying, drying and quantitative analysis for NPK. This process is highly time consuming and takes nearly 10 days for obtaining the results. It also is not economical in farmer's point of view as cost of soil testing is more. Hence, many farmers often do not tend to test their soil and apply the best fertilizers. Instead they apply some inappropriate fertilizer and end up loosing the yield. Another issue associated with fertilizer is the amount of fertilizer to be used, and over usage of fertilizer would change the composition of soil and affects the crop growth. Also under usage of fertilizer would not provide required amount of nutrients for the crop growth. The proposed system would predict the best-suited fertilizer along with optical amount to be used for best yield.

Another issue that farmers face is the over and under irrigation which would lead to loss of the crop. The proposed method makes use of conductivity-based sensor to monitor the water quantity in the soil. The measured value is constantly compared with a threshold, which corresponds to an optimal value of water level that must be maintained in farm. If value is less than threshold, then farmer is informed to enable the irrigation system for the farm. Also, most of the present-day automation systems use methods that require constant Internet connectivity. The data communication requires usage of the servers and hence needs active Internet connectivity. This causes an economic constraint for farmers. The undertaken work provides a handheld device which works without Internet. The Bluetooth module used in the system uses wireless data communication to transfer data between sensors and android mobile application. The proposed system aims at providing cost effective and time saving complete agricultural automation.

2 Previous Research

This section comprises details regarding the existing methodologies in the domain of the proposed system. Various research papers in the field of agriculture automation, soil nutrient measurement techniques and irrigation control systems are referred, analyzed and overview for each is given below.

The automation of the irrigation procedure using a centralized control device is explained in paper [1]. This work emphasizes on the usage of discrete nodes that collect the information about the water levels present in different parts. The centralized controller decides the necessity of the irrigation and actuates accordingly. Similar irrigation automation procedure can be done with the help of IOT and sensor networks [2]. This method is useful in supplying optimal amount of water to the farm. The large amount of data is gathered and computed using big data and machine learning technologies. Another automation procedure is described using Raspberry PI [3] as controller and IOT nodes. Precision algorithm along with cloud computing is used for communication.

Soil nutrients nitrogen (N), phosphorus (P) and potassium (K) levels can be predicted by colorimetry and pH [4]. The work uses color sensor (TCS3471) along with light-dependent resistor to determine the levels or NPK in the soil. Statistical tool like chi-squared classification is equipped to find the soil nutrient levels. The idea about quantifying the NPK in optical domain is demonstrated in paper [5]. The optical transducer involves conversion of signals from electrical to optical and vice versa. Hence by gaging the absorption in the optical domain, the values of NPK are quantified. The prediction of NPK nutrients depending upon the location [6] is performed using random forest (RF) algorithm. This work involves obtaining the geographical location of farmer and by using data classifiers and RF algorithm the NPK nutrients and hence crop are predicted.

An efficient method for crop management, agricultural finance management and web-based services for agriculture using advanced technologies such as IoT, big data computation and cloud computing is described in paper [7]. A diversified network connected using IoT nodes is used for data handling. Suitable classifiers with data processing are specified for effective irrigation. The work also uses drone imaging for controlling pest attacks. Soil analysis and fertility can be predicted using pH, and suitable fertilizer can be applied. This method is explained in paper [8] where authors claim that its time effective solution requires no human intervention. Authors of [9] conduct a comparative study of various advanced technologies like maximum livelihood classification (MLC), support vector machine (SVM), neural networks for plant area extraction, crop development analysis and estimation of yield.

3 Materials and Methods

This section explains the methodology and working procedure of the proposed work. Block diagram and flowcharts are incorporated in this section to explain the working principles.

3.1 *Block Diagram and Working*

This section explains the overall working procedure of the proposed work with the help of a block diagram. The proposed work can be divided into two divisions based on the functionality. First being the soil nutrients measurement system and latter being irrigation control system. Though both are individual procedures, but they work parallelly (distinguished machine cycle of microcontroller) and provide results to Bluetooth module which collectively sends the values using frequency interleaving to the mobile application. Each of these procedures has hardware and software blocks for implementing their respective functionalities. Overall working is shown in Fig. 1.

The optical transducer is fed with three wavelengths at which three nutrients (NPK) have maximum absorption. The initial and final intensities before and after

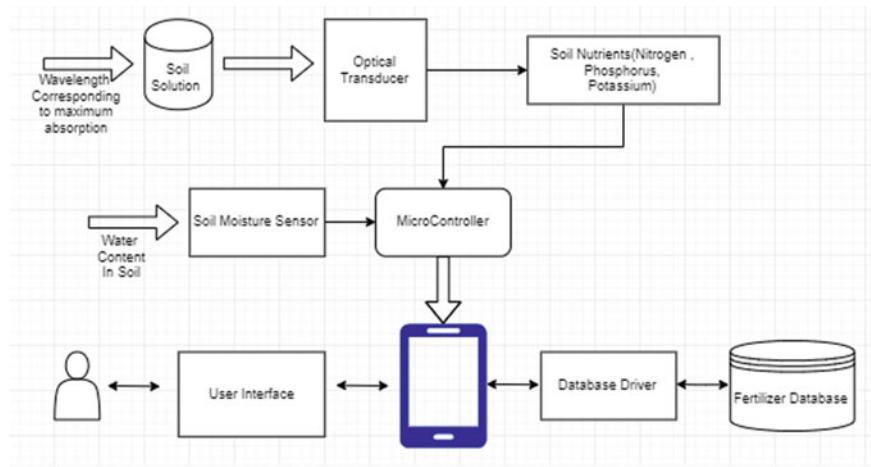


Fig. 1 Block diagram of the proposed work

absorption are measured by LED and LDR, respectively. The difference of intensities is computed by measuring difference of voltages across LED and LDR. Hence by gaging the difference in the intensities, amount of light absorbed is calculated. Absorbed intensity is applied to Beer-Lamberts law to determine the concentration of the nutrients in the solution. Further data processing is done by microcontroller. The data are compared with respect to two different threshold levels and classified into respective levels (high, medium, low). The values of absorbance less than threshold 1 are assigned as low. Values between first and second threshold are assigned as medium. The absorbance more than both the thresholds is assigned as high. The work uses dynamic thresholding algorithm to set thresholds. The thresholding levels are decided at run-time based on type of parameter measured. These values are passed to Android application via Bluetooth module for further processing.

Latter functionality of the proposed work is the irrigation control by measuring the soil moisture. This process incorporates soil moisture sensor to measure amount of water content in the soil. The soil moisture sensor works on the principle of conductivity. This sensor is inserted into the soil where soil moisture level is to be measured. Depending on water content, the resistance and hence voltage drop across the sensor change. Hence, amount of water content is determined using voltage drop across sensor. The measured value is passed to microcontroller, and after processing it is communicated to mobile app via Bluetooth. The measured value is compared with a threshold, and if found its less, then farmer is notified to turn on his irrigation system. The threshold value is optimal amount of water needed for healthy growth of plant.

3.2 Formulae Used

Formula to calculate amount of light intensity is given using Beer–Lambers law,

$$A_{\text{NPK}} = -\log\left(\frac{I_1}{I_0}\right) \quad (1)$$

where

I_1 Absorbed intensity of light by soil nutrient

I_0 Initial intensity of light transmitted.

Modified version of above formula (1) is used for computation. Thousand twenty-four is maximum value since a 10-bit ADC is used to represent the value of transmitted light intensity given in Eq. (2).

$$A_{\text{NPK}} = \log_{10}\left(\frac{1024}{I_1}\right) = 3 - \log([I_1]) \quad (2)$$

$$A_{\text{NPK}} = -\log\left(\frac{V_1}{V_0}\right) \quad (3)$$

where

V_1 is voltage across LDR, and V_0 is voltage across LED.

Equation (3) is a modified version of Eq. (1), since ratio of voltages across LDR and LED is directly proportional to ratio of intensities at LDR and LED, respectively.

3.3 Implementation

This section explains the development of system for the proposed work. Work was divided into two phases, first being the hardware section consisting of the nutrient measurement system and soil moisture sensor and latter being the software section which serves as the UI and used for database and fertilizer calculations-related computations.

Figure 2 picturizes the implementation of the proposed work. The large covered part encloses optical transducer consisting of three LEDs and one LDR. The space between LED and LDR is occupied by the testtube which contains the soil solution. LED and LDR are connected to the input and output analog pins of the microcontroller, respectively. Three LEDs of color red, green and blue (RGB) are used which correspond to maximum wavelength of NPK nutrients, respectively. The LEDs are sequentially triggered and absorbance corresponding to each of the nutrient is measured using Beer–Lambers law. Meanwhile in parallel machine cycle, soil moisture sensor measures the water content of the soil. Bluetooth module sets up



Fig. 2 Hardware implementation

a TCP-IP link between hardware and mobile having android application. Both the data are transferred, and respective functions are performed in the mobile application. Android mobile application was developed in order to visualize the results. Functionality can be divided into two parts: fertilizer recommendation and irrigation control system. The fertilizer recommendation system takes NPK values from sensor. Farmer is provided with a dropdown to select the crop. The measured NPK values are compared with a standard fertilizer database to give best-suited fertilizer with optimal amount. In irrigation control procedure, the measured value is compared with a threshold (set as 50% of maximum value). If measured soil moisture falls below the threshold, farmer is notified to turn on his irrigation system.

4 Results and Conclusion

The developed sensor was tested for ten samples from different places in Karnataka, India. The variation in absorption of soil nutrients NPK was noted and analyzed. The results were validated by comparison with results with standard laboratories and developed sensor. Irrigation monitoring was performed for a month, variation of soil moisture levels was noted, and irrigation pattern was analyzed. The graphical plot of results is shown in Fig. 3.

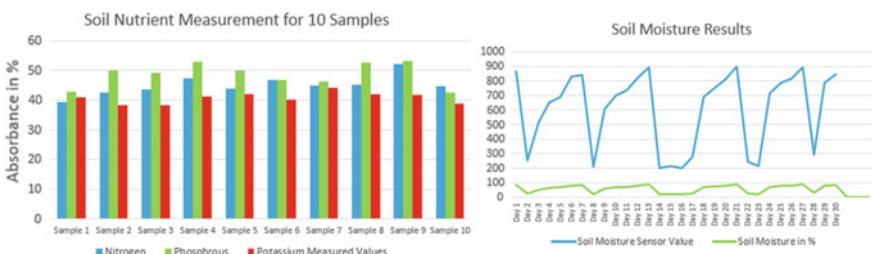


Fig. 3 Measured NPK and soil moisture values

4.1 Validation of Sensor

The developed sensor was tested and validated by comparison with soil nutrient values from standard soil testing laboratories. The variation pattern of each nutrient (NPK) was observed and analyzed. Various statistical tools like standard deviation and T-test were used to measure the accuracy of the sensor. T-test was used to determine the variation of standard deviation between two independent data sets (optical and chemical). The equation used to find standard deviation using T-test at 95% confidence interval is given in Eqs. (4) and (5).

$$t = \frac{M_1 - M_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} \quad (4)$$

$$S = \sqrt{\frac{\sum(x - M)^2}{n - 1}} \quad (5)$$

where

M is mean of individual data set

S is standard deviation of individual data set

n is number of samples of data set.

The graphical plot of values measured by the sensor and the laboratory results is showed in Fig. 4 for nitrogen, phosphorus and potassium, respectively.

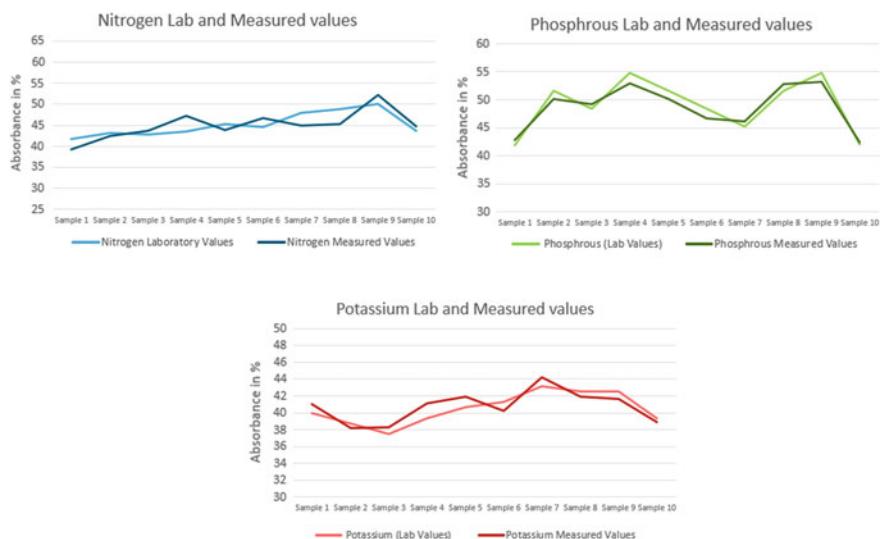


Fig. 4 Laboratory and measured values comparison

Table 1 T-test results of NPK

Nutrient	Nitrogen	Phosphorus	Potassium
T-test standard deviation	1.3862	1.963	0.8314

The T-test results for NPK nutrients are shown in Table 1. T-test was computed from Eq. (4) and (5), and standard deviations for each of the soil nutrients were calculated. Net standard deviation was calculated from individual standard deviations as shown in Eq. (6).

$$\text{Net SD} = \frac{S_N + S_P + S_K}{3} = \frac{1.3862 + 1.963 + 0.8314}{3} = 1.39 \quad (6)$$

Error in sensor exists only when there is transition between high-medium or medium-low. Relative error corresponding to this variation is 0.2 because a value + 20% or -20% is added or subtracted in fertilizer quantity.

4.2 Novelty

In agricultural sector, there is no existing complete automation method that provides both nutrient measurement system and irrigation monitoring. Proposed work provides on-spot measurement of soil nutrient values along with fertilizer recommendation and optimal amount. The developed system is handheld, portable and simple to use. The Android app does not require Internet connection since it works locally on Bluetooth wireless communication. The developed nutrient measurement system is highly cost effective and one-time deployment.

5 Conclusion

The proposed method and existing system is compared with respect to time and cost parameter, and results are provided in Table 2. We can conclude that the proposed method is time saving and cost effective when compared to the existing systems in the industry.

Table 2 Results comparison table

Parameter	Existing systems	Proposed system
Time	12–14 days for testing	Less than a minute
Cost	₹800 for one sample ^a	One-time deployment costing around ₹600

^aValues from the existing systems used in the industry and laboratories for soil testing

References

1. J. Muangprathub, N. Boonnam, S. Kajornkasirat, N. Lekbangpong, A. Wanichsombat, P. Nillaor, IoT and agriculture data analysis for smart farm. *Comput. Electron. Agric.* **156**, 467–474 (2019)
2. A. Goap, C. Rama Krishna, D. Sharma, A.K. Shukla, An IoT based smart irrigation management system using Machine learning and open source technologies. *Comput. Electron. Agric.* **155**, 41–49 (2018)
3. R.N. Rao, B. Sridhar, IoT based smart crop-field monitoring and automation irrigation system, in *2018 2nd International Conference on Inventive Systems and Control (ICISC)*, Coimbatore, pp. 478–483 (2018). <https://doi.org/10.1109/ICISC.2018.8399118>
4. R.G. Regalado, J.C. Dela Cruz, Soil pH and nutrient (Nitrogen, Phosphorus and Potassium) analyzer using colorimetry, in *2016 IEEE Region 10 Conference (TENCON)*, Singapore, pp. 2387–2391 (2016). <https://doi.org/10.1109/TENCON.2016.7848458>
5. M. Masrie, M.S.A. Rosman, R. Sam, Z. Janin, Detection of nitrogen, phosphorus, and potassium (NPK) nutrients of soil using optical transducer, in *2017 IEEE 4th International Conference on Smart Instrumentation, Measurement and Application (ICSIMA)*, Putrajaya, pp. 1–4 (2017). <https://doi.org/10.1109/ICSIMA.2017.8312001>
6. A.G. Mohapatra, B. Keswani, S.K. Lenka, Soil N-P-k prediction using location and crop specific random forest classification technique in precision agriculture. *Int. J. Adv. Res. Comput. Sci.* **8**(7) (2017)
7. P. Srinivasulu, M.S. Babu, R. Venkat, K. Rajesh, Cloud service oriented architecture (CSoA) for agriculture through internet of things (IoT) and big data, in *2017 IEEE International Conference on Electrical, Instrumentation and Communication Engineering (ICEICE)*, Karur (2017)
8. K. Abhang, S. Chaughule, P. Chavan, S. Ganjave, Soil analysis and crop fertility prediction. *Int. Res. J. Eng. Technol. (IRJET)* **05** (2018)
9. M. Wei, B. Qiao, J. Zhao, X. Zuo, Application of remote sensing technology in crop estimation, in *2018 IEEE 4th International Conference on Big Data Security on Cloud (BigDataSecurity), IEEE International Conference on High Performance and Smart Computing, (HPSC) and IEEE International Conference on Intelligent Data and Security (IDS)*, Omaha, NE, pp. 252–257 (2018). <https://doi.org/10.1109/BDS/HPSC/IDS18.2018.00061>

Performance Evaluation of Perovskite Solar Cells at Elevated Temperatures



Sumanshu Agarwal , Archana Kumari Munda, Vedika Pandey, and Kundan Kumar 

Abstract The last decade has witnessed enormous research in the field of perovskite solar cells, and state-of-the-art efficiency is approaching the performance limit. However, the effect of ambient temperature on the performance of these solar cells has not been studied thoroughly. Given that, here, we evaluate the performance of perovskite ($(\text{MA}_{0.13}\text{FA}_{0.87})\text{PbI}_3$) solar cells as a function of temperature in the range of 0–50 °C. We perform self-consistent temperature-dependent drift–diffusion simulations for the same. It is interesting to observe that despite an increase in the bandgap with the temperature, the open-circuit voltage reduces. Moreover, the open-circuit voltage reduces linearly at the rate of approximately 0.4 mV/°C. The result is contrasting to Si solar cells, where the rate of reduction of open-circuit potential is 2 mV/°C.

Keywords Modeling · Device physics · Open-circuit voltage

1 Introduction

Solar photovoltaic (PV) systems are widely used to generate electricity and thereby reducing the carbon footprint [1]. Increased demand for photovoltaics has also invited researchers to look for alternative PV technologies that are efficient yet cost effective and easy to fabricate [2–4]. In that context, exponentially growing research in the field of perovskite solar cells has shown promising results from high efficiency [5] to roll-to-roll fabrication [6]. The technology is considered to have the potential to capture a significant market share of solar cells in the future [7]. Therefore, it is vital to understand the effect of environmental factors such as the clearness index of the sky [8], air mass [9], temperature [10], and humidity [11] on the performance of this technology. Out of these, the first two items affect the performance by reducing

S. Agarwal  · A. K. Munda · V. Pandey · K. Kumar

Department of Electronics and Communication Engineering, Institute of Technical Education and Research, Siksha ‘O’ Anusandhan (Deemed to be University), Bhubaneswar 751030, Odisha, India

e-mail: sumanshuagarwal@soa.ac.in

solar insolation on the device, and hence, changes the short-circuit current (J_{SC}) [8]. Therefore, the effect of these two factors can readily be evaluated by parameterizing the illumination intensity. However, a thorough study is required for the factors that affect the performance upon variation in the ambient conditions like temperature and humidity.

Given that, there exist numerous reports that explore the phase transition of perovskite (used in PV) when exposed to a humid environment [11, 12]. Withal, there are only a few reports that discuss the effect of temperature on the material/device characteristics of perovskite solar cells [13, 14]. Moreover, many of them fail to conclusively address the effect of temperature on the performance metrics of perovskite solar cells. It is worth to mention here that the literature suggests that variation in temperature changes the dynamics of the charge carriers in semiconductor devices. Accordingly, injection current and hence the open-circuit voltage (V_{OC}) mainly get affected by the change in the temperature, which is indeed true for Si solar cells [10]. However, temperature-dependent phase transformation in hybrid perovskite materials (used for solar cells fabrications) has different implications like an increase in the bandgap with the temperature [13] (which is opposite to the case of Si). Because the technology is yet to be matured, it is essential to understand the material parameters that are crucial for the technology to blossom in the market.

In that context, the present work intends to explore the functional role of various parameters on the performance of perovskite solar cells if the operating temperature is varied. Specifically, this paper identifies the effect of temperature-induced bandgap variation of perovskite on V_{OC} of the device in the temperature range of 0–50 °C. For the study, we consider $(\text{MA}_{0.13}\text{FA}_{0.87})\text{PbI}_3$ [13] as perovskite material or the active layer in the solar cell (see Fig. 1 for the schematic of the device). Our simulation results indicate that—(a) saturation current density in dark increases in the device,

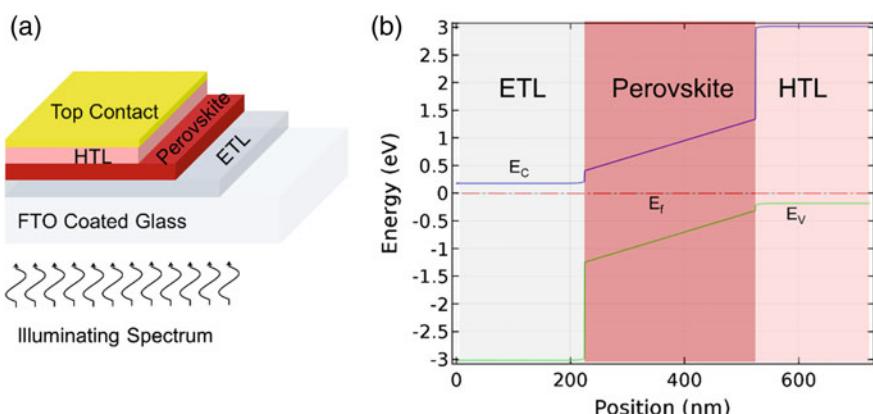


Fig. 1 Schematic and band diagram of a perovskite solar cell under discussion. **a** Schematic of NIP design of perovskite solar cell that is being investigated in the work. Equilibrium energy band diagram of the cell at 30 °C is shown in **(b)**. Here, E_C indicates minimum level in the conduction band, E_F indicates fermi energy level, and E_V indicates maximum energy level of valence band

(b) ideality factor is almost unchanged, and (c) V_{OC} decreases with an increase in the temperature in the range under discussion. It is interesting to note that while the bandgap of the active layer increases with the temperature (which is in contrast to Si), V_{OC} reduces with the temperature, a result in line with Si solar cells. However, the reported value for the rate of change of V_{OC} with temperature for Si is $\sim 2 \text{ mV/}^{\circ}\text{C}$ [15], while the same is $\sim 0.4 \text{ mV/}^{\circ}\text{C}$ for the perovskite solar cell under investigation. The future work will include the temperature dependency of other material parameters, for example, dielectric constant, mobility, interface kinetics, contact injection, etc., in the model.

2 Model System

The schematic of the perovskite solar cell being studied in the paper is shown in Fig. 1. Here, $(\text{MA}_{0.13}\text{FA}_{0.87})\text{PbI}_3$ is the perovskite material, which acts as the active layer in the solar cell. We have considered TiO_2 and spiro-MeOTAD as the electron transport layer (ETL) and hole transport layer (HTL), respectively. The material properties are adapted from the literature [16]. In addition, we assume a temperature-dependent bandgap of perovskite as a linear function between 1.62 eV at 0 °C and 1.68 eV at 50 °C [13]. We take a doping density of 10^{17} cm^{-3} in ETL (n-type) and HTL (p-type). The contacts are taken as ohmic. Using these parameters and the following system of equations, we perform self-consistent simulations in COMSOL multiphysics.

Continuity equation: The continuity equation is based on the law of conservation of charges. Accordingly, change in the current (J) is governed by generation (G) or annihilation/recombination (R) of the charge carriers. Therefore,

$$J_\xi = q \left(-D \frac{dn_\xi}{dx} \pm \mu_\xi n_\xi E \right), \quad (1)$$

$$\frac{dJ_\xi}{dx} = q(G_\xi(x) - R_\xi(x)), \quad (2)$$

where q is the elementary charge, D is the diffusion coefficient, n is the charge carrier density, μ is the mobility of the charge carrier, E is the electric field, and x is the position. Here, ξ stands for electron and hole. Note that the positive sign with the second term on the right-hand side of Eq. 1 stands for holes and the negative sign for electrons.

Poisson's equation: We observe in Eq. 1 that charge carrier density, its gradient, and electric field altogether define the current density in the device. However, the electric field is dependent on the net charge density. The functional form is known as Poisson's equation and accordingly,

$$\frac{dE}{dx} = \frac{\rho(x)}{\epsilon \epsilon_0}, \quad (3)$$

where $\rho(x)$ is the density of charge carrier at point x , ε is the dielectric constant, and ε_0 is the permittivity of the free space.

Here, the temperature dependency has been crept in by (a) temperature-dependent bandgap, (b) thermal voltage ($V_{th} = kT/q$), and (c) temperature-dependent extinction coefficient (reflected in the generation rate). Because the literature indicates that the extinction coefficient of the active layer does not vary significantly, we assume that it is (and hence the photogeneration rate of the charge carriers) invariant with the temperature. The assumption is also helpful in understanding the functional dependency of V_{OC} on the temperature, independently to the generation rate. Further, a sufficiently large diffusion length as compared to the thickness of the active layer allows us to take a uniform generation rate in the active layer [17].

3 Results and Discussions

The simulated current–voltage characteristics at different temperatures are shown in Fig. 2. Figure 2a shows the characteristics in the dark. We observe the exponential characteristics of the nature $J = J_0 \left[\exp\left(\frac{qV}{\eta kT}\right) - 1 \right]$, where J_0 is the reverse saturation current, and η is the ideality factor. We find an increase in the saturation current with the temperature; however, a rather temperature-independent ideality factor. A further discussion on the same is available later in the paper.

Figure 2b shows the current–voltage characteristics under illumination. It is evident from the figure that J_{SC} is independent of temperature. The main reason behind that is the temperature-invariant assumption of the generation rate in the simulations. Furthermore, we do not observe any significant change in the equilibrium

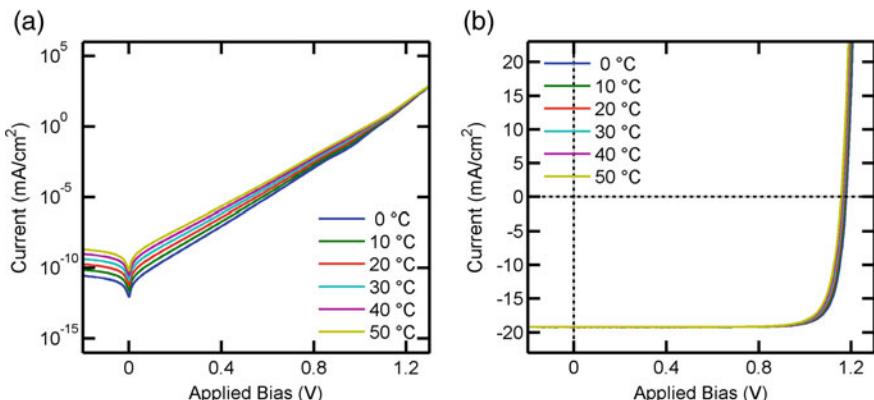


Fig. 2 Temperature-dependent current–voltage characteristics of the solar cell. **a** The dark characteristics. An increase in the saturation current density with the increase in the temperature is evident. **b** The characteristics under illumination. We observe a rather small decrease in V_{OC} with the temperature

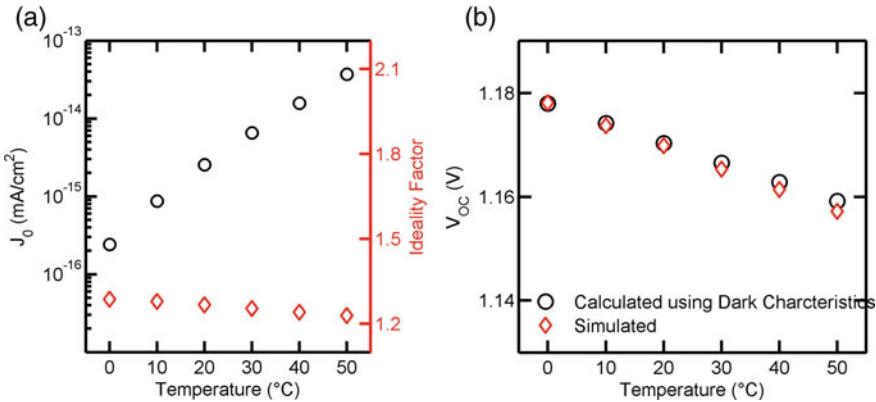


Fig. 3 Characterization parameters of the solar cell as a function of temperature. **a** Variation of J_0 and ideality factor with temperature. Exponential relation between J_0 and temperature and almost temperature-independent ideality factor is quite apparent. **b** V_{OC} as a function of temperature. Values calculated using J_0 and ideality factor corroborate well with the simulated results

band diagram; and hence, the electric field in the active layer (which is responsible for drifting out the photogenerated carriers at 0 V applied bias) is temperature-independent. Therefore, a constant J_{SC} with the temperature is observed. However, there is a reduction (though small) in V_{OC} with the temperature (see Fig. 2b and 3b). The change in V_{OC} in the temperature range from 0 to 50 °C is a mere 20 mV. For c-Si solar cells, a similar temperature range results in a change in the V_{OC} of the order of 100 mV [15]. This deviation in the characteristics in the two technologies is mainly due to an increase in the bandgap of the perovskite with the temperature, while it decreases in c-Si.

To further understand the (a) functional dependency of V_{OC} on the dark current and (b) charge carrier dynamics and device physics, we extract the temperature-dependent reverse saturation current and ideality factor from the dark characteristics of the device at 1.15 V applied bias (a value comparable to V_{OC}) [18]. The results are shown in Fig. 3a. We observe that while the ideality factor is almost constant (~1.2, indicates the presence of a mix of trap assisted and band-to-band recombination mechanisms), reverse saturation current density (J_0) follows an exponential relation with the temperature. Mathematically,

$$J_0 \propto \exp(T). \quad (4)$$

Accordingly, by using the relation for V_{OC} as [17]

$$V_{OC} = \frac{\eta k T}{q} \ln\left(\frac{J_{SC}}{J_0}\right), \quad (5)$$

we find that V_{OC} decreases with an increase in the temperature. Further, the calculated values for V_{OC} using Eq. 5 are in close agreement with the simulated values of V_{OC} (see Fig. 3b). It is also evident from Fig. 3b that V_{OC} decreases linearly with the temperature.

4 Conclusions

To summarize, we have performed self-consistent temperature-dependent drift-diffusion simulations for perovskite solar cells. The results indicate a nearly temperature-independent ideality factor and an increase in the reverse saturation current density with the temperature. Interestingly, despite an increase in the bandgap of the active layer with the temperature, we observe a decrease in V_{OC} in the range of 0 to 50 °C. We note that V_{OC} reduces by ~0.4 mV/°C in the above said temperature range, while the same rate for Si solar cells is ~2 mV/°C.

References

1. K. Seyboth, P. Matschoss, S. Kadner, T. Zwickel, P. Eickemeier, G. Hansen, S. Schlömer, C. von Stechow, Renewable Energy Sources and Climate Change Mitigation Special Report of the Intergovernmental Panel on Climate Change, New York (2012)
2. M. Li, K. Gao, X. Wan, Q. Zhang, B. Kan, R. Xia, F. Liu, X. Yang, H. Feng, W. Ni, Y. Wang, J. Peng, H. Zhang, Z. Liang, H.-L. Yip, X. Peng, Y. Cao, Y. Chen, Solution-processed organic tandem solar cells with power conversion efficiencies >12%. *Nat. Photonics* **11**, 85–90 (2017)
3. M.I. Ahmed, A. Habib, S.S. Javaid, Perovskite solar cells: potentials, challenges, and opportunities. *Int. J. Photoenergy* **2015**, 1–13 (2015)
4. Grätzel, M.: Photovoltaic and photoelectrochemical conversion of solar energy. *Philos. Trans. R. Soc. London A Math. Phys. Eng. Sci.* **365**, 993–1005 (2007)
5. X. Gao, W. Luo, Y. Zhang, R. Hu, B. Zhang, A. Züttel, Y. Feng, M.K. Nazeeruddin, Stable and high-efficiency methylammonium-free perovskite solar cells. *Adv. Mater.* **32**, 1905502 (2020)
6. D. Burkitt, R. Patidar, P. Greenwood, K. Hooper, J. McGettrick, S. Dimitrov, M. Colombo, V. Stoichkov, D. Richards, D. Beynon, M. Davies, T. Watson, Roll-to-roll slot-die coated P-I-N perovskite solar cells using acetonitrile based single step perovskite solvent system. *Sustain. Energy Fuels* **4**, 3340–3351 (2020)
7. M.A. Green, A. Ho-Baillie, H.J. Snaith, The emergence of perovskite solar cells. *Nat. Photonics* **8**, 506–514 (2014)
8. S. Nann, C. Riordan, Solar spectral irradiance under clear and cloudy skies: measurements and a semiempirical model. *J. Appl. Meteorol.* **30**, 447–462 (1991)
9. F. Kasten, A.T. Young, Revised optical air mass tables and approximation formula. *Appl. Opt.* **28**, 4735 (1989)
10. E. Radziemska, The effect of temperature on the power drop in crystalline silicon solar cells. *Renew. Energy* **28**, 1–12 (2003)
11. Z. Song, A. Abate, S.C. Wattage, G.K. Liyanage, A.B. Phillips, U. Steiner, M. Graetzel, M.J. Heben, Perovskite solar cell stability in humid air: partially reversible phase transitions in the $\text{PbI}_2\text{-CH}_3\text{NH}_3\text{I}\text{-H}_2\text{O}$ System. *Adv. Energy Mater.* **6**, 1600846 (2016)
12. Q. Tai, P. You, H. Sang, Z. Liu, C. Hu, H.L.W. Chan, F. Yan, Efficient and stable perovskite solar cells prepared in ambient air irrespective of the humidity. *Nat. Commun.* **7**, 11105 (2016)

13. Chen, H.-W., Gulo, D.P., Chao, Y.-C., Liu, H.-L.: Characterizing temperature-dependent optical properties of (MA0.13FA0.87) PbI₃ single crystals using spectroscopic ellipsometry. *Sci. Rep.* **9**, 18253 (2019)
14. I. Mesquita, L. Andrade, A. Mendes, Temperature impact on perovskite solar cells under operation. *ChemSusChem* **12**, 2186–2194 (2019)
15. J.J. Wysocki, P. Rappaport, Effect of temperature on photovoltaic solar energy conversion. *J. Appl. Phys.* **31**, 571–578 (1960)
16. S. Agarwal, P.R. Nair, Device engineering of perovskite solar cells to achieve near ideal efficiency. *Appl. Phys. Lett.* **107**, 123901 (2015)
17. S. Agarwal, P.R. Nair, Performance loss analysis and design space optimization of perovskite solar cells. *J. Appl. Phys.* **124**, 183101 (2018)
18. S. Agarwal, M. Seetharaman, N.K. Kumawat, A.S. Subbiah, S.K. Sarkar, D. Kabra, M.A.G. Namboothiry, P.R. Nair, On the uniqueness of ideality factor and voltage exponent of perovskite-based solar cells. *J. Phys. Chem. Lett.* **5**, 4115–4121 (2014)

Implementation of Transfer Learning Technique for the Detection of COVID-19



Amit Prakash Sen and Nirmal Kumar Rout

Abstract The present world is suffering from the severe attack of corona virus disease (COVID-19). With the view to minimize the spread of this deadly disease, testing and analysis of tremendous amounts of suspected cases for isolation of such individual and further treatment are a need. Pathogenic laboratory testing is the analytic best quality level; however, it is tedious with critical false negative outcomes. Fast and exact characteristic techniques are fundamentally expected to fight the deadly disease. Considering COVID-19 radio graphical changes in CT images, the paper centers to develop a deep learning model that could isolate COVID-19 cases in order to give a clinical end before the pathogenic test, along these lines sparing urgent time for affliction control.

Keywords COVID-19 · Deep CNN · Transfer learning · Medical images · Infection detection

1 Introduction

The COVID-19 has astonished the whole research community with its quick spread, possible destructiveness, with likely significant by and large effect on the lives of billions of individuals from both health hazard and financial viewpoint. First distinguished in 2019 in Wuhan, China, it has since gotten a worldwide pandemic. As per the present data of July 12, 2020, there are around 12,839,566 affirmed instances of which 567,574 death, a death rate of 4.42% [1]. One significant obstacle in controlling the spreading of this deadly disease is the incompetence and lack of tests. The present-day tests are generally centered on reverse transcription polymerase chain

A. P. Sen (✉) · N. K. Rout
School of Electronics, KIIT University, Bhubaneswar, India
e-mail: amittata.sen@gmail.com

N. K. Rout
e-mail: routnirmal@rediffmail.com

reaction (RT-PCR). It takes 4–6 h to get results, which is quite a long period in comparison with the fast spreading pace of COVID-19. The determination of COVID-19 depends on the accompanying measures such as clinical manifestations, epidemiological history, positive pathogenic testing and positive CT images. COVID-19 affects different person in several ways. In contrast with the early phase of the infection, presently the clinical qualities of COVID-19 change hugely from patient to patients which takes into account of respiratory effects, fever, cough, throat pain and pneumonia [2–5]. The less regular indications are a throbbing painfulness, sore throats, diarrhea, loss of taste or smell, migraine and rashes on skin or staining of fingers or toes while the genuine indications include problems in breathing or shortness in breathing, serious chest pain and loss of speech. However, these indications are indefinite, as constantly some new kind of cases are recognized.

Presently, many COVID positive patients were found with no clinical symptoms but having positive COVID-19 results. At the point when somebody is perceived as an individual under examination, at that point lower respiratory samples thereby accumulated for further pathogenic testing. This testing depends on reverse transcription polymerase chain reaction (RT-PCR) and sequencing of nucleic acid from the disease [6–8]. From the earliest starting point of the pandemics, the effectiveness of RT-PCR testing has been subject to different elements which includes the accessibility and availability of the testing packs in the influenced area. All the more significantly, the outcomes acquired through RT-PCR testing are flawed because of quality, constancy and reproducibility of the detection kits. Besides inefficiency, RT-PCR test kits are in huge shortage. This motivates us to focus on alternative testing manners, which are possibly faster, economical and more accessible than RT-PCR, but are as precise as RT-PCR.

Fang et al. [9] related the effectiveness of the non-contrast chest CT with RT-PCR which recognizes the viral nucleic acid and is considered as the present-day reference standard for the detection of this deadly disease. A total of 51 patients were considered having travel history or reside in prevalent areas and having the common respiratory indications of unidentified reason in the reported article. RT-PCR testing of such patients was done with repetition. Their gold standard was a possible affirmed conclusion of the pandemic disease by sequential RT-PCR testing. As per the author, a sensitivity of non-differentiate chest CT for finding of COVID-19 virus is about 98% as contrasted with early RT-PCR affectability of 71%. This deficiency of sensitivity on initial RT-PCR testing is also reported by Xie et al. [10] in their work. As per the reported article, author claims that 3% of 167 patients had negative RT-PCR for the infection while chest CT revelations deviate away from the norms of ordinary lung, proposing the usage of chest CT to reduce incorrect negative laboratory analyses. With the latest assessed viral reproduction number (R_0), the infection transmitted by an infected individual to others or in a totally non-invulnerable populace is an average of about 3.77 [11] shows a quick spread of the illness is inevitable. It is extremely a critical task to recognize a disease-ridden person as ahead of schedule as feasible in order to isolate and further apply the treatment strategies. When a choice has been made to utilize CT images as these ongoing examinations recommend for the diagnosis or screening of patients, a need promptly ascends to quickly assess

possibly extremely enormous quantities of imaging studies. Computer-based artificial intelligence technology, specifically deep learning approach, can potentially be implemented to help radiologists in order to sort infected individuals for treatment, evaluation and pattern investigation of the information. Computer-based artificial intelligence technology can possibly dissect different cases corresponding to recognize whether chest CT uncovers any variations from the norm in the lung. In [12], a new transfer learning approach called DenseNet201 is used to identify COVID-19 cases. The model shows a good results in detecting the COVID-19 cases using the pre-trained network. In [13], the author uses GAN network to generate more images and further implements GoogleNet for the detection of COVID-19.

The paper focuses on the design of a deep learning algorithm for CT images to screen COVID-19 cases during the deadly flu season to humanity. To test this notion, 369 CT images were used which are of pathogen-confirmed COVID-19 cases along with 369 CT images previously diagnosed typical viral pneumonia. On the off chance that the model recommends a probability of infection, the case can be hailed for additional consideration for further analysis and possible treatment/isolate. Such structure once checked and tried can become key benefactors in the identification and control of patients with the infection.

The rest of the paper is organized as follows. Section 2 highlights the data creation work while Sect. 3 elaborates the proposed model. Section 4 is experiments and results, and Sect. 5 concludes the paper.

2 Data Creation

In this section, construction of the COVID CT dataset elaborated. Initially, 1260 preprints are gathered about COVID-19 from medRxiv1 and bioRxiv2, presented till May 15. A large number of these preprints report quiet instances of COVID-19, and some of them show CT scans in the reports. CT scans are related to subtitles elaborating the clinical findings in the CTs. We utilized PyMuPDF3 to remove the low-level structure data of the PDF files of preprints and found all the embedded figures. The quality of figures is very much protected. From the structure data, we likewise recognized the inscriptions related to figures. Given these extricated figures and subtitles, we first physically select all CT scans. At that point for every CT scans, we read the related inscription to judge whether it is sure for COVID-19. If cases are not able to decide from the subtitle, we found the content investigating this figure in the preprint to settle on a choice. For any figure that contains numerous CT scans as sub-figures, we physically split it into singular CTs. Finally, after dropping down some of the very low-quality CT scans, 369 CT scans named as being positive for COVID-19 are obtained. These CT pictures have different sizes. All the CT scans are resized to 224×224 . These scans are from 171 patient cases. Figure 1 gives a instance of the COVID-19 CT scans. Further 369 non-COVID CT scans are included in the dataset which has symptoms of pneumonia.

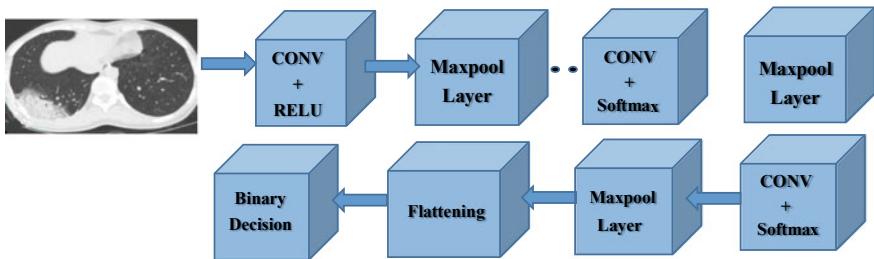


Fig. 1 Block diagram of the proposed COVID-19 detection network

3 Proposed Model

The proposed model for the detection of COVID-19 uses transfer learning approach by incorporating VGG-19 model. The depth of the network is 19. Convolutional filter of about 64 in number with size 3×3 is set. The convolution filter layers are followed by pooling layer. The complete architecture of the model is shown in Fig. 1. In order to generalize the model, data augmentation is done so as to increase the dataset. Data augmentation is an approach generally implemented to increase the number of data by some of the techniques such as cropping, padding and flipping. It makes the model robust to some variations and thereby checks over-fitting of the model to some extent. Softmax activation function is used in the output layer while ReLU function is used in other layers as output of the Softmax function is probability distribution which maps the total sum to 1. ReLU is used to introduce more nonlinearity in the network, prevent over-fitting and speedup training process. Binary cross-entropy loss function is used as the output which consists of only two classes. The block diagram is shown in Fig. 1.

4 Experiment and Results

All the experiments are carried out using Python Spider 3.1 with GTX 1080 T graphic processor and cuDNN-v7.1 deep learning library is incorporated in order to boost computation process all through the training of the network. The binary classification model is trained for predicting the possibility of a CT image to be COVID or non-COVID using 738 CT images. The created dataset consists of 369 COVID and 369 non-COVID CT images. The dataset is split into training set, validation set and test set with 66%, 17% and 17%, respectively. All the images are resized to 224-by-224. The hyper parameters are tuned on the validation set. The weight parameters in the network were optimized utilizing Adam optimizer. A batch size of 32 is chosen. The obtained graph after training the model is shown in Fig. 2. The model achieves training accuracy of 98.6 while the validation accuracy is 79.8. The model has a training loss of 0.0524 while the validation loss is 0.2975. From the result, it is seen that the model

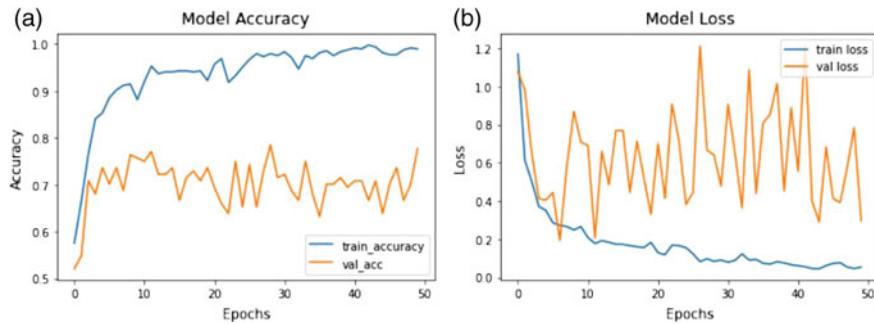


Fig. 2 **a** Model accuracy and **b** Model loss

Table 1 Performance (%) metrics

Sensitivity	Specificity	Accuracy
0.92	0.97	0.94

is efficient in predicting the COVID positive cases eight out of ten cases accurately while it fails for every two cases. The loss in the model is minimized to the extreme. Table 1 shows the performance metrics. Accuracy, sensitivity and specificity obtained are notable. The performance can be further increased by increasing the dataset and implementing further some hybrid technique. The result is promising but as the model is designed for the deadly COVID detection, more scope is there for improvement of results. Some more optimizing technique can be used in future to improve the result for far better results.

5 Conclusion

The proposed model uses transfer learning along with data augmentation technique for the detection of COVID positive cases. The dataset is created by the publically available data. The model works well in detecting the COVID positive cases approximately eight out of ten cases. The accuracy on training set is good but the accuracy on the validation set needs to be improved further for better results. The obtained data were of low quality with some induced noises. Some noise removal strategy may be incorporated further to improve the quality of the CT scan for better training of the model. Since the model is designed for the detection of the present epidemic which is severe to the mankind, more precision is required. The future work is to further improve the model with the implementation of some hybrid technique for far better predication efficiency.

References

1. Worldometers.info, Covid-19 Coronavirus Pandemic, Dover, Delaware, USA, 12 July 2020
2. D. Wang, B. Hu, C. Hu, F. Zhu, X. Liu, J. Zhang, Clinical Characteristics of 406,138 Hospitalized Patients With 2019 Novel Coronavirus-Infected Pneumonia in Wuhan, China (2020)
3. N. Chen, M. Zhou, X. Dong, J. Qu, F. Gong, Y. Han, Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. Lancet (2020)
4. Q. Li, X. Guan, P. Wu, X. Wang, L. Zhou, Y. Tong, Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia. New Engl. J. Med. (2020)
5. C. Huang, Y. Wang, X. Li, L. Ren, J. Zhao, Y. Hu, Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. Lancet (2020)
6. V.M. Corman, O. Landt, M. Kaiser, R. Molenkamp, A. Meijer, D.K. Chu, Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR. Euro surveillance bulletin Europeen surles maladies transmissibles. Eur. Commun. Disease Bull. **25**(3) (2020)
7. D.K.W. Chu, Y. Pan, S.M.S. Cheng, K.P.Y. Hui, P. Krishnan, Y. Liu, Molecular diagnosis of a novel coronavirus (2019-nCoV) causing an outbreak of pneumonia. Clin. Chem. (2020)
8. N. Zhang, L. Wang, X. Deng, R. Liang, M. Su, C. He, Recent advances in the detection of respiratory virus infection in humans. J. Med. Virol. (2020)
9. Y. Fang, H. Zhang, J. Xie, M. Lin, L. Ying, P. Pang, W. Ji, Sensitivity of chest CT for COVID-19: comparison to RT-PCR. Radiology 2020 Feb. 19:200432. [Epub ahead of print]. <https://doi.org/10.1148/radiol.2020200432>
10. X. Xie, Z. Zhong, W. Zhao, C. Zheng, F. Wang, J. Liu, Chest CT for typical 2019-nCoV pneumonia: relationship to negative RT-PCR testing. Radiology 2020 Feb. 7 (Epub ahead of print). <https://doi.org/10.1148/radiol.2020200343>
11. Y. Yang, Q. Lu, M. Liu, Y. Wang, A. Zhang, N. Jalali, Epidemiological and clinical features of the 2019 novel coronavirus outbreak in China. medRxiv [Preprint]. 2020 medRxiv [posted February 21]. <https://doi.org/10.1101/2020.02.10.20021675>
12. A. Jaiswal, N. Gianchandani, D. Singh, V. Kumar, M. Kaur, Classification of the Covid-19 infected patients using DenseNet201 based deep transfer learning. J. Biomol. Struct. Dyn. (2020). <https://doi.org/10.1080/07391102.2020.1788642>
13. M. Loey, F. Smarandache, N.E.M. Khalifa, Within the lack of chest covid-19 X-ray dataset: a novel detection model based on GAN ad deep transfer learning. MDPI **12**(4) (2020)

Blockchain Enabled and Changeable Threshold-Based Group Specific Multiple Keys' Negotiation



Anindya Kumar Biswas and Mou Dasgupta

Abstract The increase in networking along with communications taking place through the open network have resulted in multifarious intensification of rancorous and vengeful blitzkrieg. Our sensitive data, hence, is always at high risk of getting misused. In this paper, we present our proposed method of changeable threshold-based secrecy scheme, where, whenever required, the threshold can be easily changed. For this, we modify the original Shamir's (t, n) scheme because it has some inherent flaws of cheating and also because it was not possible to change the threshold there. The transpiring blockchain technology has favorable properties like it being very secure, and so, we incorporate it in our technique. We make use of the high computation dependent proof-of-work to provide reliable security to our mechanism. Our method is capable to procure multiple secure keys for an individual group. However, all the participants must collaborate together. All different groups will have unique secret keys. Not only adversaries are prevented from cheating, but also fairness of dealer is ensured by preventing from getting involved in partiality.

Keywords Blockchain · Threshold cryptography · Multi-key generation · Cheating proof · Changeable threshold

1 Introduction

Secrecy techniques, also known as information security refers to the ways by which any sensitive data is securely exchanged among a congregation of participants, by the application of cryptography. A (t, n) threshold technique always mandates that

A. K. Biswas (✉)

PhD Scholar, Department of Computer Application, National Institute of Technology Raipur, Raipur, Chhattisgarh, India

e-mail: anindya.kr.bws@gmail.com

M. Dasgupta

Assistant Professor, Department of Computer Application, National Institute of Technology Raipur, Raipur, Chhattisgarh, India

e-mail: elle.est.mou@gmail.com

minimum t out of n parameters would always be necessary to reconstruct secret. Only an authorized set of participants are allowed to reconstruct back the real secret. A chain of blocks linearly linked together through some cryptographic parameters results in the formation of blockchain. The decentralized and distributed nature make it popular and is available in the peer-to-peer network.

This paper is organized as follows: Sect. 1 (here) gives the introductory background, Sect. 2 gives the related works, Sect. 3 presents the preliminaries, our proposed method with discussions is presented in Sect. 4, and the paper is finally concluded in Sect. 5 below.

2 Related Works

There exist multiple applications associated with secret sharing schemes (SSS), like multi-party computations [1, 2], access control [3], attribute-based encryption [4], and oblivious transfer [5] in addition to threshold cryptography. However, it was Shamir [6] and Blakley [7], who independently provided significant work in threshold cryptography. The primordial Shamir's technique is based on polynomials and interpolations, became popular and lot of research work [8–10] has been done, based on it. The work [11] showed how cheating can be detected and discussed possibilities of cheater identification.

Blockchain provided the background security mechanism that made the Bitcoin [12] system popular. It was further improved [13] to address some technical flaws and make it more appropriate, and has several applications in academic as well as industrial sectors including secrecy policies. It does not rely on any third party but distributed consensus [14]. Our proposed method makes use of both blockchain and threshold schemes to achieve reliable and changeable threshold-based unique multi-key generation for each separate group. It becomes increasingly difficult to indulge in cheating or forgery and favoritism as the chain grows longer.

3 Preliminaries

3.1 Shamir's (t, n) Threshold Scheme

The general idea is that two points can uniquely define a line, but, only one point is incapable to tell anything about that particular line. This scheme involves the following steps:

1. A mutually trusted dealer D selects a prime number p , creates set $Z_p = \{1, 2, \dots, p - 1\}$, a $t - 1$ degree polynomial $f(x) = c_0 + c_1x + \dots + c_{t-1}x^{t-1}$, where $c_i \in Z_p$, $i = 1, 2, \dots, n$ and c_0 is the secret. Shares $(i, f(i))$ is created by D and distributed to all the n participants.

2. At least t or more members exchange their shares among themselves and use Lagrange Interpolation formula $f(x) = \sum_{i=1}^t f(i) \prod_{j=1, j \neq i}^t \frac{x-j}{i-j}$ to reconstruct back the polynomial and the secret $s = c_0$.

Here, cheaters can easily exchange forged shares, thereby fooling honest participants always. So, we modify it to remove the drawbacks.

3.2 Blockchain Technology

The original released blockchain has validated transactions stored. Transactions taking place during periodic time tine interval is accumulated together in candidate-block = (Hash(last block of blockchain)|Merkleroot(Tx_{set})|Timestamp), where Tx_{set} are the transactions which have been accumulated during a 10 minute interval, Merkleroot(Tx_{set}) is the root of the Merkle hash tree, and Timestamp is the moment when the block was formed. In order to solve the proof-of-work, all the interested miners compete to obtain a value $R = \text{Hash}(\text{candidate-block}|\text{nonce})$, by changing the value of nonce. This is the BTC hard problem, where different values of R are continuously found out by modifying the parameter nonce. The R should be less than a pre-specified value which is followed by some zeros, and the hash function employed here is the SHA-256 which outputs 64 hexadecimal digits or 256 bits. Our blockchain stores split shares, each block containing shares belonging to a specific group only.

4 Proposed Method

Here, we present our own scheme. The trusted dealer creates a new polynomial every time a group requests for a key negotiation. Each group, using our method, will have multiple unique keys, different from other groups, however, maintaining variable threshold property, as groups can be of different sizes determining the threshold.

4.1 Our Proposed Genesis and Data Block

Any blockchain starts with the *genesis* block and then a chain of record/data blocks chronologically and linearly linked. The *genesis* block is created by the mutually trusted dealer D . In our proposed method, D is in possession of sufficiently high computation power (like miner). Our modified *genesis* block structure is given in Fig. 1.

The above figure comprises the following parameters:

ID_D	$D_l CERT_D$	$M = \text{Hash}(ID_D D_l CERT_D)$
Body: $E(D_l CERT_D) M$		

Fig. 1 Proposed *genesis* block structure

ID_{BL_i}	$\text{Hash}(\text{previous } BL_i)$	Timestamp
Body: $E_{D_{PR}}(U = \text{Hash of } v_t \text{ shares } (i, f(i)))$		

Fig. 2 Proposed *data* block structure

1. ID_D : The identity (ID) of the dealer D stored in the header field.
2. $D_l CERT_D$: It refers to the digital (D_l) certificate ($CERT_D$) of the dealer D issued from a certificate authority.
3. M : The hash of D 's identity concatenated with D 's digital certificate.
4. Body: This is the body section of the *genesis* block which comprises the encrypted (E) record of D 's digital certificate concatenated with M value.

The block shown in Fig. 1 is the root of our blockchain system, which will be trailed by other data blocks. As requests come to D from various groups regarding secret key negotiation, the number of blocks in the chain will continue to grow. Each block holds vital information for a relevant group only and is managed by D . Provisions to incentivize D can exist once a block is validated and passes through the proof-of-work. The general block structure or the record/data block structure is given in Fig. 2.

The below parameters are presented in the above figure. The body part here contains two fields. The parameters are:

1. ID_{BL_i} : The identification number of an i th Block (BL). Each block in our method will have a unique identification, decided by D . It helps in quick access when the chain has considerably grown.
2. Hash(previous BL_i): This helps link a block to its previous block after undergoing some computation intensive process. For linking a genesis block to the first data block, the specification is $BL_i = ID_D || D_l CERT_D ||$ body and hash of it is taken out. However, the linking of one data block to the next in the chain would follow- $BL_i = ID_{BL_i} || \text{Hash}(\text{previous } BL_i) || \text{Timestamp} || \text{Nonce}$. Hash of it in entirety is to be taken, to follow a specific output and then cryptographically recorded in the next consecutive data block of the chain.
3. Timestamp : It refers to the time when a block is getting generated.
4. $E_{D_{PR}}$: It is the private key of the D (D_{PR}) and is used to encrypt (E) the concatenation of all the unique and random shares arranged in ascending order

of a variable group size. In our method, the group size is not fixed unlike other threshold-based schemes, where the group size is always fixed.

5. v_t : It refers to the variable threshold. If 6 participants (not necessarily consecutive) needs to have a common key, they come together and form a 6 – party group and so the threshold becomes 6 and accordingly D will have to create a new 6 – 1 degree polynomial. Similarly, if 10 participants decide for a common secret, then the threshold becomes 10, but the new polynomial created by the D would be of 9 degree, randomly generated.

4.2 Our Proposed Method

There exists a total of n participants, P_1, P_2, \dots, P_n and they form groups of different sizes. In our method, the threshold is decided by the D based on the group size. So, threshold changes whenever the number of participants in a group change.

Our proposed method adheres to the following steps:

Group Management Phase:

1. The participants who needs to have a common secret key can come together and inform the D about the same. Let, v_t participants come together. D now generates a new $v_t - 1$ degree polynomial after selecting a large prime number p and creating the set $Z_p = \{1, 2, 3, \dots, p - 1\}$. This value v_t is the number of participants in a group and it changes with the increase or decrease in members.
2. Split shares as $(i, f(i))$ is created by D , for $i = v_t$ number of values, representing each participant's *id* (may or may not be sequential) and is then appended in a unique data block (Fig. 2) after solving the computationally intensive proof-of-work (Sect. 3.2). This ensures validity and helps authenticate any block added to the existing chain.
3. Now, the shares are distributed to each member of the particular group securely. Each member receiving only a single share.

Forgery Detection Phase:

1. The group members exchange the shares among themselves, access the distributed blockchain and extract data from the relevant block using the block id- ID_{BL_i} . The body content is retrieved and decrypted using D 's public key D_{PU} and U is obtained. Each participant is in possession of v_t shares after receiving $v_t - 1$ shares from other participants and finds out $U' = \text{Hash}$ (all the v_t shares, arranged in ascending order).
2. If $U' = U$, then forgery or illegitimate activity have not taken place and they can proceed for secret generation. Else, if $U' \neq U$, this implies forgery of shares by malicious participants. So, an honest participant detects suspicious activity. Occurrences of any illegal activity is rare as it is computationally very difficult to alter any record in blockchain without getting caught.



Fig. 3 Proposed *general* blockchain

Secret Generation Phase:

1. Each group (having member count either same or different) consists of a group-controller G and is the member having the lowest i value from received $(i, f(i))$ shares. Using Lagrange Interpolation formula, the polynomial $f(x)$ is reconstructed by each member.
2. Decision is to be taken by the group controller G about the position of two coefficients c_j and c_k to serve as two secret keys, that is, secrets, $s_1 = c_j$ in $c_j x^j$ and $s_2 = c_k$ in $c_k x^k$ where $j, k = \text{any } v_t$ number of distinct values and depending on participant's id's in the group under consideration. j, k denotes the position of coefficients c_j and c_k and $j \neq k$. So, the value of j and k is determined by G and negotiated with the group members. Each member can now find out the two secrets based on the received value of j, k , as $s_1, s_2 = c_j, c_k$ in $c_j x^j, c_k x^k$, respectively.

Comparison/highlights: (i) Our method supports threshold changeability, unlike other methods where threshold is fixed. (ii) Although, dealer is assumed to be trusted, there is possibility of cheating and favoritism. Our method enforces restrictions on dealer, which eliminates any chance of malicious activities. (iii) Incorporation of blockchain in our method ensures that group members do not cheat each other. (iv) Our method also prevents generation of fake blocks.

Figure 3, outlines the general blockchain structure consisting of genesis block (Gen. Bl.) and various data blocks (Da. Bl.) portraying variable threshold-based groups.

5 Conclusion

The secret in our scheme can be any value, even the constant c_0 , if $j = 0$ is selected by G , the group controller. If necessary, the group size can be made as large as required and even *three* or a greater number of secret keys can be obtained per group. Any honest participant can also verify D 's public key as the necessary information is provided in the *genesis* block, along with other relevant parameters which secures it. Certificate authority is well established, and hence, any discrepancy, if exists, becomes visible.

References

1. M. Ben-Or, S. Goldwasser, A. Wigderson, Completeness theorems for noncryptographic fault-tolerant distributed computations, in *Proceedings of the 20th ACM Symposium on the Theory of Computing*, pp. 1–10 (1988)
2. D. Chaum, C. Crépeau, I. Damgård, Multiparty unconditionally secure protocols, in *Proceedings of the 20th ACM Symposium on the Theory of Computing*, pp. 11–19 (1988)
3. M. Naor, A. Wool, Access control and signatures via quorum secret sharing, in *3rd ACM Conference on Computer and Communications Security*, pp. 157–167 (1996)
4. B. Waters, Ciphertext-policy attribute-based encryption: an expressive, efficient, and provably secure realization, in *PKC 2011*, ed. by D. Catalano, N. Fazio, R. Gennaro, A. Nicolosi. LNCS, vol. 6571 (Springer, Heidelberg, 2011), pp. 53–70
5. T. Tassa, Generalized oblivious transfer by secret sharing. *Des. Codes Crypt.* **58**(1), 11–21 (2011)
6. A. Shamir, How to share a secret . *Commun. ACM* **22**, 612–613 (1979)
7. G.R. Blakley, Safeguarding cryptographic keys, in *Proceedings of AFIPS'79 National Computer Conference, Scientific Research*, vol. 48, pp. 313–317 (1979)
8. L. Harn, C. Lin, Detection and identification of cheaters in (t, n) secret sharing scheme. *Des. Codes Cryptogr.* **52**, 15–24 (2009)
9. Al. Sutijjana, Subanar, Suparna, A generalization of Shamir's secret sharing scheme. *J. Algebra* **9**(6), 283–290 (2015)
10. Y.-x. Liu, Z.-x. Wang, W.-y. Yan, Linear (k, n) secret sharing scheme with cheating detection. *Ubiquitous Comput. Commun.* IEEE Computer Society, 1942–1947 (2015)
11. A.K. Biswas, M. Dasgupta, Cryptanalysis and enhancement of Harn-Lin's secret sharing scheme with cheating detection, in *2020 First International Conference on Power, Control and Computing Technologies (ICPC2T)*, Raipur, India, pp. 27–32 (2020). <https://doi.org/10.1109/ICPC2T48082.2020.9071470>
12. S. Nakamoto, Bitcoin: A Peer-to-Peer electronic cash system (2008). <https://bitcoin.org/bitcoin.pdf>
13. A.K. Biswas, M. Dasgupta, Bitcoin cryptocurrency: its cryptographic weaknesses and remedies. *Asia Pac. J. Inf. Syst.* **30**(1), 21–30 (2020). <https://doi.org/10.14329/apjis.2020.30.1.21>
14. M. Crosby, P. Pattanayak, S. Verma, V. Kalyanaraman, Blockchain technology: beyond bitcoin. *Appl. Innov.* **2**(6–10), 71 (2016)

Speech to Sign Language Conversion Using Neuro Fuzzy Classifier



Pooja Narshetty and H. Y. Vani

Abstract Speech communication is the central interaction between the humans for the procreation and survival. Recently, researchers coordinated toward the advancement of robotized and clever investigation of human expressions. With expanded enthusiasm of human-PC/human–human connections, speech to sign language is an important another research direction. In this paper, we study the speech word to the sign image conversion. Since speech is inconsistent and noisy, fuzzy concept is introduced. The neuro fuzzy classifier is used to increase the classification rate with scaled conjugate gradient.

Keywords Adaptive neuro fuzzy inference system (ANFIS) · MFCC · Audio · Speech · Fuzzy rule · Fuzzy entropy · Membership function

1 Introduction

As per the survey, there are about 70 million people around the world who have the speech and hearing disabilities, the reason could be either at birth or by an accident occurred later on. Hence, it is very difficult for a normal abled person to interact with them because of the unfamiliar communication means used. Sign language (SL) is a common form of communication which is widely used by the speech and hearing impaired. Thus, probably the only way of easier communication/interaction with them is by learning their language—the sign language. One of the key things which can be highlighted is that all the technologies which are made cater to the normal people without any disabilities, there is a need of a system which can cater to the special class of people with which they can interact do their tasks as normal individuals, and this scope enough is the biggest reason for this kind of system to be

P. Narshetty · H. Y. Vani

Department of Information Science and Engineering, Sri Jaya Chamarajendra College of Engineering, JSS Science and Technology University, Mysore, India
e-mail: poojanarshetty7394@gmail.com

H. Y. Vani
e-mail: vanihy@sjce.ac.in

fully developed. And not only the special class but it also is very much useful for the normal people with no hearing disability.

Communication assumes a significant job in our regular daily existence. The capacity of hearing and talking has a significant job in everyday human–human collaboration. For hearing-impaired persons, the gesture-based communication is an alternative. Gesture-based communication medium is not a successful method of correspondence for the handicapped people with an ordinary one. The proposed framework provides a tool to communicate with the hearing-impaired persons. The adaptive neuro fuzzy systems [1] are utilized to develop a model, which helps to communicate with hearing-impaired persons.

2 Related Works

This section presents the investigation of existing works done so far in this area.

Silairbi et al. [2] and Vani [3] in their work propose the hybrid technique of neural network and fuzzy systems. These works are related as neuro fuzzy computing system.

Sankar K. Pal and Deba Prasad Mandal proposed speech recognition using approximate reasoning for Telugu vowels. The authors in their work propose a neuro fuzzy classifier [4] for the spoken words of Guajarati and English datasets which was developed to demonstrate the robustness of the noisy signals.

In this work [5], the importance of neuro fuzzy classifier is used to recognize phonemes and syllable informational index. Kohonen and LVQ systems are for compaction and arranging the information and the neuro fuzzy framework is for ordering. The exploratory outcomes are exhibited with acceptable precisions up to 95–96% for ANFIS.

This paper [6] presents a neuro fluffy classifier for words uttered in Thai language in clean environment.

This paper [7] provides the importance of neuro fuzzy classifier with linguistic hedges. And this work is proposed for noisy and clean speech classification. The linguistic hedges are used to improve the meaning of fuzzy rules up to secondary level. Fuzzy entropy is applied to select optimal features of MFCC for framing the rules for designing the fuzzy inference system. Results obtained from the proposed classifier is compared over conventional and neuro fuzzy classifier. In these papers [8–10] the Authors provided the information about Multilayer Neural Network, Pattern recognition using Fuzzy and ANFIS.

3 Proposed Work

The proposed algorithm for speech to sign language conversion shows that the traditional acoustic features such as MFCC mean and standard deviation and their fusion

(MEAN + MFCC) are used for feature extraction. The classifier and decision rule use the adaptive neuro fuzzy inference system. The performance of the existing state-of-the-art speech word recognition algorithm will be compared to other algorithms in terms of recognition rate and computational time. Based on the result analysis, the algorithm could be further optimized.

3.1 Speech Database

The Speech Commands Dataset was obtained from Google. The dataset has 65,000 one-second long articulations of 30 short words, by a huge number of various individuals, contributed by individuals from general society through the AIY site. It is discharged under a Creative Commons BY 4.0 permit, and will keep on developing in future discharges as more commitments are gotten. The digits ‘zero’ to ‘ten’ utterances of 100 instances are considered for our implementation.

3.2 Feature Extraction

Feature processing for speech recognition is an efficient tool to be used for model building and recognition which can be done by extracting the information.

3.3 MFCC

The Mel-frequency cepstral coefficients (MFCC) are one of the most used algorithm for extracting features from audio files. The working of MFCC is as shown in Fig. 1.

1. Framing and windowing: In this step, the signal is divided into frames of ‘n’ samples. Windowing $w(n)$ for $0 \leq n \leq N - 1$. Here, N is the total number of samples. It is performed to reduce the discontinuities among the frames.

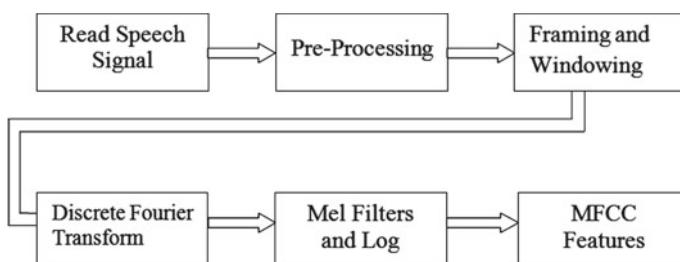


Fig. 1 MFCC extraction process

$$w(n) = 0.54 - 0.46 \cos\left(\frac{2\pi n}{N-1}\right), \quad 0 \leq n \leq N-1 \quad (1)$$

2. The frames with the windowing process are represented in a frequency spectrum using FFT.

$$D_k = \sum_{m=0}^{N_m-1} D_m e^{\frac{-j2\pi km}{N_m}} \quad (2)$$

where $k = 0, 1, 2, \dots$

3. Mel-scale is computed using the following equation

$$m(f) = 2595 \log 10\left(\frac{f}{700} + 1\right) \quad (3)$$

4. Discrete Cosine transform:

To convert back to time domain, discrete cosine transform is used. This transform helps in representing the values human understanding.

$$c_n = \sum_{k=1}^k (\log D_k) \cos\left[m\left(k - \frac{1}{2}\right)\frac{\pi}{k}\right] \quad (4)$$

5. Compute the mean for the features obtained.
 6. Twelve MFCC features are considered the implementation.

4 ANFIS Classifier

ANFIS: ANFIS is a hybrid algorithm, with the features from neural network and fuzzy systems. In this, Sugeno type is used to identify the membership function parameters of single output. It incorporates both least squares and back propagation techniques which are used for building a robust system to handle overlapping. The ANFIS has five layers; each layer has its designated characteristics as it is discussed.

5 Implementation

The implementation consists of following modules.

- Step1: Consider the audio samples, and these audio files are called as dataset.
- Step2: Once the dataset is considered, the next step is feature extraction.
- Step3: MFCC algorithm is used for the feature extraction process.
- Step4: The next step is to classify the results using ANFIS classifier.

Step5: The ANFIS classifies the results, and speech to sign image is displayed.

6 Results and Observations

The data set mentioned above was considered for obtaining results. The features were extracted using MFCC MEAN and MFCC STD, the results were observed for both, also a module which converts speech command to the resultant sign was developed, and the hand signs were obtained as per standard ASL sign language. Below are the results for both (Table 1).

The above table gives the information about the different results by changing epoch, split range, and numMF's. When the number of words increases, accuracy is less. And from the results, we can observe that as we increase epoch and split range we will get better results (Table 2).

In Table 2, it gives the comparative results of ANFIS and SCG_NFC. By using ANFIS, the results are efficient, whereas using SCG_NFC, results are highly efficient because it is suitable for large-scale problems [11]. This classifier is used at the fifth layer for optimal results.

Table 1 Results by changing epoch, split range, and numMF's

Epoch	Num MFs	Split range	Number of words	Accuracy MFCC std	Accuracy MFCCs mean
50 200	2	2	2	92.50	94
	4	2		94.50	96
	3	2		94.00	92
	2	5		100.0	100
	3	2		99.50	99
200 300 500 100	2	4	4	63.75	67.75
	2	4		63.75	65.75
	2	2		44.25	43.25
	3	2		58.25	51.25
200 100	2	5	8	31.87	37.87
	2	6		49.38	51.38
50 100 200 300 500	2	2	10	58.20	62.20
	2	2		60.30	65.30
	3	2		66.10	67.10
	2	3		61.80	61.80
	2	8		70.00	72.00

Table 2 Comparative results

ANFIS	SCG_NFC
MFCC_STD:72%	MFCC_STD:72%
MFCC_MEAN:74%	MFCC_MEAN:76%

6.1 Observations

- ANFIS-LH handles overlapped data in a better way
- Scaled conjugate is to improve the speed of computation
- DNN requires huge amount of data.

7 Conclusions

In this work, an assessment of speech word recognition is done. The extraction of the features of acoustic signals is an overwhelming mission since it impacts the recognition execution. In this regard, the effect of cepstral coefficient like MFCC and obtaining its mean and standard deviation are tried. From the results, it is observed that MFCC mean gives the better results for the data set compared to standard deviation, also a separate module was developed to display the corresponding sign for the recognized speech, and this acts as an application. As a future scope, adding a deep neural Network (DNN) along with ANFIS would still obtain a better result.

References

1. C.T. Sun, J.S.R. Jang, A neuro-fuzzy classifier and its applications, in *Proceedings of IEEE International Conference on Fuzzy Systems*, San Francisco, vol. 1, pp. 94–98 (1993)
2. Silairbi, B. Abiderrahmane, A. Benyeittou, Fluffy network adaptive based inference system for speech recognition through subtractive clustering. *Int. J. Artif. Intell. Appl. (IJAIA)* **5**(6) (2014)
3. H.Y. Vani, M.A. Anusuya, A neuro fuzzy classifier with linguistic hedges for speech recognition. Research Article EAI Endorsed Transactions on Internet of Things (2018)
4. S.K. Pal, D.P. Mandal, Fuzzy logic and approximate reasoning: an overview. *Electronics and Communication Sciences Unit, Indian Statistical Institute, Calcutta, 700035, India* (2016)
5. N. Helmi, B.H. Helmi, speech recognition with fuzzy neural network for discrete words, in *Fourth International Conference on Natural Computation*. <https://doi.org/10.1109/icnc.2008.666>
6. K. Srijiranon, N. Eiamkanitchat, Thai speech recognition using Neuro-fuzzy system, in *2015 12th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON)*
7. H.Y. Vani, M.A. Anusuya, Fuzzy speech recognition: a review. *Int. J. Comput. Appl. (0975-8887)* **177**(47) (2020)
8. A. Ghosh, N.R. Pal, S.K. Pal, Self-organization for object extraction using a multilayer neural network and fuzziness measures. *IEEE Trans. Fuzzy Syst.* **1**(1), 54–69 (1993)
9. K.S. Ray, J. Ghoshal, Approximate reasoning approach to pattern recognition, in *Electronics and Communication Sciences Unit, Indian Statistical Institute, Calcutta, 203.B.T Road, 700035, India* (1992)
10. T.A. de Lima, An investigation of Type-1 adaptive neural fuzzy inference system for speech recognition

11. B. Cetili, A. Barkana, Speeding up the scaled conjugate gradient algorithm and its application in neuro-fuzzy classifier training. *Soft. Comput.* **14**(4), 365–378 (2010)
12. N.K. Kasabov, Learning fuzzy rules and approximate reasoning in fuzzy neural networks and hybrid systems (1996)

Design and Development of an Ultra-wideband Millimetre-Wave Antenna for Short-Range High-Speed Communication



Debraj Dhang, Satyadeep Das , and Sudhakar Sahu

Abstract An annular microstrip antenna for millimetre-wave applications. This is a novel antenna for a cost-sensitive millimetre-wave of ultra-wide bandwidth application. This proposed antenna is a 16 mm × 16 mm miniaturized structure imparting very high gain and ultra-wide bandwidth. The simulations and results show that this antenna is suitable for increasing the impedance bandwidth and to provide more flexible orientation for high bandwidth communication in a short-range. Investigations depict the antenna's operating bandwidth of 98.50–118.7 GHz with highly directive radiation.

Keywords Microstrip patch antenna (MPA) · Ultra-wideband (UWB) · Coaxial fed · Mm-wave

1 Introduction

In the modern and technology-based era, the speed of data transmission is increasing and its demand too. Nowadays, antennas are the most suitable medium for wireless data transmission. And antennas are totally dependable on the frequency. Depending upon the frequencies, antennas are classified as lower frequency antenna and higher frequency antenna. Higher frequencies antennas are more suitable for high data transmission [1, 2]. Because high frequency provides more bandwidth, so that we can send huge data in less time. High-frequency antennas find their wide applications in marine and aircraft communications, radio for disaster prevention and administration, ham radio, etc.

D. Dhang · S. Das · S. Sahu

School of Electronics Engineering, KIIT-DU, Bhubaneswar, Odisha, India

e-mail: satya626auro@gmail.com

D. Dhang

e-mail: 1707362@kiit.ac.in

S. Sahu

e-mail: ssahufet@kiit.ac.in

Aperture antennas are widely in use for many mm-wave applications due to their low profile, simple structure, higher gain and broadband [3, 4]. Generally, 30% enhancement of the impedance bandwidth can be achieved by various methods [5, 6], e.g., conical cavity structures, microstrip patch radiators [7, 8] H-shaped slot exciter [9], sleeve dipole [10], etc., but at the cost of gain drop and poor impedance matching. Millimetre-wave frequencies gain importance for compact high-performance transmitters, receivers, automotive collision avoidance radars (ACAR), imaging arrays and mobile communication systems [11, 12].

The organization of the communication includes the design of the annular ring of copper as a patch over a lossy substrate of FR-4 epoxy as discussed in Sect. 2. Followed by the analysis of simulation results in Sect. 3 and finally the conclusion.

2 Design Ideation

We have started our antenna designing with a lossy substrate FR4-epoxy of dielectric constant $\epsilon_r = 4.4$ of height 1.6 mm and a chamfered annular square ring as a radiator. Our main aim is to attain the maximum gain and simultaneously maintaining ultra-wide bandwidth. The copper layer is located on both sides of the substrate, so the other side behaves as the infinite ground plane. The proposed antenna is coaxially fed at an offset location instead of symmetricity (Fig. 1 and Table 1).

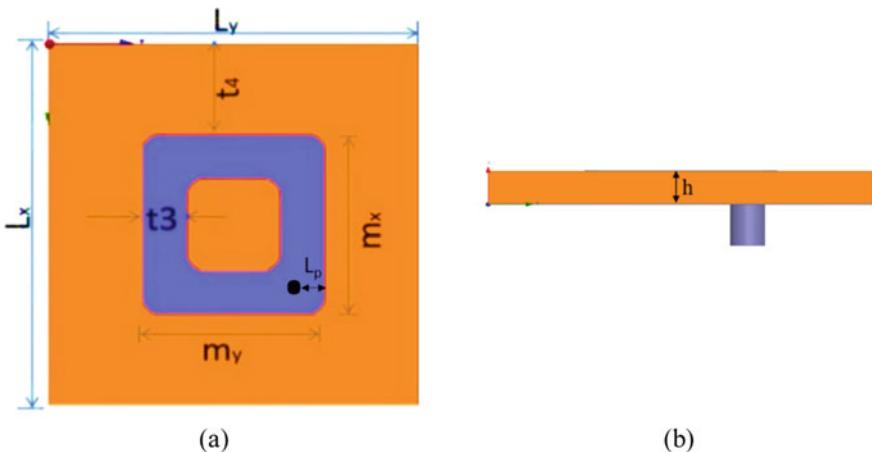


Fig. 1 Schematic of the proposed antenna: **a** top view, **b** side view

Table 1 Physical dimension of the proposed design

Parameter	L_x	L_y	m_x	m_y	t_3	t_4	L_p	h
Dimension (mm)	16	16	8	8	2	4	1	1.6

3 Simulations Analysis

All the simulations have been carried out by finite elements method-based HFSS software. The simulated performances of the proposed mm-wave antenna, specifically reflection coefficient (S_{11}), Smith chart and far-field patterns, has been discussed as follows.

3.1 *Reflection Coefficients (S_{11})*

Figure 2 shows the simulated reflection coefficient (S_{11}) versus frequency of the proposed antenna. From this simulation, it is shown that the achieved impedance bandwidth is 19.8 GHz ranging from 98.7 to 118.5 GHz with a fractional impedance bandwidth of about 18.7%. The best matching of -18.75 dB is obtained at 109.7 GHz. The impedance bandwidth validates the claim for UWB.

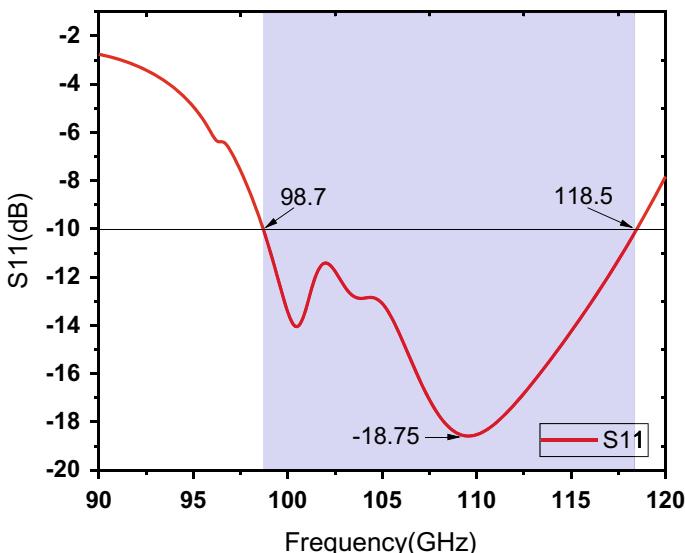


Fig. 2 Simulated S_{11} versus frequency of the proposed mm-wave antenna showing the operating bandwidth from 98.7 to 118.5 GHz

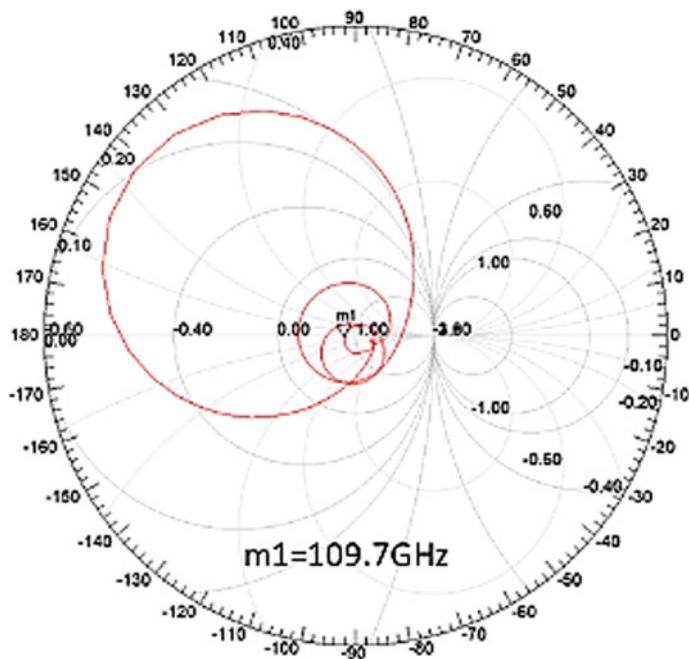


Fig. 3 Smith chart showing 50Ω impedance at 109.7 GHz

3.2 Smith Chart

Figure 3 shows the simulated Smith chart curve that exactly satisfies the 50Ω impedance at 109.7 GHz, the best matching operating frequency obtained from the S11 curve.

3.3 Radiation Pattern

Figure 4 depicts the radiation pattern of the proposed mm-wave antenna at 109.7 GHz in E-plane and H-plane for the claimed UWB frequency ranging from 98.7 to 118.5 GHz. The patterns have its main beam of radiations in +ve z-direction with low back lobe which justifies to the concept of broadside radiations.

Due to high frequency of operation, the 3D polar plot response of the proposed design shows abrupt yet consistent near-field gain of 2.8 dBi in the broadside direction as depicted in Fig. 5.

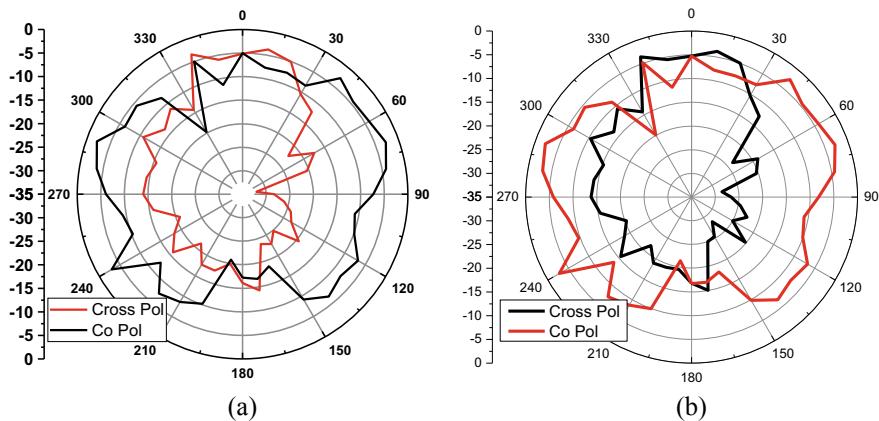


Fig. 4 Directive radiation pattern of the proposed mm-wave antenna: **a** E-plane, **b** H-plane

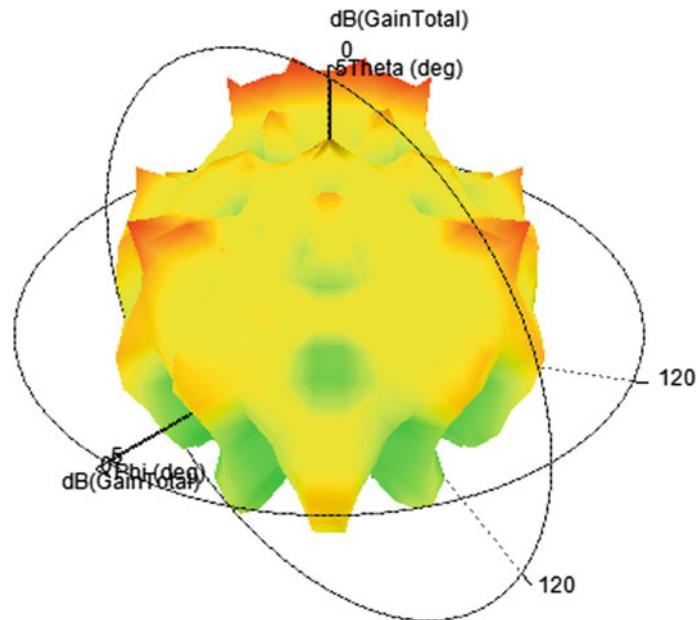


Fig. 5 3D polar plot depicting the peak gain

4 Conclusion

A novel mm-wave UWB microstrip annular ring patch antenna having a frequency range from 98.7 to 118.5 GHz, made on FR4epoxy substrate, is designed. This proposed antenna can be used for short-range high bandwidth communication.

This antenna has achieved an impedance bandwidth of 19.8 GHz. The proposed design fulfils the required specifications. The obtained spectrum from is particularly promising for it is sufficient enough to handle high data rates up to multiple gigabits per second for a short distance communication.

References

1. O.M. Haraz, M. Abdel-Rahman, S.A. Alshebili, A.R. Sebak, A novel 94-GHz dipole bow-tie slot antenna on silicon for imaging applications, in *2014 IEEE Asia-Pacific Conference on Applied Electromagnetics (APACE 2014)—Proceeding*, pp. 59–62 (2015)
2. R. Nesti, G. Pelosi, S. Pilia, S. Selleri, Design of a 67–116 GHz corrugated circular horn for the ALMA radio telescope, in *2017 IEEE Antennas and Propagation Society International Symposium. Proceedings*, vol. 2017-Jan, pp. 565–566 (2017)
3. C.A. Balanis, *Modern Antenna Handbook*, 2011th edn. (Wiley, Hoboken, NJ, USA, 2007)
4. W.L. Stutzman, G.A. Thiele, *Antenna Theory and Design*, 3rd edn., vol. 23, no. 5 (Wiley, Hoboken, NJ, USA, 1981)
5. S.-W. Qu, B.N. Kung, C.H. Chan, Waveguide Fed broadband millimeter wave short backfire antenna. *IEEE Trans. Antennas Propag.* **61**(4), 1697–1703 (2013)
6. S.W. Qu, Broadband millimeter-wave short backfire antenna with bowtie exciter. *IEEE Antennas Wirel. Propag. Lett.* **11**, 850–853 (2012)
7. D.P. Gray, L. Shafai, Parametric study of short backfire antennas with different cavity profiles, in *Proceedings of the IEEE Antennas and Propagation Society International Symposium*, vol. 3, pp. 1314–1317 (2000)
8. D. Gray, H. Tsuji, Short backfire antenna with microstrip Clavin feed. *Microw. Antennas Propag.* **3**(8), 1211–1216 (2009)
9. P. Hazdra, M. Mazanek, Wideband shorted backfire antenna with L-probe fed microstrip patch, in *Proceedings of the 4th European Conference Antennas Propagation*, pp. 1–3 (2010)
10. S. Srikanth, G. Behrens, A new broadband short-backfire antenna as a prime focus feed single and dual band, in *Proceedings of IEEE Antennas and Propagation Society International Symposium*, 2007, pp. 3684–3687.
11. G.P. Gauthier, S. Raman, G.M. Rebeiz, A 90–100 GHz Double-Folded Slot Antenna. *IEEE Trans. Antennas Propag.* **47**(6), 1120–1122 (1999)
12. Z.J. Shao, Y.P. Zhang, and J.F. Mao, A 94 GHz array antenna for 45° linear polarization in LTCC technology, in *2018 IEEE International Symposium on Antennas and Propagation Society International Symposium and USNC-URSI Radio Science Meeting—APS/URSI 2018—Proceedings*, pp. 1033–1034, 2018.

A Comparative Analysis of BTC Variants



D. Nayak, K. B. Ray, and T. Kar

Abstract The goal of image compression is to reduce the amount of information required to represent a digital image necessary for efficient storing, transmission, uploading, and downloading with less cost and time and at comparatively lower bandwidth. The block truncation coding method of image compression method is the simplest among all, in this paper, we have presented a comparative study of five popular BTC variants namely traditional BTC, AMBTC, two-parameter training method, and edge-based BTC and multi grouping BTC, in terms of objectives measures such as PSNR, SSIM, FSIM, and CR between original and reconstructed images as well as subjective measures through visible artifacts. All the algorithms rely on a common process of dividing the image into non-overlapping blocks and differ in the way of evaluating the quantization levels of the BTC technique. The different objective measures are evaluated on standard database images.

1 Introduction

With the advancement of multimedia technology, there is an easy availability of low cost, compact, and easy to operate image acquisition tools that lead to impulsive growth of multimedia data. This requires large digital space to store and larger bandwidth to transmit which is costlier as well as has lesser bandwidth efficiency. So, there is an urgent need to store and transmit these huge volumes of image data efficiently. This creates the necessity of image compression that tends to reduce the amount of image data needed to store and transmit. So, the image compression methods are gaining popularity in multimedia devices. Image compression techniques are categorized under two different categories, i.e., (i) Lossy and (ii) Lossless image compression. As the name suggests, in lossy image compression techniques, there is a loss of some information during compression and reconstruction of the input image. In contrast to lossy compression, lossless image compression has minimum

D. Nayak (✉) · K. B. Ray · T. Kar
KIIT Deemed to be University, Bhubaneswar, Odisha, India
e-mail: dibyalekhanayak@gmail.com

loss or no loss during compression and reconstruction process. Generally, lossy image compression has good compression ratio, whereas lossless image compression produces better perceptual quality of the image with lower compression ratio.

Lossy image compression methods can be either direct or indirect type. The direct method of image compression are comparatively simpler, and the compression is achieved in spatial domain itself whereas in indirect method, the image compression is achieved via manipulation of the frequency domain coefficients of the image data. BTC, fractal compression, and vector quantization are examples of direct method of compression and discrete cosine transform (DCT), discrete wavelet transform (DWT), and Walsh Hadamard transform (WHT) are few examples of indirect method of image compression. Among all methods, block truncation coding (BTC) which is a direct method of compression has gained popularity over past few years due to its low complexity of implementation. In basic BTC [1], a block-wise bit map was generated as the compressed image with 1 and 0 using mean and variance-based thresholding approach. To get the reconstructed image, the bit map was replaced with two-level quantization which was also known as moment preserving BTC (MPBTC). Lema et al. proposed the absolute moment-based BTC (AMBTC) having higher performance than the MPBTC [2]. Wu et al. developed a hybrid technique by combining the discrete cosine transform, vector quantization, and BTC to reduce the bit rate of the image [3]. Pasi et al. developed entropy-based BTC and was able to achieve lower bits per pixel (bpp) without substantially increasing the complexity [4]. Guo et al. included the void and clustering methods to improve the performance of the BTC [5]. Joshi et al. attempted optimization of the compression algorithms using BTC, DCT, and DWT. Chuang et al. developed an improved multi grouping AMBTC methods [6] and achieved good performance. An adaptive block truncation coding was developed by considering the edge feature information of the pixel of the block to get the variable CR [7]. Ryan et al. [8] developed a k-d tree-based BTC method to reduce the bit rate. An adaptive multi-level block truncation coding was developed by Nadeem using level selection process based on the absolute difference which led to an improved architecture [9]. Chen et al. proposed a two-parameter modeling for the quantization level [10] of BTC. Xiang et al. developed a multi grouping where the grouping has been done based on the difference between two-level quantization values with the predefined threshold. BTC variants have been developed to improve the quality of the image based on the PSNR and the bpp [11] but at a lower complexity. Motivated by this in the current work, we made a comparative evaluation of the performances of traditional and few state-of-the-art BTC variants through subjective and objective fidelity criteria.

The rest of the paper has been organized as follows: Sect. 2 discusses different BTC variants, Sect. 3 focuses on simulation and results from the analysis of different BTC variants followed by conclusions in Sect. 4.

2 Different BTC Methods

2.1 Traditional BTC

In traditional BTC, the grayscale image is divided into 4×4 non-overlapping blocks. Then, the block-wise mean is compared to the pixel value in the block to generate the binary bitmap image as given in (1).

$$B(i, j) = \begin{cases} 0, & x(i, j) < \mu \\ 1, & x(i, j) \geq \mu \end{cases} \quad (1)$$

where μ is the mean of the block. The reconstructed image is obtained by replacing 1's and 0's of the pixel value by the two quantization values x_h and x_l of the block, respectively, evaluated as given in (2) and (3), respectively.

$$x_h = \mu + \sigma \sqrt{\frac{p}{q}} \quad (2)$$

$$x_l = \mu - \sigma \sqrt{\frac{p}{q}} \quad (3)$$

where σ : standard deviation of the block, p is the total number of 0's in the image block and q : is the total number of 1's in the image block.

For the reconstruction of the compressed image, the bit map image block $B(i, j)$ is encoded with the x_h and x_l values for each block as given in (4)

$$y(i, j) = \begin{cases} x_h, & B(i, j) = 1 \\ x_l, & B(i, j) = 0 \end{cases} \quad (4)$$

This method is not suitable for the hardware implementation, because of the square root term in evaluation of the x_h and x_l values. To get the hardware implementation and better bit rate, different simple BTC methods have been proposed.

2.2 Absolute Moment Based BTC (AMBTC)

AMBTC is a simple lossy color image compression method with an appreciable compression ratio, suitable for real-time application. In this technique, the image is divided into 4×4 non-overlapping blocks similar to the traditional BTC method, and the block-wise mean is evaluated as given in (5).

$$\text{Mean} = \frac{\sum_{i=1}^m \sum_{j=1}^m x(i, j)}{m \times m} \quad (5)$$

where $m = 4$ and represents size of the block and $x(i, j)$ represents the value of the pixel of the image block. If the pixel value exceeds the mean, then the value of the pixel will be replaced by the 1 else it will be replaced by 0 which will generate a bit map image block similar to the traditional BTC. The x_l and x_h values are evaluated as given in (6) and (7)

$$x_l = \frac{1}{K} * \sum_{\substack{x(i, j) \leq \text{Mean}}}^n x(i, j) \quad (6)$$

$$x_h = \frac{1}{n - K} * \sum_{\substack{x(i, j) > \text{Mean}}}^n x(i, j) \quad (7)$$

where K denotes the total numbers of 0s.

The reconstructed image is obtained by replacing the 1's and 0's of the bit map image with x_h and x_l , respectively.

2.3 Two Parameter Method (2p BTC)

Digital half toning-based BTC has been proposed based on two-parameter quantization. The bit map image block of size 4×4 is replaced by the high level (x_h) and low level (x_l) quantization value similar to traditional BTC. The high level and low level quantization are based on the two parameters α_1 and α_2 , and the values of the parameters are decided based on the PSNR training method. The x_h and x_l are evaluated as given in (8) and (9), respectively.

$$x_h = \max - \alpha_1 (\max - \text{mean}) \quad (8)$$

$$x_l = \min + \alpha_2 (\text{mean} - \min) - 1 \quad (9)$$

where max and min are maximum and minimum value of the block.

2.4 Edge Based AMBTC Method (EBTC)

In this method, an edge map of the image is generated by using the canny filter. The edge map of the non-overlapping blocks of size 4×4 are categorized in to edge informative block and non-edge informative block. Each edge informative blocks are divided into 3 clusters by k-means algorithm which gives the encoded values of '00,' '01,' and '10' based on the cluster to which it belongs to in the block, and in the non-edge informative blocks, the normal AMBTC is used to get the encoded

image value of 0 and 1. By this, the authors claimed to achieve variable compression rate with improved bpp and CR than the traditional BTC and AMBTC methods. For the reconstruction of the image, all the encoded data is replaced by the mean of the clusters for the edge informative block and mean of the non-edge informative block.

2.5 *Multi Grouping BTC (MGBTc)*

In this method, an adaptive and dynamic multi grouping technique is used for the AMBTC. In this method, a predefined threshold is set and compared with the difference between the quantization values x_h and x_l calculated by AMBTC method. If the x_h and x_l value exceeds the threshold, then the blocks are grouped else they are left without grouping. In this technique, groups 0 and 1 further grouped into 00, 01, 10, and 11 groups. With multi grouping BTC the bpp and the CR of the image is improved with higher perceptual quality of the reconstructed image.

3 Simulations and Result Analysis

For simulation of different BTC variants, we have considered 6 basic images such as lena, peppers, girl, airplane, girl with toy, boat, and twenty three images from the Kodak dataset [12]. The different BTC-based compression algorithms are compared based on various objective criteria's such as PSNR [7], CR [7], Bpp, and SSIM [13]. PSNR of an image is the most important objective of the quality assessment criteria. For the good quality of reconstructed image, the mean square error (MSE) value should be as low as possible and PSNR should be as high as possible. However, a compression algorithm also demands to have high compression ratio or low bpp. The compression ratio is the ratio between the size of original image to the compressed image. The bpp is the bits per pixel present in the compressed image.

The quality and the rate of compression of the reconstructed image are evaluated based on PSNR, SSIM, CR, and FSIM [14]. All the results were simulated in MATLAB 2017a environment, and the results were given in Table 1. The average performance measure is placed in Fig. 1. It is observed from Table 1 and Fig. 1 that MGBTc has the highest average CR, EBTC has the highest average FSIM and SSIM and 2p BTC has the highest average PSNR measure. The highest average measures are indicated by bold face in Table 1.

The original image and reconstructed image of Lena (BMP image of size 512×512) for the considered BTC variants are shown in Fig. 2. It is observed from Fig. 2 that 2p BTC has the best perceptual visual quality among all the considered methods.

Table 1 PSNR, SSIM, and FSIM measure of different BTC methods

Image		BTC [1]	AMBTC [2]	2p BTC [10]	EBTC [7]	MG BTC [11]
Lena	PSNR	35.163	35.958	55.82	36.99	35.609
	SSIM	0.934	0.8818	0.1887	0.9632	0.8353
	FSIM	0.959	0.8668	0.5737	0.9678	0.8838
Airplane	PSNR	34.52	35.576	52.116	32.854	36.478
	SSIM	0.9402	0.875	0.1977	0.924	0.8742
	FSIM	0.9504	0.8671	0.6189	0.9925	0.8409
Girl	PSNR	33.82	34.504	51.304	32.451	34.634
	SSIM	0.9151	0.8476	0.2295	0.945	0.7976
	FSIM	0.9456	0.8631	0.6355	0.9769	0.7869
Peppers	PSNR	34.91	36.195	53.347	32.03	35.601
	SSIM	0.9352	0.8798	0.2108	0.9523	0.825
	FSIM	0.9456	0.8871	0.6019	0.9683	0.8873
Boat	PSNR	34.04	34.898	54.099	34.456	34.974
	SSIM	0.9187	0.8481	0.1484	0.954	0.7876
	FSIM	0.9587	0.9012	0.6352	0.954	0.8455
Girl with toy	PSNR	33.98	34.814	55.142	32.198	33.654
	SSIM	0.9053	0.8294	0.1336	0.9725	0.6839
	FSIM	0.9653	0.8891	0.6615	0.9849	0.8804
Kodak image set(Avg)	PSNR	35.223	36.177	55.814	33.979	34.85
	SSIM	0.916	0.8566	0.1118	0.9536	0.6807
	FSIM	0.9407	0.9328	0.6354	0.9616	0.8856
Average	Average PSNR	35.446	34.522	53.949	33.565	35.115
	Average SSIM	0.9256	0.8598	0.1744	0.9604	0.7835
	Average FSIM	0.9456	0.8867	0.6232	0.9671	9.3382
	Average CR	4	4	4	2.91	5.32

4 Conclusions

In this paper, we have compared five variants of BTC based on the quality parameters such as PSNR, SSIM, FSIM, and CR. From simulation results, it is found that the two-parameter BTC has the highest average PSNR value of 53.94. The edge-based BTC is having the highest average SSIM and FSIM values of 96.04% and 96.71%. While considering the compression ratio of BTC variants, the multi-level BTC method

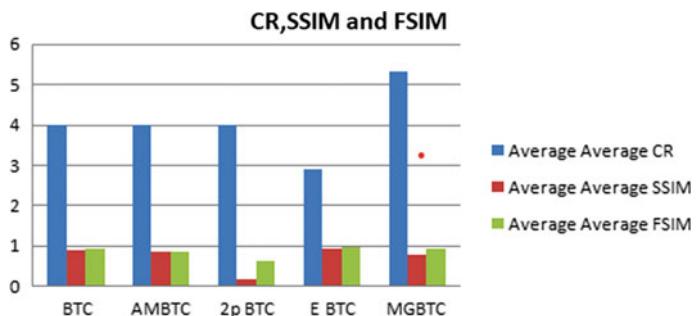


Fig. 1 The comparison of average values of CR, SSIM, and FSIM values of BTC variants



Fig. 2 a Original image of size 512×512 and reconstructed image of b traditional BTC, c AMBTC, d Edged-based BTC, e Two parameter BTC, f Multi grouping BTC

has superior performance in terms of the highest CR of 5.32 among the considered methods.

References

1. E.J. Delp, O.R. Mitchell, Image compression using block truncation coding. *IEEE Trans. Commun.* **27**(9), 1335–1342 (1979)
2. M.D. Lema, R. Mitchell, Absolute moment block truncation coding and its application to color images. *IEEE Trans. Commun.* **32**(10), 1148–1157 (1984)
3. Y. Wu, D. Coll, BTC-VQ-DCT hybrid coding of digital images. *IEEE Trans. Commun.* **39**(9), 1283–1287 (1991)
4. P. Franti, O. Nevalainen, Block truncation coding with entropy coding. *IEEE Trans Commun.* **43** (1995)
5. J.M. Guo, M.F. Wu, Improved block truncation coding based on the void-and-cluster dithering approach. *IEEE Trans. image Process.* **18**(1) (2009)
6. J. Chuang, Y. Hu, C. Chen et al., Adaptive grayscale image coding scheme based on dynamic multi-grouping absolute moment block truncation coding. *Multimed. Tools Appl.* (2020). <https://doi.org/10.1007/s11042-020-09325-3>
7. J. Mathew, M.S. Nair, Adaptive block truncation technique using edge based quantization approach. *Comput. Electr. Eng.* (2015). <https://doi.org/10.1016/j.compeleceng.2015.01.001>
8. R. Rey, M. Daga, Improved k-d tree-segmented block truncation coding for color image compression, in *IEEE 2nd International Conference on Signal and Image Processing* (2017)
9. Zainul, N.A., A. Jaffery, Irshad, An approach to color image coding based on adaptive multilevel block truncation coding, in *Applications of Artificial Intelligence Techniques in Engineering*, pp. 597–606 (2018)
10. S.L. Chen et al., VLSI implementation of an ultra-low-cost and low power image compressor for wireless camera networks. *J. Real-Time Image Process.*, 1–10 (2015)
11. Z. Xiang, C.Y. Hu, Adaptive and dynamic multi-grouping scheme for absolute moment block truncation coding. *Multimed. Tools Appl.* (2019)
12. Kodak lossless true color image suite. <https://www.r0k.us/graphics/Kodak/>. Accessed April 2019
13. Z. Wang, A.C. Bovik, H.R. Sheikh et al., Image quality assessment: from error visibility to structural similarity. *IEEE Trans. Image Process* **13**(4), 600–612 (2004)
14. L. Zhang, X. Mou, D. Zhang, FSIM Matlab source code. <https://www4.comp.polyu.edu.hk/~cslzhang/IQA/FSIM/Files/FeatureSIM.m>. Accessed April 2019

Crosstalk Noise Reduction in Long Wire Interconnects Using MTCMOS Inverters



Jayashree Mallidu and Saroja V. Siddamal

Abstract The scaling in the CMOS technology has derived a new route to the miniaturization of semiconductor devices. As technology is changing day by day, more and more transistors are integrated into a single chip. As the transistor increases in number, the operating speed should also increase. Power consumption of the circuit increases with the increase in the count of transistors on the die. The paper addresses the effect of MTCMOS in reducing power consumption. MTCMOS technique is a power or gating structure used to reduce leakage current in the standby state and also reduced static power consumption. Experimental results show that 60–19% leakage power reduction compared to the CMOS technique. The investigation is further continued in reducing crosstalk noise. In this paper, the crosstalk noise which occurs in between two wire interconnects due to coupling capacitance is reduced by using MTCMOS repeaters. The simulation is carried in the Cadence tool at 180 nm technology. MTCMOS repeater uses low power 41.323 pW compared to CMOS repeater of 17.872 nW.

Keywords MTCMOS · Repeaters · Crosstalk noise · Leakage current · Interconnects

1 Introduction

The scaling down of CMOS technology has led to more functions in single die. This has increased the number of long global interconnects [1]. This has led to degradation of voltage due to interconnect delay. To reduce such interconnect, delay repeaters are in use. The authors in paper [1] have proposed encoding algorithm to reduce capacitive crosstalk noise and total power. The authors have used high threshold

J. Mallidu · S. V. Siddamal (✉)
KLE Technological University, Hubballi, India
e-mail: sarojavs@kletech.ac.in

J. Mallidu
e-mail: jayashree.mallidu@kletech.ac.in

transistors with buffer using novel bus encoding scheme. The authors claim 26% of total power and 42% leakage power eliminating capacitive crosstalk. The authors in paper [2, 3] have worked on insertion of repeaters to reduce crosstalk noise. The authors have used MTMOS to minimize power dissipation. They have calculated the size, distance of repeaters, and sleep transistors to reduce power. The contribution is to reduce the power in global bus interconnects. The authors in paper [4] highlight of the effect of buffer insertion and increase of leakage power due to buffer insertion. The authors have implemented a bus system in 65 nm SOI technology in which they claim 45% reduction in system power.

In paper [5], discussion is done on the role of interconnects in designing integrated circuits. The paper highlights on various techniques and approaches used by researchers on interconnects. In paper [6], the authors have proposed a novel data bus system which reduces crosstalk noise. A set of crosstalk noise aware codes are used to minimize the data transmission. The system is implemented in 40 nm CMOS technology. The authors claim 12.5% leakage power reduction and 18.26% active power consumption. The authors in [7] have proposed a high-speed low-power repeater. The proposed repeater has capacity to drive large loads with reduced delay and power dissipation. It is implemented in 45 nm CMOS technology. Crosstalk contributes major to the amount of noise in IC design. In paper [8], the authors have discussed mitigation techniques to impact the crosstalk. They use techniques like repeaters, skewing, shielding, etc., for crosstalk mitigation in long wire interconnect. Spice simulation is performed to show the effects of techniques. Semiconductor technology had undergone a lot of change. Crosstalk is one of the major issues faced. The authors in [9] have developed a flow to analyze the effect of crosstalk. Synopsis prime tool is used to calculate delay and to address the issues of crosstalk.

2 Crosstalk Noise

In deep submicron technology (>130 nm), the presence of capacitive coupling in between the nets leads to logic failures and timing degradation in VLSI circuits. Crosstalk is a phenomenon, by which logic transmitted in a VLSI circuit or a net or wire creates undesired effects on the neighboring circuit or nets or wires. Due to capacitive coupling, the switching characteristics of nets are affected by the simultaneous switching of nets that are nearby (see Fig. 1). The net under analysis that suffers from coupling noise is termed as the victim, and all the neighboring nets which contribute to coupling noise is referred on the victim are termed as aggressors.

Delay noise contributes significantly for high performance of VLSI circuits. Therefore, it becomes necessary to accurately model the delay noise while performing timing analysis for VLSI chips in the ‘nm’ process technology.

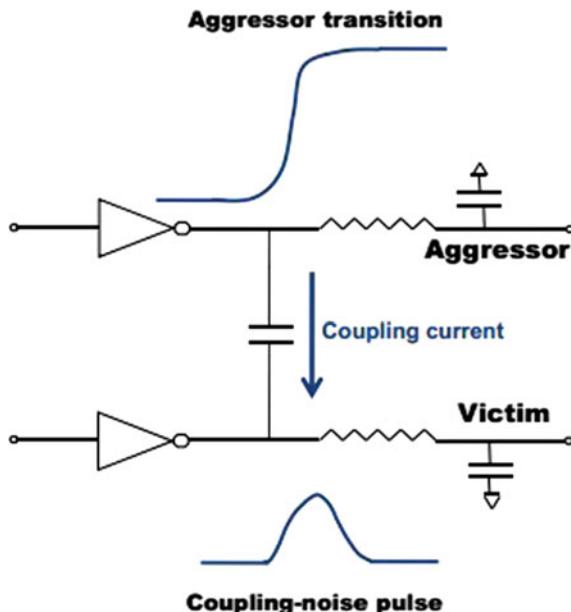


Fig. 1 Effect of coupling capacitor

2.1 Effects of Varying Coupling Capacitance and Crosstalk Generation

In the deep submicron, the coupling capacitance becomes more dominant than any other interlayer capacitance. Hence, this capacitive coupling leads to the logic failure and timing degradation in circuits. Crosstalk is the phenomenon in which the logic transmitted in one interconnects will cause timing degradation or logic failure in other interconnects. The variation of the output at the victim node due to change in coupling capacitance is shown below with the waveforms. The circuit of two interconnects with the coupling capacitance C_c is as shown in Fig. 2.

Figure 3 shows the effect of coupling capacitance on victim node. The capacitance

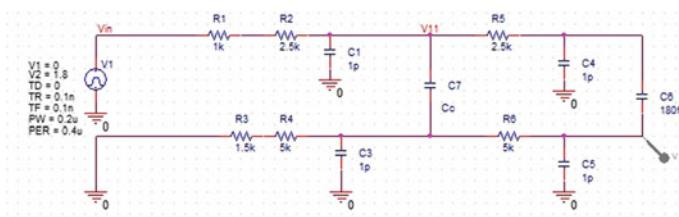


Fig. 2 Interconnects with the coupling capacitance C_c

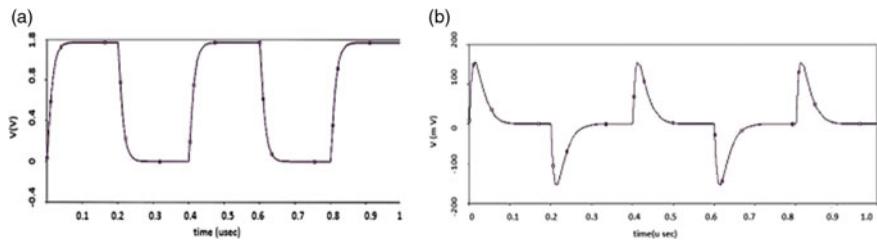


Fig. 3 Effect of coupling capacitance on victim node: **a** Output waveform at the victim node for $C_c = 180 \text{ nF}$. **b** Output waveform at the victim node for $180 \times 10^{-17} \text{ F}$

Table 1 Comparison of theoretical and practical peak voltage

R_1	$R_2 = R_5$	R_3	$R_4 = R_6$	$C_1 = C_4$	$C_3 = C_5$	C_c	$V_p (\text{T})$	$V_p (\text{P})$
1 k	2.5 k	1.5 k	5 k	1p	1p	180fF	0.15	0.15
20 k	2.5 k	50 k	5 k	1p	3p	7.5fF	2.5 m	3 m
10 k	40 k	20 k	80 k	120f	200f	360fF	0.642	0.75
1 k	20	150	100	60	120f	180	0.173	0.17

C_c is varied from nano farad (nF) to atto farad (aF).

The peak voltage at the victim node is theoretically calculated using Eq. (1)

$$V_p = (C_c / t_r) * R_V \left(1 - e^{\frac{-t_r}{R_V * (C_C + C_V)}} \right) \quad (1)$$

The spice simulation is done for Fig. 2 for various values of R , C , and C_c along with the interconnect. Table 1 shows the comparison between theoretically and practical peak voltage.

3 Proposed Methodology

In digital circuits, the subthreshold current is normally viewed as the parasitic leakage which is ideally having no current. But in deep submicron technology, due to scaling in size of the transistor, the threshold voltage of the transistor is reduced. Because of the reduction in threshold voltages, the leakage current of the circuit is increased exponentially. Due to this, the leakage current in the circuit cannot be neglected. In this work, to reduce the leakage power consumption in the circuit, the authors propose multithreshold CMOS (MTCMOS) technique. MTCMOS is a power gating structure to reduce leakage current in the circuit.

MTCMOS reduces the leakage current during standby mode. When the device is in standby mode, low threshold voltage transistors are isolated from V_{dd} and ground by using high threshold voltage transistors. As a case study, NAND and NOR gates

Table 2 Comparison of leakage power for CMOS and MTCMOS based techniques

Use case	CMOS P_{leakage} (W)	MTCMOS P_{leakage} (W)	% Reduction
Inverter	0.02	0.02	–
NAND	0.04	0.0244	60%
NOR	0.046	0.0087903	19%

are designed using MTCMOS and compared with NAND and NOR gated designed using CMOS. Leakage current is calculated using the below equations.

$$S = V_{\text{th}} \ln 10 \quad (2)$$

$$I_o = \mu C_{ox} \frac{W}{L} V_{\text{th}}^2 e^{1.8} \quad (3)$$

$$I_{\text{subth}} = I_o 10^{\left[\frac{V_{\text{gs}} - V_t}{S} \right]} \quad (4)$$

DC analysis is performed to calculate the leakage current practically.

$$P_{\text{leakage}} = I_{\text{subth}} * V_{\text{dd}} \quad (5)$$

Power dissipated due to leakage current is obtained using Eq. (5). Table 2 shows the P_{leakage} comparison of the CMOS and MTCMOS based techniques. It is observed from Table 2 that using MTCMOS technique 60% of leakage power consumption for NAND gate and 18% leakage power consumption for NOR is gate observed. Since in deep submicron technology, leakage power is a significant factor. A reduction of leakage power is required to enhance circuit performance.

3.1 MTCMOS Repeater Insertion to Minimize Crosstalk

Two major impacts of crosstalk on the circuits are crosstalk-induced delay which causes propagation change in the circuit, and it results in degradation of signal by inducing voltage spikes. The authors have emphasized on repeater insertion technique to reduce crosstalk noise.

Repeater insertion: Repeater insertion divides the interconnects into equal sections. Each section is driven by equal size repeaters. The primary objective of the repeater insertion is to minimize the time required for the signal to propagate through a long interconnect. As the time delay is reduced and the signal is restored after each repeater, the crosstalk noise is also reduced. The authors propose MTCMOS based repeaters between long interconnects to overcome crosstalk noise and reduce power consumption see Fig. 4 (Table 3).

Fig. 4 MTCMOS based repeaters insertion

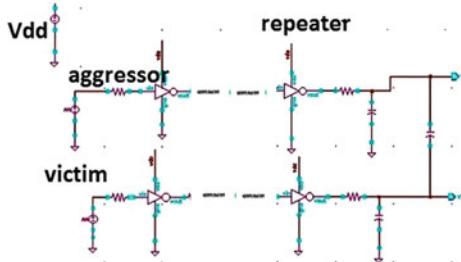


Table 3 Peak voltage values for output at victim node

V_{in}	V_p without MTCMOS repeaters	V_p with MTCMOS repeaters	Leakage power	
			CMOS	MTCMOS
1.8 V	0.15 V	1.73 V	17.872 nW	41.323 pW

4 Conclusion

The coupling capacitance (C_c) present in between the aggressor and the victim line causes an undesirable effect on the output of the victim line. The paper presents the design of repeaters using MTCMOS. Results show that MTCMOS reduces leakage power by 60–19% based on the design. Further the MTCMOS is used as repeaters to reduce the leakage power, voltage degradation, and effect of coupling capacitance. Results show that using MTCMOS as repeaters in long wire interconnect, the voltage degradation is reduced by 96%. The leakage power reduction is paid by increase in delay due to increase in no. of transistors in MTCMOS design.

References

1. H. Deogun, R. Rao, D. Sylvester, D. Blaauw, Leakage-and crosstalk-aware bus encoding for total power reduction, pp. 779–782. <https://doi.org/10.1145/996566.996776>
2. H. Fatemi, B. Amelifard, M. Pedram, Power optimal MTCMOS repeater insertion for global buses, in *ISLPED'07*, 27–29 August 2007, Portland, Oregon. Copyright 2007. ACM. 978-1-59593-709-4/07/0008
3. S.V. Siddamal, R. Hotkar, Design of long wire interconnect using MTCMOS repeater with low power and delay constraints, in *International Conference on Innovative Ideas in Engineering and Technology (ICIET-2018)*, RBYMEC, Ballari, Karnataka, 12–13 April 2018
4. H. Singh, R. Rao, K. Agarwal, D. Sylvester, R. Brown, Dynamically pulsed MTCMOS with bus encoding for reduction of total power and crosstalk noise. *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.* **18**(1), 166–170. <https://doi.org/10.1109/TVLSI.2009.2031290>
5. P. Karthikeyan, S. Mallick, Optimization techniques for CNT based VLSI interconnects—a review. *J. Circ. Syst. Comput.* **26**(03), 1730002 (2017)
6. H. Jiao, R. Wang, Y. He, Crosstalk-noise-aware bus coding with low-power ground-gated repeaters. *Int. J. Circ. Theory Appl.* **46**(2), 280–289 (2018). <https://doi.org/10.1002/cta.2378>

7. A. Karthikeyan, P.S. Mallick, High-speed and low-power repeater for VLSI interconnects. *J. Semiconductors* **38**(10) (2017)
8. R. Amira, M. Razif, S. Marwang, M. Maharum, A.H. Hasani, Z. Mansor, Mitigation techniques for crosstalk in ICs, in *The Electronic Packaging Interconnect Technology Symposium 2019. IOP Conference Series: Materials Science and Engineering*, vol. 701, 24–25 November 2019, Penang, Malaysia
9. R. Sridevi, P. Chandra Sekhar, B.K. Madhavi, Cross talk delay reduction in system on chip, in *Microelectronics, Electromagnetics and Telecommunications. Lecture Notes in Electrical Engineering*, vol. 471 (Springer, Singapore, 2018)

GPS and GSM Enabled Smart Blind Stick



Sourodip Ghosh , Moinak Bose , and Ankit Kudeshia

Abstract Blind sticks are the supporting devices to help visually impaired individuals in scanning their surroundings and identifying the obstacles without the need of assistance from external agents. The current mechanical form of the blind sticks needs upgradation for better support to the subject. Therefore, this paper proposes a modification to these sticks with the accession of sensors, microcontrollers, and buzzers that aids the subject better navigate their environment/surroundings and thus provides enhanced mobility experience. The sensors detect the hurdles and obstacles from a safe distance, and the microcontroller activates a buzzer to alert the user in case of threats. The model employs another unit to send location updates to kith and kins of the subject in emergency conditions with the help of GPS and GSM units. The microcontroller in the second module is connected to a sensor placed at his/her hand that monitors the health parameters such as pulse rate. Any abnormal changes in these readings result in an automated text message sent to the mentioned contacts along with the GPS location. The model is aimed to be a cost-effective and user-friendly device for aiding the visually challenged people that guarantees highly reliable navigation performance and greater user experience by offering various additional features.

Keywords Arduino UNO · Sensors · GPS · GSM · Smart devices

1 Introduction

According to an NCBI PubMed report [3], the number of visually impaired people in India is projected to be 31.6 million in 2020. Globally, this number has reached 285 million in the year 2019, and out of which, 39 million people have completely lost the ability to achieve optic recovery. The common causes of blindness include traumatic injuries, infections of the retina, macular degeneration, and diabetes. The

S. Ghosh · M. Bose · A. Kudeshia
KIIT University, Bhubaneswar, India
e-mail: sourodip.ghosh02@gmail.com

above reasons demand to support their vision with alternative means, if not medically. The various works in the literature towards offering aid to the visually challenged individuals have been presented next.

1.1 Related Works

Yusro et al. [10] proposed a system design of a white, GPS enabled smart environment explorer stick (SEES). This helped visually impaired subjects to roam around freely, both indoors and outdoors. Further, Jameson et al. [5] designed a wearable system for warning visually impaired subjects to avoid any possible collisions. In comparison with other features, this device throws a notification alert when any hazard is detected. It uses two ultrasonic transducers for locating objects. The authors also worked on optimizing the performance of the device in terms of obstacle range localization and reduction of false alarms maintaining low power consumption. Liarokapis et al. [7] use a combination of computer vision technique in conjunction with GPS technology to create an AR/VR interface which offers a very efficient navigation tool for the concerned subjects. Furthermore, Loomis et al. [8] conducted a test where visually impaired subjects were given a backpack containing computing facilities, each containing GPS modules, and headphones. They also attached a motion-detector module specifically for assessing head movements. A GPS enabled audio interfaced system was also integrated with it to facilitate the subject to hear and know its current location. Laurent et al. [6] made use of a SONAR model that mimicked the echolocation property of bats as the blind mobility aid. It enhanced the reliability of navigation by using methods involving spatial hearing. Several such directional aids for these subjects can be found in the works proposed in [1, 2, 4, 9].

The above systems have proved to offer an effective aid to the visually impaired subjects. However, employing them is practically infeasible due to the high computational resource requirements of the system and the financially exorbitant nature of the model. Hence, this paper proposes a smart stick model that is financially affordable for a larger section of such people with added features compared to the above-proposed models, thus giving them a better user experience and making the lives of vision-impaired subjects easy and smooth.

1.2 Contributions of the Paper

The previous methods involving smart sticks have scopes of incompatibility in terms of lesser features or cost-efficient medium, thereby compromising the accuracy of the device. Some products currently in the sale are more expensive thus less affordable in all levels of the community. This paper introduces concepts as follows.

1. Development of the smart stick using prominent sensors, buzzer, microcontrollers, a pair of GSM and GPS module.
2. The prototype of the proposed model is generated using the software, and the simulations are performed to showcase the working of the smart stick for different surrounding conditions.
3. A microcontroller-based circuit has also been designed that adds various features to the stick and automates its functionality which is also tested for its real-time implementation.
4. The proposed system allows more features incorporated in the stick than the currently available models, offering a lower cost, thus making it affordable to almost all sections of the people.

2 Proposed Smart Stick Model

This section details the various components of the proposed smart stick and their integration to get the final model. The first component is the set of sensors that are used for providing real-time information about different parameters. These include four different types of sensors that are ultrasonic sensors, infrared sensors, a water sensor, and the pulse sensor which capture the signals from the surroundings of the subject. Though a cost-efficient water sensor can be used by the implementation of wires, however, that would result in inaccurate detection disapproving the purpose of its presence. Therefore, a more sophisticated and slightly expensive sensor is used for water detection purpose which does not notably increase the cost of the model. For processing the sensor data to control the actions of the device, Arduino microcontroller has been used due to its re-programmable feature and also for being financially affordable out of all available microcontrollers such as Atmel. The use of the sound system while outdoor usage of the stick would always be limited to the subject's sense of hearing, as their focus would have to be on the voice command from the speakers. The target user might be less comfortable with this feature, and thus, a buzzer would be a more suitable option for them.

2.1 Components

The various components shown in Fig. 1 have diversity with reference to their usability that has been explained next.

Sensors Sensors pick up changes in environment due to the presence of external stimuli around the device and communicate it to the microcontroller. Thus, the sensors play an important role in the initial phase of the device with which it triggers functioning.

1. **Ultrasonic Distance Sensor (HC-SR04):** It uses ultrasonic sound to calculate the distance of obstacle in front of it. It emits an ultrasonic sound which then

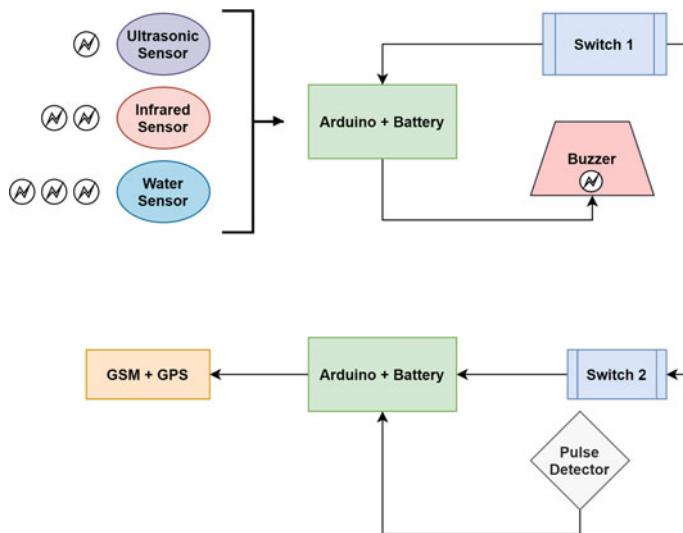


Fig. 1 Proposed system workflow

travels back after hitting the obstacle to the device. The period of the operation is calculated, and the distance is estimated. It works similar to the principle of sound navigation ranging (SONAR) and other such devices. Being ultrasonic in nature, the signal travels faster than normal sound and does not affect or get affected by the surrounding noise as well. With a range of around 20 metres, the sensor is perfect for detecting any sudden block in the path.

2. **Water Sensor:** This sensor has been used at the bottom of the stick to detect the presence of any kind of liquid on the walking surface. Making the user alert about any such presence would help avoid any such instance of slipping or falling.
3. **Pulse Detector:** Installed in the holding region of the stick connected to the wrist of the subject, this sensor stays attached with the handle of the stick. The users entangle their wrist around this belt for a better grip of the stick. This sensor would be connected with the GSM + GPS module of the device and would send a notification to the contacts with user's live location when any kind of change in pulse rate is observed.
4. **Passive Infrared Sensor (HC-SR501):** This sensor detects emission of light around it. The primary function of this sensor is to detect the presence of any sudden light source around the device. It can be in the form of a car, or a person. This sensor particularly makes the device effective for being used in darkness.

Buzzer and Notification Receiver Buzzers are mechanical, electro-mechanical, or piezoelectric devices which are typically used for notification purpose and has been used for the same here as well. With different buzzing patterns for the different sensors, the device receives a command from the microcontroller, informing the user about their surroundings. Global system for mobile communication (GSM)

(SIM900A) module reduces the data received and forms a digital packet of it to be sent to mobile devices, so that information stays loaded beforehand. Global positioning system (GPS) (NEO-6M) module collects data from satellite and stores the location data based on latitudes and longitudes in a format which when sent can be accessed by any common user to trace the module location. These two modules, along with the accession of the Arduino microcontroller, monitor the system location and send text messages when necessary.

2.1.1 Micro-controller (Arduino UNO R3)

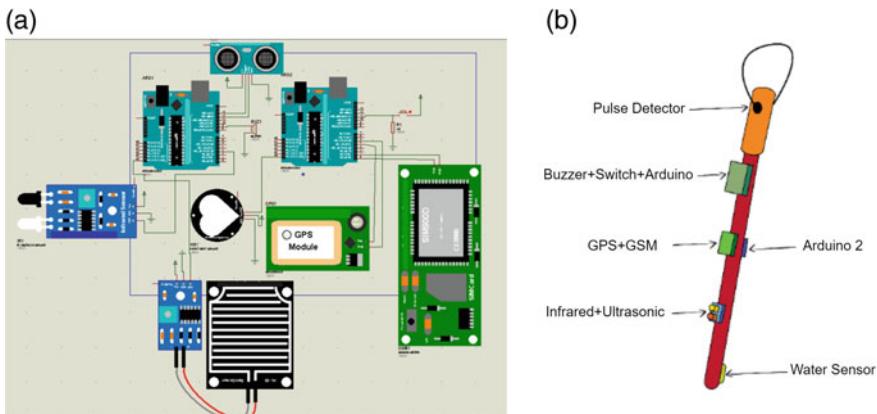
The microcontroller used here is the standard Arduino UNO microcontroller, which is an open-source board manufactured based on Microchip Atmega328P. It has 14 input and output pins in the form of both analog and digital. It has the extensive function of commanding the buzzer in one module and commanding the GSM and GPS in another module. It receives responses from the sensors, switch, and commands the other components to act according to the program installed in it. Being re-programmable, Arduino comes in handy for a large expanded region of usage. Arduino was considered due to these reasons; first, it serves as a cheap and sustainable microcontroller available at our disposal. Secondly, other controllers available contain in-built RAM which makes them both expensive and more prone to short-circuiting.

2.2 System Integration

The complete system can be divided into two extensive modules. With two modules each consisting of an Arduino UNO microcontroller of its own, the overall load of work gets distributed thus increasing the efficiency. One of the modules function with the ultrasonic sound sensor, passive infrared sensor, water sensor, and the Arduino UNO of this module is connected with the buzzer of the device which is activated with different forms of buzzing as a response of each different sensors. The other module is connected with a switch that is present on the handle of the stick and with a pulse detector that is present in the wrist band associated with the handle of the stick. Response from any of the sensors or the switch makes the GSM send a notification to the uploaded contacts in it along with the live location from the GPS. The circuit diagram of the modules integrated together is shown in Fig. 2a. The method involves the integration of a buzzer and sensors with a microcontroller device, and by taking advantage of GPS and GSM modules, more features are accessed, and the system is improved from the previous methods. Thus, the system proves to be more advanced and accurate when tested in a real-time scenario. The software used to build a simulation of the model is Proteus Professional 7. It is developed by Labcenter Electronics and is preferred by many developers. The activity of individual components is discussed as shown in Table 1. The modules are arranged in a way which allows correct positioning of all the components and would facilitate cost

Table 1 Description of activity of various components

Components	Status	Description
Ultrasonic sensor	HIGH	Detects sudden obstacle on path, notifies Arduino
Passive Infrared sensor	HIGH	Detects change in light intensity, notifies Arduino
Water sensor	HIGH	Detects presence of liquid, notifies Arduino
Pulse sensor	HIGH	Detects change in pulse rate, notifies Arduino
GPS module	HIGH	Receives live location
GSM module	HIGH	Sends live location; text to emergency list

**Fig. 2** a Circuit diagram for the proposed system. b Smart stick prototype

reduction. This ensures that the mounting of all devices on the stick stands proper, maintaining functionality with optimized performance. A simulation prototype of the final device is shown in Fig. 2b.

3 Conclusion

A smart walking stick for visually impaired people has been proposed in this work which comprises sensors integrated with the microcontrollers and the output devices such as buzzer and GSM/GPS-based notification receiver. The prototype of the device is presented which shows its comfortable design and the efficacy for its usage by the needy people. The integration of all the appropriately chosen I/O and microcontroller components makes the proposed system affordable and smart to use.

References

1. M.P. Agrawal, A.R. Gupta, Smart stick for the blind and visually impaired people, in *2018 Second International Conference on Inventive Communication and Computational Technologies (ICICCT)* (IEEE, 2018), pp. 542–545
2. D. Dakopoulos, N.G. Bourbakis, Wearable obstacle avoidance electronic travel aids for blind: a survey. *IEEE Trans. Syste. Man Cybern. Part C (Appl. Rev.)* **40**(1), 25–35 (2009)
3. L. Dandona, R. Dandona, R.K. John, Estimation of blindness in India from 2000 through 2020: implications for the blindness control policy. *Nat. Med. J. India* **14**(6), 327–334 (2001)
4. N. Dey, A. Paul, P. Ghosh, C. Mukherjee, R. De, S. Dey, Ultrasonic sensor based smart blind stick, in *2018 International Conference on Current Trends towards Converging Technologies (ICCTCT)* (IEEE, 2018), pp. 1–4
5. B. Jameson, R. Manduchi, Watch your head: a wearable collision warning system for the blind, in *SENSORS, 2010 IEEE* (IEEE, 2010), pp. 1922–1927
6. B. Laurent, T.N.A. Christian, A sonar system modeled after spatial hearing and echolocating bats for blind mobility aid. *Int. J. Phys. Sci.* **2**(4), 104–111 (2007)
7. F. Liarokapis, Location-based mixed reality for mobile information services. *Adv. Imaging-Fort Atkinson* **21**(4), 22–25 (2006)
8. J.M. Loomis, Y. Lippa, R.L. Klatzky, R.G. Golledge, Spatial updating of locations specified by 3-d sound and spatial language. *J. Exp. Psychol. Learn. Mem. Cogn.* **28**(2), 335 (2002)
9. S. Sharma, M. Gupta, Kumar, A., Tripathi, M., Gaur, M.S.: Multiple distance sensors based smart stick for visually impaired people, in *2017 IEEE 7th Annual Computing and Communication Workshop and Conference (CCWC)* (IEEE, 2017), pp. 1–5
10. M. Yusro, K. Hou, E. Pissaloux, H. Shi, K. Ramli, D. Sudiana, Sees: concept and design of a smart environment explorer stick, in *2013 6th International Conference on Human System Interactions (HSI)* (IEEE, 2013), pp. 70–77

Prediction of Speed for Smart Insulin Pump Utilizing Adaptive Neuro-fuzzy Inference System and ANN



J. V. Alamelu and A. Mythili

Abstract Smart infusion and insulin pumps are widely used these days. The prediction and decision making are indeed an important task within the cyber physical system environment. This paper is an attempt to provide the expert system for the optimal speed for the motor employed in infusion pump systems, according to the flow rate of the appropriate substance to be administered to the patient. The optimum flow of the dosage is necessary, and it can be performed by the appropriate speed of the motor. Hence, the motor speed and its revolution for the corresponding flow rate are calculated for all the glycemic conditions. In this research, the implementation of the artificial neural network (ANN) and adaptive neuro-fuzzy inference system (ANFIS) is utilized for prediction. The model is assessed by root mean square error (RMSE), mean square error (MSE), mean absolute error (MAE), mean absolute percentage error (MAPE). The performances in ANFIS has been compared with different membership functions such as trapezoidal, Pi, Gaussian, Gaussian bell, triangular membership function. The RMSE is obtained as $4.96899e-06$ and MAPE is $1.33e-06$ for trapezoidal membership function which is less when compared to the other membership functions as well as with ANN. The predictive decision analysis will be the supportive functionality for continuous subcutaneous insulin infusion (CSII) and continuous glucose monitoring (CGM) systems.

Keywords Infusion pump · Flow rate · ANFIS

J. V. Alamelu (✉) · A. Mythili
SENSE, VIT University, Vellore, Tamil Nadu, India
e-mail: jvalamelu@gmail.com

A. Mythili
e-mail: mythili.asaithambi@vit.ac.in

J. V. Alamelu
M S Ramaiah Institute of Technology, Bangalore, India

1 Introduction

An infusion pump is a device that dispenses fluids, nutrients, medications to the patient. These devices are designed to deliver the required volume of substance within a specified time duration precisely. The precise and timely delivery of the substance are a primary objective for patient health care [1]. Modernization and inventions of the latest technology in computing, control, and information technology lead to the findings of smart health care devices for monitoring and control of patient health. Much smart monitoring and delivery devices are available to treat different illnesses caused to the human being. The recent delivery devices like smart infusion and insulin pump devices indicate the supervisors, caregivers in the hospital environment about the dosage requirement, errors, and provide alarms. Smart infusion pumps are widely used as part of anesthesia machines, ventilator systems, artificial pancreas, and insulin pumping systems to treat diabetic conditions. These devices are widely adopted in cyber physical environment. Diabetic therapy or insulin treatment enables patients to control the blood glucose level [2]. Nowadays, to dispense insulin to the body, a smart wireless insulin pump is used by the diabetic patient to infuse the dosage as per the glucose level sensed by the device. The implementation of CSII and CGM systems have grown widely in diabetes care [3]. The support of these devices is used in clinical scenarios to monitor, control glycemic conditions, and risk of hypoglycemia for type 1 and type 2 diabetes is reduced [4]. The usage of smart devices in hospital scenario is very important as the caretaker should follow the guidelines as specified by the smart pump manufacturer and further follow the dosage requirement according to the patient's condition [5]. Research in continuous glucose monitoring, insulin infusion is done as per the data acquired from the patient with sensors connected to the human body as closed loop optimal control. The device assists in transmitting basal rates, carbohydrate ratio, infusion of insulin, correct dosage with insulin settings. This information will be available to the patient, care taker, and physician [6].

In a real-time situation, the smart infusion system has two subsystems along with an inbuilt controller which is programmed. One subsystem has blood glucose sensors and another has an actuation unit with pump and motor arrangement [7]. The optimal control of the actuation unit is needed to infuse the fluid. The device is a complete closed loop control with decision making for the optimal amount of infusion of the substance to the human body [8]. From the literature, it is observed that the decision making of the insulin to be infused as per the blood glucose sensing value has been focused predominantly. To implement this different modeling decision making, optimal control strategies and methods are deployed [9, 10]. The focus on flow rate along with the decision of the optimal speed of motor used in the actuation unit is still immature. In this paper, an attempt has been made to perform the prediction and decision making of the optimal speed of the motor with ANN and ANFIS. The key concern is every drop count from infusion pump counts is related to flow rate and the precise operation of the electric motors for the pumps.

2 System Design

The closed loop control of the smart infusion-based insulin pump is shown in Fig. 1. It consists of subsystem 1, a sensing unit with a blood glucose sensor that senses the glucose level from the patient, accessed by the controller connected to it. The controller is programmed to perform decision making optimally to regulate the glucose level by delivering the required dosage to the patient via the subsystem 2 which is an actuation unit. The actuation unit is the focus of this paper as the optimal speed control and the decision of the required speed for different flow rate conditions are controlled. The information will be further updated to the caretaker, physician for further accessing of electronic health record (EHR) for the patient.

2.1 Methodology

To implement the decision making for the speed of the DC motor [11] for the required flow rate of the insulin, the predefined conditions from the insulin infusion protocol is considered for different glycemic conditions. For these conditions, the speed of the DC motor to actuate the pump is calculated. In this paper, the data with different glycemic values, insulin flow rate, and the rotational speed of the motor are considered. The inputs are blood glucose level (BG) from the patient, the number of units of insulin (In) to be infused to the patient in terms of its flow rate, and the output is optimal pump speed with DC motor (op) integrated into the unit is determined. The dataset has been created based on the insulin protocol as prescribed by the practitioners. The inputs are given to the ANN and ANFIS. The test and train data have been created, and further, the prediction and error have been identified with ANFIS [12] and compared with the basic ANN model [13].

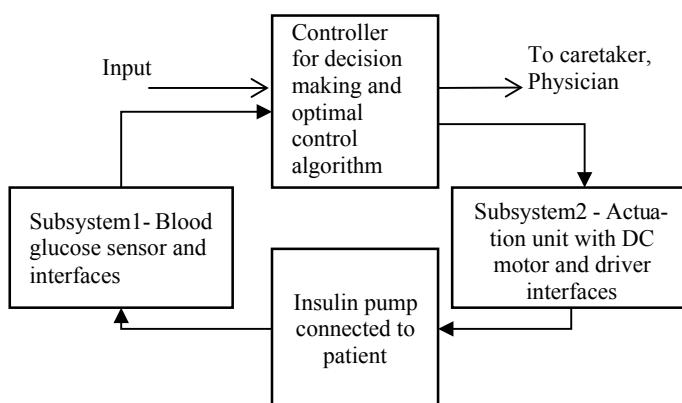


Fig. 1 Closed loop control of smart insulin pump

ANN Model

ANN is a network made up of many neurons that are interconnected and effectively work to transfer data to each other to perform a specific task [14]. In this paper, the backpropagation algorithm has been considered for testing the infusion data of the insulin pump. The algorithm uses a forward and backward pass. The input is given to the neurons of the network and propagates through every layer, and the actual response is obtained as an output set. The weights are fixed in the forward pass and adjusted in the backward pass based on the error correction rules incorporated and the error is found. The training and prediction has been observed, MSE has been plotted.

ANFIS Model

ANFIS is a multilayer feed-forward network in which each node (neuron) performs a particular function on incoming signals. Determining the fuzzy logic structure involves model inputs, membership function, and fuzzy rules generation [15].

The generic ANFIS model structure with the Sugeno system has been considered for the implementation [16]. ANFIS is a hybrid learning algorithm with backward pass learning and the forward pass. In the forward pass, signals move from layer 1 to layer 4 and the output parameters such as motor speed are determined. The error rates are propagated backward in backward pass further the parameters are updated using the gradient descent method [17].

Smart Insulin Performance

In this work, an adaptive neuro-fuzzy interface strategy is utilized. Neuro-fuzzy system has the properties of neural networks and fuzzy expert system [18]. Neural networks learn the data, interpret the acquired knowledge, as meaning associated with each neuron and weight. The fuzzy logic models are combined with the fuzzy set with its membership function, rules for further decision making of the output where it is the optimal speed of DC motor in this work. To carry out this neuro-fuzzy-based closed loop control system is designed and shown in Fig. 2. The two inputs BG, In are provided to the neuro-fuzzy interface system to perform computations to obtain the speed of pumping unit.

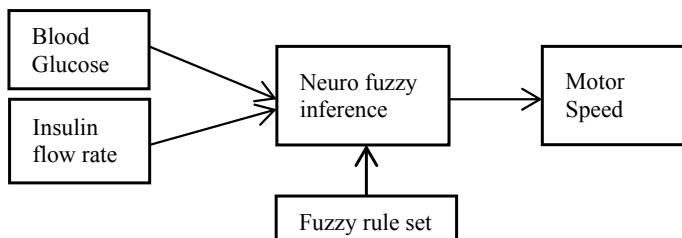


Fig. 2 Neuro-fuzzy model for insulin pump

The inputs BG and In are the inputs which are x_1, x_2 as per ANFIS provided to the neural network model where the data is trained, tested with ANFIS command, and evaluated to obtain the error [19]. The model can be developed to calculate the performance of the proposed system with different membership functions. In this work pi, triangular, Gaussian, g-bell, and trapezoidal membership functions are considered to determine the required statistical parameters. Rules are formed as per the ranges of blood glucose levels. Impaired fasting glucose is present when the fasting level is ≥ 110 and ≤ 125 mg/dl and the 2-h value is < 140 mg/dl. The insulin required for diabetes is 1 unit to 6 units for the range of glucose is from 180–400 mg/dl. If it exceeds these range, physician instruction will be followed. These conditions are framed as a fuzzy rule set and decision making of the speed of the pump is controlled optimally. The relationship between the flowrate and speed of the motor is given as,

Speed of the motor = flowrate*10 rev per ml/60 min per h. Here, the flow rate indicates the number of units to be infused. Based on this, the complete adaptive neuro-fuzzy inference engine is developed for the insulin pump system as per the infusion protocol [20]. The knowledge-based rules are provided to obtain the desired speed of the pumps' motor and in-turn infuses the desired flow of the fluid or substance to the patient. With the error obtained from the ANFIS model, the following statistical metrics are found to evaluate the performance of the model. The statistical metrics considered here are:

$$\text{Mean Square Error MSE} = \frac{1}{n} \sum_{i=1}^n (x_i - y_i)^2 \quad (1)$$

$$\text{Root Mean Square Error RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - y_i)^2} \quad (2)$$

$$\text{Mean Absolute Error MAE} = \frac{1}{n} \sum_{i=1}^n |x_i - y_i| \quad (3)$$

$$\text{Mean Root Error MRE} = \sum_{i=1}^n ((x_i - y_i)/x_i) \quad (4)$$

$$\text{Mean absolute Prediction Error MAPE} = \frac{1}{n} \sum_{i=1}^n ((x_i - y_i)/x_i) * 100 \quad (5)$$

where x_i is the actual output and y_i is the predicted output and n is the number of the outputs.

3 Results and Discussion

The statistical metrics evaluated for the proposed model for the smart insulin pump with different membership functions are shown in Table 1. It is inferred that the trapezoidal membership function provides lesser value in terms of MAE, MSE, RMSE, MRE, MAPE, and better prediction accuracy is obtained which can be inferred from MAPE. The trained data based on glucose for the respective speed of the insulin pump resultant with the test data, and the predicted data plot is shown in Fig. 3. The outputs shown here are based on the trapezoidal membership function. The observed and the predicted values are almost nearer to each other. The MSE plot for ANN is also shown in Fig. 3. The performance error obtained on validation is $5.73e-11$ which is higher when compared to the MSE value obtained from ANFIS with trapezoidal MFs. The statistical parameters are calculated and compared with ANN and different membership functions in ANFIS.

Table 1 Statistical parameters for different MFs in ANFIS, ANN model

	ANFIS—pi MFs	Gaus-mf	Trap-mf	g-bell mf	Tri mf	ANN model
MAE	3.14×10^{-5}	3.29×10^{-5}	7.41×10^{-7}	5.18×10^{-6}	0.00044	1.17×10^{-6}
MSE	4.45×10^{-8}	4.70×10^{-8}	2.47×10^{-11}	1.21×10^{-9}	9.007×10^{-6}	5.73×10^{-11}
RMSE	0.000213	0.0002	4.97×10^{-6}	3.48×10^{-5}	0.0030	9.51×10^{-6}
MAPE	5.68×10^{-5}	5.89×10^{-5}	1.33×10^{-6}	9.37×10^{-6}	0.0088	12.309
MRE	2.555×10^{-5}	2.64×10^{-5}	6.025×10^{-7}	4.21×10^{-6}	0.000363	8.1240

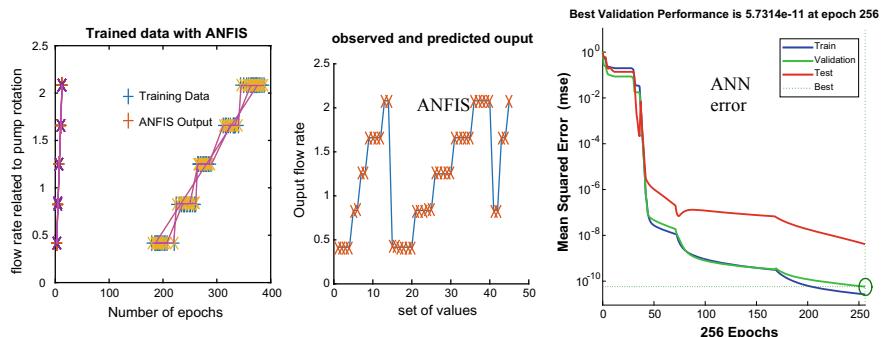


Fig. 3 Trained data with observed and predicted output for ANFIS and ANN

4 Conclusions

Precise movement of the motor is needed and hence for a proper clinical decision of the device expert system for the diabetic patient is implemented. In this work, for different glycemic conditions, its corresponding dosage, the flow rate is considered and further speed of the motor is calculated. The results show that the trapezoidal membership function achieves lesser statistical errors and better prediction accuracy when compared to other membership functions as well as with ANN model. It is further observed that the prediction accuracy is better for trapezoidal membership function. This method could be incorporated for all types of infusion pumps in the cyber physical system incorporated with the health care system. The modeling and prediction can be incorporated for smart medical pumps that communicate to the cloud.

References

1. D. Yu, K. Hsu, J.H. Kim, P. Delaurentis, Infusion pump informatics approach to quantify impact of alerts and alarms on healthcare delivery, pp. 681–685 (2017)
2. A.Y. Ben Sasi, M.A. Elmalki, A fuzzy controller for blood glucose-insulin system, **2013**(May), 111–117
3. J.C. Pickup, Y. Reznik, A.J. Sutton, Glycemic control during continuous subcutaneous insulin infusion versus multiple daily insulin injections in type 2 diabetes: individual patient Data meta-analysis and meta-regression of randomized controlled trials. *Diabetes Care* **40**(5), 715–722 (2017)
4. G.E. Umpierrez, D.C. Klonoff, Diabetes technology update: use of insulin pumps and continuous glucose monitoring in the hospital, pp. 1–11 (2018)
5. G. Cocha, O. Rodriguez, L. Codapli, C. Amorena, H. Mazzeo, Intelligent insulin pump design, pp. 7–10 (2018)
6. H. Anhalt, N.J.V. Bohannon, Insulin patch pumps: their development and future in closed-loop systems. *Diabetes Technol. Ther.*, **12**(S1), S-51-S-58 (2010)
7. V. Iacovacci, L. Ricotti, P. Dario, A. Menciassi, Design and development of a mechatronic system for noninvasive refilling of implantable artificial pancreas. *IEEE/ASME Trans. Mech.* **20**(3), 1160–1169 (2015)
8. E. Engineering, An automatic insulin infusion system based on LQG control technique Akshaya Kumar Patra and Pravat Kumar Rout **17**(3), 252–275 (2015)
9. C. Series, Fuzzy logic implementation for diagnosis of Diabetes Mellitus disease at Puskesmas in East Jakarta Fuzzy logic implementation for diagnosis of Diabetes Mellitus disease at Puskesmas in East Jakarta (2018)
10. A.K. Patra, P.K. Rout, Adaptive continuous-time model predictive controller for implantable insulin delivery system in Type I diabetic patient (2016)
11. P. Santhosh, P. Vijayakumar, Performance study of BLDC motor used in wireless medical applications. *Wirel. Pers. Commun.* **94**(4), 2451–2458 (2017)
12. N. Mathiyazhagan, Soft computing approach for predictive blood glucose management using a fuzzy neural network, in *2014 IEEE Conference Norbert Wiener 21st Century*, pp. 1–3, (2014)
13. M. Heidari, H. Homaei, Design a PID controller for suspension system by back propagation neural network. *J. Eng. (United States)*, **2013** (2013)
14. M. Kaveh, V. Rasooli, R. Amiri, ANFIS and ANNs model for prediction of moisture diffusivity and specific energy consumption potato, garlic and cantaloupe drying under convective hot air dryer. *Inf. Process. Agric.* **5**(3), 372–387 (2018)

15. S. Alby, A prediction model for type 2 diabetes using adaptive neuro-fuzzy interface system (2018)
16. N. Mathur, I. Glesk, A. Buis, Comparison of adaptive neuro-fuzzy inference system (ANFIS) and Gaussian processes for machine learning (GPML) algorithms for the prediction of skin temperature in lower limb prostheses. *Med. Eng. Phys.* **38**(10), 1083–1089 (2016)
17. E. Habibi, M. Salehi, G. Yadegarfar, A. Taheri, Optimization of the ANFIS using a genetic algorithm for physical work rate classification. *Int. J. Occup. Saf. Ergon.* **0**(0), 1–8 (2018)
18. M. Ün, Ç. Çiftcioglu, Position and speed control of infusion pump actuator for biomedical applications **0869**(6), 19–25 (2018)
19. D.U. Campos-Delgado, M. Hernández-Ordoñez, R. Femat, A. Gordillo-Moscoso, Fuzzy-based controller for glucose regulation in type-1 diabetic patients by subcutaneous route. *IEEE Trans. Biomed. Eng.* **53**(11), 2201–2210 (2006)
20. S. Tiwari, R. Babbar, G. Kaur, Performance evaluation of two ANFIS models for predicting water quality index of river Satluj (India) **2018** (2018)

Curvelet Transform and ISODATA Thresholding for Retinal Vessel Extraction



Sakambhari Mahapatra, U. R. Jena, and Sonali Dash

Abstract Researchers says eyes are good indicator of many diseases like Diabetic retinopathy, Glaucoma, Hyper tension, cardiac disease and many age related abnormalities. Here we aim to put forward a reliable, fast and computerized method to get the vessel network of fundus image, that can assist the ophthalmologist to diagnose the disease in early stage and also help in keeping the track of progress of treatment. Here curvelet transform is applied to the green channel of retinal image. It highlights the curves and edges of the vessel. Then CLAHE is implemented to improve the contrast and take care of the background inhomogeneity. Then the vessels are extracted by a simple ISODATA thresholding. This method gives 95.12% accuracy and average sensitivity, precision and specificity values 72.14%, 73.52% and 97.35% respectively. The execution time is also very low i.e. 1.36 s. This approach is very fast and shows a very high value of accuracy while retaining comparable values of sensitivity specificity and precision.

Keywords Curvelet transform · ISODATA thresholding · Retinal image

1 Introduction

Retinal vessel plays important role in early diagnosis and treatment of many diseases like Glaucoma, Diabetic Retinopathy, Cataract, Macular degeneration and Hyper tension [1]. The changes in vessel width, branching angle, tortuosity and shape assist

S. Mahapatra (✉) · U. R. Jena

Department of Electronics and Telecommunication, Veer Surendra Sai University of Technology, Burla 768018, India

e-mail: mahapatra.shakambhari@gmail.com

U. R. Jena

e-mail: urjena@rediffmail.com

S. Dash

Department of Electronics and Communication Engineering, Raghu Institute of Technology, Vishakhapatnam, India

e-mail: sonali.isan@gmail.com

the ophthalmologists in diagnosing disease and tracking the effect of treatments. However there are numerous challenges which interrupt the processes of getting the vessel network structure from the fundus image, such as non-vessel structures like optic disk, fovea, epithelium, exudates, the background inhomogeneity. Although manual segmentation of vessels is possible but it is time consuming, tedious and error prone specially for large no of cases. Hence there is need for automatic blood vessel extraction.

Recently various authors have suggested numerous approaches for blood vessel extraction by taking various databases. In almost every method the fundus image is first pre-processed to get the enhanced image and then a suitable segmentation method is implemented to get the vessel network. Usually pre-processing helps in removing the noise, making the background homogeneous and improving the contrast between the vessel and non-vessel region and also highlighting the edges and fine details of the vessels. These enhancement methods can be pixel based like: gamma correction, histogram equalization, morphological operations or it may be transform based like: DWT, Curvelet transform, Gabor transform etc. Similarly various segmentations methods are also implemented to extract the vessels in retinal image. Thresholding methods are commonly implemented to separate the vessel and non-vessel regions, however SVM, and other supervised methods are also been proposed for vessel segmentation. In [2] the authors implemented CLAHE and Gaussian filter to enhance retinal image and modified iterative self-organizing data analysis technique (MISO-DATA) for extracting blood vessel. A supervised classification method through ANN has been implemented by Dharmateja et al. [3], the features obtained by Zernike Moments co-efficient of fundus image are the inputs to ANN for classification. In [4] Zhang et al. have segmented the retinal vessels by thresholding the Match Filter response of the image. The threshold value is obtained through the image response to first-order derivative of Gaussian(FDOG). The authors in [5] have combined CLAHE and anisotropic diffusion filter for fundus image enhancement and Kirsch's template for segmentation. Alli et al. [6] demonstrated the effectiveness of Gabor filter for fundus image enhancement. The Gabor filtered image is segmented by Otsu method, ISODATA and K-mean clustering approach. Frangi matched filters are applied for fundus image enhancement and AdaBoost classifier for vessel network extraction is carried out by N Memari et al. [7]. In [8] CLAHE improved retinal image combined with morphological opening operation is segmented using ISODATA method to get the blood vessel tree. The authors in [9] have suggested adaptive local thresholding for large vessel extraction and the thin vessels are extracted by Support Vector Machine (SVM) and a tracking growth is implemented get the whole vasculature.

Azzopardi et al. in [10] have implemented two B-COSFIRE filters namely symmetric and asymmetric on retinal image and the two responses are summed up and thresholded to produce the blood vessel structure.

This paper presents a three step analysis of retinal vessel extraction-Image pre-processing, segmentation and post processing. In the pre-processing step the ability of Curvelet transform to highlight the edges along the curvature is explored to enhance the vessel. Then the segmentation is carried out in two steps first the optic disk is draw out by subtracting the morphological opening operated figure from the enhanced

figure and then applying a simple ISODATA thresholding, the image is binaries. Finally a morphological cleaning is done to remove the small artifacts and to produce the final segmented output. The remaining part of the paper is outlined as follows. The proposed vessel segmentation work is discussed in Sects. 2 and 3 does the result analysis and Sect. 4 culminates the report.

2 Proposed Method of Vessel Segmentation

The retinal vessel segmentation is carried out in three successive stages Pre-processing: Segmentation and Post processing. Figure 1, illustrates the different steps of this approach in sequence.

2.1 Pre-processing

Because of non-uniform illumination and low quality accusation device the captured fundus image suffers from low contrast, blur, intensity inhomogeneity and poor visibility of thin vessels. Hence image pre-processing becomes an elementary step before segmentation. Here Curvelet transform followed by Contrast limited adaptive histogram equalization (CLAHE) is implemented as a pre-processing step to improve the contrast of fundus figure. The green channel of fundus image shows better contrast, so it is implemented for image enhancement and further segmentation process. Curvelet transform [11] is a multi-scale scale geometrical transform. It can represent the edges and curve singularities much more efficiently as compared to traditional wavelet transform, Rizlet transform and others. Hence Curvelet transform seems to be more suitable for enhancing vessel structure edge. The fast discrete Curvelet transform (FDCT) is implemented in two ways unequally spaced fast Fourier transform and wrapping function. Although both the methods give similar result, rapping based FDCT is faster and easier so here wapping based method is implemented. In general the discrete Curvelet transform can be expressed as [12]:

$$(s, \theta, k_1, k_2) = \sum_{0 < i < M} f(i, j) \varphi s, \theta, k_1 k_2(i, j) \quad 0 < j < n \quad (1)$$

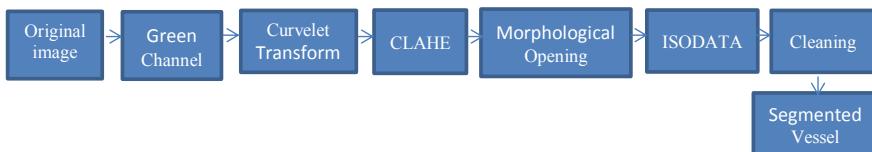


Fig. 1 Flow diagram of the suggested approach

Here ‘ s ’ is the scale or no of decomposition level, ‘ θ ’ is the orientation, ‘ $k1$ ’ and ‘ $k2$ ’ are spatial location of Curvelet, and ‘ $f(i,j)$ ’ is the image in spatial domain. As the decomposition level increases the Curvelets becomes thinner and sharper. The Curvelet coefficients are modified to enhance the edges in an image, and improves the image contrast. Usually the coefficients are modified through a nonlinear function in such a way that details of the small amplitude are enlarged at the expense of the larger ones. Then the inverse FDCT with the modified coefficients gives the enhanced image. Thus image enhancement through Curvelet transform consists of the following steps.

1. Find the Curvelet coefficient $C_{s,\theta}$ of the green channel of the image for different orientation (θ) and level of decomposition (s).
2. For decomposition level $s = 1$, set all coefficients to zero.
3. For $s = 2, 3 \dots 6$, multiply 2 with each coefficient $C_{s,\theta}$.
4. Now reconstruct the enhanced image from the modified Curvelet coefficients by applying inverse FDCT.

It emphasizes the edges and curve singularities of the image in different orientation and improves the contrast as well.

The distinction of the image is further improved by CLAHE. It improves the contrast in the vicinity of a given pixel value called the tiles and then the adjacent tiles are combined by binary interpolation to remove the artificial boundaries. It takes care of the over amplification of contrast and reduces the amplification of noise. The output of the Curvelet transform and CLAHE are illustrated in Fig. 2.

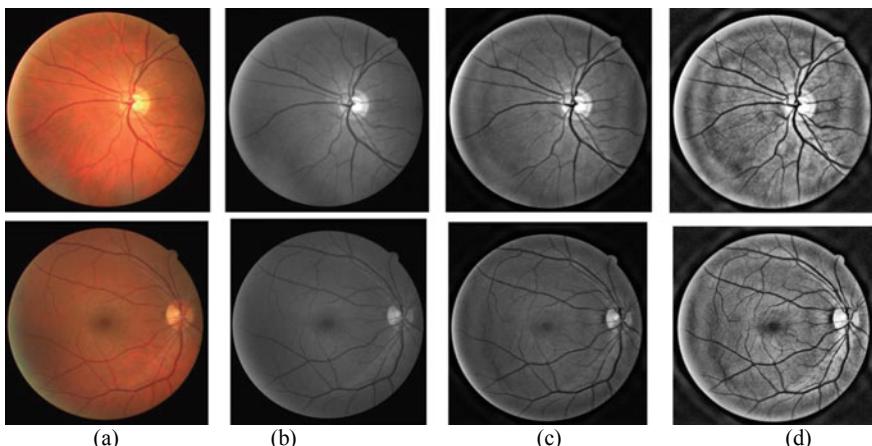


Fig. 2 Image output at different stages. **a** Original Image, **b** Green Channel **c** Curvelet enhanced image **d** CLAHE output

2.2 Segmentation

The blood vessels segmentation is carried out in two steps. First the optic disk is removed through morphological operation and then the vessels are extracted by ISODATA thresholding [2]. To remove the optic disc morphological opening operation is done on the enhanced image with a suitable structuring element. Disk shape structuring element (SE) with 8 pixcel size are found to be suitable for the above purpose as it matches with the shape of the optic disk. The opening operation on image ' f ' with SE ' B ' is simple erosion followed by dilation operation and is defined below.

$$f^\circ B = (f \ominus B) \oplus B \quad (2)$$

The optic disk removed image is thresholded using ISODATA which can technique can efficiently trace the vessels and distinguish them from the background. The ISODATA thresholding processes as follows:

1. Choose the average gray value of the image as the initial threshold value T_0
2. Using T_0 , the image is divided into two regions, (C_1 and C_2). C_1 : consist of all pixels with gray value lesser than T_0 . And C_2 consist of all pixels with intensity value higher or equal to T_0 .
3. Compute the mean gray values of the region C_1 and C_2 .
4. Calculate the new threshold value T_i

$$T = \frac{m_1 + m_2}{2} \quad (3)$$

Here m_1 and m_2 are the mean value of the region C_1 and C_2 respectively.

5. Repeat the steps from 2 to 4, with the current threshold value until the difference in two successive T_i is below some predefined value.

2.3 Post processing

It is observed that the thresholded image contains some unwanted pixels forming dots and very thin lines emerged as noise, which can be misclassified as noise. So a morphological opening operation efficiently removes these out growths and produces the final segmented image. Figure 3 shows the segmented image and post processed image and compares it with the ground truth.

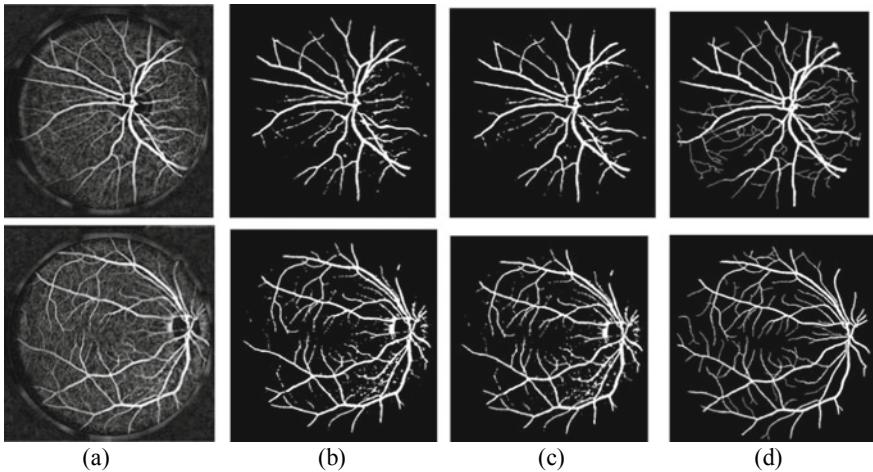


Fig. 3 Segmented image compared to Ground truth image. **a** Optic disk removed **b** ISODATA thresholded Image **c** After Post processing **d** Ground truth image

3 Result and Discussion

The suggested method is analysed and apprised in DRIVE data base. The data base has two sets of images training data base and testing data base. Each set contains 20 colour images, corresponding mask and the ground truths. Here the testing data set images are used. The segmented image of the proposed approach is compared with the ground truth and analysed using five different performance measures i.e. Sensitivity, Accuracy, Precision and Specificity [13].

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (4)$$

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (5)$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (6)$$

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}} \quad (7)$$

TP: No of times a pixel is correctly identified as a vessel is termed as True positive
 TN: No of times it is correctly identified as a background is termed as true negative

FP: Number of times a pixel is incorrectly identified as vessel is termed as false positive
 FN: Number of times a pixel is incorrectly identified as background is false negative

Sensitivity quantifies the techniques ability to detect the vessel pixel correctly, while specificity is a measure of the ability to detect the background pixels and the

accuracy gives conformity of the segmentation result. MCC is the percentage of rate of binary categorization. Table 1 shows the values of this parameter for the proposed approach on DRIVE data base. The proposed method archives average Sensitivity 72.14, Accuracy 95.12, Precision 73.52, Specificity 97.35. Table 2 shows a comparison of the proposed approach with few state of art methods discussed here. The accuracy obtained through the proposed approach is higher as compared to the methods

Table 1 Performance measures of the proposed approach on DRIVE data base

Images	Accuracy	Sensitivity	Precision	Specificity
Image 1	95.7771	73.6175	77.8484	97.9479
Image 2	95.8774	72.8322	84.7656	98.5066
Image 3	94.6715	75.1528	72.4311	96.8327
Image 4	96.0898	64.6735	90.0092	99.2727
Image 5	95.7407	64.3892	86.7277	98.9814
Image 6	94.9451	67.2593	77.7994	97.9305
Image 7	94.4411	71.2457	68.9542	96.7739
Image 8	93.8941	75.4518	61.9111	95.6302
Image 9	95.0224	72.1813	68.2363	97.0368
Image 10	96.085	65.9817	82.9614	98.7847
Image 11	94.4102	70.4323	68.1796	96.7679
Image 12	95.3155	73.0186	72.8065	97.4226
Image 13	95.1103	62.2803	83.5142	98.6678
Image 14	93.1355	82.3856	55.0427	94.0811
Image 15	94.0214	78.2417	55.8778	95.2377
Image 16	95.4882	70.9174	77.2468	97.9268
Image 17	94.2272	74.4399	63.4774	96.0514
Image 18	95.3849	75.6541	69.0559	97.0828
Image 19	96.7078	79.9715	80.2677	98.2217
Image 20	96.0538	72.7467	73.3666	97.9038
Average	95.11995	72.14365	73.52398	97.35305

Table 2 Comparison of parameters for various methods with the proposed method

	Accuracy	Sensitivity	Precision	Specificity
Match filter and Gaussian filter [4]	93.82	71.20	–	–
B-COSFIRE [13]	94.22	74.99	–	96.21
Gabor +ISODATA [6]	94.48	73.7	–	97.5
CLAHE + ISODATA [8]	94.6	67.5	82.1	98.8
Zernike features [3]	94.86	62.98	–	98.39
Proposed approach	95.119	72.143	73.523	97.353

discussed here. The methods in [4, 6, 8] and [13] implement thresholding techniques for vessel extraction and the method in [3] is a supervised method which implements ANN for vessel and background classification. All these approaches uses transfer methods for fundus image enhancement. The B-COSFIRE [13] shows higher value of sensitivity but the accuracy and specificity values are lower. The Gabor transforms and ISODATA thresholding method in [6], though shows a comparable value of sensitivity and specificity but the accuracy obtained is lower as compared to the suggested approach. The ISODATA thresholding method on CLAHE enhanced image produces lower accuracy of segmentation and lower value for sensitivity parameter also. However Curvelet transformed fundus image in combination with CLAHE produces better contrast enhanced output and when segmented through ISODATA thresholding produces high accuracy and comparably good values of sensitivity, precision and specificity. This method is also very time efficient. It achieves the segmentation in 1.36 s.

4 Conclusion

Here we proposed a new and efficient approach of retinal vessel extraction technique through Curvelet transform and ISODATA thresholding. The Curvelet transform is implemented to highlight the edges and curvature of retinal vessels in fundus images. A simple co-efficient manipulation i.e. equating the first level decomposition coefficients to zero and doubling other higher level Curvelet coefficients adds novelty to the approach and sharpen the vessel boundaries efficiently. The segmentation results on DRIVE data set shows an average accuracy of 95.12, and average sensitivity, precision and specificity values 72.14, 73.52 and 97.35 respectively. The segmentation time is also quite low i.e. 1.36 s. Hence this method gives a reliable segmented result in very few seconds hence can be treated as a novel alternative for vessel network extraction. In future optimal curvelet co-efficient can be used to better represent the curves and singularities. Further this method can be extended for disease identification and artery and vein classification also.

References

1. R.J. Chalakkal, W.H. Abdulla, Improved vessel segmentation using curvelet transform and line operators, in *Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*. IEEE (2018)
2. K.B. Khan, et al. An efficient technique for retinal vessel segmentation and denoising using modified ISODATA and CLAHE. IIUM Eng. J. **17**(2), 31–46 (2016)
3. D. Adapa, et al., A supervised blood vessel segmentation technique for digital Fundus images using Zernike Moment based features. Plos one **15**(3), e0229831 (2020)
4. B. Zhang, et al. Retinal vessel extraction by matched filter with first-order derivative of Gaussian. Comput. Biol. Med. **40**(4), 438–445 (2010)

5. J. Dash, N. Bhoi, *Retinal blood vessel extraction using morphological operators and Kirsch's template* (Soft Comput. Sig. Proc. Springer, Singapore, 2019), pp. 603–611
6. A. Ali, A. Hussain, W.M.D.W. Zaki, Segmenting retinal blood vessels with Gabor filter and automatic binarization. *Int. J. Eng. Technol. (UAE)* **7**(4), 163–167 (2018)
7. N. Memari, et al. Supervised retinal vessel segmentation from color fundus images based on matched filtering and AdaBoost classifier. *PloS one* **12**(12), e0188939 (2017)
8. J. Dash, N. Bhoi, Retinal blood vessels extraction from fundus images using an automated method, in *4th International Conference on Recent Advances in Information Technology (RAIT)*. IEEE (2018)
9. L. Xu, S. Luo, A novel method for blood vessel detection from retinal images. *Biomed. Eng. Online* **9**(1), 14 (2010)
10. G. Azzopardi, et al., Trainable COSFIRE filters for vessel delineation with application to retinal images. *Med. Image Anal.* **19**(1), 46–57 (2015)
11. S.S. Kar, S.P. Maity, Blood vessel extraction and optic disc removal using curvelet transform and kernel fuzzy c-means. *Comput. Biol. Med.* **70**, 174–189(2016)
12. D.L. Donoho, M.R. Duncan, Digital curvelet transform: strategy, implementation, and experiments. *Wavelet applications VII*. vol. 4056. International Society for Optics and Photonics (2000)
13. J. Dash, N. Bhoi, An unsupervised approach for extraction of blood vessels from fundus images. *J. Digit. Imaging* **31**(6), 857–868 (2018)

Design a T-Shape Cantilever Beam Using by Scilab and COMSOL



Vasagiri Suresh, Burra Rajesh Kumar, and Vankara Jyothi

Abstract Cantilever plays a major part in the detection of gases, upgrading existing sensors with fabricating technologies. The paper describes displacement of piezoelectric cantilever sensor of a T-shaped for MEMS technology for chemical/biodetection. The microelectromechanical systems (MEMS) sensors give compactness and robustness as well as enhanced lifespan. Thus, the concept of the microcantilever beam layered with piezoelectric materials is to monitor and analyze the density of gases. Various properties for cantilever beam have been mentioned in this paper. Configurations of cantilever beams are simulated and designed with COMSOL Multiphysics and theoretical simulations are made in Scilab. The design was built and computed to determine the corresponding displacement for the pressure being applied. Analysis is repeated by changing its structure with same measurements. Deflections are measured and tabulated to choose the optimal configuration for good sensitivity of a gas sensor.

Keywords MEMS sensors · Piezoelectric materials · Micro cantilever

1 Introduction

Air contamination poses a major danger to climate. It is the existence of harmful pollutants and other contaminants that modify the air quality. Pollution levels impact standard of living including illnesses such as skin inflammation, bronchitis, emphysema, asthma, and muscle weakness. This impacts not only humans and even plants by destroying their chlorophyll and disturbing the photosynthesis cycle. To track the accumulation of such pollutants on an ongoing basis, it is useful to position an apparatus in the location from where these gases emit. The advancement of IC technology

V. Suresh (✉)
SRF, GITAM (Deemed to be University), Visakhapatnam, India
e-mail: svasagir@gitam.edu

B. Rajesh Kumar · Vankara Jyothi
Associate Professor, GITAM (Deemed to be University), Visakhapatnam, India

opened the door for various developments with space-saving benefit. MEMS also developed a new system for high precision manufacture of smaller and more flexible instruments. This also provides real-time analysis through periodic transmission of necessary data. MEMS often offers a wide base of materials depending on the application needed. Microcantilever beam developed and calculating the target analyte by covering it with a film of sensing materials. Once the gas is moved across the beam to be measured, it becomes adsorbed by the sensing sheet. The piezoelectric material with beam uses static measuring method, in which the displacement [1] in turn produces pressure on the beam. It makes displacement shifts depending on the accumulated mass on beam.

Briscoe and Dunn [2] described “electric energy which accumulates with response for the mechanical stress applied to materials with non-centrosymmetric structures,” as piezoelectricity whereas Erturk and Inman [3] described piezoelectricity that “a type of correlation with electrical and mechanical behaviours in crystals and ceramics of any sort.” Greek source of “piezoelectricity” means “squeeze/push,” that defined as the ability of piezoelectric materials to produce an electrical charge when a mechanical pressure is applied.

A piezoelectric element detects a pressure shift due to deflection of the cantilever and reacts with a shift in the electrical signal mechanical flexibility [4, 5] that contributes to greater flexibility of a cantilever caused by surface stress that results in increased output and responsiveness of sensor. The design of several efficient piezoelectric microcantilever beam as detection systems and sensors has been demonstrated by several researchers. These rely on two distinct actuation concepts; one is static and the other is dynamic. This static method tests a beam deflection due to adsorption detecting materials on single side of cantilever. Such adsorption leads to unequal surface tension and consequent beam bending. A Dynamic configuration evaluates cantilever’s resonance. The functional concept of a microcantilever-based sensor requires the gas exchange of mechanical energy which takes place from microstructured part deflections. The calculated performance is due to shift in deflection of the cantilever. As the adsorbent was not evenly spread over cantilever, the location of adsorbent thus impacts the dynamic reaction in various ways, based on the position of the mass relative to cantilever. The analysis was carried out on various structures for the microcantilever beam which maintains constant length, width, and thickness.

The goal is to measure the environmental levels of several gases. This work varies from typical microgas sensors throughout the sensing, which tests displacement directly with piezoelectric materials. The cantilever beam simulations of the various configuration were performed using COMSOL Multiphysics. The difficulty of this issue is that energy will be derived through the thin layer of piezoelectrical substrate. In this work, several modeled cantilever structures are developed using COMSOL. The software chosen for modeling and simulating this system is COMSOL Multiphysics. This is an efficient, immersive environment focused to solve on differential equations. Scilab is free, open-source computing software as well as a programming language of a greater extent, mathematically driven kind. This may be used for the interpretation of information, mathematical analysis,

image improvement, simulation of fluid mechanics, computational engineering, and modeling, simulation of dynamic structures and symbolic modulation. Scilab even offers a free module named Xcos (a Modelica-based language) to design and simulate of systematic as well as stipulated complex processes, for both discrete and continuous subsystems.

2 Design Parameters

Parametric analysis for FEM modeling is heavily dependent on the following considerations.

- Review of physical issue through qualitative interpretation of the systemic solution anticipated
- Expertise of basic engineering concepts
- Good understanding of the finite-element analysis procedures.

To represent with such detail is possible and the original structure or area, the option of form, scale, amount, and configuration for a component must be addressed without growing the analytical work needed by the solution.

- (i) Type of elements
- (ii) Size of elements
- (iii) Location of nodes
- (iv) Number of elements
- (v) Simplifications afforded by the physical configuration of the body
- (vi) Finite representation of infinite bodies
- (vii) Node numbering scheme
- (viii) Automatic node generation

In Eq. 1, which is also known as Stoney's formula, deflection in the cantilever beam is measured. This refers to the applied stress "T" as the cantilever final deflection. The displacement in a cantilever beam is due to variation of the voltage of the upper surface and the lower surface. The interpretation is

$$\delta = 3\delta \left(\frac{1-v}{E} \right) \left(\frac{L}{T} \right)^2 \quad (1)$$

δ = deflection.

δ = stress applied.

E = Young's modulus.

T = cantilever beam thickness.

V = poison's ratio.

R = Cantilever radius of curvature.

δ = stress applied.

An open edge of a microcantilever deflects vibrational amplification that creates when a piezoelectric substance pushes the system externally. The intensity of the pulse was the function of frequency at which the resonant frequency peaks. A set of simple geometric forms is used in CAD to construct a periodic design and its functions. It allows for the creation of the geometry structure by transforms and Boolean operators. This design method is performed consecutively, rendering modification simple. The 3D design is planned, as its displaced study is regarded in its structure, thus allows the simulator ideal for studying electrostatics. Substrate size is taken according to the configuration of the system and the criteria to be regarded. Included in COMSOL content libraries with a wide variety for predetermined materials and specific items may be produced from some of such products. A few of the substances in the library are semiconductors, ceramic, piezoelectric, etc. While choosing material for its structure framework, COMSOL searches for material parameters which are needed by a domain-governing mathematical system (Fig. 1).

When the material refused to maintain the requirements, the applicant can present the measured values necessary for material property. For simulation, the predefined substance is used. Stress load formula of piezoelectric substances needs vector elastic properties, coefficient of coupling, and magnetic permeability [6]. Specifically, substrate parameters needed by microcantilever will be length, Young's modulus and Poisson's ratio. Once the structure has been developed as well as the content has been allocated to a domain, next move would be to determine the dynamics correlated with construct. In this stage, initial conditions like forces, factors like (structural conditions), temperatures, and voltages are defined. The related calculations in COMSOL using semiconducting unit. Piezoelectric materials to first be grouped to three directions were described. Within $x-y$ axis, the 2D model is built which causes piezoelectric being poled outward [7, 8]. By using the above parameters, coupling matrix and elastic matrix were calculated in Scilab in Fig. 2. By using the structures of the materials, we can calculate the matrices.

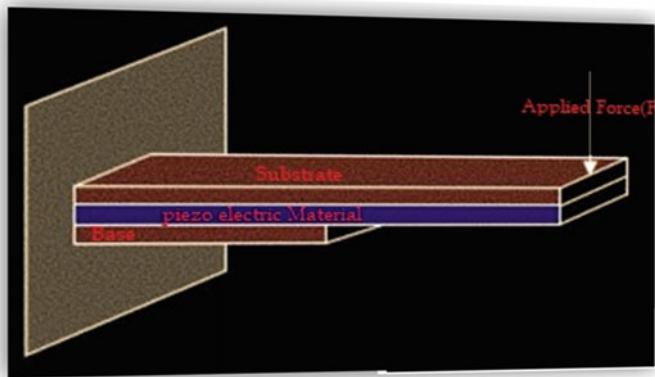


Fig. 1 Piezoelectric cantilever beam in base mode

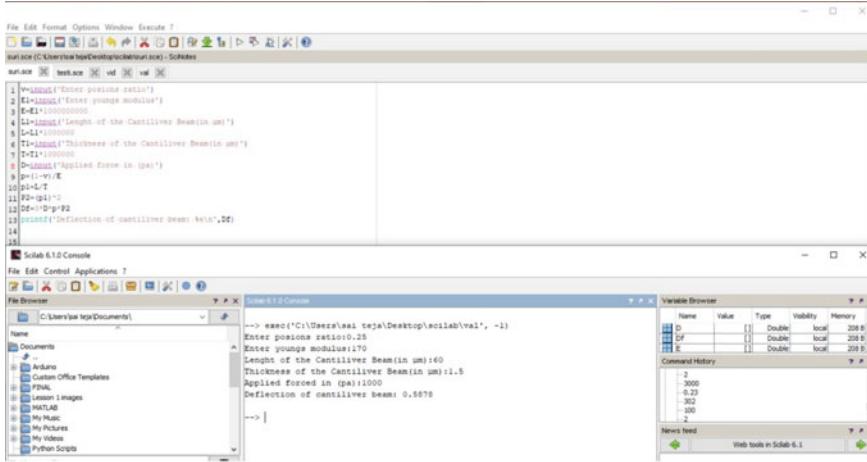


Fig. 2 Theoretical calculation in Scilab

3 FEM Analysis of Gas Sensors

In this paper, simulation and construction of a mems-based microcantilever for activity are under the conditions and environment. The efficient cantilever concept is focused upon on connectedness of specific electrical and mechanical [9] specifications and thereby its reliance upon material chosen and the geometry chosen. So, the effect depends on cantilever's electrical and mechanical tolerances. The cantilever method is focused on both material collection and geometric aspect determination. For this conceptual design, an aluminum substrate upon silicon substratum is chosen for cantilever construction. Systematic moves toward maximizing dimensions involve basic theoretical calculations for dimensions, accompanied by simulation and study of finite elements [10] of these cantilevers under the applied electric field. The demonstrated geometry shows variability for various gases on microcantilever beam sensor, thereby the suggested cantilever may be a good gas sensor. In this section, the FEA analysis is performed when the upper and lower surfaces are piezoelectric material, but the inner layer is copper material (Fig. 3).

Therefore, the finite element model (FEM) has been developed and optimized in this paper to examine gas sensor based on piezoelectric microcantilever for its material properties. COMSOL, a popular finite element modeling, results are used for creation and designing of a microcantilever finite element analysis. Piezoelectric [11] methods used to evaluate finite elements turn a vector structure through its single counterpart method. Piezoelectricity gives the relation between charge density equations and elasticity equations by electromagnetic coefficients. Let cantilever beam of length (L) and then method of measurement of pressure charge are described as

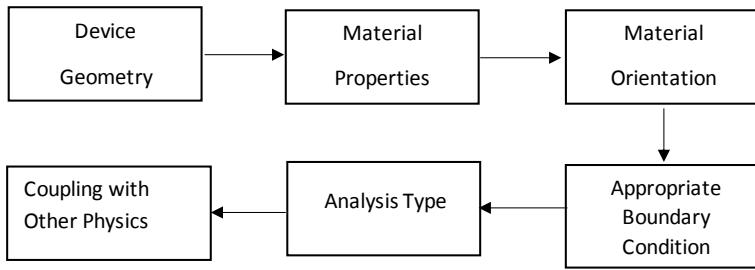


Fig. 3 Flowchart of piezoelectric simulation in FEM

$$\epsilon = S^E \sigma + d_t E'$$

$$D = d\sigma + \eta E'$$

where ϵ = strain vector,

S^E = elasticity matrix,

σ = stress vector,

d = coupling matrix,

d_t = transpose of coefficient matrix,

E' = electric field vector,

D = displacement vector, and

η = permittivity matrix.

FEM Analysis of T-Shaped Cantilever.

- Space Dimension: 3D
- Modules: there will be two modules used for this observation, one is for solid materials observed by structural mechanics-piezoplane stress-static analysis and the other for piezoelectric materials observed by COMSOL Multiphysics-structural mechanics-plane stress-static analysis. Geometry: The cantilever beam consists of three layers as shown in Fig. 1 with the dimensions stated in Table 1.
- Boundary settings: Because the study requires piezoelectric material, there will be two separate boundary conditions that must be set by the end user. The first is

Table 1 Dimensions and material properties of cantilever beams

Parameters	Substrate + Base	Piezoelectric material
Length	15	12
Width	9	3
Thickness	0.16	0.03
Material	copper	Lithium niobite
Density (kg/m ³)	2330	4700
Young's modulus (Pa)	140G	170G
Poisson's ratio	0.34	0.25

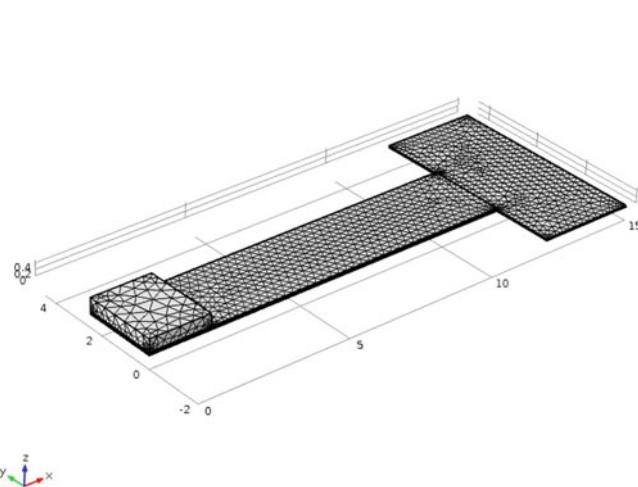


Fig. 4 T-shaped piezoelectric cantilevers mesh simulations

mechanical condition, and the second would be electrical condition. In mechanical condition, all three vertical layers on left side of the frame are limited by zero movement when all other areas were free.

- External load: free end of the beam, i.e., the right side, is applied to an arbitrary external force in the position. The assignment of load could be set in COMSOL.
- Material properties: top and bottom layers are described as materials. The material parameters were described with in stress-charge type where the user may determine the relative permittivity matrix, the coupling matrix, the elasticity matrix, and the density.
- Meshing: A cantilever beam will be meshed, and using a regular meshing, the optimum mesh density is calculated by increasing the mesh density steadily beginning with the coarse mesh and obtaining the end results.
- Solving: developed model is demonstrated for static analysis (Figs. 4 and 5).

4 Simulation and Results

The cantilever beam's gas analyzer works. Once mass/pressure is applied to the cantilever beam's active layer, it is deviated from location of equilibrium. It causes displacement differences based on the strain imposed on a beam. For the mechanical study of piezoelectric systems, the pressure parameters are varied to determine the microcantilever beam's displacement characteristics. Different pressure percentages vary by keeping suitable temperature and maximum gas analyzer pressure at room temperature. Figure 3 shows the mesh parameters must also be to achieve a strong convergence and time balance for the simulation required [5]. meshing-controlled

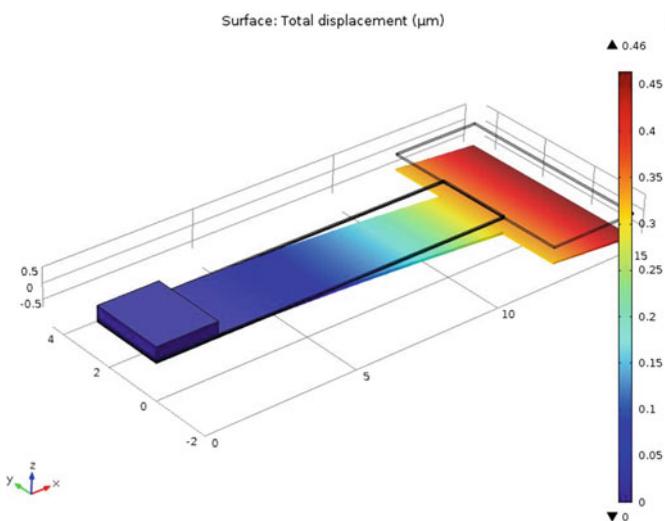


Fig. 5 T-shaped piezoelectric cantilevers displacement simulations

mechanics was used to optimize the analysis, and mesh density was restricted to default amount. This cantilever architecture analysis was performed to establish the harmonic distortion for load applied. Thus, the cantilever tests achieve optimum maximum stress, displacement, and flexibility of the MEMS cantilever system for evenly applied strain (Table 2).

To determine the displacement of the applied pressure, the T-shaped model was estimated. The simulation was performed by changing the same dimensional design. Then, deflection is measured and tabulated to pick an optimal configuration with high efficiency analysis of a gas sensor.

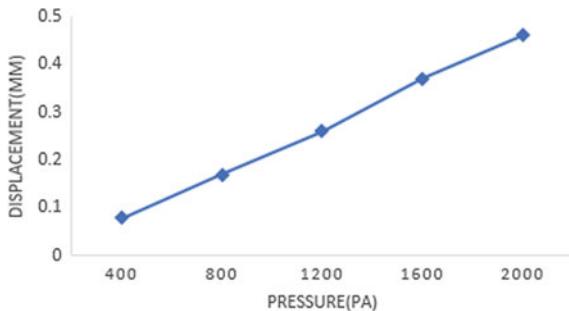
The load pressure applied is varied to obtain the relation with displacement with all four different structures and values. This relationship between the load pressure applied and the T-shaped structure displacement is plotted, and Fig. 6 indicates linear displacement difference for the pressure applied. Applying load pressure can simulate its adsorbed mass of a gas analyte on to cantilever frame.

Strain: The highest strain can occur at the fixed end of the beam, as predicted. But based on the structure of the beam, just at end of its bottom piezoelectric layer there is another region of high stain related to the immediate shift in the cross section of the beam. The study indicated that the largest strain exists at the fixed end of the beam, approximately, so the strain at the boundary layer transition area is marginally

Table 2 Comparison of theoretical and practical results for T-shape

Design	Pressure (Pa)	Deflection
Theory values (Scilab)	2000	0.40
Practical	2000	0.46

Fig. 6 Plot of displacement versus load pressure



higher than the corresponding area, but it is not higher than that of the fixed end of the beam. This is seen in Fig. 6.

5 Conclusion

Microcantilever displacement sensor finite element analysis was performed using COMSOL Multiphysics tool and Scilab Analysis is performed suitable for specific load strain. The architecture that was designed functions as a beam so that the intended gas molecules become embedded in copper coating material. Eventually, due to higher volume redistribution on surface load field, it gives out displacement volume. Varying displacement is done with the substrate and active surfaces. This will not enhance its reactivity and presence to gas particles compared to conventional materials. For various levels of pressure across ambient temperature, these complex microcantilever systems experience specific displacements. The analytical model mentioned shows similar displacement applies to with the same force. So, it is used to fabricate these models depending on effectiveness usage. Thus, the association between gas mass and the displacement of the cantilever beam is almost regular. Likewise, a detector can be developed just by using the corresponding medium and efficient thicknesses for gas. Therefore, the relationship of displacement and gas mass is mostly consistent with theoretical measurements obtained from Scilab.

Acknowledgements This work was executed with the support of Department of Science and Technology-Science and Engineering Research Board (DST-SERB) New Delhi, India for providing [Project No: ECR/2017/001183] under ECRA-SERB scheme.

References

1. P. Graak, A. Gupta, S. Kaur1, P. Chhabra, D. Kumar, A. Shetty, Design and simulation of various shapes of the cantilever for piezoelectric power generator by using comsol, in *Comsol conference in Pune* (2015)

2. J. Briscoe, S. Dunn, 2.1 Background, in *Nanostructured Piezoelectric Energy Harvesters*, 1st edn. (SpringerInternational Publishing, Berlin/Heidelberg, Germany, 2014), pp. 3–4
3. A. Erturk, D.J. Inman, 1.4 Summary of the Theory of Linear Piezoelectricity, in *Piezoelectric Energy Harvesting*, 1st edn. (John Wiley & Sons: Hoboken, NJ, USA, 2011), pp. 9–12
4. M. Bhuyan et al., Development of a fluid actuated piezoelectric micro energy harvester: finite element modeling simulation and analysis. *Asian J. Sci. Res.* **6**(4), 691 (2013)
5. M. Bailoni, Y. Wei, L. Norum, Mathematical modelling and simulation of magnetostrictive materials by comsol multiphysics, in *Comsol Conference*, Norwegian University of Science and Technology (2008)
6. K. Yang, Z. Li, D. Chen, Design and fabrication of a novel T-shaped piezoelectric ZnO cantilever sensor. Institute of Microelectronics, Chinese Academy of Sciences (IMECAS), Beijing 100029, China Volume 2012
7. P.K. Panda, Review: environmental friendly lead-free piezoelectric materials. *J. Mater. Sci.* **44**(19), 5049–5062 (2009)
8. K. Prashanthi, M. Naresh, V. Seena, T. Thundat, V. Ramgopal Rao, A novel photoplastic piezoelectric nanocomposite for MEMS applications. *J. Microelectromechanical Syst.* **21**(2), 259–261, Article ID 6118300 (2012)
9. W. Zhou, A. Khaliq, Y. Tang, H. Ji, R.R. Selmic, Simulation and design of piezoelectric microcantilever chemical sensors. *Sens. Actuators A* **125**(1), 69–75 (2005)
10. M.N. Uddin, Md. Shabiul Islam, J. Sampe, S.A. Wahab, S.H. Md Ali, Vibration based T-shaped piezoelectric cantilever beam design using finite element method for energy harvesting devices, in *2016 IEEE International Conference on Semiconductor Electronics (ICSE)*
11. V. Monika Reddy, G.V. Sunil Kumar, Design and analysis of microcantilever with various shape using COMSOL multiphysics software presented , IJETAE publication, vol. 3, issue 3, ISSN 2250 – 2459, March 2013

Design and Analysis of PZT-Based Piezoelectric Speakers



Vasudha Hegde, Christina David, E. Shirley Jesseca, Shreya Nadgouda, and S. M. Vrinda

Abstract Piezoelectric speakers work on the inverse piezoelectric effect to produce sound. When voltage is applied to the speaker, it will create waves of sound. When acoustic voltage signals are put across the ceramic part of the piezoelectric speakers, it leads to the oscillation of the material. In the following paper, an analysis for the design of efficient piezoelectric-based speaker is carried out. A novel simulation has been executed to find the behaviour of the piezo structure under the influence of magnetic field. Displacements of diaphragm under different magnetic field strengths have been analysed. The fundamental design requirements like oscillations in the audible frequency range and maximum surface displacement have been considered. A circular PZT disc is modelled and analysed practically for variation in thicknesses and diameters of its plate on COMSOL Multiphysics. For studying the stationary and dynamic performances of the PZT disc, simulation of the disc for a set range of frequency using COMSOL Multiphysics is done. The magnetic response of the PZT disc of 40 mm diameter has also been studied. This proposed analysis can further assist in the conversion of anything into a speaker.

Keywords Audible frequency · Diaphragm-based piezoelectric speaker · Magnetic field · Natural frequency · Vibration

1 Introduction

The working of speakers and buzzers that are made of piezo-based materials is driven by the inverse piezoelectric effect to produce sounds. These sound waves will be generated whenever a voltage is applied across a speaker or piezo buzzer. When a voltage signal is applied to the piezoelectric ceramic, it causes vibration in the material. These vibrations give rise to sound that travels outside the speaker.

V. Hegde · C. David · E. S. Jesseca · S. Nadgouda (✉) · S. M. Vrinda

Department of Electrical and Electronics Engineering, Nitte Meenakshi Institute of Technology, Bangalore, India

e-mail: shreyanadgouda@gmail.com

Piezoelectric speakers are usually used in alarm clocks or other mechanical devices in order to obtain a high-grade audio signal. This is because they are restricted to a narrow range of frequency response. A circular piezoelectric disc has been selected in this work as this geometry produces a maximum surface displacement and also has a higher sensitivity [1, 2].

The performance of piezoelectric diaphragm is studied through the dynamic analysis in terms of the frequency response and the surface displacement. A report showing the response of a circular PZT diaphragm in a magnetic environment is illustrated.

2 Modelling

2.1 Theoretical and Simulation Analysis

The resonant frequency at which the diaphragm must oscillate in order to achieve best performance is the natural frequency of the object. Whenever a vibrating force is applied at a resonant frequency of a given system, the system then oscillates or vibrates at greater amplitudes. To compare the values obtained during simulation and to notice the changes of frequencies with its diameter and thickness of PZT disc, theoretical analysis is done using the mathematical formula. The relationship between the natural frequency with diameter and thickness of a circular disc is given in the Eq. (1)

$$f = \frac{c}{2\pi} \sqrt{\frac{Et^3}{mR^4(1 - \vartheta^2)}} \quad (1)$$

For PZT-5A: f is the natural frequency of the PZT circular disc in Hertz, C is the dielectric loss factor ($C = 1.8$), E is Young's modulus in Newton/m² ($E = 5.33$), t is the thickness in mm, m is in mass/kg² ($m = \rho t$, where $\rho = 7750$ kg/m³), R is the radius of circular PZT disc in mm, ν is the value of Poisson's ratio ($\nu = 0.31$) [3, 4].

From (1), it is evident that the natural frequency depends on the two variables, i.e. radius and thickness. Hence, these two parameters have been chosen for analysing the variation of frequency [5]. The effect of the changes in diameters and thicknesses of PZT disc have been examined through COMSOL software.

The piezoelectric material lead zirconate titanate (PZT) possesses high stability, sensitivity and resonance frequency. It can also produce electrical response to the dynamic mechanics. PZT 5A and PZT 5H are the two forms of PZT crystals that are commonly used. PZT-5H has considerable properties of piezoelectric materials, but it is highly influenced by temperature change. PZT 5H has a slightly lower temperature range. PZT 5A is the best in case of applications involving high temperature and vibrations. The performance of PZT 5A is also said to remain constant over widely varying temperatures. Thus, PZT 5A was chosen for our simulation. Simulation

is done for a circular disc of 1 mm thickness, and PZT-5A has been used as the piezoelectric material for the structure. During simulation, all the peripheral edges are hinged, fixed constraint is given to the edges, and 2.5 V is applied to the PZT. Also, PZT materials are known for their large piezoelectric charge coefficient, moderate primitives and high coupling factors, so PZT material is chosen. For finite element analysis, a normal mesh is selected. Stationery analysis is done for various thicknesses and diameters.

2.1.1 Variations with Thickness

The natural frequency of a PZT disc and the thickness of the disc are directly proportional to each other. By fixing a constant diameter of 40 mm, the thickness was varied in order to observe a relation between the frequencies and thicknesses. The simulated values and the analytical values are as shown in Table 1.

Figure 1 shows the relationship between the frequency and thickness based on Table 1. From Table 1 and the graph in Fig. 1, it is very clear that frequency is directly proportional to thickness. The result obtained is in accordance with the reference paper [6]. The simulation results of PZT disc having different thickness are as depicted in Fig. 2.

Table 1 Variations of frequency with thickness

Thickness (in mm)	Simulated frequency (in Hz)	Analytical frequency (in Hz)
0.5	2126.3	1804.38
0.6	2553.1	2208.38
0.7	2979.4	2576.44
0.8	3405.7	2944.5
1	4256.4	3680.63

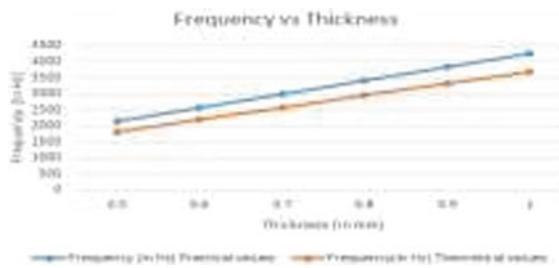


Fig. 1 Frequency versus thickness

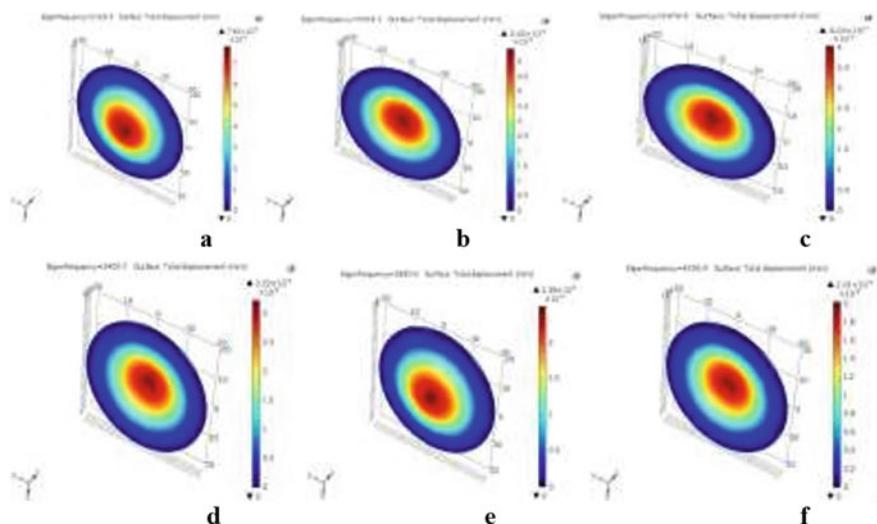


Fig. 2 Simulation of PZT discs with thickness (in mm): **a** 0.5 **b** 0.6 **c** 0.7 **d** 0.8 **e** 0.9 **f** 1

2.1.2 Variations with Diameter

The natural frequency of the PZT disc and the diameter of the disc are inversely proportional to each other. By fixing a constant value for thickness as 1 mm, the diameters were varied to choose a disc with frequency lying in audible range. The analytical and simulated values are depicted in Table 2.

Figure 3. represents the relationship between the frequency and diameter based on Table 2.

From Table 2 and Fig. 3, it is very clear that frequency is inversely proportional to diameter. The result obtained is in accordance with the reference paper [6]. The simulation results of PZT disc with variation in the diameter are as shown in Fig. 4a-e.

Though the 45 mm diameter disc gave a better frequency value, 40 mm diameter disc was chosen because of the difficulties in getting the facilities needed in fabricating a PZT disc of custom-made dimension and availability of PZT disc of 40 mm diameter in market. Thus, PZT disc of 40 mm diameter has been chosen for simulation also.

Table 2 Variations of frequency with diameters

Diameter (in mm)	Simulated frequency (in Hz)	Analytical frequency (in Hz)
30	7556.9	6643.35
35	5557.6	4807.35
40	4256.3	3680.63
45	3363.9	2908.15

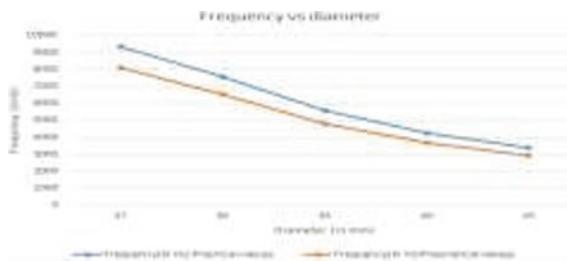


Fig. 3 Frequency versus diameter

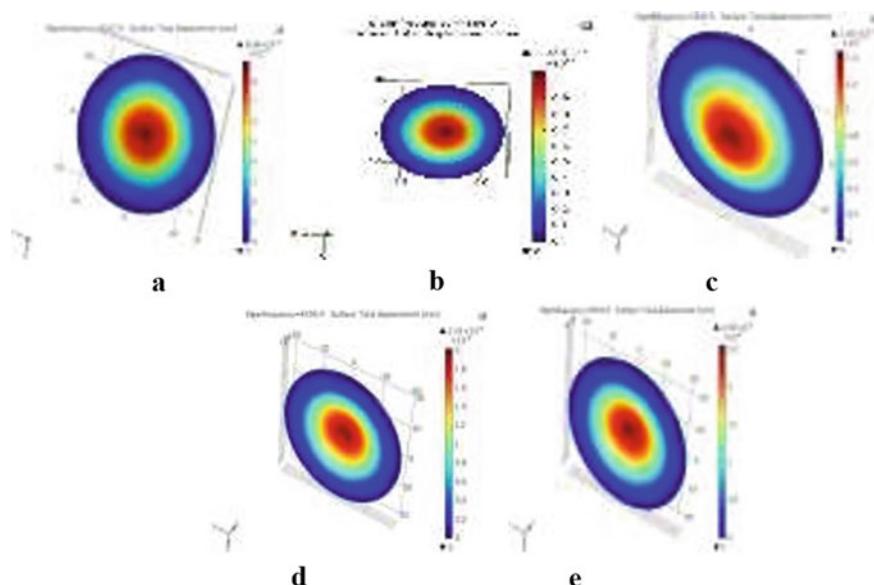


Fig. 4 Simulation of PZT discs with diameters (in mm): **a** 27 **b** 30 **c** 35 **d** 40 **e** 45

3 Simulation of PZT Structure

From Tables 1 and 2, it is clear that there is a difference in values of simulated and theoretical values that is because the simulation is done only for the piezo-based disc alone and not the whole structure including the metal base. Also, the piezoelectric material parameters depend on the fabrication process also. Based on the analysis done, it was found that PZT discs of diameter 40 and 45 mm gave maximum displacements. Since there were many difficulties in fabricating a PZT disc (like procuring high end tools and specific chemicals), it was easier to opt for the PZT disc that was readily available, i.e. 40 diameter and 1 mm thickness PZT disc. When compared to Barium Titanate (BaTiO_3), PZT materials show a better sensitivity, and they also have a greater operating temperature. PZT 5-A and PZT

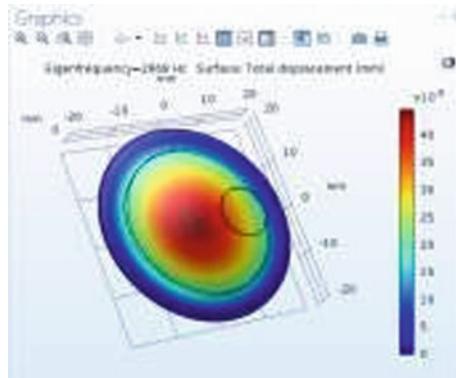


Fig. 5 Static response of the complete disc

5-H are the two variants of PZT crystals that are commonly used. The performance of PZT-5A is said to remain constant over widely varying temperatures, and also, it is the best in case of applications involving high vibrations. Thus, a PZT-5A disc of diameter 40 mm and thickness 1 mm has been chosen for the simulation. The circular PZT disc that has been selected is stuck in-between a metal base of nickel-aluminium-bronze and an electrode. For simulating the structure, i.e. a circular piezo material with the metallic base and electrodes, the solid mechanics is used as the Physics; Eigen frequency is chosen for the stationary analysis. Frequency domain is selected for the dynamic analysis in the study.

3.1 Static Response

Physics-controlled mesh with a normal size of element having a frequency range of 20 Hz to 20 kHz was chosen for simulating the static response of the structure. Figures 5 and 6 show results that were obtained. From this simulation, it is evident that a maximum displacement occurs at the central portion of the disc with 2668 Hz Eigen frequency that falls in the acoustic frequency range [6].

3.2 Dynamic Response

The first dynamic frequency and the data set were taken for simulating the dynamic response. Figures 7 and 8 depict the results of the dynamic response that were obtained. For a circular disc, maximum displacement occurs at the centre. A point graph is plotted for displacement versus frequency range to check the frequency at which maximum displacement occurs. From Figs. 8 and 9, it is evident that the piezo

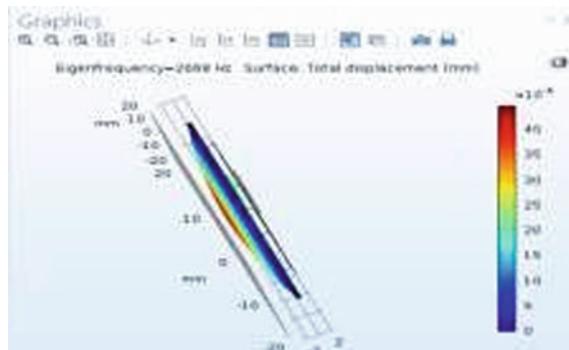


Fig. 6 Deflection obtained in static response structure including metal electrodes

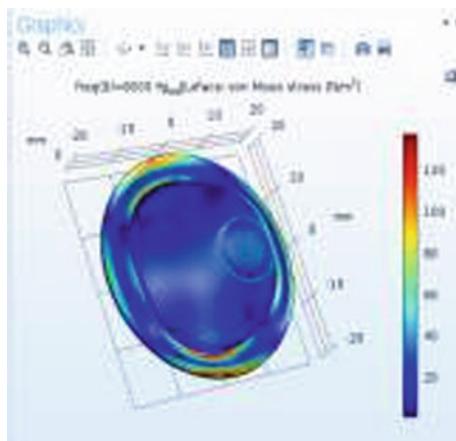


Fig. 7 Dynamic response

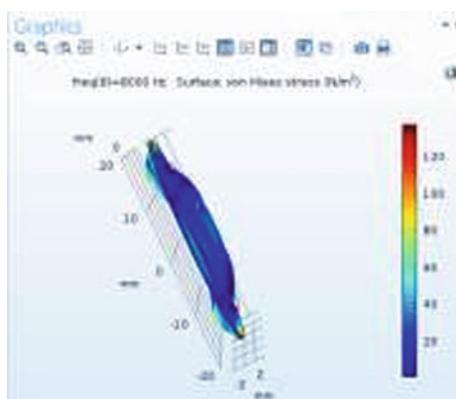


Fig. 8 Deflection obtained from the dynamic response

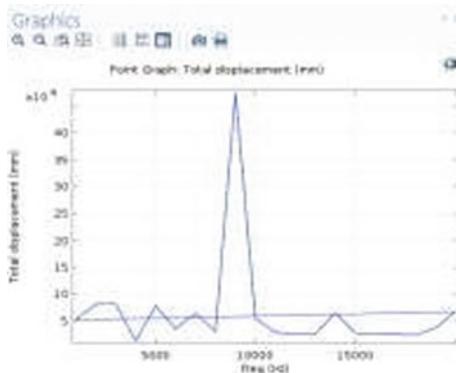


Fig. 9 Point graph

disc exhibits the highest displacement at the central region with 8 kHz frequency which falls within the acoustic range of frequency.

3.3 Magnetic Field Response

The PZT is enclosed into a cube of dimensions $200 \times 200 \times 200$ made of the material air. The dimension of the cube is chosen to be much higher than that of the dimension of the disc. All the six faces of the cube are subjected to magnetic fields of different magnitudes to observe the response of the PZT in the magnetic field. The corresponding response and variations in surface displacement are studied [7].

The simulation of the whole structure in the magnetic field and Eigen frequency is done with 5, 10, 15, 20 and 25 Wb/m (along the x , y , z plane) magnetic fields, respectively, which were applied across all the six cubical faces as depicted in Fig. 10. It is noticeable that the highest deflection occurs at the central region and reduces near PZT edges. The surface displacement in mm is noted at central region and has been tabulated in Table 3.

The graph depicting the relationship between displacement of the PZT and the magnetic field is shown in Fig. 11. From Table 3, it is evident that as the magnetic field increases, the surface displacement decreases. This is due to the eddy current effect.

4 Results

The design factor of circular piezoelectric disc made of PZT-5A piezoelectric material that can be used as speaker is simulated with the help of COMSOL software. Dynamic

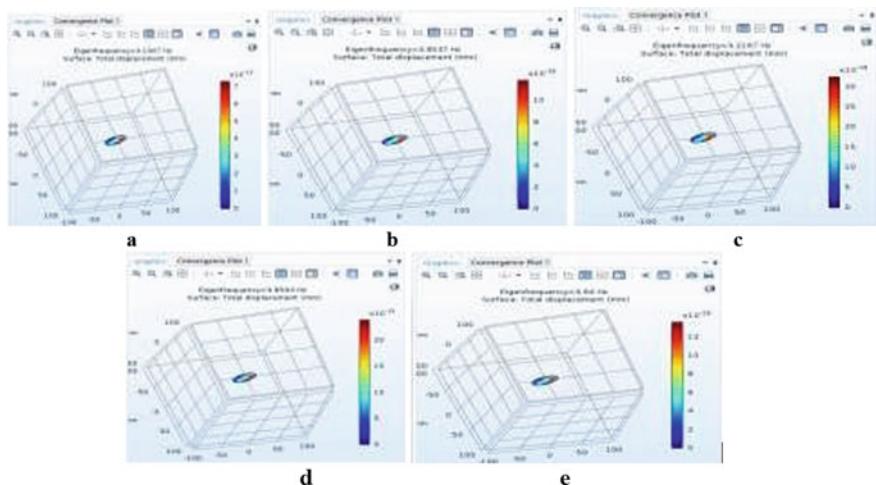


Fig. 10 Simulation of PZT disc with magnetic field **a** 5Wb/m **b** 10 Wb/m **c** 15Wb/m **d** 20 Wb/m **e** 25Wb/m

Table 3 Variations of the surface displacement with magnetic field

Magnetic field (in Wb/m)	Surface displacements (mm)
5	7×10^{-13}
10	10×10^{-16}
15	20×10^{-19}
20	30×10^{-19}
25	12×10^{-20}

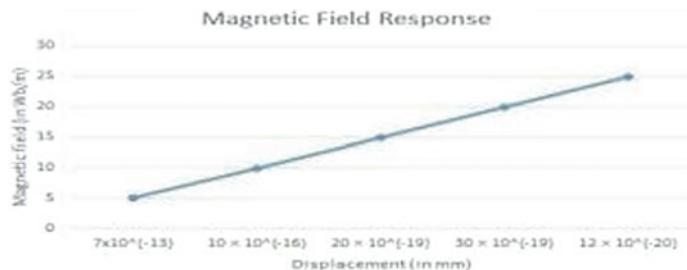


Fig. 11 Magnetic field response

and static properties like the displacement of circular PZT disc and the frequency have been examined.

Fixing a constant value of thickness and changing the values of the diameters, the various Eigen frequencies were tabulated to compare the analytically computed

values. Thus, it proves evidently from results of frequency versus diameter plot that frequency and diameter show inverse relationship. Given a selected 40 mm diameter, thickness values were changed, and the frequencies were tabulated for the results of simulation and also the analytically computed values. It is clearly proven from the results of the frequency versus thickness plot that frequency and thickness show direct proportionality. The results obtained for variations of frequency with diameter and thickness are in accordance with the proved results. The static response showed the highest displacement occurring at the central portion of the disc and PZT disc has the natural frequency in the audible range. Dynamic response of the structure was acquired by using the analysis of frequency domain, and the maximum value of displacement was observed at 8000 Hz that lies in the audible range of frequencies. The static response for the PZT structure in a magnetic field environment was also achieved. From this analysis, it is clear that surface displacement decreased when the magnetic field was increased due to eddy current effect. A feasible application of the analysis made is that it can be applied for the conversion of any object into a speaker. In this case, whenever driven with the help of an audio amplifier, piezoelectric diaphragm will resonate to sound waves leading to the object's vibration [8].

References

1. V. Hegde, S. Veena, H.M. Ravikumar, S. Yellampalli, Piezoelectric acoustic pressure sensor diaphragm design for energy harvesting, in *International Conference on Advances in Energy Conversion Technologies (ICAECT), Manipal*, (2014, 23–25 Jan 2014). <https://doi.org/10.1109/ICAECT.2014.675705>.
2. A. Kumar, Periasamy, Simulation study and performance optimization of piezoelectric mems device for acoustic sensor applications. *Int. J. Electrical Electronics Data Commun.* ISSN: 2320-2084
3. N. Guo, The Vibration Characteristics of Piezoelectric discs. PhD Thesis, Imperial College of London (1989)
4. CTS Electronic components data sheet for PZT 5A. https://venividikiwiki.ee.virginia.edu/mediawiki/images/9/95/PZT_5Aand5H.pdf
5. S. Bhuvana, P. Hoode, V.T. Sindhu, V. Hegde, Design and analysis of piezoelectric cantilever based vibrationsensor, in *IEEE International Conference on System, Computation, Automation and Networking, Pondichery* (2018), <https://doi.org/10.1109/ICSCAN.2018.8541161>
6. V. Hegde, S. Gatade, C.K. Neha, M. Hegde, V. Aher, V. Sai, Design and analysis of piezoelectric cantilever based vibration energy harvester for low power devices, in *Global Conference for Advancement in Technology (GCAT), Bangalore* (2019). <https://doi.org/10.1109/GCAT47503.2019.8978326>
7. B. Ali Rachedi, A. Babouri, F. Berrouk, A study of electromagnetic field generated by high voltage lines using Comsol, in *International Conference on Electrical Sciences and Technologies in Maghreb* (2014). <https://doi.org/10.1109/CISTEM.2014.7076989>
8. Turn Anything into Speaker. <https://www.instructables.com/id/Turn-Nearly-Anything-into-a-Speaker>

Drowsiness and Yawn Detection System Using Python



Surabhi Kumari, Kumari Akanksha, Sasmita Pahadsingh, Swati, and Supriya Singh

Abstract Nowadays, the number of road accidents is increasing day by day. One of the major reasons behind it is feeling drowsy while driving or drunk driving. Drunk driving and drowsiness alone lead to 36% of accidents. Therefore, by detecting the drowsiness, the number of fatigue-related vehicle accidents can be minimized. So, the key objective of this paper is to design a prototype of drowsiness and yawn detection system using Python and Dlib model. It is a real-time system which will detect the drowsiness among car drivers by capturing image continuously and will warn the driver whenever they will feel sleepy. The innovation of the present work is established on blinking of eyes and yawn frequency. The per closure value of eye is examined for detection of drowsiness, and whenever it exceeds a certain value, then the driver is recognized to be sleepy. Similarly, we will inspect the yawn value to detect the drowsiness and whenever it exceeds its minimum threshold value it will give yawn alert.

Keywords Python programming · Dlib · OpenCV · Face detection · Eye detection · Yawn detection

S. Kumari (✉) · K. Akanksha · S. Pahadsingh · Swati · S. Singh
KIIT University, Bhubaneswar, India
e-mail: kumarisurabhi120@gmail.com

K. Akanksha
e-mail: akankshakumari954@gmail.com

S. Pahadsingh
e-mail: spahadsinghfet@kiit.ac.in

Swati
e-mail: swatisharma00659@gmail.com

S. Singh
e-mail: supriyasingh9803@gmail.com

1 Introduction

Driving with drowsiness or in alcoholic state are the two main reasons causing traffic accidents. Driver exhaustion not just effects the readiness and reaction time of the driver, but it also increases the chances of being involved in car accidents. According to many researches, drowsiness is related to thousands of traffic accidents each year. National Highway Traffic Safety Administration (NHTSA) analysis data indicate that driving while drowsy is a contributing factor to 25–27% of car crashes and the driving while drowsy results in five-to-six times higher near—crash/crash risk relative to alert driver [1]. And 20% have also admitted that at some point in their driving careers they have fallen asleep while driving [2]. This high accident rate is due to the fact that sleepy driver fail to take correct action prior to a collision. This important problem is often ignored by the driver. In this way, the utilization of helping frameworks that screen a driver's degree of cautiousness is vital to forestall street mishaps. Drowsiness is something which can be described in people by behavioral changes in their facial expressions like the eyes will start getting closed, frequent yawning and the neck inclines. Therefore, physiological signals are much appropriate to notice drowsiness with few untrue positives, making it achievable to alert a drowsy driver in a punctual manner and thereby preclude more road accidents. As far as the literature is concerned for face detection, huge amount of algorithms and methodologies are available. From various literature, it is found that mostly three techniques are used for the detection of drowsiness. We used image processing as it is better than these two since it is simple, easy to use and cost effective. In this system, open CV is used which focus on real-time application for the detection of yawn count and eye blink, whereas dlib is used for face detection and mapping. The objective of this system is to monitor continuously the yawning and eye blink of the driver. When the limit of eye blink and yawn count exceeds the threshold value, it alerts the driver by generating an alarm. The Dlib prototype is prepared to recognize 68 facial tourist spots. As appeared in Fig. 1, the languor highlights are removed and the driver is alarmed in case of laziness being distinguished. The prototype does not need any earlier data of the person who is trying it. Principle prerequisites are Python programming and OpenCV.

Fig. 1 Block diagram of real-time drowsiness and yawn detection system

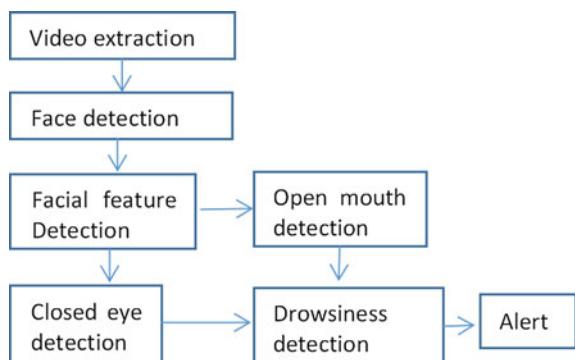


Table 1 Advantages and limitations of each technique

Measures	Parameters	Advantages	limitations
Subjective measures	Questionnaire	Subjective	Not possible in real time
Vehicle based measures	Deviation from lane and loss of control	Non-intrusive	Unreliable
Physiological measures	Heart rate, pulse rate Body temperature, etc.	Reliable, accurate	Intrusive
Behavioral measures	Yawning eye closure Eye blink, head pose	Easy to use	Lighting condition Background

1.1 State-of-Art Method

Already, there has been lot of research and papers have been put forward to detect the fatigue among drivers. In some papers, traditional methods are used like by measuring heart rate, brain waves and respiration [3]. And some require more than one camera to monitor only yawning patterns of the driver [4]. Most common reason behind the midnight accidents are driver's drowsiness, and this is shown in [5] using their embedded system proposal. In this paper, [6] they have used machine learning approach for detection purpose, and it basically do face detection in various conditions like with glasses or without glasses. In this paper, [7] the approach which has been taken required external camera to take the images as it is a night vision camera, so it will be easier to detect in poor light also. Very recently, in 2019, this paper [8] has also proposed a technique in which a drowsiness detector is implemented using OpenCV and Python. As, all these techniques show slightly less accuracy so to overcome all the problems, this paper proposed a technique which is not even complex and will show less error as compare to others.

2 Comparison of Different Techniques of Drowsiness Detection

Generally, drowsiness detection techniques are divided into four main categories (Table 1).

3 Implementation

This section will describe the different parts which are used to implement the drowsy driver detection system. Here, we have used openCV and Python to implement the system. The different phases used in the system are.

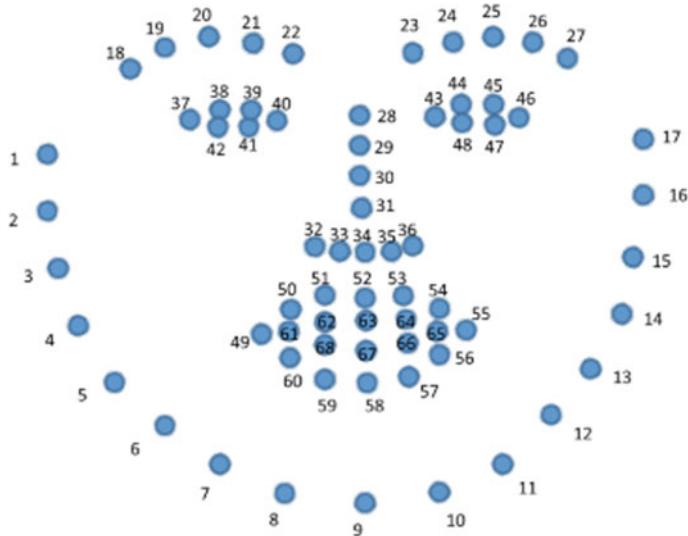


Fig. 2 Process of mapping 68 facial marks over detected face

3.1 Face Detection

Dlib library in Python has a pre-trained facial indicator detector which is helpful to locate the 68 points which map to the facial structure. The location of 68 points is shown in Fig. 2.

Location of detected facial parts:

- (A) Left eye is detected by points 43–48.
- (B) Right eye is detected by points 37–42.
- (C) Mouth is detected by points 49–68.

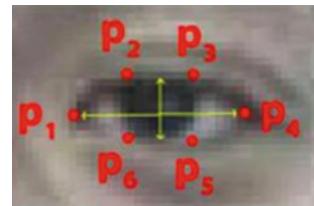
3.2 Eyes Detection

Eyes are detected using the facial landmark, and then, their EAR ratio is calculated using the Euclidean distance. If the EAR ratio is less than the threshold value of EAR (as stated in the code) for consistent 30 frames (as given in the code), then a alarm will be raised to wake the driver so as to prevent the driver from any accident. The formula to calculate the EAR ratio is given below [9] (Fig. 3):

$$\text{EAR} = \frac{(|P_2 - P_6| + |P_3 - P_5|)}{2 * |P_1 - P_4|} \quad (1)$$

When the eyes are closed, the numerator decreases and the denominator value remains same it would be with eyes open so the value of EAR decreases.

Fig. 3 Landmark of eye to calculate the EAR



The video is processed 30 frames at a time so that the blink will not be considered as drowsy.

3.3 Yawning Detection

The mouth is detected using the dlib's facial landmark. If the distance between the upper lip and lower lip is more than the threshold value (as stated in code) for consistent 30 frames (as taken in code), then an alarm will be raised for the driver to wake up so as to prevent the accident caused due to drowsiness of driver. We have calculated a single point for upper lip by calculating the mean of all six coordinates of top lip, and then the same was calculated to get the value for the low lips. To get the yawn value, we subtracted the mean of low lip from top lip. The formula to calculate yawn is given below:

$$\text{Distance} = \text{top_mean} - \text{low_mean} \quad (2)$$

The following steps were used to test the system:

Step 1: As the recording starts, the system starts reading the frame, and then, it is re-scaled and transformed to gray scale images.

Step 2: cv2 detector is carve up. The position of the face is detected by it.

Step 3: Dlib predictor determine facial landmark and find position of eyes and mouth.

Step 4: The coordinates of eyes and mouth are taken to calculated the EAR ratio and the yawn distance value to find whether the driver is drowsy or not.

Step 5: The ratio and distance is calculated and compared with the threshold value which was set as $\text{EAR} = 0.3$ and $\text{yawn} = 20$ to determine whether the person is sleepy or not. If the EAR is below the threshold, it is determined that driver is sleepy and so if the yawn value is more than the threshold.

Step 6: If the value remain consistent for given number of frames, the alarm is raised so as to prevent accident.

4 Algorithm

When the code will run firstly, it will load the predictor and detector so as to detect the face and then in it the eyes and ears, respectively. The video stream will start; after that, it will then process into the frames, and the predictor will detect the face.

When the mouth is detected, the lip distance is calculated; if the lip distance is more than 20 for continuous 30 frames, then the driver is regarded as sleepy; hence, the yawn count will increase to 1, and a yawn alert will be issued for the driver to tell him to take some fresh air.

When the eyes are detected. The eyes are detected if EAR is more than 0.3 than next set of videos will be processed as frames. If EAR is less than 0.3 for consistent 30 frames, then EAR count becomes 1 and drowsiness alert is issued telling the driver to wake up. If we want to stop the video stream than will terminate the program and if we want to continue to run the system, then next set of videos will be processed as frames (Fig. 4).

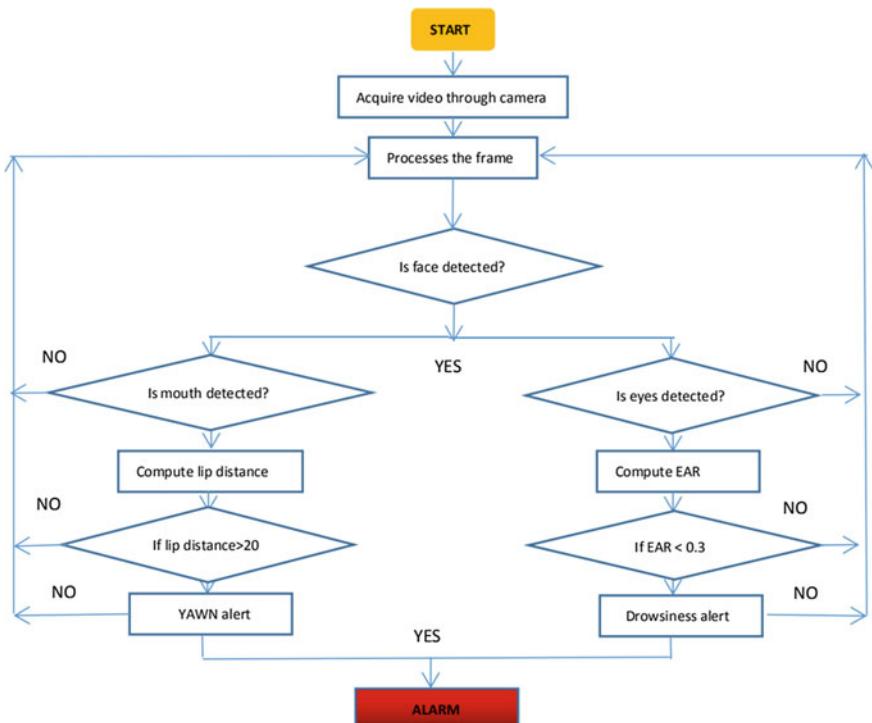


Fig. 4 Flowchart describing the algorithm

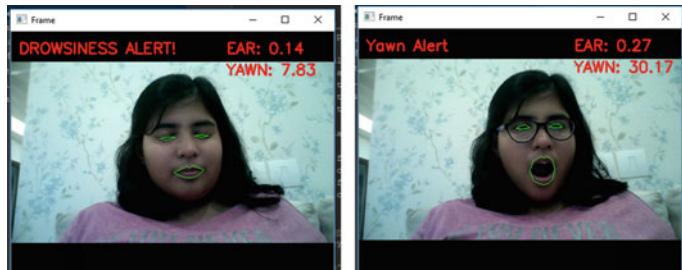


Fig. 5 Showing drowsiness alert and yawn alert

5 Result and Output

The below figures shows the real-time computational result on the basis of two procedure (eye closure and yawn detection). The obtained output shows the YAWN ALERT if the distance between the lips are increased and shows DROWSINESS ALERT if the EAR is decreased. As we can see that the EAR value is 0.14 which is less than the threshold value which is 0.3, so the system will give drowsiness alert saying “wake up sir.” And we can see the yawn value is 30.17 which is more than the threshold value set for yawn which was 20; so the system will give a yawn alert saying “take some fresh air sir” (Fig. 5).

6 Conclusion

This research completely shows the drowsiness detection and is very useful in avoiding road. The system detects the face of the driver, and then, it calculates the number of eye blink and yawn count; if the number exceeds the threshold value, it generates an alarm which alerts the driver and helps in avoiding road accidents. It gives almost accurate result except in some cases like in poor lightening condition. This system can be modified in future so that it can overcome the problem of not working properly under poor light.

References

1. ARau P. Drowsy Driver Detection and Warning System for Commercial Vehicle Drivers: Field Operational Test Design, Analysis, and Progress. National Highway Traffic Safety Administration; Washington, DC, USA: 2005
2. <https://www.nsc.org/road-safety/safety-topics/fatigued-driving>
3. P. K. Stanley, T. Jaya Prakash, S. Sibin Lal and P. V. Daniel, “Embedded based drowsiness detection using EEG signals,” IEEE Intern. Conf. on Power, Control, Signals and Instrumentation Engineering, Chennai, India, Sept. 2017, pp. 2596–2600

4. L. Li, Y. Chen and Z. Li, Yawning detection for monitoring driver fatigue based on two cameras, in 12th Intern. IEEE Conf. on Intelligent Transportation Systems, St. Louis, USA, Oct. 2009, p. 1–6
5. Fouzia, R. Roopalakshmi, Jayantkumar A. Rathod, Ashwatha S. Shetty And K. Supriya, “Driver Drowsiness Detection System Based On Visual Features, in 2018 Second International Conference On Inventive Communication And Computational Technologies (ICICCT), 20–21 April 2018
6. Cyun-Yi Lin, Paul Chang, Alan Wang and Chih-Peng Fan, “Machine Learning and Gradient Statistics Based Real-Time Driver Drowsiness Detection”, IEEE International Conference on Consumer Electronics-Taiwan (ICCE-TW), 19–21 May 2018
7. IEEE International Conference on Intelligent Techniques in Control, Optimization and Signal Processing (INCOS)
8. S. Mohanty, S.V. Hegde, S. Prasad, J. Manikandan, Design of real-time drowsiness detection system using dlib, in *2019 IEEE International WIE Conference on Electrical and Computer Engineering (WIECON-ECE)*, Bangalore, India, 2019, p. 1–4, <https://doi.org/10.1109/wiecon-ece48653.2019.9019910>
9. https://www.researchgate.net/publication/328911666_Real-Time_Driver's_Drowsiness_Monitoring_Based_on_Dynamically_Varying_Threshold

Intelligent Analysis of X-Ray Images for Detecting Bone Abnormality in Upper Extremities



Puja Kalivarapu, Rufina Flora George Rajan, and B. Niranjana Krupa

Abstract In this paper, two levels of classification have been carried out, the first being upper extremity recognition followed by bone abnormality detection. Support vector machines (SVMs), VGG16 and Xception V3 convolutional neural networks (CNNs) were implemented for this purpose, and their performances have been compared. The conventional method (SVM) uses the grey level co-occurrence matrix (GLCM) and local binary pattern (LBP) methods for feature extraction. In the first level, X-ray images from the MURA dataset were classified into seven categories. Next, the images were classified into two classes based on the presence of an abnormality. The VGG16-based model performed the best in the first level of classification with an average accuracy of 94.4% over three trials. In the second level, the SVM model performed the best with an accuracy of 66.33%.

Keywords X-ray · SVM · VGG16 · Xception V3 · MURA

1 Introduction

Bone abnormalities are frequently missed during radiological examinations, a study misinterpreted as normal can lead to several complications. To avoid such issues, the usage of machine learning (ML) in the field of medicine, to aid radiologists, has gained popularity rapidly.

In [1], content-based medical image retrieval (CBMIR) is done using CNNs. X-ray images were taken from the IRMA dataset, which has 55 different categories. The images were enhanced using adaptive histogram equalisation (AHE), then classified

P. Kalivarapu (✉) · R. F. George Rajan · B. Niranjana Krupa

PES University, Bangalore, KA 560085, India

e-mail: pkalivarapu@gmail.com

R. F. George Rajan

e-mail: rufinaflora2303@gmail.com

B. Niranjana Krupa

e-mail: bnkrupa@pes.edu

using two pre-trained models—AlexNet and GoogleNet. The CNNs were used as feature extractors, with the last few layers replaced for classification. The performance of both the networks was compared based on precision, recall and F1-scores calculated from confusion matrices generated from five trials.

An enhanced CBMIR system was developed in [2], using the IRMA dataset. Median filtering was applied to low performing classes to improve classification accuracy, and all images were then enhanced using AHE. Features were extracted using principle component analysis (PCA), wavelet transform and speeded up robust feature (SURF) descriptors. A database was created consisting of codewords generated from the features extracted. The performance of three different classifiers—SVM, K-nearest neighbour (KNN) and relevant vector machine (RVM)—was compared over a fivefold cross validation approach.

The main objective of the paper was to use different ML algorithms to detect bone abnormalities and classify regions of the upper extremity recognised in X-ray images and compare their performances. The performances of three image classification models on the MURA [3] dataset, namely SVM-based classifier, VGG16 and Xception V3-based CNN classifiers, were compared. The regions of upper extremity considered include—fingers, wrist, forearm, elbow, humerus, hand and shoulder. An abnormality is detected if a fracture, hardware or cancerous tissue is detected, other abnormalities and skeletal regions have not been considered as they are not present in the dataset.

2 Methodology

2.1 Description of Dataset

The paper has made use of the MURA dataset, which comprises 40,561 images of the upper extremity, where each study (Fig. 1) was manually labelled by radiologists as either normal or abnormal. The abnormalities include fractures, hardware inserted in from previous surgeries, degenerative joint diseases and lesions [3]. The train set of the MURA dataset was used for training and testing the model, detailed information of which is shown in Table 1.

2.2 Conventional Approach

Image Enhancement

An image read from the dataset was first enhanced to improve image quality (Fig. 2a(i)–(v)). Image normalisation using the NORM_MINMAX method was carried out to change the domain of pixel intensity values, followed by pixel equalisation using contrast limited AHE (CLAHE), to improve the contrast. Noise was filtered

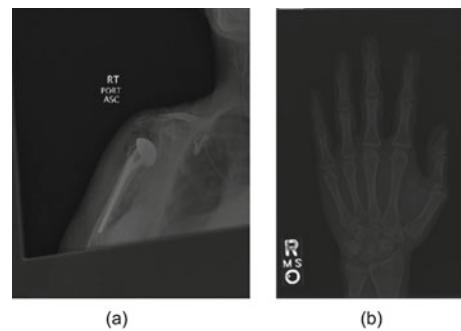


Fig. 1 **a** Abnormal Shoulder X-ray image, **b** Normal Hand X-ray image [3]

Table 1 Description of the “Train set” from the MURA dataset [3]

Study	Normal	Abnormal	Total
Elbow	1094	660	1754
Finger	1280	655	1935
Forearm	590	287	877
Hand	1497	521	2018
Humerus	321	271	592
Shoulder	1364	1457	2821
Wrist	2134	1326	3460
Total No.	8280	5177	15167

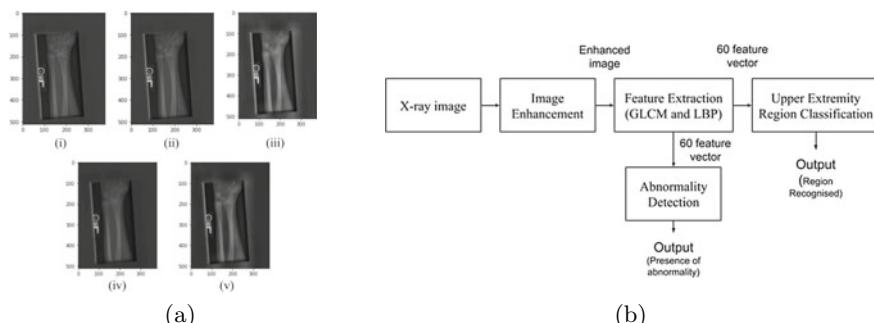


Fig. 2 **a** (i) Original X-ray image [3], (ii) Normalised image, (iii) CLAHE image, (iv) Denoised image, (v) Sharpened image, **b** SVM approach block diagram

using non-local means denoising, and finally, image sharpening was done to make details in the X-ray more prominent.

Feature Extraction

The texture of an image captures a lot of important information stored in the image which plays a vital role in image classification. The GLCM [4] extracts correlation, energy, homogeneity, dissimilarity, contrast and angular second moment features of each X-ray image. The 8-bit variant of LBP [5] is used, considering only eight points in the neighbourhood of a pixel, at 3. A histogram of patterns extracted as binary codes is appended to the feature vector obtained from GLCM to form a 60-feature vector. These vectors were used for both upper extremity recognition and bone abnormality detection.

SVM Implementation

The kernel used in both the SVM models is the radial basis function (RBF) kernel. It is popularly used with nonlinearly separable data, and it is very effective on datasets with a large number of features. The upper extremity recognition model (Fig. 2b) classified images into seven classes, while the bone abnormality detection model classified the data into two classes—positive and negative—based on the detection of an abnormality. Both models used a gamma value of 0.0001 and C value of 1000, obtained using grid Search.

2.3 CNN Approach

The two CNN approaches implemented are VGG16 and Xception V3.

VGG16 Model

Upper Extremity Recognition Model The input images were first reshaped to (400,500,1) and then passed into the model (Fig. 3a). The VGG16 CNN [6] extracted the most distinguishing features of the image, and the final layers classify the image into seven categories. The model was initially trained with the learning rate parameter set as 0.01, and the classification accuracy obtained was 26.66%. On decreasing learning rate to 0.001 of the initial value, the accuracy shot up to 97.6% at the end of 32 epochs. Final parameters for the model were 0.0001 learning rate, for 32 epochs and batch size of 24.

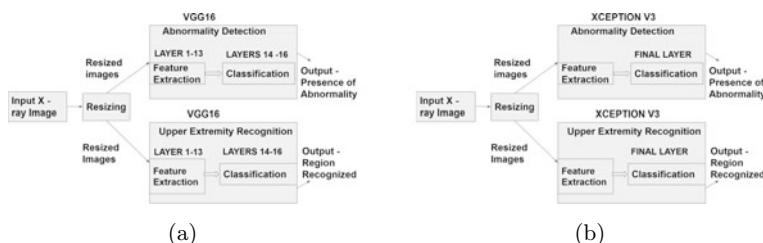


Fig. 3 Block diagrams: **a** VGG16 Approach, **b** Xception V3 Approach

Bone Abnormality Detection Model This model was built identical to the previous one with the addition of dropout (dropout rate = 0.5) and batch normalisation layers after each original layer in VGG16 [6], to avoid over fitting and data skewing. The images were reshaped and classified into two classes. The model initially classified all inputs as normal. Upon the introduction of additional layers, the classification accuracy improved to 60%. Final parameters for this model were 0.01 learning rate, for 64 epochs and batch size of 24.

Xception V3 Model

Upper Extremity Recognition Model The Xception V3 model [7] was implemented same as the VGG16 model. The preprocessed data was fed directly into the model (Fig. 3b). Convolution and separable convolution layers extract features along with batch normalisation. The images were classified into seven classes. The model was trained using batch size of 15 for 20 epochs with a learning rate of 0.0001.

Bone Abnormality Detection Model This model was implemented similar to previous model, and the images were classified into two classes. The model was trained using batch size of 15 for 30 epochs with a learning rate of 0.01.

3 Results and Discussion

The results of this paper were obtained by testing the models in three trials, with an 80:20 split, each consisting of a different subset of the MURA Train dataset. From Table 2 (XCP - Xception V3), it was noted that the SVM model has classified Elbow data with the highest recall, precision and F1-score [8], and the lowest values were obtained for Forearm data. Shoulder images were best classified by the VGG16 model, and the lowest values were obtained for the Forearm images. The same can be seen with the Xception V3 model, and the best results were seen with Shoulder data and worst with Forearm data.

Table 3 displays the results generated while validating the abnormality detection models. It can be observed that in the negative category, the VGG16 model has the highest precision, recall and F1-score values, whereas in the positive category, the SVM-based model has performed the best.

In Table 4, it can be seen that the VGG16 model has performed the best in recognising the upper extremity region, with an average accuracy of 94.4%, and the SVM model has performed the best in bone abnormality detection with an average accuracy of 66.33%. The performance of all three models in bone abnormality detection could be improved by feature engineering and data augmentation, increasing classification accuracy. Although direct comparison cannot be made, and in [3] and [8], the authors have tried to detect only bone abnormalities using the MURA dataset. In [8], they have used SVM classifier with LBF kernel to achieve an accuracy of 62%, whereas in [3], they have employed DenseNet169 and achieved 81.5% sensitivity and 88.7% specificity.

Table 2 Upper extremity recognition classification report

Region	Precision			Recall			F1-Score		
	SVM	VGG16	XCP	SVM	VGG16	XCP	SVM	VGG16	XCP
Elbow	0.6567	0.89	0.81	0.5433	0.91	0.93	0.5633	0.90	0.87
Finger	0.9133	0.91	0.95	0.9167	0.91	0.84	0.9133	0.91	0.90
Forearm	0.5833	0.79	0.72	0.48	0.81	0.63	0.5067	0.80	0.67
Hand	0.7867	0.95	0.93	0.7677	0.95	0.94	0.7667	0.95	0.93
Humerus	0.78	0.86	0.76	0.8166	0.77	0.70	0.8167	0.81	0.73
Shoulder	0.6933	0.96	0.95	0.7267	0.95	0.96	0.7167	0.95	0.96
Wrist	0.7533	0.94	0.92	0.75	0.94	0.93	0.7433	0.94	0.93

Table 3 Abnormality detection classification report

	Precision			Recall			F1-Score		
Abnormality	SVM	VGG16	XCP	SVM	VGG16	XCP	SVM	VGG16	XCP
Negative	0.61	0.6566	0.58	0.48	0.61	0.54	0.5367	0.6324	0.56
Positive	0.6833	0.4967	0.70	0.795	0.5367	0.73	0.7333	0.5159	0.72

Table 4 Accuracy comparison of models implemented

Model	Upper extremity recognition (%)	Abnormality detection (%)
SVM classifier	79	66.33
VGG16	94.4	60.21
Xception V3	90.61	65.48

4 Conclusion

For the SVM model, the work began with image enhancement, followed by feature extraction using the GLCM and LBP methods. Validation of the SVM classifiers was exacted by the RBF kernel. The pre-trained CNN models extract features, and additional layers act as classifiers. All models were successfully trained and tested, on the MURA dataset. The CNN approach using VGG16 performed best in recognition of the upper extremity with an average accuracy of 94.4%. Whereas, the SVM model yielded better results in bone abnormality detection with an average accuracy of 66.33%, as compared to the CNN approaches. Hence, it can be concluded that the VGG16 model outperformed the SVM and Xception V3 models in upper extremity recognition and the SVM model performed the best in bone abnormality detection. As future work, the scope can be expanded to other regions of the human skeleton and detection of other types of abnormalities.

References

1. P. Arti, A. Agrawal, A. Adishesh, V. Lahari, K.B. Niranjana, Convolutional neural network models for content based x-ray image classification, in *2019 IEEE International WIE Conference on Electrical and Computer Engineering (WIECON-ECE)*. (IEEE, 2019), pp. 1–4
2. M. Lahari, B.N. Krupa, Intelligent content based x-ray image retrieval using speeded up robust feature descriptors, in *2017 IEEE International WIE Conference on Electrical and Computer Engineering (WIECON-ECE)* (IEEE, 2017), pp. 70–73
3. P. Rajpurkar, J. Irvin, A. Bagul, D. Ding, T. Duan, H. Mehta, B. Yang, K. Zhu, D. Laird, R.L. Ball, et al. Mura: large dataset for abnormality detection in musculoskeletal radiographs. arXiv preprint [arXiv:1712.06957](https://arxiv.org/abs/1712.06957) (2017)
4. M. Hall-Beyer, Glcm texture: a tutorial. Nat. Counc. Geogr. Inf. Anal. Remote Sens. Core Curriculum **3** (2000)
5. Zare, M.R., Mueen, A., Seng, W.C., Awedh, M.H.: Combined feature extraction on medical x-ray images, in *2011 Third International Conference on Computational Intelligence, Communication Systems and Networks* (IEEE, 2011), pp. 264–268
6. K. Simonyan, A. Zisserman, Very deep convolutional networks for large-scale image recognition. arXiv preprint [arXiv:1409.1556](https://arxiv.org/abs/1409.1556) (2014)
7. F. Chollet, Xception: deep learning with depthwise separable convolutions, in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1251–1258 (2017)
8. P.K. Mall, P.K. Singh, D. Yadav, Glcm based feature extraction and medical x-ray image classification using machine learning techniques. in *2019 IEEE Conference on Information and Communication Technology*, pp. 1–6 (2019)

Swarm Optimization of Multiple UAV's for Resource Allocation in Humanitarian Aid and Disaster Relief Operations



Anant Shukla , Rishav Choudhary , and Malavika Prabhuram

Abstract This paper presents a solution for Humanitarian Aid and Disaster Relief (HADR) using a swarm of UAV's which are capable of mapping the affected area and spotting human and livestock in the disaster struck area. It will send the surveyed information to the ground station and relief packages can be dropped at specific locations. The swarm consists of hybrid VTOL UAV's capable of vertical take-off and landing in rough terrain, in the absence of a landing strip. The flight time achieved is greater than an hour with a range of 50 km, which is unattainable with a conventional multirotor. The automated operation is implemented by pre-programming the system with a framework specific to the application. A learning network can be implemented to allow for better efficiency of system in multiple disaster scenarios. Flight stabilization is done with the help of a system of redundant IMU sensors computed on an on-board computer. The UAS is capable of localization and navigating in GPS enabled environments as well as, GPS denied environments with the help of a GPS, compass and inertial navigation fused with Visual Odometry. We use PSO to implement the swarm UAV's, it is required to mitigate mid-air collisions between UAV's in airspace and allocate work among them.

Keywords Swarm robotics · UAV's: unmanned aerial vehicles · HADR: humanitarian aid and disaster relief · Particle swarm optimization · Wireless mesh topology

A. Shukla · R. Choudhary · M. Prabhuram
SRM Institute of Science and Technology, Kattankulathur, Chennai, India
e-mail: as4698@srmist.edu.in

R. Choudhary
e-mail: rishavchoudhary_skc@srmuniv.edu.in

M. Prabhuram
e-mail: mp3602@srmist.edu.in

List of Abbreviations

UAV	Unmanned Aerial Vehicles
UAS	Unmanned Aerial Systems
HADR	Humanitarian Aid and Disaster Relief
LRS	Long Range System
VTOL	Vertical Take-off and Landing
PSO	Particle Swarm Optimization
GPS	Global Positioning System
IMU	Inertial Measurement Unit
AHRS	Attitude Heading Reference System
AMSL	Above Mean Sea Level
DOF	Direction of Freedom
IC	Internal Combustion
GCS	Ground Control System

1 Introduction

1.1 *Swarm UAV's in an HADR Operation*

Since our environment and population size is changing continuously there has been an increase in cases of natural as well as man caused disaster. The disasters which occur are varied in nature and complexity. To battle this, we have to be dynamic in our HADR approach. For this reason, use of UAV's is proposed.

To implement swarm UAV's, it is required to mitigate mid-air collisions between UAV's in airspace and allocate work among them. The other parts of the UAS are the ground station for control and data analysis, and the communication network between the operator and drones. We use Particle Swarm optimization in order to allocate the work and coordinate between each UAV. Moreover, the empty weight of the UAV should be less so more of the relief materials can be carried in one sortie. The UAV's should also be resilient enough to carry-out these tasks. To facilitate this, UAV's have been manufactured using materials like Carbon fibre, Glass Fibre and high-quality wood which are robust and have higher endurance value, enabling the UAV to operate in extreme conditions.

This paper discusses a swarm-based solution for UAV's which can be deployed in HADR. These disasters may vary in severity from Earthquakes, Floods, Forest Fires, Chemical spills to a wide range of disasters. These drones are capable of mitigating the disaster, searching for survivors, mapping the entire area or even dropping relief material.

2 Proposed Methodology

2.1 Development of Swarm System

For the final deployment platform, we have chosen a VTOL fixed wing hybrid design of 2.6-m wingspan. For the flight electronics, we have developed a GPS based autopilot, an AHRS and a 6 DOF Inertial Measurement Unit. We are also using a Long-Range System for providing telemetry and video communication up to 40 kms. This works on 915 and 433 MHz. This telemetry transmits data from payload sensors, position, flight information and autopilot.

A completely distributed approach is used in swarm intelligence-based strategy. Each UAV operates autonomously while collaborating with nearby UAV's to explore the disaster struck area and generate the map of the place, report the damage assessment and inform where relief is required. This occurs in three phases.

The first phase consists of a spread-out phase after take-off. This is done to ensure the UAVs are well placed to explore the area. Here we maximize the minimum distance between the UAVs, simultaneously flying in a defined radius. Next comes the monitoring phase, where the UAVs transmit the sensed data via communication link to other agents. As soon as one of the UAV's detect damage or a human in distress, the third phase begins. This is the search stage where the planes begin collaborating with each other to search the surroundings for other humans in distress, such as humans below building rubble. To select the optimal direction for continuing, each plane uses the data received from N nearest neighbours along with its own data.

To a certain extent, group behaviour is controlled by the decision gap parameter that in turn controls the ratio of exploration to exploitation for search technique. The plane displacement between changes in direction is small for a low value of decision gap parameter, which increases its dependency on other planes in the team. If the parameter is higher, the displacement between direction change is also higher. Hence, the exploration capacity of the strategy is increased which in turn increases probability of detecting more promising areas and parameter D represents the ratio of exploration exploitation to some extent. As a result, when there is a lower number of UAV's, the algorithm is benefitted by an increased D parameter in order to increase its exploratory capabilities and avoid premature convergence to sub-optima (Figs. 1 and 2).

2.2 Flight Control System and Communication Stack

The Ground Control Station communicates between the ground station controller and the UAV's and receives the visual data from the on-board cameras. The commands for UAV's are sent from the Ground Control Station. We are running 2 PC's as Ground Control Computers. One of them is responsible for receiving telemetry data and running scripts for the UAV's pertaining to path planning, etc. The second one

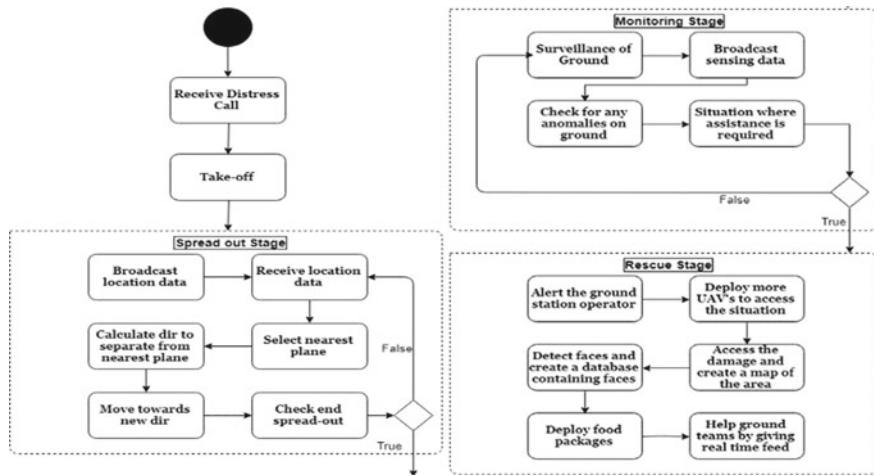


Fig. 1 The three phases in which the HADR action will be taken



Fig. 2 Ground communication station

is responsible for image and video acquisition. The script on the second running PC is responsible for image processing and facial recognition.

2.3 GPS Inaccessible Environment

UAVs are reliant on precise localization for navigating safely and for locating the destination. Generally, UAV's use GPS for complementing inertial sensing in order to provide navigation without drift. Even though this is an easy answer, GPS is not

an entirely dependable sensing technique. This is because obstruction and multi-paths created by skyscrapers and mountains may break down the satellite coverage. Strategies such as utilizing apriori maps and simultaneous localization and mapping (SLAM) for building maps on the fly are to remain localized with onboard sensors.

3 Mechanical Design

3.1 Overview

The UAS which we have proposed consists of a UAV which is a hybrid VTOL airplane. The characteristics of the structure include high manoeuvrability in narrow areas and flight time achieved is over an hour, that is unattainable in conventional multirotor carriers. By only using the rotors for take-off, landing, and hovering, the aircraft becomes more robust to cross winds. The drone has been designed so as to achieve a service ceiling of 4000 ft AMSL with an operating range of 50 km.

3.2 Frame

It has a customized frame of carbon fibre and glass fibre with wings made of wood. These materials help in increasing the power to weight ratio. Hence, the shape of the aircraft body is optimized to improve lift (Fig. 3).

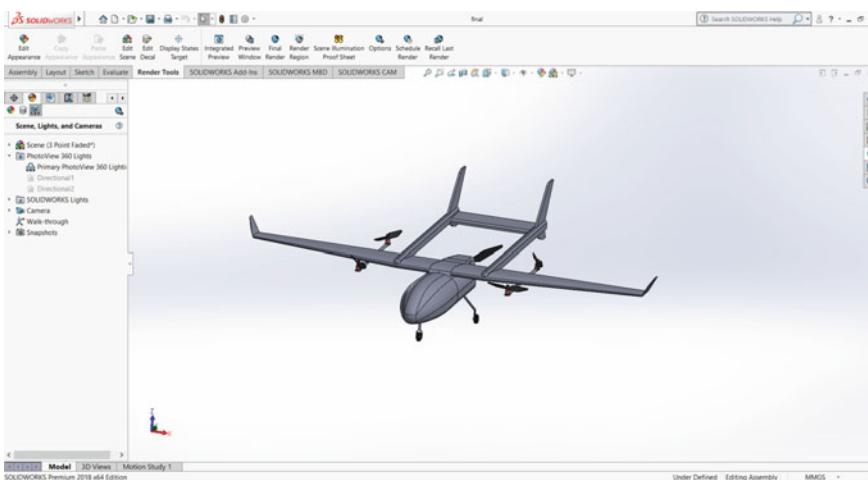


Fig. 3 The proposed design of UAV

3.3 VTOL System

The UAV utilizes a VTOL system which facilitates take off from a level surface and there is no requirement of a landing strip which makes this UAV ideal to be used in HADR region. VTOL is facilitated by incorporating a combination of four motors in H configuration and an overhead boom for flight stabilization.

3.4 Propulsion

The four motors are mounted below the fuselage to provide a vertical thrust and a single 30 cc IC engine is mounted on the fuselage in a pusher configuration to provide the necessary forward thrust.

3.5 Modes of Operation

The aircraft operates in three modes, VTOL, Transition and Fixed wing.

VTOL: In this phase, the UAV uses the set of four motors to lift, powered by lithium polymer batteries, which provide the required thrust to lift the UAV.

Transition phase: In this phase, the UAV switches from vertical to forward flight and vice versa.

Fixed wing: Once the gasoline/petrol engine starts, it propels the UAV in the forward direction and the four rotors are stopped.

3.6 Electronics Housing

Electronics components are housed in fuselage and all components have been placed accordingly for the proper weight distribution. The bottom of the fuselage has the holder for camera to give a live feed of the area under surveillance. Another bracket is designed which houses the drop mechanism to drop the relief materials (Table 1).

4 Conclusion

This paper presents a UAV swarm solution for coordinating multiple UAV's that can be deployed for Humanitarian Aid and Disaster Relief. With the help of this system, we can deploy a swarm of drones which can do the surveillance of the disaster

Table 1 UAV specification chart

Parameter	Value
Wingspan	2600 mm
Wing area	0.88 m ²
Empty weight	6.5 kg
Maximum take-off weight	15 kg
Payload	4 kg
Cruse speed	120 km/h
Stall speed	35 km/h
Endurance	~90 min
Service ceiling	4000 ft AMSL
Power system	Gasoline electric hybrid

struck area, broadcast their status to surrounding UAVs and the ground station and determine the most promising direction to move in.

Hardware Design of a Turbo Product Code Decoder



Gana C. Nair, B. Yamuna, Karthi Balasubramanian, and Deepak Mishra

Abstract In a communication channel, message transfer happens through a noisy medium which can introduce errors in the transmitted message. It is required to maintain an acceptable bit error rate for a reliable transmission. This can be handled by an efficient error control coding scheme. This paper discusses the hardware design of an error control decoder, namely turbo product code decoder. The iterative Chase-Pyndiah decoding algorithm has been used in the turbo decoder design. The decoder has been designed using both MATLAB and Verilog, and its performance has been analysed. To study the robustness of the decoder in dealing with different data types, performance analysis is done with both the received data and the encrypted version of the same. Advanced Encryption Standard (AES) has been used for the encryption process, and it is shown that there is only a negligible difference in the performance for both the data sets.

Keywords Turbo product code · Advanced encryption standard · Chase-Pyndiah decoder

1 Introduction

An efficient wireless communication system should facilitate error-free reception and data security across a communication channel. This can be achieved to an extent

G. C. Nair (✉) · B. Yamuna · K. Balasubramanian

Department of Electronics and Communication Engineering, Amrita School of Engineering,

Coimbatore, Amrita Vishwa Vidyapeetham, India

e-mail: ganacnair@gmail.com

B. Yamuna

e-mail: b_yamuna@cb.amrita.edu

D. Mishra

Digital Communication Division (DCD), Space Application Center (SAC),
ISRO, Ahmedabad, India

e-mail: deepakmishra@sac.isro.gov.in

by combining error control coding (ECC) with an efficient network security protocol. Error control coding enables error detection and correction and helps in retrieving the transmitted message that has been corrupted by noise.

A good communication system should be intelligent enough to detect and minimize the error rate in the transmitted message that can occur due to the channel noise. This can be achieved using error control coding techniques such as Hamming codes, cyclic codes, BCH codes, Reed Solomon codes (RS), linear block codes (LBC), convolution codes, turbo codes, turbo product codes and low-density parity check codes [1]. These codes differ from each other in terms of their error correcting capability and complexity level. Cryptographic encryption techniques are basically used to ensure secure data transmission across a network so as to protect the data from the reach of any unauthorized third party or a hacker. Subjecting encrypted data to ECC ensures a secure reliable communication without performance degradation.

One of the widely used error control codes in communication systems is the turbo product code (TPC). The iterative Chase-Pyndiah algorithm is widely used for decoding turbo product codes [2]. An improved low complex and optimized Chase decoding algorithm for BPSK modulation was proposed and implemented on field programmable gate array in [3]. Hardware-oriented TPC and a method for construction of TPC have been proposed in [4], where the limitations of the hardware and the reuse of hardware modules have been taken into account in the decoding process. In order to reduce the TPC decoding complexity, an efficient syndrome-based hybrid decoder for TPC has been proposed in [5]. A new decoding algorithm to enhance the performance of the decoder and to significantly reduce the decoding complexity considering encoding and decoding aspects for an extended TPC has been proposed in [6]. An innovative ultra-high-speed block turbo decoder architecture with efficient memory usage using an extended BCH code has been proposed in [7]. An iterative decoding algorithm with flexibility in performance-complexity trade off for block turbo codes has been discussed in [8].

Data theft is a matter of concern when it comes information transmission. If the transmission is not secure, data can be misused and can be used for unethical practices. Data encryption is one of the cryptographic techniques used to ensure secure communication. AES is a standard secure technique for encryption and decryption and is widely used in cryptographic applications [9]. Systems to ensure both reliable and secure communication by using turbo codes on AES cipher text have been proposed in [10]. Further, it has been shown in [11] that the use of TPC and AES brings about an improved BER performance.

This work involves hardware design of a TPC decoder using Chase-Pyndiah algorithm and its performance analysis with AES encrypted data. The paper is arranged as follows. Section 2 briefly introduces decoding of turbo product codes. The hardware design of the Chase decoder is discussed in Sect. 3 followed by results and analysis in Sect. 4. The paper concludes in Sect. 5 with prospective future work.

2 TPC Decoding

The iterative Chase-Pyndiah decoding algorithm involves row decoding followed by column decoding or vice versa as shown in Fig. 1. The completion of one row and one column decoding operation represents one full iteration. The decoder shows improved performance due to the iterative process involved [12].

3 Hardware Design of TPC Decoder

The encoded data of TPC $(31, 21)^2$ BCH is arranged in a matrix of dimension $[31 \times 31]$, on which BPSK modulation is performed and is transmitted over an AWGN channel. The noisy encoded data is converted to 32-bit floating point representation using IEEE 754 single precision format and given as input to the decoder.

The block diagram for the hardware realization of the TPC decoder is shown in Fig. 2.

The top-level design includes a memory module along with the main TPC decoder module. The decoder consists of different sub-modules for computations of least reliable bits (LRB), hard values, test pattern generation, candidate codeword generation, syndrome decoding, Euclidean distance, and extrinsic information. The received data is stored in a memory and is later fetched by the decoder module to perform the row and column decoding.

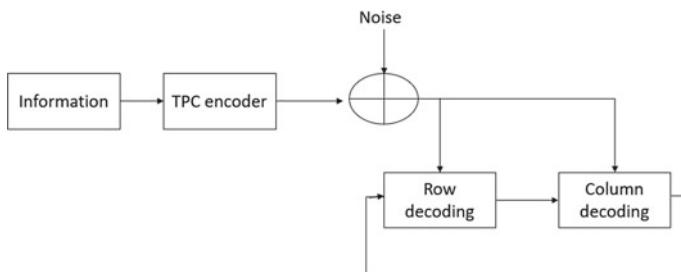


Fig. 1 Iterative Chase-Pyndiah TPC decoder[11]

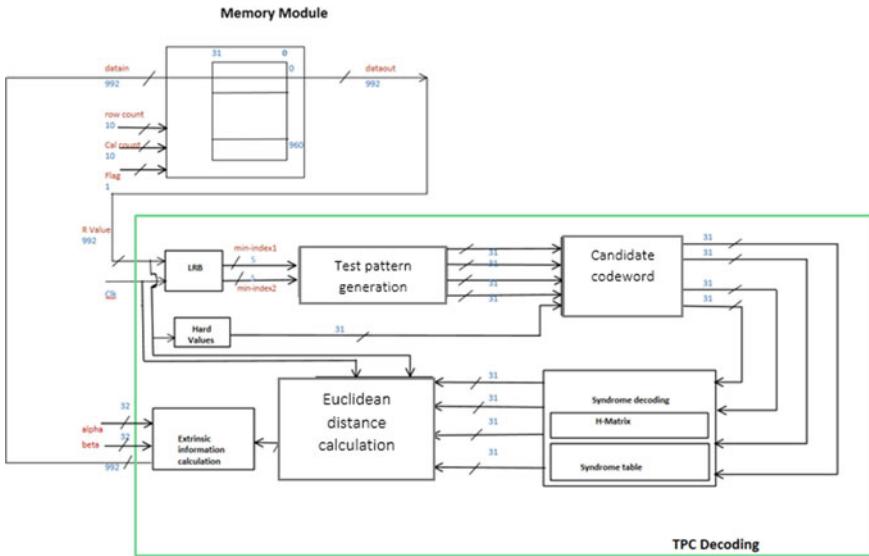


Fig. 2 Hardware design of Chase decoder

The decoding operations include the extrinsic information calculation and the updation of the decoded values. Two half iterations are performed, where the alpha and the beta values for each iteration are passed correspondingly. Row decoding is done in the first half iteration and column decoding in the second half iteration. After each half iteration, the decoded values are updated in the memory. The 31 row/column values are expressed as 32-bit floating point numbers and during each row/column decoding, and they are accessed from the memory for processing.

In each of the row/column decoding first, p least reliable bit positions are determined (where $p = d_{\min}/2$) from the received sequence by identifying the p least reliable input bit values. For the constituent code BCH(31, 21), d_{\min} is 5 and hence 2 least reliable positions have to be determined. The received data R is then converted to hard values, by making use of the sign bit of the data. If sign bit is 1, the hard value is 0 and vice versa. For p least reliable positions, 2^p test patterns are generated by placing all combinations of zeros and ones in the p positions and filling all other bit positions with zeros. For the case of BCH (31, 21)² constituent code, four test patterns are generated, with four combinations of zeros and ones in the p least reliable positions. The 2^p test patterns obtained are XORed with the hard values to generate 2^p candidate codewords. In this case, four candidate codewords are generated. These candidate codewords are decoded using syndrome decoding. This is followed by calculation of the Euclidean distance between each of the candidate codewords and the received data, and the codeword with the least distance is taken as the decoded codeword D . Then, extrinsic information is calculated using prede-

fined values of alpha and beta, and the R value is correspondingly updated [2]. This updated R value is written back in to the input memory for the next half iteration. This set of operations is repeated for all the rows and columns of R .

4 Results and Analysis

The information to be transmitted is encoded using BCH(31,21) turbo product code. This data is BPSK modulated and transmitted through a noisy channel. At the receiver side, BPSK demodulation is done, and the demodulated data is decoded using iterative Chase-Pyndiah algorithm. The simulation of the TPC decoder is first done in MATLAB, and BER performance is analyzed. The decoder is then designed in hardware using Verilog and tested with the noisy data generated from the MATLAB simulation set-up [13]. This entire procedure is repeated with AES encrypted data to analyse how the decoder's performance gets affected by security enhanced data. The results obtained with the two simulation set-ups are discussed below. Figure 3 shows the BER plot of the TPC decoder in MATLAB. The BER plots have been shown for both encrypted and regular data, and it can be seen that the decoder performs equally well under both set-ups. Figure 4 shows the comparative BER plots of the TPC decoder designed in MATLAB and Verilog. Using AES encrypted data for verification, it can be seen that the performance of the hardware model matches exactly with the MATLAB simulation model, thus validating the design of the hardware model.

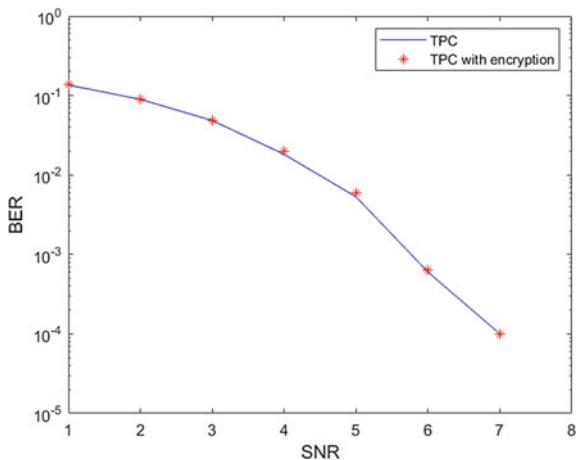


Fig. 3 BER plot of TPC with and without encrypted data showing negligible difference in performance for both systems

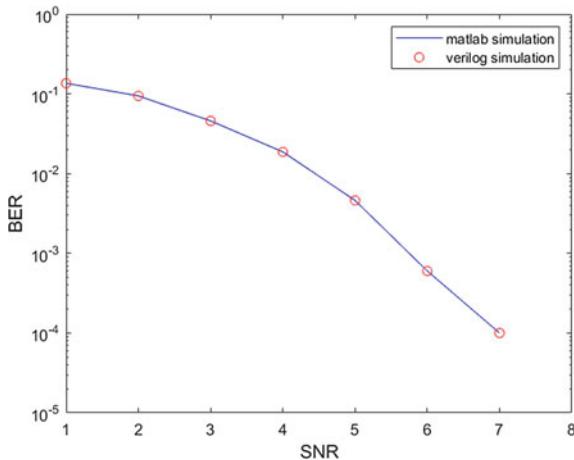
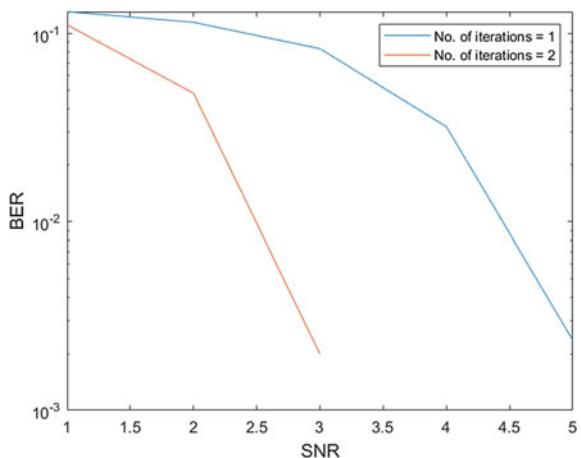


Fig. 4 Comparative BER plots of TPC with encrypted data for Matlab and Verilog simulation. Both the plots overlap perfectly, thus establishing the correctness of the decoder implementation

Fig. 5 BER plots of TPC decoder with AES encrypted data for multiple iterations. As iteration increases BER performance shows a considerable improvement. Beyond iteration 2, the BER is zero



4.1 Effect of the Number of Iterations

To understand the effect of the number of iterations of the TPC decoder on its performance, analysis was done with multiple iterations. Figure 5 shows the BER plots of the TPC decoder with multiple iterations. It can be seen that there is a considerable improvement in BER performance on increasing the number of iterations. By the third iteration, it is shown that BER is completely zero for all range of SNRs.

5 Conclusion and Future Work

This paper presented MATLAB and hardware models of a turbo product code decoder using Chase-Pyndiah algorithm. Simulation analysis was performed using the received data and also with the AES encrypted version of the same. Similar results with both data sets show that the decoder is robust enough to handle encrypted data also. It is also shown that there is an improvement in performance as the number of iterations increases. It is envisaged that this design may be realized on a hardware platform like FPGA and analysed in real time. Turbo product code with five different modes of advanced encryption standard can also be analysed and implemented as possible extensions of this work.

References

1. T.K. Moon, Error correcting Codes-Mathematical Methods and Algorithms (Wiley, NJ, 2005)
2. R.M. Pyndiah, Near-optimum decoding of product codes: block turbo codes. *IEEE Trans. Commun.* **46**(8), 1003–1010 (1998)
3. W. Kuang, R. Zhao, Z. Juan, FPGA implementation of a modified turbo product code decoder, in *2017 IEEE 9th International Conference on Communication Software and Networks (ICCSN)* (IEEE, 2017), pp. 71–74
4. Y. Krainyk, V. Perov, M. Musiyenko, Y. Davydenko, Hardware-oriented turbo-product codes decoder architecture, in *2017 9th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS)*, vol. 1. (IEEE, 2017), pp. 151–154
5. E.-H. Lu, P.-Y. Lu, A syndrome-based hybrid decoder for turbo product codes, in *International Symposium on Computer, Communication, Control and Automation (3CA)*, vol. 1. (IEEE, 2010), pp. 280–282
6. C. Xu, Y. Liang, W.S. Leon, A low complexity decoding algorithm for extended turbo product codes. *IEEE Trans Wireless Commun* **7**(1), 43–47 (2008)
7. C. Leroux, C. Jégo, P. Adde, M. Jézéquel, High-throughput block turbo decoding: from full-parallel architecture to fpga prototyping. *J. Sig. Proc. Syst.* **57**(3), 349–361 (2009)
8. S. Dave, J. Kim, S.C. Kwatra, An efficient decoding algorithm for block turbo codes. *IEEE Trans. Commun.* **49**(1), 41–46 (2001)
9. N.-F. Standard, Announcing the advanced encryption satandard (AES). *Federal Inf. Proc. Standards Publ.* **197**(1–51), 5–47 (2001)
10. P. Samundiswary, Performance analysis of turbo coding with AES for CCSDS standard. *Int. J. Comput. Appl.* **96**(12), 18–21 (2014)
11. B. Mishra, T. Chaware, Turbo product code and AES encryption for wireless communication, in *2013 Nirma University International Conference on Engineering (NUiCONE)*. (IEEE, 2013), pp. 1–5
12. S.N. Vaniya, N. Kumar, C. Sacchi, Performance of iterative turbo coding with nonlinearly distorted ofdm signal. in *IEEE Annual India Conference (INDICON)* (IEEE, 2016), pp. 1–5
13. A. Ambat, K. Balasubramanian, B. Yamuna, D. Mishra, Fpga implementation of an efficient high speed max-log-map decoder. in *2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI)* (IEEE, 2018), pp. 747–751

Multiple Object Detection and Tracking Using Deep Learning



Shreyas Burde and Suneeta V. Budihal

Abstract The paper proposed a framework to design and develop object detection and tracking using learning based algorithms. It is an important event in security and surveillance systems using images or videos. YOLO V3 model is used for object detection and is trained using coco dataset. The objects detected are Humans, bottles, Drilling machine, Air powered saw, etc. The model detect the objects from camera field view, i.e. objects that are clearly viewed. The proposed model with pre-processing, provides an average accuracy of around 96% with different test cases.

Keywords Object recognition and tracking · Deep learning · YOLO V3

1 Introduction

Image recognition and tracking system plays a major role in the day to day life with a wide spread application in hotels, factory, warehouses, airports, streets, malls, etc. Unattended or missing object detection becomes more difficult due to high volume of data, different weather conditions, illumination, occlusion and motion backgrounds through image processing. The traditional object recognition system did not provide the flexibility to analyse the objects, which makes it hard to detect objects both in real time as well as off line. Traditional surveillance system was manually controlled and was always affected by human errors. A survey demonstrated that, object detection and tracking with Machine learning/ Deep learning solutions progressed major in computer vision and graphics domains.

The state of the art literature survey is carried out on object detection and tracking. The authors [1], used new YOLO architecture for object detection and it was framed as regression problem in order to separate bounding boxes and it's probabilities. In

S. Burde · S. V. Budihal (✉)
KLE Technological University, Hubballi, Karnataka, India
e-mail: suneeta_vb@kletech.ac.in

S. Burde
e-mail: shreyasburde@icloud.com

paper [2], the authors provided a background research information of deep learning and YOLO architecture and created a protective image database for training and testing. The authors [3] tried to encounter the problems that occurred during detection such as slow processing, usage of light weight data sets, etc. and in order to overcome these issues, YOLO V2 model was used. The authors [4] introduced a YOLO 9000 architecture for object detection, capable of detecting over 9000 categories of objects. Authors [5] compared two Neural Networks(NN) for ball detection in non staged condition for their performances. Authors [6] proposed various improved structural frameworks for the YOLO algorithm and also proposed a new structural network framework for YOLO-R architecture.

The authors [7] used YOLO algorithm, for counting vehicles for real time traffic management. The authors [8] proposed a novel algorithm for the pedestrian detection with YOLO and Gaussian Mixture Model for background detection leading to improved object detection. In paper [9] for multi class object detection YOLO V2 was used, capable of detecting multiple classes. The authors [10] proposed usage of YOLO algorithm for the purpose of Chinese traffic sign detection. The literature survey revealed that the dataset requirement of the proposed architecture is more with the accuracy which can be enhanced further using a better model.

2 The Proposed Framework

A maximum of four classes were considered for multiple object detection and tracking. Sample objects considered for multi class classifier are bottle, mobile, plat, tools, etc. The total number of images used are distributed shown in Table 1. It provides the complete information regarding 5000 images used for training and validation. Every image augmented output is a subset of original dataset. The variation of images helped the model during training phase to learn the detailed features of object to cater under-fitting and over fitting [11], but lowers the accuracy. The images are annotated by putting images of different class into different folders. YOLO v3 is the third generation of the YOLO [12] family, which mainly used logistic regression for the prediction of the objects in the frame. During counting up of classification losses, the

Table 1 Summary of the number of datasets used for training and testing

Class	Training	Testing	Total
Mobile	698	60	758
Human	1020	150	1170
Bottle	430	90	520
plant	923	50	973
Drilling machine	624	150	774
Air powered saw	732	150	882
Total	4427	650	5077

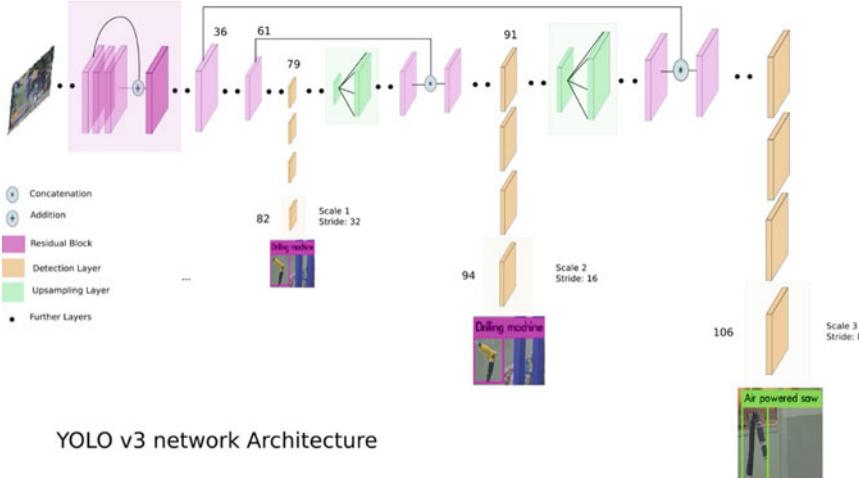


Fig. 1 YOLO V3 architecture

model used binary cross entropy loss for each label and reduced computational twist by avoiding SoftMax features [13]. In the case of YOLO v3 model as the prediction of the output takes place at 3 different sectors or stages and has been illustrated in Fig. 1. Each grid make 3 predictions using 3 anchor boxes, so there are total 9 pre-defined anchor boxes and these are chosen using K-means clustering. To compute the loss for true positive, only one of them is responsible for the object. Hence, highest IoU (Intersection over Union) with the ground truth is selected. Every prediction improves at predicting certain sizes and aspect ratios. The loss function comprises of the following losses, i.e. localization loss, classification loss and confidence loss.

$$\sum_{i=1}^{S^2} \mathbb{P}_i^{\text{obj}} \sum_{c \in \text{classes}} (P_i(c) - \hat{P}_i(c))^2 \quad (1)$$

Classification loss at each cell is squared error of the class conditional probabilities for each class, when an object is detected. The localization loss is the error in the predicted boundary box locations and sizes [14, 15].

$$\lambda_{\text{coord}} \sum_{i=0}^{S^2} \sum_{j=0}^B \mathbb{P}_{ij}^{\text{obj}} [(x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2] + \lambda_{\text{coord}} \sum_{i=0}^{S^2} \sum_{j=0}^B \mathbb{P}_{ij}^{\text{obj}} [(\sqrt{w_i} - \sqrt{\hat{w}_i})^2 + (\sqrt{h_i} - \sqrt{\hat{h}_i})^2] \quad (2)$$

$$\sum_{i=0}^{S^2} \sum_{j=0}^B \mathbb{P}_{ij}^{\text{obj}} (C_i - \hat{C}_i)^2 \quad (3)$$

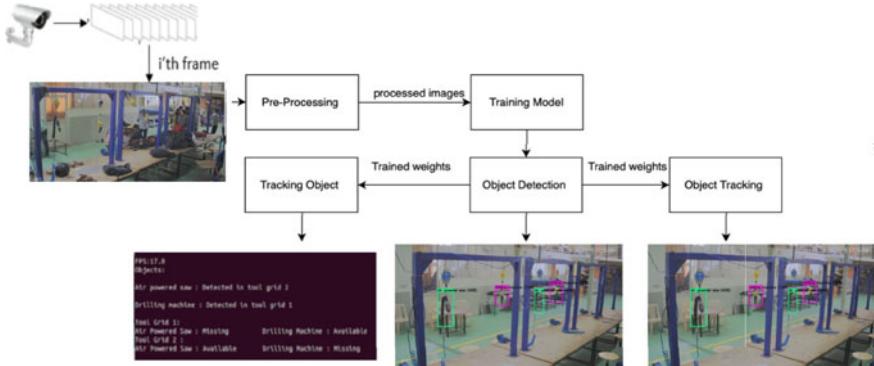


Fig. 2 The detailed block schematic representing the sub blocks of the proposed framework

$$\lambda_{\text{noobj}} \sum_{i=0}^{S^2} \sum_{j=0}^B \mathbb{1}_{ij}^{\text{noobj}} (C_i - \hat{C}_i)^2 \quad (4)$$

The final multi-path loss adds to all the four equations, i.e. Equations 1, 2, 3 and 4. The localization loss measures the errors in the predicted boundary box locations and sizes. Object tracking is the process of tracking an object trajectory in a plane moving around a scene. Initially, an object assigned with an alphabet or a number is identified. That alphabet or number for each object throughout the video sequence is maintained by keeping a track of them while moving. If the object with the alphabet or number from previous frame is changed then particular object is being moved. The detailed block schematic of the proposed framework is as shown in Fig. 2.

3 Results and Analysis

For object detection and tracking video surveillance of a laboratory is considered. Initially, the v3 model is trained to detect custom objects, i.e. drilling machine, air powered saw, man, etc. First v3 model used Darknet system which is made up of 53 layers and is capable of object detection and also has additional 53 layers stacked onto it, and thus it gives 106 layers fully convolutional architecture. The model detects at three different scales. The model is a CNN model and its output is generated at feature map by applying a 1×1 kernel on to it, and this occurs at three different sizes at three different places in the network. Initially, the model is trained with YOLO v3 architecture. Transfer learning and fine tuning techniques are used to train the model. On applying fine tuning techniques from pre trained Imagenet, accuracy is increased. In fine tuning, the whole base model was not trained, instead the pre-trained weights of ImageNet were considered and trained the model by giving the generated dataset.

3.1 Results of Pre-processing

Initially, a video is read and converted it into sequence of frames. Then histogram equalization is used to improve the contrast. Image per processing from the surveillance video is shown in Fig. 3. we have conducted several training tests in order to attain maximum accuracy in the case of detection by using random 10 trained images and results are shown in Table 2. It is clearly observed that with pre-processing, improved accuracy is attained. The images used for training the system are taken from video surveillance camera. Target objects for training were human, mobile, plant, bottle, Drilling Machine, Air powered saw. Number of images used for training are 5077.

3.2 Results of Object Detection and Tracking

We address the problem of detection and tracking and make following contributions. First one, we proposed a framework to train and test the object detection model with pre-processed images in the dataset. Second one, we proposed the method to find missing object from the field of view of IP camera. To find the missing object from the frame, Initially, the video surveillance frame is resized to 1920×1080 . The system first segregates the tool grids based on their coordinates in the frame. The results are as shown in Figs. 4 and 5 for the various trained objects.



Fig. 3 Results of pre-processing

Table 2 Results of pre-processing for 10 random images

Number of images	Without pre-processing (%)	With pre-processing (%)
Drilling machine 1	88	97
Air powered saw 1	86	92
Drilling machine 2	59	77
Air powered saw 2	89	95

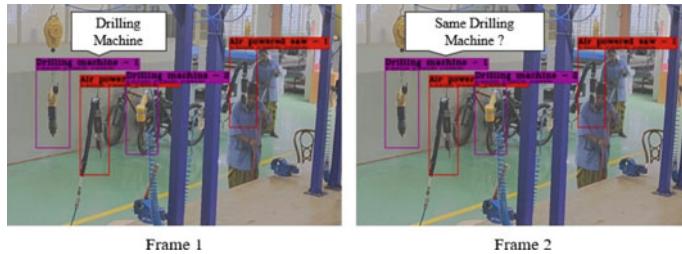


Fig. 4 Object tracking



Fig. 5 Results of object detection and tracking

4 Conclusion

The paper proposed a framework to detect the tools in the video frame and identify the missing tools in a laboratory using Yolo v3 object detection model and also notify the user through mail. The model is tested for different test cases. The system gives expected output of all the test cases accurately with an accuracy of 95%. This work can further be developed to detect multiple tools from all the tool grids using multiple IP cameras.

References

1. J. Redmon, et al., You only look once: unified, real-time object detection, in *The IEEE Conference on CVPR* (2016), pp. 779–788
2. S. Wang, et al., Research on image recognition of insulators based on YOLO algorithm, in *International Conference on PST*, pp. 3871–3874 (2018)
3. A.P. Jana, et al. YOLO based detection and classification of objects in video records, in *3rd International Conference on RTEICT, India* (2018), pp. 2448–2452
4. J. Redmon, et al., *The IEEE Conference on CVPR*, pp. 7263–7271, 2017
5. M. Buric, et al. Ball detection using yolo and mask R-CNN, *International Conference on CSCI, USA*, pp. 319–323
6. W. Lan, et al. Pedestrian detection based on YOLO network model, in *ICMA*, pp. 1547–1551 (2018)

7. C.S. Asha, et al. Vehicle counting for traffic management system using YOLO and correlation filter, in *IEEE ICECCT* (Bangalore, 2018) pp. 1–6
8. Q. Peng, et al. Pedestrian detection for transformer substation based on GMM and YOLO, in *8th International Conference on IHMSC*, pp. 562–565 (2016)
9. K. Jo, et al. A real-time multi-class multi-object tracker using YOLOv2, in *2017 IEEE International Conference on SIPA* (Kuching, 2017), pp. 507–511
10. J. Zhang, et al. A real-time Chinese traffic sign detection algorithm based on modified YOLOv2 *Algorithms* **10** *4*, 127 (2017)
11. J. Redmon, A. Farhadi, *The IEEE Conference CVPR*, pp. 7263–7271 (2017)
12. R. Girshick, et al. *The IEEE Conference on CVPR*, pp. 580–587 (2014)
13. J.R.R. Uijlings et al., Selective search for object recognition. *Int. J. Comput. Vis.* **104**, 154–171 (2013)
14. S. Pavaskar et al., Real-time vehicle-type categorization and character extraction from the license plates. *CISC* 557–565 (2019)
15. V.B. Suneeta, P. Purushottam, K. Prashantkumar, S. Sachin, M. Supreet, Facial expression recognition using supervised learning, in *Computational Vision and Bio-Inspired Computing*, ed. by S. Smys, J. Tavares, V. Balas, A. Iliyasu. ICCVBIC 2019. Advances in Intelligent Systems and Computing, vol 1108 (Springer, Cham, 2020). https://doi.org/10.1007/978-3-030-37218-7_32

Blood Donation Management System Using Android Application



Ahana Bandyopadhyay , Ishita Kundu , Anwita Chakraborty , Rajat Kumar , Ayush Kumar , and Sukanta Sabut

Abstract Availability of blood during a crisis is crucial for every human being. Numerous nations are confronting difficulties in keeping up with an adequate supply of blood and guaranteeing a swift association between a donor and recipient, thereby serving as a disadvantage in the current framework. In this study, accentuation has been laid out on application improvement for portable devices utilizing available software and hence is valuable for clients having low device memory. In our proposed application, time to perform the task is minimal, and all authentic information about the donors is notified in the recipient's mobile devices instantaneously, thereby outlining an escalated participation and proper communication between the recipient and donor. Our proposed framework of blood management through an Android application would be at ease, and wastage of blood could be minimized at a higher extent.

Keywords Donor · Recipient · Android application · Blood management

1 Introduction

Blood is a critical aspect of survival for every living organism and constitutes 8% of the total human body weight. A comprehensive analysis suggests that 119 out of 195 countries face a staggering shortage of blood. Further for maternal mortality, injuries and prevalence of inherited blood disorders like thalassaemia and sickle-cell anemia, blood transfusion plays a significant role. However, contrasting practices among blood banks can steer hazardous results. Blood wastage is also a major problem causing healthcare providers battle an acute shortage of blood units for various medical procedures. Given this backdrop, the significance of deploying technologies that encourage efficient management of blood banks cannot be understated. Therefore, to provide better assistance to the person in search of blood, a highly efficient and systematic management framework needs to be devised which will reduce the time required to look out for blood donors.

A. Bandyopadhyay () · I. Kundu · A. Chakraborty · R. Kumar · A. Kumar · S. Sabut
KIIT University, Bhubaneswar, India
e-mail: ahanabandyopadhyay1998@gmail.com

2 Related Works

Blood is an essential healthcare asset which portrays an expert usage in clinical services. Muralidaran et al. [1] designed a smart blood bank as a service on cloud which involves a Web application hosted on cloud to provide easy access to reports and store the data related to the donors inside the cloud database. To address the issue of effective blood quality management, Jiang et al. [2] proposed a dynamic blood information management system based on RFID that guarantees the quality of donated blood and efficiency of operation management. Rahman et al. [3] took a similar approach to provide a location-aware mobile phone-based blood donor recruitment technique which aims at promoting the quality and efficiency of quality management. To make the ends meet for both the donor and the recipient, Fahim et al. [4] developed an Android-based blood donation application as mHealth solutions to create a link between the donor and the recipient. In [5], Abhijeet Moharkar et al. proposed a method to find a composed blood pack with its blood parts (RBC, platelet, plasma) and offer it to the patient/customer in the required time range. The framework comprises Android application for customers/patients and Raspberry Pi B+ as a PC for blood donation administrator. In [6], Moh. Nabil et al. proposed a Cloud medical monitoring and Web-based blood donation system which permits the recipients to request blood from the blood donation centers. In [7], J. Vaqué proposed a system of blood donors recording framework. It is an informative framework upheld by a central computer to which are associated the terminals arranged in the separate blood donation centers. Ilias Fotopoulos et al. [8] developed an information system that uses cloud computing to empower blood donation centers and blood assortment offices to effectively enroll a sound blood benefactor or donor's pool and radically increment the degree of consistency by utilizing the concept of gamification. In [9], Soo-Jung Kim et al. proposed a blood observing and its proper management framework which can be deployed for usage in emergency clinics or hospitals. This framework will constantly report the temperature of the blood donation center fridge, track the area of a blood pack to build staff activity effectiveness, and can affirm that the doled out blood sack was shipped properly without any error to the recipient for timely transfusion.

3 Proposed Workflow

In this study, an Android-based application is designed to facilitate the process of inspection for the nearest donor by integrating Google Maps. The database is implemented in Firebase to store necessary information and retrieve real-time data which eases communication with the recipient through an Android-Firebase push notification. The flowchart for the application is shown in Fig. 1.

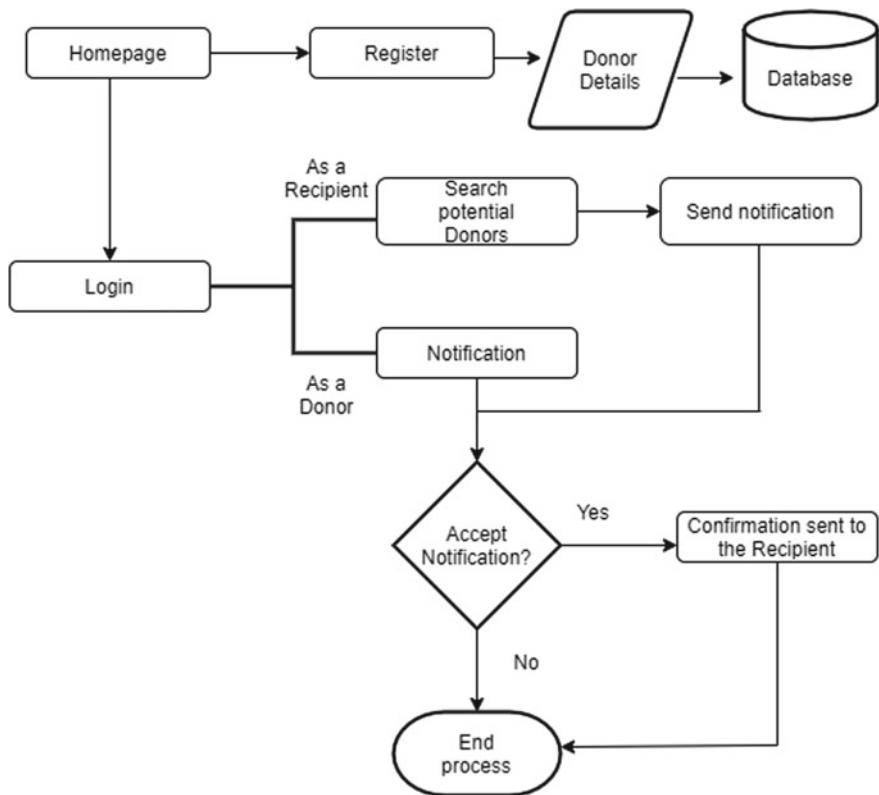


Fig. 1 Flowchart for the proposed workflow

3.1 Application Layout

The application layout is developed in Android Studio using XML which is a markup language that characterizes a set of rules for encoding records in a configuration that is both comprehensible and machine-coherent. Various views have been incorporated for designing the frontend namely ImageView, TextView, and Buttons which are placed inside a Relative Layout. The ImageView is utilized to showcase the images using numerous styling components and to give an appropriate position in a specific page. TextView is a user interface component that displays text to the client. In case of drop-down menus, Spinners are used to accelerate the process of selecting one incentive from the set. Intents have been used to link the different pages and to successfully integrate with the back-end, Kotlin, a cross-platform, statistically typed general-purpose programming language has been used. The accessibility of the designed application is escalated by integrating several Web technologies such as Firebase Cloud Messaging and APIs for push notification and real-time GPS tracking.

3.2 Firebase Real-Time Database

A Firebase real-time database has been employed to assist the process of collection, management, and modification of user data. It is a cloud-facilitated NoSQL-based database where the structured data gets reserved in terms of JSON documents. The real-time functionalities give synchronization across associated devices and are accessible when there are no system networks through a local cache. To sync the project with Gradle and integrate Firebase, our application was associated with the Firebase application in the Firebase console which is redirected to a connection dialog where a recent Firebase project can be created and a real-time database containing insights regarding the benefactor and its related blood groups can be correlated. To fetch and display the list of all potential donors and their blood groups in a particular arena, a layout for the list view is established, and hence, a new layout resource file is developed. The application further actualizes an onStart strategy which empowers execution every time any information is changed inside the database.

3.3 Google Maps

The fundamental thought process of consolidating Google Maps into the application is to diminish the time taken to look for donors and get an extensive idea of the number of contributors present nearby. The Google Maps application programming interface (API) is utilized to get to Google's servers. An Android restricted API key is framed which is free and supports a number of clients, and a fragment element is employed to the layout file for the activity. The OnMapReadyCallback interfaces with the assistance of the on MapReady callback technique that is executed to obtain a Google Map object portraying a real-time map. To set the type of map to normal, Google Map object's setMapType() method is called, and one of the type constants characterized in Google Map is passed.

3.4 Firebase Cloud Messaging

The proposed application uses Android push notification to send relevant notifications to the registered donors of that locality. Firebase Cloud Messaging is a cross-platform where notifications can be sent utilizing the Firebase Admin Software Development Kit (SDK) or Firebase Cloud Messaging (FCM) server by generating a notification message. The application acts distinctively based on the fact whether it is available on background or foreground and sent notifications to targeted users. When the application is operated in the Background, Firebase Software Development Kit automatically processes the message and displays it as a notification in the donor device's framework. If the application receives an FCM message while its in the

foreground, then by default this message will not be exhibited to the user. To follow up on the messages when the application is in the foreground, Firebase Messaging Service should be extended, and the onMessageReceived technique lastly recovers the message substance by utilizing getNotification.

4 Result and Comparative Analysis

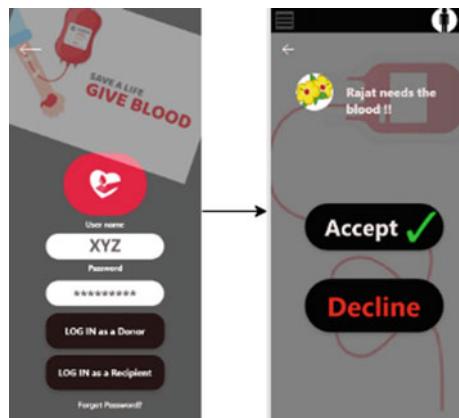
The suggested architecture of the application in terms of donor and recipient has been visualized in Figs. 2 and 3, respectively. The introductory page redirects the user to either the login or registration page, according to the need. The user is advocated to provide the fundamental details that will get added in the database and will aid in faster information processing. Google Maps recognize the nearness of registered potential donors in the desired location and display the necessary subtleties utilizing which the recipient can easily contact the donor. The recipient has the freedom to select any contributor as per their choice following which a notification will be sent to that individual giver following which the donor will receive a blood donation request. Upon acknowledgement, the recipient receives a unique identification number and eases the communication process. The primary purpose of creating the database using Firebase real-time database is that it entails a code to change the database on the client-side, while on the contrary, SQL expects a server-side code to be developed with server programming languages. Further, the advantage of using FCM can be conceptualized from the fact that the performance of the sent notification can be accessed from the dashboard directly, irrespective of regular monitoring by the donors. It also additionally provides the facility to statistically visualize the execution of the notification.

Nabil et al. [6] outlines a mobile application where users can complete the process of blood donation with minimum effort. The data from patients are obtained using specific sensors with their hardware interfaces connected to the Arduino. Although Arduino provides flexibility in terms of cost, its performance in terms of accuracy



Fig. 2 Application layout for the recipient

Fig. 3 Application layout for the donor



may affect the quality of data collected, thus hampering the overall effectiveness of the application. Since the use of such devices has been avoided in our application, a seamless inter-connectivity is guaranteed between the donor and the recipient. The fact that our proposed application eliminates the use of Raspberry Pi not only makes our application cost-effective, but also optimizes the utility of available resources by reducing the complexity.

5 Conclusion

The paper contributes to acknowledge people with blood requirements by creating an user-friendly interface that will connect all donors and recipients in a unique and selective framework. The employment of Google Maps in the application will recognize the donors who are available in the vicinity of the recipient and hence establish effective communication within a brief period of time. This redesigns the existing customary strategy to a helpful and client steady system which is financially effective and can save thousand lives. Also, enriching the accessibility of the proposed mobile application without the requirement of Internet to suffice the need of blood in provincial areas is seen as a future scope of this exploration.

References

1. B. Muralidaran, A. Raut, Y. Salve, S. Dange, L. Kolhe, Smart blood bank as a service on cloud. *IOSR J. Comput. Eng* **18**(2), 121–124 (2016)
2. M. Jiang, B. Xing, Z. Sun, P. Fu, H. Chen, M. Chen, P. Deng, G. Wang, Y. Xu, and Y. Wang, A dynamic blood information management system based on rfid, in *2005 IEEE Engineering in Medicine and Biology 27th Annual Conference* (IEEE, 2006), pp. 546–549

3. M. S. Rahman, K. A. Akter, S. Hossain, A. Basak, and S. I. Ahmed, Smart blood query: a novel mobile phone based privacy-aware blood donor recruitment and management system for developing regions, in *2011 IEEE Workshops of International Conference on Advanced Information Networking and Applications* (IEEE, 2011), pp. 544–548
4. M. Fahim, H. I. Cebe, J. Rasheed, and F. Kiani, mhealth: Blood donation application using android smartphone, in *2016 Sixth International Conference on Digital Information and Communication Technology and its Applications (DICTAP)* (IEEE, 2016), pp. 35–38
5. A. Moharkar, A. Somani, Automated blood bank using embedded system. *Int. J. Innov. Res. Sci. Eng. Technol* **7**(1) (2018)
6. M. Nabil, R. Ihab, H. El Masry, S. Said, and S. Youssef, A web-based blood donation and medical monitoring system integrating cloud services and mobile application. *J. Phys. Conf. Ser.* **1447**, 012001, IOP Publishing (2020)
7. J. Vaque, Computer system for blood donors management,” in *Medical Informatics Europe*, vol. 78 (Springer, 1978), pp. 775–781
8. I. Fotopoulos, R. Palaiologou, I. Kouris, and D. Koutsouris, Cloud-based information system for blood donation, in *XIV Mediterranean Conference on Medical and Biological Engineering and Computing 2016* (Springer, 2016), pp. 802–807
9. S.-J. Kim, S. K. Yoo, H.-O. Kim, H.-S. Bae, J.-J. Park, K.-J. Seo, and B.-C. Chang, Smart blood bag management system in a hospital environment, in *IFIP International Conference on Personal Wireless Communications* (Springer, 2006), pp. 506–517

Reinforcement Learning for Improving Coherence of Multi-turn Responses in Deep Learning-Based Chatbots



D. G. Suhaas Kiran, Swapneel, Safal Deepak Pansare, and B. N. Krupa

Abstract Chatbots are still far behind in their ability to hold meaningful conversations. The objective of the work is to implement and improve the multi-turn responses of deep learning-based chatbots. Multi-turn response is the ability of a chatbot to give coherent and sensible responses in successive turns. Firstly, sequence to sequence (Seq2Seq) model was built, and its responses were analyzed by varying training parameters. Secondly, the reinforcement learning (RL) method using the Seq2Seq model was implemented, and it is demonstrated that this improves coherence in multi-turn conversations. The RL model performed better than the Seq2Seq model in terms of BiLingual Evaluation Understudy (BLEU) score with a score of 0.3334 compared to 0.2336 of the Seq2Seq model. Average conversation length was found to increase with RL with 3.75 turns compared to 3.05 turns with Seq2Seq.

Keywords Chatbot · Seq2seq · Reinforcement learning · Multi-turn responses · Dialogue systems · BLEU score

1 Introduction

Task-oriented chatbots have a variety of applications, especially in the customer-service sector. Chatbots can be very effective in handling customer queries as they are accessible round-the-clock, cost-effective and can handle large number of conversations simultaneously. However, most customers still find the human interaction more convenient. It cannot give appropriate responses to arbitrary questions and also lacks coherence and context [1].

As per the survey, Paikari [2] examines the current state of chatbots and identifies directions of future work. Young [1] discusses the current trends and developments in dialogue systems. Jwala [3] gives an insight into the different methods to develop the chatbot and the classification of chatbots.

D. G. Suhaas Kiran · Swapneel (✉) · S. D. Pansare · B. N. Krupa

Department of Electronics and Communication Engineering, PES University, Bangalore, India
e-mail: swapneel500@gmail.com

Sequence to sequence model (Seq2Seq) was first proposed by Google in Sutskever [4]. The authors propose a comprehensive method for learning sequences, making very few assumptions about the structure of the sequence. This paper uses Seq2Seq method for English to French translation. Kyunghyun [5] proposes a novel neural network model called RNN encoder–decoder that consists of two recurrent neural networks (RNN), encoder and decoder. Kingma [6] proposes Adam, an order one gradient-based optimization method of objective functions which are stochastic in nature. Adam optimizer was used to update the weights in the network. Pascanu [7] describes the challenges one faces during training RNNs namely exploding and vanishing gradient problem. Pham [8] applies the concept of dropout to RNNs with LSTMs. The paper shows that dropout when applied to the problem of unconstrained handwriting recognition increases the performance greatly. Richard [9] studied open domain generative-based chatbots, Seq2Seq model and encoder–decoder LSTM architecture and has done literature survey of several publications related to chatbots. In Papineni [10], authors propose a method of automatic machine translation evaluation, BiLingual Evaluation Understudy (BLEU).

In this work, firstly, a Seq2Seq model is built that uses encoder–decoder architecture [4]. To overcome the problem of lack of coherence in multi-turn conversations generated by Seq2Seq model, an improvement over the Seq2seq model is developed by using RL. For the evaluation of the single-turn responses of the dialogue systems, BLEU is used [10]. However, for the multi-turn conversations, other evaluation metrics were defined and used in addition to BLEU, for measuring coherence, type-to-token ratio and average conversation length were used which contribute to the novelty of this paper. By using the evaluation metrics specifically proposed for multi-turn conversations, it is shown that reinforcement learning method improves coherence in multi-turn conversations.

2 Methodology

In this section, a detailed explanation on the proposed methodology is given. This includes the datasets used, the Seq2Seq and the RL models employed and validation. The parameters of the models are chosen to optimize the performance of the chatbots [6, 8, 11].

2.1 Datasets

In this paper, two datasets are used, Cornell movie dialog corpus and Twitter dataset. Cornell corpus has large and metadata rich collection of fictional conversations extracted from raw movie scripts. There are a total of 2,20,579 conversations between 10,292 pairs of movie characters. In total, there are 3,04,713 utterances. Twitter

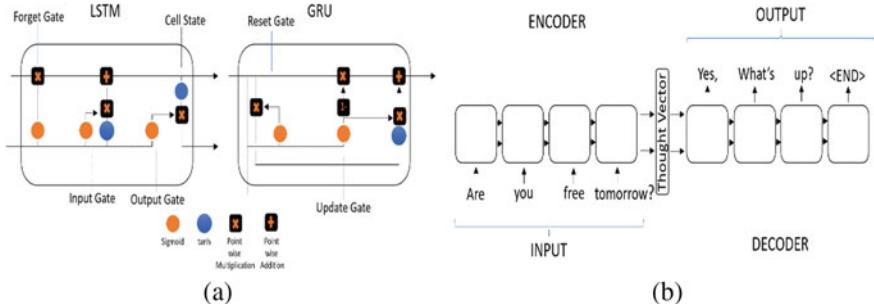


Fig. 1 Architecture of **a** LSTM and GRU, **b** Seq2Seq—encoder and decoder

dataset contains 7,25,000 raw tweets randomly mined from Twitter. The tweets contain special characters, symbols, usernames and emojis.

2.2 Sequence to Sequence Model (*Seq2Seq*)

Given the complexity of the problem of natural language processing, patterns in the data and number of layers in the network, deep learning is preferred for chatbots. Seq2Seq takes input as a sequence of words and generates an output sequence of words by using recurrent neural networks (RNN). To overcome the problem of vanishing gradient, long short-term memory (LSTM) is used in our model, and they are computationally more intensive, accurate and work better with large datasets (Fig. 1a).

Training and Testing. The size of the dataset was 4000 sentences. The vocabulary threshold was set to zero. The number of epochs used in training was 25, 50, 75 and 100. The batch size was kept as 10.

During testing, questions from the datasets and random questions are asked, and the answers generated by the chatbot were saved which were evaluated later.

2.3 Reinforcement Learning (*RL*)

The RL model is based on the deep reinforcement learning method for open-domain dialogue [11]. The various parts of our reinforcement learning method are defined as follows. An action a is next sentence to generate. A state s is defined by previous dialogue turn p . Policy is determined by LSTM encoder–decoder model. Rewards are defined based on two parameters.

Ease of Answering. This is a measure of generating a response by using the negative log likelihood of probability of responding to utterance with a dull response. A set of generic responses was created that were to be avoided in our conversations.

$$r1 = -\frac{1}{N_D} \sum_{d \in D} \frac{1}{N_d} \log p_{\text{seq2seq}}(d|a) \quad (1)$$

where D is the dull response set. N_D denotes the cardinality of D , and N_d denotes the number of tokens in the dull response d .

Semantic coherence. This is used to maintain coherence between conversations and prevent them from becoming meaningless. It considers the forward probability of generating a response given the previous dialogue, but also the backward probability of generating the previous dialogue given the response.

$$r2 = \frac{1}{N_a} \log p_{\text{seq2seq}}(a|q_i) + \frac{1}{N_{q_i}} \log p_{\text{seq2seq}}^{\text{backward}}(q_i|a) \quad (2)$$

where $p_{\text{seq2seq}}(a|q_i)$ denotes the probability of generating an answer a given the previous response q_i , and $p_{\text{seq2seq}}^{\text{backward}}(q_i|a)$ is the backward probability of generating previous response q_i given the answer a .

Training and Testing. Cornell movie dialog corpus was used for training because it has coherent and connected dialogues. The size of the dataset used is large (70,000 sentences). To decrease the number of training parameters, a vocabulary threshold value was defined. All the words in the corpus whose frequency in the corpus was less than the threshold value were removed. The dialogues are transformed into vectors and fed into the LSTM encoder-decoder model. For the backward probability function, a reverse dataset was created, in which the previous sentence was treated as answer, instead of the next sentence. The reinforcement learning model was initialized with the general response generation policy of predicting the response given the previous sentence using the Seq2Seq model. Thus, training was done with Seq2Seq model for 30 epochs and then 25 epochs using policy gradient for optimizing rewards. The batch size used was 100, and learning rate was 0.0001.

For testing, two ways of generating responses were developed. The first method was Q/A type, where generic random questions were asked and the model generated answer based on the question asked. In the second method, the model was provided with the first sentence as a conversation starter, and the model simulated the conversation on its own by taking the previously generated response as input.

2.4 Validation

BLEU [10] is used to evaluate the single turn responses in Seq2Seq model and RL model. However, BLEU does not give any idea about the coherence, which is the most important factor to be taken into consideration for evaluation of long conversations. Hence, different metrics are used for multi-turn responses.

Type-To-Token Ratio (TTR). TTR of a sentence is defined as the number of distinct words divided by the total number of tokens in the given sentences. The closer the TTR value is to 1.0, the greater the diversity of words in that sentence.

Average Conversation Length. The length of a conversation is determined by the number of meaningful responses in a conversation. The conversation is considered meaningless when it starts generating dull and non-connected responses. To find the average length of a conversation, the averages of the lengths of coherent responses from the conversation were determined.

3 Results and Discussions

For comparing the Seq2Seq model for the two datasets for different epochs, 21 questions were selected which comprised of sentences from the dataset and few random generic questions chosen by us, such as “Hello” and “How are you”. The average response lengths were evaluated for 25, 50, 75 and 100 epochs, respectively, for both the datasets. Reference answers were set for all the questions, and the average of BLEU scores for individual questions was calculated (Table 1).

It was found that the overall BLEU score of responses and average length of responses increased with the number of epochs (Table 1). However, the responses generated by model trained on Cornell dataset were more meaningful, as the input sentences were in form of dialogues. For multi-turn responses, the TTR was slightly greater in Seq2Seq model. But the average length of meaningful conversations increased in RL model. From Table 2, it can be concluded that RL model shows more coherent responses when compared to the multi-turn responses generated by the Seq2Seq model. In Fig. 2, the boxes represent the coherent responses. Jiwei [11] has obtained 2.68 turns compared to 3.05 turns for our work for Seq2Seq model. For RL, they obtained 4.48 turns compared to 3.75 from our work. It can be validated that turns increase with the help of RL model. While going from Seq2Seq to RL, Jiwei [11] reports an increase of 174% for TTR for unigram. In our case, TTR remains the same.

Table 1 Epoch versus avg. response lengths and BLEU scores for the datasets used

Epoch	Cornell corpus		Twitter dataset	
	BLEU score	Avg. response length	BLEU score	Avg. response length
25	0.0432	1.0	0.0502	3.0
50	0.2008	2.571	0.3672	2.95
75	0.4949	3.0	0.4925	3.19
100	0.7368	3.571	0.5765	3.24

Table 2 Models used versus BLEU scores, TTR and avg. conversation length

Model	BLEU score	TTR	Avg. conversation length
Seq2Seq	0.2336	0.4913	3.0536
RL	0.3334	0.4851	3.7455

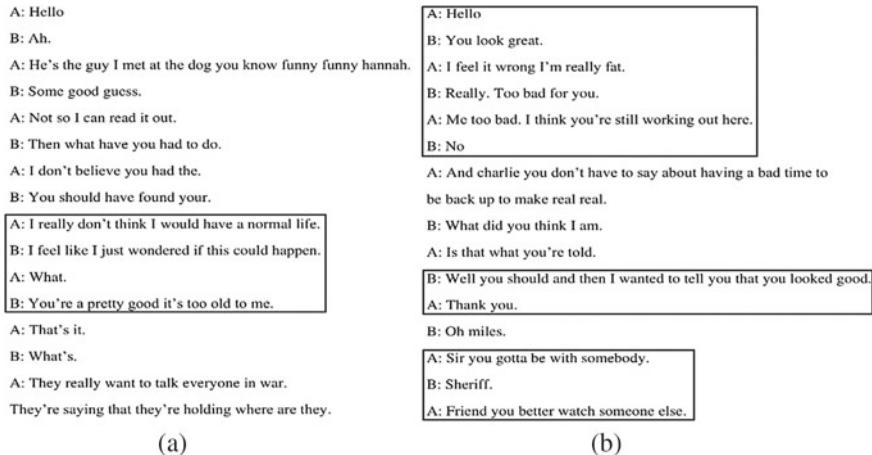


Fig. 2 Example of multi-turn responses using **a** Seq2Seq model, **b** RL model

4 Conclusion

The open-domain-based chatbot was implemented using Seq2Seq and RL model. The RL model improves coherence as the average conversation length which was found to be 3.75 was higher in case of RL compared to 3.05 for Seq2Seq. The RL model performed better than the Seq2Seq model in terms of BLEU score with a score of 0.3334 compared to 0.2336 of the Seq2Seq model. The rewards were defined with an aim to address the problem of repetitive responses and lack of coherence. RL method simulates the conversation, tries different sequences of words as responses and also uses trial and error to maximize the rewards which have been defined.

References

1. T. Young, D. Hazarika, S. Poria, E. Cambria, Recent trends in deep learning based natural language processing. arXiv preprint [arXiv:1708.02709](https://doi.org/10.1109/mci.2018.2840738) (2017). <https://doi.org/10.1109/mci.2018.2840738>
2. E. Paikari, A. Van der Hoek, A framework for understanding chatbots and their future, in *2018 IEEE/ACM 11th International Workshop on Cooperative and Human Aspects of Software Engineering (CHASE)*, Gothenburg, 2018, p. 13–16. <https://doi.org/10.1145/3195836.3195859>
3. K. Jwala, G.N.V.G. Sirisha, G.V. Padma Raju, Developing chatbot using machine learning. Int. J. Recent Technol. Eng. (IJRTE) **8**(1S3) (2019). ISSN: 2277-3878
4. I. Sutskever, O. Vinyals, Q.V. Le, Sequence to sequence learning with neural networks. *Adv. Neural. Inf. Process. Syst.* **2014**, 3104–3112 (2014). <https://doi.org/10.5555/2969033.2969173>
5. K. Cho, B. Van Merriënboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, Y. Bengio, Learning phrase representations using RNN encoder-decoder for statistical machine translation. arXiv preprint [arXiv:1406.1078](https://doi.org/10.3115/v1/d14-1179) (2014). <https://doi.org/10.3115/v1/d14-1179>
6. D.P. Kingma, J. Ba Adam: a method for stochastic optimization. arXiv preprint [arXiv:1412.6980](https://doi.org/10.48550/arXiv.1412.6980) (2014)

7. R. Pascanu, T. Mikolov, Y. Bengio, On the difficulty of training recurrent neural networks. arXiv preprint [arXiv:1211.5063](https://arxiv.org/abs/1211.5063) (2012)
8. V. Pham, T. Bluche, C. Kermorvant, J. Louradour, Dropout improves recurrent neural networks for handwriting recognition, in *14th International Conference on Frontiers in Handwriting Recognition*, Heraklion, p. 285–290 (2014). <https://doi.org/10.1109/icfhr.2014.55>
9. R. Csaky, Deep learning based chatbot models. arXiv preprint [arXiv:1908.08835](https://arxiv.org/abs/1908.08835) (2019). <https://doi.org/10.13140/rg.2.2.21857.40801>
10. K. Papineni, S. Roukos, T. Ward, W.J. Zhu, BLEU: a method for automatic evaluation of machine translation, in *Proceedings of the 40th Annual Meeting on Association for Computational Linguistics. Association for Computational Linguistics*, pp. 311–313 (2002). <https://doi.org/10.3115/1073083.1073135>
11. J. Li, W. Monroe, A. Ritter, M. Galley, J. Gao, D. Jurafsky, Deep reinforcement learning for dialogue generation. arXiv preprint [arXiv:1606.01541](https://arxiv.org/abs/1606.01541) (2016). <https://doi.org/10.18653/v1/d16-1127>

Intensification of Bandwidth, Return Loss and Gain of Ultra-Wideband Microstrip Antenna with Single-Band-Notch (U-Slot) Characteristics



Manisha Mohanty and Bikram Choudhury

Abstract A trailblazing improvised band-notch characteristics by altering the frequencies in a microstrip-fed UWB antenna is introduced. Two slots have been inserted on the ground plane, which gives much extensive impedance bandwidth. To endanger the frequency band stop execution, a refashioned U-shaped conductor-backed plane is used with variable dimensions. The proposed antenna has a tiny size of 22 mm × 22 mm and operates for well-founded VSWR < 2 in the frequency range of 3.1–14 GHz thus demonstrating the band elimination output in the frequency range of 5.1–5.9 GHz. The antenna manifests a data transmission capability of 10 dB return loss over the entire frequency band considered for UWB. FR4 substrate has been taken for the antenna design and fortified with 50 Ω impedance microstrip line. From the validated designs, it is realized that the frequency band is dependent on the partial ground width and U-shaped slot. The antenna design and simulation based analysis is conducted using simulation tools from CST Microwave Studio. The main motto of this paper is to obtain better results in terms of efficiency, bandwidth, return loss and gain by modifying a conventional microstrip antenna with help of band-notch techniques (U-Slot) in order to eliminate the influence of opposition from limited bands like WLAN and WiMAX (5.15–5.825 GHz) over the total 7.5 GHz UWB, i.e. from 3.1 to 10.6 GHz. Indeed, we have received better performance in the U-Slot UWB MSA.

Keywords Frequency band-notch function · Ultra-wideband (UWB) microstrip patch antenna (MSA) · Gain · Bandwidth · VSWR and return loss

M. Mohanty
College of Engineering and Technology, Bhubaneswar, Odisha, India
e-mail: manisharifa1234@gmail.com

B. Choudhury (✉)
CYIENT LTD, Hyderabad, India
e-mail: bikram.suraj1994@gmail.com

1 Introduction

Ultra-wideband compatible wireless system has accumulated remarkable deliberation in the wireless communication due to its inflated exactness, distinctive focal points, including low preternatural power thickness and particularly swift information rate. In 2002, the frequency band ranging from 3.1 to 10.6 GHz had been designated as unlicensed band by Federal Communication Commission (FCC), divided and shared the frequency band for UWB requisition. In the above assigned UWB frequency band, there exists WLAN (which occupies the frequency bands ranging from 5.150 to 5.825 GHz) which obstructs the UWB pursuit [1–4]. That's why various band-notch characteristics are formulated for UWB wireless systems not only to mitigate the potential blockage but also to minimize the prerequisite of supplementary band-stop channels in wireless system for UWB [5–9]. To overcome all these issues, the proposed notched antenna with help of U-Slot is designed and evaluated by the help of CST Microwave Studio, which is basically an electromagnetic software. The suggested antenna is constructed and experimentally studied by using band-notch characteristics, which yields promising suitable outcomes. The design and necessary mathematical expressions used to obtain the fundamental parameters of an MSA antenna are presented in Sect. 2. Section 3 discusses results and comparisons of the system models developed in CST with the help of U-Slot to attain higher gain, bandwidth and return loss and finally followed by concluding remarks in Sect. 4.

2 Antenna Design of UWB MSA with U-Slot Band Notch

For the initial design of microstrip patch antenna without band notch, FR-4 substrate has been taken to be 1.6 mm thick and 4.4 dielectric constant. The feedline width is perceived to be 1.8 mm long and $50\ \Omega$ input impedance is fed to it. A rectangular patch of dimension $10\text{ mm} \times 13.5\text{ mm}$ is created on the top of the substrate. At the bottom portion is the partial ground plane. The feedline fused into the partial ground plane to provide the antenna with an impedance (Fig. 1).

In Fig. 2, the plot frequency vs S11 is obtained by etching two rectangular positions in the halfway ground plane with a length of 7.3 mm (L_s) and a width of 2 mm (W_s) to follow UWB (ultra-wideband) requirements. The S11 parameter stands for the reflected power from the antenna, and therefore, it is also considered as the return loss or reflection coefficient. However within this frequency band, WLAN operates on 5.15–5.825 GHz, which may create interference with UWB operation. To overcome this issue, we have implemented U-Slot band-notch technique, which is illustrated on Sect. 2.1 (Table 1).

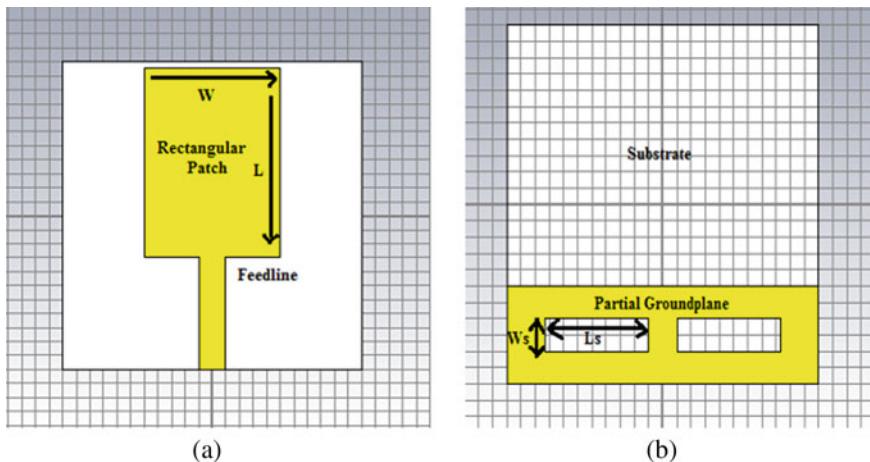


Fig. 1 **a** Front view and **b** rear view of UWB MSA

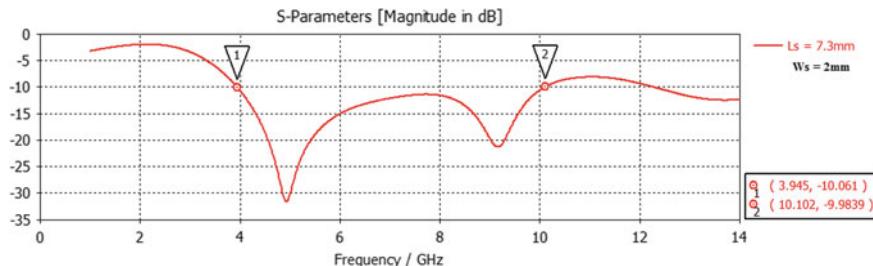


Fig. 2 Frequency versus S₁₁ plot for the optimum outcome of L_s and W_s without notch

Table 1 Parameter table for conventional UWB antenna and U-Slot UWB antenna

Type of antenna	Conventional UWB antenna		UWB antenna with U-Slot	
Parameters	Length in mm	Width in mm	Length in mm	Width in mm
Ground plane (W _s , L _s)	22	6	22	6
Patch	13.5	10	13.5	10
Substrate	22	22	22	22
Feedline	8	1.8	8	1.8
U-Slot	—	—	10	1.8

The substrate thickness = 1.6 mm

The substrate height = 0.02 mm

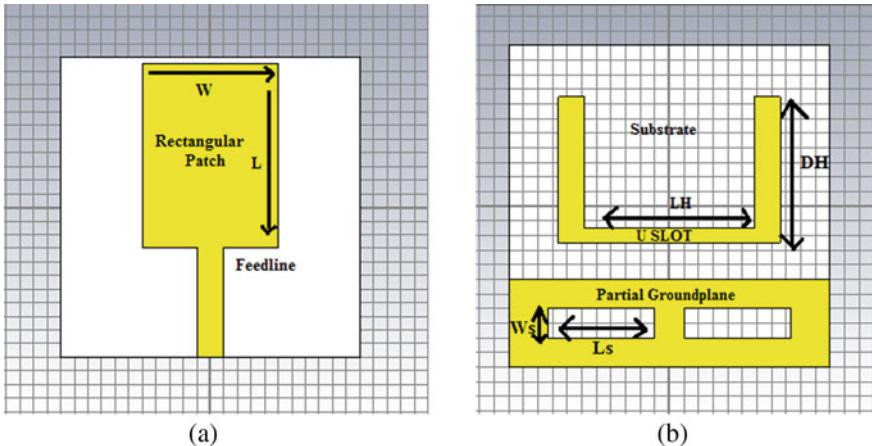


Fig. 3 **a** Front view and **b** rear view of UWB MSA with U-Slot

2.1 Parameter Table

From the design of UWB MSA with U-Slot, the patch length is introduced as ‘ L ’ and the patch width is introduced as ‘ W ’ in Fig. 3a. The yellow colour indicates feedline and patch, which is basically a material with high conductivity; here we have chosen copper. The white colour from the design signifies the substrate. The back view of the UWB antenna in Fig. 3b shows the partially covered ground plane by substrate. To achieve the ultra-wideband result, two cuts have done on ground plane. ‘ L_s ’ designates the length of the cut and ‘ W_s ’ specifies the width of the cut. The design has a U-Slot, the width and length of the slot are determined by DH and LH. In this antenna design, two cut has been made which is also known as defected ground structure. Generally, it is created to improve the bandwidth of related antenna. The basic concept is that a slot in the ground plane behaves as a load which when added to the antenna can be used to bring the input impedance point closer to the characteristic impedance (e.g. 50Ω). The structure or shape of defected ground structure (DGS) is rather simpler, and there is no need of a large area. The core concept of this arrangement is that the shield current distribution is disturbed by DGS in the ground plane, which affects the input impedance and current flow of the antenna, thus increasing the bandwidth.

3 Results and Analysis

See Fig. 4.

An U-shaped slot is formed on the ground plane after attaining the UWB attributes by combining two parameters LH (length of the U-Slot) and DH (U-Slot width). The

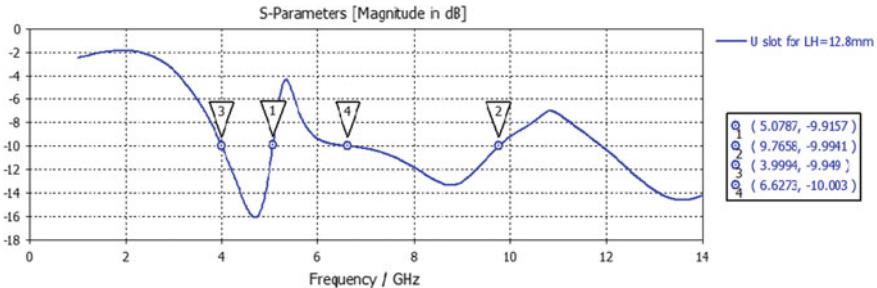


Fig. 4 Frequency versus S11 graph for the optimum outcome with U-Slot

basic fundamental design of MSA covers the WLAN frequency range, i.e. 5.150–5.825 GHz as it totally covers the entire complete range from 3.945 to 10.102 GHz presented in Fig. 2, which is not acceptable as it creates interference and overlapping but when we implemented the U-Slot-notch technique in Fig. 3, we got two frequency ranges, i.e. 3.99–5.07 GHz and 6.627–9.7658 GHz, which discards the interference caused by WLAN.

In Fig. 5, several variations have been done with LH in the U-Slot which is depicted to give the accurate result at LH = 12.8 mm. As the length of U-Slot increases, simultaneously antenna radiation follows the same. The frequency band rejection occurs in the middle from 4.9 to 6 GHz as the band-notch properties reject the band from 5.1 to 5.9 GHz. The refusal of the frequency band varies with the duration of increase for LH. By taking a ratio of $\lambda/4$, the length of LH is calculated.

Figure 6 represents VSWR curve showing the optimum value at LH (12.8 mm), which is the optimal outcome of the antenna wrapping the entire ultra-wideband with the band-notch characteristics implemented. VSWR is a calculation of how effectively radio frequency power is transmitted through a transmission line from a power source to a load.

The antennae's electromagnetic waves directionally depend on the strength preferred by the radiation pattern. The antenna gain is correlated with the antenna's directivity and the electrical efficiency that is expressed by the far-field radiation

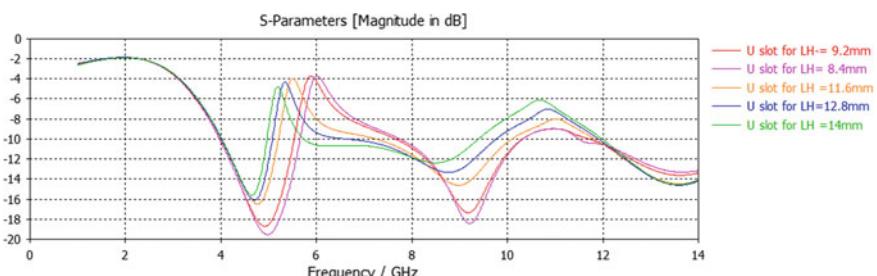


Fig. 5 Frequency versus S11 plot by changing LH in U-Slot

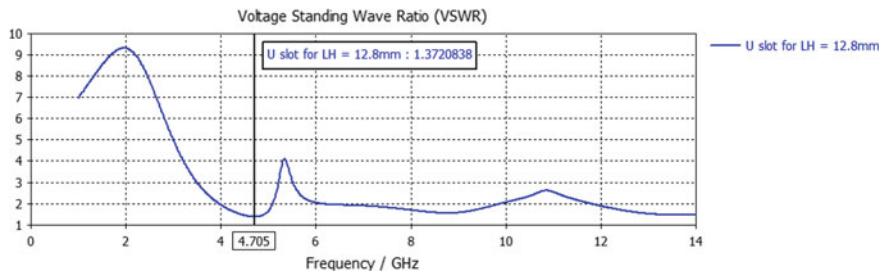


Fig. 6 VSWR curve of suggested antenna with U-Slot (LH = 12.8 mm)

pattern in the 3D diagram. The radio waves are converted from the input power through the antenna headed in a specified direction defined as the gain. 2.03 dB is the antenna gain here at frequency 4.9 GHz, and the gain at 9.1 GHz is 3.7 dB without notch, which is represented in Fig. 7.

When we applied the U-Slot-notch technique, we got better gain at frequency 4.7 and 9 GHz, which is demonstrated on Fig. 8.

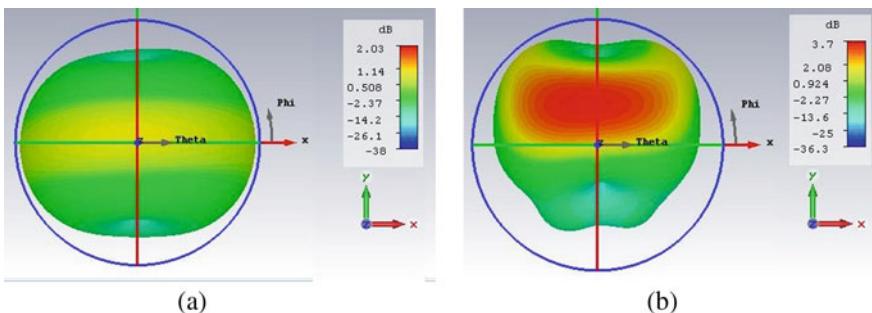


Fig. 7 Radiation pattern at **a** 4.9 GHz and **b** 9.1 GHz of UWB MSA without U-Slot

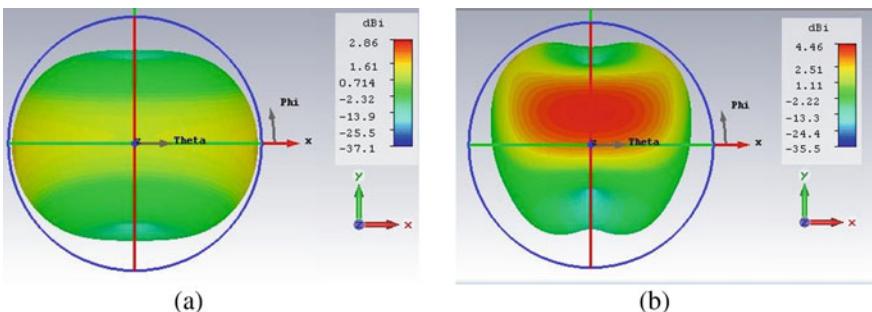


Fig. 8 Radiation pattern at **a** 4.7 GHz and **b** 9 GHz of UWB MSA with U-Slot

Table 2 Final compression analysis between UWB antenna with and without notch

Parameters	UWB without band notch	UWB with band notch
Bandwidth	6.157 GHz	P1 = 1.08, P2 = 3.14 GHz
Resonant frequency	P1 = 4.9 GHz, P2 = 9.1 GHz	P1 = 4.7 GHz, P2 = 9 GHz
Gain	2.03 dB, 3.7 dB	2.86 dB, 4.46 dB
Return loss	P1 = -32 dB, P2 = -21 dB	P1 = -16 dB, P2 = -13 dB

Table 2 renders the comparison and improvement we have achieved with the help of notch characteristics. Bandwidth specifications for WLAN and other short-range communication can be removed for a microstrip UWB antenna by cutting a fractional portion in the radiator of the antennas. That is achieved by adjusting location and opening distance, so we can supervise bandwidth of notched band and centre frequency. The polarization of antenna is generated by the radiated fields which is evaluated by far field of antenna.

4 Conclusion

After the design and description of a UWB antenna, we had used the band-notch technique to obtain microstrip patch antenna's band-notch characteristics. Through building a U-Slot, it was calculated that the frequency falls between 5.1 and 5.9 GHz. By making narrow slots of different shapes into the radiating patch of the antenna, we can achieve the band-notch characteristic easily which will affect the current flow of the patch, and it will help to get the band notch in its desire frequency band. The simulation findings for the band-notch antenna show that the antenna proposed fulfilled the characteristics of a single band notch for different frequency range and demonstrated the return loss and radiation pattern in the fascinated UWB antenna.

References

1. R. Zaker, C. Ghobadi, J. Nourinia, Novel Modified UWB planar monopole antenna with variable frequency band-notch function. *IEEE Antennas Wirel. Propag. Lett.* **7** (2008)
2. R.C. Jain, M.M. Kadam, Band notching methods used in UWB antennas. *Int. Res. J. Eng. Technol. (IRJET)* **2** (2015)
3. Z. Ahmed, G. Perwasha, S. Shahid, H. Zahra, I. Saleem, S. Muzahir Abbas, Ultra wideband antenna with WLAN band-notch, in *Characteristic 2013 IEEE*
4. Kang, H. Zhang, Z. Li et al., A band-notched UWB printed half elliptical ring monopole antenna. *Prog. Electromagn. Res.* **35**, 23–33 (2013)

5. Y.-Z. Cai, H.-C. Yang, Wide monopole antenna with three band-notched characteristics. *IEEE Antennas Wirel. Propag. Lett.* **13** (2014)
6. W.-S. Lee, D.-Z. Kim, K.-J. Kim, J.-W. Yu, Wideband planar monopole antennas with dual band-notched characteristics. *IEEE Trans. Microw. Theory Tech.* **54** (2006)
7. C.A. Balanis, *Antenna Theory: Analysis and Design* (WILEY, 2009)
8. S.K.R.M. Satyabrata Maiti, Design of a compact ultra wide band microstrip patch antenna. *IEEE* **5** (2014)
9. A.K. Nayak, D. Gountia, B.K. Turuk, S.B. Panda, Compact UWB antenna for S, C, and X bands applications, in *2018 IEEE International Symposium on Smart Electronic Systems (iSES) (Formerly iNiS)*, Hyderabad, India, 2018, p. 183–186

Design of Multi-utility Bottleneck Horn Antenna for Ku-, K-, and Ka-Band Applications



Yagnik Mehta, Nisarg Patel, Ronak Vashi, and Jagdish Rathod

Abstract In this paper, we propose the design to accommodate for the Ku-band (12–18 GHz), K-band (18–27 GHz), and Ka-band (26.5–40 GHz) frequency used cases as high-fidelity broadcasting for military applications, short-range terrestrial communication, and advanced weather forecasting applications by implementing the design structure with horn antenna having bottleneck profile. To have a ubiquitous application embedded with a single antenna, this structure was developed using advanced design analysis to determine the paragon shape responsible for high radiation efficiency. This constraint is evaluated in terms of return loss, radiation pattern, electromagnetic field patterns, and current distribution. The simulated impedance bandwidth of the bottleneck horn antenna is 100% showing that the antenna is operational at the entire intended frequency band range. The accounted gain is of 23.18 dB and directivity of 23.23 dB having 99.78% efficiency and low operating power of 0.5 W making the antenna high-line ideal for the desired satellite communication.

Keywords Ku band · K band Ka-band · Military communications · Advanced weather forecasting · Terrestrial application bottleneck horn antenna

1 Introduction

For a few decades, space agencies and military experts around have been using Ku-, K-, and Ka-band for multidomain applications. Ku-band constraining the frequency range of 12–18 GHz has been used for radar applications, military intelligence, high-fidelity satellite-based television, and global data broadcasting, and most notably for microwave towers and radio astronomy [1–3]. Due to its high flexibility, it was open for smaller dish size DBS (direct broadcasting satellite) and power-efficient terrestrial applications. But due to the observable attenuation of electromagnetic waves causing rain fade phenomenon, previous versions of this antenna were facing

Y. Mehta (✉) · N. Patel · R. Vashi · J. Rathod
Birla Vishvakarma Mahavidyalaya Engineering College, V.V. Nagar, Gujarat 388120, India
e-mail: yagnikmehta999@gmail.com

limitations in their used cases. Hence to overcome limitations, researchers had to develop separate antenna working in K- and Ka-band spectrums.

Ironically issue faced in the Ku-band antenna is found to apply in advance weather forecasting performed with Ka-band. Ka-band is used by space agencies for unlinking data to the terrestrial satellite with 30–40 GHz. NASA uses this band specifically for communications related to deep space mission such that in the famous KEPLER mission [4–6]. Currently, it is being utilized by SpaceX Star-link Satellite Systems which aims to supply high Internet covering the entire globe. To have the necessary applications currently being used, we have approached to development of bottleneck horn antenna which provides high gain and good radiation pattern symmetry while retaining the antenna's simplicity in construction which is essential for satellite communications systems [7, 8]. However, in general, the reduced dielectric and insertion losses, lower weight, and cost efficiency increase the applications of this antenna.

Recently, the bottleneck horn has received considerable interest. The performance of the common horn antenna is improved by modifying the overall shape and inserting a layer of epsilon-near-zero material at a certain thickness along the antenna wall [9]. Considering the above, a comparative study of the Ku-band, K-band, and Ka-band bottleneck horn with metallic material based on measurements and simulations is reported in the present work. This work is constructed as follows: Sect. 2 presents the design of the antenna. The results of the presentations are presented in Sect. 3. The last section summarizes the main conclusions [10–12].

2 Proposed Antenna Design

The aim here was to create a multifrequency operating bottleneck horn antenna working on Ku, K, and Ka-band, respectively, with maximum gain and high efficiency that eventually takes a lot less power to radiate at the desired frequency for the further user in satellite applications [5]. We augmented our model of the antenna with few different materials like copper showcasing alternate results like E-Field, H-Field, Fairfield (RCS), current density, power flow, etc. The antenna is designed to keep the design considerations and the mathematical formulas with the necessary values to be set for the model.

The dimensions of the bottleneck horn antenna are augmented with formulas of input radius (a_i), output radius (a_o), length of the bottleneck profile (L_1), and length of the cylindrical waveguide (L_2) as shown in Table 1.

The calculated dimensions for the proposed horn antenna are output radius (a_o) = 34.912 mm, input radius (a_i) = 8.947 mm, length of bottleneck profile (R1) = 200 mm, and length of the cylindrical waveguide (L_2) = 10.00 mm. These dimensions are calculated at an operating frequency of 12–18 GHz. Our antenna follows the path of smooth bottleneck shape with a hollow structure having the lumped port assigned at the lower end of the antenna. Below are the perspective, top and side views of the final modelled copper materialized bottleneck horn antenna: Fig. 1

Table 1 Mathematical analysis for bottleneck antenna

Sr. No.	Parameter	Formula
1	Input radius a_i	$a_i = \frac{3\lambda_{center}}{2\pi}$
2	Output radius a_o	$a_o \cong 1.723\lambda_c$
3	Length of bottleneck profile L_1	$L_1 = \frac{2a_o}{\sin 20^\circ}$
4	Length of the cylindrical waveguide L_2	$L_2 = 2\lambda_{g,avg}$

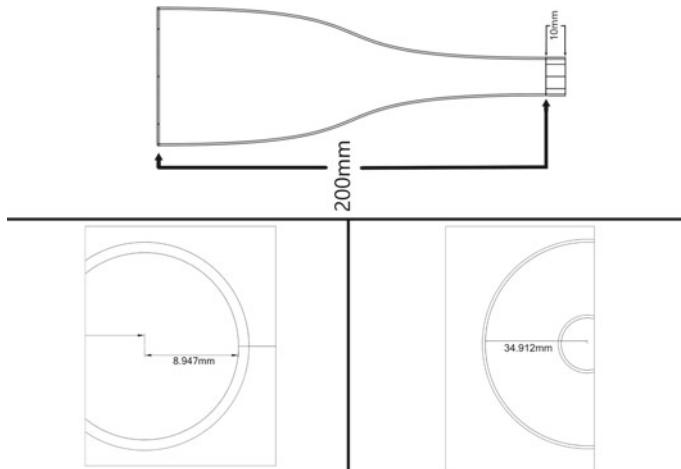


Fig. 1 Overall length of the antenna (top), lower/input radius of the antenna (bottom left), and the upper/output radius of the antenna (bottom right)

3 Results and Analysis

3.1 VSWR (*Voltage Standing Wave Ratio*) Results of Modelled Antenna

By performing the simulation and doing analysis of the designed model, we observed the VSWR for Ku-band (12–18 GHz) is ≤ 1.15 dB, for K-band (18–267 GHz) is ≤ 1.1 dB, and for Ka-band (26.5–40 GHz) is ≤ 1.05 dB which confirms that the designed model has accurately matched impedance along with giving evidence of working at three bands, respectively, which is essential for the antenna to rectify the effect of reflection of power from the antenna. The VSWR below 1.33 is meant to be ideal for satellite communications or military applications.

Figure 2 represents the division of VSWR at various frequencies, 12–18 GHz, 18–27 GHz, and 26.5–40 GHz to be precise, for copper, respectively.

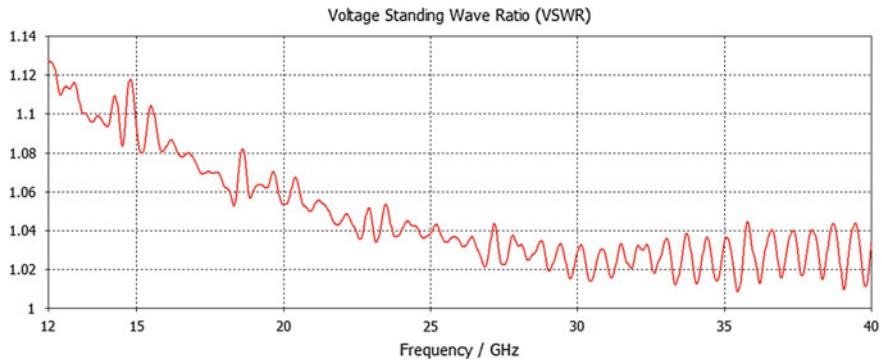


Fig. 2 VSWR of aluminium, copper, and gold material for modelled bottleneck antenna

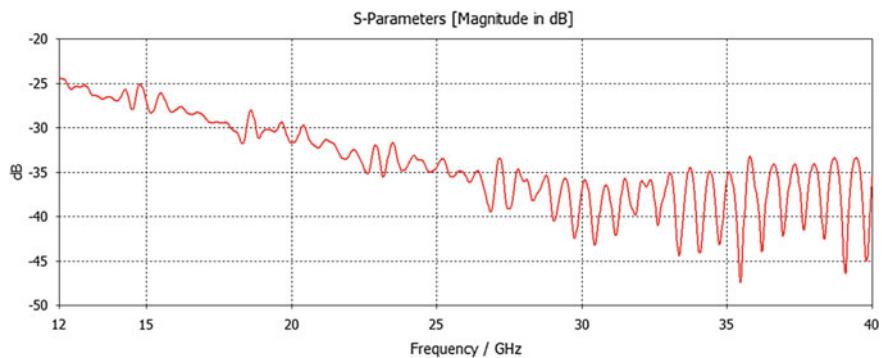


Fig. 3 S_{11} parameters of aluminium, copper, and gold material for modelled bottleneck antenna

3.2 *S_{11} Parametric Studies of Modelled Antenna*

The simulation results of the S_{11} parameters for the design obtained are as shown that includes S -parameter against the frequency Fig. 3 represented below perspicuously portray that for the materials of copper, the S parameter—return loss—is beneficially lesser than -15 dB over the Ku-, K-, and Ka-band frequency.

3.3 *Gain and Directivity of Modelled Antenna*

The measurement results of the modelled antenna for IEEE gain are presented in Fig. 4 along with the simulation results for the directivity of the antenna as shown in Fig. 5. As shown from the results in Fig. 4, the gain of the antenna is measured to be 23.18 dB at the almost whole band of 12–40 GHz and from Fig. 5, and the directivity of the antenna is measured to be 23.23 dB. In conclusion, the simulated gain is close

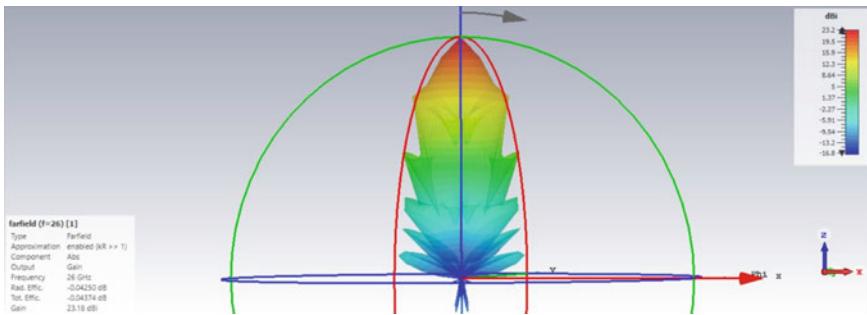


Fig. 4 Gain of copper material for bottleneck antenna

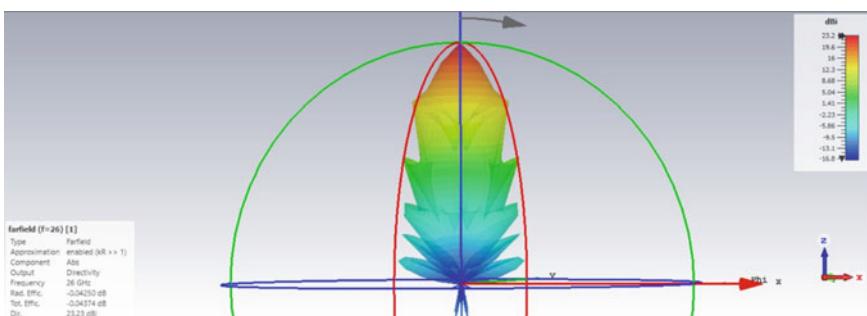


Fig. 5 Directivity of copper material for bottleneck antenna

to the directivity as seen in the figures below. The efficiency of the designed model is over 95%. Such results are prudently due to the reason for uniformity in the thickness and sinusoidally shaped antenna.

4 Conclusion and Future Scopes

In an epitome, this paper demonstrates the design and analysis of a multi-utility bottleneck horn antenna to be utilized in the frequency spectrum of Ku-, K-, and Ka-band. With a conical radiation pattern, for optimum gain, this design also demonstrates fundamental characteristics defined for the antenna to work for satellite communications, given the lightweight profile of the antenna and the low power usage and observed results. From the perspective of the overall design and the results obtained from the horn antenna, we can approach the more complex design of the antenna with higher practical gain and greater power efficiency. Using metamaterials and sophisticated 3D printing technology, a novel way to quickly create low-cost and lightweight bottleneck horn antennas in spectrum is modified. This is especially

useful for direct-broadcast satellites (DBS) or radio astronomy since the received power does not exceed the maximum power of the antenna.

References

1. X. Yan, Y. Fang, Z. Lu, Design of a broadband Ku-band lens horn antenna for SAR applications, in *Proceedings of 2014 3rd Asia-Pacific Conference on Antennas and Propagation*, Harbin, 2014, p. 29–32
2. C.Y. Tan, K.T. Selvan, A performance comparison of a Ku-band conical horn with an inserted cone-sphere with horns with an integrated dielectric lens and metamaterial loading [Antenna Designer's Notebook]. *IEEE Antennas Propag. Mag.* **53**(5), 115–122 (2011)
3. A. Heiman, A. Badescu, Design of a conventional horn antenna for Ku band, in *2020 International Workshop on Antenna Technology (iWAT)*, Bucharest, Romania, 2020, p. 1–4
4. M. Patriotic, F.N. Ayoub, C.G. Christodoulou, S. Jayaweera, A K/Ka-band frequency reconfigurable transmit/receive antenna array, in *2019 13th European Conference on Antennas and Propagation (EuCAP)*, Krakow, Poland, 2019, p. 1–4
5. A. Shishlov, I. Vilenko., V. Denisenko., I. Zimin, V. Klassen, Ku/Ka-band antenna terminals for satellite communications, in *2019 International Conference on Engineering and Telecommunication (EnT)*, Dolgoprudny, Russia, 2019, p. 1–5
6. M. Ferrando-Rocher, J.I. Herranz-Herruzo, A. Valero-Nogueira, B. Bernardo-Clemente, Full-metal K-Ka dual-band shared-aperture array antenna fed by combined ridge-groove gap waveguide. *IEEE Antennas Wirel. Propag. Lett.* **18**(7), 1463–1467 (2019)
7. R. Marino, R. Mashiah, H. Matzner, E. Levine, Investigation of a multi-waveguide fed horn antenna. *J. Wirel. Netw. Commun.* **2**(1), 5–8 (2012)
8. J.D. Kraus, R.J. Marhefka, *Antennas—For All Applications*, 2nd edn, T.A. Milligan, *Modern Antenna Design*, 2nd edn (Wiley), p. 447–459
9. B. Zhang, L. Wu, Y. Zhou, Y. Yang, H. Zhu, Y. Wu, A 3D printed aluminum alloy K-band square stepped doubled ridged horn antenna, in *2018 IEEE International Conference on Computational Electromagnetics (ICCEM)*, Chengdu, 2018, p. 1–2
10. J.M. Rathod, Y.P. Kosta, Development of conical horn feed for reflector antenna. *Int. J. Eng. Technol.* **1**(1), 71 (2009)
11. P. Duangtang, P. Mesawad, R. Wongsan, Performance improvement of conical horn antenna by using wire medium structure and dielectric load 24 (2017)
12. C. Eric, An open standard for Ka-band interoperable satellite antennas, in *MILCOM 2019—2019 IEEE Military Communications Conference (MILCOM)*, Norfolk, VA, USA, 2019, p. 76–81

Cyberbully Detection Using 1D-CNN and LSTM



Sourodip Ghosh , Aunkit Chaki , and Ankit Kudeshia

Abstract This paper proposes deep learning-based solutions for cyberbullying issue that is becoming increasingly common in the modern era of social media and digital connect. The early detection and identification of such events can curb the menace of this unethical practice. Toward this, 1D-CNN, LSTM and bidirectional LSTM (BiLSTM) networks are utilized in this work that are able to detect and classify texts into six different cyberbully classes. The dataset used in our training and testing procedure contains $159k$ input examples comprising a variety of texts belonging to both non-bullying and bullying sentiments. Our results show that the proposed models achieve an overall test accuracy of 0.9633, 0.9412 and 0.9745 using 1D-CNN, LSTM and BiLSTM networks, respectively, thereby making BiLSTM a suitable network for cyberbully detection purpose.

Keywords Cyberbully · CNN · LSTM · BiLSTM · Text classification

1 Introduction

Cyberbullying [10] is a type of harassment by using electronic means which includes toxic comments, indicating obscene languages, hate speeches or threats related to one's identity. According to various surveys, cyberbully has a global dominance in countries like India, Brazil and the United States, which is a major concern in terms of ensuring social and mental well-being of individuals. This topic is of utmost significance and has acquired serious attention from researchers throughout the past two decades. Computer-aided tools in this modern digital era have been developed by researchers to automate the cyberbully detection process in order to make analysis more reliable. This paper proposes an analysis of CNN and LSTM models for cyberbully text detection for six primary bully text classes and presents a comparison

S. Ghosh · A. Chaki · A. Kudeshia
KIIT University, Bhubaneswar, India
e-mail: sourodip.ghosh02@gmail.com

of their performances. This work uses a huge dataset of $159k$ texts with an increased number of classes and offers a superior classification result from past researchers in terms of model diversity.

1.1 Literature Survey

Cyberbully, a phenomenon formally identified in 2003, has been posing a severe threat to the vastly growing social media platforms in modern age of Internet and digital connect. The first work on automated cyberbully detection was done in 2006 [8]. They conducted multiple online surveys of people with different age groups, race, gender and country. They also studied about various situations or response related to cyberbully. Since then, multiple researchers have attempted to address this issue using computer-aided tools. It was categorized as a text classification problem. Some used machine learning (ML) algorithms to preprocess data, use feature engineering to detect target words and classify bully texts from normal texts, and the results were evaluated using F1-Score and accuracy parameters. However, the precision of these models is highly abstract, as the use of swear words and the high reliability of these models on handcrafted features act as a disadvantage to the overall reliability of these models. Moreover, it gets more difficult for models to classify between sarcasm and texts which are truly considered as a cyberbully.

Recently, deep learning approaches have also been considered for cyberbully detection in [1]. Karthik et al. [2] in 2011 used multiple machine learning algorithms, namely Naive Bayes, SVM, J48 and JRip algorithms with a data size of $4.5k$ text examples to classify between cyberbully topics like sexuality, intelligence and culture and race, with an accuracy of 80.2%, 70.39% and 68.3%. Djuric et al. in 2015 [3] proposed a continuous BOW (CBOW) model to classify between hate speech and clean texts, from Yahoo data of size around $951k$ texts, with an AUC score of 0.80. The data features they used were distributed representations of comments (paragraph2vec). Van Hee et al. used Ask.fm data and used SVM on $91k$ pre-processed texts to classify between “Threats”, “Defense”, “Sexual talks”, “Insult”, “Defamation”, “Encouragements” and “Swear”, with an F1-Score of 55.39. Though this approach offered low model evaluation scores, it was one of the first work on classification into a high number of inter-related classes. They used word features like unigram and bigram bag-of-words, character trigram bag-of-words and sentiment lexicon features (comment2vec). The Yahoo data was again used by Nobata et al. [7] to detect between abusive and non-abusive classes, where they used a Vowpal wabbit regression model on two types of datasets, namely finance ($759k$) and news ($1390k$). They achieved an AUC score of 0.90. In 2017, Badjatiya et al. [1] used Twitter data of size $16k$ examples, divided between classes “Sexism”, “Racism” and “None”. They used deep learning models such as Fast Text, CNNs and also complex neural networks like LSTMs and GBDT and achieved an overall F1-Score of 0.93. Furthermore, Wulczyn et al. work in [12] used a Wikipedia dataset of size about $100k$ using MLP and LR models offering an AUC score of 0.9659.

2 System Design

The proposed method uses a comparison analysis of 1D-CNN with LSTM architectures as shown in Fig. 1, on text semantics to detect words that symbolize bullying. The models propose further advancement in contrast to previous works involving automatic cyberbully detection using deep learning.

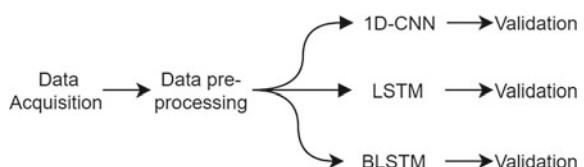
2.1 Dataset and Preprocessing

The dataset used in this work was provided by the research team at Jigsaw LLC of 159, 571 text examples which comprises 6 different target classes. These classes belong to *threat*, *insult*, *toxic*, *severe toxic*, *identity hate* and *obscene* category of text. Thus, this dataset enables the model to classify text between non-bully and bully with one of the categories given above. The text examples are first converted into unique tokens using the tokenization process where each token is a unique number that is used to feed into the neural network. A set of 40k most common words have been considered for training and testing by discarding rest of the rare words in the dataset. This has been done to expedite the training process. Padding has been used to make the text length uniform for all the text examples by truncating them to contain maximum 200 words. The dataset has been randomly sampled into training, validation and testing sets with 80%, 10% and 10% text examples, respectively. It is used by the model to learn from the data, extract features from the words, derive sentiment on the texts and finally evaluate it on the unseen data.

2.2 Neural Network Architectures

Several neural network architectures are considered for cyberbully detection purpose which are explained next.

Fig. 1 Flowchart for the proposed model



2.2.1 1D-CNN-Based Detection Network:

A sequential CNN model with 1D conv layer, dense layer and global max pool layer is used. For convolutional and initial dense layers, “ReLU” activation function is used, whereas the last dense layer uses a “sigmoid” activation function, in accordance with the multi-class classification outputs. The model uses an embedding layer with maximum length of 200. The embedding size and the number of maximum features have been set to 128 and 40k, respectively. Two dropout layers were used with dropout rate of 0.2. The CNN architecture is considered as follows: Embedding \Rightarrow Conv1d \Rightarrow Global Max Pool \Rightarrow Dropout \Rightarrow Dense \Rightarrow Dropout \Rightarrow Dense \Rightarrow Softmax(6). The model employs categorical cross-entropy loss function which is computed as

$$CE = - \sum_i^C t_i \log(s_i) \quad (1)$$

Adam optimizer [6] is a combination of the Adagrad [4] and RMSprop [11] optimizer, thus it combines their advantages and offers a superior performance, and therefore, it has been used for updating the weights and bias parameters as

(i) Weight initialization:

$$\rho_m \leftarrow 1, \rho_v \leftarrow 1, \mathbf{m} \leftarrow \mathbf{0}, \mathbf{v} \leftarrow \mathbf{0} \quad (2)$$

(ii) Parameter update rules:

$$\rho_m \leftarrow \beta_m \rho_m \quad (3)$$

$$\rho_v \leftarrow \beta_v \rho_v \quad (4)$$

$$\mathbf{m} \leftarrow \beta_m \mathbf{m} + (1 - \beta_m) \nabla_w J \quad (5)$$

$$\mathbf{v} \leftarrow \beta_v \mathbf{v} + (1 - \beta_v) (\nabla_w J \odot \nabla_w J) \quad (6)$$

$$\mathbf{w} \leftarrow \mathbf{w} - \alpha \left(\frac{\mathbf{m}}{\sqrt{\mathbf{v}} + \epsilon} \frac{\sqrt{1 - \rho_v}}{1 - \rho_m} \right) \quad (7)$$

where m, v represents first and second moment vector, respectively. Similarly, β_m and β_v represent exponential decay rate for first and second moment vector, respectively. ρ_m and ρ_v specifies the adaptive learning rate time decay factor. This parameter is similar to momentum and relates to the memory for prior weight updates. α in Eq. 7 represents learning rate or step size. $\nabla_w J$ represents gradient of cost function, J . ϵ in Eq. 7 is a small value to prevent division by zero condition. In Eq. 6, \odot refers to element-wise multiplication, and in Eq. 7, the operations under the root are also handled element-wise. Furthermore, the model was complied, and training with only 2 epochs offered sufficiently improved model evaluation performance.

2.2.2 LSTM

Long short-term memory (LSTM) [5] is an recurrent neural network containing feed-forward neural networks and feedback connections and is explicitly known for remembering patterns for longer periods of time. LSTM units contain cells, each divided into an input gate, output gate and a forget gate. They are used to tackle the vanishing gradient problem commonly faced during training conventional RNN models. The governing expressions determining the functional parameters are discussed in Eqs. (8)–(13), where $\sigma(.)$ denotes logistic sigmoid function, the quantities p, q, r, C denote the input, forget, output gates, cell activation vectors, respectively. The LSTM layer architecture is considered as follows: Embedding \Rightarrow LSTM \Rightarrow Global Max Pool \Rightarrow Dropout \Rightarrow Dense \Rightarrow Dropout \Rightarrow Dense \Rightarrow Softmax(6).

$$p_t = \sigma(x_t U^p + k_{t-1} W^p), \quad (8)$$

$$q_t = \sigma(x_t U^q + k_{t-1} W^q) \quad (9)$$

$$r_t = \sigma(x_t U^r + k_{t-1} W^r) \quad (10)$$

$$e_t = \tanh(x_t U^g + k_{t-1} W^g) \quad (11)$$

$$C_t = \sigma(q_t * C_{t-1} + p_t * e_t) \quad (12)$$

$$k_t = \tanh(C_t) * r_t \quad (13)$$

The proposed LSTM model was created with the combination of the base LSTM model, an added embedding layer, a dense layer and a dropout layer with dropout factor equal to 0.2. The model contains 5.16 million mac, which makes the complete model more sophisticated in terms of computation time and was trained only for 2 epochs.

2.2.3 Bidirectional LSTM (BiLSTM)

BiLSTM [9] is a slightly advanced version of regular LSTM. The main difference between these two is that BiLSTM will have its input in two ways. First is from the starting to the end of a sequence and then the last being from the end to the starting of the sequence. Due to this extra feature, the model is able to preserve information both from both past and future. This extra feature makes the LSTM much more efficient for text data. It is observed that various long text sequences hold valuable information at the end, and in such cases, BiLSTM are of best use. The model we created using BiLSTM had a total of 5, 218, 854 trainable parameters. The complete architecture is structured as follows: Embedding \Rightarrow Bidirectional \Rightarrow Global Max Pool \Rightarrow Dropout \Rightarrow Dense \Rightarrow Dropout \Rightarrow Dense \Rightarrow Softmax(6). The model was trained with 2 epochs. The results and relevant discussions are presented next.

3 Results and Discussions

The bidirectional LSTM model outperformed the other two models in terms of classification accuracy. However, in terms of training time, the *1D*-CNN model performed considerably better. The BiLSTM has more parameter than the LSTM because the number of parameters in the LSTM layer of BiLSTM is double to that of the unidirectional LSTM. The *1D*-CNN is ideal for systems where computational capacity is a crucial factor. Although the number of trainable parameters are maximum in the *1D*-CNN, it takes least time to train, and we achieved an accuracy significantly higher than LSTM. The best performance results are also highlighted in Table 1.

It can be noted that BiLSTM is best suited when performance is the primary concern, whereas *1D*-CNN requires significantly lesser training time with slightly lower performance than BiLSTM. The *1D*-CNN has a test accuracy of 0.9633, while LSTM and BiLSTM offer an accuracy of 0.9412 and 0.9745, respectively. It is also an observation that, the model shows errors while detecting sarcasm text from real threats, therefore accuracy and precision performance get slightly deteriorated.

4 Conclusion

In this paper, we proposed a comparison analysis between three popular deep learning techniques termed as CNN, LSTM and BiLSTM. From the proposed research, our results show that BiLSTM outperforms other models in terms of accuracy (0.9745), yet a significant difference in training time is observed. The BiLSTM model takes around 65x more time than the *1D*-CNN model, with slightly lesser test accuracy. Therefore, it is evident that *1D*-CNN can also be used where there is limitation in computational requirements. The *1D*-CNN and RNN models show high performance in detecting tokenized words, and thus, further research in text classification using these methods can be encouraged.

Table 1 Comparison analysis of evaluation parameters

Evaluation parameter	<i>1D</i> -CNN	LSTM	BiLSTM
Accuracy	0.9633	0.9412	0.9745
Precision	0.9651	0.9481	0.9701
Recall	0.9650	0.9438	0.9766
F-Measure	0.9650	0.9459	0.9733
Computation time	133s	4429s	8697s
MAC	5, 300, 934	5, 169, 654	5, 218, 854

Bold text represents best performance, out of three

References

1. P. Badjatiya, S. Gupta, M. Gupta, V. Varma, Deep learning for hate speech detection in tweets, in *Proceedings of the 26th International Conference on World Wide Web Companion* (2017), pp. 759–760
2. K. Dinakar, R. Reichart, H. Lieberman, Modeling the detection of textual cyberbullying, in *Fifth International AAAI Conference on Weblogs and Social Media* (2011)
3. N. Djuric, J. Zhou, R. Morris, M. Grbovic, V. Radosavljevic, N. Bhamidipati, Hate speech detection with comment embeddings, in *Proceedings of the 24th International Conference on World Wide Web* (2015), pp. 29–30
4. J. Duchi, E. Hazan, Y. Singer, Adaptive subgradient methods for online learning and stochastic optimization. *J. Mach. Learn. Res.* **12**(7) (2011)
5. S. Hochreiter, J. Schmidhuber, Long short-term memory. *Neural Comput.* **9**(8), 1735–1780 (1997)
6. D.P. Kingma, J.B., Adam, *A Method for Stochastic Optimization*. arXiv preprint [arXiv:1412.6980](https://arxiv.org/abs/1412.6980) (2014)
7. C. Nobata, J. Tetreault, A. Thomas, Y. Mehdad, Y. Chang, Abusive language detection in online user content, in *Proceedings of the 25th International Conference on World Wide Web* (2016), pp. 145–153
8. J.W. Patchin, S. Hinduja, Bullies move beyond the schoolyard: a preliminary look at cyberbullying. *Youth Violence Juvenile Justice* **4**(2), 148–169 (2006)
9. M. Schuster, K.K. Paliwal, Bidirectional recurrent neural networks. *IEEE Trans. Signal Process.* **45**(11), 2673–2681 (1997)
10. R.L. Servance, Cyberbullying, cyber-harassment, and the conflict between schools and the first amendment. *Wis. L. Rev.*, 1213 (2003)
11. T. Tieleman, G. Hinton, Lecture 6.5-rmsprop: divide the gradient by a running average of its recent magnitude. COURSERA: Neural Netw. Mach. Learn. **4**(2), 26–31 (2012)
12. E. Wulczyn, N. Thain, L. Dixon, Ex machina: personal attacks seen at scale, in *Proceedings of the 26th International Conference on World Wide Web* (2017), pp. 1391–1399

Comparative Study of Skin Lesion Segmentation and Feature Extraction in Different Color Spaces



A. Veeramuthu, A. Anne Frank Joe, B. S. Sathish, L. M. I. Leo Joseph,
P. Ganesan, and V. Elamaran

Abstract Skin lesion is the most common skin disorder in human. This may be cancerous (malignant) or benign (non-cancerous). Sometimes, it can be very dangerous as life threatening. In this paper, the skin lesion is segmented from the test image and extracted the most important features for further analysis. In skin lesion image segmentation, the whole data (image) is huddled into amount of small data groups (clusters) in a significant comportment. All the image pixels in the same cluster should share the common and unique characteristics among them. The accomplishment of the higher-level image processing techniques solely depends on the outcome of the segmentation process. Most of the skin lesion segmentation methods utilized the RGB color images and it has a lot of limitations. So in our work, the test image

A. Veeramuthu

Department of Information Technology, Sathyabama Institute of Science and Technology,
Chennai, India

e-mail: aveeramuthu@gmail.com

A. Anne Frank Joe

Department of Electronics and Instrumentation Engineering, Sathyabama Institute of Science and
Technology, Chennai, India

e-mail: annefrankjoe@gmail.com

B. S. Sathish

Department of Electronics and Communication Engineering, Ramachandra College of
Engineering, Eluru, Andhra Pradesh, India

e-mail: subramanyamsathish@yahoo.co.in

L. M. I. Leo Joseph

Department of Electronics and Communication Engineering, S.R. Engineering College,
Warangal, Telangana, India

e-mail: leojoel@srcwwarangal.ac.in

P. Ganesan (✉)

Department of Electronics and Communication Engineering, Vidya Jyothi Institute of
Technology, Aziz Nagar, C.B. Post, Hyderabad, India

e-mail: gaganeshnathan@gmail.com

V. Elamaran

Department of ECE, School of EEE, SASTRA Deemed University, Thanjavur, Tamil Nadu, India

e-mail: elamaran@ece.sastra.edu

is renovated to other spaces and then segmentation is performed. The investigational outcome clearly demonstrated the competency of the proposed method.

Keywords Segmentation · Skin lesion · Image quality measures · Feature extraction

1 Introduction

In clustering, the whole data is huddled as an amount of small data groups (clusters) in a significant comportment [1–3]. All the image pixels in the same cluster should share the common and unique characteristics among them [4]. These characteristics may be color, texture, or intensity [5, 6]. The accomplishment of the higher-level image processing techniques solely depends on the outcome of the segmentation process. It is very difficult to decide on the optimal segmentation method. Especially for the noisy images, segmentation becomes more difficult one [7]. This is due to both the image and noisy pixels are considered as the same category. This is the major reason the segmentation process is failed to achieve its goal, i.e., leads to either under or over segmentation [8]. For image segmentation, number of methods developed but most of them are application oriented [9]. Skin lesion is the most common skin disorder in human. This may be cancerous (malignant) or benign (non-cancerous). Sometimes, it can be very dangerous as life threatening. In this paper, the skin lesion is segmented from the test image and extracted the most important features for further analysis. Most of the skin lesion segmentation methods utilized RGB color images and it has a lot of limitations. So, in our work, the test data is renovated to other spaces and then segmentation is performed. The investigational outcome clearly demonstrated the competency of the proposed method.

2 Skin Lesion Segmentation—Proposed Method

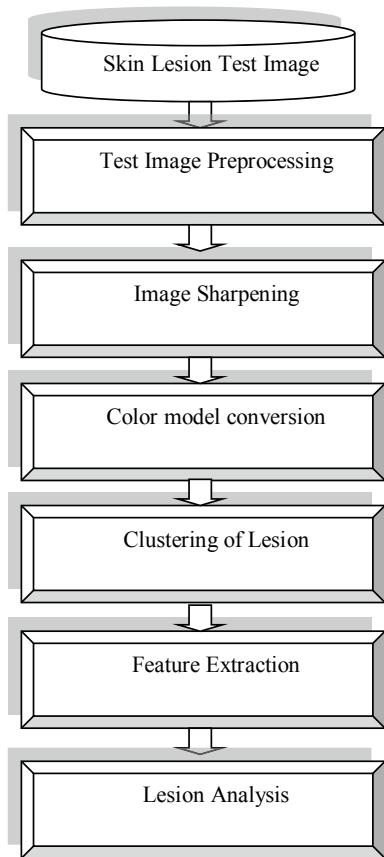
The suggested methodology for skin lesion clustering is depicted in Fig. 1. The outline of the proposed method is revealed as

Step 1: The test data is originally larger in size. So, the image is resized to lower the computational cost.

Step 2: The test data in RGB enhanced by image sharpening method. It is necessary to emphasize the fine details and edges in the test image. The proposed method utilized the Sobel detector of kernel size 3.

Step 3: MFCM is utilized to perform the clustering of the image in RGB space. The data is divided into three clusters as background, central part of the lesion, and spreaded area.

Fig. 1 Skin lesion clustering methodology



Step 4: Most of the skin lesion segmentation methods utilized RGB color images and it has a lot of limitations. So, in our work, the test image is renovated to other spaces and then segmentation is performed.

Step 5: The proposed method is evaluated using some image quality parameters as PSNR, MSE, and execution time.

Step 6: The necessary features are extracted and they are very useful to analyze the skin lesion images.

A simple but standard FCM clustering is mainly based on two parameters, centroid (cluster centers) and membership, to divide the whole image as distinct groups [8, 10, 2, 5]. The detachment among the pixels and centroid computed using Euclidean distance [9, 11] as illustrated in (1)

$$F_m(U, V) = \sum_{i=1}^c \sum_{k=1}^n \mu_{ik}^m \|x_k - v_i\|^2 \quad (1)$$

$$v_i = \frac{\sum_{k=1}^n \mu_{ik}^m X_k}{\sum_{k=1}^n \mu_{ik}^m} \quad (2)$$

$$\mu_{ik} = \left\{ \sum_{j=1}^c \left\{ \frac{\|x_k - v_i\|}{\|x_k - v_j\|} \right\}^{2/(m-1)} \right\}^{-1} \quad (3)$$

The foremost weakness of FCM is that does not afford spatial information which is noteworthy for image segmentation [4, 5, 12]. So, the MFCM is included this in the form of a weighted summation of the membership function.

$$S_{ij} = \sum_{k \in W(X_j)} U_{ik} \alpha_{k1} + \frac{\sum_{k \in (X_j)} U_{ik} \alpha_{k2}}{\sum_{t=1}^c \sum_{k \in W(X_j)} U_{tk}} \quad (4)$$

The membership function is modified as in (5)

$$U_{ij(new)} = \frac{U_{ij}^p * S_{ij}^q}{\sum_{k=1}^c U_{kj}^p * S_{kj}^q} \quad (5)$$

It is eminent that all pixel has a weight(W_{ji}) related to clusters.

$$W_{ji} = \frac{1}{1 + e^{-\left\{ \frac{\|x_j - v_i\|^2}{\sum_{j=1}^n \|x_j - v_i\|^2 (\frac{c}{n})} \right\}}} \quad (6)$$

Now, the new objective function of MFCM becomes (7)

$$MF = \sum_{k=1}^n \sum_{i=1}^c (U_{ik}^m W_{ji}^m) \|X_k - V_i\|^2 \quad (7)$$

3 Investigational Outcomes and Discussion

The test image which is employed to assess the competency of recommended methodology is exemplified in Fig. 2. The original size (2048 * 1526, 1.01 MB) of the image is resized to 300 * 245 (26.5 KB) to reduce the computational cost.

The test image in RGB is enhanced by image sharpening method. It is necessary to emphasize the fine details and edges in the test image as depicted in Fig. 3. The proposed method utilized the Sobel detector of kernel size 3.

MFCM is utilized to perform the separation of the skin lesion image in RGB space. It is segmented into three clusters as background, central part of the lesion

Fig. 2 Input image to assess the competency of the suggested method



Fig. 3 Enhanced (sharpened) version

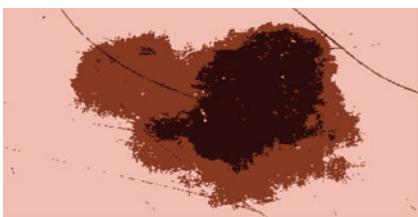


(dark brown color), and spreaded area (light brown color). The segmented image is represented in Fig. 4a. The error image is the dissimilarity between the sharpened image and its segmentation result. This is displayed in Fig. 4b.

Most of the skin lesion segmentation methods utilized RGB color images and it has a lot of limitations. So, in our work, the test data is converted into other spaces and then segmentation is performed. The image in CIELab is illustrated in Fig. 5.

The segmented image is portrayed in Fig. 6a. The image is segmented into three clusters as background, central part of the lesion, and spreaded area. The error image is the dissimilarity between the sharpened image and its segmentation result. This is displayed in Fig. 6b.

The image in CIELuv is demonstrated in Fig. 7. The segmented image is depicted in Fig. 8a. It is divided into three clusters as background, central part of the lesion, and spreaded area. The error image is displayed in Fig. 8b.



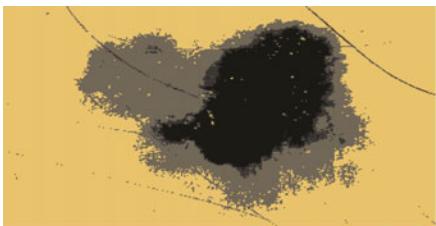
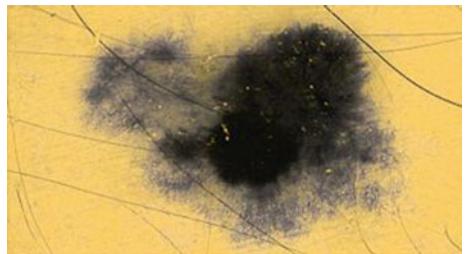
(a) segmented image



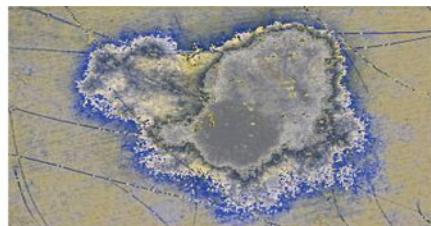
(b) error image

Fig. 4 Outcome of test image (RGB color space) for the proposed method. **a** Segmented image and **b** error image

Fig. 5 Test image in CIELab color space



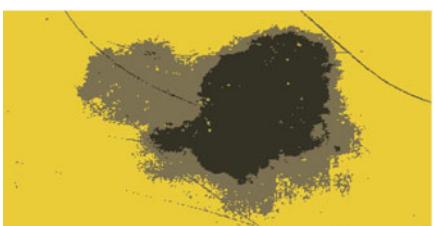
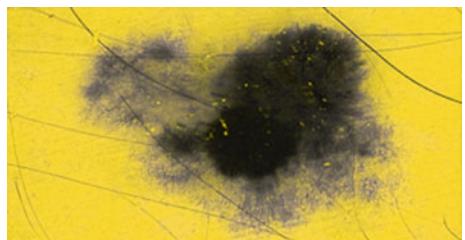
(a) segmented image



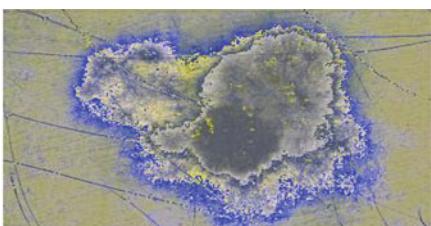
(b) error image

Fig. 6 End result of skin lesion image (CIELab color space) for the proposed method. **a** Segmented image and **b** error image

Fig. 7 CIELuv version of test image



(a) segmented version



(b) error image

Fig. 8 End result of skin lesion image (CIELuv color space) for the proposed method. **a** Segmented version and **b** error image

Fig. 9 HSI version of the image

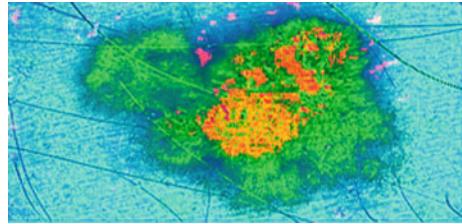


Figure 9 illustrated HSI version of the image. The segmented image is depicted in fig. 10a. Here, it is splitted as four clusters. The error image is demonstrated in Fig. 10b.

Figure 11 demonstrated the HSV version of the image and it is clustered as 5 segments. The segmented image is depicted in Fig. 12a. The error image is presented in Fig. 12b.

Table 1 illustrates the result of the suggested technique for skin lesion image segmentation in different color spaces. The proposed method is evaluated using some image quality measures such as PSNR, MSE, and execution time. The segmented image in HSV color space has very good PSNR (37.61) and MSE (3.36). However, when it comes to computational cost, CIELuv color space outperformed all other color spaces.

Table 2 represents the extracted features from the segmentation result. The extracted features are very useful to analyze the skin lesion images.

The illustrative representation of the impact of the proposed method for skin lesion clustering is displayed in Fig. 13.

4 Conclusion

The proposed method investigated the comparative analysis of skin lesion segmentation in different color systems. The trial image in RGB is transformed into other color systems. Skin lesion image in various color methods is segmented using modified FCM. The proposed method is evaluated using some image quality measures such as PSNR, MSE, and execution time. The segmented image in HSV color space has very good PSNR and MSE. However, when it comes to computational cost, CIELuv color space outperformed all other color spaces.

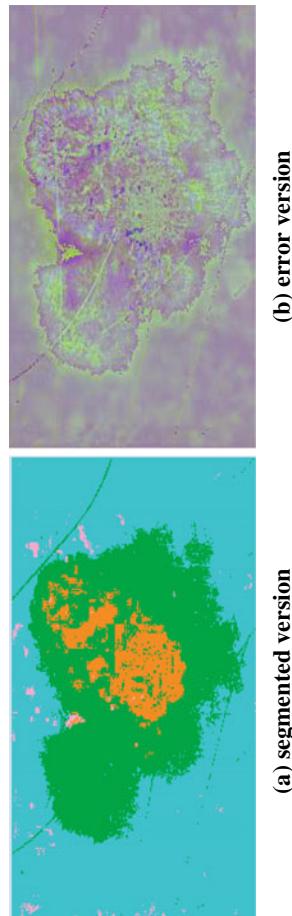


Fig. 10 Upshot of skin lesion image (HSI color space) for the proposed method. **a** Segmented version and **b** error version

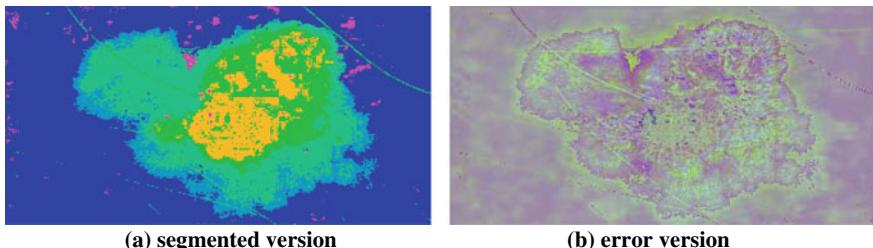
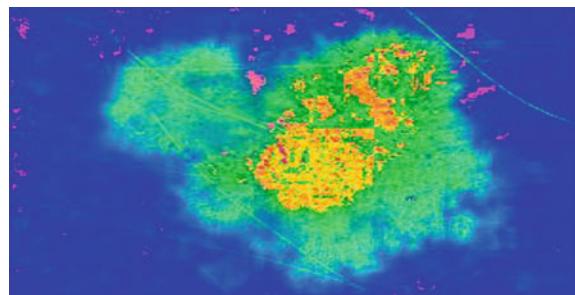
Fig. 11 HSV version**Fig. 12** End result of skin lesion image (HSV color space). **a** Segmented version, **b** error version**Table 1** Skin lesion image clustering in different color spaces

Image quality measures	RGB	CIELab	CIELuv	HSI	HSV
PSNR	22.66	21.93	21.68	36.22	37.61
RMS	18.74	20.41	21.00	3.94	3.36
Execution time (in s)	3.617	1.391	1.128	1.724	1.506
No. of clusters	3	3	3	4	5

Table 2 Extracted features from the segmentation result

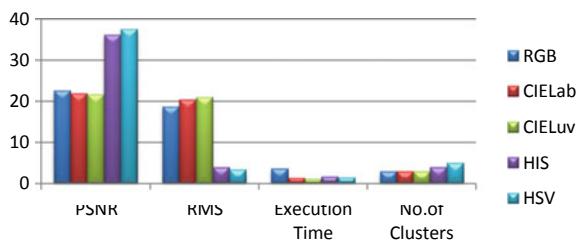
Features	RGB	CIELab	CIELuv	HSI	HSV
Area	37,989	35,688	35,003	37,427	31,323
Centroid	143	128	127	143	125
Orientation	-88	76	76	-88	77
Perimeter	1016	2030	2015	1026	1965
Euler number	-119	-162	-149	-233	-172
Mean	243.27	235.24	236.19	7.6357	7.5801
Standard deviation	13.445	18.417	19.937	3.2738	3.2363
Skewness	-1.7159	-1.5219	-1.4943	-0.0797	-0.0948

(continued)

Table 2 (continued)

Features	RGB	CIELab	CIELuv	HSI	HSV
Energy (histogram)	0.0675	0.0334	0.0288	0.0867	0.0876
Entropy	4.3837	5.4826	5.6252	3.7207	3.7095
Texture energy	0.0003	0.0006	0.0008	0.0002	0.0003
Inertia	488.91	330.07	388.01	1028.2	663.24
Correlation	0.8363	0.9547	0.9511	0.5062	0.7523
Texture entropy	12.633	12.2518	11.995	12.985	12.577

Fig. 13 Pictorial representation of the color space for skin lesion clustering



References

1. N.R. Pal, K. Pal, J. Bezdek, A Possibilistic Fuzzy C Means clustering algorithm. *IEEE Trans. fuzzy Syst.* **13**(4) (2005) 517–530
2. K.S. Chuang, H.L. Tzeng, J. Wu, T.J. Chen, Fuzzy c-means clustering with spatial information for image segmentation. *Comput. Med. Imag. Graph.* **30**(1), 9–15 (2006)
3. R. Krishnapuram, J. Keller, The possibilistic c means algorithm: Insights and recommendations. *IEEE Trans. Fuzzy Syst.* **1**(2), 98–110 (1993)
4. V. Kalist, P. Ganesan, B.S. Sathish, J.M.M. Jenitha, Possiblistic-fuzzy C-means clustering approach for the segmentation of satellite images in HSL color space. *Procedia Comput. Sci.* **57**, 49–56 (2015)
5. T. Havens, J.C. Bezdek, C. Leckie, L.O. Hall, M. Palaniswami, Fuzzy c-means algorithms for very large data. *IEEE Trans. Fuzzy Syst.* **20**(6), 1130–1146 (2012)
6. S. Jia, C. Zhang, Fast and robust image segmentation using an superpixel based FCM algorithm, in *IEEE International Conference on Image Processing (ICIP)*, 2014, p. 947–951
7. P. Ganesan, K. Palanivel, B.S. Sathish, V. Kalist, K.B. Shaik, Performance of fuzzy based clustering algorithms for the segmentation of satellite images—a comparative study, in *IEEE Seventh National Conference on Computing, Communication and Information Systems (NCC CIS)*, 2015, p. 23–27.
8. S. Krinidis, V. Chatzis, A robust fuzzy local information C-means clustering algorithm. *IEEE Trans. Image Process.* **19**(5), 1328–1337 (2010)
9. C. Correa, *A Comparison of Fuzzy Clustering Algorithms Applied to Feature Extraction on Vineyard*. Lecture Notes in Computer Science (Springer, 2012), p. 56–65
10. T. Lei, X. Jia, Y. Zhang, L. He, H. Meng, A.K. Nandi, Significantly fast and robust fuzzy C-means clustering algorithm based on morphological reconstruction and membership filtering. *IEEE Trans. Fuzzy Syst.* **26**(5), 3027–3041 (2018)
11. J.K. Parker, L.O. Hall, J.C. Bezdek, Comparison of scalable fuzzy clustering methods, in *Proceedings of the IEEE International Conference on Fuzzy Systems*, 2012, p. 1–9
12. Z. Wang, Q. Song, Y.C. Soh, K. Sim, An adaptive spatial information-theoretic fuzzy clustering algorithm for image segmentation. *Comput. Vis. Image Understand.* **117**, 1412–1420 (2013)

Recognition of Handwritten Digit in Free Space



Manas Chandan Behera, B. Shivalal Patro, and Ritik Gupta

Abstract In the area of computer vision, the handwritten digit recognition is gaining massive popularity. It showcases our work through this paper by putting a better and accurate method to understand manually written digit from 0 (zero) to 9 (nine) by a pen in mid-air/free space (3D). Our project identifies the nib of the pen in the user's hand and tracks the flow of the nib of the pen [which writes the digit in mid-air/free space (3D)]. Then, it matches the dataset to recognize the digit written by the user. Several experiments on the MNIST dataset have been done to obtain specified results. A whooping training set of 60,000 primary examples along with a test set of 10,000 examples has been obtained from the MNIST's database. Normalized size centred in fixed-size imaging is obtained from the database. The validation and effectiveness of the algorithm for handwritten digit inscribed in mid-air/free space (3D) with the use of the pen have been successfully obtained from our experiment outcomes. The concept of these projects can be effectively used in computer games, automatic teller machine (ATM), televisions, etc.

Keywords Handwritten digit · MNIST dataset · Convolution neural network (CNN) · OpenCV

1 Introduction

In this modern era, the increasing demand for deep learning technology has given us a striking kink in the sphere of machine learning making it more artificially intelligent. Deep learning's usage has been noteworthy in a large range of fields because of its

M. C. Behera · B. Shivalal Patro (✉) · R. Gupta

School of Electronics Engineering, KIIT Deemed to Be University, Bhubaneswar, India

e-mail: shivalalpatro@gmail.com

M. C. Behera

e-mail: manaschandan79@gmail.com

R. Gupta

e-mail: ritikgupta8936@gmail.com

numerous varieties of applications in surveillance, health, medicine, etc. In deep learning, the convolution neural network (CNN) has been extensively used in pattern recognition, speech recognition, text recognition, and handwritten digit recognition. CNN has been a powerful tool in the field of handwriting digit recognition field. In 2003, Simard et al. brought a preferred CNN structure for optical report analysis and weeding out the complex approach of neural network training [1].

In recent times, advancement in computing technology and person interfaces has caused a wonderful boom in interactive usages, such as gesture interface in a virtual environment. In today's lifestyle, the gesture equipment which records hand movement to control the interactive interface has shown an imminence growth in the past years. The person carries out the motion of the hand in the visibility range of the device so that it can interact with the program. Hence, this is the limitation. Also, there are many limitations associated with vision-based technology such as sensitivity to light being one of the main issues which are followed by the cost [2–6].

The neural network has been categorized into three subsections based on their shapes. Those are input, output, and intermediate (hidden) layers. The input layer takes the input dataset on the handwritten of the handwritten digit. The input layer must be composed of a proliferation of the selected image pixel size that is present in the dataset. The next layer, which is the hidden layer, receives its input from the first layer that allows us to understand the input [7].

2 Methodology

Human being handwriting varies from each other which makes human being handwriting recognition a prime subject of study and research. Due to such variations in written digit, it has become an extremely tough venture for a person to undergo a handwritten digit. So, the question arises what will happen if a human being is replaced by a computer? It will be a safer way to train our computers and predict those human written digits. As a result, it will be much easier to understand this when compared to manually written digits [8, 9].

We should make an accurate system if we want to anticipate the correct output by the system. A right database must be furnished to the device so that we can train the device to recognize the digit and get an accurate result. There are two fundamental principles to run the recognition system. Firstly, it relies upon how appropriate our function help extract the digit from the input faded. Secondly being a high-quality classifier [10], there is a sufficient amount of database present on the internet. In this work, we have preferred the MNIST database. Here, we preferred the MNIST database because it is open source and freely available on the Internet [11].

2.1 OpenCV

OpenCV's full form is open-source computer vision library. It is an open-source computer vision and machine learning software library in Python. OpenCV was built to furnish a common framework for the computer vision software and to accelerate the usage of device notion within the commercial products. The library has a greater than 2500 optimized algorithm, which includes a complete set of each conventional and today's computer vision and machine learning algorithms. OpenCV is very useful in detecting and recognizing the movement of humans in videos, their faces, etc. [12, 13]. In our project, we took the help of the OpenCV library for tracking the movement of the pen in mid-air/free space(3D).

2.2 Convolution Neural Network

Deep learning has been efficaciously carried out to a wide variety of machine learning programs. Convolution neural network (CNN) has wide applications such as identifying faces, handwritten digit, tumours, and a lot of different fields [14]. Convolution layer is the core building block of CNN. The parameters of the layer consist of a set of learnable filters having small receptive areas, however making a bigger through the whole depth of the enter volume for computing the dot product and generating a two-dimensional activation map of that filter. Each filter is convolved through the width and peak of the input quantity [15]. Figure 1 is an example that suggests the user moves his/her hand along the direction of the arrow as shown. The model used in the project is built using the open-source TensorFlow 2.2.0 library. The architecture of the model is proposed in this work, and any pretrained model is not being used. The model is built using the various functions of the TensorFlow library which include convolution 2-D layers, batch normalization layers, max pooling layers, a single dropout layer, and dense layers. In accordance to get maximum accuracy, the various hyperparameters like several layers, activation function, etc. are tuned using a general iterative approach.



Fig. 1 Handwriting trajectory

2.3 Flow Chart

The detailed flow chart of the project implemented in this paper is shown in Fig. 2. As the program starts, the main OpenCV connected to webcam opens up (i.e. the mainframe starts). Once the project is opened, we would have three options to do. First, use the pen functionality to draw something. Second, use the eraser function to erase something. The third is the prediction button. There is another chain associated with the prediction button, i.e. if passed it will do the preprocessing; thereafter compare the input with dataset do the prediction and display the result in the terminal window. If pressed the buttons, the program will terminate.

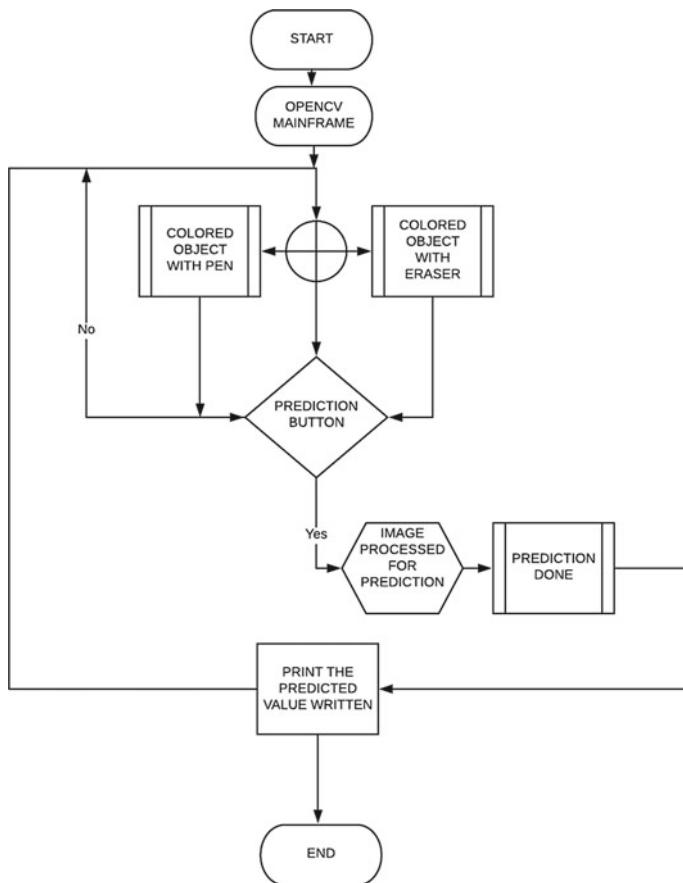


Fig. 2 Flow chart of the implemented project

3 Project Description

The initial task in the project is to identify the colour range of the target object with the open-source OpenCV library. As to easily operate in the image, it has been converted to HSV format and particulars of HSV range were found for the coloured target.

- The proceeding step is to create a mask in that particular HSV range for the mainframe, which would identify the target coloured object in the mainframe.
- For tracking the objects that are coloured within the frame, the algorithm used was to find contours in the masked frame and to identify the maximum contour (area). Then for getting out the coordinates of this contour, a bounding rectangle has been drawn around the contour, and coordinates of the rectangle are obtained.
- To draw in the air, after detecting the coordinates of the target coloured object, the algorithm merges the mainframe and a black canvas where the digits are drawn. The drawing is done based on drawing the lines between the coordinates of the bounding rectangle of contour from the previous frame (F-1) and the current frame (F).
- The rubber functionality was achieved by drawing in black colour on the canvas; thus after merging, it will serve as an eraser. For the prediction of the digital number, the canvas frame was used. It is then rescaled and made greyscale as part of preprocessing.
- The button functionality is achieved by background-image subtractor, which will identify the number of 255 pixels in a particular region and thus be used to run conditional statements (for pen/eraser and prediction).
- The prediction is done on the preprocessed canvas, and optimum results are observed in the terminal window.

The performance matrix for our project is shown Table 1, and the classification report is shown in Table 2.

Table 1 Performance matrix

Number	Precision	Recall	F1-score	Support
0	0.99	1.00	1.00	980
1	0.99	1.00	1.00	1135
2	0.99	1.00	0.99	1032
3	0.99	1.00	1.00	1010
4	1.00	0.99	0.99	982
5	1.00	0.99	0.99	892
6	1.00	0.99	0.99	958
7	1.00	0.99	0.99	1028
8	1.00	0.99	1.00	974
9	0.99	0.99	0.99	1009

Table 2 Classification report

	Precision	Recall	F1-score	Support
Accuracy	–	–	0.99	10,000
Micro average	0.99	0.99	0.99	10,000
Weighted average	0.99	0.99	0.99	10,000

Figure 3 shows an example how a digit is written in the window. And after preprocessing how the result is printed in the terminal window. Figure 4 symbolizes the loss plot which is plotted against loss reached vs several iterations (epoch). As the implementation of CNN algorithm is too complicated and the training time is very long, that's why only epoch 30 was set in the experiment. The model loss value is 0.0295 (or 2.95%). Here the training set loss (which is represented by the blue curve) and the test set loss (represented by the orange curve) are shown in our CNN model.

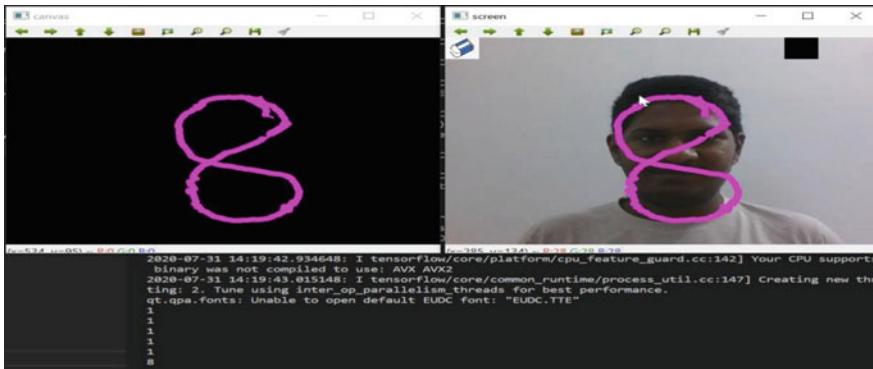
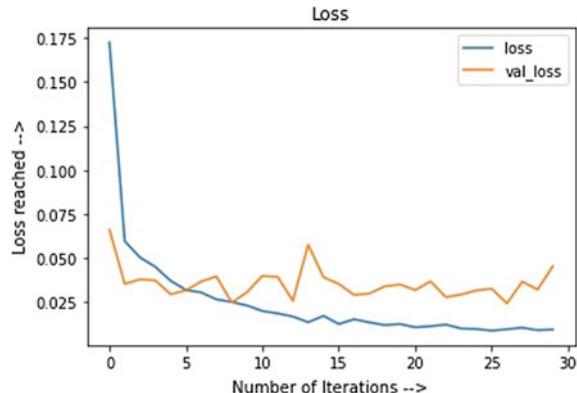
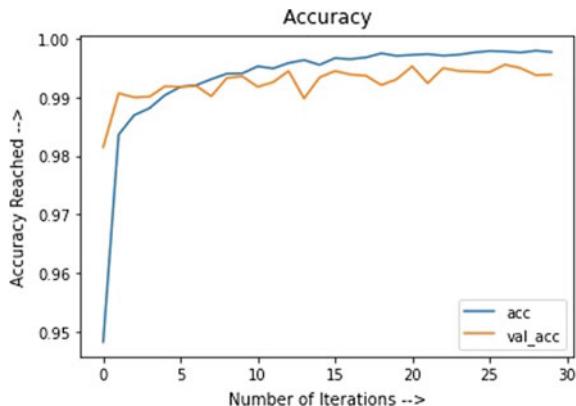
**Fig. 3** A sample snapshot of the terminal window**Fig. 4** Loss plot

Fig. 5 Accuracy plot

The model is trained on the MNIST dataset having 60,000 images which gave a very good result, and the performance matrix along with the classification report on the test data of 10,000 images gives proof that the model will be able to handle new data with high accuracy. Since while running the project, the images need to be normalized so that the model will be able to handle new image data. Furthermore, while increasing the scope the model would be made deeper to make the model robust to any new image data.

Figure 5 symbolizes the accuracy plot which is plotted against accuracy reached vs several iterations (epoch). Looking at the intricacies and the longevity in the training time of the application of CNN algorithm, only 30 was henceforth set in the experiment. The model accuracy value is 0.9953 (or 99.53%). The model accuracy value is 0.9953 (or 99.53%). Here the training set accuracy (which represents blue curve) and the test set accuracy (represents by the orange curve) are shown in our CNN model.

4 Conclusion

This paper has justified a scientific algorithm model that can help construct an impelling classifier for gesture recognition in mid-air/free space (3D). Using deep learning technology for handwritten digit recognition has been the crux of the entire paper. We have experimented with using CNN using the MNIST dataset. In this paper, we obtained very promising results with a very high level of accuracy. Making use of a CNN with TensorFlow has shown a successful result with an accuracy rate of 0.9953 (or 99.53%) which reflects that there is a negligible amount of error in our project. Further, this trained model has a very good scope in alphabets recognition and also can be compared with the related works.

References

1. S. Ahlawat, A. Choudhary, A. Nayyar, S. Singh, B. Yoon, Improved handwritten digit recognition using convolution neural network (CNN). *Sensor* **20**, 3344 (2020). <https://doi.org/10.3390/s20123344>
2. J.-S. Wang, F.-C. Chuang, An accelerometer-based digital pen with a trajectory recognition algorithm for handwritten digit and gesture recognition. *IEEE Trans. Ind. Electron.* **59**(7), 2998–3007 (2012)
3. E. Sato, T. Yamaguchi, F. Harashima, Natural interface using pointing behaviour for human-robot gesture interaction. *IEEE Trans. Ind. Electron.* **54**(2), 1105–1112 (2007)
4. A.D. Cheok, Y. Qiu, K. Xu, K.G. Kumar, Combined wireless hardware and real-time computer vision interface for tangible mixed reality. *IEEE Trans. Ind. Electron.* **54**(4), 2174–2189 (2007)
5. T.S. Ajay, Handwritten digit recognition using convolution neural networks. *Int. Res. J. Eng. Technol. (IRJET)* **4**(7), 2971–2976 (2017)
6. S. Patil, D. Kim, S. Park, Y. Chai, Handwriting recognition in free space using WIMU-based hand motion analysis. *J. Sens.* (2016)
7. T. Makkar, Y. Kumar, A. Dubey, A. Rocha, A. Goyal, Analogizing time complexity of KNN and CNN in recognizing handwritten digit, in *2017 Fourth International Conference on Image Information Processing (ICIIP)* (2017)
8. Priya, R. Singh, S. Changlani, Review on handwritten digit recognition. *Int. J. Novel Res. Dev.* **2**(4), 1–13 (2015)
9. T. Makkar, Y. Kumar, A.K. Dubey, A. Rocha, A. Goyal, Analogizing time complexity of KNN and CNN in recognizing handwritten digits, in *2017 Fourth International Conference on Image Processing (ICIIP)*
10. M.Z. Alom, P. Sidike, T.M. Taha, V.K. Asari, Handwritten Bangla digit recognition using deep learning. *Comput. Vis. Pattern Recognit. (cs.CV)* [arXiv:1705.02680](https://arxiv.org/abs/1705.02680) (2017)
11. MNIST dataset link. <http://yann.lecun.com/exdb/mnist/>
12. I. Culjak, D. Abram, T. Pribanic, H. Dzapo, M. Cifrek, A brief introduction to OpenCV. MIPRO 2012, 21–25 May 2012
13. Brief explanation about OpenCV <https://opencv.org/about/#:~:text=OpenCV%20>
14. H.A. Alwzmary, H.M. Albehadili, Y.S. Alwan, N.E. Islam, Handwritten digit recognition using convolution neural networks. *Int. J. Innov. Res. Comput. Commun. Eng.* **4**(2) (2016)
15. M.A. Hossain, M.M. Ali, Recognition of handwritten digit using convolution neural network (CNN). *Glob. J. Comput. Sci. Technol. (D)* **XIX**(II) (2019). Version I

An Area-Efficient 1.23 V Current-Mode Bandgap Reference with Start-Up Circuit



Aditya Kumar Hota and Kabiraj Sethi

Abstract A 1.23 V bandgap reference circuit with a startup circuit operating for $-40\text{--}120\text{ }^{\circ}\text{C}$ temperature range is presented in this paper. By the use of a cascode current mirror circuit, the supply voltage dependency is reduced here. A start-up circuit is proposed to come out from the zero-current condition of the reference circuit. The power consumed by the circuit is $94\text{ }\mu\text{W}$, whereas the on-chip area occupied by the circuit is $69 \times 87\text{ }\mu\text{m}^2$. The circuit is designed in UMC 180 nm CMOS mixed-mode technology.

Keywords Bandgap voltage reference (BGR) · Complementary to absolute temperature (CTAT) · Proportional to absolute temperature (PTAT) · Start-up circuit

1 Introduction

All types of circuits, whether a mixed-signal, digital or analog, require a constant supply voltage irrespective of voltage, temperature, and process variations. The bandgap reference (BGR) circuit is the one that meets all these requirements. Different BGR circuits have been proposed by many authors for constant output. Bogoda et al. [1] have used a switched current technique and current memory technique to reduce the area.

Omran et al. [2] has proposed a look-up table-based approach to design BGR that eliminates the need of simulator for a required specification. But this approach requires a lot of data to meet the accurate result. Arne et al. [3] have proposed a BGR without resistors keeping all the transistors in a strong inversion region to make it temperature-insensitive. But, one of the major problems in the BGR circuits is the zero-current condition in which the circuit never reaches the bias condition and fails to provide the necessary reference voltage. This problem can be solved by using

A. K. Hota · K. Sethi
Veer Surendra Sai University of Technology, Burla, Odisha 768018, India
e-mail: akhota_etc@vssut.ac.in

a start-up circuit which helps to initialize the BGR without disturbing its normal operation.

The paper organization is as follows: Sect. 2 presents the literature survey of the conventional BGR circuits, Sect. 3 describes the proposed BGR with start-up circuit, the simulation results and discussion are carried out in Sect. 4, and finally, Sect. 5 concludes the paper.

2 State of Art BGR Circuit

The conventional BGR circuit is shown in Fig. 1a, where the bipolar junction transistors (BJTs) behave as diodes. The output reference voltage is the difference voltage between two forward-biased diodes with different ratioed currents that is

$$V_{\text{ref}} = \alpha V_{\text{EB}1} + \beta \Delta V_{\text{EB}}, \text{ where } \Delta V_{\text{EB}} = V_{\text{EB}1} - V_{\text{EB}2} = V_T \ln(N)$$

$$\text{Thus } V_{\text{ref}} = (\alpha + \beta) V_{\text{EB}1} - \beta V_{\text{EB}2} \quad (1)$$

Equation 1 presents the first order voltage reference neglecting the higher-order dependencies. α and β are two positive constants, V_T is the thermal equivalent voltage (kT/q , T = absolute temperature, k = Boltzmann constant, and q = charge of an electron) and N is the number of diodes that transistor Q2 represents. Thus, ΔV_{BE} has a positive temperature coefficient (TC) and behaves as a proportional to

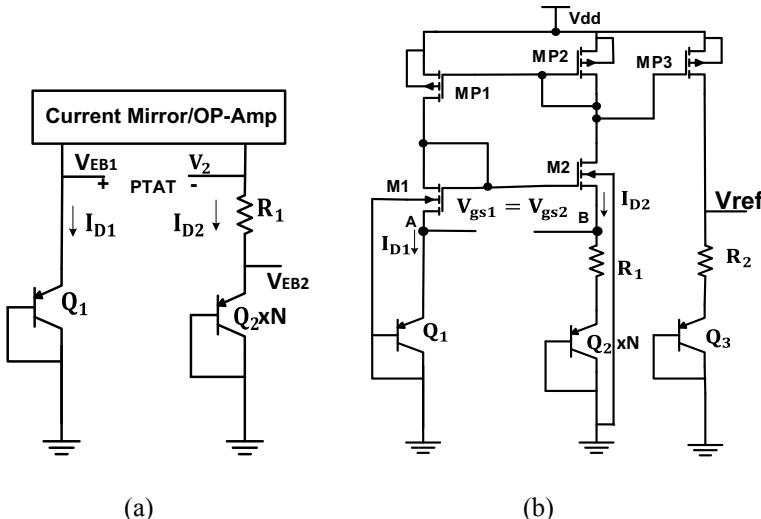


Fig. 1 The basic bandgap reference circuit

absolute temperature (PTAT). $V_{BE1} = (q/kT) \ln (I_{D1}/I_S)$, on the other hand, has a negative temperature coefficient and behaves as a complementary to absolute temperature (CTAT). The summation of the PTAT and CTAT terms leads to a temperature-independent output reference voltage. The PTAT and CTAT have the TCs as $\frac{\partial}{\partial T} V_T = 0.085 \text{ mV/}^\circ\text{C}$ and $\frac{\partial}{\partial T} V_{EB1} = -1.6 \text{ mV/}^\circ\text{C}$, respectively. The variation of a constant voltage, its variation concerning temperature should be zero ($\frac{\partial}{\partial T} V_{ref} = 0$) that is

$$\alpha_1 \frac{\partial}{\partial T} V_{BE1} + \alpha_2 \frac{\partial}{\partial T} V_T = 0 \quad (2)$$

where $\alpha_2 = \alpha_1 \cdot \beta$. For $\alpha_1 = 1$ the value of α_2 is roughly 18.82. Thus

$$V_{ref} = V_{BE1} + \alpha_2 V_T = 0.7 \text{ V} + 18.82 \times 26 \text{ mV} = 1.2 \text{ V} \quad (3)$$

In many papers, the V_{BE1} is reported between 0.7 and 0.75 V. So, the V_{ref} also reported from 1.2 to 1.25 V. The two voltages V_{EB1} and V_2 (in Fig. 1a) should be the same and this can be achieved by either implementing a current mirror circuit [4–6] or an operational amplifier (OP-Amp) or a combination of both [7, 8]. In this paper, we focus on the current mirror BGR and the basic current mirror BGR circuit is shown in Fig. 1b. Here, M1 and M2 have the same size to have $I_{D1} = I_{D2}$.

3 Proposed BGR Circuit

The proposed BGR circuit (Fig. 2) consists of a core circuit, a current mirror circuit, and a start-up circuit. The BGR core comprises of BJTs Q1–Q4 and resistors R1 and R2 ($R2 \gg R1$). Here, N is taken as 2 to reduce the area and power of the BGR circuit, although the symmetry may not be achieved to reduce the parasitic components. The current in each of the branches are the same as

$$I_{D1} = I_{D2} = I_{D3} = \frac{V_T \ln (2)}{R1} \quad (4)$$

Here, cascode current mirrors circuits are used to reduce the supply voltage dependency on the reference voltage. The current mirror circuit consists of MP1–MP6 (PMOS) and M1–M7 (NMOS) transistors. The start-up circuit consists of NMOS transistors M5–M7 only. The aspect ratio of M7 is large so that it will not allow the start-up circuit to disturb the normal operation of the BGR after start-up. M6 acts as a resistor not high as R1 and R2 and so it is a small transistor. M5 helps to drop the voltage at the gate of MP2 to conduct it and supply current to the subsequent circuit. Similarly, M7 helps to build a voltage at the gate of M8 to turn it on supplying the required current. Table 1 shows the specifications of all the components used in the BGR circuit.

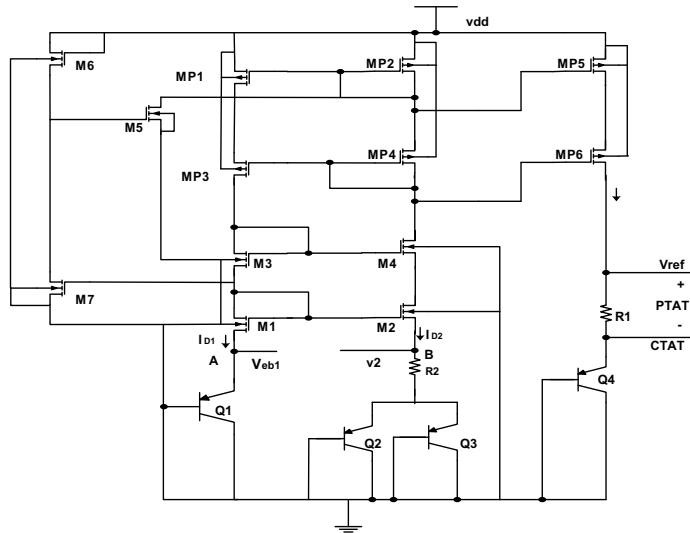


Fig. 2 Proposed BGR circuit with start-up circuit

Table 1 Components' specification of BGR circuit

Component	Specification	Component	Specification
MP1–MP6	Length = 1 μm , Width = 10 μm	M7	Length = 1 μm , Width = 100 μm
M1–M4	Length = 1 μm , Width = 10 μm	Q1–Q4	Length = 10 μm , Width = 10 μm
M5	Length = 1 μm , Width = 2 μm	R1	1.9 K
M6	Length = 1 μm , Width = 25 μm	R2	54 K

4 Post-layout Simulation Result

The BGR circuit is designed in United Microelectronics Corporation (UMC) 180 nm CMOS mixed-mode technology. All the simulations results are obtained by the SPECTRE simulator.

Figure 3 shows the layout diagram of the BGR circuit. It occupies an area of $69 \times 87 \mu\text{m}^2$.

The current drawn by the circuit is $9.4 \mu\text{A}$ in each branch and the total power consumed is $97 \mu\text{W}$.

Figure 4 shows that the output is constant at 1.23 V for a supply greater than 2.8 V . Figure 5 shows the maximum V_{ref} at 27°C as 1.2325 V and a minimum voltage of 1.2313 V at 120°C , resulting in a voltage difference of 1 mV . Thus, the TC of the reference voltage is $13.3 \text{ ppm}/\text{C}$. The effect of process variation is shown in Fig. 6. It is observed that for fast-fast (FF) and slow-slow (SS) corner, there is a maximum deviation of 3 mV at 120°C . For other corners fast-n slow-p (FNSP) and slow-n fast-p (SNFP), the deviation is negligible. Figure 7 shows the transient response of

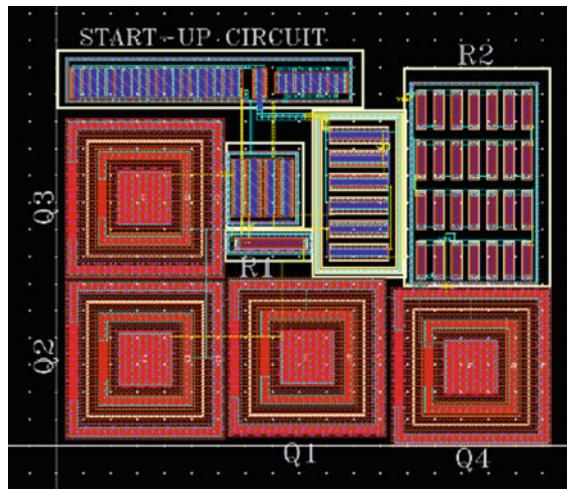


Fig. 3 The layout of the BGR circuit

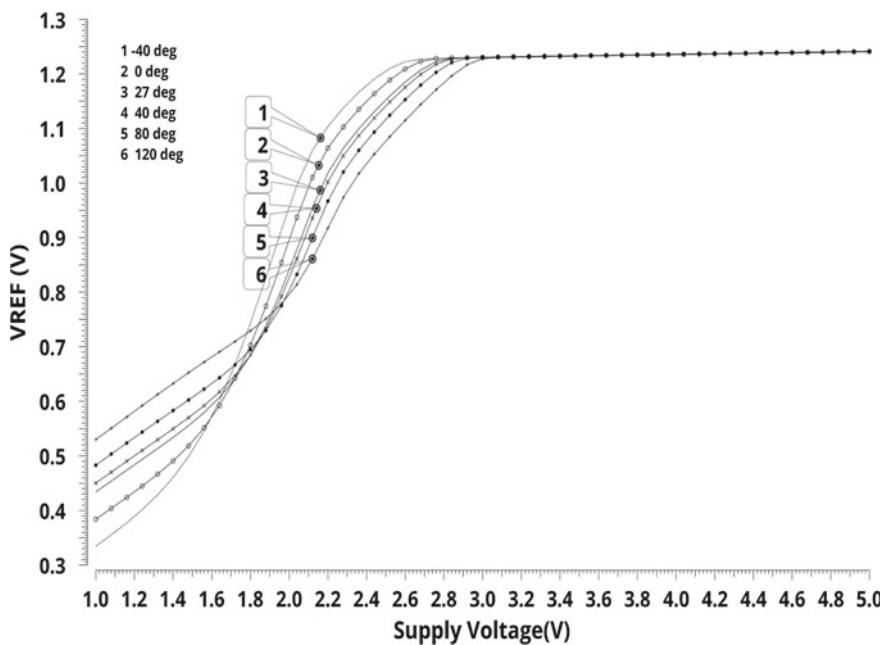


Fig. 4 Reference voltage versus supply voltage and temperature variations

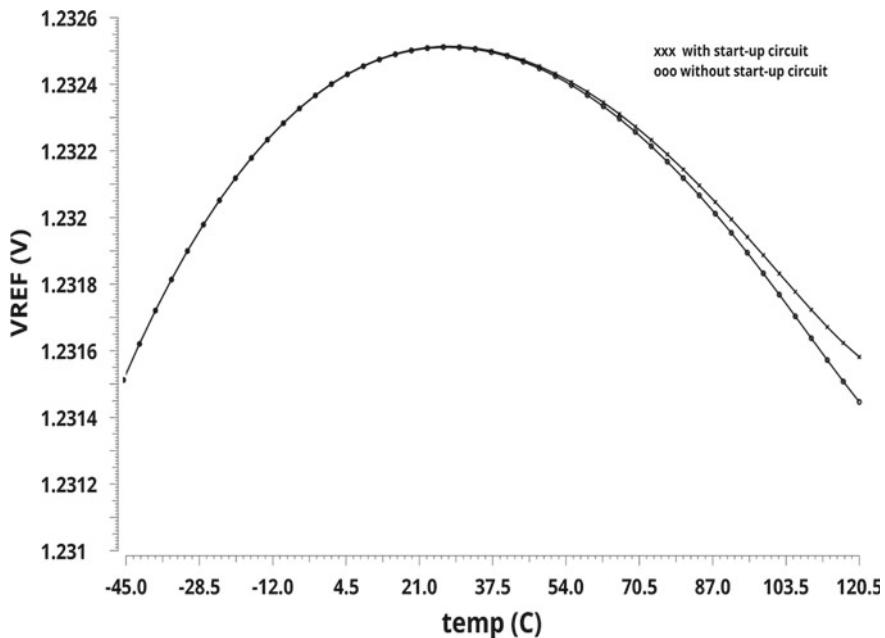


Fig. 5 Reference voltage versus temperature variation for a supply voltage of 3.3 V

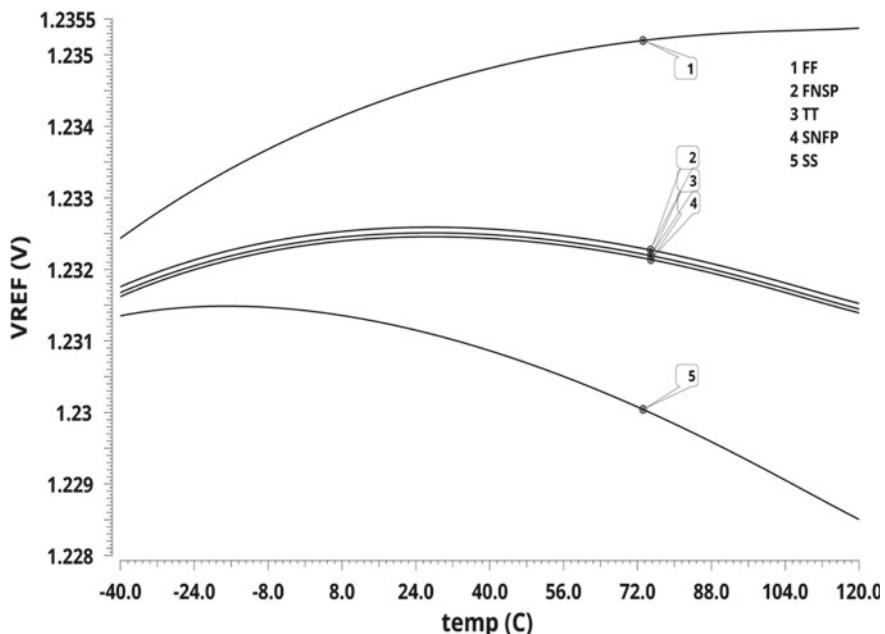


Fig. 6 Reference voltage for process variations

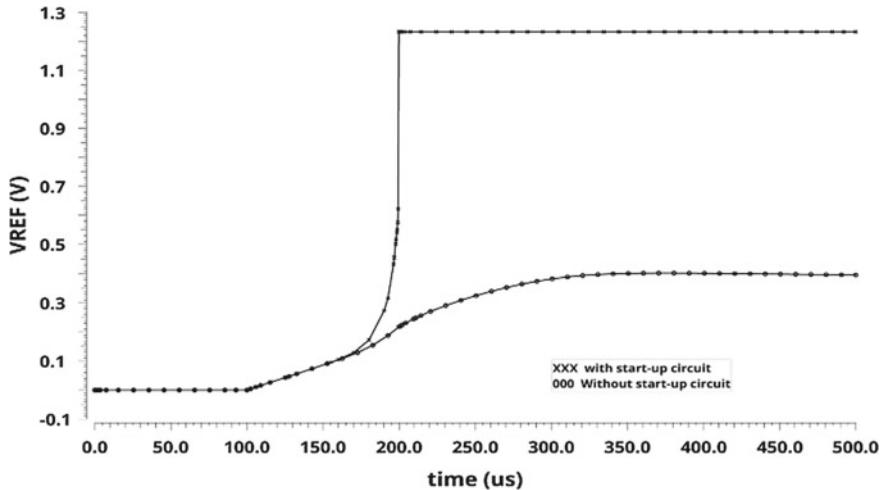


Fig. 7 Transient response of the reference voltage

Table 2 Performance comparison

References	Output (V)	Technology (nm)	Supply voltage (V)	TC (ppm/°C)	Temperature range (°C)	Power (μW)	Area (mm ²)
[5]	0.54	40	1.425–3.6	41.5	−40 to 85	–	0.0094
[7]	1.09	180	1.2–1.8	147	−40 to 120	–	0.0294
[8]	1.24	130	3.3	11.5	−40 to 125	600	–
This work	1.23	180	2.8–4.2	13.3	−40 to 120	97@3.3 V	0.0060

the BGR output with and without the start-up circuit. The response without start-up circuit is as such due to the zero-current condition that is eliminated by the use of start-up circuit.

Table 2 presents the comparison of the proposed work with the other works. It is found that the temperature coefficient and power dissipation of the proposed work is better than other BGR circuits.

5 Conclusion

This article presents a BGR circuit which can resist the supply, temperature, and process variations. The implementation of the cascode current mirror works very well in canceling the supply voltage dependency, and the start-up circuit helps the BGR to come out from the zero-current condition. The area and power consumptions

of the proposed circuit are also small making the BGR more suitable for different circuits.

References

1. A.I.U. Bogoda et al., An area-efficient CMOS bandgap reference utilizing a switched-current technique. *IEEE Trans. Circuits Syst.* **57**(10), 762–766 (2010)
2. H. Omran et al., Systematic design of bandgap voltage reference using precomputed lookup tables. *IEEE Access* **7**, 100131–100142 (2019)
3. T.R. Visvanathan, A CMOS bandgap reference without resistors. *IEEE J. Solid-State Circuits* **37**(1), 81–83 (2002)
4. R. Akshaya, Design of an improved bandgap reference in 180 nm CMOS process technology, in *IEEE International Conference On Recent Trends in Electronics Information & Communication Technology (RTEICT)*, 2017, p. 521–524
5. H. Chen et al., A Sub-1 ppm/ $^{\circ}\text{C}$ precision bandgap reference with adjusted-temperature-curvature compensation. *IEEE Trans. Circuits Syst.* 1–10 (2017). <https://doi.org/10.1109/TCSI.2017.2658186>
6. K.K. Lee et al., A sub-bandgap reference circuit with an inherent curvature-compensation property. *IEEE Trans. Circuits Syst.* **62**(1), 1–9 (2015). <https://doi.org/10.1109/TCSI.2014.2340553>
7. Y. Osaki et al., 1.2-V supply, 100-nW, 1.09-V bandgap and 0.7-V supply, 52.5-nW, 0.55-V subbandgap reference circuits for nanowatt CMOS LSIs. *IEEE J. Solid-State Circuits* **48**(6), 1530–1538 (2013). <https://doi.org/10.1109/JSSC.2013.2252523>
8. A. Pajkanovic, Bandgap voltage reference in 130 nm: design and schematic level simulation, in *International Symposium on Industrial Electronics INDEL*, p. 89–93 (2014)

Design of AlGaAs/InGaAs/GaAs-Based PHEMT for High Frequency Application



Geeta Pattnaik and Meryleen Mohapatra

Abstract The high electron mobility transistor popularly known as HEMT comes in the category of hetero-structure field effect transistor. With time, it gained popularity by exhibiting improved RF and DC performance in higher power and higher frequency applications. Another important advantage offered by HEMT is low noise. Operation of HEMT is related to the formation of a two dimensional electron gas (2-DEG) in the channel layer. The 2-DEG region helps in minimizing the scattering and collision of electrons thereby offering high mobility of electrons, high electron velocity, and lower noise. Keeping in mind, the applications and advantages offered by HEMT, we propose an AlGaAs/InGaAs/GaAs Pseudomorphic HEMT (PHEMT) structure. The design is simulated in SILVACO ATLAS tool. Input-output characteristic, transconductance, maximum frequency, and cut-off frequency of the proposed PHEMT is observed and reported. The maximum drain current obtained is 720 mA/mm at $V_{ds} = 2$ V and transconductance of nearly 591 mS/mm. While evaluating the frequency performance of the proposed structure, a cut-off frequency of 22.5 GHz and maximum frequency of 87.5 GHz are observed while simulation.

Keywords pHEMTs · Maximum frequency · Cut-off frequency · Transconductance

1 Introduction

HEMT have proved its capability and efficiency in high power, high frequency, high speed, and low noise applications. RF application such as satellites, cellular telecommunications, radars, direct broadcast receivers have already witnessed the advantage of including HEMTs [1–6]. At very low gate voltages, HEMT is able to

G. Pattnaik (✉) · M. Mohapatra

Department of Electronics and Communication Engineering, ITER, Siksha ‘O’ Anusandhan
Deemed to be University, Bhubaneswar, India
e-mail: geetapattnaik2@gmail.com

M. Mohapatra
e-mail: meryleenmohapatra@soa.ac.in

report high values of current due to high electron mobility and minimized scattering in 2-DEG region [7]. HEMTs have found usage in millimeter wave applications and microwave design owing to its efficiency. Toward the end of 70s decade, enhancement in molecular beam epitaxy (MBE) technology and introduction of quantum well structures was witnessed [8]. Alongwith his team at Fujitsu, T. Mimura faced issues related to high-density of the surface states close to the interface while dealing with GaAs MESFETs hence incorporation of modulation doped hetero-junction super-lattice idea came into play [9]. Those structures, however, still faced some issues, for which the idea of controlling the electrons in the super-lattice. Mimura practically achieved this by introducing a Schottky contact on a hetero-junction. With that, the GaAs/AlGaAs HEMT was first introduced. The hetero-junction formed by GaAs/AlGaAs helps to achieve the maximum electron mobility by preventing the backflowing of channel charge carriers and minimizing the electron scattering in the 2-DEG channel layer. Among the various catagories of HEMT, PHEMTs have gained their importance by improving the barrier between channel and substrate and for which it is able to prevent the leakage of mobile carriers into the GaAs substrate [10].

We focused on designing a PHEMT that will exhibit higher drain current at low drain voltage, higher transconductance value, high values of cut-off frequency and maximum frequency, which will make the HEMT suitable for RF applications. We plotted the drain current for various gate voltages and transconductance by varying the drain voltage, maximum frequency and cut-off frequency for different drain and gate voltages.

2 Device Structure

Pseudomorphic HEMT (PHEMT) structure using AlGaAs/InGaAs/GaAs is reported in this paper.

From the design perspective of our device, the structure starts with a 517 nm thick p-type GaAs substrate. A 170 nm thick buffer layer composed of $\text{Al}_{0.22}\text{Ga}_{0.78}\text{As}$ lies on the substrate layer. A 18 nm undoped $\text{In}_{0.18}\text{Ga}_{0.82}\text{As}$ channel layer is taken on the buffer layer. Above the channel layer, a 20 nm undoped $\text{Al}_{0.22}\text{Ga}_{0.78}\text{As}$ spacer layer is placed. The spacer layer separates the channel layer from the dopant impurities and help form the 2-DEG layer effectively which results in high mobility of electrons. Above the spacer layer lies the 30 nm $\text{Al}_{0.22}\text{Ga}_{0.78}\text{As}$ supply layer. Between the supply and spacer layer, a 1 nm thin and heavily doped delta doping layer is taken. Two GaAs cap layers present above the supply layer material. The cap layers help in reducing the access resistance. Tables 1 and 2 are tabular representation of the dimensional parameters and doping concentrations of the layers of proposed HEMT structure.

Table 1 Tabular demonstration of doping concentration of layers of proposed design

Device parameter	Values	Width of layers
Supply layer doping	$10^{17}/\text{cm}^3$	30 nm
Delta doping layer	$10^{18}/\text{cm}^3$	1 nm
Spacer layer doping	Undoped	2 nm
Buffer layer doping	$10^{17}/\text{cm}^3$	170 nm
Cap layer doping	$3.5 \times 10^{18}/\text{cm}^3$	

Table 2 Dimensional demonstration of the proposed design

Device parameters	Values
Gate length	$0.35 \mu\text{m}$
Drain and source length	$0.12 \mu\text{m}$
Distance between gate and source	$0.06 \mu\text{m}$
Distance between drain and gate	$0.08 \mu\text{m}$
Work function of gate	4.43 eV

3 Optimization Techniques Used in the Proposed HEMT Design

With increasing demand, the RF and DC performance of HEMT requires more improvement. This enhancement in performance is achieved by optimizing the HEMT device. Several optimization techniques for HEMT are available. Depending on our requirements, we need to incorporate any single or combination of techniques to achieve the desired result. We have considered the below techniques to make our proposed structure suitable for RF applications.

3.1 Gate Length

A short gate length helps to achieve higher microwave and radio frequency. A shorter gate length, however, gives rise to short channel effects that eventually hampers the HEMT's performance. Hence, the gate length of HEMT must be optimized with the view of improving the HEMT performance in terms of drain current, transconductance, cut-off frequency, and maximum oscillation frequency, while avoiding the short channel issues.

3.2 *Channel Thickness*

Channel thickness has a remarkable contribution toward the RF and DC performance of HEMT devices. The relation between thickness of channel and performance of HEMT is reported that by reducing the thickness of channel short channel effects are improved but the on-current and the transconductance are reduced.

3.3 *Indium Concentration*

It is observed from various studies that increasing the content of indium in the channel will help enhance the HEMT performance. Inspite its advantage, impact ionization occur in the channel by increasing indium concentration leading to issues like kink effects on the HEMTs output characteristics. The HEMT parameters like drain current, transconductance, maximum oscillation frequency and cut-off frequency are affected by the concentration of Indium in the channel. Hence, optimization of the concentration of indium in the channel plays a major role.

3.4 *Silicon δ -Doping Technology*

Silicon δ -doping is used between spacer layer and barrier layer. It is a very thin but heavily doped layer that helps by supplying large number of electrons to channel. By increasing the concentration of silicon δ -doping, electron charge density in the channel increases leading to improvement in drive current, transconductance, and cut-off frequency.

4 Result and Discussion

The structure was simulated and verified using the Silvaco Atlas TCAD software. Figure 1b shows the simulated PHEMT structure. We have used some specific models while simulating the design in Silvaco Atlas tool. The in-built equations for the specific models help us obtain the results of the parameters reported in this paper. The mobility models used are CONMOB, FLDMOB, CONSRH and Gummel, Newton convergence methods. CONMOB can be used only by GaAs and silicon-based structure, the CONMOB uses a lookup table from where it gets the information about the electron and hole mobility for specific concentration. FLDMOB model (Parallel Electric Field Dependence) reports the electric field dependent mobility and CONSRH model (Concentration dependent Shockley-Read Hall), holds the information about occurrence of recombination due to trapping using the minority carrier

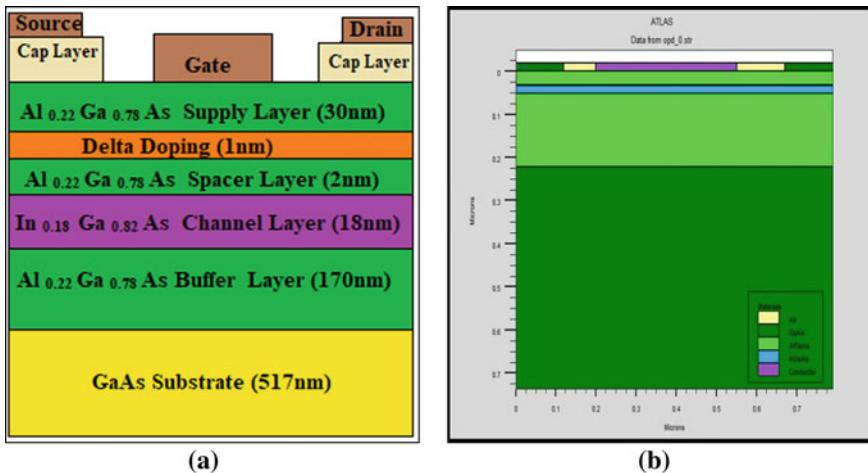


Fig. 1 **a** Schematic diagram of AlGaAs/InGaAs/GaAs Pseudomorphic HEMT, **b** simulated proposed structure in Silvaco Atlas Tool

lifetimes. As the transconductance (g_m) and drain current (I_d) are considered significant parameters for device characterization, we have plotted them at various voltages. We have used an optimized short gate length of $0.35\ \mu\text{m}$ along with a silicon δ -doping of $10^{18}/\text{cm}^3$ and Indium concentration of $x = 0.18$ that helped us to obtain a high drain current that eventually increased the transconductance, cut-off frequency, and maximum oscillation frequency. The optimized channel thickness helped to mitigate the short channel effects without disturbing the current and transconductance.

Figure 2 shows the I_d versus V_{ds} plot for $V_{gs} = 0, -0.4$, and $-1.8\ \text{V}$ while Fig. 3 represents the energy band diagram of the proposed HEMT. Figure 4 represents I_d versus V_{gs} for $V_{ds} = 6, 4$, and $2\ \text{V}$. High drain current is obtained due to the inclusion

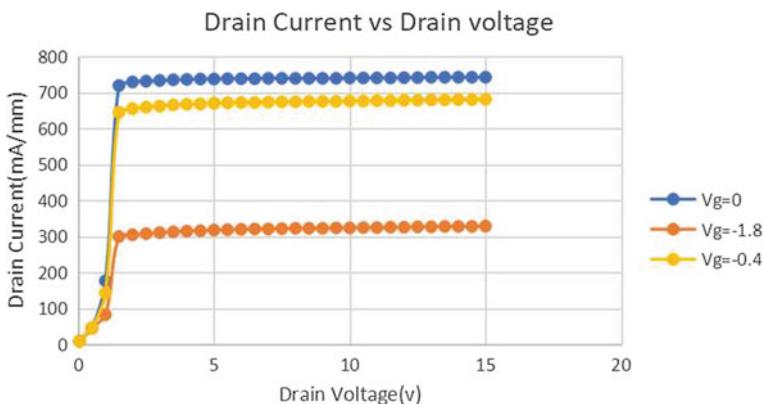


Fig. 2 I_d versus V_{ds} curve for different V_{gs} values

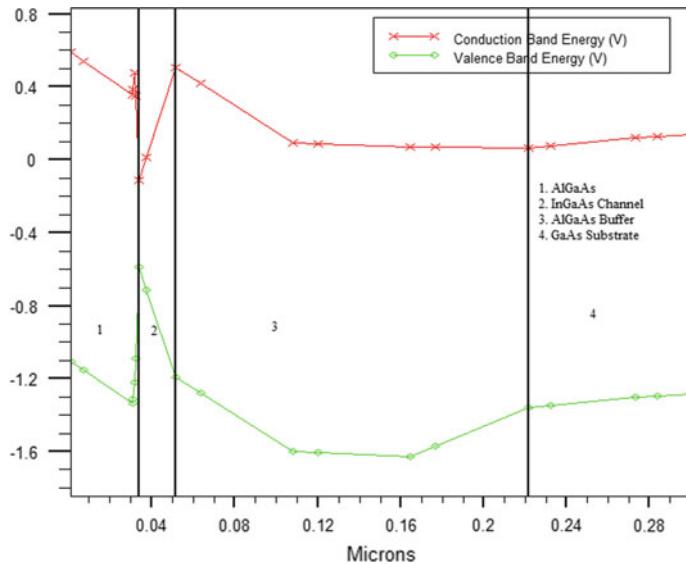


Fig. 3 Valence and conduction band diagram

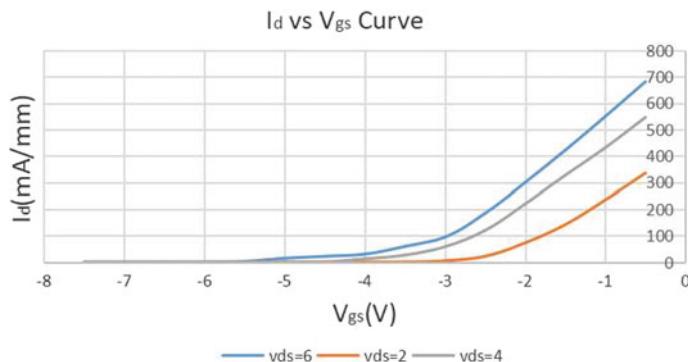


Fig. 4 I_d versus V_{gs} curve for different V_{ds} values

of Silicon δ-doping layer Fig. 5 shows the transconductance (g_m) versus V_{gs} curve for V_{ds} values = 6, 4, and 2 V. The maximum transconductance obtained is nearly 591 mS/mm for $V_{ds} = 6$ V. The graphs showing the cut-off frequency and maximum frequency are represented by Figs. 6 and 7, respectively. By analyzing the graphs, it is seen that the achievable maximum frequency and cut-off frequency of proposed HEMT are 87.5 GHz and 22.5 GHz, respectively. Table 3 shows a comparison between the results obtained in this paper with [10] paper.

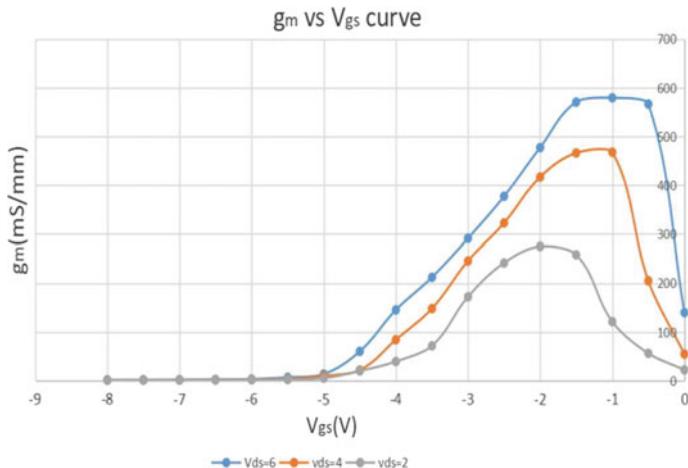


Fig. 5 g_m versus V_{gs} curve with varied V_{ds} values

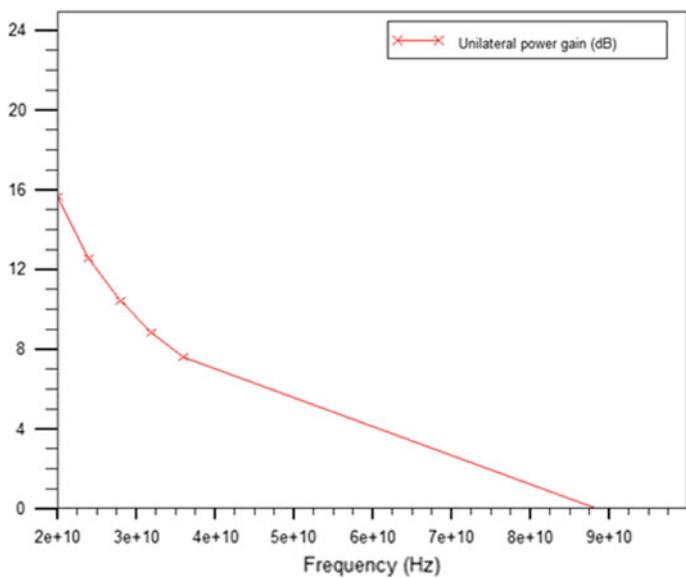


Fig. 6 Frequency curve of HEMT with $f_{\max} = 87.5$ GHz

5 Conclusion

We thus designed a AlGaAs/InGaAs/GaAs-based PHEMT structure using the optimization techniques. For evaluating the proposed HEMT's RF & DC performance,

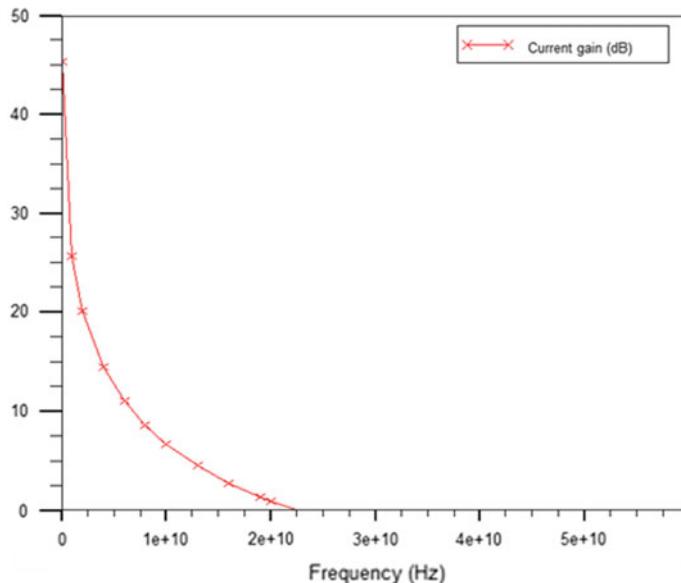


Fig. 7 Frequency curve of HEMT with $f_t = 22.5$ GHz

Table 3 Comparision of result

Papers	Drain current (mA/mm)	Transconductance (mS/mm)	F_{\max} (GHz)
This paper	720	591	87.5
Paper [10]	680	485	84

observation and analysis of its characteristics were performed. The simulated structure exhibited good output current characteristics and high transconductance of nearly 591 mS/mm. This improvement is obtained with increase in the width of the channel layer and reduction in the distance between gate and channel. Improvement in drain current is achieved with increase of doping concentration of the delta doping layer and supply layer.

High frequency efficiency of a device is evaluated by its maximum frequency of oscillation (f_{\max}) and cut-off frequency (f_t). Frequency performance can be enhanced with increase in electron velocity and transconductance. The RF performance of HEMT is affected by the transconductance. f_{\max} and f_t are directly proportional to the transconductance. Transconductance in turn is directly proportional to drain current. Due to inclusion of delta doping & InGaAs as channel layer the 2-DEG is enhanced, and hence, high electron mobility and high electron velocity are achieved resulting in higher drain current and ultimately high frequency.

References

1. Y. Chen, C. Chen, C. Chiong, H. Wang, A compact 40-GHz Doherty power amplifier with 21% PAE at 6-dB power back off in 0.1- μm GaAs pHEMT process. *IEEE Microw. Wirel. Compon. Lett.* **29**(8), 545–547 (2019)
2. M. Meghdadi, A. Medi, Design of 6–18-GHz high-power amplifier in GaAs pHEMT technology. *IEEE Trans. Microw. Theory Tech.* **65**(7), 2353–2360 (2017)
3. D.P. Nguyen B.L. Pham, T. Pham, A. Pham, A 14–31 GHz 1.25 dB NF enhancement mode GaAs pHEMT low noise amplifier, in *IEEE MTT-S International Microwave Symposium (IMS)* (2017), p. 1961–1964
4. Y. Chen Y. Wang, C. Chiong, H. Wang, An ultra-broadband low noise amplifier in GaAs 0.1- μm pHEMT process for radio astronomy application, in *IEEE International Symposium on Radio-Frequency Integration Technology (RFIT)* (2017), p. 80–82
5. B.R. Bennett, T.F. Chick et al., Strained InGaAs/InAlAs quantum wells for complementary III–V transistors. *J. Cryst. Growth* **388**, 92–97 (2014)
6. Z. Yinghui, W. Xiantai, S. Yongbo, C. Yuxiong et al., An 88 nm gate-length $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}/\text{In}_{0.52}\text{Al}_{0.48}\text{As}$ InP-based HEMT with f_{\max} of 201 GHz. *J. Semicond.* **33**, 0740041–07400414 (2012)
7. N. Waldron, H.D. Kim, J.A. Alamo, A self-aligned InGaAs HEMT architecture for logic applications. *IEEE Trans. Electron Devices* **57**, 297–304 (2010)
8. H. Morkoç, L.C. Witkowski, T.J. Drummond, C.M. Stanchak, A.Y. Cho, B.G. Streetman, Growth conditions to achieve mobility enhancement in $\text{Al}_x\text{Ga}_{1-x}\text{As}-\text{GaAs}$ heterojunctions by m.b.e. *Electron. Lett.* **16**(19), 753–754 (1980)
9. M. Mohapatra, N. Shukla, A.K. Panda, Delay analysis of ultra high speed InAlAs/InGaAs high electron mobility transistor. *Int. J. Nano Biomater.* (2014)
10. S.-J. Cho, C. Wang, R.K. Maharjan, N.-Y. Kim, An optimized 0.1 μm T-gate AlGaAs/InGaAs/GaAs PHEMT power amplifier MMIC for Ka-band applications, in *Proceedings of APMC 2012*, Kaohsiung, Taiwan. (2012)

Signal Propagation Modelling on the Inner Surface of Human Arm



Deepak Kumar Rout^{ID}, Asha Rani Mohapatra, Rishu Bhardwaj,
Raunak Kumar, and Vedant Raj Singh

Abstract One of the newer and most promising ways of wireless patient monitoring is by using wearable sensors. Owing to this there has been a gradual increase in the use of electrical invasive/non-invasive devices which have empowered the development of wireless body area networks. Wireless body area network (WBAN) helps in continuous health monitoring of patients in their on the go avoiding hospital visits, except in alarming health conditions. A wearable WBAN provides communication links among sensors on the body surface which is connected with a health monitoring system without hampering the daily life of the patient. Some innovative technologies supporting WBAN have proved their efficacy in some applications like biofeedback and remote health monitoring. The signal propagation on the inner surface of human arm is studied in this paper. Pathloss and the propagation characteristics play crucial role in data transmission in the wireless channel. A pathloss-based propagation model is discussed and a channel model is developed by evaluating the received power at various distances from the transmitter. The mean and median pathloss are estimated to study the behaviour of signal propagation.

Keywords Body area network · Channel modelling · Signal propagation

1 Introduction

Comprehensive development of wireless communication is going on with the different generations of wireless networks. One of the major domains in the application of wireless communication to novel use is the wireless body area network (WBAN) for patient monitoring which is essential for health and disease management. The present times of pandemic and emphasized the necessities of such monitoring manifold. In WBAN, there are sensors on or inside the body which sense crucial health data which may be in the form of electrocardiogram (ECG), electroencephalogram (EEG), blood oxygen saturation, blood pressure, etc., so that disease

D. K. Rout (✉) · A. R. Mohapatra · R. Bhardwaj · R. Kumar · V. R. Singh
School of Electronics, KIIT University, Bhubaneswar, India
e-mail: deepak09@live.com

or abnormality may be detected in an early stage [1–3]. The information gathered by the sensors is transmitted to a medical monitoring centre (MMC) where the patient's health conditions and whereabouts are tracked and traced by the algorithms implemented on the servers. In the COVID-19 situation, such monitoring could be highly useful to track the deteriorating health of people in quarantine and health workers in the frontline of duty.

However, there are a few challenges that need to be addressed while implementing WBAN-based remote health monitoring [2]. Modelling the signal propagation is one of the major challenges in WBAN. The absorption and shadowing effects of the human body tissues make the signal propagation highly challenging in the periphery of the human body. Pathloss measurements can be used to understand signal propagation in and around the human body. Analysis of the signal propagation will help in the design of transceivers which will enable the WBAN communications. Thus, the channel modelling is the starting point for effective and body-centric communications.

In the succeeding sections, the article discusses an on-body propagation scenario along the length of a person's arm to model the on-body path loss in the 900 MHz band, and models are derived. The parameters obtained from the results of the corresponding measurements are summarized for the channel models.

2 Related Work

Channel models for BAN consider the effect of the human body on the radio wave propagation and antenna characteristics. Several studies on BAN have been performed aimed at developing channel models for in-body, on-body, and off-body scenarios. Channel modelling plays a crucial role in the development of any wireless communication system. Smith et al. performed experimental setup and presented that a novel long-term channel predictor is suited to the on-body area communications channels [4]. The main characteristics of a WBAN and description of the major issues in the WBAN design, namely the radio channel, the power consumption and the coexistence with other RF-based systems are presented by Cavallari et al. [5]. In [6, 7], the authors present the wideband channel models which are more complex than the narrowband channel models. In [4, 5], statistical analysis or modelling of on-body channel at 2.45 GHz is performed. In [4], measurements of signal propagation on real human bodies are performed with different settings. In [5], the authors measured path loss on-body channels in different settings and the result is compared with anechoic environment. Channel modelling in a ferryboat environment is explored for an off-body narrowband and ultra-wideband channel model for body area networks with statistical analysis of the system loss model parameters [8]. A physical-statistical based channel model for off-body wireless communications is presented by Cwalina et al. [9]. The authors developed a model using first and second-order statistics utilizing 2.36 GHz measurement data for two different propagation environments. The Institute of Electric and Electronics Engineers (IEEE), in

its IEEE802.15.6 suggests operations in the 400, 600 and 900 MHz bands. An article by Januszkiewicz et al. describes two simplified models of human body for WBANs, however, most of their studies are for impedance matching of wearable antennas and focused at electromagnetic simulations [10]. Pathloss is measured in the 900 MHz and 2.4 GHz bands of body area networks in the CM3 channel model by Sharma et al. [11]. The study, however, does not formulate any channel model. The modelling of the measured pathloss data is important to know the statistical parameters related to the signal propagation in the medium.

A study of the research literature in the area, presented above, investigates the nature of signal propagation in the periphery of the human body. However, none of authors explored the development of a proper channel model in 900 MHz band. The few articles which have addressed the 900 MHz band have not developed a channel model that considers the statistical nature of the signal propagation. Thus, it is essential that we investigate the pathloss characteristics in this band in order to know the impact of signal propagation in the inner surface of human arm.

In this paper, we have developed a statistical path loss model by studying the propagation through the inner surface of human arm in the 900 MHz band.

3 Channel Modelling Setup

The channel modelling has a significant role in emulating the real-life scenarios prevalent in the periphery of the human body. The setup consists of a generic signal generator capable of generating frequency up to 4.8 GHz. A generic spectrum analyser capable of operating in the same band was used as the receiver. The transmitter and receiver were tuned to a frequency of 900 MHz. Miniature transmitting and receiving antennas with dimension of 46 mm × 6 mm with gain of 2 dB and VSWR of less than 2 mm were used in the transmitter and receiver respectively. These specifications of the antennas make them suitable to be mounted safely on the human body without hindering the activities. The transmitter is kept at inner side of the wrist. The receiving antenna is movable and is shifted away from the transmitter for each measurement. In each step, 10 mm of separation was increased. The initial T-R separation was 10 mm and it is successively increased. As discussed prior, the antennas were positioned on the human arm, and the transmitting and receiving antenna were connected to the signal generator and spectrum analyser, respectively. The received power measurements were recorded. Figure 1 shows the transmit and receive antennas mounted on the human arm. The subject is a 36 year male of height 172 cm and weight 71 kg, average build Asian.

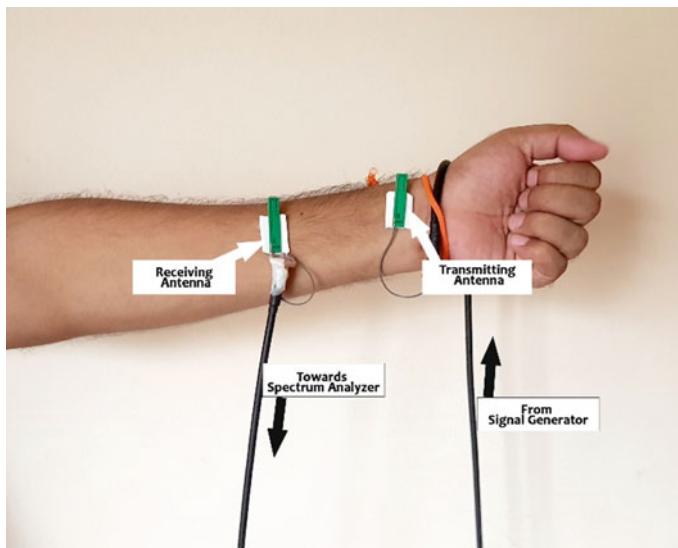


Fig. 1 Channel modelling setup consisting of the transmit and receive antennas

4 Result and Observations

The received power is observed by checking the signal power level in the spectrum analyser obtained from the receiving antenna. The transmitter was placed on the human wrist and the receiving antenna was moved in the steps of 10 mm each time. The receiving antenna was varied up to the distance of 120 mm. After detailed recording of received power up to 120 mm, the deviation in pathloss decreased. In every 10 mm, the deviation was negligible, thus, the next measurements were conducted at 150, 200 and 250 mm. To keep the transmitter power within safe limits, the maximum distance of varying was limited to 250 mm. Therefore, the communication on the human arm at 900 MHz is limited to a maximum transmitter-receiver separation of 250 mm. For each transmitter-receiver (T-R) separation, 30 received power measurements were observed. A total of more than 500 received power measurements were taken. The measurements obtained were tabulated and the various statistical parameters were studied. The pathloss was calculated, along with average pathloss and median pathloss for each T-R separation and further fitted by curve fitting tool of MATLAB[®] to develop the channel model.

The path loss and received power (in dB) were calculated for various distances on the human arm in 10–250 mm range. Figure 1 shows the collected path loss in dB for the various T-R separations. A total of 30 different paths are considered for the graph. It is easily found from the graph that the signal propagation on the human body surface is highly attenuating, an average 2.9 dB change in attenuation is observed in the mean path loss for each 10 mm change in the displacement. In case of median path loss, the average change is 2.93 dB per change of 10 mm. Similarly, the mean

Table 1 Mean and median path loss for transmitter-receiver separation of 10–250 mm in the 900 MHz band

Distance (mm)	Mean pathloss (dB)	Median pathloss (dB)
0	36.72	36.31
20	40.57	39.86
30	44.16	43.20
40	46.56	45.59
50	49.79	49.60
60	52.39	52.27
70	58.52	57.83
80	61.91	61.46
90	63.22	62.61
100	66.33	67.86
110	68.26	68.15
120	68.26	68.02
150	71.53	71.49
200	72.00	72.06
250	71.65	71.87

path loss varies by an average of 6.986 dB for every change in distance of 50 mm, this figure for median path loss is 7.112 dB. The signal attenuation in the human body scenario is found to be highly attenuating. The attenuating nature of the human body can be attributed to the absorption by the human body tissues.

The mean and median path loss for the 15 different locations on the human arm for the T-R separation of 10–250 mm with increments of 10 mm have been shown in Table 1. The mean path loss ranges between 37.55 and 72.00 while the median path loss varies between 37.55 and 72.06 dB. As the T-R separation increases an obvious increase in the path loss is observed.

It is observed from the received power measurements as shown in Figs. 2 and 3 that the path loss in the 900 MHz band ranges from 37.55 to 72.00 dB across the entire range. It is also detected from the data that while the pathloss increases as the separation between the transmitter and receiver increases, the increase in non-uniform, which can be attributed to the effect of the shadowing. Shadowing results from the absorption in the human body tissues.

A path loss model can be formulated by studying the pathloss data measured from the measurement campaign. On applying the curve fitting techniques to the data, it is found that the expression shown below best fits the mean and median path loss. The expression in Eq. (1) corroborates with the fitting models of CM3 channel models [12]. Thus, the path loss in the inner surface of the human arm can be expressed in Eq. (1) as

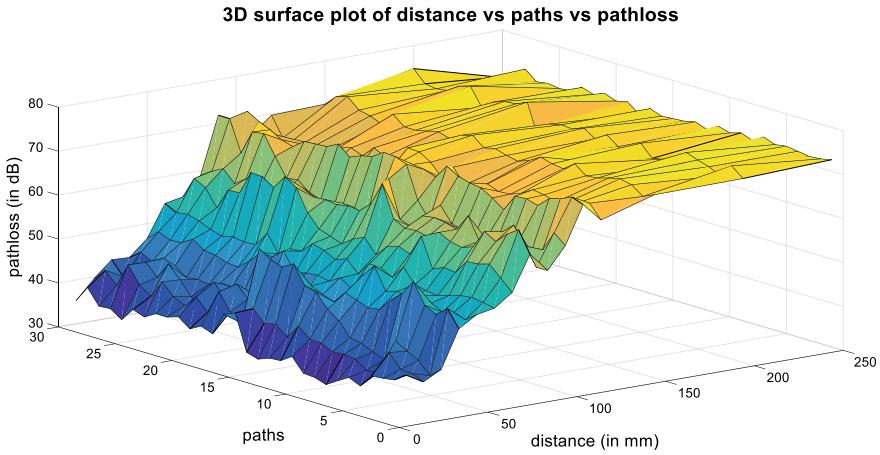


Fig. 2 3D plot between distance, paths and pathloss

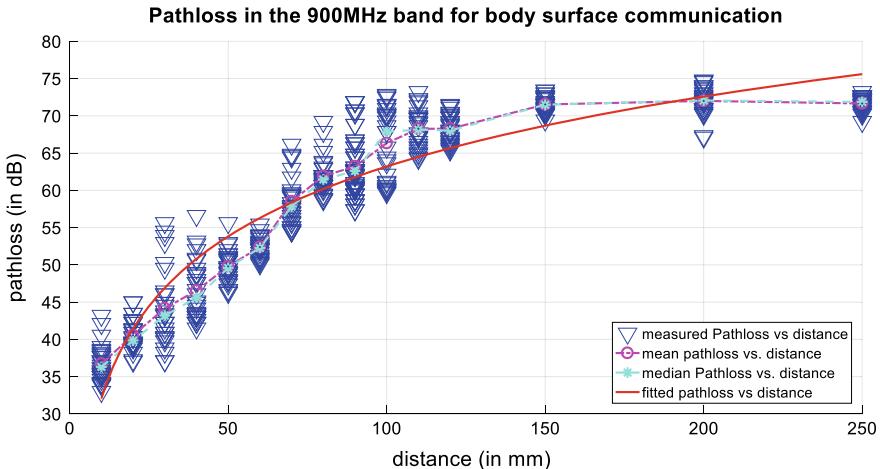


Fig. 3 Path loss measurement samples along with their average and fitted plots for human arm WBAN

$$\text{PL}(d)(\text{in dB}) = a \times \log_{10}(d) + b + N \quad (1)$$

where a and b are the parameters of linear curve fitting, N is a random variable that emulates shadowing and is conveyed by its standard deviation σ_N . The values of a , b and σ_N are evaluated by analysing the fitting parameters as 31.15, -34.7 and 5.9666, respectively.

The measured path loss, mean path loss, median pathloss and fitted path loss at different T-R separations are shown in Fig. 3.

5 Conclusion

In this article, we study the signal propagation in the vicinity of human body. The path loss on the inside of the human arm is modelled as a path loss model. Received power collected from a measurement campaign has been plotted, analysed and modelled. The statistical parameters associated with the data such as median and mean have been evaluated and presented. The attenuating nature of the human body channel is explored and a path loss model is developed. The work will be useful for researchers developing transceivers operating the said frequency band of 900 MHz. The future work in this area will involve the development of channel models in other body positions and other frequency bands.

References

1. M.M. Nezhad, M. Eshghi, Sensor single and multiple anomaly detection in wireless sensor networks for healthcare, in *2019 27th Iranian Conference on Electrical Engineering (ICEE)*, p. 1751–1755 (2019). <https://doi.org/10.1109/IranianCEE.2019.8786588>
2. K. Hasan, K. Biswas, K. Ahmed, N.S. Nafi, M.S. Islam, A comprehensive review of wireless body area network. *J. Netw. Comput. Appl.* **143**, 178–198 (2019). <https://doi.org/10.1016/j.jnca.2019.06.016>
3. M.R. Yuce, Implementation of wireless body area networks for healthcare systems. *Sens. Actuators, A Phys.* **162**, 116–129 (2010). <https://doi.org/10.1016/j.sna.2010.06.004>
4. D.B. Smith, L.W. Hanlen, D. Miniutti, Transmit power control for wireless body area networks using novel channel prediction, in *2012 IEEE Wireless Communications and Networking Conference (WCNC)*, p. 684–688 (2012). <https://doi.org/10.1109/WCNC.2012.6214456>
5. R. Cavallari, S. Member, F. Martelli, R. Rosini, S. Member, C. Buratti, R. Verdone, A survey on wireless body area networks: technologies and design challenges **16**, 1635–1657 (2014)
6. J. Dong, D. Smith, Joint relay selection and transmit power control for wireless body area networks coexistence, in *2014 IEEE International Conference on Communications*, p. 5676–5681 (2014). <https://doi.org/10.1109/ICC.2014.6884226>
7. J.M. Choi, H. Kang, Y. Choi, A study on the wireless body area network applications and channel models, in *2008 Second International Conference on Future Generation Communication and Networking*, p. 263–266 (2008). <https://doi.org/10.1109/FGCN.2008.216>
8. K.K. Cwalina, S.J. Ambroziak, P. Rajchowski, An off-body narrowband and ultra-wide band channel model for body area networks in a ferryboat environment. *Appl. Sci.* **8**, 988 (2018)
9. M. Mohamed, M. Cheffena, A. Moldsvor, F.P. Fontan, Physical-Statistical Channel Model for Off-Body Area Network. *IEEE Antennas Wirel. Propag. Lett.* **16**, 1516–1519 (2017). <https://doi.org/10.1109/LAWP.2016.2647323>
10. Ł. Januszkiwicz, S. Hausman, P.D. Barba, Human body modelling for wireless body area network optimization, in: *2020 14th European Conference on Antennas and Propagation (EuCAP)*, p. 1–5 (2020). <https://doi.org/10.23919/EuCAP48036.2020.9136084>
11. P.K. Sharma, T.V. Sai, Human Body Modelling for Wireless Body Area Network Optimization, in *International Conference on Intelligent Computing and Smart Communication*, p. 565–573 (2019)
12. IEEE Standards Association: *IEEE Standard for Local and Metropolitan Area Networks—Part 15.6: Wireless Body Area Networks* (2012). <https://doi.org/10.1109/IEEESTD.2012.6161600>

Detection of Brain Tumor from MR Images Using BWT and SOM-SVM with Authentication



Nilesh Bhaskarrao Bahadure, Sidheswar Routray, and Arun Kumar Ray

Abstract This paper does the qualitative analysis of the watermarking technique to provide an improved healthcare system. It addresses two main problems of brain tumor detection: one is to find the region of interest, i.e., to detect brain tumor from the medical images and second is to authenticate patient information on the medical images. The system combines a region-based segmentation using self-organizing map algorithm and secure authentication using a watermarking technique based on the Berkeley wavelet transformation (BWT) algorithm. In recent times, a significant amount of interest has been developed in encouraging radiologists or clinical experts to use the technology-based system for health care. The proposed approach combines BWT for watermarking and self-organizing map and support vector machine algorithms for region-based segmentation to improve the objectives of better treatment. The experimental result of the proposed method is evaluated for performance and quality analysis on magnetic resonance imaging (MRI) images, resulted well in terms of quality parameters and imperceptibility.

Keywords Berkeley wavelet transformation (BWT) · Self-organizing map (SOM) · Support vector machine (SVM) · Magnetic resonance imaging (MRI) · Watermarking

N. B. Bahadure

Department of Electronics Engineering, School of Technology, Sanjay Ghodawat University, Kolhapur, India

S. Routray (✉)

School of Computer Science and Communication Engineering, Jiangsu University, Zhenjiang, People's Republic of China
e-mail: sidheswar69@gmail.com

A. K. Ray

School of Electronics Engineering, Kalinga Institute of Industrial Technology (KIIT), Bhubaneswar, India

1 Introduction

In this research work, we address two main problems: one is to segment the brain tumor from MR images for detection of brain tumor and other is to authenticate patient's information using watermarking. Detection of tumor from the brain MRI or from other modern imaging modalities plays an important role in deciding right therapy at the right time. Similarly, authenticating patient's data or information in the medical images is also equally important for protecting patient's identity. Providing healthcare solutions to the remote areas is always a big challenge. If we can able to authenticate the services without disclosing the patient's identity and transmit the information on low bandwidth by watermarking, then it is possible to provide healthcare services to the remote areas. Brain tumor detection is an important, challenging and interesting concept, in particular, when it involves the extraction of information from MR images, computed tomography (CT) images or from other modern medical imaging modalities [1]. Brain tumor segmentation comprises of separation of the tumor tissues from the processed image; the first part of image contains the normal brain cells like edema, necrosis, gray matter, white matter, and the cerebrospinal fluid (CSF), whereas the second part contains the tumor cells of the brain [2]. A separate and specific patient care is done for the patients with brain tumor, through the evaluation of advance medical imaging [3]. In recent years, many imaging modalities, such as X-rays, ultrasonography, magnetoencephalography, computed tomography, and MRI, have been developed to show details of brain tumor and thus, help doctors to study brain tumor with an aim of providing improved treatment to patients [4].

The SOM-based technique is popular for region-based segmentation similar to fuzzy clustering-based technique [5]. For better healthcare system with full authentication on patient's information, watermarking scheme is needed to embed on medical image modalities. Viswanathan et al. [6] proposed a joint FED watermarking framework using spatial fusion and addresses two problems such as authentication of patient information and fingerprint features to maintain multiple documents. Kamran et al. [7] investigated information preserving watermarking scheme, which claims to be a right blend of watermarking technique to authenticate patient information by means of e-healthcare technology, whereas Walia et al. [8] presented a watermarking technique for image authentication based on Weber's law. Vellaisamy et al. [9] proposed a new technique for watermarking of patient information in medical images based on inversion attack which preserves the fidelity of the host image and works on different medical imaging modalities. Chen et al. [10] presented image watermarking scheme based on reversible image watermarking which is very sensitive to image quality and thus, very useful in military and medical imaging authentication. Kabayashi et al. [11] presented a novel approach for providing integrity and information authenticity in DICOM images using cryptographic approach, which claims to be more trustworthy of medical images and providing better integrity and authenticity without compromising image quality. To verify the reliability of region of interest (ROI) extracted from the medical images and to detect tampered block

if any inside ROI, Eswaraiah et al. [12] investigated robust watermarking technique based on integer wavelet transform (IWT).

It is observed from above discussion that some of the methods are proposed to obtain segmentation only and some of the methods are proposed to obtain classification only, and some of the methods are proposed only for watermarking [13]. Therefore, feature selection and patient's information authentication have not been addressed in previous published research articles. Here, we combine BWT and SOM to improve diagnostic accuracy to classify benign from malignant tumor slices [13]. The objective is to perform feature extraction from the segmented tumor region followed by classification of healthy and infected-tumor tissues and to authenticate patient's information using watermarking. The segmentation of the tumor region can be accomplished by using an SOM-SVM technique and watermarking by using a BWT [14].

2 Development of Proposed Algorithm

Algorithm1: SOM based segmentation

1. Select an input image
2. Define white pixels (M)
3. Select white pixels ($w_k(0); 1 \leq k \leq M$)
4. Set iteration count $t = 0$
5. if not converged
6. **while** loop **do**
7. Select the next p (p is pixel value of white pixels)
8. Calculate weight of the white pixels $d(p, w_k(t))$
9. Select $M^* = \min_p d(p, w_k(t))$
10. Find $w_k(t+1) = \begin{cases} w_k(t) + \alpha(p - w_k(t)) & (k \in N) \\ w_k(t) & (k \notin N) \end{cases}$
11. Update M^* and its neighborhood
12. if not converge
13. then $t=t+1$
14. **end while**

Figure 1 shows the proposed algorithm based on segmentation using SOM-SVM and watermarking based on Berkeley wavelet transformation [14, 15]. The host images are MR images, which are first processed and converted to grayscale images and then, segmented using SOM-SVM algorithm. The edges of the segmented image are preserved using edge detection and morphology operation. Once the segmentation operation is completed, then it reads the watermark image. It consists of the patient information and embeds this image on the segmented MR image. If the message length is greater than the maximum permitted message, then read the

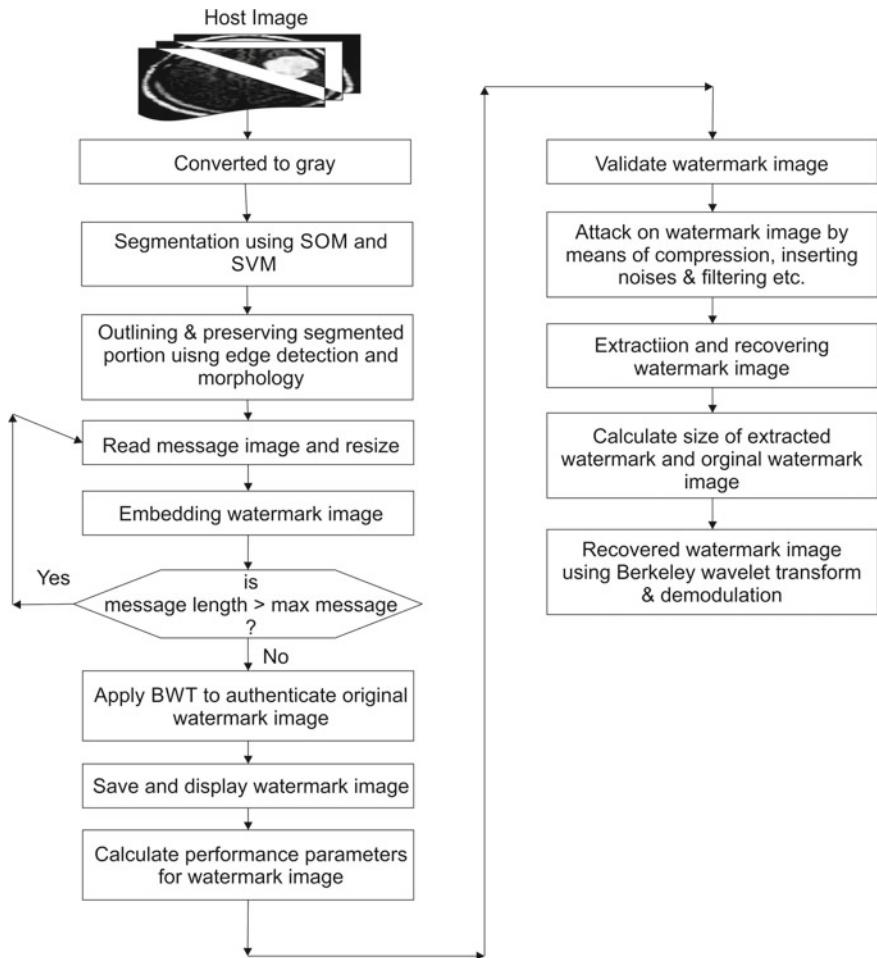


Fig. 1 Proposed methodology

message image again with resizing the message image. To authenticate the information on MR images, apply Berkeley wavelet transformation, and calculate the performance parameters for further evaluation and confirms that no attack has been done on the processed watermark image. Once the validation of the watermark image is completed, then extract and recover the original watermark image. The proposed algorithm uses SOM for segmentation and SVM for the classification of brain tumor from MR images [14]. The deductive segments are automatically cluster by SOM without knowing the characteristics of input data. The process used to implement self-organizing map on selecting the segments is described in Algorithm 1.

3 Results and Discussion

In order to validate the performance of the proposed detection algorithm, we use two benchmark datasets such as DICOM dataset [16] and Brain Web dataset [17]. Figure 2 shows the different stages of brain tumor detection and extraction. Key point shown in Fig. 2 helps the radiologists to identify the key point on the tumor and tumor looks area from the infected images. Table 1 shows the tumor identities and information obtained for the images shown in Fig. 2. The valuable information is extracted from the tumor infected images such as number of tumors, position of largest tumor, and approximated diameter of the largest tumor in square cm. These identities are helped to understand and differentiate between the images on benign and malignant characteristics. The comparison of our approach with existing approaches

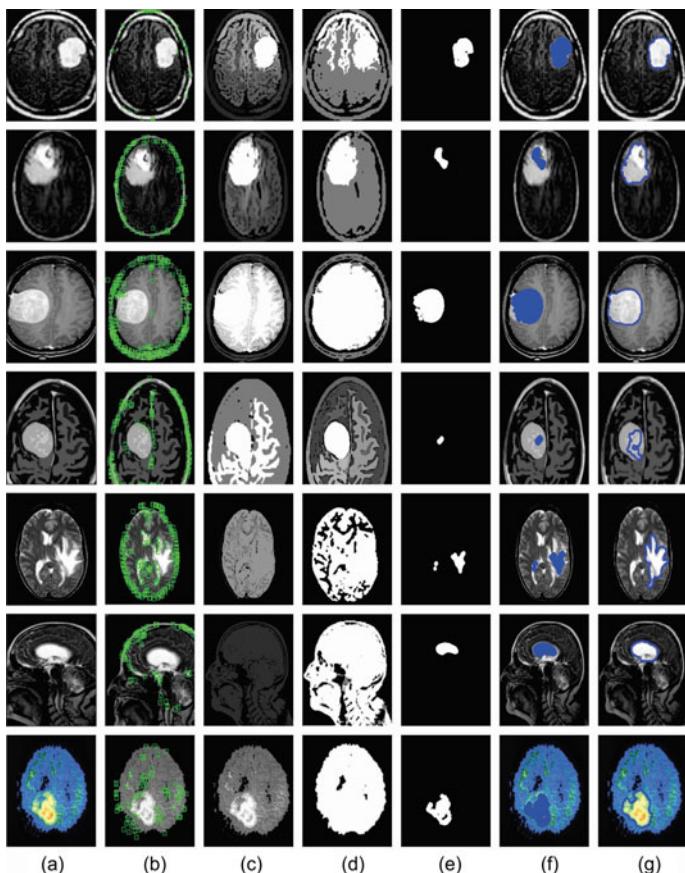


Fig. 2 Result of brain tumor detection and extraction: **a** original image, **b** finding key points, **c** reconstructed image, **d** threshold image, **e** tumor extracted image, **f** segmentation marking, **g** tumor identification

Table 1 Tumor identities and information from Fig. 2

Images	No. of tumors	Position of largest tumor	Diameter of largest tumor
Image-1	1	Right temporooccipital lobe	3.62
Image-2	1	Left frontal lobe	2.43
Image-3	1	Right temporooccipital lobe	4.33
Image-4	1	Left temporooccipital lobe	1.06
Image-5	2	Right temporooccipital lobe	2.49
Image-6	1	Right frontal lobe	2.71
Image-7	1	Left temporooccipital lobe	3.75

is shown in shown in Table 2 in terms of TP, TN, FP, FN, accuracy, sensitivity, and precision.

Table 2 Comparison and measure of overlap for segmented tissues

Datasets	Techniques	TP	TN	FP	FN	Accuracy	Sensitivity	Precision
1	Proposed method	94.7	0	0	5.3	94.7	94.7	100
	KIFCM [1]	90.5	0	0	9.5	90.5	100	90.5
	IFWSVM [18]	392	NA	346	42	NA	90.95	NA
	DWT [19]	NA	NA	NA	NA	94.8	100	100
2	Proposed method	99.9	0	0	0.1	99.9	99.9	100
	KIFCM [1]	100	0	0	0	100	100	100
	IFWSVM [18]	679	NA	1523	28	NA	96.03	NA
3	Proposed method	99.8	0	0	0.2	99.8	99.8	100
	KIFCM [1]	100	0	0	0	100	100	100
	IFWSVM [18]	377	NA	2050	41	NA	90.19	NA

4 Conclusion

In this study, we addresses two problems, one is to detect the brain tumor and second is to watermark the patient information on the medical images. The proposed technique uses Berkeley wavelet transformation for watermarking. It can be used to detect region of interest and to authenticate watermarked image. It ensures secure authentication with the help of attacks from compression, inserting noise, averaging images, etc. It can authenticate the watermarked image even if it is compressed by 10–15% of compression ratio and extreme low visibility due to high noise. In future, we examine watermarking medical images for authentication as well as for automatic report generation technique.

References

1. E. Abdel-Maksoud, M. Elmogy, R. Al-Awadi, Brain tumor segmentation based on a hybrid clustering technique. *Egypt. Inf. J.* **16**, 71–81 (2015)
2. M. Ahmed, S. Yamani, N. Mohamed, A. Farag, T. Moriarty, A modified fuzzy c-means algorithm for bias field estimation and segmentation of MRI data. *IEEE Trans. Med. Imaging* **21**(3), 193–199 (2002)
3. I.M. Alwan, E.M. Jamel, Digital image watermarking using Arnold scrambling and Berkley wavelet transform. *Al-Khwarizmi Eng. J.* **12**(2), 124–133 (2015)
4. S. Routray, A.K. Ray, C. Mishra, MRI Denoising using sparse based Curvelet transform with variance stabilizing transformation framework. *Indonesian J. Electr. Eng. Comput. Sci.* **7**(1), 116–122 (2017)
5. C. Benson, V. Lajish, Morphology based enhancement and skull stripping of mri brain images, in *IEEE International Conference on Intelligent Computing Applications*, Coimbatore, Tamil Nadu, India, pp. 254–257 (2014)
6. P. Viswanathan, P.V. Krishna, A joint fed watermarking system using spatial fusion for verifying the security issues of teleradiology. *IEEE J. Biomed. Health Inf.* **18**(3), 753–764 (2014)
7. M. Kamran, M. Farooq, An information-preserving watermarking scheme for right protection of EMR systems. *IEEE Trans. Knowl. Data Eng.* **24**(11), 1950–1961 (2012)
8. E. Walia, A. Suneja, Fragile and blind watermarking technique based on webers law for medical image authentication. *IET Trans. Comput. Vis.* **7**(1), 9–19 (2012)
9. S. Vellaisamy, V. Ramesh, Inversion attack resilient zero-watermarking scheme for medical image authentication. *IET Trans. Image Process.* **8**(12), 718–727 (2013)
10. M. Chen, Z. Chen, X. Zeng, Z. Xiong, Model order selection in reversible image watermarking. *IEEE J. Select. Topics Signal Process.* **4**(3), 592–604 (2010)
11. L.O.M. Kobayashi, S.S. Furuike, P.S.L.M. Barreto, Model order selection in reversible image watermarking. *IEEE Trans. Inf. Technol. Biomed.* **13**(4), 582–589 (2009)
12. R. Eswaraiah, E.S. Reddy, Robust medical image watermarking technique for accurate detection of tampers inside region of interest and recovering original region of interest. *IET Trans. Image Process.* **9**(8), 615–625 (2015)
13. A. Padma, R. Sukanesh, Segmentation and classification of brain CT images using combined wavelet statistical texture features. *Arab. J. Sci. Eng.* **39**, 767–776 (2014)
14. N.B. Bahadure, A.K. Ray, H.P. Thethi, Comparative approach of MRI-based brain tumor segmentation and classification using genetic algorithm. *J. Digit. Imaging* **31**, 477–489 (2018)
15. N.B. Bahadure, A.K. Ray, H.P. Thethi, Image analysis for MRI based brain tumor detection and feature extraction using biologically inspired BWT and SVM. *Int. J. Biomed. Imaging* **2017**, 1–12 (2017)

16. DICOM, DICOM samples image sets (2018)
17. Brainweb, Brain web: Simulated brain database (2018)
18. L. Guo, L. Zhao, Y. Wu, Y. Li, G. Xu, Q. Yan, Tumor detection in MR images using one-class immune feature weighted SVMs. *IEEE Trans. Magn.* **47**(10), 3849–3852 (2011)
19. G. Yang, T. Nawaz, T.R. Barrick, F.A. Howe, G. Slabaugh, Discrete wavelet transform-based whole-spectral and subspectral analysis for improved brain tumor clustering using single voxel mr spectroscopy. *IEEE Trans. Biomed. Eng.* **62**(12), 2860–2866 (2015)

Wavelet Transform for Signal Compression in Sparse Algorithms



Rosalin, Nirmal Kumar Rout, and Debi Prasad Das

Abstract In this paper, the proportionate normalized least mean square (PNLMS) and its variants, μ -law PNLMS (MPNLMS), improved PNLMS (IPNLMS), and filter PNLMS (FPNLMS) algorithms are simulated for different types of input signals for a sparse system identification problem. The input signals such as a Gaussian random signal, band-limited signal, a speech signal, a color signal, and a uniform random signal are analyzed in this paper. The input is compressed by a wavelet transform before application. The simulations are carried out to examine the behavior of Gaussian noise on a sparse physical system. The simulation further justifies the robustness of the FPNLMS algorithm to the change in the input signal and noise level.

Keywords Sparse · Compression · Gaussian · Wavelet

1 Introduction

The compression of data focuses on compressing the amount of information present in a signal with or without loss of data in it. The frequently used lossy compression method is the wavelet transform. Wavelet transform allocates the simultaneous location of time and frequency of the signal.

In real-world applications, the systems are encountered where signal transmission has a sparse impulse response [1] like echo canceller [2], wireless multipath channel [3], underwater communications [4], and HDTV [5]. Adaptive algorithms

Rosalin (✉) · N. K. Rout

School of Electronics Engineering, KIIT University, Bhubaneswar, India
e-mail: rosalin.kiit15@gmail.com

N. K. Rout
e-mail: routnirmal@rediffmail.com

D. P. Das
Department of Process Modeling and Instrumentation, CSIR-Institute of Minerals and Materials Technology, Bhubaneswar, India
e-mail: dpdasimmt@gmail.com

were developed for systems with sparse impulse response as least mean square (LMS) and normalized LMS (NLMS) algorithms failed for these systems [6]. The notable algorithms formulated are the proportionate NLMS (PNLMS) [7] and its variants, improved PNLMS (IPNLMS) [8], μ -law PNLMS (MPNLMS) [9], and filter PNLMS (FPNLMS) [10]. At the initial stage, the PNLMS attains a faster convergence speed, but the pace decreases as the filter adapts. The initial convergence rate of the IPNLMS is slower compared to PNLMS, which provided better performance for the medium sparse impulse response system. The MPNLMS maintains the convergence speed during the adaptation but with an increase in computational load. The FPNLMS provides a faster convergence rate compared to PNLMS, IPNLMS, and MPNLMS for sparse systems with a comparable computational load. PNLMS, IPNLMS, MPNLMS, and FPNLMS algorithms further provide a better performance after compressing the input signal which is Gaussian random in nature with mean zero and variance one [10].

The ever-growing amount of data present in a wireless network is the problem faced during signal transmission. So, the internet of things (IoT) systems and 5G are outlined to bear with a large network. The requirement of such extensive connectivity in the network leads to a large amount of data to be produced [11]. Therefore, signal compression is required before application. One such application is to identify the sparse physical system after compressing the input signal. In this paper, different types of input signals are analyzed such as a Gaussian random signal, band-limited signal, a speech signal, a color signal, and a uniform random signal. Here, the sparse system $h(n)$ is identified based on Fig. 1. In Fig. 1, $x(n)$ is the compressed input signal and $\eta(n)$ is the input Gaussian noise. The output of the sparse system $(n) = x(n) * h(n) + \eta(n)$. The error signal, $e(n) = y(n) - \hat{y}(n)$, where $\hat{y}(n)$ is the output of the adaptive filter, $\hat{h}(n)$. The study of the PNLMS algorithm and its modifications in this context is carried out in this paper. The rest of the paper is arranged as follows. Brief introduction to wavelets is described in Sect. 2. Section 3 provides the simulation for different SNR values for different types of input signals, and Sect. 4 presents the conclusion.

2 Data Compression by Using Wavelet Transform

Wavelets are mathematical operations that divide the data into frequency elements, and the elements are analyzed for a resolution matched to its scale [12]. The discrete wavelet transform (DWT) is used for operations on discrete data for non-stationary signals. In wavelet transform, a signal is represented as a set of wavelets or basis functions. This basis functions is called as a mother wavelet [15]. Numerous wavelets are being used now a days are Daubechies, bi-orthogonal, coiflets, symlets, and dmey. For a signal in time domain $x(t)$, the transform is

$$X(a, b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} \psi\left(\frac{t-b}{a}\right) x(t) dt \quad (1)$$

where a is the scaling parameter, b is time parameter and $\psi\left(\frac{t-b}{a}\right)$ is the mother wavelet function. The Daubechies wavelet is the base function to represent solutions of integral equations [12]. The Daubechies wavelets have two unique properties. First, there are a finite number of nonzero scaling coefficients resulting in db1, db2...db45 wavelets. The second property is that the first $M - 1$ moment of the wavelets is zero for order- M . In DWT a time-scale representation of a signal is obtained by digital filtering techniques. The DWT is computed by the Mallat algorithm [16]. The Mallat algorithm is a process of filtering and decimation which is continued until the desired value is reached.

3 Simulation Results

The simulations are carried out as per Fig. 1 for a system identification problem. The physical system is identified by an adaptive finite impulse response filter (FIR). The DWT is applied to the input signal for compressing it. Wavelet decomposition is done by db3 wavelet which is a third-level wavelet of Daubechies wavelets. The input signal after compression is passed through a sparse system having a transfer function $h(n)$. $h(n)$ is an exponentially decaying vector of dimension N_u [13].

$$h = [S^T U^T F]^T \quad (2)$$

$$S = \left[1, \exp\left(-\frac{1}{\tau}\right), \exp\left(-\frac{2}{\tau}\right), \dots, \exp\left(\frac{N - N_u}{\tau}\right) \right] \quad (3)$$

$\tau > 0$ is decay constant and N_u is the bulk delay of the number of active filter coefficients present. U , is a mean zero Gaussian sequence and F is a diagonal matrix

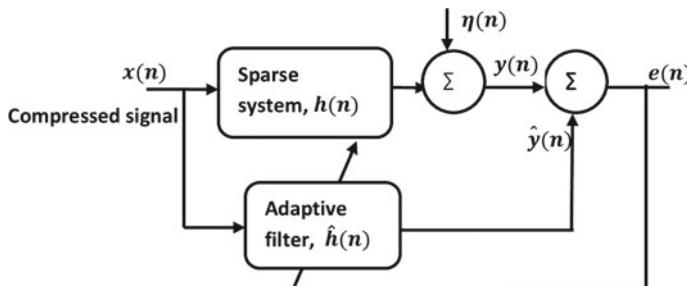


Fig. 1 Sparse System Identification block diagram

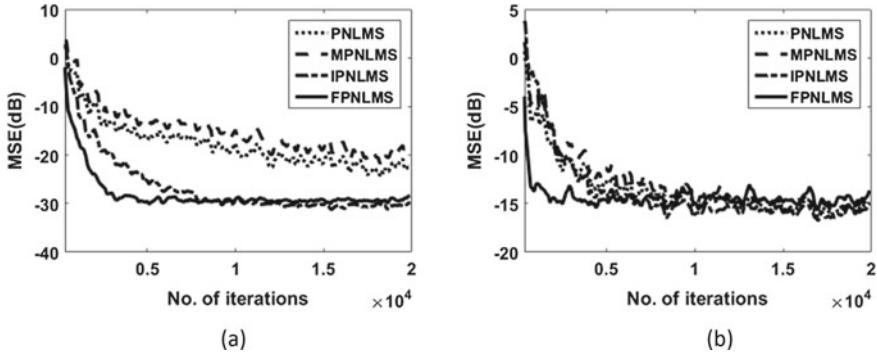


Fig. 2 Learning curve of a Gaussian signal at **a** $\eta_p = -5 \text{ dB}$, **b** $\eta_p = 10 \text{ dB}$

of dimension, $N_u \times N_u$ with elements from a Gaussian sequence with mean zero. N is the total number of filter coefficients present in the physical system. Simulations are carried out at $\tau = 8$ and $N_u = 16$ where $N = 512$ is used to generate a sparse impulse response in this paper. A white Gaussian noise, $\eta(n)$ of power, $\eta_p = -5 \text{ dB}$, 10 dB is added to the output of the physical system to get the desired signal. The simulations are carried out for 20,000 iterations with step size, $\mu = 0.9$. The compressed signal is power normalized before application.

3.1 Gaussian Random Input

DWT is applied to the input signal which is Gaussian random in nature with mean zero. As seen from Fig. 2, the FPNLMS algorithm shows better steady-state behavior with faster convergence time compared to PNLMS, IPNLMS, and MPNLMS algorithm. However, the FPNLMS algorithm provides the same steady-state MSE as the IPNLMS algorithm with an increase in noise.

3.2 Colored Input

In this experiment, DWT is applied to the colored signal, $x(n)$ which is generated by using the Auto-Regressive (AR) model to a Gaussian signal $x_1(n)$ [14].

$$x(n) = 1.79x_1(n-1) - 1.85x_1(n-2) + 1.27x_1(n-3) - 0.41x_1(n-4) + u(n) \quad (4)$$

$u(n)$ is a normally distributed noise. As seen from Fig. 3, the IPNLMS and FPNLMS algorithms show better steady-state performance compared to PNLMS

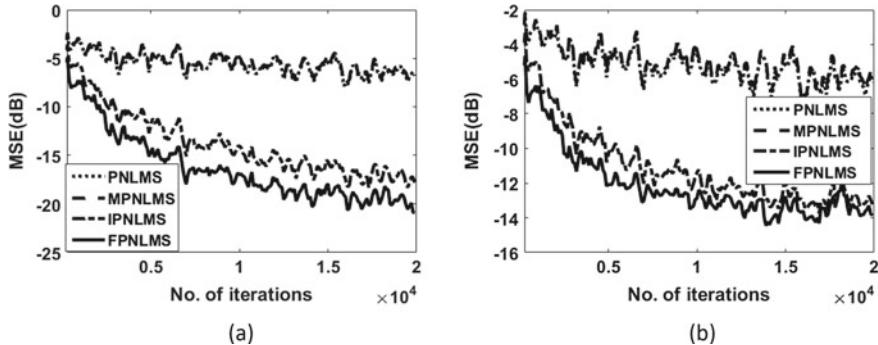


Fig. 3 Learning curve of a colored signal at **a** $\eta_p = -5 \text{ dB}$, **b** $\eta_p = 10 \text{ dB}$

and MPNLMS algorithms. Also, as noise increases, the IPNLMS and FPNLMS have the almost same performance. Lastly, FPNLMS attains lower MSE for both low and high SNR noise.

3.3 Speech Signal

A speech signal is recorded and given as input after compressing it using a DWT to the physical system. The simulation is done for 10,000 samples as it attains convergence very fast. As inferred from Fig. 4 all, the algorithms provide a fixed initial MSE. With the increase in noise, the steady-state error reduces for IPNLMS, MPNLMS, FPNLMS, and PNLMS. All algorithms attain convergence at the same time but FPNLMS provides the lowest MSE at all SNR values.

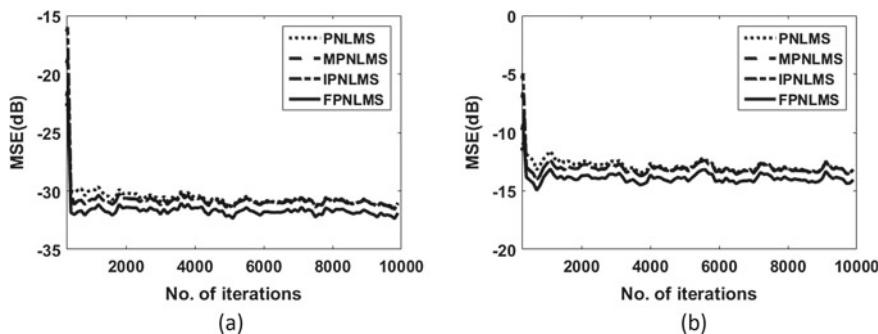


Fig. 4 Learning curve of a speech signal at **a** $\eta_p = -5 \text{ dB}$, **b** $\eta_p = 10 \text{ dB}$

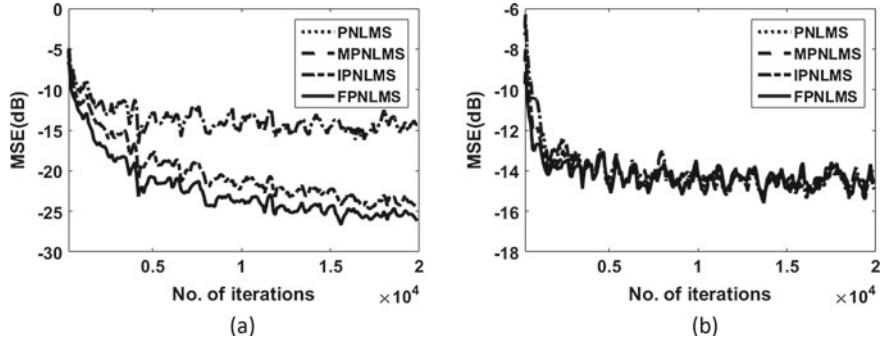


Fig. 5 Learning curve of a band-limited signal at **a** $\eta_p = -5$ dB, **b** $\eta_p = 10$ dB

3.4 Band-Limited Signal

The band-limited signal applied as input to the physical system for its identification undergoes compression by DWT. It is produced by transmitting a Gaussian random signal through a third-order Butterworth low-pass filter. As seen in Fig. 5, FPNLMS outperforms existing algorithms for low noise power. As noise increases, all algorithms attain a fixed steady-state MSE and convergence time.

3.5 Uniform Random Signal

A uniformly distributed random signal within a range of 0 to 1 is used as an input signal to the physical system. As inferred from Fig. 6, the FPNLMS algorithms steady-state MSE is low compared to PNLMS, MPNLMS, and IPNLMS algorithms as noise increases. PNLMS and MPNLMS algorithms do not converge for

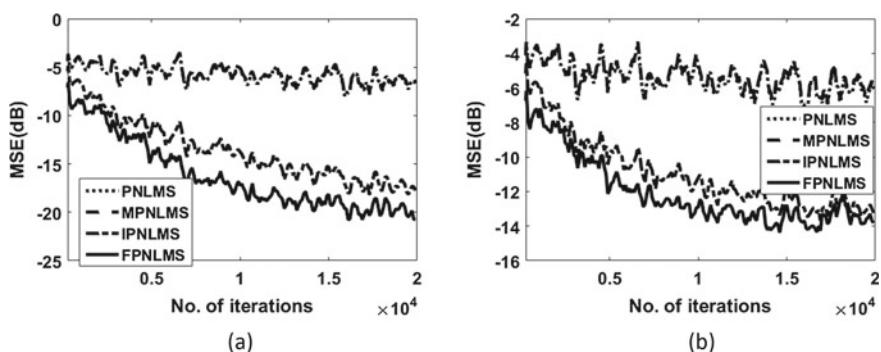


Fig. 6 Learning curve of a uniform random signal at **a** $\eta_p = -5$ dB, **b** $\eta_p = 10$ dB

a uniformly distributed random signal while IPNLMS and FPNLMS converge. The IPNLMS behaves similarly to the FPNLMS algorithm in terms of convergence time but gives a low MSE during adaptation.

4 Conclusion

Data compression before applying it to a sparse system that is affected by a Gaussian noise is identified by PNLMS, IPNLMS, MPNLMS, and FPNLMS algorithm in this paper. In low noise, all algorithms convergence curve and steady-state error show acceptable performance but the performance deteriorates as noise increases. Further, the FPNLMS algorithm is robust to change in signal and noise variation compared to PNLMS, MPNLMS, and IPNLMS algorithm.

References

1. J.W. Yoo, J.W. Shin, P.G. Park, An improved NLMS algorithm in sparse systems against noisy input signals. *IEEE Trans. Circ. Syst. II: Express Briefs* **62**(3) (2015)
2. J. Radecki, Z. Zilic, K. Radecka, Echo cancellation in IP networks, in *Proceedings of the 45th Midwest Symposium on Circuits and Systems*, vol. 2, pp. 219–222 (2002)
3. W. Bajwa, J. Haupt, G. Raz, R. Nowak, Compressed channel sensing, in *Proceedings of the IEEE CISS*, pp. 5–10 (2008)
4. B.K. Das, M. Chakraborty, Sparse adaptive filtering by an adaptive convex combination of the LMS and the ZA-LMS algorithms. *IEEE Trans. Circ. Syst. I: Regular Papers* **61**(5) (2014)
5. W. Schreiber, Advanced television systems for terrestrial broadcasting. *Proc. IEEE* **83**(6), 958–981 (1995)
6. B. Jelfs, D.P. Mandic, J. Benesty, A class of adaptively regularised PNLMS algorithms, in *Proceedings of the 15th International Conference on Digital Signal Processing*, Cardiff, UK, pp. 19–22 (2007)
7. D.L. Duttweiler, Proportionate normalized least mean square adaptation in echo cancellers. *IEEE Trans. Speech Audio Process.* **8**(5), 508–518 (2000)
8. J. Benesty, S.L. Gay, An improved PNLMS algorithm, in *Proceedings of the IEEE ICASSP*, Orlando, FL, USA, May 2002, vol. 2, pp. 1881–1884
9. H. Deng, M. Doroslovacki, Improving convergence of the PNLMS algorithm for sparse impulse response identification. *IEEE Signal Process. Lett.* **12**(3), 181–184 (2005)
10. Rosalin, N.K. Rout, D.P. Das, Filter proportionate normalized least mean square algorithm for a sparse system. *Int. J. Adapt. Control Signal Process.* **33**, 1695–1705 (2019)
11. Z. Qin, J. Fan, Y. Liu, Y. Gao, G. Ye Li, Sparse representation of wireless Communications-a compressive sensing approach. *IEEE Signal Processing Magazine*, May-2018
12. C.S. Lai, Compression of power system signals with wavelets, in *2014 International Conference on Wavelet Analysis and Pattern Recognition* (2014)
13. W.H. Khong, P.A. Naylor, Efficient use of sparse adaptive filters, in *Proceedings of the 40th Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, CA, Oct 2006, pp. 1375–1379
14. Th.S.D. Singh, A. Chatterjee, A comparative study of adaptive algorithms for nonlinear system identification based on second-order Volterra and bilinear polynomial filters. *Measurement* (2011)

15. M. A. Rahman, M. Jannatul Ferdous, M.M. Hossain, M. Rashedul Islam, M. Hamada, A lossless speech signal compression technique, in *2019 1st International Conference on Advances in Science, Engineering and Robotics Technology (ICASERT)*, Dhaka, Bangladesh, pp. 1–7 (2019)
16. A.K. Sinha, S.K. Dutta, B. Dev, Jaydipta, R. Ranjan, R. Kumari, Wavelet-based Speech Coding technique using median function thresholding, in *2014 International Conference on Electronics and Communication Systems (ICECS)*, Coimbatore, pp. 1–4 (2014)

Performance Evaluation Based on Classification of Web Log Data: A Machine Learning Approach



Varun Malik, Ruchi Mittal, Jaiteg Singh, and Pawan Kumar Chand

Abstract The present study contributes to the understanding of artificial intelligence by assessing the importance of data in the generation of meaningful information and its prompt retrieval. Web mining is studied with the effective use of the WEKA tool. Web mining mainly deals with the extraction of knowledge flow and data present on the Web. Web mining can be categorized into three different categories Web structure mining, Web content mining, and Web usage mining. The present study emphasizes on Web usage mining that follows data processing under the three levels such as preprocessing, pattern discovery, and pattern analysis. The study findings reveal various algorithms can be adopted for data feeding and classification of data over some preprocessed method and comprised of the results. The findings of the study will be helpful for the users in acclimatizing with vivid algorithms that are associated with the Web data mining process which further gives them the sumptuous knowledge on how a dataset can be managed on Web usage according to their need of hours for generalizing the valid results.

Keywords Web mining · Web usage mining · Machine learning classifiers

1 Introduction

Machine learning is a category of techniques that is very effectively used for predictive analytics [1–5]. Web mining, also known as machine learning for Web applications, is an umbrella of techniques of data mining that are used to extract relevant patterns from Weblogs. It is used to identify and service the requirements of Web applications [6, 7]. In Web usage mining, the host identifies the location and other general information of users including their browsing data. Web usage mining is

V. Malik · R. Mittal (✉) · J. Singh
Chitkara University Institute of Engineering and Technology, Chitkara University, Rajpura,
Punjab, India
e-mail: ruchi.mittal@chitkara.edu.in

P. K. Chand
Chitkara Business School, Chitkara University, Rajpura, Punjab, India

a procedure that uses the techniques of data mining to examine and find relevant information from a user's log files. As the user surfs a Web site, the log files store the data related to the users' activities. Web mining can be categorized into three main components, namely Web content mining that discovers knowledge from hyperlinks, documents, and Web content; Web structure mining to discover information from the Website using hyperlinks; and Web usage mining that involves the activity of automatic discovery of user access patterns from Web servers located in different locations [8]. In Web usage mining, there are mainly three different stages such as preprocessing, pattern discovery, and pattern analysis [9].

This paper attempts to analyze the data extracted from Web server log files to compare the performance of various machine learning classifiers.

2 Literature Review

These studies explores Web mining from the machine learning perspective which makes this work very significant [10–12]. Authors in a study [6] presented that the term Web mining has been used in two distinct ways, namely, Web content mining and Web usage mining. In their paper, the authors stated that Web mining is used for techniques that relate to a wide range of issues. Despite being meaningful, this has led to different meanings of Web mining for different people and thus emphasized the need for common terminology.

Authors in [13] presented a Web text mining process capable of extracting knowledge in an environment, which involved multiple organizations having distributed and heterogeneous data. The Web text mining process consists of four steps given in a flexible architecture. In their prototype, the Web text mining process is initiated by a step where job offers are recovered. Using this process, input data was considered for classifying the job offers. This information mainly consisted of the details of the organization offering a job and the skills that were required for the job profile. For applying this step, the SOA and the miners model have been used to manage interactions between building blocks with well-defined tasks and of a reusable structure. These building blocks are recommended with the capability to interoperate with each other and create other services as a result. The significant advantage of this structure is that it helps a quick design of a given process to be carried out as per business needs. Further, new services can be created without the need of making many changes.

Authors in [14] stated that Web usage mining is used to extract relevant information from Web server log files, which contain details of Web user activity on Web sites so that a static Web site is analyzed and minimal changes are made to the Web site according to user needs. In this paper, the authors' objective was to incur minimal search costs in retrieving desired Web pages from Web sites. Thus, an algorithm termed Max-Heap with farthest first clustering approach was proposed. This algorithm was used to group frequently used Web pages, and then, these pages were reorganized with the help of the algorithm. The authors considered a Web site with Web pages with Web server log files, preprocessed log files, time spent by users,

and frequency of access by each user on the Web site. Based upon the prediction of accessing Web pages, a Max-Heap tree was created to process the Web site structure. After implementing the algorithm, experimental results were drawn out using Java. At last, the authors concluded that the recommended algorithms suggested Web pages to users based on the current visit and past navigation patterns.

According to [15], authors presented the data mining application called Web miner (Web mining), which helps in distance education useful to instructors for teaching their students, and the profiles and progress reports of individual students maintained as part of the distance education mode of teaching. In this paper, the authors aimed to make students aware that they do not need knowledge of data mining to use Web mining. Further, the authors presented how the tool uses the KDD process is based on Web services and is also used by external software.

3 Methodology

3.1 Preprocessing

The data has been derived from the server log files consisting of user's actions, and data preprocessing has been performed to improve the quality and effectiveness of data by cleaning and removing graphics and multimedia contents. Secondly, user sessions (total time spent by the user) are identified, and then, path completion is done to find the page references in a log file that are lost. This step of preprocessing is beneficial for identifying cyber-attacks and cybercrimes with the source of Web server log files [16].

3.2 Pattern Discovery

The data is then processed using the classification tab in WEKA, a Web usage mining tool used for data mining. The classification technique was applied to the existing dataset that included the contents of server log files accessed from google analytics and executed the dataset on different classification algorithms under the WEKA tool. In the process of classifying the dataset, we used different classification algorithms and analyzed the output of these algorithms on the same dataset. We have run the classifier using Weka based on different algorithms such as Naïve Bayes, Lazy IBK, decision tree (J48), function multilayer perception, and random forest. According to [17], today most of the research in this direction is being done using supervised methods, while unsupervised methods are not finding application in a widespread manner (Fig. 1).

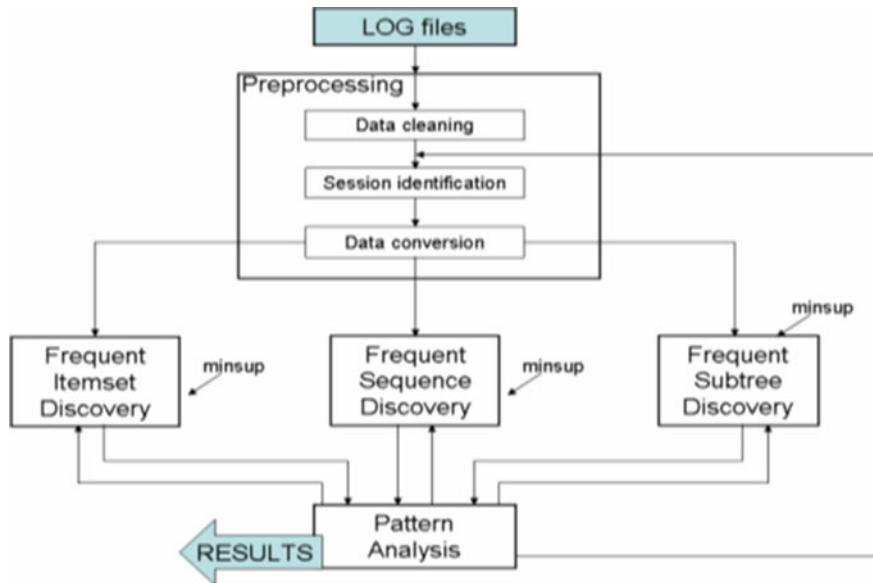


Fig. 1 Process of Web usage mining [9]

3.3 Pattern Analysis

When data over the Web is accessed, it is very hard to fathom the meaning of the result, and the quantum of results mined can be large, and therefore, it has to be understood as to how to identify relevant patterns from this data. Also, one has to apply and evaluate the result. All of these problems can be solved using visualizations tools, and several such tools exist to handle them [17].

4 Results and Analysis

The result of classification of data set using different data mining algorithm using WEKA tool is given in Table 1 along with the summary of the values of different parameters associated with the classification algorithm. In this experiment, five major classification algorithms have been used and compared on the dataset.

In the first algorithm, the Naïve Bayes classifier is used giving an accuracy of 81.4815%, wherein correctly instances are 4092 instances out of a total of 5022 instances and 930 incorrectly classified instances as shown in Fig. 2a. The second classification algorithm, Lazy IBK algorithm, has classified 4737 instances correctly out of total 5022 instances and 285 incorrectly classified instances giving a predictive accuracy of 94.3250% as in Fig. 2b.

Table 1 Results of classification of a dataset based on various algorithms

Parameter algorithm	Naive Bayes	LazyIBK	treeJ48	fxn multilayer perception	Random forest
Correctly classified instances	4092	4737	4933	4800	4956
Incorrectly classified instances	930	285	89	222	66
Kappa statistics	0.4868	0.8548	0.9544	0.8855	0.9663
Mean absolute error	0.1548	0.0383	0.0173	0.0373	0.0178
Root mean square error	0.2972	0.1931	0.1029	0.1547	0.0895
Related absolute error (%)	58.9383	14.5953	6.6041	14.1968	6.7933
Root relative square error (%)	82.0194	53.295	28.3938	42.7105	24.7120
Total no of instances	5022	5022	5022	5022	5022
TP rate	0.815	0.943	0.982	0.956	0.987
FP rate	0.343	0.087	0.031	0.066	0.021
Precision	0.804	0.943	0.981	0.951	0.986
Recall	0.815	0.943	0.982	0.956	0.987
F-Measure	0.805	0.943	0.981	0.952	0.986
MCC	0.525	0.862	0.961	0.901	0.972
ROC area class	0.904	0.929	0.981	0.970	0.994
PRC area	0.889	0.921	0.977	0.960	0.991
Correctly classified instances (%)	81.4815	94.3250	98.2278	95.5795	98.6858
Incorrectly classified instances (%)	18.5185	5.6750	1.7722	4.4205	1.3142

In the next step of classification of the dataset, we used the tree-based classification algorithm, i.e., J48 algorithm, and it gave the predictive accuracy of 98.2278% as 4933 instances out of total 5022 instances were correctly classified and 89 were incorrectly classified instances as in Fig. 2a. In a different classification approach, function multilayer perception algorithm was applied which has classified 4800 instances correctly and 222 instances incorrectly, thereby giving the predictive accuracy of 95.5795% as shown in Fig. 2b.

The last classification algorithm that was applied on the dataset is random forest, and it gives a predictive accuracy of 98.685%, and it has 4956 instances out of total 5022 instances as correctly classified and 66 incorrectly classified instances as shown in Fig. 2a, b.

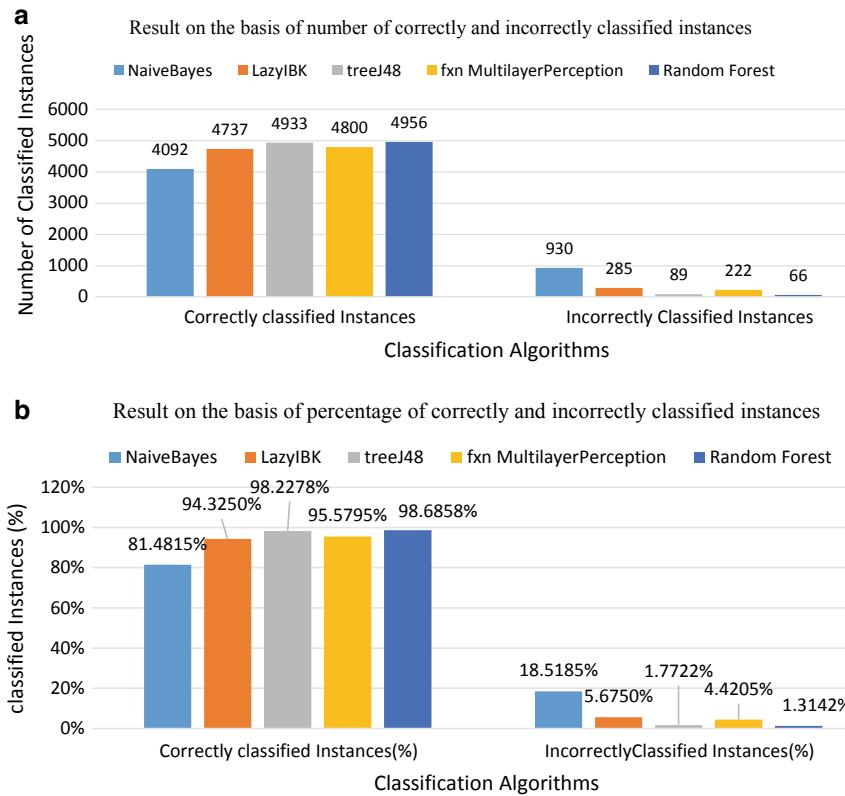


Fig. 2 **a, b** Summary of the classification of algorithms

5 Conclusion

The study succeeded in managing the huge flow of data on the Web for generating the information and the prompting of user-friendly retrievals of information. The Web usage mining connotes that server log files are occupied with huge amounts of unwanted data which needs processing to filter out the necessary information for instant usage and to reinforce the decision-making process of the users [18]. The study has successfully tested the vivid datasets from server log files preprocessed them and has executed the data on different algorithms using the WEKA tool. The findings of the study show that the Naïve Bayes algorithm process 81.4815% correctly instances while executed Lazy IBK, tree J48, function multilayer perception, and random forest algorithm with 94.3250%, 98.2278%, 95.5795%, and 98.6858%, respectively, as shown in Fig. 2a, b. Further, the findings of the study also conclude that random forest algorithm instances have the highest number of correctly classified instances when compared with Naïve Bayes, Lazy IBK, J48, and function multilayer perception algorithms.

References

1. A. Mittal, A. Aggarwal, R. Mittal, Predicting university students' adoption of mobile news applications: the role of perceived hedonic value and news motivation. *Int. J. E-Serv. Mobile Appl. (IJESMA)* **12**(4), 42–59 (2020)
2. S. Gupta, R. Mittal, A. Mittal, Modelling the intentions to adopt APIs: a PLS-SEM approach, in *2019 6th International Conference on Computing for Sustainable Global Development (INDIACom)* (IEEE, 2019), pp. 246–250
3. R. Mittal, Multivariate regression predictive modeling in analyzing student performance: a data mining approach. *J. Comput. Theor. Nanosci.* **16**(10), 4362–4366 (2019)
4. R. Mittal, Prediction of heart diseases using Hayes process macro serial mediation model 6. *Indian J. Public Health Res. Dev.* **10**(10), 538–543 (2019)
5. R. Mittal, Identification of salient attributes in social network: a data mining approach, in *International Conference on Recent Developments in Science, Engineering and Technology* (Springer, Singapore, 2019), pp. 173–185
6. T. Srivastava, P. Desikan, V. Kumar, Web mining—concepts, applications and research directions, in *Foundations and Advances in Data Mining* (Springer, Berlin, Heidelberg, 2005), pp. 275–307
7. H. Chen, M. Chau, Web mining: machine learning for web applications. *Ann. Rev. Inf. Sci. Technol.* **38**(1), 289–329 (2004)
8. L.K. Grace, V. Maheswari, D. Nagamalai, Analysis of web logs and web user in web mining (2011). arXiv preprint [arXiv:1101.5668](https://arxiv.org/abs/1101.5668).
9. R. Ivánky, I. Vajk, Frequent pattern mining in web log data. *Acta Polytechnica Hungarica* **3**(1), 77–90 (2006)
10. A. Gudivada, N. Tabrizi, A literature review on machine learning based medical information retrieval systems, in *2018 IEEE Symposium Series on Computational Intelligence (SSCI)* (IEEE, 2018), pp. 250–257
11. M. Rath, Machine learning and its use in e-commerce and e-business, in *Handbook of Research on Applications and Implementations of Machine Learning Techniques* (IGI Global, 2020), pp. 111–127
12. M. Rath, S. Mishra, Advanced-level security in network and real-time applications using machine learning approaches, in *Machine Learning and Cognitive Science Applications in Cyber Security* (IGI Global, 2019), pp. 84–104
13. M. Castellano, G. Mastronardi, A. Aprile, G. Tarricone, A web text mining flexible architecture. *World Acad. Sci. Eng. Technol.* **32** (2007)
14. K. Shyamala, S. Kalaivani, An effective web page reorganization through heap tree and farthest first clustering approach, in *2017 IEEE International Conference on Power, Control, Signals and Instrumentation Engineering (ICPCSI)* (IEEE, 2017), pp. 2259–2262
15. D. García-Saiz, M.E. Zorrilla, E-learning web miner: a data mining application to help instructors involved in virtual courses, in *EDM*, pp. 323–324
16. V. Chitraa, D. Davamani, A. Selvadoss, A survey on preprocessing methods for web usage data. arXiv preprint [arXiv:1004.1257](https://arxiv.org/abs/1004.1257)
17. V. Anitha, P. Isakkki, A survey on predicting user behavior based on web server log files in a web usage mining, in *2016 International Conference on Computing Technologies and Intelligent Data Engineering (ICCTIDE'16)* (IEEE, 2016), pp. 1–4
18. G. Stumme, A. Hotho, B. Berendt, Semantic web mining: state of the art and future directions. *J. Web Semant.* **4**(2), 124–143 (2006)

Energy and Direction Aware Routing Protocol for Flying Ad Hoc Networks



Sudesh Kumar , Ram Shringar Raw, and Abhishek Bansal

Abstract Recently, unmanned aerial vehicles (UAVs) have gained popularity in various applications such as military, commercial, and civilian domains. The communication between multiple UAVs in the sky can naturally form a flying ad hoc network (FANETs). However, due to limited energy and high mobility of UAVs, the performance of the network can degrade therefore; designing an efficient routing protocol for FANETs becomes a very challenging issue. In this paper, a new energy and direction aware location aided routing protocol (ED-LAR) has been proposed for FANETs environment, based on a utility function which considers the residual energy and movement direction of UAVs. The proposed routing scheme improves the link duration time between the UAVs as well as network lifetime by selection of best next hop among the UAVs in the route discovery process. The simulation result indicates that our scheme performs well as compared to state-of-the-art schemes.

Keywords UAVs · FANETs · ED-LAR · LAR · GPS

1 Introduction

The collaboration between the unmanned aerial vehicles (UAVs) in flying ad hoc networks (FANETs) provides a wide range of applications such as traffic monitoring, military purpose, entertainment industries, product delivery, civilian sectors, emergency management, etc. [1–4]. In FANETs, multiple UAVs commonly known as the flying vehicle or drone can communicate with each other for establishing an ad

S. Kumar · A. Bansal

Department of Computer Science, Indira Gandhi National Tribal University, Amarkantak, MP, India

e-mail: sudesh.kumar@igntu.ac.in

A. Bansal

e-mail: abhishek.bansal@igntu.ac.in

R. S. Raw

Department of Computer Science and Engineering, NSUT (East Campus), New Delhi, India
e-mail: shringarjnu@gmail.com

hoc network for different applications scenario. But the other side these applications require highly accurate real-time data transmission among the UAVs to finish the mission in the time bound period through efficient way. Further, the high mobility and restricted energy of UAVs is the main reason of frequent topology changing of FANETs resulting in intermittent link among the UAVs. Therefore, it is very difficult for topology-based routing (TBR) protocols to maintain stable route among the UAVs in FANETs environment [5], since TBR schemes forward the packet based on the route obtained previously [6]. As compared to TBR, the position-based routing (PBR) protocols highly concern with the selection of the best next hop in the route discovery process. Therefore, the PBR schemes are more impactful for the dynamic nature of FANETs. The PBR protocols include the location information of the moving UAVs through the global positioning system (GPS), and further, this information is used to recognize the route for forwarding the packets between the source and destination [7]. FANETs basically derived from the traditional ad hoc networks such as mobile ad hoc network (MANET) and vehicular ad hoc network (VANET). But due to the dynamic behaviour of FANETs, the existing routing protocols of MANETs and VANETs are not viable in this environment. Therefore, many of the improvement are required in routing strategy [8]. In addition, energy constraint of UAVs is one of the major issues for FANETs environment. Many of the energy-aware schemes proposed such as CBLADR [9], RGR [10], and EALC [11], but these schemes are weakening into link durability criteria due to high mobility of UAVs. The authors [12] introduced PBR scheme such as LAR for MANET environment; the proposed scheme partitions the network into two zones; expected zone and request zone. This protocol controls the flooding and overhead by restricting other mobile nodes in the network via zone concept. After that, LAR has inspired extensions in VANET environment also with more concepts such as distance, movement, inter vehicle communication link. [13]. All these schemes proposed under the low mobility of nodes and less concern about the energy constraint. In reference [14], the authors' proposed UAVR scheme for VANET which is assisted by the UAVs for VANET, but this scheme fails to deal with link stability criteria between the UAVs. Furthermore, the authors [15] implemented LAR protocol in a FANETs environment with high mobility as well as a verity number of UAVs through the simulation. But due to energy limitation of UAVs, the minimum link duration time and network lifetime can degrade the performance of FANETs. Therefore, in LAR, some enhancements are required to make more suitable of this protocol for FANETs environment. Considering the above-mentioned issues of energy, link durability, and network lifetime, in this paper, the author's proposed ED-LAR scheme with following important contributions:

- Residual energy ratio and movement direction concepts have been defined to improve the routing process for selection of the best next hop UAV in the forward direction of destination UAV.
- The proposed approach improves the link duration time between the UAVs, as well as the network lifetime.

2 Proposed Routing Strategy

Due to the unique characteristics such as very high mobility of UAVs, restricted energy of UAVs, regularly network topology varying, some improvements are required in the existing LAR protocol. Therefore, some assumption listed as follows.

- Every UAV like U_i is aware of its own 3D coordinate (X_i, Y_i, Z_i) via using GPS and exchanges the useful information such as UAV-identity, UAV-location, UAV-direction as well as UAV-residual energy of one-hop neighbours periodically and update the neighbour list table through *HELLO* messages.
- All UAVs randomly deployed to initially have equal energy with a same transmission range (R), and links among the UAVs are bidirectional.

The proposed ED-LAR scheme is based on two key system parameters: the residue energy ratio (η), and the movement direction of UAV (θ), which indicating the angle of moving UAVs from the baseline draw from the forwarding UAV to destination UAV. Furthermore, in this proposed scheme, the neighbour UAV having a minimum value of utility function (UF) is selected as a next hop for further transmission of the data packets on the basis of η as well as angle θ . The UF is defined as

$$\text{UF} = \alpha(1 - \eta) + (1 - \alpha)\theta \quad (1)$$

where α is a tuneable weight for η, θ .

2.1 Route Discovery Process

The proposed ED-LAR scheme firstly utilized expected zone (EZ) and the request zone (RZ) as similar to the existing LAR protocol. Figure 1 shows the route discovery process for neighbour selection in FANETs scenario. When the route request (RREQ) is initiated, the forwarder UAV let U_S broadcast a *RREQ* message to all its neighbours UAV (RZ) such as U_1 and U_2 . In this process, firstly, if more than one neighbour UAVs have equal θ but different η , then consider $\alpha = 1$ in Eq. (1) and after that select the next hop UAV. Secondly, if more than one neighbour UAV has equal η but different θ , then consider $\alpha = 0$ in Eq. (1). Furthermore, if all neighbour UAVs have different η and θ , then put $\alpha = 0.5$ in Eq. (1) and select the next hop UAV accordingly. In Fig. 1, the forwarder UAV U_S selects U_2 as a next hop UAV, because U_2 has a minimum value of *UF* as compared to all other neighbour of U_S . Once the U_2 receives the *RREQ* from previous forwarder, the U_2 becomes the new next forwarder and follows the same procedure until the data packet received by the $U_D(U_S \rightarrow U_2 \rightarrow U_4 \rightarrow U_9 \rightarrow U_{13})$.

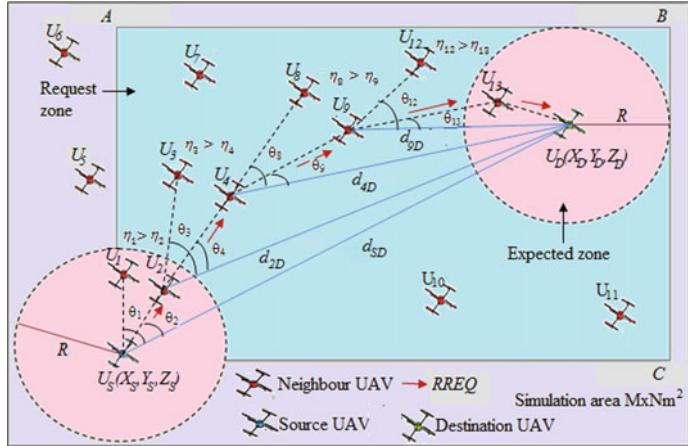


Fig. 1 Route discovery process through proposed ED-LAR

3 Mathematical Analyses

The proposed ED-LAR scheme considers the residual energy as a key parameter, which plays the important role in the selection of the appropriate next hop UAV because, if the residual energy level [16] of any UAV node is low (threshold = 30 J), that UAV cannot participate in the routing process, and hence, the UAV will dead. Let $E^{\text{int}}(U_i)$ as the initial energy of the any UAV node U_i , then the residual energy $E^{\text{res}}(U_i)$ of U_i can calculate as follows.

$$E^{\text{res}}(U_i) = E^{\text{int}}(U_i) - E^{\text{con}}(U_i) \quad (2)$$

where $E^{\text{con}}(U_i)$ refers to the energy consumption by U_i for transmitting as well as receiving data. So, the residual energy ratio η_i of the U_i can be expressed as:

$$\eta_i = \frac{E^{\text{res}}(U_i)}{E^{\text{int}}(U_i)} \quad (3)$$

In addition, the second key parameter for the ED-LAR scheme is the direction of UAV movement, which is measured for selecting the next hop UAV has a minimum threshold (45°) angle from the baseline drawn from the forwarding UAV to destination UAV. Let θ_i is an angle formed between a neighbour U_i , sender U_S , and destination U_D , and then, θ_i can be calculated by the following question:

$$\theta_i = \cos^{-1} \left(\frac{(d_{Si})^2 + (d_{SD})^2 - (d_{iD})^2}{2d_{Si} \cdot d_{SD}} \right) \quad (4)$$

Furthermore, let the coordinate values of U_S , U_i , and U_D are (X_S, Y_S, Z_S) , (X_i, Y_i, Z_i) , and (X_D, Y_D, Z_D) , respectively. Then, distance among the UAVs can be calculated as follows:

$$d_{SD} = \sqrt{(X_S - X_D)^2 + (Y_S - Y_D)^2 + (Z_S - Z_D)^2} \quad (5)$$

$$d_{Si} = \sqrt{(X_S - X_i)^2 + (Y_S - Y_i)^2 + (Z_S - Z_i)^2} \quad (6)$$

$$d_{iD} = \sqrt{(X_i - X_D)^2 + (Y_i - Y_D)^2 + (Z_i - Z_D)^2} \quad (7)$$

These above-discussed parameters play a very important role to improve the link duration time (LDT) among the UAVs. The LDT is simply defined by the period of time in which two UAVs connected in the network. The LDT can be calculated with the motion parameters of two neighbour UAVs within the transmission range R . Let (X_i, Y_i, Z_i) and (X_j, Y_j, Z_j) denote the position of two UAVs like U_i and U_j and assumed that $(Z_i = Z_j = 0)$. Furthermore, let V_i and V_j is the speed of UAVs ($V_i = V_j = V$) along with the θ_i and θ_j as movement direction, respectively, towards the destination UAV. Then, the LDT_{ij} between the two UAVs such as U_i and U_j can be calculated as equation below:

$$LDT_{ij} = \frac{-(pq + rs) + \sqrt{(p^2 + r^2)R^2 - (ps - qr)^2}}{(p^2 + r^2)} \quad (8)$$

where $p = V \cos \theta_i - V \cos \theta_j$, $q = X_i - X_j$, $r = V \sin \theta_i - V \sin \theta_j$, $s = Y_i - Y_j$.

4 Simulation Set-up and Result Analysis

In this section, the authors evaluate the impact of ED-LAR by comparing the link duration time and network lifetime between the proposed ED-LAR and existing UAVR as well as LAR scheme with weight coefficient $\alpha = 0|0.5|1$ through ns-2 simulator. In the beginning of FANETs scenario, all UAVs (05-25) are randomly distributed in the area of $800 \text{ m} \times 700 \text{ m}$ with a transmission range of each UAV is fixed to 250 m and used 802.11 g as a mac layer. The velocity of each UAV is fixed as 50 m/s. The initial energy of every UAV is 100 J. The simulation time is 200 s with 512 bytes packet size, and traffic type is CBR.

It can be seen clearly from Fig. 2, the proposed ED-LAR is greater link duration time as compared to the UAVR and existing LAR protocol. Basically, the link duration time between the UAVs depends on the transmission range as well as residual energy of UAVs. In addition, if the link duration time is high, then the UAVs can communicate more time as a result transmit more data packet which increases the performance of the FANETs. In this FANETs scenario, seven number of UAVs move with fixed

Fig. 2 LDT with communication range

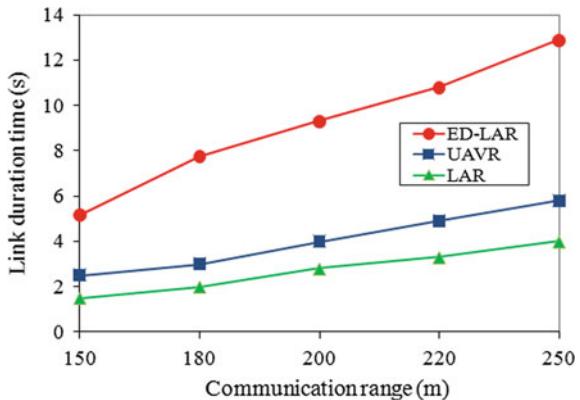
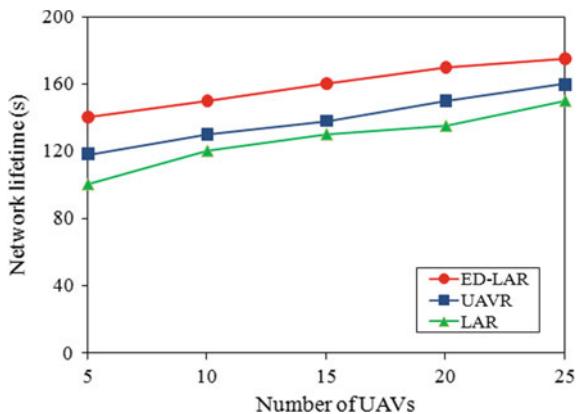


Fig. 3 Network lifetime with UAVs



velocity as a 50 m/s towards the destination UAV. Furthermore, the network lifetime can be calculated as the time from the commandment of the simulation until the first node in the network is dead. Figure 3 shows the network lifetime of ED-LAR with a different number of UAVs moving with the same speed as a 50 m/s. The result presents that the proposed scheme is performing well against the other two protocols, because it is adopting the residual energy thresholds to avoid the participation of lower residual energy UAVs.

5 Conclusion

FANETs require precise routing protocols for the efficient route between the UAVs due to the challenges faced by highly dynamic environment. In this paper, the authors present a new ED-LAR protocol based on the residual energy and the movement direction of UAV for the FANETs environment. Furthermore, on the basis of utility

function, the proposed ED-LAR find the optimal route through the selecting of best next hop UAV. To verify the effectiveness of this scheme, we simulate the ED-LAR and compared with UAVR and LAR protocol. The simulation results show that, the ED-LAR outperforms state-of-the-art protocols in terms of link duration time as well as network lifetime.

References

1. I. Bekmezci, O.K. Sahingoz, S. Temel, Flying Ad-Hoc Networks (FANETs): a survey. *Ad Hoc Netw.*, 1254–1270 (2013). <https://doi.org/10.1016/j.adhoc.2012.12.004>
2. S. Kumar, A. Bansal, R.S. Raw, Health monitoring planning for on-board ships through flying ad hoc network, in *Advanced Computing and Intelligent Engineering. Advances in Intelligent Systems and Computing*, ed. by B. Pati, C. Panigrahi, R. Buyya, K.C. Li, vol. 1089 (Springer, Singapore, 2020), pp. 391–402. https://doi.org/10.1007/978-981-15-1483-8_33
3. C.R. Panigrahi, J.L. Sarkar, B. Pati, B. Bakshi, E3M: an energy efficient emergency management system using mobile cloud computing, in *IEEE International Conference on Advanced Networks and Telecommunications Systems (ANTS)*, Bangalore, pp. 1–6 (2016). <https://doi.org/10.1109/ANTS.2016.7947862>
4. M. Rath, B. Pati, B.K. Pattanayak, Mobile agent-based improved traffic control system in VANET, in *Integrated Intelligent Computing, Communication and Security. Studies in Computational Intelligence*, ed. by A. Krishna, K. Srikanthaiah, C. Naveena (Springer, Singapore, 2019), p. 771. doi.org/10.1007/978-981-10-8797-4_28
5. O.M. Oubbati, A. Lakas, F. Zhou, M. Günes, M.B. Yagoubi, A survey on position-based routing protocols for Flying Ad hoc Networks (FANETs). *Veh. Commun.* **10**, 29–56 (2017). <https://doi.org/10.1016/j.vehcom.2017.10.003>
6. S. Kumar, A. Bansal, Performance investigation of topology-based routing protocols in flying ad-hoc networks using NS-2, in *IoT and Cloud Computing Advancements in Vehicular Ad-Hoc Networks*, ed. by R. Rao, V. Jain, O. Kaiwartya, N. Singh (IGI Global, Hershey, PA, 2020), pp. 243–267. <https://doi.org/10.4018/978-1-7998-2570-8.ch013>
7. L. Lin, Q. Sun, J. Li, F. Yang, A novel geographic position mobility oriented routing strategy for UAVs. *J. Comput. Inf. Syst.* **8**, 709–716 (2012)
8. O.S. Oubbati, M. Mozaffari, N. Chaib, P. Lorenz, M. Atiquzzaman, A. Jamalipour, ECaD: energy-efficient routing in flying Ad hoc networks. *Int. J. Commun. Syst.*, e4156 (2019). <https://doi.org/10.1002/dac.4156>
9. N.C. Shi, X. Luo, A novel cluster-based location-aided routing protocol for UAV fleet networks. *Int. J. Digit. Content Technol. Its Appl. (JDCTA)* **6**(8) (2012). <https://doi.org/10.4156/jdcta.vol6.issue18.45>
10. R. Shirani, M. St-Hilaire, T. Kunz, Y. Zhou, J. Li, L. Lamont, On the delay of reactive-greedy-reactive routing in unmanned aeronautical ad-hoc networks. *Procedia Comput. Sci.* **10**, 535–542 (2012). <https://doi.org/10.1016/j.procs.2012.06.068>
11. F. Aadil, A. Raza, M.F. Khan, M. Maqsood, I. Mahmood, S. Rho, Energy aware cluster-based routing in flying ad-hoc networks. *Sensors (Basel, Switzerland)* **18**(5), (2018). <https://doi.org/10.3390/s18051413>
12. Y. Ko, N.H. Vaidya, Location-aided routing (LAR) in mobile ad-hoc networks. *Int. J. Wirel. Netw.* **6**(4), 307–321 (2000)
13. R.S. Raw, D.K. Lobiyal, S. Das, S. Kumar, Analytical evaluation of directional-location aided routing protocol for VANETs. *Int. J. Wirel. Pers. Commun.* **82**(3), 1877–1891 (2015). <https://doi.org/10.1007/s11277-015-2320-7>
14. O.S. Oubbati, A. Lakas, N. Lagraa, M.B. Yagoubi, UVAR: an intersection UAV-assisted VANET routing protocol, in *IEEE Wireless Communications and Networking Conference*, Doha, pp. 1–6 (2016). <https://doi.org/10.1109/WCNC.2016.7564747>

15. S. Kumar, A. Bansal, R.S. Raw, Analysis of effective routing protocols for flying ad-hoc networks. *Int. J. Smart Veh. Smart Transp. (IJSVST)*. **3**(2), 1–18 (2020). <https://doi.org/10.4018/IJSVST.2020070101>
16. O.S. Oubbat, M. Atiquzzaman, P. Lorenz, M.H. Tareque, M.S. Hossain, Routing in flying ad hoc networks: survey, constraints, and future challenge perspectives. *IEEE Access* **7**, 81057–81105 (2019). <https://doi.org/10.1109/ACCESS.2019.2923840>

Performance Comparison of Tree-Based Machine Learning Classifiers for Web Usage Mining



Ruchi Mittal, Varun Malik, Vikas Rattan, and Deepika Jhamb

Abstract Web usage mining plays a very important role in finding new patterns and recognition from the web server log file. It is the subcategory of web mining that is used to mine data over the web. In this paper, the authors classified the server log file, identified new patterns of data, and analyzed the data over the tree-based classification algorithms. Authors accessed the tree algorithms (treeJ48, RandomTree, RandomForest, and REPTree) to classify the weblog data and analyzed the results. This paper summarizes the results and compared the tree-based algorithms classification data and check which algorithm gives a better result.

Keywords Web mining · Web usage mining · Machine learning classifiers

1 Introduction

The concept of web mining can be considered as storing the information related to the users' activities on computer and web usage mining is one of the categories of web mining which works on weblog data to extract patterns based on the way the user browse the information on the web [1–5]. The process of web mining consists of three independent stages, namely web content mining, web structure mining, and web usage mining. All these stages help in pointing out the users' access patterns on the web, and such knowledge can be useful for catering to the varying needs of the organizations. The task of discovering these patterns can be accomplished using various web mining techniques such as classification, clustering, associations, and visualization. The previous research in the area of web mining has potentially identified the patterns in the log files present on the web to understand the users'

R. Mittal · V. Malik (✉) · V. Rattan

Institute of Engineering and Technology, Chitkara University, Rajpura, Punjab, India
e-mail: varun.malik@chitkara.edu.in

D. Jhamb

Chitkara Business School, Chitkara University, Rajpura, Punjab, India

behavior and the way they access the websites [6]. Web mining, similar to other data mining techniques, is a very effective tool used for predictive analysis [7–9].

The key aim of the paper is to perform the classification techniques on weblog data and to understand the performance of the various data mining classifiers in terms of their predictive accuracy. Our approach is mainly based on applying the tree-based classifiers such as J48, random tree, random forest, and REP tree for classifying the dataset [10].

Web usage mining is to analyze the data present on the web pages in the form of log records. These log records contain a variety of important information such as URL accessed by the user, IP address used by the user at the time of web access, and the time at which the information was accessed, and so on. Mining such information to discover hidden and novel patterns in these log files has a lot of potential in developing customers understanding to improve the quality of services offered to the customers [11]. While visiting and browsing a website, users leave a lot of information that can be processed to gain important insights about the users' preferences and behavior on the web to customize the web sites based on users' choices and preferences and to improve their experience on the web [12]. The web mining techniques can assist to collect, pre-process, analyze, and interpret the data, and the models can be built to have better website structures and improved business operations [13, 14]. The authors in this study proposed to compare the performance of various tree-based classification algorithms such as decision tree J48, random tree, random forest, and REP tree in terms of their predictive accuracy using Weka tool.

2 Related Work

According to [15], it is defined that web usage mining plays a crucial role in designing websites and web pages. In this paper, the authors discuss server log files and other relevant information on the web, which led to useful mining of data for future use. Further, authors describe a purposed architecture online MINER which led to the identification of those pages automatically which are not visited a single time, and it also analyzes the traffic volume on the website. In this paper, the authors describe that web usage mining or patterns, and data mining combined will be the basis for future work and future research which will apply data mining algorithms and the KDD process.

Authors described that web mining is a body of knowledge related to web data in the form of such techniques as web content, web structure, and web usage mining. In this paper, the authors identify the main problems in creating an intelligent tool available for mined knowledge and provide in all three categories of web mining. The paper also gives a brief overview of research in web usage mining and their architecture of usage of extracted data and their development efforts in the field of web mining. Authors state that web mining is a fast-growing area of research nowadays. The authors also suggested a starting point for identifying opportunities for future research work on web usage mining [16].

In a different study, the authors implemented a technique in web usage mining for the banking system that helped a company to handle web performance issues. The authors used the pattern discovery phase and proposed a classification technique based on the K-nearest algorithm using Euclidean distance to classify patterns, and the results aided the company to extract useful knowledge from web server logs. The authors concluded that this technique can be used to identify activity data of surfers in web server logs. This allowed the company to create relevant data that they used for analyzing web applications. In the paper, the K-nearest neighbor algorithm showed good results in comparison with Bayesian classification; therefore, it can further be used in web usage mining. Further, researchers can work on another data mining algorithm and can combine different data to get the desired results [17].

According to [18], the authors have defined that the process of finding valuable information and knowledge from the web data is called web usage mining. In this paper, the authors explained about web usage mining that identification of patterns is of high importance in various fields such as business intelligence, e-learning, personalization, etc. This paper described the application areas of web usage mining in-depth and also described the current issues of web usage mining. In this paper, the authors explained two major issues in web usage mining. The first issue is security for users, and the second issue is including semantics in web content. Further, the authors suggested the various areas in web usage mining which will lead to future trends in research. Further, analysis is performed on the results of mining to extract information, to help designers handle users' needs, i.e., individualizing website usage experience and better organization of websites.

Authors presented web usage mining as a data mining mechanism to extract knowledge from log data. Web usage mining is a data mining technique for pattern discovery of new patterns from web server log files for research. The web includes a large collection of data or information for academics and analyzed the data for organizations and institutes for better services in academics to improve the performance of the website by processed or accessed the log files. This paper mainly focused on data collection over web servers for academic education by using Weblog Expert lite 9.3 tools. The authors concluded that webpages are a beneficial advertisement tool for an institution and other sectors like government, etc. The huge benefit can be obtained from web advertisements, website design, and content. This is an area of data mining and pattern recognition which can lead to a good impact in research work [19].

3 Research Methodology

The data mining tool, Weka was selected to process web server log data. The data has been sourced from a leading university in India and has been used in a different study by [10]. The web server log dataset consisted of the weblog files having attributes such as session, new session, bounce rate, pages per session, and class variable having the values based on the relevant, irrelevant, and most relevant instance for a particular

row in the dataset. Initially, the data was pre-processed to handle the missing values. For this, in the pre-process tab, ReplaceMissingValue method is selected to convert the missing and zero value to NULL values for the application of classification technique on the dataset.

The web server log data is run through the J48 algorithm initially for classification. The experiment was run twice by keeping the percentage split ratio as 60:40, 70:30, 80:20, and 90:10 for specifying the training and testing data, respectively. A similar process was followed with other classification algorithms such as random forest, random tree, and REP Tree algorithms. Then, the results were observed and evaluated based on the correctly classified and incorrectly classified instances for all these algorithms.

4 Experiment and Results

This paper mainly focuses on the classification of data, where data is derived from web server log files. These log files may contain a lot of noise or unwanted data that is not relevant to the analysis according to the end-users perspective; therefore, before classification, data is pre-processed using ReplaceMissingValues filter to set the blank values to NULL values in the dataset. After the pre-processing step, various classification methods are implemented on the dataset using Weka, a data mining tool. The dataset consisting of 5022 instances, is run through various tree-based classifiers, namely treeJ48, random forest, random tree, and REP tree. The percentage split (%) method is used as a test category to classify these instances to check the performance of different algorithms on the same dataset using the WEKA tool. By default, the percentage split value is set to 66%.

The first experiment was conducted using a treeJ48 classifier by setting percentage split (%) values to 60%, 70%, 80%, and 90%, respectively, on 5022 instances of weblog files. With 60:40 split ratio on total 5022 instances, classification algorithm J48 resulted in 1968 correctly classified instances and 41 incorrectly classified instances were incorrectly instances. With 70:30 split ratio, the same algorithm correctly classified 1478 instances and 29 instances were incorrectly classified, i.e., 98.0756% correctly classified instances and 1.9244% incorrectly classified instances. In another case, 1004 instances are classified by setting the split value to 80%. Out of 1004 instances, 986 instances are correctly classified and 18 instances are incorrectly classified instances, i.e., 98.2072% correctly classified and 1.7928% incorrectly classified instances. When the percentage split is kept as 90:10, the J48 algorithm correctly classified 493 instances and 9 instances were incorrectly classified out of total 502 instances, giving an accuracy of 98.4064%. This clearly indicates that the J48 algorithm is giving the maximum accuracy on the dataset when the classification criteria in the form of percentage split ratio was kept to be 90:10 as shown in Table 1.

The second experiment was conducted using random forest classifier by setting percentage split (%) value to 60%, 70%, 80%, and 90%, respectively, on 5022

Table 1 Classification of dataset over tree algorithms with different percentage split

Classification algorithms	Percentage split (%)	Total number of instances executed	Correctly classified instances	Incorrectly classified instances	Correctly classified instances (%)	Incorrectly classified instances (%)
treeJ48	60	2009	1968	41	97.9592	2.0408
	70	1507	1478	29	98.0756	1.9244
	80	1004	986	18	98.2072	1.7928
	90	502	494	8	98.4064	1.5936
Random forest	60	2009	1977	32	98.4072	1.5928
	70	1507	1484	23	98.4738	1.5262
	80	1004	991	13	98.7052	1.2948
	90	502	493	9	98.2072	1.7928
Random tree	60	2009	1954	55	97.2623	2.7377
	70	1507	1468	39	97.4121	2.5879
	80	1004	983	21	97.9084	2.0916
	90	502	490	12	97.6096	2.3904
REP tree	60	2009	1965	44	97.8099	2.1901
	70	1507	1471	36	97.6111	2.3889
	80	1004	977	27	97.3108	2.6892
	90	502	492	10	98.0080	1.9920

instances of weblog files. First of all percentage split ratio was kept as to 60%, and thus, total 2009 instances were used out of total 5022 instances, and out of which, 1977 instances were correctly classified, and 77 instances were incorrectly classified. With 70% split, 1507 instances out of total 5022 instances were used for classification, and in the result, 1484 instances were correctly classified, and 23 instances were incorrectly classified, i.e., 98.4738% correctly classified instances and 1.5262% incorrectly classified instances. In another case, 1004 instances are classified by setting the split value to 80%. Out of 1004 instances, 991 instances are correctly classified, and 13 instances are incorrectly classified instances, i.e., classification accuracy of 98.7052%. Finally, the percentage split is kept as 90:10, and total 502 instances were used for classification; out of which, 493 instances were correctly classified, and 9 instances were used incorrectly classified here and thus giving a classification accuracy of 98.2072%. Thus, random forest works best on the dataset giving maximum accuracy of 98.7052% by keeping the percentage split ratio to 80:20 as in Table 1.

The third experiment was conducted using random tree classifier by setting percentage split (%) value to 60%, 70%, 80%, and 90%, respectively, on 5022 instances of weblog files. Firstly, when the percentage split is 60%, total 2009 instances were classified out of a total of 5022 instances, and of these, 1954 instances were correctly classified, and 55 instances were incorrectly classified. With a 70%

split ratio, 1507 instances out of total 5022 instances were used for classification, and in the result, 1468 instances were correctly classified, and 39 instances were incorrectly classified, i.e., 97.4121% correctly classified instances and 2.5879% incorrectly classified instances. In another case, 1004 instances are classified by setting the split value to 80%. Out of 1004 instances, 983 instances are correctly classified, and 21 instances are incorrectly classified instances, i.e., 97.9084% correctly classified and 2.0916% incorrectly classified instances. When the percentage split is 90%, total instances examined are 502; out of which, 490 instances were correctly classified, and 12 instances were incorrectly classified giving the classification accuracy of 97.6096% as shown in Table 1.

The fourth experiment was conducted using REP tree classifier by setting percentage split (%) value to 60%, 70%, 80%, and 90%, respectively, on 5022 instances of weblog files. When the percentage split ratio is set to 60%, total 2009 instances were used for classification, and out of these, 1965 instances were correctly classified, and 65 instances were incorrectly classified. With 70% split ratio, 1507 instances out of total 5022 instances were used for classification, and in the result, 1471 instances were correctly classified, and 36 instances were incorrectly classified, i.e., 97.6111% correctly classified instances and 2.3889% incorrectly classified instances. In another case, 1004 instances are classified by setting the split value to 80%. Out of 1004 instances, 977 instances are correctly classified, and 27 instances are incorrectly classified instances, i.e., 97.3108% correctly classified and 2.6892% incorrectly classified instances. When the percentage split ratio is set to 90:10, total 502 instances were used for classification; out of which, correctly classified instances were 492, and incorrectly classified instances were 10 giving classification accuracy of 98.0080% as shown in Table 1 (Figs. 1, 2, 3, and 4).

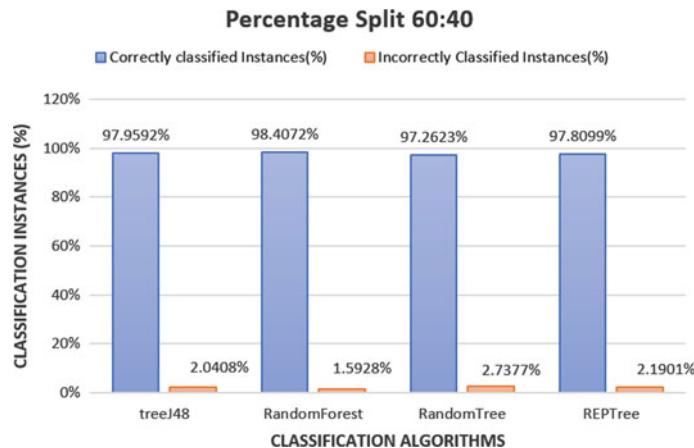


Fig. 1 Classification based on 60:40 percent split

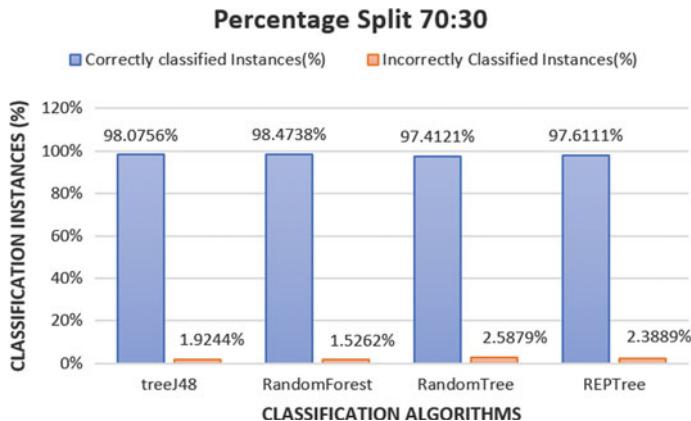


Fig. 2 Classification based on 70:30 percent split

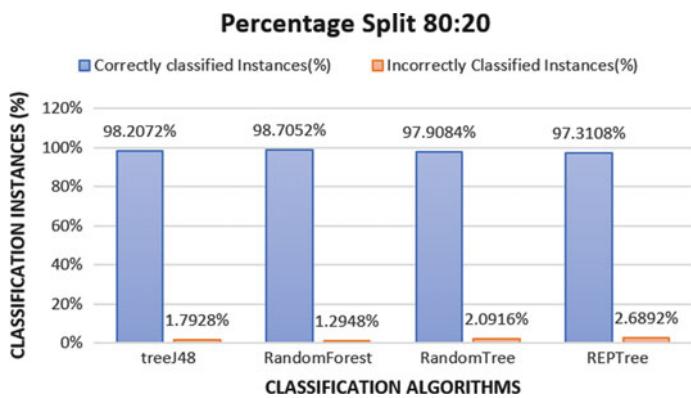


Fig. 3 Classification based on 80:20 percent split

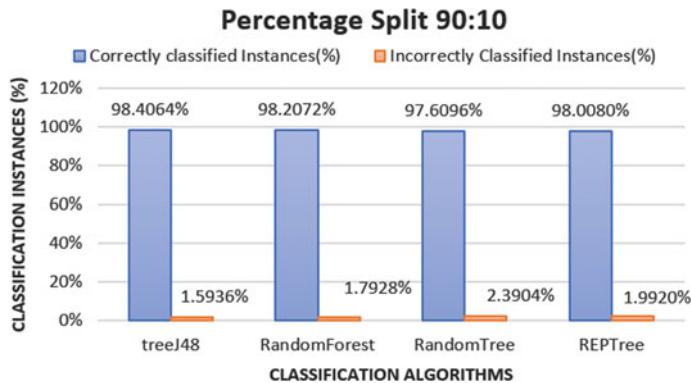


Fig. 4 Classification based on 90:10 percent split

5 Conclusion

In this paper, the authors concluded the results of tree-based algorithms under the classification category of web mining using the data mining tool WEKA. The authors have summarized the result of different tree-based algorithms to check these for maximum efficiency on the web server log dataset. Authors have performed the analysis using various algorithms such as treeJ48, random tree, random forest, and REP tree and observed the results. Tree-based J48 gives maximum accuracy of 98.2072% with 90:10 split ratio, random tree gives maximum accuracy of 97.9084% with 80:20 split ratio, random forest gives maximum accuracy of 98.7052% with 80:20 split ratio, and REP tree algorithm gives maximum accuracy of 98.0080% with 90:10 split ratio. With this experimentation of classifying the data, the authors have concluded that random forest algorithm gives the best result with a classification accuracy of 98.7052% correctly classified instances when the percentage split ratio is kept at 80% while REPTree gives minimum accuracy by correctly classifying 97.3108% instances at the same percentage split ratio, i.e., 80%. In conclusion, the authors discovered that random forest gives better results between tree algorithms section while using the WEKA tool while REP tree gives the lowest classified instances when we compared over different percentage split ratio.

References

1. M.F. Arlitt, C.L. Williamson, Internet web servers: workload characterization and performance implications. *IEEE/ACM Trans. Netw.* **5**(5), 631–645 (1997)
2. S. Miller, M. Crystal, H. Fox, L. Ramshaw, R. Schwartz, R. Stone, R. Weischedel, BBN: Description of the SIFT system as used for MUC-7, in *Seventh Message Understanding Conference (MUC-7): Proceedings of a Conference Held in Fairfax, Virginia, April 29-May 1, 1998*
3. B. Jarboui, M. Cheikh, P. Siarry, A. Rebai, Combinatorial particle swarm optimization (CPSO) for partitional clustering problem. *Appl. Math. Comput.* **192**(2), 337–345 (2007)
4. C.H. Lee, Y.H. Fu, Two levels of prediction model for user's browsing behavior, in *Proceedings of the International MultiConference of Engineers and Computer Scientists*, vol. 1 (2008)
5. T. Hussain, S. Asghar, S. Fong, A hierarchical cluster based preprocessing methodology for Web Usage Mining, in *2010 6th International Conference on Advanced Information Management and Service (IMS)* (IEEE, 2010), pp. 472–477
6. R. Khanchana, M. Punithavalli, Web usage mining for predicting users' browsing behaviors by using FPCM clustering. *Int. J. Eng. Technol.* **3**(5), 491 (2011)
7. R. Mittal, Multivariate regression predictive modelling in analysing student performance: a data mining approach. *J. Comput. Theor. Nanosci.* **16**(10), 4362–4366 (2019)
8. R. Mittal, Prediction of heart diseases using Hayes process macro serial mediation model 6. *Indian J. Public Health Res. Dev.* **10**(10), 538–543 (2019)
9. R. Mittal, Identification of salient attributes in social network: a data mining approach, in *International Conference on Recent Developments in Science, Engineering and Technology* (Springer, Singapore, 2019), pp. 173–185
10. V. Shrivastava, N. Gupta, Performance improvement of web usage mining by using learning based k-mean clustering. *Int. J. Comput. Sci. Its Appl.* ISSN 2250-3765

11. T.P. Hong, M.J. Chiang, S.L. Wang, Mining weighted browsing patterns with linguistic minimum supports, in *IEEE International Conference on Systems, Man and Cybernetics*, vol. 4 (IEEE, 2002), 5 p
12. R. Cooley, B. Mobasher, J. Srivastava, Data preparation for mining World Wide Web browsing patterns. *Knowl. Inf. Syst.* **1**(1), 5–32 (1999)
13. H. Chen, M. Chau, Web mining: machine learning for web applications. *Ann. Rev. Inf. Sci. Technol.* **38**(1), 289–329 (2004)
14. G. Stumme, A. Hotho, B. Berendt, Semantic web mining: state of the art and future directions. *J. Web Semant.* **4**(2), 124–143 (2006)
15. D.S. Babu, S.A. Nabi, M.A. Ali, Y. Raju, Web usage mining: a research concept of webmining. *Int. J. Comput. Sci. Inf. Technol.* **2**(5) (2011)
16. R.K. Malviya, M.C. Malviya, V.K. Soni, R. Joshi, P. Purohit, Survey of web usage mining. *Int. J. Comput. Sci. Technol. (IJCST)* **2**(3) (2011). ISSN:2229–4333(Print)| ISSN: 0976–841(Online)
17. S. Suharjito, D. Diana, H. Herianto, Implementation of classification technique in web usage mining of banking company, in *2016 International Seminar on Intelligent Technology and Its Applications (ISITIA)*, pp. 211–218 (2016)
18. S. Jain, R. Rawat, B. Bhandari, A survey paper on techniques and applications of web usage mining, in *2017 International Conference on Emerging Trends in Computing and Communication Technologies (ICETCCT)* (IEEE, 2017), pp. 1–6
19. S.P. Singh, Analysis of web site using web log expert tool based on web data mining, in *2017 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS)* (IEEE, 2017), pp. 1–5

Classification and Analysis of Water Quality Using Machine Learning Algorithms



Amandeep Kaur, Meenu Khurana, Preetinder Kaur, and Manpreet Kaur

Abstract This research work revolves around the development of supervised machine-learning models that can automatically classify the quality of river water. The original dataset is transformed and binned into two (swimming, boating) class types. Using this data, an exploratory study of the machine-learning models has been done as to construct a generic water quality classifier. At the same time, it was found that there is an imbalance in dataset. To overcome this problem, SMOTE algorithm was applied and the exploratory analysis of the machine-learning algorithms was done. The performance analysis of the various classifier algorithms show initially that there is a need for customization of the machine learning so that generalized classifier can be build. Deeper analysis of the study showed that the correlation-based features are helpful in RF and CART. At the same time, the PCA data projection shows a higher level of accuracy (0.989) with the neural network algorithm.

Keywords River water quality · Neural network · Random forest · PCA algorithm · Classification

A. Kaur (✉) · M. Khurana · M. Kaur

Chitkara University Institute of Engineering and Technology, Chitkara University, Rajpura, Punjab, India

e-mail: amandeep@chitkara.edu.in

M. Khurana

e-mail: meenu.khurana@chitkara.edu.in

M. Kaur

e-mail: manpreetk.arora@chitkara.edu.in

P. Kaur

Western Sydney University International College, Sydney, Australia

e-mail: p.kaur2@westernsydney.edu.au

1 Introduction

The quality of the river water can be mainly judged by the three parameters [1]. These are named as biological, chemical, and physical parameters. The ratio of living creatures such as algae, bacteria, fungi, etc., indicates the health of the river water body. The ratio of these organisms in the rivers and other water bodies depends on the dissolved nutrition and oxygen content in the water [2]. A higher level of chemical and pollutants may lead to ecological imbalance [3]. Hence, it is essential to monitor the health of rivers and other water bodies especially, when the human activity around the water bodies is high [1]. Physical parameters such as color, temperature, turbidity, specific conductivity, pH value, and oxygen play an important role in determining the chemical and physical parameters of water. India is running multiple such programs to evaluate the quality of rivers such as Ganga [4–6]. However, all these efforts of monitoring are expensive and difficult to thoroughly implement due to the vast area it needs to cover. Researchers have been applying their mind to overcome these problems by using technological advancements [7, 8]. Many of these researches are trying to build generic models of predictions so that, same algorithms can be used to assess the quality of other river bodies of the same category. It would be possible only if machine learning or deep learning models are applied to this problem [9]. Hence, this research work primarily focuses on building an automated process to overcome such challenges [10–12].

2 Review

Almost all governments in the world assess water quality as per the usage. In this section, a tabular summary of the contemporary work done in the context of water quality assessment and automation of the process is tabulated in Table 1, after consulting many articles from high impact journals.

It was found that the manual process of analysis of water quality is tedious, expensive, and time-consuming. With the help of a machine-learning algorithm, a fully automated system is the need of the hour. However, the first step is to select relevant factors/variables that can accurately map the quality of classes (Swimming and Boating) of the river water. The second challenge that can be noted is about the availability of data. It was found that water quality data related to most of Indian rivers is not available for research purposes. Publicly available data [22] has unbalanced instances [23, 24] of good and bad quality. Hence, for any fully automated classifier to be successful, there is a need to handle the unbalanced instance issue. Lastly, the process of automation must be repeatable and should be of high performance. The next section, explains all the steps and process used to overcome these challenges.

Table 1 Water quality assessment state of art

Ref	Authors/water body	Problem	Algorithm(s)	Key finding/learning outcomes
[13]	All water bodies under the monitoring of World Health Organization (WHO)	Report on types water quality monitoring adopted by WHO and other international water monitoring agencies	Descriptive account of all types of water quality monitoring methods and procedures	Learned about the variables, methods, pollutants, operations, Labs, and data processing required
[14]	D. Ömer Faruk/Büyükk Menderes River, Turkey	water quality prediction	ARIMA + neural network	Hybrid approach works well
[15]	K. H. Reckhow/Neuse River estuary, Canada	Water quality prediction	Bayesian probability networks	For better water quality prediction, Bayes nets have a much wider scope
[16]	K. P. Singh, N. Basant, and S. Gupta	Classify the sampling sites (spatial) and months (temporal) to group + water quality prediction	Support vector machine	Non-linear methods such as SVM are better than linear methods
[17]	S. Liu, H. Tai, Q. Ding, L. Xu, and Y. Wei	Aquaculture water quality prediction	Support vector machine + genetic algorithm optimization	The hybrid approach is better than applying SVM alone
[18]	G. Tan, J. Yan, C. Gao, and S. Yang	Prediction of water quality	least square SVM	LS SVM is highly accurate
[19]	A. Sankar and P. Pandey/River Water, Yamuna, Mathura, UP	Prediction of water quality	ANN	ANN is accurate
[20]	B. Amit and E. Bril	Prediction of water quality	Statistical approach	The novel method used in this is the hydraulic model and it has higher accuracy
[21]	E. Mcbean/Kali, River	Prediction of water quality	Statistical approach	This approach has given better view to the problems and provides better perspective to solve them
[14]	D. Sharma and A. Kansal/Yamuna	Prediction of water quality	Statistical approach	This approach has given better view to the problems and provides better perspective to solve them

(continued)

Table 1 (continued)

Ref	Authors/water body	Problem	Algorithm(s)	Key finding/learning outcomes
[11]	E. Bril	Prediction of water quality	Radial basis function	The novel method ensures zero false negative events

3 Methods and Material

This research work initiates with the Charles River dataset [22] and then proceeds further with the exploration of the dataset characteristics that are required for the research work. The section also discusses the approaches engaged in identifying the procedures to solve the issues regarding unbalanced data. The last section explains the automation of the process to solve the problem stated earlier in the paper. Figure 1 shows the work flow of the research work. The dataset consists of four-year time series dataset of the Charles River in Massachusetts collected from the United States Environmental Protection Agency (USEPA) website [25]. The dataset is updated every 15 min during important periods of monitoring. It collects the data related to the temperature, dissolved oxygen, pH, specific conductance, turbidity, chlorophyll, and phycocyanin. All of these variables are important to check the quality of river water suitable for swimming or boating. For subjecting the dataset to machine-learning models, the dataset had to be transformed into discrete-time series. Hence, by using water safety limits given by WHO, the dataset was divided into fit and unfit class for swimming and boating. The dataset composition analysis showed that the division of the dataset led to the problem of an unbalanced dataset. Hence, the dataset was further

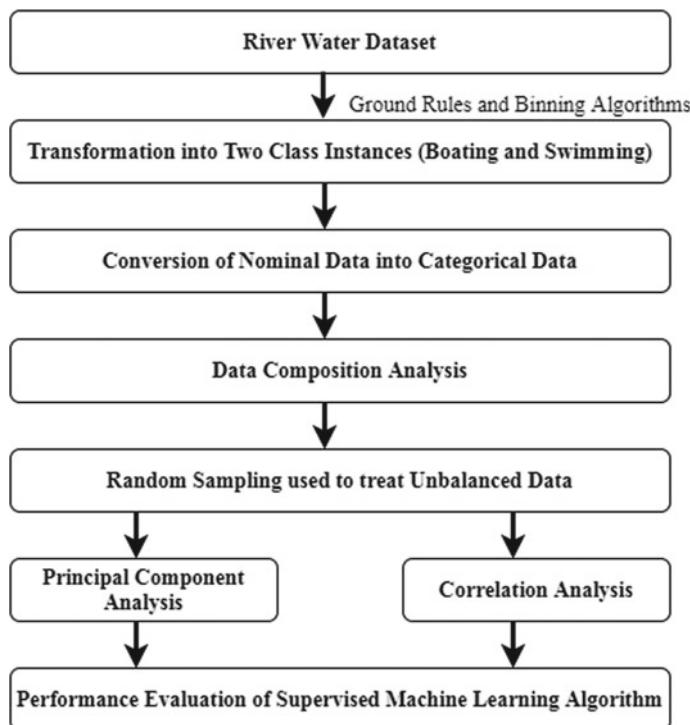


Fig. 1 Work flow of the research

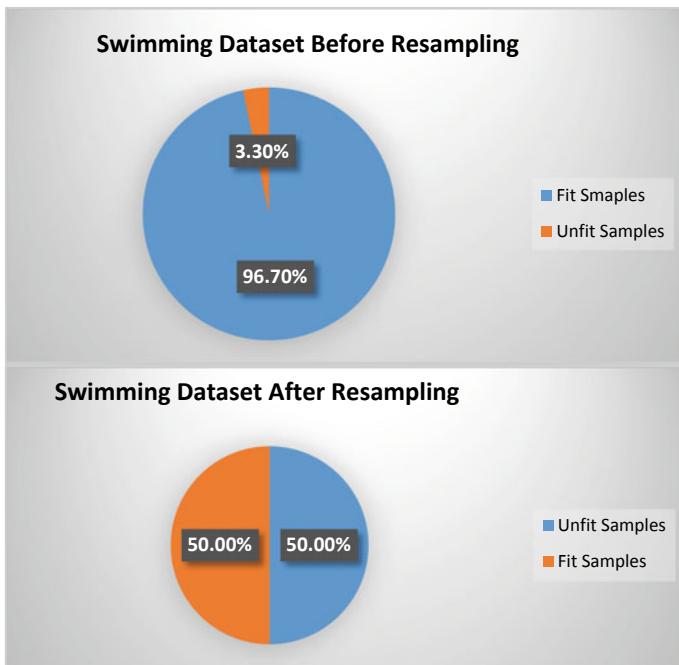


Fig. 2 Ratio of classes after SMOTE [26] algorithm

balanced between the minority and majority class by using SMOTE algorithm, as shown in Fig. 2.

4 Results and Discussion

Technically, adding more attributes can increase the power of an algorithm to categorize. At the same time, it can be observed that techniques such as random forest/decision tree suffer from issues of collinearity. Due to the high degree of connection/association between the participating parameters, the classifier may produce over-fitted output. To overcome this problem, there is normally a need to drop highly correlated attributes. Principal component analysis (PCA) is one such method that produces an orthogonal dataset [27]. Hence, in this section, a comparative analysis has been done between the correlation-based feature selection and PCA-based features to finally selecting the best algorithm for the said tasks.

4.1 Correlation Based Feature Selection

For choosing the best hypothesis, between the correlation-based feature elimination and selection, multiple experiments were conducted at different levels of correlation values on all the variables. For maintaining the brevity and readability of the paper, only the results of best-performing algorithms have been presented. The evaluation was done between six machine-learning algorithms (Neural Network, Random Forest, KNN, CART, NB, SVM) [28, 29], as shown in Tables 2 and 3.

It was found that the maximum elimination of features happened at the threshold level of 0.70. It meant that machine-learning algorithms had the least overhead. At this level, the random forest algorithm achieved maximum balanced accuracy.

The standard deviation was minimal and it was inferred that, at this level, the random forest algorithm was quite stable.

From this level of correlation analysis, as mentioned in Fig. 3, it may be concluded that the neural network algorithm was best suited for this purpose. The values of recall and precision reinforced this inference. The correlation threshold of 0.7 is valid in the case of the boating dataset. The evaluation of the machine-learning algorithms on the boating dataset showed that the performance between the random forest and KNN algorithm is quite comparable.

All these results (recall, precision, accuracy, $F1$) were computed on the level of macro and validated with the 10 K-fold method. In both cases, the random forest algorithm performed well, primarily due to similarity in data point values. However,

Table 2 Swimming dataset: Random forest performance at correlation threshold = 0.70

RF	Precision	Recall	$F1$ -score	Average accuracy	Standard deviation	Balance accuracy	Threshold
Micro average	0.5	0.5	0.5	0.499281	0.001802	0.499691	0.70
Macro average	0.34	0.5	0.49	0.499281	0.001802	0.499691	0.70
Weighted average	0.34	0.5	0.49	0.499281	0.001802	0.499691	0.70

Table 3 Boating dataset: Random forest performance at correlation threshold = 0.75

RF	Precision	Recall	$F1$ -score	Average accuracy	Standard deviation	Balance accuracy	Threshold
Micro average	0.5	0.5	0.5	0.50136	0.000542	0.5	0.75
Macro average	0.25	0.5	0.33	0.50136	0.000542	0.5	0.75
Weighted average	0.25	0.5	0.33	0.50136	0.000542	0.5	0.75

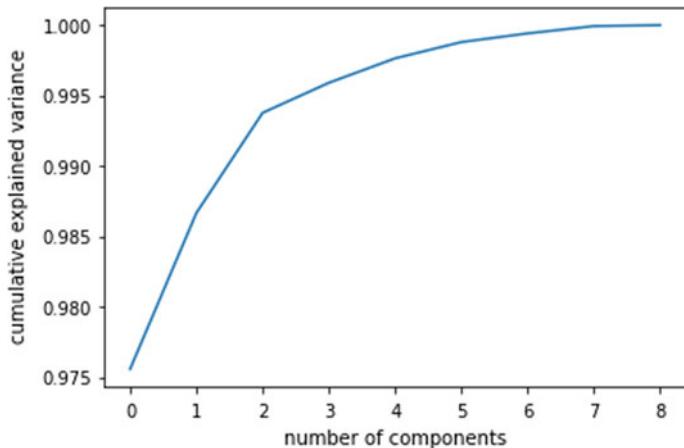


Fig. 3 Explained variance of full dataset (boating and swimming)

on an average basis, the performance matrix values acquired were lower in case of boating dataset as compared to the swimming dataset.

4.2 PCA Based Features

The application of principal component analysis not only used for the selection of features, but also used for selection of a matrix that is orthogonal. An orthogonal dataset means that all the columns of the matrix have the lowest order of correlation. In other words, a projected image or a reduced dimension is created and based on the explained variance matrix, the components (features) are selected.

The box plot in Fig. 4a shows the evaluation of all the algorithms using the K-fold method [30]. It can be observed that the average values of each performance

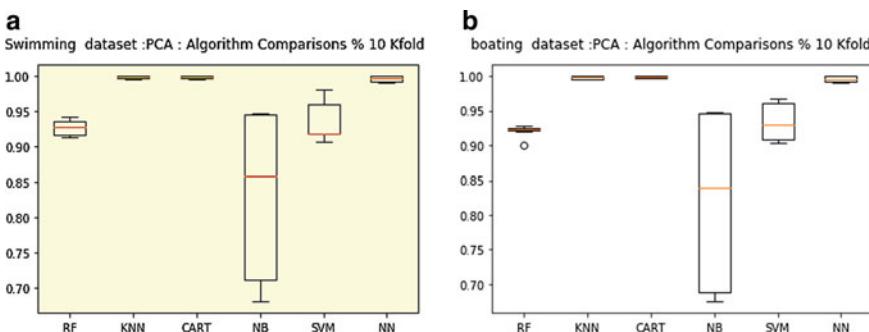


Fig. 4 PCA-based performance evaluation of algorithms for swimming and boating dataset

matrix increased a bit, as compared to the correlation-based feature selection. Due to this fact, the neural network algorithm performed better than the decision tree-based (CART & RF) algorithms, as shown in Fig. 4b. The size of the box plot of the neural network algorithm also shows that the neural network method had the lowest standard deviation; hence, it had maximum stability. All this research work was carried out using Google colab environment and python libraries. The next section concludes the finding this research work.

5 Conclusion and Future Scope

In this research work, the prediction and classification of river water quality were done. The application of SMOTE algorithm on both the datasets produced unbiased results in the evaluation of all algorithms. It removed the skewness in the dataset. The process of transforming the original dataset into a categorical dataset helped to build an automated system that could do water assessment at boating level and swimming level. This was possible owing to the fact that acceptance criterion of water quality for boating is less stringent as compared to that applicable for swimming. It is also found that the correlation-based feature engineering is useful to produce a high order of accuracy in most of the algorithms. The correlation-based evaluation shows that RF and CART are the best-suited algorithms for this research work. However, the PCA-based evaluation shows a higher level of accuracy with the neural network algorithm. There is an urgent need in India to improve the water quality of all its water bodies. The main limitation of this work is the non-availability of public datasets on the water of Indian rivers. Hence, for future directions, it is suggested that this work may extended on the India's river datasets.

References

1. R. Das Kangabam, M. Govindaraju, Anthropogenic activity-induced water quality degradation in the Loktak lake, a Ramsar site in the Indo-Burma biodiversity hotspot. *Environ. Technol. (United Kingdom)* **40**(17), 2232–2241 (2019)
2. T. Poonam, B. Tanushree, C. Sukalyan, Water quality indices—important tools for water quality assessment: a review. *Int. J. Adv. Chem.* **1**(1), 15–28 (2013)
3. D.S. Bhargava, Water quality variations and control technology of Yamuna River. *Environ. Pollution. Ser. A, Ecol. Biol.* **37**(4), 355–376 (1985)
4. D. Swaroop Bhargava, Use of water quality index for river classification and zoning of Ganga river. *Environ. Pollution. Ser. B, Chem. Phys.* **6**(1), 51–67 (1983)
5. S. Bhatnagar, Repercussions of tourism on water quality of River Ganga in Lower Himalayas, pp. 90–103 (2018)
6. R.K. Srivastava, A.K. Sinha, D.P. Pande, K.P. Singh, H. Chandra, Water quality of the River Ganga at Phaphamau (Allahabad)—effect of mass bathing during Mahakumbh. *Environ. Toxicol. Water Qual.* **11**(1), 1–5 (1996)
7. R. Jeyakumar, S. Parimalarenganayaki, L. Elango, River bank filtration for natural treatment of water in India: a review. *Int. J. Civ. Eng. Technol.* **8**(8), 1203–1212 (2017)

8. R. Paliwal, P. Sharma, A. Kansal, Water quality modelling of the river Yamuna (India) using QUAL2E-UNCAS. *J. Environ. Manage.* **83**(2), 131–144 (2007)
9. X. Wang, F. Zhang, J. Ding, Evaluation of water quality based on a machine learning algorithm and water quality index for the Ebinur Lake Watershed, China. *Sci. Rep.* (2017)
10. K. Chen et al., Comparative analysis of surface water quality prediction performance and identification of key water parameters using different machine learning models based on big data. *Water Res.* (2020)
11. Z.Y. Wu, M. El-Maghraby, S. Pathak, Applications of deep learning for smart water networks. *Procedia Eng.* (2015)
12. A. Najah Ahmed et al., Machine learning methods for better water quality prediction. *J. Hydrol.* (2019)
13. V. Kimstach, R. Helmer, Chapter 2—Strategies for water quality assessment, vol. 5 (1996)
14. D. Ömer Faruk, A hybrid neural network and ARIMA model for water quality time series prediction. *Eng. Appl. Artif. Intell.* **23**(4), 586–594 (2010)
15. K.H. Reckhow, Water quality prediction and probability network models. *Can. J. Fish. Aquat. Sci.* **56**(7), 1150–1158 (2011)
16. K.P. Singh, N. Basant, S. Gupta, Support vector machines in water quality management. *Anal. Chim. Acta* **703**, 152–162 (2011)
17. D. Li, H. Tai, L. Xu, S. Liu, Q. Ding, Y. Wei, A hybrid approach of support vector regression with genetic algorithm optimization for aquaculture water quality prediction. *Math. Comput. Model.* **58**(3–4), 458–465 (2011)
18. G. Tan, J. Yan, C. Gao, S. Yang, Prediction of water quality time series data based on least squares support vector machine. *Procedia Eng.* **31**, 1194–1199 (2012)
19. A. Sarkar, P. Pandey, River water quality modelling using artificial neural network technique. *Aquat. Procedia* **4**, 1070–1077 (2015)
20. B. Amit, E. Brill, New approach for estimation of detention time and prediction of quality in water networks. *Water Qual. Res. J.* **53**(2), 72–85 (2018)
21. N.C. Ghosh, E.A. Mcbean, Water quality modeling of the Kali River, India. *Water. Air. Soil Pollut.* **102**(1–2), 91–103 (1998)
22. Puneet Arora, Charles River data water quality sensor data. Mendeley Data, p. 123 (2020)
23. G. E. A. P. A. Batista, R. C. Prati, and M. C. Monard, A study of the behavior of several methods for balancing machine learning training data. *ACM SIGKDD Explor. Newsl.* (2004)
24. J.M. Johnson, T.M. Khoshgoftaar, Survey on deep learning with class imbalance. *J. Big Data* (2019)
25. E. Gov, Reports Charles River. epa.gov (2020). [Online]. Available: <https://www.epa.gov/charlesriver/reports-and-documents-charles-river>
26. J. Mathew, C.K. Pang, M. Luo, W.H. Leong, Classification of imbalanced data by oversampling in kernel space of support vector machines. *IEEE Trans. Neural Networks Learn. Syst.* (2018)
27. S. Karamizadeh, S.M. Abdullah, A.A. Manaf, M. Zamani, A. Hooman, An overview of principal component analysis. *J. Signal Inf. Process.* (2013)
28. S. Shalev-Shwartz, S. Ben-David, Understanding machine learning: from theory to algorithms (2013)
29. Q. He, N. Li, W.J. Luo, Z.Z. Shi, A survey of machine learning algorithms for big data. *Moshi Shibiai yu Rengong Zhineng/Pattern Recognition and Artificial Intelligence* (2014)
30. J. Brownlee, A Gentle Introduction to k-fold Cross-Validation. machinelearningmastery.com (2019)

Extracellular Conductivity and Nerve Signal Propagation: An Analytical Study



Satyabrat Malla Bujar Baruah, Biswajit Das, and Soumik Roy

Abstract The nerve membrane is one of the most researched area in understanding the flow of signal in nervous system. Extensive studies have already been undertaken in this field, and more complex issues are yet to be addressed. The nerve fiber facilitates in the flow of information; thus in order to study the different aspects of signal propagation, a vivid understanding of this structure is necessary. In this study, the effect of loss or attenuation of the signal as it traverse down the nerve is analyzed. The nerve fiber consists of the internal and the external cytoplasmic fluids; here, study is conducted on how the extracellular cytoplasm or the extracellular space influences the nerve signal propagation as it travels down a nerve fiber. The standard cable equation depicting the nerve as an equivalent electrical circuit is used to carry out the study. The analytical study highlights some important aspect of ectoplasm or the extracellular space and its role in signal propagation.

Keywords Nerve conduction · Extracellular conductivity · Fiber diameter · Cable · Extracellular · Ectoplasm · Endoplasm

1 Introduction

The nerve fiber is surrounded by the extracellular space around it which contains different ions responsible for the generation of action potential. The extracellular space is of utmost importance as it plays a very significant role in signal propagation [1–6]. A nerve signal may undergo leakage as the nerve membrane is not a perfect

S. M. B. Baruah · B. Das (✉) · S. Roy

Department of ECE, Tezpur University, Tezpur University Campus, Napaam, Sonitpur, Tezpur, Assam 784028, India

e-mail: das.biswa700@gmail.com

S. M. B. Baruah

e-mail: baruah.satyabrat@gmail.com

S. Roy

e-mail: oumik@tezu.ernet.in

insulator [7, 8]; hence, understanding these losses is of utmost importance in understanding the overall morphology of the fiber. The current study emphasized on how the nerve signal undergoes attenuation due to the effect of the extracellular space around it. Studies have shown that if the extracellular space decreases, then the resistance increases which in turn causes the mobility of ions to decrease hence causing less signal attenuation [9]. But when the extracellular space increases, the resistance of the spacing gets decrease which in turn causes the mobility of ions to increase causing more attenuation of the signal. The electrical parameters concerning a neuron are well expressed by the cable equation [10–13]; it shows the flow of currents and voltages between the nearby nerve fibers which help in understanding the functionalities of signal propagation between the neurons. In this study, analysis is carried out on the proposed cable model to understand the different aspects of signal attenuation.

2 Related Work

Bedard and Destexhe [2] have shown that voltage attenuation is highly observed in the presence of ionic diffusion in the nerve membrane; similar to [2] the nerve model under consideration in this study is a conventional open circuit (FO) model, where only forward transfer of ions between adjacent neurons is allowed. They suggest that larger the extracellular space more is the attenuation of signals leading to more compact dendrites [2]. The nerve fiber can be either uniformed, tapered [14, 15], or flared; in this study, these nerve fibers are taken into consideration.

3 Proposed Model

Cable equation or the cable model is considered to show the electrical properties of a nerve fiber as shown in Fig. 1. The dotted line represents the boundary between outside and the inside part of the fiber; each of the tank circuit is the equivalent Hodgkin and Huxley (HH) [16–18] model showing the nerve membrane by its equivalent electrical circuits.

The internal and external membrane potentials can be shown as:

$$\frac{\partial V_i}{\partial x} = -R_i I_i(x) \text{ and } \frac{\partial V_e}{\partial x} = R_e I_e(x) \quad (1)$$

where $R_i = \frac{4r_i}{\pi D_i^2} \Delta x$ and $R_e = \frac{4r_e}{\pi D_e^2} \Delta x$. The transmembrane voltage can be given as:

$$\frac{\partial V_m}{\partial x} = -I(R_e + R_i) \quad (2)$$

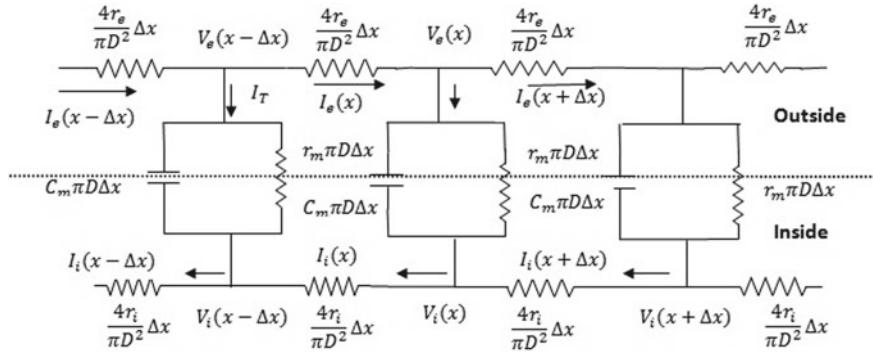


Fig. 1 Proposed model using cable equation [10–13]

Here, $(R_e + R_i)$ is the total axial resistance. Considering the current mobility inside and outside of the membrane to be same, i.e., $I_i = I_e = I$ which is the axial current; therefore, the external and the internal currents become equal to the transmembrane current I_T as:

$$\frac{\partial I_e}{\partial x} = \frac{\partial I_i}{\partial x} = \frac{\partial I}{\partial x} = I_T \quad (3)$$

Differentiating Eq. (2) with respect to x and substituting the values of Eq. (3), the resultant wave equation thus obtained is:

$$\frac{d^2 V_m}{dx^2} + C_m(R_e + R_i) \frac{dV_m}{dt} + \frac{(r_e + r_i)}{r_m}(V_m + E_l) = 0 \quad (4)$$

Here, $C_m = c_m \pi D l$ and $R_m = r_m \pi D l$ are the equivalent total membrane capacitance and resistance, respectively. Dividing the nerve fiber into a number of equipotential cylinder [19], i.e., $\frac{dV}{dx} = 0$, the solution for the wave equation becomes:

$$\frac{dV_m}{dt} = \left[\frac{V_{in} - V_m}{C_m(R_e + R_i)^2} - \frac{(V_m - E_l)}{R_m C_m} \right] \quad (5)$$

Here, the term $C_m(r_e + r_i)^2$ can be further expanded as:

$$C_m(R_e + R_i)^2 = \frac{16c_m l^2}{\pi} \left(\frac{r_i^2 l}{D_i^3} + \frac{r_e^2 l D_i}{D_e^4} + \frac{2r_e r_i}{D_e^2 D_i} \right) \quad (6)$$

4 Results and Discussion

4.1 Simulation Considerations

Constant extracellular space = $10 \mu\text{m}$, constant diameter = $2 \mu\text{m}$, $r_m = 1 \Omega/\text{mm}^2$, $c_m = 1 \mu\text{F}/\text{mm}^2$, length of fiber are varied from 5 to $200 \mu\text{m}$. $V_r = -65 \text{ mV}$, $E_l = 10.6 \text{ mV}$, and the simulation is done using PYTHON.

4.2 Fiber with Constant Extracellular Space and Varying Diameter

For a tapered and a flared fiber having a constant extracellular space around it; the resultant plot is obtained by using Eq. (5) and plotting it on python, which is provided in Fig. 2a, b, respectively. It is seen here that the impact of extracellular space on the signal attenuation is similar in both the kind of nerve fibers.

Figure 2a shows a tapered fiber having constant extracellular space, and Fig. 2b shows a flared fiber having constant extracellular space. $V_1(t)$ is the initial action potential; $V_{out1}(t)$, $V_{out2}(t)$, $V_{out3}(t)$, $V_{out4}(t)$, $V_{out5}(t)$ are taken at a distance of 5, 37.5, 78.125, 151.25, 183.75 μm , respectively; the diameter is linearly increasing with a slope factor of 0.002 for Fig. 2a and is linearly decreasing with a slope factor of 0.002 for Fig. 2b.

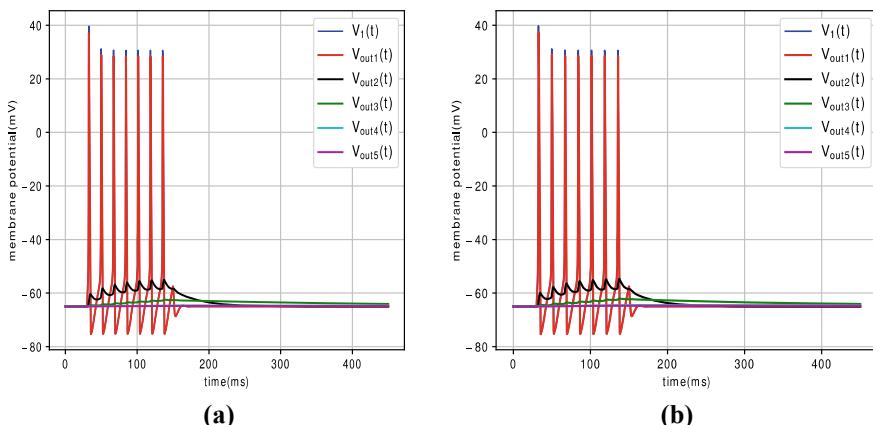


Fig. 2 **a** Tapered fiber having a constant extracellular space. **b** Flared fiber having a constant extracellular space

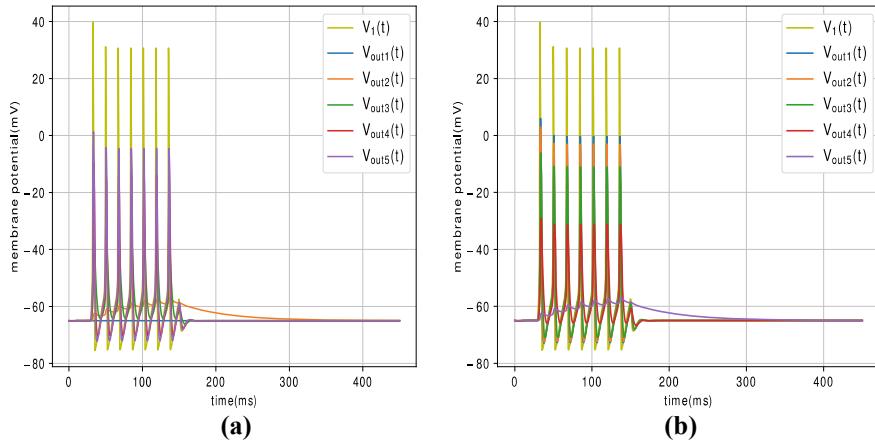


Fig. 3 **a** Uniform fiber with increasing extracellular space. **b** Uniform fiber with decreasing extracellular space

4.3 Fiber with Constant Diameter and Varying Extracellular Space

Considering a fiber of constant diameter and increasing extracellular space and using Eq. (5) and plotting it on python, the resultant plot obtained is provided in Fig. 3a. Similarly, for a fiber of constant diameter decreasing extracellular space around it, and using Eq. (5) and plotting it on python, the resultant plot obtained is provided in Fig. 3b. In both the cases, the spacing around the fiber is considered within a range of 1.5–15 μm .

Figure 3a shows a uniform fiber with increasing extracellular space, and Fig. 3b shows uniform fiber with decreasing extracellular space; $V_1(t)$ is the initial action potential, and $V_{\text{out}1}(t)$, $V_{\text{out}2}(t)$, $V_{\text{out}3}(t)$, $V_{\text{out}4}(t)$, $V_{\text{out}5}(t)$ are taken at a distance of 70 μm . The point of measurement for Fig. 3a is surrounded with extracellular spaces of 0.15, 0.95, 1.96, 2.97, 4.19 μm , respectively, and for Fig. 3b 4.19, 2.97, 1.96, 0.95, 0.15 μm , respectively.

From the results obtained in Fig. 2a, b, it can be concluded that if the extracellular space is of constant nature, and the diameter of the fiber is a tapered or a flared one; the variation in shape of the fiber does not affect the strength of the signal much. Now, for a uniform diameter fiber and having a increasing extracellular space around it, it is seen from Fig. 3a that the signal undergoes more attenuation. So, a uniform fiber enclosed within a larger extracellular space gives the ions more advantages in terms of mobility to the extracellular space and hence results in a decreased axial resistance [9] against the flow of signal, thereby causing more attenuation, whereas, if the extracellular space decreases, then mobility of ions reduces significantly to the extracellular space causing the signal to get less attenuated as it travels down the nerve fiber which is seen in Fig. 3b.

Table 1 Notation table for different parameters

S. No.	Notation	Interpretation
1	c_m and r_m	Characteristic membrane capacitance and characteristic membrane resistance
2	C_m and R_m	Equivalent total membrane capacitance and equivalent total membrane resistance
3	I and $(R_e + R_i)$	Axial current and axial resistance
4	I_T and V_m	Transmembrane current and transmembrane voltage
5	I_i and I_e	Internal and external current
6	D_i and D_e	Internal and external diameter of fiber
7	R_i and R_e	Volumetric axial resistances

5 Conclusion and Future Scope

From the results obtained, it can be concluded that if the extracellular space is decreasing in nature, then the signal undergoes less attenuation proportionate to the decrease in the extracellular space due to increased resistivity [9] against the axial flow. Similarly, for an extracellular space which is increasing in nature, the signal undergoes more attenuation compared to the decreasing counterpart which in turn leads to more compact dendrites as suggested in the standard literature [2]. For a constant extracellular space, the loss in signal is approximately equal whether the fiber is a tapered type or a flared type; thus the results obtained show the significance of the extracellular space in nerve conduction when it is of varying nature. The current model can also be extended for a myelinated fiber, and the results can be compared with the current study. Also, the effect of extracellular space on velocity of impulse can also be investigated further (Table 1).

References

1. C.S. Barros, S.J. Franco, U. Müller, Extracellular matrix: functions in the nervous system. *Cold Spring Harbor Perspect. Biol.* **3**(1), a005108 (2011)
2. C. Bédard, A. Destexhe, Generalized cable theory for neurons in complex and heterogeneous media. *Phys. Rev. E* **88**(2), 022709 (2013)
3. C. Nicholson, S. Hrabětová, Brain extracellular space: the final frontier of neuroscience. *Biophys. J.* **113**(10), 2133–2142 (2017)
4. B.A. MacVicar, H.B. Choi, Astrocytes provide metabolic support for neuronal synaptic function in response to extracellular K⁺. *Neurochem. Res.* **42**(9), 2588–2594 (2017)
5. J.M. Resnick, G.E. O'Brien, J.T. Rubinstein, Simulated auditory nerve axon demyelination alters sensitivity and response timing to extracellular stimulation. *Hearing Res.* **361**, 121–137(2018)
6. C. Gold, et al., On the origin of the extracellular action potential waveform: a modeling study. *J. Neurophysiol.* **95**(5), 3113–3128 (2006)

7. A.L. Hodgkin, A.F. Huxley, Potassium leakage from an active nerve fibre. *J. Physiol.* **106**(3), 341–367 (1947)
8. B. Katz, R. Miledi, Transmitter leakage from motor nerve endings. *Proc. R. Soc. Lond. Ser. B. Biol. Sci.* **196**(1122), 59–72 (1977)
9. J. Clark, R. Plonsey, The extracellular potential field of the single active nerve fiber in a volume conductor. *Biophys. J.* **8**(7), 842 (1968)
10. C. Koch, Cable theory in neurons with active, linearized membranes. *Biol. Cybern.* **50**(1), 15–33 (1984)
11. I. Tasaki, G. Matsumoto, On the cable theory of nerve conduction. *Bull. Math. Biol.* **64**(6), 1069 (2002)
12. W. Rall, Core conductor theory and cable properties of neurons. *Comprehensive Physiol.*, 39–97 (2011)
13. W. Rall, Electrophysiology of a dendritic neuron model. *Biophys. J.* **2**(2), 145–167 (1962)
14. R. Hodes, Linear relationship between fiber diameter and velocity of conduction in giant axon of squid. *J. Neurophysiol.* **16**(2), 145–154 (1953)
15. M. Fujii, et al., Evidence for the tapering of nerve axons. *Showa Univ. J. Medi. Sci.* **6**(2), 179–184 (1994)
16. A.L. Hodgkin, A.F. Huxley, A quantitative description of membrane current and its application to conduction and excitation in nerve. *J. Physiol.* **117**(4), 500–544 (1952)
17. A.L. Hodgkin, A.F. Huxley, Currents carried by sodium and potassium ions through the membrane of the giant axon of Loligo. *J. Physiol.* **116**(4), 449–472 (1952)
18. E.M. Izhikevich, Simple model of spiking neurons. *IEEE Trans. Neural Netw.* **14**(6), 1569–1572 (2003)
19. E. Syková, C. Nicholson, Diffusion in brain extracellular space. *Physiol. Rev.* **88**(4), 1277–1340 (2008)

EBN-Net: A Thermodynamical Approach to Power Estimation Using Energy-Based Multi-layer Perceptron Networks



Koustav Dutta, Rajarshi Pal, and Rajendra Prasad

Abstract Energy forecasting at the generation site in the smart grid furnishes extensive precious data, which further facilitates unprecedented options in smart power applications in emerging smart cities. At power generation facilities excess generated power requires energy storage, and standby power generation results in huge economic losses, whereas under the generation of power lead to fluctuations at the distribution side. In this paper, the power estimation at the combined cycle power plant has been done taking into consideration the various factors affecting the process with the help of multi-layer perceptron networks, but the prime focus of this paper is the presentation and explanation of the working of multi-layer perceptron networks in a way that has never been done before as per the literature survey conducted by us. The core working and the propagation of weights and related processes in the neural networks have been explained with the help of second law of thermodynamics and kinetic theory of gases taking into consideration the works of eminent scientists like Maxwell, Gibbs and Boltzmann. The explanations and mechanisms of the working of the various layers of the multi-layer perceptron network with the viewpoint of thermodynamics presented in this paper opens a new dimension in the interpretation of the working of neural networks which ultimately optimizes the ways of thinking and explainable of the different neurons and synapses of the neural network. The proposed algorithm helps to improve the forecasting with Multi-Layer Perceptron Networks (MLPN) essentially to predict power generation output using the minimum number of input variables. The results, which can be considered highly satisfactory, demonstrate the MLPN's prediction accuracy with a normalized root mean square error for all conditions of less than 5% and with practically no deviation. We demonstrate how beneficial matching of two already proven techniques can bring about spectacular results in energy generation prediction.

Keywords Energy-based model · Multi-layer perceptron network · Neural networks · Power estimation · Combined cycle power plants · Artificial neural networks · Boltzmann distribution · Entropy

K. Dutta · R. Pal · R. Prasad (✉)
KIIT Deemed to be University, Bhubaneswar, India
e-mail: rprasad@kiit.ac.in

1 Introduction

Multi-Layer Perceptron Networks (MLPNs) are non-linear and non-parametric data fitting methods, which can capture non-linear dependencies between parameters far better than the conventional linear models do [1]. For a set of input and output data, the corresponding tuning parameters (weights) of MLPN should be set appropriately, which is referred to as training methods [2]. In this case, the inputs are historical time series lags, and the single output is the one step ahead forecasted value. In this study, a three layer (one input layer, one output layer and one hidden layer) feed-forward MLPN is used, and the weight parameters are calculated by scaled conjugate gradient backpropagation algorithm in an iterative procedure [3]. For the sake of preventing the overfitting over the data, a percentage of data is reserved for validation. If the value of error does not decrease for validation data in a predefined number of consecutive iterations, then the termination criterion will be satisfied, and the training will stop. Eventually, the sigmoid transfer function is used in hidden layer neurons, and the linear transfer function is used in output neuron [4]. This paper focuses on the development and application of deep artificial neural networks for accurate estimation of the power generated from a combined cycle power plant taking into consideration various ambient variables. A Combined Cycle Power Plant (CCPP) is composed of steam turbines (ST), gas turbines (GT) and heat recovery steam generators as a single unit. In a CCPP, the electricity is produced by gas and steam turbines, which are included in the same cycle and is transferred from one turbine to another. The vacuum collected from the system has an effect on the steam turbine and the other three of the ambient variables affect the GT performance.

2 Proposed Algorithm

The paper deals with the problem of inaccurate and adulterated estimation of power generated from a Combined Cycle Power Plant (CCPP). This paper thus brings about a phenomenal work by estimating the power generated in an optimized and precise manner. The paper incorporates the dataset containing 9568 data points collected from a combined cycle power plant over 6 years (2006–2011) when the power plant was set to work with a full load [5]. Features consist of hourly average ambient variables such as temperature (T) in the range 1.81 °C and 37.11 °C, Ambient Pressure (AP) in the range 992.89–1033.30 millibar, Relative Humidity (RH) in the range 25.56–100.16% and Exhaust Vacuum (V) in the range 25.36–81.56 cm Hg to predict the net hourly electrical energy output (EP) of the plant. The architecture of the MLPN network is as follows:

1. The Multi-Layer Perceptron Networks (MLPN) developed in the paper consists of four input neurons in the input layer, i.e. Temperature, pressure, humidity and vacuum which plays the important role of estimation of the power generated from a CCPP.

2. The input layer is connected with the first hidden layer with the incorporation of various weights and biases. The first hidden layer consists of 6 neurons. In this process, the detailed features, inter-correlation and relationship among the different variables are learnt, and feature maps are developed. The weights used help in learning the importance and dependence of each of the variables used in the input layer, and the biases play a role as an adjustment factor, thus involving delay in the learning representation. The activation function used in this first hidden layer is Rectified Linear Unit (ReLU) [6] [ReLU: max (0, x)] which helps in extracting the non-linear features from the data. Since ReLU is zero for all negative inputs, it is likely for any given unit to not activate at all.
3. Next, the neurons of the first hidden layer are connected densely with the neurons of the second hidden layer, which consists of 6 neurons, thus forming a stacked representation of the deep neural network. Again, the weights and biases pass through the interconnection in between the two layers, thus helping to extract and understand the detailed inter-relationships in between the various factors which help in the learnt representation of the input vectors in the neural network. The activation function used in this layer is a Leaky ReLU [ReLU: max (0.01, x)] which doesn't get affected by the problem of the negative values of the input vectors, thus preventing the process of overfitting of the neural network. The design of the stacked multi-layer perceptron networks is given in Fig. 1.

This beauty of the paper lies in the fact that this paper brings about an interpretation and explanation of a neural network mechanism with the help of laws of thermodynamics and kinetic theory of gases taking into consideration the works of eminent scientists like Maxwell, Gibbs and Boltzmann. The explanations and mechanisms of the working of the various layers of the multi-layer perceptron network with the view point of thermodynamics presented in this paper opens a new dimension in the interpretation of the working of neural networks which ultimately optimizes the ways of thinking and explainables of the different neurons and synapses of the neural network.

The explanation of the multi-layer perceptron networks [7] used in the paper is done in the form of an Energy-Based Model (EBM) which is solely based upon the attainment of minimum energy state by gas molecules in surrounding space and is based upon the concept of kinetic theory of gases [8] and principle of thermodynamics

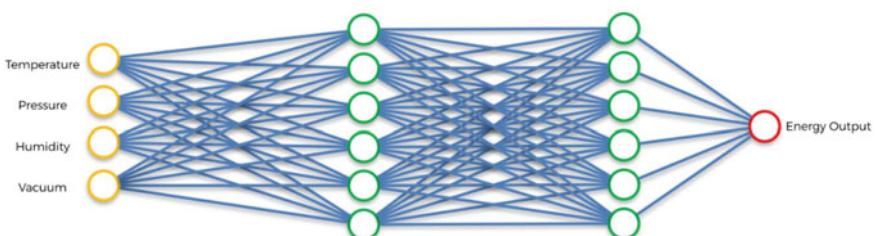


Fig. 1 Architecture of multi-layer perceptron network

[9] of all particles in the universe. In fact, the concept is based upon the fact of attaining zero entropy by all the particles in the universe by minimum energy state attainment. The concept can be explained by the Boltzmann distribution [10] for gaseous particles but is applicable for any particle present in any space in the universe, similarly, the neurons present in the neural network also attain the state of minimum energy in the network by optimization of the learnable parameters used in the neural network with the help of optimization functions and thus reaching the state of minimum energy which is mathematically analogous to point of Global Minima, i.e. Cost function (Loss function for complete training sample) becomes zero. Thus, the concept of the world and particles present in the universe can all be regarded as a neural network all of which are in the process of attainment of minimum entropy, thus reaching minimum energy state.

The equation used in the interpretation of the multi-layer perceptron networks as an Energy-Based Model (EBM) is explainable by creating an analogy and is given by the Boltzmann distribution in Eq. 1:

$$p_i = \left[\frac{e^{-\varepsilon_i/kT}}{\sum_{j=1}^M e^{-\varepsilon_j/kT}} \right] \quad (1)$$

where p_i is the probability of a system being in a certain state, ε_i is the energy of the system at state i , k is the Boltzmann constant (1.380649), and T is the temperature of the system.

The Boltzmann distribution tells that the probability of a system remaining in the state i is very low if the energy of the system is high, i.e. there is much randomness or excitation present in the system (The intensity of Brownian Motion [11] of the molecules of the system is high); thus, probability of a system being in a certain state is inversely proportional to the energy of the system in that state.

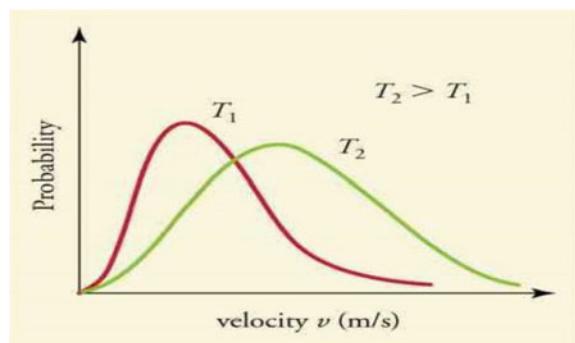
The excitation or randomness of the molecules of the system is also explained in terms of entropy of the molecules of the system. The velocity of the molecules of the system undergoing Brownian motion due to increase in temperature (entropy also increases) also increases to a large extent, and thus, the probability of the molecules of the system to remain in a particular state also decreases which can be shown by the Maxwell–Boltzmann distribution curve in Fig. 2.

In case of a multi-layer perceptron network, the weights and biases of the neurons or synapses dictate the system to remain in the lowest energy state, where the energy equation depends upon the update of the weights via the optimization process which is given by Eq. 2:

$$E(v, h) = - \sum_i^n a_i v_i - \sum_j^m b_j h_j - \sum_i^n \sum_j^m v_i w_{i,j} h_j \quad (2)$$

where a_i , b_j are the biases, v_i are the visible nodes, h_j are the hidden nodes, and $W_{i,j}$ are the weights between v_i and h_j .

Fig. 2 Maxwell–Boltzmann probability distribution curve



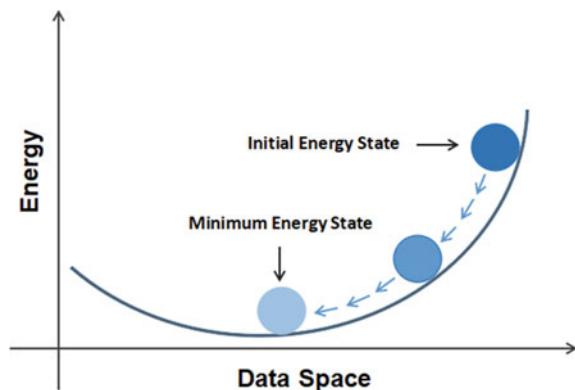
The process of update of parameters causes the change in energy of the system, and thus, the system slowly reaches the lowest energy state which can be shown by the Fig. 3:

The entire architecture of the Multi-Layer Perceptron Networks (MLPN) with the thermodynamical approach of interpretation used for estimation of the power generated has been developed with complex differential equations, linear algebraic representations and the concept of graph theory intertwining with the underlying equations of derivatives and vector algebra. The mathematical equations and developments used in the paper is in alignment with the thermodynamic approach presented in the paper and suffices the interpretation and explanation in all possible ways. The development of the multi-layer perceptron networks is shown in Fig. 4.

where

- W_{ij}^i Weights used in each of the connections in between the neurons in the i th layer
- X_i Input vectors incoming into the multi-layer perceptron networks
- f_{ij} Each of the neurons in the hidden layers
- O_{ij} Output vectors outgoing from each of the neurons of the layers
- \hat{y} The final predicted output neuron

Fig. 3 Energy estimation curve



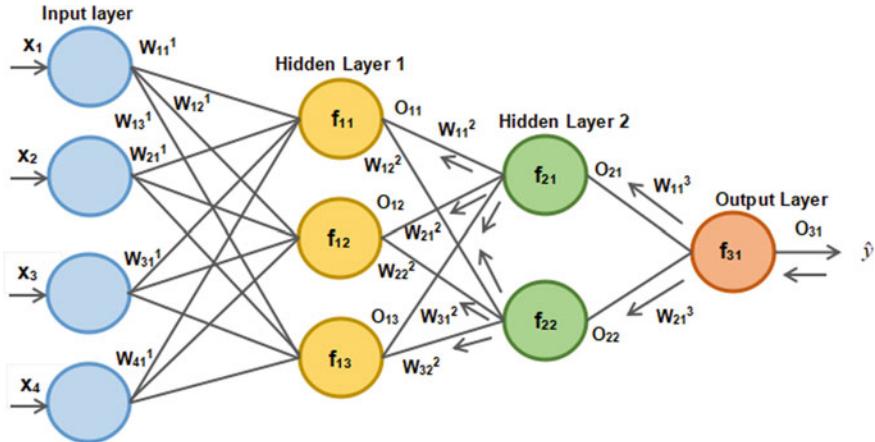


Fig. 4 Detailed architecture and working of multi-layer perceptron network

y Original output (used in the training samples of the MLPN)

→ Represents the direction in which the optimization is done in the MLPN.

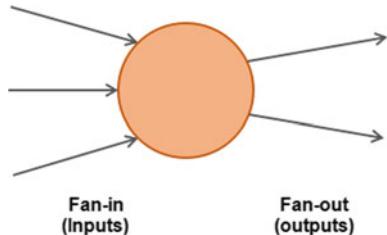
In the architecture of the above neural network, the weight initialization technique used is He-Normal initialization as well as He-Uniform initialization [11]. Let us consider the neuron as shown in Fig. 5.

In the above figure, fan-in represents the input vectors incoming into the neuron and fan-out the output vectors outgoing from the neuron. In this method of He-Initialization, the weights are initialized keeping in mind the size of the previous layer which helps in attaining a global minimum of the cost function faster and more efficiently. The weights are still random but differ in range depending on the size of the previous layer of neurons. This provides a controlled initialization hence the faster and more efficient gradient descent [12].

The equation of the He-Normal initialization is given in Eq. 3 which incorporates a normal or Gaussian distribution.

$$W_{ij} \approx N(0, \sigma), \text{ where } \sigma = \sqrt{\frac{2}{\text{fan - in}}} \quad (3)$$

Fig. 5 Diagram of a neuron



The equation of the He-Uniform initialization is given in Eq. 4 which incorporates a uniform distribution.

$$W_{ij} \approx u \left[-\sqrt{\frac{\sigma}{\text{fan} - \text{in}}}, \sqrt{\frac{\sigma}{\text{fan} - \text{in}}} \right] \quad (4)$$

The working of the neural network by incorporation of weights and biases with the addition of the activation function can be explained in Figs. 6 and 7.

In the process of getting an accurate estimation of the generated power, the neural network has to minimize the lost function (difference between the predicted and original output) during the training procedure of the neural network by updating the various learnable parameters used in the multi-layer perceptron networks like the weights and biases. This process of optimizing the learnable parameters during the training procedure of the MLPN is done with the help of optimization function. The optimization function used in this paper is Adaptive Momentum Optimizer (Adam) [13] which basically follows a backpropagation algorithm, i.e. it goes back into the neural network like a feedback mechanism and thus updates the parameters which in turn goes on minimizing the loss function and reaching a precise and accurate estimation of the output. The cost function used in the neural network is a mean-squared error loss [14].

The process of optimization can also be explained by the fact that while updating the parameters in the network via the chain rule derivative functions, the optimizer reaches the global minima or lowest energy state of the gradient descent curve [15] (analogically of the energy curve).

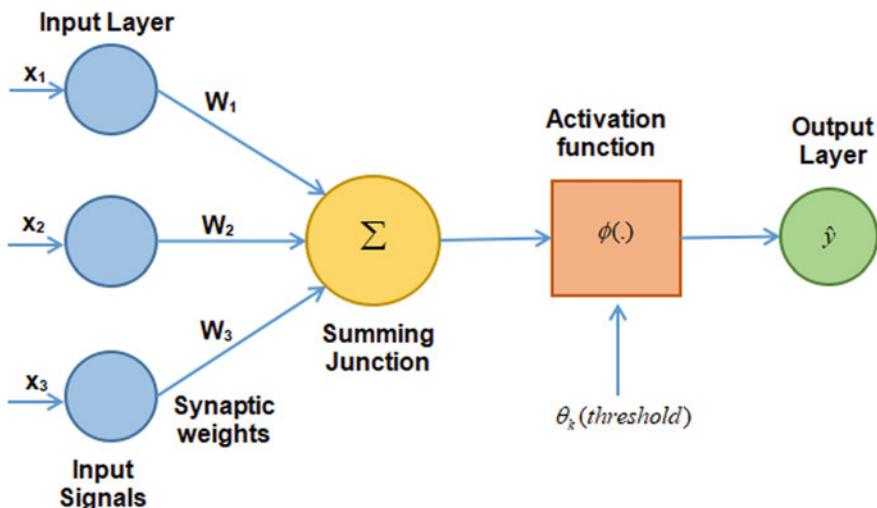
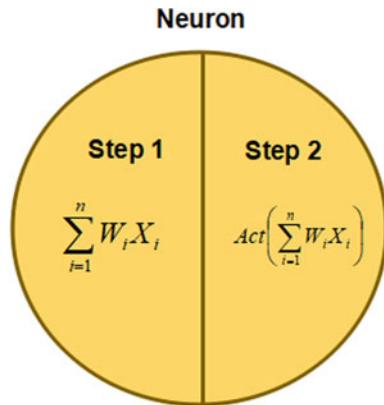


Fig. 6 Working of multi-layer perceptron network

Fig. 7 Detailed mechanism inside perceptron (neuron)



The learning rate value used in this paper is 0.001. Thus, basically, the learning rate parameter helps to decide the rate at which the derivative of the loss function of the neural network will reach zero and attain the global minima position.

The optimization and training of the neural network are done in batches for efficient computation and quick learning. The batch size used in the MLPN is 16, and the number of epochs used in the training process are 150. An epoch is a measure of the number of times all of the training vectors are used once to update the weights. The gradient descent curve is shown in Fig. 8.

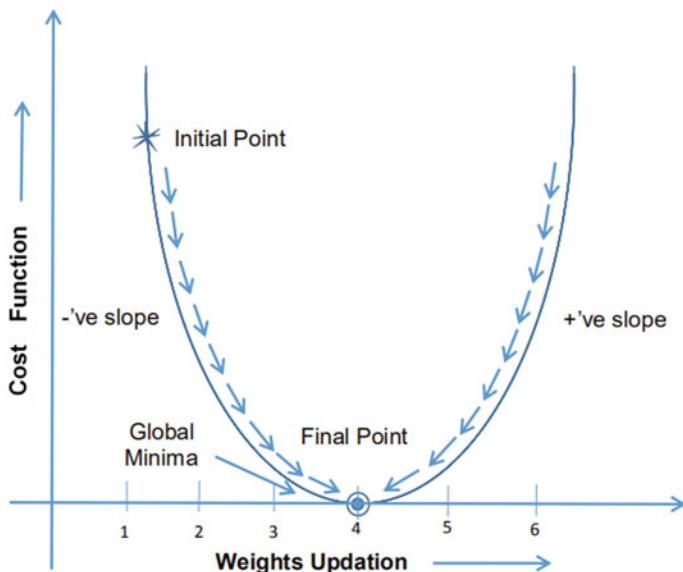


Fig. 8 Gradient descent algorithmic curve

From the figure above, the continuous process of update of the parameters, thus reaching the global minima point is clearly shown. In this way, the neural network (basically the cost function) reaches the lowest energy state having lowest entropy and randomness, and thus, the probability of remaining in this very state is maximum which is in alignment with Maxwell–Boltzmann distribution curve interpretation. The update of the weight parameter in the neural network is given in Eq. 5:

$$W_{\text{new}} = \left(W_{\text{old}} - \eta \frac{\partial L}{\partial W_{\text{old}}} \right) \quad (5)$$

where

- W_{old} Initial weight of the neural network;
- W_{new} Updated weight due to optimization.
- L Cost function of the neural network;
- η Learning rate.

The equation of the Cost Function is given in Eq. 6:

$$\text{Loss} = \frac{1}{N} \sum_{i=1}^N (y - \hat{y})^2 \quad (6)$$

where

- N number of training samples;
- y original output;
- \hat{y} predicted output.

The entire process of the backpropagation algorithm and continued chain rule process for the update of the parameters via the optimization function as shown in the equations below with respect to the architecture of the multi-layer perceptron networks shown in Fig. 1.

$$\begin{aligned} W_{11\text{new}}^3 &= W_{11\text{old}}^3 - \eta \frac{\partial L}{\partial W_{11\text{old}}^3} \\ \frac{\partial L}{\partial W_{11}^3} &= \left(\frac{\partial L}{\partial O_{31}} \cdot \frac{\partial O_{31}}{\partial W_{11}^3} \right); \quad \left(\frac{\partial L}{\partial W_{21}^3} \right) = \left(\frac{\partial L}{\partial O_{31}} \cdot \frac{\partial O_{31}}{\partial W_{21}^3} \right) \\ W_{11\text{new}}^2 &= W_{\Delta\Delta\text{old}}^2 - \eta \frac{\partial L}{\partial W_{11}^2} \\ \left(\frac{\partial L}{\partial W_{11}^2} \right) &= \left(\frac{\partial L}{\partial O_{31}} \cdot \frac{\partial O_{31}}{\partial O_{21}} \cdot \frac{\partial O_{21}}{\partial W_{11}^2} \right) + \left(\frac{\partial L}{\partial O_{31}} \cdot \frac{\partial O_{31}}{\partial O_{22}} \cdot \frac{\partial O_{22}}{\partial W_{11}^2} \right) \end{aligned}$$

The application of the process of contrastive divergence approach which is specifically used in Gibbs sampling process (special case of Markov-Chain Monte-Carlo

Table 1 Comparison of accuracies of different network optimization techniques

Optimization technique or algorithm	Accuracy score (%)
Adaptive momentum	98.93
Adaptive gradient	96.23
Adaptive delta	92.45
RMSProp	91.34
Stochastic gradient descent with momentum	87.67

Technique) can also be done in this case due to the interpretation of the multi-layer perceptron networks as an energy-based model, and thus, due to the application of contrastive divergence, the learning process of the algorithm during training phase can be optimized to a large extent, and the point of minimum energy or global minima can be reached very fast and efficiently with more accurate and proper updated learnable parameters in order to predict or estimate the output with highest level of accuracy.

3 Results and Conclusion

Accuracy is used as a metric in order to analyze the performance of the multi-layer perceptron networks for estimation of the power generated in the CCPP. A comparative analysis report of accuracies achieved with the help of various network optimization technique is given in Table 1.

Therefore, from the table, it is deduced that the Adaptive Momentum (Adam) optimization algorithm which is actually a combination of Stochastic Gradient Descent with Momentum (SGDM), and RMS prop algorithms help in the best optimization of weights and parameters in order to achieve an accuracy of 97.93%, which is by far the most robust and efficient algorithm to achieve such a ‘State-of-Art’ accuracy in the process of estimation of the power generated in the CCPP.

References

1. H. Nascimento Camelo, P.S. Lucio, J.B.V.L. Junior, P.C.M. de Carvalho, A hybrid model based on time series models and neural networks for forecasting wind speed in the Brazilian northeast region. *Sustain. Energy Technol. Assess.* **28**, 65–72 (2018)
2. G. Sadeghi, S. Nazari, M. Ameri, F. Shama, Energy and energy evaluation of the evacuated tube solar collector using Cu₂O/water nanofluid utilizing MLPN methods. *Sustain. Energy Technol. Assess.* **37**, 100578 (2020)
3. C. Khadse, M. Chaudhari, V. Borghate, Conjugate gradient back-propagation based artificial neural network for real-time power quality assessment. *Int. J. Electr. Power Energy Syst.* **82**, 197–206 (2016)

4. N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, R. Salakhutdinov, Dropout: a simple way to prevent neural networks from overfitting. *J. Machine Learn. Res.* **15**(1), 1929–1958 (2014)
5. H. Kaya, P. Tüfekci, F.S. Gürgen, Local and global learning methods for predicting power of a combined gas & steam turbine, in *Proceedings of the International Conference on Emerging Trends in Computer and Electronics Engineering ICETCEE*, pp. 13–18 (2012)
6. K. Hara, D. Saito, H. Shouno, Analysis of function of rectified linear unit used in deep learning, in *2015 International Joint Conference on Neural Networks (IJCNN)* (IEEE, 2015), pp. 1–8
7. P. Tüfekci, Prediction of full load electrical power output of a base load operated combined cycle power plant using machine learning methods. *Int. J. Electr. Power Energy Syst.* **60**, 126–140 (2014)
8. E. Wild, On Boltzmann's equation in the kinetic theory of gases, in *Mathematical Proceedings of the Cambridge Philosophical Society* (Vol. 47, No. 3, pp. 602–609). Cambridge University Press: A. Kleidon, R.D. Lorenz (eds.), *Non-equilibrium Thermodynamics and the Production of Entropy: Life, Earth, and Beyond* (Springer Science & Business Media, 2004)
9. J.S. Rowlinson, The Maxwell-Boltzmann distribution. *Mol. Phys.* **103**(21–23), 2821–2828 (2005)
10. G.E. Uhlenbeck, L.S. Ornstein, On the theory of the Brownian motion. *Phys. Rev.* **36**(5), 823 (1930)
11. T.A. Sai, H.H. Lee, Weight initialization on neural network for neuro PID controller—case study, in *2018 International Conference on Information and Communication Technology Robotics (ICT-ROBOT)* (IEEE, 2018), pp. 1–4
12. R. Lenka, K. Dutta, A. Khandual, S.R. Nayak, Bio-medical image processing: medical image analysis for malaria with deep learning, in *Examining Fractal Image Processing and Analysis* (IGI Global, 2020), pp. 158–169
13. D.P. Kingma, J. Ba, Adam: A method for stochastic optimization (2014). arXiv preprint [arXiv: 1412.6980](https://arxiv.org/abs/1412.6980)
14. C.M. Theobald, Generalizations of mean square error applied to ridge regression. *J. Roy. Stat. Soc.: Ser. B (Methodol.)* **36**(1), 103–106 (1974)
15. L. Bottou, Large-scale machine learning with stochastic gradient descent, in *Proceedings of COMPSTAT'2010* (Physica-Verlag HD, 2010), pp. 177–186

Deep Convolutional Neural Network-Based Knee Injury Classification Using Magnetic Resonance Imaging



R. Pandey, S. Mitra, A. Prajapati, A. K. Singh, and V. K. Shrivastava

Abstract Radiologists tend to possess human error during pathologies scans for any abnormalities; thus, the introduction of automation shall has a great impact on accurately detecting injuries and more musculoskeletal examinations. Magnetic Resonance Imaging (MRI) scans are an effective method to detect injured sections but the process tends to be time-consuming and prone to human error. Therefore, our paper highlights the automation impact on knee MRI for abnormal, anterior cruciate ligament (ACL), meniscal damages using deep learning. Deep learning model such as Convolutional Neural Network (CNN) can infer the representation of images due to its properties: local connection and shared weights. We have explored two deep CNN models (VGG16 and VGG19) with transfer learning approach to predict knee injury on MRNet dataset collected by Stanford University. The classification accuracy obtained for abnormal, ACL tear, and meniscal tear are 83.19%, 84.96%, and 70.80%, respectively, using VGG16, whereas 84.07%, 84.96%, and 74.34%, respectively, using VGG19 model.

Keywords Knee injury · MRI · Deep learning · CNN · Transfer learning

1 Introduction

Knee injury is one of the most common injuries which are caused when an unwanted bending force applied at the knee joint from accidents or falls. Swelling and pain are the symptoms of knee injuries. The most common knee injuries include tear of anterior cruciate ligament (ACL) and meniscus. The ACL is a tissue that joins the femur to the large bone tibia of the lower leg at the knee joint [1]. It is one of the major ligaments of the knee and can get ruptured due to sudden jerks while playing sports like basketball, volleyball, football, etc. Diagnosis of an ACL tear is executed by Magnetic Resonance Imaging (MRI) scans to figure out the signs of

R. Pandey · S. Mitra · A. Prajapati · A. K. Singh · V. K. Shrivastava (✉)

School of Electronics Engineering, Kalinga Institute of Industrial Technology (KIIT),
Bhubaneshwar 751024, India

e-mail: vimal.shrivastavafet@kiit.ac.in

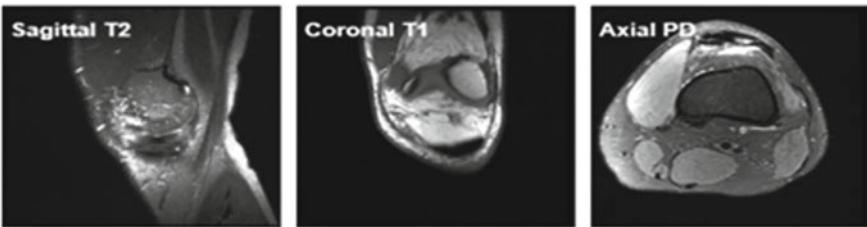


Fig. 1 Knee MRI in three planes: sagittal, coronal, axial

injuries and damaged tissue in the knee, including cartilage. The meniscus tear is one of the most conventional cartilage injuries. The role of the meniscus is to act as a shock absorber that predominantly maintains knee joint stability during weight-bearing activities. Initial diagnosis of meniscal tear involves physically examining the injury by palpating the joint to look for the tender areas and examining the motion of the knee joint. A surgical repair is usually recommended in the circumstance that it doesn't repair itself.

MRI captures the images in three planes: sagittal, coronal, and axial [2]. Figure 1 provides an illustration of knee MRI in these three planes. Here, the first row provides a black and white view of the knee MRI in the three planes, namely sagittal, coronal, and axial, whereas the second row provides the same views in colored format. The images in Fig. 1 are taken from the MRNet dataset [3] considered in this paper for experiment.

Diagnosis of these injuries is done either by an X-ray or MRI as per the medical requirements. An MRI primarily envisions components of the joint, this includes the articular cartilage, menisci, and intra-articular ligaments. However, diagnosis of injury using MRI is time-consuming and prone to human error. Therefore, a predictive model that classifies the knee injury can help to radiologists. However, a limited work on this has been reported in the literature to the best of our knowledge. In [4]; a semi-automated approach has been proposed for ACL tear detection using histogram of oriented gradient (HOG) descriptors and gist descriptors with support vector machine and random forest classifiers. In [5], segmentation of knee injury has been proposed using active contour method. Bien et al. [3] have proposed a deep learning approach for detection of ACL tear, meniscal tear, and abnormal class using knee MRI. Along this line, we have presented here transfer learning-based model capable of automatically determining whether an injury is abnormal, ACL tear, or meniscal tear.

2 Methodology

Deep learning is analogous to a function that precisely focuses on mimicking the human brain and how the brain processes data which drives the decisions taken by

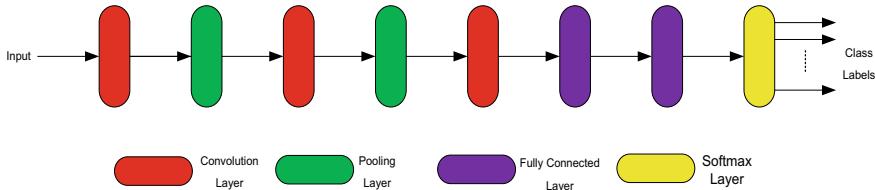


Fig. 2 A general deep CNN architecture

it. Deep learning has the capability to learn from raw unstructured data. The huge data is available now, for example: the data present in the online media, social media, e-commerce which simply termed as big data altogether. However, it might be time-consuming for human comprehension but can be handled using deep learning. A class of deep learning is convolutional neural network (CNN) which finds its application in visual imagery [6]. CNN is inspired by biological process between the neurons and the numerous connected layers. It mainly has three layers: convolution, pooling, and fully connected layer. Figure 2 represents the general deep CNN architecture, where the input images of the dataset are convolved and pooled subsequently as they are processed. However, the limitation of CNN is its high computational cost and requirement of large dataset to train effectively. Hence, we have used transfer learning approach [7] that is vital in certain applications and contexts, where obtaining large dataset for training neural networks is either costly or impossible. A brief description of transfer learning has been provided in next section.

2.1 Transfer Learning

It is an exploratory analysis in the machine learning field which primarily stores the knowledge received while solving a particular problem so that it can be applied on another contrasting, but related problem. For example, the knowledge gained while recognizing cars, so that it can be further used for the identification of trucks. Here, we have used transfer learning approach for knee injury classification, where the initial weights used for the model are the weights obtained after training the model on ImageNet dataset [8]. Further, all the layers are frozen in transfer learning approach except last few fully connected layers, i.e., the weights of only unfrozen layers are updated while training. Figure 3 shows the concept of transfer learning approach.

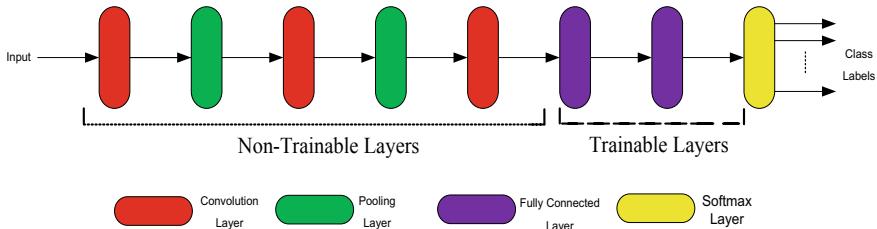


Fig. 3 Concept of transfer learning approach

3 Results and Discussion

3.1 Dataset Collection and Preparation

The dataset used in this paper is MRNet that is acquired from the Stanford University Medical Center [3]. It consists of 1130 training set in each category of abnormal, ACL tear, and meniscal tear. Further, there are 120 test set in each category of abnormal, ACL tear, and meniscal tear. Each data contains three types of planes, namely sagittal, coronal, and axial. The model is trained on each injury type and each plane separately.

3.2 Experimental Setup

In this paper, we have applied two pre-trained deep CNN models: VGG16 and VGG19. These models were pre-trained using ImageNet dataset, and trained weights were loaded initially for knee MRI classification. Based on the concept of transfer learning, all the layers of model are frozen except last layer while training. The training set has been divided into training and validation set with the ratio of 90:10, i.e., training set consists of 1017 images and 113 validation set. The test set has been separately provided having 120 images. Due to large dataset and model complexity, we have further divided the training set into few parts and fed to model one part at a time. The trained weights are stored and loaded to the model for next part of training. We have used Adam optimizer to update the weights. Further, 30 epochs were considered along with early stopping strategy [9] to prohibit over-fitting of the data. Here, training is stopped if pre-defined criterion meets, and model is trained till full epoch count if early stopping criterion does not meet. We have set the criterion as if validation loss does not decrease. The experiments have been performed in Google Colaboratory that provides Intel(R) Xeon(R) CPU@2.30 GHz, 13 GB RAM, and NVIDIA Tesla K80 GPU. We have used python 3.6 with Keras framework having a Tensorflow backend.

Table 1 Classification accuracy obtained using VGG16

	Sagittal (%)	Axial (%)	Coronall (%)
Abnormal	80.53	83.19	81.42
ACL tear	84.96	80.53	84.96
Meniscal tear	67.26	70.80	66.37

3.3 Performance Evaluation

The main objective of this research is to develop a model for three types of knee injury: Abnormal, ACL tear, and Meniscal tear. We have experimented with two pre-trained deep CNN models (VGG16 and VGG19) with transfer learning approach for each plane: sagittal, coronal, and axial. In this way, each model is trained nine times individually with crisscross combination of three knee injuries (abnormal, ACL tear, meniscal tear) and three planes (sagittal, coronal, axial). Table 1 presents the test accuracy obtained using VGG16 model corresponding to each knee injury and each plane. Similarly, Table 2 shows the classification accuracy achieved using VGG19. We can observe that the highest classification accuracy obtained among three planes for abnormal, ACL tear, and meniscal tear are 83.19%, 84.96%, and 70.80%, respectively, using VGG16, whereas 84.07%, 84.96%, and 74.34%, respectively, using VGG19 model. Further, we have shown the receiving operating characteristic (ROC) curve in Fig. 4 for VGG16 and VGG19. Tables 3 and 4 show the confusion matrix obtained using VGG16 and VGG19, respectively.

Table 2 Classification accuracy obtained using VGG19

	Sagittall (%)	Axial (%)	Coronall (%)
Abnormal	81.42	83.19	84.07
ACL Tear	84.96	84.96	84.96
Meniscal Tear	66.37	74.34	63.72

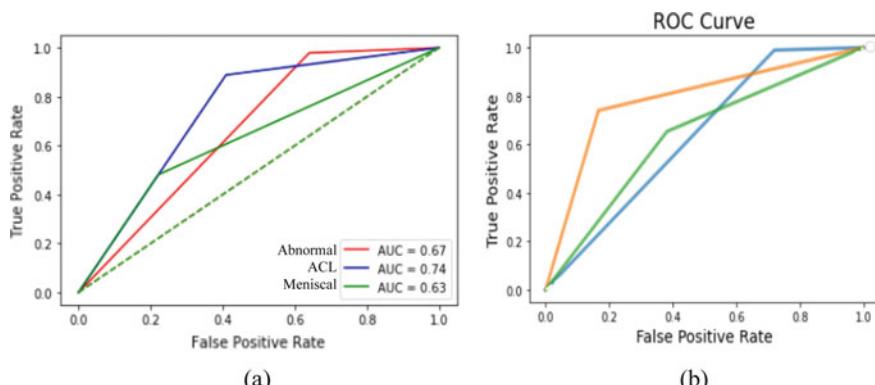
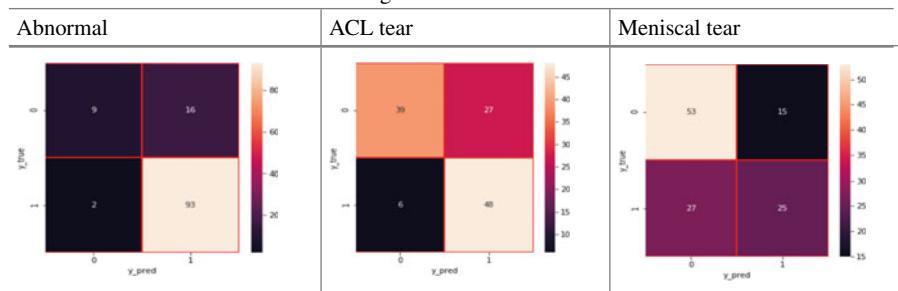
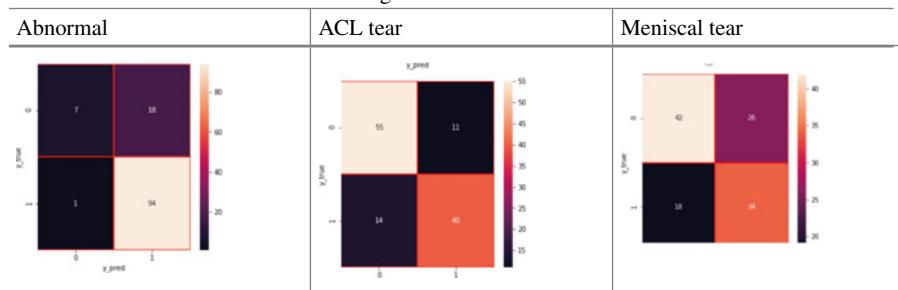
**Fig. 4** ROC curve obtained using: **a** VGG16 and **b** VGG19

Table 3 Confusion matrix obtained using VGG16**Table 4** Confusion matrix obtained using VGG19

4 Conclusion

The paper presents a transfer learning-based approach using VGG16 and VGG19 model for detection of knee injury on MRI datasets. The experiments have been performed in large dataset consists of 1130 training set and 120 test set with three class of injuries (abnormal, ACL tear, meniscal tear) for three planes (sagittal, coronal, axial). Hence, a total of nine models have been implemented by making a crisscross combination of three injuries and three planes. The highest classification accuracy obtained among three planes for abnormal, ACL tear, and meniscal tear are 83.19%, 84.96%, and 70.80%, respectively, using VGG16, whereas 84.07%, 84.96%, and 74.34%, respectively, using VGG19 model. In future, we intend to work on other state-of-the-art models. Further, multi-class classification can be a future work of this paper.

References

1. D.M. Daniel, M.L. Stone, B.E. Dobson, D.C. Fithian, D.J. Rossman, K.R. Kaufman, Fate of the ACL-injured patient: a prospective outcome study. *Am. J. Sports Med.* **22**(5), 632–644 (1994)
2. T.W. Hash, Magnetic resonance imaging of the knee. *Sports Health* **5**(1), 78–107 (2013)

3. N. Bien, P. Rajpurkar, R.L. Ball, J. Irvin, A. Park, E. Jones, M. Bereket, B.N. Patel, K.W. Yeom, K. Shpanskaya, S. Halabi, Deep-learning-assisted diagnosis for knee magnetic resonance imaging: development and retrospective validation of MRNet. *PLoS Med.* **15**(11), e1002699 (2018)
4. I. Štajduhar, M. Mamula, D. Miletic, G. Ünal, Semi-automated detection of anterior cruciate ligament injury from MRI. *Comput. Methods Programs Biomed.* **140**, 151–164 (2017)
5. R. Kaur, M.N. Aggarwal, Classification of knee MRI images. *Indian J. Comput. Sci. Eng.* **2**(3), 356–363 (2011)
6. V.K. Shrivastava, M.K. Pradhan, S. Minz, M.P. Thakur, Rice plant disease classification using transfer learning of deep convolution neural network, in *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences* (2019)
7. G. Pattnaik, V.K. Shrivastava, K. Parvathi, Transfer learning-based framework for classification of pest in tomato plants. *Appl. Artif. Intell.*, 1–13 (2020)
8. J. Deng, W. Dong, R. Socher, L.J. Li, K. Li, L. Fei-Fei, Imagenet: a large-scale hierarchical image database, in *2009 IEEE conference on computer vision and pattern recognition* (IEEE, 2009), pp. 248–255
9. R. Caruana, S. Lawrence, C.L. Giles, Overfitting in neural nets: backpropagation, conjugate gradient, and early stopping, in *Advances in Neural Information Processing Systems*, pp. 402–408 (2001)

Performance Comparison of LSTM Models for SER



Tanushree Swain, Utkarsh Anand, Yashaswi Aryan, Soumya Khanra,
Abhishek Raj, and Suprava Patnaik

Abstract Speech emotion recognition is essentially a sequence analysis task. Therefore, deployment of LSTM models is an appropriate benchmark for automatic emotion recognition of speech. This work is an attempt to compare the performance of stacked CNN-LSTM versus stand-alone LSTM architecture for recognition of emotions. The key contribution of this work is exploitation of the stacked CNN-LSTM architecture and augmentation of training data so as to get robust and reliable performance. Results are shown for the RAVDESS database. MFCCs from preprocessed raw audio files are considered as input to the models. Accuracy and other metrics indicate that hybrid CNN-LSTM achieves improved recognition accuracy compared to the stand-alone LSTM architecture. Augmentation of data supports better learning and robustness.

Keywords Automatic speech emotion recognition · CNN · LSTM · MFCC

T. Swain · U. Anand · Y. Aryan · S. Khanra · A. Raj · S. Patnaik (✉)
School of Electronics Engineering, KIIT Deemed to be University, Bhubaneswar, India
e-mail: suprava_patnaik@yahoo.com

T. Swain
e-mail: tanushrees98@gmail.com

U. Anand
e-mail: ananutkarsh@gmail.com

Y. Aryan
e-mail: yashaswi9903@gmail.com

S. Khanra
e-mail: soumyakhanra6@gmail.com

A. Raj
e-mail: abhishek.c.raj@gmail.com

1 Introduction

Recently virtual agents have begun to permeate our daily lives in many different ways. Therefore, there is a rising need for intelligent human and machine interactive systems. Although some of these systems have mastered the art of Speech Emotion Recognition (SER) and transcription, they still lack human-like aspects of the conversation. Apart from phoneme content, people perceive a lot from variations in rhythm, loudness, and tonal textures. Auditory variation leading to dynamism in rhythm, dialect, pitch, etc., influence the interpretation of underlying emotions. Recently, deep Convolutional Neural Networks (CNN) model has captured attention for their remarkable performance on image data. On the other hand, recurrent frameworks such as Long Short-Term Memory (LSTM) have demonstrated the state-of-the-art results in sequence analysis tasks, where cues carry temporal dependency. In this work, we seek to evaluate the performance of LSTM and stacked CNN-LSTM architectures for speech emotion recognition. We also have included data augmentation techniques to enhance the performance of these architectures.

2 Related Works

The earlier works on SER are focused on conventional statistical machine learning models by using Hidden Markov Models and/or Gaussian mixture models for sequence analysis along with support vector machines and K-nearest neighbors as classifiers. In 2014, Han et al. [1] presented one of the first deep learning frameworks for SER. Their approach was to interpret underlying phrase in hierarchical order by using densely connected neural network. In continuation of his work, Lee et al. [2] tried with RNN and LSTM by using a 32-dimensional vector consisting of pitch, probability of voiced amplitude, rate of zero cross, MFCC with log energy, and other derivatives of each frame. They used the probabilistic approach to learning and the aggregation of local probabilities into a global feature vector and Extreme Learning Machine (ELM) at the top. CNN-LSTM hybrid architectures are also being studied for SER in [3, 4]. Zhao et al. [4] created a multidimensional CNN-LSTM and used global spectral as well as temporal features as emotional cues. Etienne et al. [6] have suggested the concept of data augmentation.

3 Overview of Architectures

The central principle of RNN is based on the usage of sequential data. LSTM network is an upgradation of RNN structures that uses “memory cell state” instead of a hidden unit to learn long-term dependencies which the RNN is incapable of. This work is based on an CNN-LSTM model consisting of four Local Feature Learning

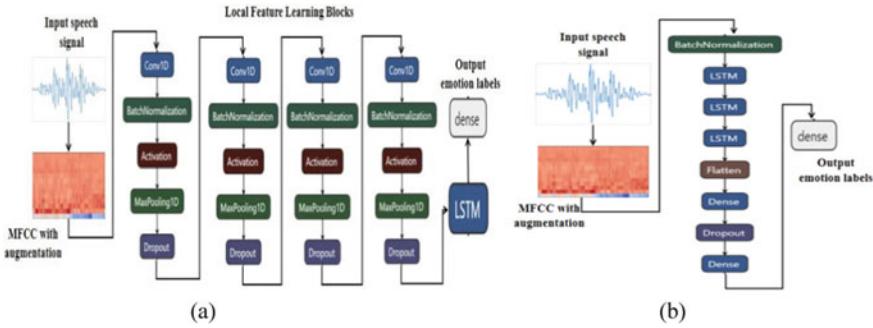


Fig. 1 Model architectures: **a** LSTM, **b** Stacked CNN-LSTM

blocks (LFLB) that learn from the input local features and an LSTM cell at the end for long-term dependency computing. The LFLB blocks consist of convolution filters, batch normalization, and Exponential Linear Unit (ELU) activation function. Pooling and dropout layers are used for dimension reduction. Model training is done inputs in batches of 64, “Adam” as an optimizer and learning rate of $1e-3$. Unlike Rectified Linear Unit (ReLU) ELU has negative activation, and it pushes the mean closer to zero. This causes faster learning and convergence. It performs identify operation on non-negative inputs and exponential non-linearity on negative inputs. The LSTM architecture is formed by three LSTM cells followed by a flattened layer. This flattened layer is the input to a dense layer with 285 neurons and ReLu activation followed by a dropout and dense layer that gives us the final predictions. Both the architectures have an output layer with fully connected softmax activation that gives us the prediction of the emotion label of the speech. Batch normalization is used for better and faster training. Schematic of LSTM and CNN-LSTM stacked architecture is shown in Fig. 1.

4 Proposed Framework

4.1 Dataset

Comparative analysis is carried out on the Ryerson Audio-Visual Database of Emotional Speech (RAVDESS) [5], which comprises of eight emotions spoken by 12 female and 12 male speakers (Fig. 2).

The purpose of including a feature extraction stage in the SER pipeline is to ensure that only informative data gets processed, and model performance is not affected due to redundant or noisy input. Mel Frequency Cepstrum Coefficient (MFCC) imitates the human auditory experience by considering the facts that human auditory perception does not obey a linear response [6]. Each tone has a real frequency determined by “hertz;” each tone has a subjective frequency “pitch” calculated by

Index	Description of Index
Mode	01: Video along with audio, 02: Only Video mode, 03: Only Audio mode
Data	01: Audio of Speech, 02: Audio of Song
Emotion	08:Surprise, 07:Disgust, 06:Fearful, 05:Angry,04:Sad, 03:Happy, 02:Calm, 01:Neutral
Pitch	02: high (intense emotion), 01: low (normal emotion)
Content	02: Dogs are sitting by the door, 01: Kids are talking by the door
Repeat	02: 2 nd time, 01: 1 st time,
Actor	01:1 st Actor-to-.25: 25 th Actor(If odd Male speaker else Female speaker) 03-01-06-02-01-02-12.wav: Audio-Speech-Fearful-Low-Dogs...-1st time -12 Actor(Female)

E

Fig. 2 Annotations of RAVDESS database

a metric named the mel scale. The main purpose of the subjective frequency is to capture the significant phonetic characteristics.

4.2 Data Augmentation

Audio-based models often encounter the issue of temporal overfitting, possibly due to less number of training data collected for selected speakers. One of the ways to resolve this is data augmentation. Augmentation of speech data can be achieved in following five ways: involving deformation by stretching or reducing the duration, adding background noise, manipulating pitch, and amplitude, all of these but only for interim tampering. Corrupting the speech signals with external factors does the training in more generalized manner forcing the model to learn relevant features so that it can be used for ubiquitous data. The methods that we have adopted for augmentation of the speech signals are: (1) Noise tuning: adding synthetically generated noise randomly but within the range 0.005 of maximum amplitude, (2) Time shifting: random shifting of sequence restricted to 2500 samples on either ways, (3) Rhythm shifting: stretching of sequence limited to the range 0.8–1.2 times the length of original sequence, (4) Tonal shifting: shifting the formant randomly but limited up to 2-semitones, and (5) Amplitude tuning: random change of amplitude from a uniform distribution. The objective is to increase the number of training samples realistically. Original RAVDESS consists of 1440 samples which were increased to 6914. Neutral and calm speech samples are not considered for intensity variation in the original dataset, therefore are not processed for tonal or amplitude tuning.

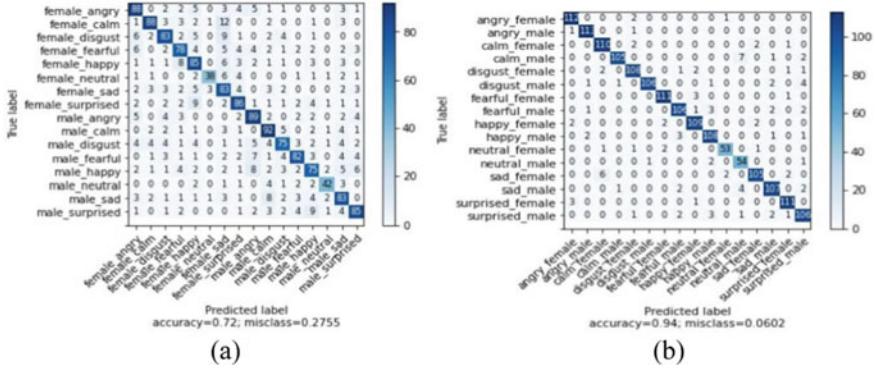


Fig. 3 Confusion matrix 16 classes **a** LSTM, **b** CNN-LSTM

5 Results

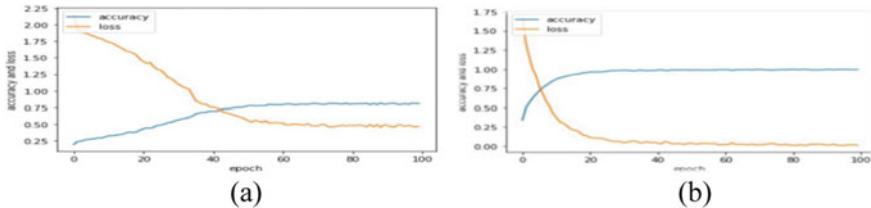
There exist notable gender dimorphism in the vocal apparatus of male and female, which affect both glottal frequency and filter aspects of voice production. Expressiveness emotional responses from men and women are very different. Often the expressions from women are more intense on disgust and horror, whereas men are found to be stronger on anger. Results are shown for 16 classes with gender abstraction and 8 classes with gender-based grouping. Results are evaluated for confusion matrix. Hypothetically of diagonal elements of confusion matrix are the probability of confusion among the classes and are wrong predictions. For 16 classes, LSTM architecture gives us an accuracy of 72% on the test data, whereas the CNN-LSTM architecture gives us an accuracy of 94% on the test data. Confusion matrix and other metrics are shown in (Fig. 3) and Table 1. We also check the performance by taking male and female speakers separately leading to 8 classes each. The LSTM results an accuracy of 78% for both female and male speakers. This shows 6% improvement for LSTM model. CNN-LSTM model establishes to be equally reliable to mixed gender and gender separated classifications. Besides being unbiased toward gender CNN-LSTM achieves acceptable accuracy in only 20 epochs. Therefore, its training is faster in terms of epochs. For all the experiments accuracy of neutral class, emotion less speech is detected with only 50% accuracy. Confusion rate is more for low arousal emotions (Fig. 4). Confusion matrix and other metrics corresponding to 8 classes are shown in (Fig. 5).

6 Conclusion

We observe that CNN-LSTM performs better than the LSTM, particularly for with both male and female speakers. We also discussed the impact of data augmentation on

Table 1 Precision, recall, and F1-Score with 16 classes

Label	Precision	Recall	F1-Score	Precision	Recall	F1-Score
	LSTM			Stacked CNN-LSTM		
Female angry	0.72	0.77	0.74	0.95	0.97	0.96
Female calm	0.84	0.76	0.80	0.92	0.96	0.94
Female disgust	0.74	0.72	0.73	0.94	0.94	0.94
Female fearful	0.71	0.68	0.69	0.97	0.97	0.97
Female happy	0.66	0.74	0.70	0.94	0.95	0.94
Female neutral	0.88	0.66	0.75	0.95	0.91	0.93
Female sad	0.66	0.72	0.66	0.95	0.91	0.93
Female surprised	0.74	0.75	0.74	0.95	0.97	0.96
Male angry	0.72	0.77	0.75	0.97	0.98	0.97
Male calm	0.72	0.80	0.76	0.98	0.91	0.95
Male disgust	0.74	0.65	0.69	0.99	0.92	0.95
Male fearful	0.80	0.71	0.75	0.93	0.92	0.93
Male happy	0.68	0.65	0.66	0.92	0.94	0.93
Male neutral	0.82	0.72	0.77	0.82	0.93	0.87
Male sad	0.66	0.72	0.69	0.93	0.92	0.93
Male surprised	0.77	0.74	0.75	0.89	0.91	0.90

**Fig. 4** Loss and accuracy plot of female 8 classes **a** LSTM, **b** CNN-LSTM

the performance and reliability of the model. In future, we seek to modify this existing architecture to create an end to end pipelines that will require less preprocessing of the input and provide better performance. Study can be further elaborated to consider the impact of arousal and valence.

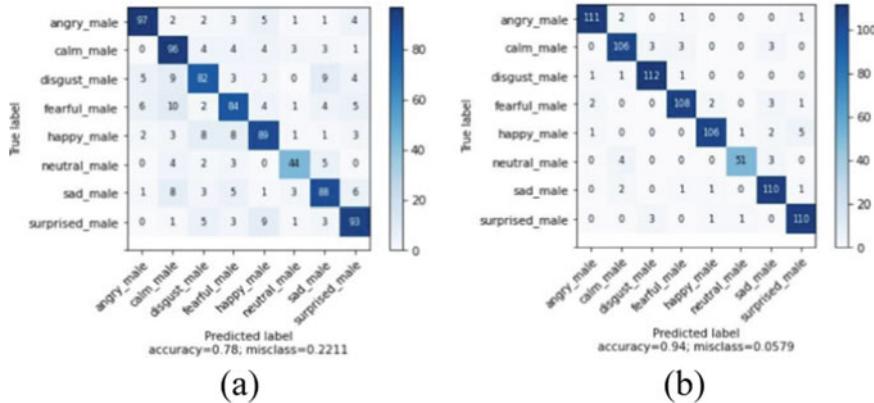


Fig. 5 Confusion matrix for 8 male speakers, **a** LSTM, **b** CNN-LSTM

References

1. K. Han, et al., Speech emotion recognition using deep neural network and extreme learning machine, in *INTERSPEECH* (2014)
2. J. Lee, I. Tashev, High-level feature representation using recurrent neural network for speech emotion recognition, in *INTERSPEECH* (2015)
3. C. Etienne, G. Fidanza, A. Petrovskii, L. Devillers, B. Schmauch, CNN+LSTM architecture for speech emotion recognition with data augmentation, in *Workshop on Speech, Music and Mind 2018* (2018). <https://doi.org/10.21437/SMM.2018-5>
4. J. Zhao, X. Mao, L. Chen, Speech emotion recognition using deep 1D & 2D CNN LSTM networks. *Biomed. Signal Process. Control* **47**, 312–323 (2019)
5. S. Livingstone, F. Russo, The Ryerson audio-visual database of emotional speech and song (RAVDESS): a dynamic, multimodal set of facial and vocal expressions in North American English. *PLOS ONE* **13**(5), e0196391 (2018). Available: <https://doi.org/10.1371/journal.pone.0196391>
6. C. Etienne, G. Fidanza, A. Petrovskii, L. Devillers, C. Schmauch, CNN+LSTM Architecture for Speech Emotion Recognition with Data Augmentation, [arXiv:1802.05630](https://arxiv.org/abs/1802.05630)

Harmony Search Optimization-Based Direct Estimation of Harmonic Components



Y. Ramya Sree, Bh. Sudharani, K. Sravan Kumar, D. J. V. Prasad, and Ch. Durga Prasad

Abstract In this paper, harmony search (HS) method has been employed for estimating the accurate harmonic components present in voltage/current waveforms of power system, since the metaheuristic algorithms are more attractive for intricate optimization to solve the problems of nonlinear in nature with the high degree of variables. Unlike conventional estimation approaches, direct curve fitting-based approach is adopted in this paper for the nearest and quick estimation of different harmonic components of distorted voltage signals. Comparative assessment of the HS algorithm with particle swarm optimization (PSO) reveals the advantages in terms of convergence and accuracy.

Keywords Harmonics · PSO · Harmony search algorithm

1 Introduction

In the modern power system, the use of microprocessor-based controllers and power electronic-based devices is enlarged. This leads to power quality problems and harmonic pollution [1]. The harmonics components will result in some adverse effects such as malfunction of relays and heating of components. To reduce harmonic

Y. Ramya Sree (✉) · Bh. Sudharani · D. J. V. Prasad · Ch. Durga Prasad
Department of Electrical and Electronics Engineering, SRKR Engineering College, Bhimavaram,
India
e-mail: ramyasreeyicherla@gmail.com

Bh. Sudharani
e-mail: sudha.bhupathiraju5@gmail.com

Ch. Durga Prasad
e-mail: chdpindia@gmail.com

K. Sravan Kumar
Department of Electrical and Electronics Engineering, Vignan's Institute of Information
Technology, Visakhapatnam, India
e-mail: sravankumar.konathala@gmail.com

pollution, it is necessary to determine harmonic parameters. The harmonic parameters include amplitude (A), phase angle (φ) and frequency (f). With these estimated parameters, the harmonics are eliminated by connecting some devices to the power system [2, 3].

Several techniques such as conventional, optimization, neural networks (NN) and hybrid algorithms have been implemented for harmonics estimation and tested. The discrete Fourier transform (DFT) [4] is one of the old conventional fast feasible algorithm. But, it does not give actual results and unstable for a few undesirable conditions. The fast Fourier transform (FFT) [5] is an advanced extension of the discrete Fourier transform (DFT). The FFT is the best technique for a static signal, and it does not perform well for time-varying signals. Kalman filtering method [6] is a simple way to estimate parameters. But it needs prior statistics of the signal. The least square techniques RLS [7] and LMS [8] are the best algorithms used so far. The values obtained by these estimation algorithms are closer to the actual values. But, these have certain drawbacks in terms of inaccuracies due to the existence of some variable parameters and noise. Some optimization techniques PSO, generic algorithm (GA) and artificial bee colony (ABC) are presented in [9–13]. These are the most effective algorithms for interharmonics, signals with noise and dynamic signals estimation. The estimation by these algorithms is carried out in one phase. So far the hybrid techniques [14–19] are considered, and these techniques are derived by combining the conventional least square algorithms with evolutionary algorithms and NN-based algorithms. The main objective of integrating the least square algorithms with other techniques is to reduce the error and to improve the accuracy of the estimation. The main drawback of these hybrid approaches is the estimation carried out in two phases. The amplitudes of harmonics are estimated by the least square technique, and the fundamental and other harmonic frequency components are estimated by other techniques.

In this paper, harmony search (HS) algorithm is implemented to estimate the parameters of the harmonics directly based on curve fitting approach by search-based mechanism. This algorithm is the best choice in terms of accuracy, computational time and convergence.

2 Problem Formulation

The general representation of distorted voltage and/or current signals of known fundamental frequency f is

$$x(t) = \sum_{h=1}^N A_h \sin(\omega_h t + \varphi_h) + \mu(t) \quad (1)$$

In Eq. (1), N = total number of harmonics, $\omega_h = h \cdot 2\pi f_0$ (f_0 = fundamental frequency), $\mu(t)$ = Additive noise. The discrete time version of the above equation

can be written as

$$x(k) = \sum_{h=1}^N A_h \sin(\omega_h k T_s + \varphi_h) + \mu(k) \quad (2)$$

In Eq. (2), T_s = Sampling period. The parameters of the harmonics have been estimated from the discrete time signal. Let the estimated parameters be A'_h and φ'_h . The distorted signal with estimated parameters is represented as

$$x(k)' = \sum_{h=1}^N A'_h \sin(\omega_h k T_s + \varphi'_h) \quad (3)$$

For estimation of exact parameters, an error is evaluated using Eq. (4) given by

$$\text{Error}, d(k) = x(k) - x(k)' \quad (4)$$

The objective function is framed with the help of error signal to estimate the parameters of harmonic signal given by

$$J = \min \sum_{k=1}^N d(k)^2 \quad (5)$$

In this paper, the objective function is minimized to estimate accurate parameters of harmonics by the harmony search algorithm whose details are provided in below section.

3 Harmony Search (HS) Optimization Algorithm

Harmony search optimization (HS) is a population-based algorithm proposed by Geem from music improvisation concept [20–23]. The first step in this algorithm is to initialize the optimization process for the objective function $f(x_1, \dots, x_n)$ with parameters such as range of variables $a_i \leq x_i \leq b_i, i = 1, 2, \dots, n$, harmony memory (HM) vector size (HMS), harmony memory rate coefficient (HMRC) and pitch adjusting rate (PAR). Next is to initialize the vector of harmony memory (HM). For this, randomly pick HMS number of vectors (x_1, \dots, x_n) and put them in the HM vector corresponding to the dominant values $f(x_1, \dots, x_n)$:

$$\text{HM} = \begin{bmatrix} x_1^1, \dots, x_n^1 & | & f(x^1) \\ \vdots & | & \vdots \\ x_1^{\text{HMS}}, \dots, x_n^{\text{HMS}} & | & f(x^{\text{HMS}}) \end{bmatrix}.$$

The next step is to update agents of HS $\mathbf{x}' = \mathbf{x}'_1, \dots, \mathbf{x}'_n$. For each $i = 1, 2, \dots, n$ the element x'_i is chosen. With the probability level to HMCR, from the total numbers x_i obtained in the vector of HM and with the probability level to 1—HMCR, randomly from the presumed limit $a_i \leq x_i \geq b_i$. If, in the prior step, the element x'_i is chosen from the vector of HM then, with the probability level to PAR, adjust the element x'_i in this way, $x'_i \rightarrow x'_i + \alpha$ (control the pitch of the note), for $\alpha = bw.u$, where bw indicates the bandwidth part of the limit of the variables and u is the randomly chosen number from range $\in [-1, 1]$ and with the probability level to 1—PAR we have nothing to do. If $f(x') < f(x^{\text{HMS}})$ then replace the element x' inside the vector of harmony memory HM instead of the element x^{HMS} and modify vector HM corresponding to the advancing values of the objective function. Updating HM vector steps are repeated until the termination criterion is satisfied.

4 Simulation Results

To analyze the performance of the harmony search algorithm for the estimation of harmonic parameters, two case studies with different signals have been taken.

Signal 1: Signal corrupted with noise along with DC component with decaying nature is considered as signal 1. In this case, a stationary signal having harmonics of order fundamental, 3rd, 5th, 7th, 11th is generated in MATLAB. The fundamental frequency (f_1) is 50 Hz, 3rd harmonic frequency $f_3 = 150$ Hz, 5th harmonic frequency $f_5 = 250$ Hz, 7th harmonic frequency $f_7 = 350$ Hz, eleventh harmonic frequency $f_{11} = 550$ Hz. The amplitude value of the fundamental harmonic component had taken a higher value compared to other harmonic components.

$$\begin{aligned} x(t) = & 1.5 \sin(\omega t + 80) + 0.5 \sin(3\omega t + 60) \\ & + 0.2 \sin(5\omega t + 45) + 0.15 \sin(7\omega t + 36) \\ & + 0.1 \sin(11\omega t + 30) + 0.5 \exp(-5t) + \mu(t) \end{aligned}$$

where $\mu(t)$ = random noise (Tables 1, 2 and Fig. 1).

Signal 2: In this case, a stationary signal having harmonics of order fundamental, 3rd, 5th, 7th, 11th, is generated in MATLAB (Table 3).

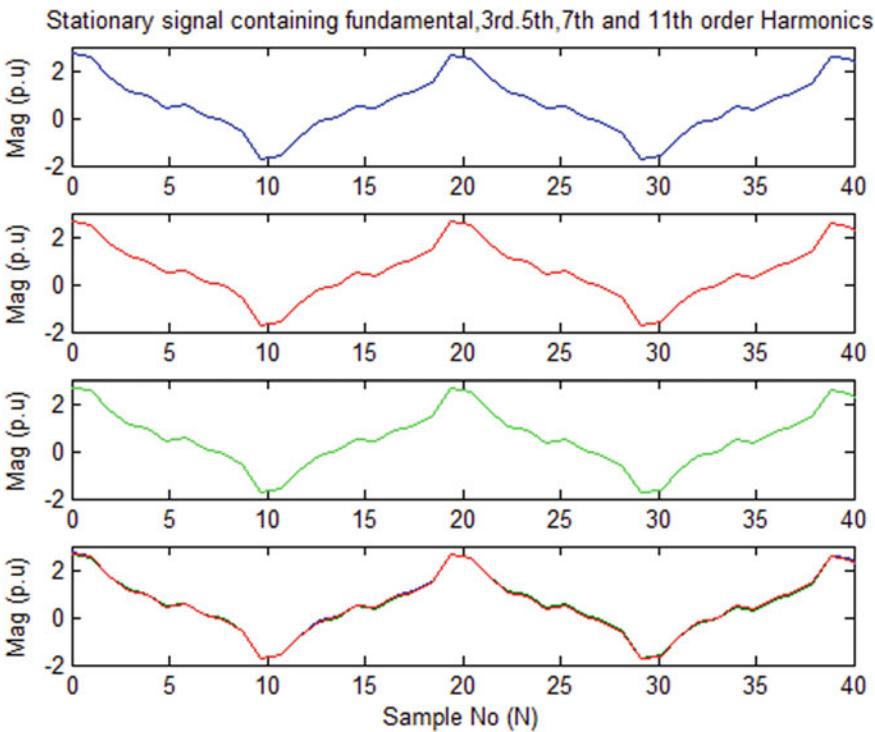
$$x(t) = 1.2 \sin(2\pi f_1 t + 75) + 0.2 \sin(2\pi f_3 t + 55)$$

Table 1 Optimal amplitude values for signal-1 obtained by HS

Technique	Quantity	Fundamental	3rd	5th	7th	11th	DC
PSO	Amplitude	1.5010	0.4996	0.2087	0.1496	0.1004	0.4984
HSO	Amplitude	1.5000	0.4994	0.2073	0.1503	0.1000	0.4989

Table 2 Optimal amplitude and phase values for signal-1 obtained by HS

Technique	Quantity	Fundamental	3rd	5th	7th	11th	DC
PSO	Amplitude	1.4897	0.5003	0.1946	0.1500	0.0998	0.5011
	Phase	77.9899	59.9691	48.6828	36.009	30.149	
HSO	Amplitude	1.5003	0.4994	0.1966	0.1504	0.1003	0.5011
	Phase	80.0028	60.0291	44.8954	36.077	29.930	

**Fig. 1** Estimation plots of signal 1 using PSO and HS**Table 3** Optimal amplitude and phase values for signal-2 obtained by HS

Algorithm	Parameter	Fundamental	3rd	5th	7th	11th
PSO	Amplitude	1.2302	0.2001	0.1945	0.1397	0.0999
	Phase	74.9938	53.0311	44.9234	38.6918	29.8427
HSO	Amplitude	1.2002	0.19997	0.2001	0.1499	0.0999
	Phase	75.0168	55.0139	44.9839	39.9947	29.9692

$$\begin{aligned}
 & + 0.2 \sin(2\pi f_5 t + 45) + 0.15 \sin(2\pi f_7 t + 40) \\
 & + 0.1 \sin(2\pi f_{11} t + 30) + \mu(t)
 \end{aligned}$$

5 Conclusions

In this paper, direct curve-based harmonic estimation concept is presented for distorted power system signals, and these harmonic components are accurately estimated using HS algorithm. This method provided direct amplitude and phase values of power system signals of different harmonic contents. The comparisons with PSO showcase the advantages of the HS in estimation.

References

- P.K. Ray, B. Subudhi, BFO optimized RLS algorithm for power system harmonics estimation. *Appl. Soft Comput.* **12**(8), 1965–1977 (2012)
- S.K. Singh, A.K. Goswami, N. Sinha, Power system harmonic parameter estimation using bilinear recursive least square (BRLS) algorithm. *Int. J. Electr. Power Energy Syst.* **67**, 1–10 (2015)
- S.K. Singh, D. Kumari, N. Sinha, A.K. Goswami, N. Sinha, Gravity search algorithm hybridized recursive least square method for power system harmonic estimation. *Eng. Sci. Technol. Int. J.* **20**(3), 874–884 (2017)
- F.J. Harris, On the use of windows for harmonic analysis with the discrete Fourier transform. *Proc. IEEE* **66**(1), 51–83 (1978)
- Z. Ren, B. Wang, Estimation algorithms of harmonic parameters based on the FFT, in *2010 Asia-Pacific Power and Energy Engineering Conference*. IEEE (2010, March), pp. 1–4
- S.K. Singh, N. Sinha, A.K. Goswami, N. Sinha, Several variants of Kalman Filter algorithm for power system harmonic estimation. *Int. J. Electr. Power Energy Syst.* **78**, 793–800 (2016)
- M. Bettayeb, U. Qidwai, Recursive estimation of power system harmonics. *Electr. Power Syst. Res.* **47**(2), 143–152 (1998)
- S.K. Singh, N. Sinha, A.K. Goswami, N. Sinha, Variable constraint based least mean square algorithm for power system harmonic parameter estimation. *Int. J. Electr. Power Energy Syst.* **73**, 218–228 (2015)
- Z. Lu, T.Y. Ji, W.H. Tang, Q.H. Wu, Optimal harmonic estimation using a particle swarm optimizer. *IEEE Trans. Power Delivery* **23**(2), 1166–1174 (2008)
- Y.N. Yin, W.X. Lin, W.L. Li, Estimation amplitude and phase of harmonic based on improved PSO, in *IEEE ICCA 2010*. IEEE (2010, June), pp. 826–831
- S. Biswas, A. Chatterjee, S.K. Goswami, An artificial bee colony-least square algorithm for solving harmonic estimation problems. *Appl. Soft Comput.* **13**(5), 2343–2355 (2013)
- Y. Kabalci, S. Kockanat, E. Kabalci, A modified ABC algorithm approach for power system harmonic estimation problems. *Electr. Power Syst. Res.* **154**, 160–173 (2018)
- S. Biswas, S.K. Goswami, A. Chatterjee, Optimum distributed generation placement with voltage sag effect minimization. *Energy Convers. Manage.* **53**(1), 163–174 (2012)
- J. Enayati, Z. Moravej, Real-time harmonics estimation in power systems using a novel hybrid algorithm. *IET Gener. Transm. Distrib.* **11**(14), 3532–3538 (2017)
- P.K. Ray, B. Subudhi, Neuro-evolutionary approaches to power system harmonics estimation. *Int. J. Electr. Power Energy Syst.* **64**, 212–220 (2015)

16. P. Garanayak, G. Panda, P.K. Ray, Power system harmonic parameters estimation using ADALINE-VLLMS algorithm, in *2015 International Conference on Energy, Power and Environment: Towards Sustainable Growth (ICEPE)*. IEEE (2015, June), pp. 1–6
17. S. Mishra, Hybrid least-square adaptive bacterial foraging strategy for harmonic estimation. IEE Proc.-Gener., Transm. Distrib. **152**(3), 379–389 (2005)
18. S.K. Singh, N. Sinha, A.K. Goswami, N. Sinha, Power system harmonic estimation using biogeography hybridized recursive least square algorithm. Int. J. Electr. Power Energy Syst. **83**, 219–228 (2016)
19. S.K. Singh, N. Sinha, A.K. Goswami, N. Sinha, Robust estimation of power system harmonics using a hybrid firefly based recursive least square algorithm. Int. J. Electr. Power Energy Syst. **80**, 287–296 (2016)
20. Z.W. Geem, J.H. Kim, G.V. Loganathan, A new heuristic optimization algorithm: harmony search. Simulation **76**(2), 60–68 (2001)
21. A. Askarzadeh, E. Rashedi, Harmony search algorithm: basic concepts and engineering applications, in *Intelligent Systems: Concepts, Methodologies, Tools, and Applications* (IGI Global, 2018), pp. 1–30
22. K.S. Lee, Z.W. Geem, A new meta-heuristic algorithm for continuous engineering optimization: harmony search theory and practice. Comput. Methods Appl. Mech. Eng. **194**(36–38), 3902–3933 (2005)
23. D. Weyland, A critical analysis of the harmony search algorithm—how not to solve sudoku. Oper. Res. Perspect. **2**, 97–105 (2015)

Di-zeta Converter-Based Solar–Wind Generating System with a Novel SMES Battery Hybrid Energy Management



K. Aseem and S. Selva Kumar

Abstract Renewable energy system plays a major role in today's world of power deficits. The renewable power technologies keep fluctuating due to their dependence on natural phenomena. Using hybrid combinations, it is possible to meet the technological and economical requirements of the energy storage process. This paper proposes a hybrid energy storage system based on the superconducting magnetic energy storage System (SMES) and the lead–acid battery to provide power stabilization in the unreliable microgrid connected to the grid. The proposed system may be a good solution for minimizing the impact during fault condition on variability in power output at the point of common coupling (PCC). The two hybrid power sources are connected using dual input single output zeta (di-zeta) converter to improve efficiency and versatility of the proposed system. Grid-linked microgrid with SMES battery was simulated using MATLAB/Simulink platform and evaluated using an energy management algorithm with and without fault condition. The proposed system is demonstrated by comprehensive simulation results.

Keywords Solar–wind hybrid system · Hybrid energy management · Di-zeta converter · SMES battery combination

1 Introduction

Hybrid energy generating systems are mainly based on non-conventional sources of energy, mainly wind and solar systems [1]. As the renewables are heavily dependent on the natural phenomenon, there will be fluctuations in the power output of PV and wind generation. In addition, the power transmitted to the load can get interrupted when there is a fault in the grid-linked microgrid. Therefore, it requires power management system to maintain the voltage and feed the load during fault condition.

K. Aseem (✉) · S. Selva Kumar

Department of Electrical and Electronics Engineering, Amrita School of Engineering, Amrita Vishwa Vidyapeetham, Coimbatore, India

e-mail: azieee@gmail.com

Modern research is now focused on grid system power smoothening using high-power energy storage technology [2]. SMES has the potential to balance the power and boost renewable energy system power quality [3, 4] and deduces the model, which takes active and reactive power as inputs of the SMIB system with SMES. Kang et al. [5] focus on the integration of superconducting coil in mitigating the fluctuations in the distribution network with plug in battery electric vehicle. The use of lead–acid battery will compensate for low-energy density of SMES.

2 Modelling of PV and Wind Energy Conversion System

PV array transmits power to a three-phase voltage source inverter (VSI) through a dual input zeta converter (di-zeta). The peak available power from the PV array used is 100 kW. The method of incremental conductance (IC) quantifies maximum level of power by comparing instant conductance of the solar photovoltaic array. The maximum power is tracked in reference to the duty cycle variations [6, 7].

Wind power transmission system consists of a wind turbine that extracts mechanical energy from wind energy. The shaft of the wind turbine is attached to the shaft of the permanent magnet synchronous generator gearbox. Equations 1 and 2 give the detailed mathematical formulation for the optimized power coefficient C_P

$$C_P(\lambda, \beta) = C_1 \left(\frac{C_2}{\lambda_i} - C_3 \beta - C_4 \right) e^{\frac{-C_5}{\lambda_i}} C_6 \lambda \quad (1)$$

$$\frac{1}{\lambda_i} = \frac{1}{\lambda - 0.08\beta} - \frac{0.035}{\beta^3 + 1} \quad (2)$$

$$C_1 = 0.2126, C_2 = 118, C_3 = 0.4, C_4 = 0.00023, C_5 = 20 \text{ and } C_6 = 0.00296.$$

The attributes of the power coefficient are a function of the ratio of tip speed ratio λ and pitch angle β .

3 Propose Di-zeta Converter-Based Solar–Wind Generating System

A dual input di-zeta convertor is used here. Even if one source is inaccessible, the other can provide the necessary or lower power, ensuring a continuous supply of power. Proposed solar–wind di-zeta fed hybrid energy system with SMES battery energy storage is shown Fig. 1. A wide range of variations in input voltage induced by various insulations is permissible in this system.

The switches Q_1 and Q_2 are switched on when wind and PV are available. During this interval, the currents rise linearly through L_1 , L_2 and L_0 . This is called the

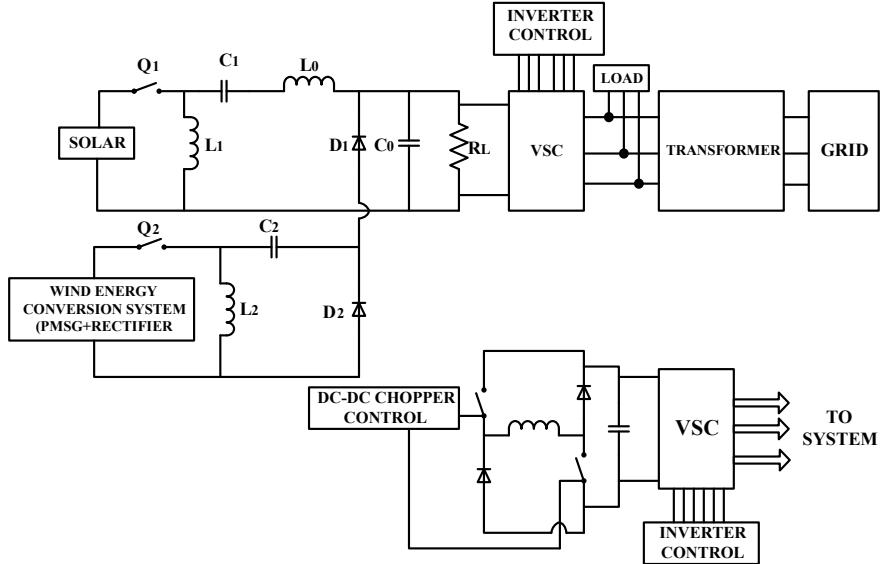


Fig. 1 Proposed solar–wind di-zeta fed hybrid energy system with SMES battery hybrid energy storage

charging interval. L_1 and L_0 release the accumulated energy to C_0 and the load when PV and wind are switched off. This is the discharging mode.

4 SMES Modelling and Control

The SMES unit consists of a power conditioner, a DC to DC converter and a semiconductor coil.

The chopper circuit is used to control coil for charging and discharging. The energy and power stored by the SMES is given by Eqs. 3 and 4

$$E = 0.5LI^2 \quad (3)$$

$$P_{SMES} = \frac{dE}{dt} = LI \frac{dI}{dt} = VI \quad (4)$$

Here, the energy contained in SMES is indicated as E , the inductance indicated as L , the current indicated as I , and the voltage is V . The SMES control diagram is demonstrated in Fig. 2.

The reference power is identified based on Eqs. 5 and 6.

$$P_{ref} = 3V_g i_{dpvref} \quad (5)$$

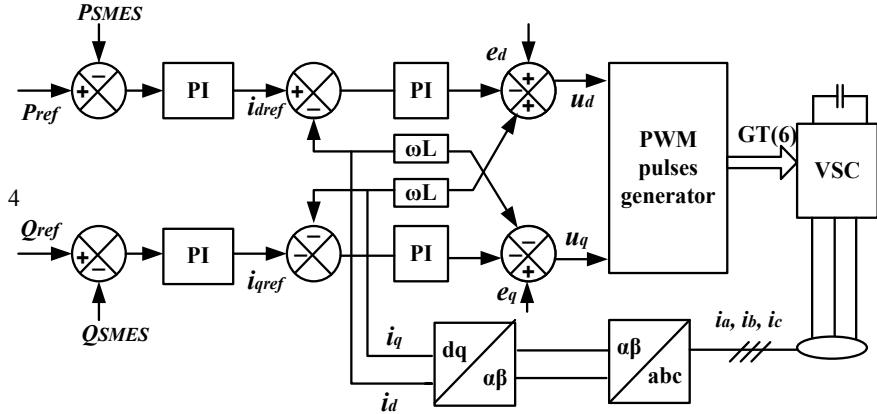


Fig. 2 Control strategy of SMES inverter

$$Q_{ref} = 3V_g i_{qp vref} \quad (6)$$

5 Battery Modelling and Control

For this work, the lead-acid battery is selected as the element of energy storage. With the aid of a bidirectional converter, real power flow can be controlled by battery. The battery reference power is dynamically produced by the difference power of the active power injection and the hybrid power generated by the inverter [8]. Charging and discharging modes are selected based on the comparison of the inverter power with a threshold power of 80 kW as in the control diagram as shown in Fig. 3.

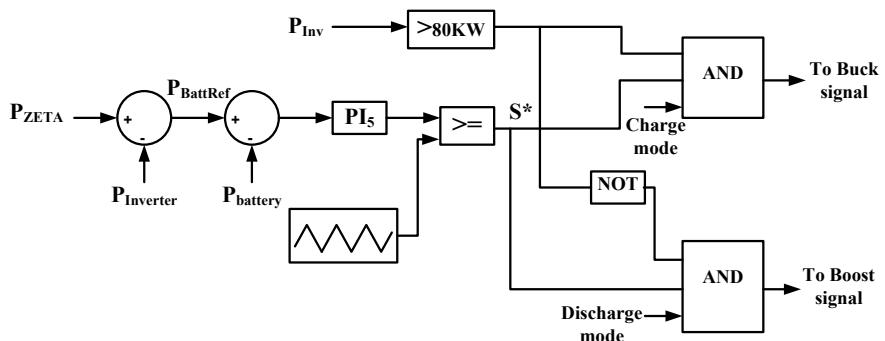


Fig. 3 Control logic of battery

6 SMES Battery Energy Management Controller

Switching between SMES battery is achieved by comparing the power available to them. SMES current is used as the deciding factor for storage system switching. The power will pass to the battery when the SMES current is below the minimum value. SMES current is given during the energy exchange phase in Eq. 7.

$$I_{\text{SMES}(t)} = \sqrt{I_{\text{ini_SMES}} + 2 \frac{\Delta E}{L_{\text{SMES}}}} \quad (7)$$

Based on the change of SMES current $I_{\text{SMES_min}}$ and $I_{\text{SMES_max}}$ from its minimum and maximum value, the capacity of SMES is calculated as follows:

$$\frac{1}{2} L_{\text{SMES}} (I_{\text{SMES}}^2 - I_{\text{SMES_min}}^2) \leq E_{\text{SMES}} \leq \frac{1}{2} L_{\text{SMES}} (I_{\text{SMES_max}}^2 - I_{\text{SMES}}^2)$$

SMES can mitigate the minor power fluctuations in the PCC during normal operating hours, and the battery is used as a contingency device. Battery will eliminate the PCC's long power fluctuations. SMES alone will recover the fault if its current is less than the maximum current. When the SMES current is closer to its minimum limit, the battery is switched ON.

7 Simulation Results and Analysis

The VSI converts the DC voltage to 260 AC voltage. The power output of the inverter is 80 kW. The variation in the power output during the interval $t = 0$ s to $t = 1$ s is due to variations in the irradiance and temperature. There is a distortion in the power output during this interval. Constant power output is reached at time $t = 1.7$ as shown in Fig. 4a.

The inverter output power P_{hybrid} and the coupled power to the main grid P_{grid} are obtained as $P_{\text{hybrid}} = 90$ K kW and $P_{\text{grid}} = 80$ kW under normal operating condition. The reference of SMES power, P_{SMES} is calculated by $P_{\text{SMES}} = 100$ kW – $P_{\text{hybrid}}(P_{\text{grid}})$, and the Q_{SMES} relation is always set as zero. If a fault happens at the main grid, the actual power balance fails, and the inverter's power output drops from its nominal value to a lower value. The SMES delivers the active power such as $P_{\text{SMES}} = P_{\text{hybrid}} - P_{\text{grid}}$. Figure 4b shows the inverter power output when SMES is not connected. During the fault condition, the power lowers to a much negative value. Figure 4c displays the inverter power output when connecting SMES with fault, and the power output has been balanced. Thus, SMES keeps the hybrid system connected to the main grid persistently. Figure 4d shows the power output of inverter while connecting SMES battery combination, and power is well balanced under the fault.

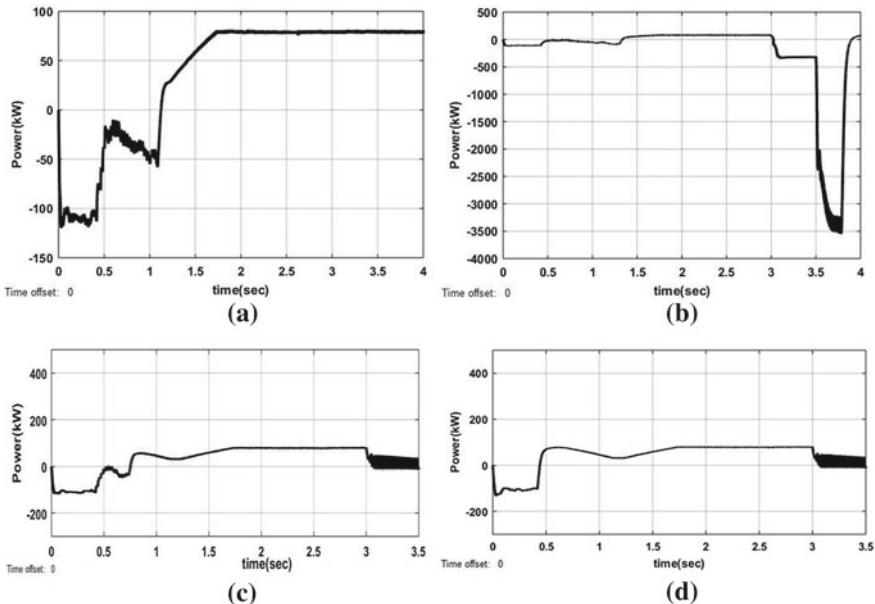


Fig. 4 **a** Power output of the inverter, **b** inverter response during fault without SMES, **c** inverter response with SMES, **d** inverter response with SMES battery

8 Conclusion

This paper discusses implementing SMES battery combination to reduce fluctuation during fault conditions in the hybrid power system. The results of the simulation confirm that the proposed system can ease the power perturbations and system stability can be regulated under fault condition. If SMES is not connected during fault condition, the power drops to a larger value. Thus, the system can survive under fault condition by connecting SMES unit. Use of the SMES allows the hybrid power system to improve fault ride through capability. The SMES battery is better for timely management of the microgrid transient fault issues than the battery alone. The two power sources are linked through di-zeta converter which enhances the efficiency of the proposed energy system.

References

1. A.-K. Daud, M. Ismail, W. Kukhun, M.M. Mahmoud, Simulation of a hybrid power system consisting of wind turbine, PV, storage battery and diesel generator: design, optimization and economical evaluation. *Int. J. Energy Eng.*
2. H.-G. Kim, D.-C. Lee, J.-K. Seok, G.-M. Lee, Stand-alone wind power generation system using vector-controlled cage-type induction generators, in *Proceedings of Inter. Conference on*

- Electrical Machines and systems*, vol. 1, Nov 2003, pp. 289–292
- 3. M. Farhadi, O. Mohammed, Energy storage technologies for high power applications. *IEEE Trans. Ind. Appl.* **52**(3), 1953–1961 (2016)
 - 4. S. Dechanu Paprittha, K. Hongesombut, M. Watanabe, Y. Mitani, I. Ngamroo, Stabilization of tie-line power flow by robust SMES controller for interconnected power system with wind farms. *IEEE Trans. Appl. Supercond.* **17**(2), 2365–2368 (2007)
 - 5. X. Shi, S. Wang, W. Yao, A. Waqar, W. Zuo, Y. Tang, Mechanism analysis and experimental validation of employing superconducting magnetic energy storage to enhance power system stability. *Energies* **8**(1), 656–681 (2015)
 - 6. K. Aseem, S. Selva Kumar, Closed loop control of DC-DC converters using PID and FOPID controllers. *Int. J. Power Electron. Drive Syst. (IJPEDS)* **11**(3), 1323–1332 (2020). ISSN: 2088-8694. <https://doi.org/10.11591/ijpeds.v11.i3.pp1323-1332>
 - 7. K. Aseem, S. Selva Kumar, Energy management controller for grid connected solar PV system with SMES-battery hybrid energy storage. *J. Adv. Res. Dyn. Control Syst.* **12**(04-Special Issue), 1385–1396 (2020). <https://doi.org/10.5373/jardcs/v12sp4/20201617>
 - 8. P. Vadana, K.K. Sasi, Dynamic energy management on a hydro-powered smart microgrid. *Adv. Intell. Syst. Comput.* **397**, 627–635 (2016)

Design and Implementation of Mixed Signal ASIC-Event Logger in TSMC 0.18 μm CMOS Technology



Saroja V. Siddamal^{ID}, Suhas B. Shirol^{ID}, Sujata S. Kotabagi,
Shraddha B. Hiremath, and Nalini C. Iyer

Abstract The authors present the design of integrated event recorder KLEEL-2020 with alarm consisting of low power on-chip RC oscillator. The technology used to implement the ASIC is TSMC 0.18 CMOS logic or mixed signal/RF, general purpose, 3.3 V/1.8 V. The on-chip oscillator produces frequency of 1.024 MHz. It draws a maximum current of 100 μA from battery, and the variation in frequency should be less than 10% for all PVT conditions. The ASIC has Low Drop Voltage regulator to provide constant supply voltage of 1.8 V to core. The Placement and Route flow is done in cadence innovous. The die size of the core is 1565 μm * 1565 μm. The power consumed for typical case (1.82 V) is 1.497 mW. The ASIC is taped out to Taiwan at mini@sic runway.

Keywords Event recorder · Event · ETC power analysis · ASIC · PVT

1 Introduction

Event recorder is the test ASIC designed and developed by team of students and faculty of KLE Technological University. This ASIC has similar features of maxium DS1683. The DS1683 [1] is an elapsed time recorder. The time is tracked and recorded

S. V. Siddamal (✉) · S. B. Shirol · S. S. Kotabagi · S. B. Hiremath · N. C. Iyer
KLE Technological University, Hubballi, India
e-mail: sarjav@kletech.ac.in

S. B. Shirol
e-mail: suhasshirol@kletech.ac.in

S. S. Kotabagi
e-mail: sujatask@kletech.ac.in

S. B. Hiremath
e-mail: shraddha_h@kletech.ac.in

N. C. Iyer
e-mail: nalinic@kletech.ac.in

based on the EVENT pin held high at the falling edge. The application is to track the power cycle of the device.

A similar lower version maxium DS 1682 [2] counts the number of events on the event pin and records them. The device also records the total time the DS 1682 was last reset to “0.” The authors in paper [3] have designed a test ASIC NKETC2019 using low power on-chip oscillator. The ASIC finds its application as elapsed time counter. The counters count at every 1 s. The oscillator generates a frequency of 32 kHz which is further downscaled to 1 Hz to run the elapsed time counters. DS 1602 [4] is a real-time counter which counts the duration in seconds when Vcc power is applied designed to count seconds when VCC power is applied. In paper [5] CMOS current starved inverter based oscillator is presented, this is used to drive the counters. The trim bits are used to change the frequency at different process corners.

The rest of the paper is organized as follows. The architecture, pin details, descriptions of ASIC are described in Sect. 2. The P and R flow and back-end implementation, timing and power analysis are described in Sect. 3 followed by conclusion and references.

2 Proposed Front-End ASIC

The architecture of the proposed event logger KLEEL2020 is as shown in Fig. 1. KLEEL2020 keeps log of events including duration and number of events occurred. The data is available to the outside world through I²C interface. This ASIC runs on single supply of 3.3 V and internally generates 1.8 V for core power using on-chip regulator. Fully integrated on-chip oscillator generates 1 MHz base clock which

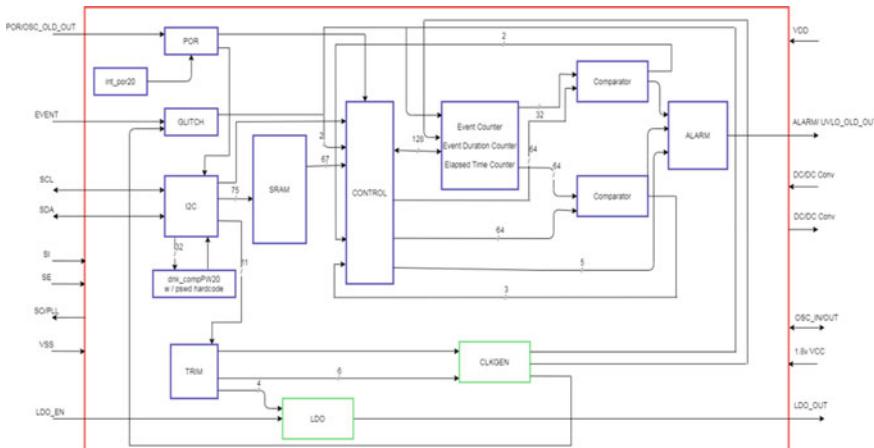


Fig. 1 Architecture of KLEEL2020

is further divided 8 kHz for time counting/recording. On power up, KLEEL2020 loads 32-bit Event Duration Counter (EDC) register and 32-bit Event Counter (EC) register from SRAM. The EDC, ETC, and EC are loaded from outside through I²C interface. In addition, there is a 32-bit Continuous Time Counter (CTC) which runs continuously irrespective of any conditions. All counters are clocked by 8 kHz clock.

2.1 Mixed Signal Design of Event Logger-KLEEL2020

The mixed signal ASIC KLEEL2020 is a 16 pin DIL, and pin description is as shown in Fig. 2.

The features of KLEEL2020 are

- Records the occurrence of events when the time event is high.
- 32-Bit, elapsed time counter to provide the total time of event duration.
- Alarm to indicate various flags and completion of events.
- 4 byte password protection.
- On-chip RC-based Oscillator.
- I²C Compatible interface.
- 16 pin DIL.

3 Back-End Implementation of KLEEL 2020

Cadence Innovous is used for physical implementation. The KLEEL2020 approximately consists of 5000 standard cell instances. The chip is designed using TSMC 0.18 CMOS logic or mixed signal/RF, general purpose technology. It consists of 6 metal routing layers. Table 1 shows the general design information.

The physical design is done using the following steps.

Floor Planning

The placement of macros and blocks is done in floor planning. The cadence tool is used to optimize the place and route to obtain minimum chip area.

Routing

The physical connection to the core is done using routing. The metal layers, vias are used to do physical connection. The actual interconnections between standard cells, pins, and pads are done here. Figure 3 shows the optimized placed and routed design.

Timing Analysis

Timing analysis verifies the timing performance of a design. All possible timing violations in possible paths are verified. The design is modified to meet the timing closure. The timing analysis is performed using the fastest industry tool Tempus. The setup and hold timing are measured for worst case.

(a)

POR/OLD_OSC_OUT	1	16	VCC
EVENT	2	15	LDO_EN
SDA	3	14	LDO_OUT
SCL	4	13	DC_DC_OUT
TESTMODE	5	12	DC_DC_IN
SI/PLL_OUT	6	11	1.8V_VDD
SO/PLL_IN	7	10	OSC_IN/OUT
GND	8	9	ALARM/SE

KLEEL2020

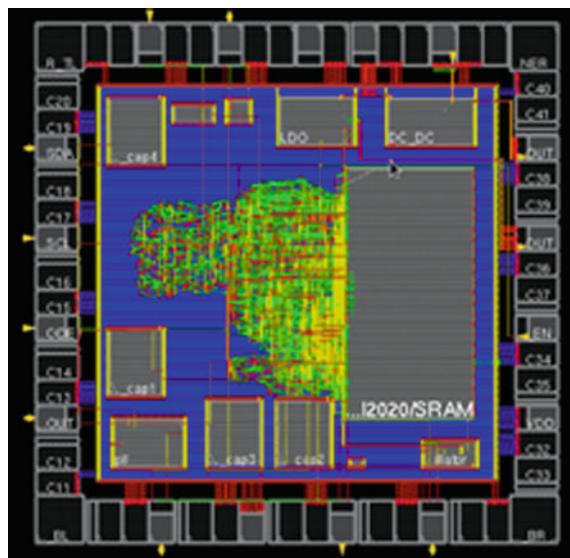
(b)

Sl.no	Pin Name	Description
1.	POR	Bidirectional pin Power on reset signal
2.	EVENT	Event Trigger Signal
3.	SDA	I ² C serial data Input/Output.
4.	SCL	I2C serial clock.
5.	TESTMODE	When '1' ,ASIC is in test mode and '0', ASIC is in Function mode
6.	SI	Test mode input
7.	SO	Scan Out in Test Mode
8.	GND	Ground
9.	ALARM/SE	Function mode Alarm output, Test mode acts as scan enable
10.	OSC_IN/OUT	This is bidirectional pin; Oscillator output and Input Clock signal .
11.	1.8 VDD	1.8V Supply for core
12.	DC-DC IN	DC/DC converter Input
13.	DC-DC_OUT	DC/DC converter Output
14.	LDO OUT	LDO output: This pin gives output 1.8V from LDO
15.	LDO_EN	LDO Enable: This pin is used to Enable or Disable the LDO block
16.	VCC	3.3V DC input voltage

Fig. 2 **a** 16 pin DIL, **b** pin description of KLEEL2020

Table 1 Design information

Design status	Routed
Design name	Chip
#Instances	5250
#Hard macro	2
#Std. cells	5193
#Pad	55
#Net	2210
#Special net	3
#IO pins	11
Average pins per net	3.608

Fig. 3 Optimized placed and routed KLEEL2020

Planning

It is necessary to analyze the power grid of the chip. The power needs to be distributed to every part of the design. Volta IC power integrity is used for power measurement. Table 2 shows the power requirement for typical voltage of 1.8 V. In the similar way, power analysis is done for worst and best voltage cases.

Table 2 Power requirement for typical voltage 1.8 V

Internal power	Switching power	Total power	Leakage power
1.199	0.2886	1.497	0.009996

Table 3 Test cases passed

S. No.	Pin name	Description
1	Reset	When POR goes low, even through the event is high, there is no alarm
2	SE	When scan enable is high, the design does not function in functional mode
...
...
34	Overflow bit set	The bit is set high if ED count is equal or exceeds the overflow limit value
35	SRAM write	Check if valid/reliable data transfer takes place at different address
36	SRAM read	Check if valid/reliable data transfer takes place at different address

Testing of IC

A complete framework end to end functional verification of KLEEL2020 using System Verilog Methodology (SVM) is done to check the functional precision of the design before the tape-out. Total of 36 test cases are tested. Few test cases are depicted in Table 3. Few design issues found were fixed in RTL.

Scan chain is inserted to check the ASIC after tape-out. The objective is to make the testing easier observing the flip/flop transition. This is done in test mode. Figure 4 shows the scan chain coverage report where 100% flip/flops are mapped.

Physical Verification

Physical verification of KLEEL 2020 is done using mentor calibre tool. The design is verified and has cleared all calibre DRC rules. The area consumed for the design is 2,449,225 μm^2 .

Figure 5 a shows the GSD II streamed out file sent to TSMC foundry Taiwan for mini@sic runway on January 23rd 2020. Figure 5b shows the fabricated mixed signal ASIC.

Scan mapping status report		
<hr/>		
Category	Number	Percentage
Scan flip-flops mapped for DFT	1590	100.00%
Flip-flops not mapped for DFT		
non-DFT scan flip-flops	0	0.00%
flip-flops not scan replaceable	0	0.00%
flip-flops not targeted for DFT	0	0.00%
<hr/>		
Totals		100.00%

Fig. 4 Scan chain coverage report

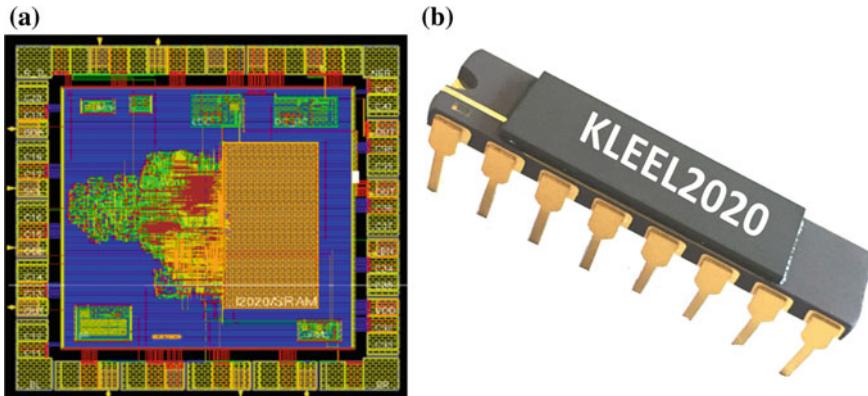


Fig. 5 a Signed off GDSII of KLEEL2020, b Fabricated KLEEL2020

4 Conclusion

The work is the effort to give complete design to chip tape-out experience for students and faculty of tire II University. The authors have tried to present the entire work in this paper. The team has used various tools like Genus, Innovous, Virtuoso, Voltas, Tempus from Cadence and calibre from mentor graphics. The ASIC, event logger is implemented in TSMC mixed signal CMOS 0.18 μm technology. The on-chip RC oscillator draws a maximum current of 100 μA from battery. The ASIC has LDO to provide constant supply voltage of 1.8 V to core. The power consumed for typical case (1.82 V) is 1.497 mW. The die size of the core is 1565 $\mu\text{m} \times 1565 \mu\text{m}$. The ASIC is taped out to Taiwan at mini@sic runway on January 22nd 2020 for fabrication.

Acknowledgements The authors would like to thank Dr. Anand Bariya, VP SiFive, for his valuable discussion on P and R flow, and Shri Shripad Annigari, SiFive for his valuable discussion on analogue blocks. The authors would like to thank Mentor Graphics for providing Calibre tool and University for funding this work under ‘Institutional Research Project’ (IRP) scheme.

References

1. <https://datasheets.maximintegrated.com/en/ds/DS1683.pdf>
2. <https://datasheets.maximintegrated.com/en/ds/DS1682.pdf>
3. S.V. Siddamal, S.B. Shirol, S. Hiremath, N.C. Iyer, Design and physical implementation of mixed signal elapsed time counter in 0.18 μm CMOS technology, in *VLSI Design and Test. VDAT 2019. Communications in Computer and Information Science*, vol. 1066 (Springer, Singapore, 2019)
4. <https://datasheets.maximintegrated.com/en/ds/DS1602.pdf>
5. S. Kalburgi, et al., Ultra low power low frequency on-chip oscillator for elapsed time counter, in *32nd International Conference on VLSI Design and 18th International Conference on Embedded Systems (VLSID-2019)*, ©2019 IEEE, pp. 251–256. <https://doi.org/10.1109/vlsid.2019.00062>

Invasive Weed Optimized Drift Parameter Involved Differential Protection Scheme



D. J. Abhishek, P. Jagadeesh, N. Srinivasu, and Ch. Durga Prasad

Abstract A new drift parameter-oriented differential positive sequence current component-based fault detection scheme is proposed in this paper. The optimal value of the drift parameter is identified with invasive weed optimization (IWO) algorithm and incorporated in cumulative sum process of fault detection index (FDI). This process makes the fault detection scheme as threshold free and also enhances the dependability of the method.

Keywords Fault detection · Invasive weed optimization (IWO) · Drift parameter · Sequence components

1 Introduction

Application of intelligent algorithms to engineering problems produces recommended solutions compared to existing approaches. These algorithms are extensively used in relaying applications such as phasor measurement unit (PMU) placement [1–3], fault classification and location [4], etc. Recently, these techniques were used in threshold identification process of fault detectors (FD's) [5–7] to enhance the overall performance during typical fault events. Particle swam optimization (PSO) was used extensively in such areas. Still several challenges are alive in this domain.

Absolute sum and cumulative sum approaches are widely used in fault detector applications. Cumulative sum-based fault detectors are superior than absolute sum-based methods. This cumulative sum approach is proposed in 2008 for quick fault

D. J. Abhishek (✉) · P. Jagadeesh · N. Srinivasu · Ch. Durga Prasad

Department of Electrical and Electronics Engineering, SRKR Engineering College, Bhimavaram, India

e-mail: abhishekguntur19@gmail.com

P. Jagadeesh

e-mail: jagadeesh.pothula@gmail.com

Ch. Durga Prasad

e-mail: chdpindia@gmail.com

detection in transmission lines [8] and performed directly on instantaneous current signal. Later, the method is used in different perspectives [9–11]. Instead of applying cumulative sum concept on time domain signals, the reliability of the schemes enhanced with other input signals [9]. In this line, sequential components of voltage and/or current-based approaches are also available in literature for accurate fault detection [12].

In this paper, a new cumulative sum of differential positive sequence current component-based fault detection scheme is implemented which is driven by IWO evaluated drift parameter. For this purpose, an objective function is composed with fault location, fault inception, and fault resistance values. These values are randomly generated by IWO and final optimal value is processed for calculation of fault detection index. Section 2 provides this information of drift parameter setting. Section 3 presents test system details and results of IWO evaluated drift parameter. Section 4 presents simulation results for various faults followed by discussions and finally, Sect. 5 ends with conclusions.

2 Proposed Method

Fundamental component of positive sequence currents on two ends of the transmission line is calculated with the help of discrete Fourier transform (DFT) to implement the proposed scheme. Let I_1^S is the positive sequence current at the sending end and I_1^R is the positive sequence current at the receiving end, and then the differential current component is given by

$$I_1^{\text{diff}} = |I_1^S - I_1^R| \quad (1)$$

Using the differential component, cumulative sum calculated to detect faults in unit in between the sending and receiving ends is given by

$$G(k) = G(k - 1) + \max(I_1^{\text{diff}} - \beta, 0) \quad (2)$$

In Eq. (2), $G(k)$ is the cumulative sum index at k th sample and β is the drift parameter. Existing approach to find β is the peak attained by the actuating signal during normal operating condition. Therefore, it leads to threshold setting problem to discriminate faults from non-fault disturbances. This β value changes with location, inception and resistance of fault and hence the problem is three-dimensional. To avoid additional threshold setting, IWO algorithm is used to find optimal value of the β so that the index rises only in case of faults with location, inception and resistance of fault as variables. Alireza Mehrabian and Caro Lucas proposed IWO algorithm from the inspiration of spreading, growing, and producing strategies of seeds in the year 2006 [13]. Initialization process of IWO is similar to other heuristic algorithms and the solutions are updated using standard deviation σ of the random function will be reduced from a previously defined initial value (σ_{initial}) to a final value (σ_{final}), in

every iteration/generation using Eq. (3) given by

$$\sigma_{\text{iter}} = \frac{(\text{iter}_{\max} - \text{iter})^n}{\text{iter}_{\max}^n} \cdot (\sigma_{\text{initial}} - \sigma_{\text{final}}) + \sigma_{\text{final}} \quad (3)$$

This searching and updating process terminates at the final iteration. The best value available at the final iteration is taken as global optimum value. Since the problem may attain global best for any of single line to ground faults, three separate runs need to be executed on the test system in order to find the optimal drift parameter which enhances the overall fault detection results.

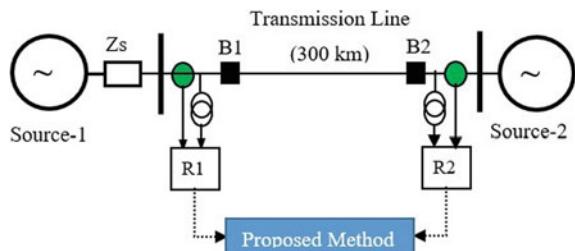
3 Optimal Drift Parameter for Test System

The process of identification of optimal drift parameter and the efficacy of the proposed method with such drift parameters are tested on a standard 2-bus power system whose model is shown in Fig. 1. Specifications of the test system are available in [5].

Fundamental positive sequence current components are extracted using DFT and differential component is calculated using Eq. (1). This component is further processed to cumulative index with drift parameter as a function of system fault variables. Line-to-ground faults are taken into account for identification of optimal drift parameter. Since identification of such peak value with case by case evaluation is difficult and there is no guarantee for actual peak arrival which provides reliable protection outputs. This is evident from Fig. 2 where three single line to ground faults are simulated by varying fault parameters and corresponding indices are plotted. Therefore, this problem is solved by IWO and the results are reported in Table 1.

Among all three optimal values of drift parameter, 0.6254 is global optimal drift parameter and which is taken for further calculation of final cumulative fault index to detect all types of faults.

Fig. 1 Test system for simulation



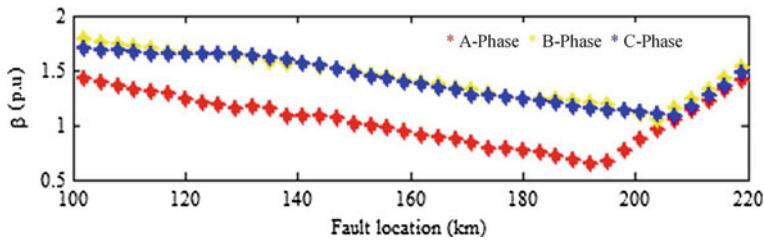


Fig. 2 Possible values of drift parameters for various faults

Table 1 Optimal drift parameter values for L-g faults

L-g	Location (km)	Time (s)	R (Ω)	β
A-g	283.3313	0.0499	75.1951	0.6254
B-g	158.4777	0.0412	12.6578	0.7335
C-g	277.9565	0.0309	62.2489	0.6573

4 Performance Assessment of Proposed Scheme

The proposed IWO-driven fault detection scheme is tested with different faults by varying the fault parameters and corresponding results are provided below.

4.1 Line to Ground Faults

Figure 3 shows the detection plot of BG fault at 180 km from measured end with 30Ω fault resistance. This fault is initiated at 0.05 s and detected by the proposed scheme at 0.054 s with a delay of 4 ms. The involvement in optimal drift parameter in the detection process provide quick and reliable decision during this fault.

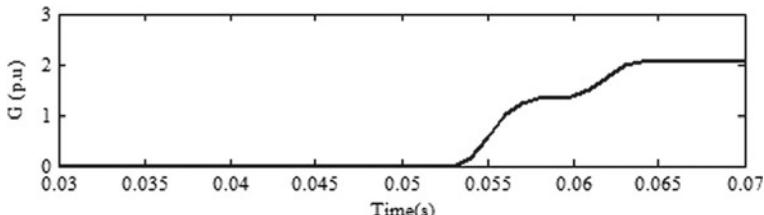


Fig. 3 Response of proposed scheme during B-g fault

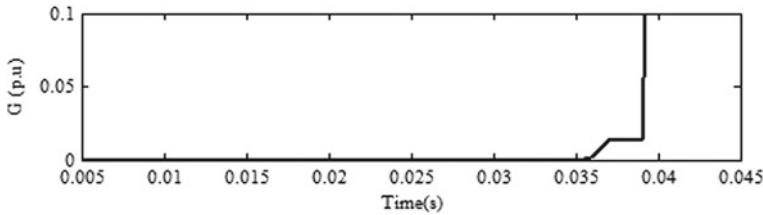


Fig. 4 Response of proposed scheme during A-B fault

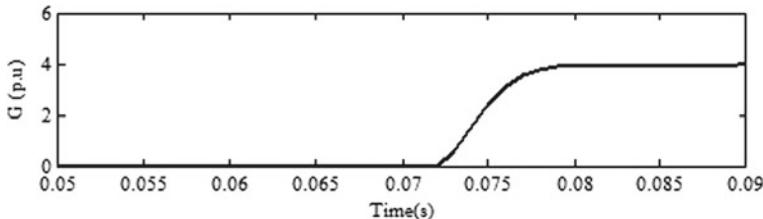


Fig. 5 Response of proposed scheme during A-B-C-g fault

4.2 Multi-phase Faults

When A-B fault is simulated with fault parameters of 220 km location, 0.034 s inception time and 5Ω fault resistance, the response of proposed technique is presented in Fig. 4.

4.3 Symmetrical Faults

Figure 5 shows the proposed method response during symmetrical fault. For this case generation, an A-B-C-g fault is simulated at 130 km from relay end with 10Ω of fault resistance initiated at 0.07 s. In this case also fault is detected within 5 ms.

4.4 High Resistive Ground Faults

To test the reliability of the proposed scheme during high resistive faults occurrence, an A-g fault is created at 80 km with fault resistance of 200Ω whose detection plot is presented in Fig. 6. The detection time is 13 ms. which is slightly high because of high fault resistance.

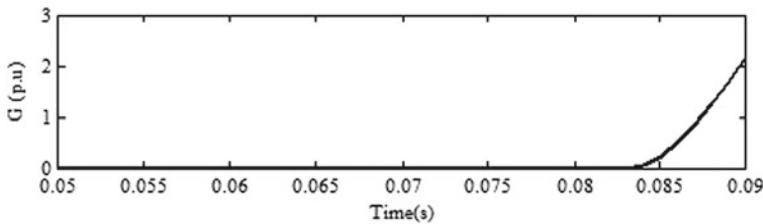


Fig. 6 Response of proposed scheme during A-g high resistive fault

Aforementioned simulation results show the impact of optimal drift parameter on quick and threshold-free decision of the FD irrespective of type of fault and there parameters.

5 Conclusions

IWO evaluated drift parameter concept is introduced in this paper for quick and reliable fault detection in transmission lines. This mechanism also eliminates the additional threshold setting to discriminate faults from other disturbances. This method enhanced the overall reliability of the algorithm and ensured quick outputs to isolate faulty components from rest of the system.

References

1. M.J. Damborg, R. Ramaswami, S.S. Venkata, J.M. Postforoosh, Computer aided transmission protection system design part I: algorithms. *IEEE Trans. Power Appar. Syst.* (1), 51–59 (1984)
2. T.K. Maji, P. Acharyya, Multiple solutions of optimal PMU placement using exponential binary PSO algorithm for smart grid applications. *IEEE Trans. Ind. Appl.* **53**(3), 2550–2559 (2017)
3. M. Hajian, A.M. Ranjbar, T. Amraee, A.R. Shirani, Optimal placement of phasor measurement units: particle swarm optimization approach, in *International Conference on Intelligent Systems Applications to Power Systems* (IEEE, 2007), pp. 1–6
4. J. Upendar, C.P. Gupta, G.K. Singh, G. Ramakrishna, PSO and ANN-based fault classification for protective relaying. *IET Gener. Transm. Distrib.* **4**(10), 1197–212 (2010)
5. C.D. Prasad, P.K. Nayak, Performance assessment of swarm-assisted mean error estimation-based fault detection technique for transmission line protection. *Comput. Electr. Eng.* **71**, 115–128 (2018)
6. C.D. Prasad, M. Biswal, P.K. Nayak, Wavelet operated single index based fault detection scheme for transmission line protection with swarm intelligent support. *Energy Syst.* (2020)
7. C.D. Prasad, M. Biswal, Swarm intelligence-based differential protection scheme for wind integrated transmission system. *Comput. Electr. Eng.* **1**(86), 106709 (2020)
8. S.R. Mohanty, A.K. Pradhan, A.A. Routray, A cumulative sum-based fault detector for power system relaying application. *IEEE Trans. Power Deliv.* **23**(1), 79–86 (2007)
9. M.A. Jarrahi, H. Samet, Modal current and cumulative sum based fault detection approach in transmission lines. *Int. J. Emerg. Electr. Power Syst.* **19**(6) (2018)

10. H. Laaksonen, New passive islanding detection method based on cumulative sum of change in voltage positive sequence phase angle. *Int. Rev. Model. Simul. (IREMOS)* **6**(5), 1482–1488 (2013)
11. J. Zafar, J. Gyselinck, *CUSUM based fault detection of stator winding short circuits in doubly-fed induction generator based wind energy conversion systems* (ICREPQ Grana-da, Spain, 2010)
12. A. Rahmati, R. Adhami, A fault detection and classification technique based on sequential components. *IEEE Trans. Ind. Appl.* **50**(6), 4202–4209 (2014)
13. S. Karimkashi, A.A. Kishk, Invasive weed optimization and its features in electromagnetics. *IEEE Trans. Antennas Propag.* **58**(4), 1269–1278 (2010)

Breast Cancer Detection and Classification Employing Convolutional Neural Networks and Support Vector Machine for Histopathological Images



K. S. Babu and N. R. Sharvani 

Abstract In recent years, the main causes of cancer deaths that are most concerning are breast cancer. Convolution neural networks (CNN) approach is deep learning that is used to overcome many problems faced on a few featured based approaches. The CNN approach methodology explains the classification of an eosin-stained breast biopsy and also hematoxylin (H&E). The extracted features by CNN could also be utilized for the training of a support vector machine classifier. The approach consists of firstly identifying the data sources where few datasets are considered of different patients and identifying the problem. Furthermore an exploratory data analysis consists of few visualization techniques that are performed using python libraries. Finally pre-processing the data here, the feature extraction is carried out. A predictive model is constructed using support vector machine (SVM) for the classification of breast tumors. The proposed system predicts if the tumor is malignant or benign and its subclasses.

Keywords Biopsy · Hematoxylin · Convolution neural network · Support vector machine

1 Introduction

Cancer is the most common public health problem in the world today. Specifically, breast most cancers are one of the most usual kinds of most cancers among women. Breast cancer is always observed in the breast when they start to grow without any control. Breast cancer can be diagnosed using various methodologies by considering medical images tested for patients like histology and radiology. The radiology tested samples of images help in identifying the areas which are abnormal and can also locate

K. S. Babu  · N. R. Sharvani

Department of Electronics and Instrumentation, SJCE, JSS Science and Technology University, Mysuru, India

e-mail: babuks@sjce.ac.in

N. R. Sharvani

e-mail: sharvanin994@gmail.com

it. But these analyses do not help in determining if the area located is cancerous. The simplest way to study if cancers are present or not wherein tissue of the affected area is studied under a microscope which is called biopsy. A data is collected and formed out of histology images of breast cancer patients which are named “BreakHis.” A CNN, i.e., convolution neural networks or also known as convnet, is one of the deep learning knowledge of algorithms that inputs an image and assigns learnable weights and biases to a number of components in the fed images for evaluation which allows in classifying among benign and malignant tumors [1, 2].

1.1 Related Work

Stenkvist [3] has done a brand-new approach is CNM: Computerized nuclear morphometric which provides a multivariate numerical score; this is related to nuclear atypia. Zhang et al. [4] adopted a novel technique VMD, i.e., variational mode decomposition for automatic diagnosis of breast cancer which applied multilevel iterative. The method takes histopathological images of datasets of multidimensional spatially evolving indicators. The average accuracy of the classifiers threefold and tenfold go-validation techniques is 89. 61% and 88. 23% respectively are resulted. Spanhol et al. [5] worked on a dataset of 7909 breast most cancers histopathology images gathered on 82 patients. Here, automatic type of sampled datasets is performed for the two classes. The consequences can be a computer-aided diagnosis device for the clinician. The accuracy of the result is between 80 and 85%.

2 Methods

The proposed work block diagram is as given in Fig. 1 which is an outline on how the classification of tissues in the breast can be implemented using the convolution neural network algorithm using the VGG16 model using a dataset collected. In this stage, the main focus is the prediction of breast tumors, and their subclasses are done. Pyramid matching and then they are classified. 125 images are taken for each of the tumor types mentioned above, out of which 70% of images are taken for training, 15% of images for validation, and 15% of images for testing the samples [6].

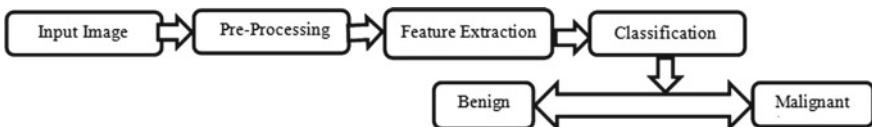


Fig. 1 Block diagram of the classification of breast cancer

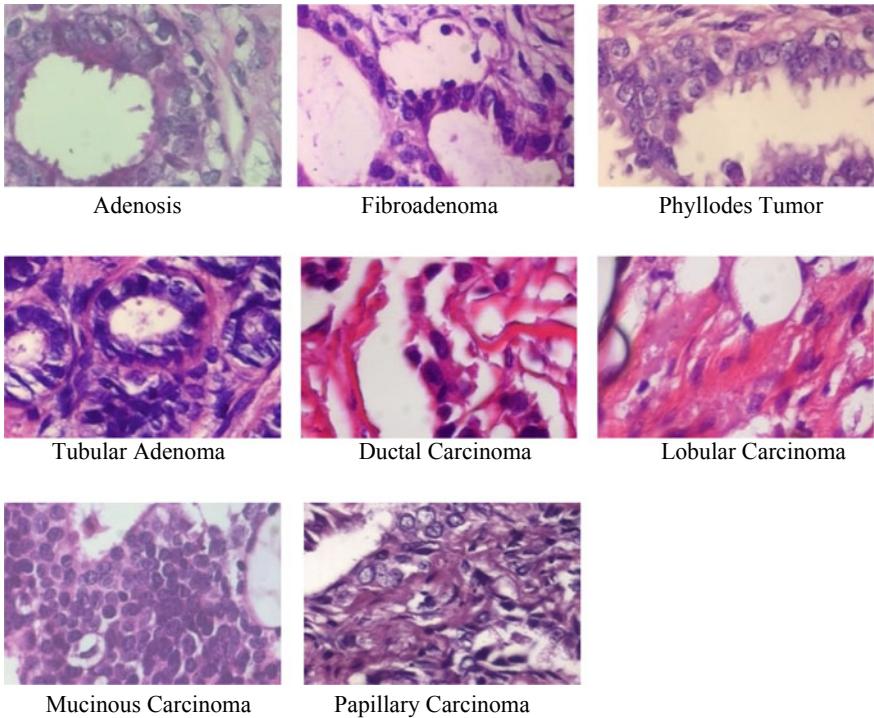


Fig. 2 Samples of histopathological images from each type of breast tumor

2.1 *Image Dataset*

The dataset consists of a number of microscopic biopsy images of benign and malignant tumors. It consists of a total of 7909 images, out of which 125 images of each tumor type are taken for training from BreakHis dataset. The slides on which the tumors collected are mandatorily stained with hematoxylin and eosin (HE) (Fig. 2) [7].

2.2 *Pre-processing*

In other pre-processing steps of keras and tensorflow, the data do not get pre-processed, instead the model training process gets utilized [8]. Image augmentation reduces memory overhead which is very much important using TensorFlow. It is especially important when many networks are running parallel on the same machine. Generally, few pre-processing steps involved are:

- Uniform aspect ratio

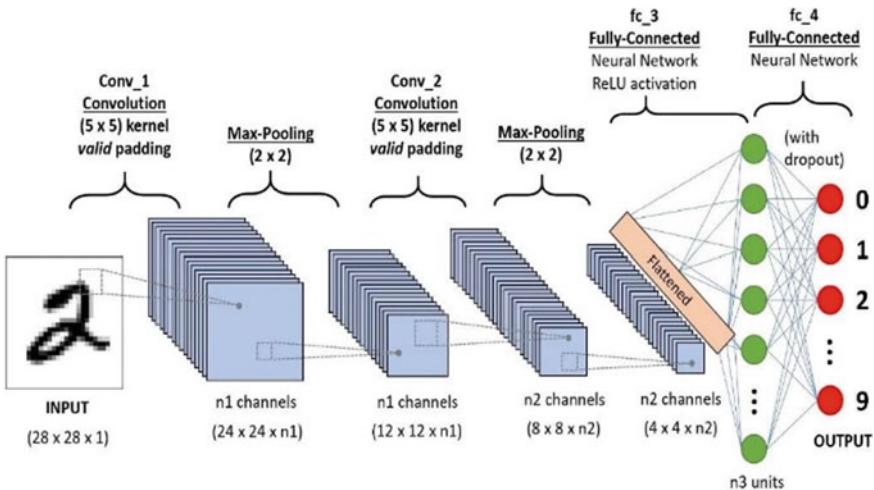


Fig. 3 Convolution neural network layers

- Image resizing
- Normalizing image inputs
- Dimensionality reduction.

2.3 Feature Extraction

In the feature extraction step, convolution neural network is used which is a deep learning technique that uses learnable weights and biases for feature extraction and fully connected layers for the classification of breast cancer which have ability to learn these filters/traits [9, 10]. Figure 3 depicts general sequence of convolution layers. The convolution layer applies the convolution operation to input image where it has a tensor shape of image length and image width [11]. The convolutional layers are shown in Table 1.

2.4 Classifiers

Five distinct classifiers had been used to evaluate the mentioned characteristic devices: SVM, k-nearest neighbor (knn), logistic regression, random forests (rf), and decision tree. A k-nearest neighbor is a genre of prevalence which is used both classification and regression that gives contribution to neighbors assigning weights of $1/d$, where d is distance to the neighbor. SVM is a well-known algorithm used for classification which performs linear and nonlinear model. Logistic regression is the

Table 1 Convolution layers

Convolution layer	Filter size	Feature maps
1st	3×3	64
2nd	3×3	96
3rd	3×3	128
4th	3×3	256
5th	3×3	256

Fully connected with 2048 hidden units

Fully connected layer with number of hidden units is equal to number of classes

Softmax layer

shape of the statistical model which makes use of logistic feature and binary established variable. A random forest area is a chorus method that incorporates decision tree predictors. The precept behind is that vulnerable inexperienced persons make a group (in this case the decision trees) and unite collectively and make a strong learner.

3 Results

The complete proposed system analysis and results depend on CNN+SVM Classification and augmentation and its performance with the applied dataset [10]. The results output is seen in accuracy, i.e., percentage terms for each of the tests. The classifier lists which are seen in the output are logistic regression, decision tree, linear SVM, and random forest along with accuracy and F1-score. The general formula for the score calculation is as below:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (1)$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (2)$$

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{TN} + \text{FN}} \quad (3)$$

The precision, recall, F1-score, and accuracy values in percent for CNN models applied for the augmented dataset for 200 epoch value are obtained from confusion matrix after the training of images are as given in Table 2.

Magnification of images increases the number of pixels, thus increasing the details of images making the features clearer, hence increasing the accuracy of the classifiers for $400\times$ magnification. The accuracies of different classifiers of different magnification values for augmented dataset in terms of percentage are in Table 3.

Table 2 Precision, recall score, F1-score, and accuracy

Magnification	Precision	Recall	F1-score	Accuracy obtained
40×	99.5	96.62	98.0	97.91
100×	1	97.4	98.7	98.7
200×	1	99.1	99.59	99.58
400×	99.7	99.5	99.52	99.71

Table 3 Accuracies of different classifiers

Magnification	Logistic regression	Linear SVM	Random forest	Decision tree	kNN
40×	76.7	73	65.8	44.9	49.8
100×	83.6	80.9	71.4	53.2	56.8
200×	80.8	79.2	68.1	46.9	55.3
400×	84.1	79.4	69.5	47.9	56.7

Once after the training, coming to the testing of samples is successfully done where the image given as input will be selected and the selected tumor type will be displayed in the output window.

4 Conclusion

In this work, different classifiers, i.e., linear SVM, logistic regression, decision tree, k-nearest neighbor, and random forest are used for breast cancer classification that is histopathological images which are collected from BreakHis dataset and are executed by way of configuring with CNN performance. Convolutional neural networks are one of the best deep learning techniques to replace traditional classifiers which helps to increase the performance by accurate feature extraction and classification.

Acknowledgements The authors of paper thank Dr. Udayashankara (Professor, JSS Science and Technology) for his support and providing assistance in work.

References

1. A. Ferrari, S. Lombardi, A. Signoroni, Bacterial colony counting with convolutional neural networks in digital microbiology imaging. *Sample Recogn.* **61**, 629–640 (2017)
2. C. Szegedy, et al., Going deeper with convolutions, in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June (2015), pp. 1–9
3. B. Stenkivist et al., Computerized nuclear morphometry as an objective method for characterizing human cancer cell populations. *Cancer Res.* **38**(12), 4688–4697 (1978)

4. Y. Zhang, B. Zhang, F. Coenen, W. Lu, Breast cancer diagnosis from biopsy images with highly reliable random subspace classifier ensembles. *Mach. Vis. Appl.* **24**(7), 1405–1420 (2012)
5. F.A. Spanhol, L.S. Oliveira, C. Petitjean, L. Heutte, A dataset for breast cancer histopathological image classification. *IEEE Trans. Biomed. Eng.* **63**(7), 1455–1462 (2016)
6. L. He, L.R. Long, S. Antani, G. Thomas, Computer assisted diagnosis in histopathology. *Sequence Genome Anal. Methods Appl.* 271–287 (2010)
7. M. Veta, J. Pluim, P. van Diest, M. Viergever, Breast cancer histopathology image analysis: a review. *IEEE Trans. Biomed. Eng.* **61**(5), 1400–1411 (2014)
8. R. Girshick, J. Donahue, T. Darrell, J. Malik, Rich feature hierarchies for accurate object detection and semantic segmentation, in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, June (2014), pp. 580–587
9. Y.M. George, H.H. Zayed, M.I. Roushdy, B.M. Elbagoury, Remote computer-aided breast cancer detection and diagnosis system based on cytological images. *IEEE Syst. J.* **8**(3), 949–964 (2014)
10. T. Araújo et al., Classification of breast cancer histology images using convolutional neural networks. *PLoS ONE* **12**(6), e0177544 (2017)
11. O. Abdel-Hamid, A.-R. Mohamed, H. Jiang, L. Deng, G. Penn, D. Yu, Convolutional neural networks for speech recognition. *IEEE/ACM Trans. Audio, Speech Language Process.* **22**(10), 1533–1545 (2015)
12. M. Kowal, P. Filipczuk, A. Obuchowicz, J. Korbicz, R. Monczak, Computer-aided diagnosis of breast cancer based on fine needle biopsy microscopic images. *Comput. Biol. Med.* **43**(10), 1563–1572

Lobe Correction of a Miniaturized Microstrip Antenna Using SIW Cavity



Seshadri Binaya Behera, Sraddhanjali Mohapatra, and Subhrakanta Behera

Abstract A miniaturized inset-fed microstrip antenna with slot and substrate integrated waveguide (SIW) has been investigated in this work. Two orthogonally overlapped slits on the radiator surface along with a SIW cavity are introduced on the patch surface. The combined effect of slot loading and SIW technique achieves around 35% of antenna miniaturization. The metallic cavity formed by the SIW technique reduces the surface waves leading to lateral radiation and improves the radiation characteristics of the antenna. The unwanted back lobe and side lobes have been mitigated, and thus, the radiation performance of the proposed structure has been improved. The improved directivity in the broadside direction leads to gain improvement. The gain response of the structure has been improved up to 4.85 dBi. The impedance bandwidth of this structure is 500 MHz over 10.5 GHz, and it is well suitable for X-band applications. The overall dimension of the designed structure is $0.6\lambda \times 0.7\lambda$.

Keywords Lobe correction · Miniaturization · Substrate integrated waveguide (SIW)

1 Introduction

The past few decades have witnessed a rapid advance in the wireless communication system which expedites the need for integrating very compact devices in the system. This demand for small devices makes antenna miniaturization a popular research consideration. Based on the dielectric substrates, with both the radiator and ground layer of metal conducting plane punctured with metallic holes, the substrate integrated waveguide (SIW) structures provide a very compact and flexible solution for integrating the circuit elements into the same planar structure. This SIW technology is a promising method which can offer the combined features of both the slot antenna like the low profile and the cavity backed antenna features like directional

S. B. Behera · S. Mohapatra (✉) · S. Behera

School of Electronics Engineering, KIIT University, Bhubaneswar, Odisha, India

e-mail: sraddhanjalimohapatra@gmail.com

properties. The authors in [1] have explained the modeling and design considerations of SIW structures for different applications. An ultra-miniaturized SIW cavity backed slot antenna with 87% miniaturization has been reported in [2]. Filter application along with high miniaturization has been achieved in [3]. A Minkowski fractal gap combined with SIW structure [4] with dual frequency and orthogonal linear polarization has been discussed. A meander lined unit cell with metallic vias [5] proposes a novel miniaturization technique with very high directivity. This metallic cavity configuration also contributes to offer high radiation efficiency, which has been discussed in [6–8]. Although antenna size reduction is very important for the modern wireless communication system, it comes with several drawbacks like degraded radiation pattern and narrow impedance bandwidth. The diffraction of surface waves at the dielectric edges gives rise to lateral radiation and degrades the antenna directivity. Numerous analyses have been carried out in the past to reduce the lateral radiation. Amendola et al. [9] provide an insight on use of a shorted patch to reduce surface waves. In this work, an inset-fed microstrip patch antenna with a cross-shaped slot has been proposed. The metallic vias placed along the edges of the proposed patch creates a SIW cavity. This metallic cavity facilitates the surface waves to transform into space waves, and thus, the lateral radiation manifested by the original structure gets directed into the broadside direction. Also, the combined effect of the slots and the SIW cavity influences the resonant frequency of the patch structure to shift backward and achieve miniaturization.

2 Proposed Antenna Configuration

The proposed antenna structure shown in Fig. 1 is designed using a single-layer Taconic TLY substrate with a substrate height of 1 mm with relative permittivity, $\epsilon_r = 2.2$, and loss tangent of 0.0009. The single-layer construction adopts a dual-side metallic laminated surface on which the antenna layout is embedded on upper surface. The initial antenna layout consists of a rectangular patch of dimension 23 mm \times 20 mm with inset feeding technique, while the ground plane dimensions are 27 mm \times 20 mm. This patch without any irregularity exhibits resonance over 14.5 GHz with an impedance bandwidth of 425 MHz. Basically, antenna miniaturization can be realized by increasing the electrical length of the current flow on the radiator surface. To attain that, two orthogonally overlapped slits are etched onto the patch surface to perturb the surface current. However, the slots alone were not sufficient to achieve antenna size reduction. Hence, SIW cavity has been introduced in antenna 3. The length and width of the feed are $L_f = 7$ mm and $W_f = 2.5$ mm, respectively. The diameter of the metallic vias is $d = 1$ mm, and the distance between the adjacent metallic vias is $h = 1$ mm. The width of the slot along the width of the patch is $L_{s1} = 16$ mm and the width of the slot along the patch width is $W_{s1} = 0.5$ mm. Similarly, the width of the slot along the length of the patch is $L_{s2} = 16$ mm, and the width of the slot along the length of the patch is $W_{s2} = 0.5$ mm. The indentation gap by the

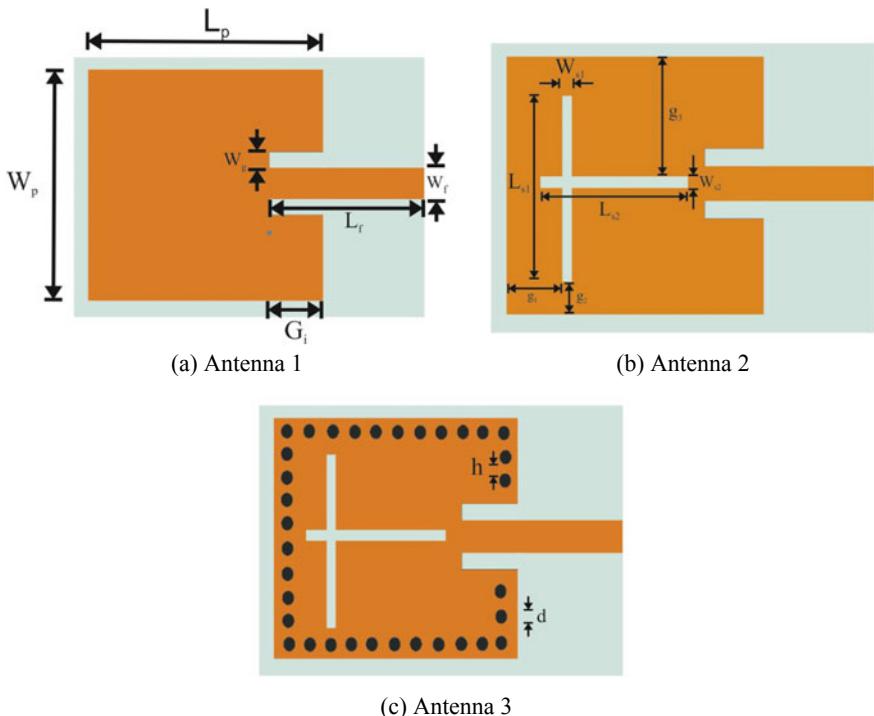


Fig. 1 Antenna layout

inset feed, G_i , is 3.2 mm. The distances between the patch edge and the slots have been labeled in the figure as $g_1 = 6.5$ mm, $g_2 = 1.95$ mm, $g_3 = 10$ mm.

3 Results and Discussion

3.1 Miniaturization of the Proposed Structure

The proposed design has been numerically analyzed using Ansys high-frequency structure simulator (HFSS) and the simulated reflection characteristics have been illustrated in this section. A comparison of the S_{11} parameters of all the three antenna configurations is depicted in Fig. 4. As it can be observed the initial antenna configuration, i.e., antenna 1 resonates at 14.7 GHz. Later by the introduction of the orthogonally overlapped slits, the resonance has shifted to 14.8 GHz. Though it is expected to obtain a frequency shift toward the lower range, the slot position might be responsible for this error. Again, substrate integrated waveguide technique has been adopted to achieve antenna size reduction. After utilizing SIW technique, the resonating

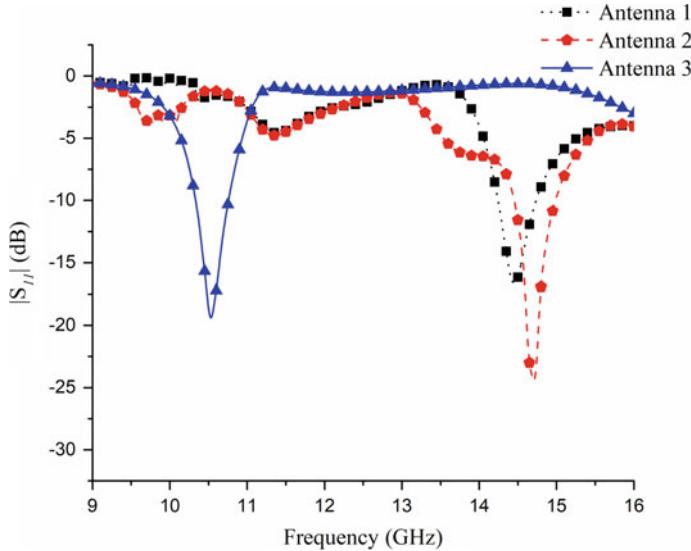


Fig. 2 S-parameter comparison of different antenna configurations

frequency shifts back to 10.5 GHz with a better gain response. The metalized holes placed along the patch edges create a cavity around the patch, and these result in a robust suppression of the inductive part of the shorted patch, consequently restricting operation over a higher frequency. Thus, 35% miniaturization has been achieved. But it can be seen that the return loss characteristics degrades a bit after miniaturization Fig. 2.

3.2 Lobe Correction of the Proposed Antenna

Figure 3 compares the radiation characteristics of all the three antenna configurations. From Fig. 3, it can be observed that antenna 1 and antenna 2 both have undesired back lobes and side lobes which indicate poor radiation characteristics. This degraded radiation is caused due to the collision of surface waves at the substrate truncations, which give rise to lateral radiation. These surface waves can be reduced and directed as space waves, and thus, the radiation characteristics can be improved. This is achieved by perforating metallic via along the edges of the patch which create a SIW cavity. This restricts the surface wave diffraction at the edges and directs it in the broadside direction. Hence, lobe correction is carried out in antenna 3 as shown in Fig. 3. The simulated gain response is shown in Fig. 4 which compares the simulated gain responses of all the antenna configuration. It can be observed that the antenna 1 has a very low gain which is around 0 dBi, and this is further improved by introducing

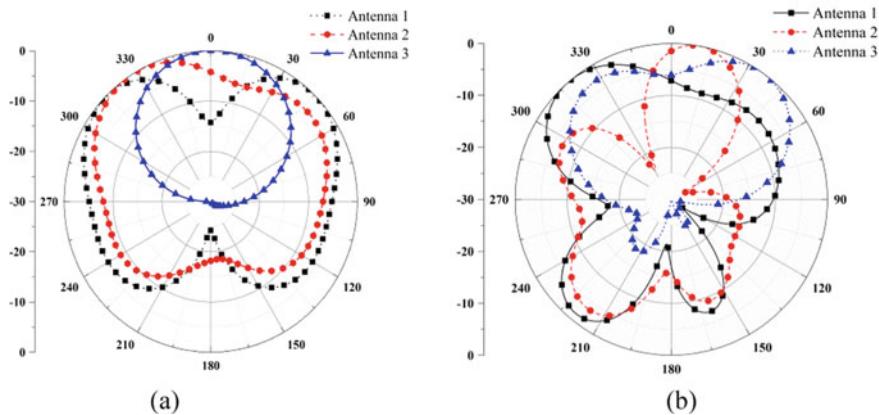


Fig. 3 Radiation pattern (a) at E-plane (b) at H-plane

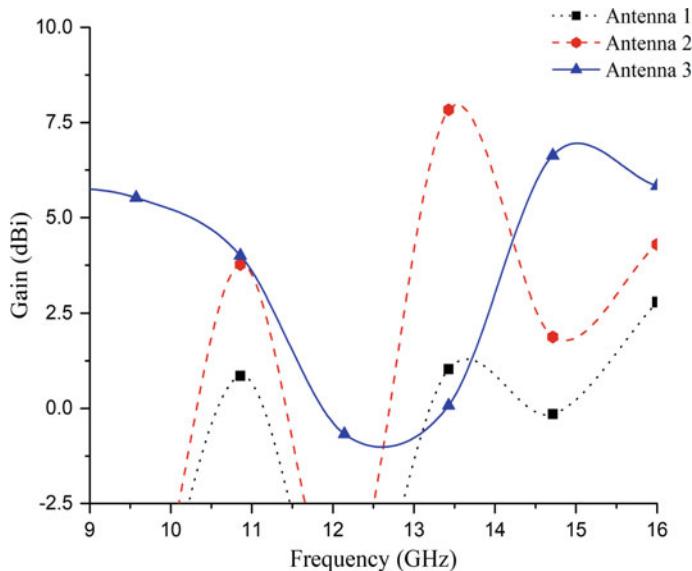


Fig. 4 Simulated gain response

the slots in antenna 2. A gain improvement of 4 dBi has been by using SIW cavity in antenna 3 as illustrated in the figure.

4 Conclusion

In this work, an inset-fed microstrip antenna with a dimension of $(0.6\lambda \times 0.7\lambda \times 0.03\lambda)$ has been proposed. Two orthogonal slits are introduced on the surface combined with metallic vias along the patch edge forming a SIW cavity. This antenna resonates over 10.5 GHz which earlier resonated over 14.8 GHz, thus achieving 35% of miniaturization. This structure exhibits impedance bandwidth of 500 MHz with gain response of 4.85 dBi. This can be useful for various radar applications like weather monitoring, air traffic control and vehicle speed tracking.

References

1. M. Bozzi, F. Xu, D. Deslandes, K. Wu, Modeling and design considerations for substrate integrated waveguide circuits and components, in *8th International Conference on Telecommunications in Modern Satellite, Cable and Broadcasting Services*, Nis (2007)
2. A.P. Saghati et al., An ultra-miniature SIW cavity-backed slot antenna. IEEE Antennas Wirel. Propag. Lett. **16**, 313–316 (2017)
3. L. Li, Z. Wu, K. Yang, X. Lai, Z. Lei, A novel miniature single-layer eighth-mode SIW filter with improved out-of-band rejection. IEEE Microwave Wirel. Compon. Lett. **28**(5), 407–409 (2018)
4. Z. Fu, et al., Dual-frequency miniaturized substrate integrated waveguide quarter-mode cavity-backed antenna based on minkowski fractal gap with orthogonal polarization radiation characteristics. Hindawi Int. J. Antennas Propag. **2019**, Article ID 1816763, 9 pages. (2019). <https://doi.org/10.1155/2019/1816763>
5. Y. Dong, T. Itoh, Miniaturized substrate integrated waveguide slot antennas based on negative order resonance. IEEE Trans. Antennas Propag. **58**(12), 3856–3864 (2010)
6. M. Asaadi, A. Sebak, High-gain low-profile circularly polarized slotted SIW cavity antenna for MMW applications. IEEE Antennas Wirel. Propag. Lett. **16**, 752–755 (2017)
7. Z. Wu, J. Liu, J. Zhang, M. Lin, X. Huang, Design of a Ka-band high-gain antenna with the quasi-annular SIW corrugated technique. IEEE Antennas Wirel. Propag. Lett. **18**(5), 1001–1005 (2019)
8. D. Guan, C. Ding, Z. Qian, Y. Zhang, Y.J. Guo, K. Gong, Broadband high-gain SIW cavity-backed circular-polarized array antenna. IEEE Trans. Antennas Propag. **64**(4), 1493–1497 (2016)
9. G. Amendola, L. Boccia, G.D. Massa, Surface wave radiation from a shorted elliptical patch antenna, in *IEEE Antennas and Propagation Society International Symposium. Digest. Held in conjunction with: USNC/CNC/URSI North American Radio Sci. Meeting* (Cat. No.03CH37450), Columbus, OH, 2003, vol. 1 (2003), pp. 601–604

Implementation of Custom DRC in TSMC 0.18 μm for Optimized Layout



B. H. Shraddha, Nagaratna Shanbhag, Saroja V. Siddamal,
and Nalini C. Iyer

Abstract This paper aims in implementation of DRC rules in TSMC 0.18 μm PDK. The main focus is on methodology employed to implement rules for optimization with respect to area and delay of the circuit under design. PDK consists of foundry specific rule files which are used in electronic design automation (EDA) tools. A prominent step in chip design process is physical verification of the layout for a particular circuit. It is critical to verify that the layout does not violate any design rules, and that it matches the schematic design. The DRC rule deck focuses on primitive rules such as spacing between different tracks and minimum widths. This work focuses on implementation of different rules such as metal width, poly, and contact and via placement which do not exist in the rule deck. Standard verification rule format (SVRF) manual is referred while writing snippets for different rules. Rules are written for TSMC 0.18 μm technology and are verified in Mentor Graphics Calibre tool.

Keywords DRC · PDK · Standard verification rule file · Calibre

1 Introduction

Large-scale integration of various subcomponents on to a single wafer is the requirement of the present era. It means that the design to be developed should have smaller area, consumes less power and functions reliably under all environmental conditions. In VLSI engineering, any design that has to be fabricated has to undergo a rule check which imposes a geometric constraint on the devices and interconnects in the entire circuit under design. A floor plan is prepared to develop a layout of the design which has the information of device placement and interconnect routing. Post-functional verification, a rigorous physical verification has to be formed on the layout. This ensures the production of IC at an acceptable yield.

B. H. Shraddha (✉) · N. Shanbhag · S. V. Siddamal · N. C. Iyer
KLE Technological University, Hubballi, India
e-mail: shraddha_h@bvb.edu

Fabrication of an integrated circuit involves drawing layout of the circuit under design. As the size of the semiconductor device is shrinking, the complexity of the layout increases motivating the manufacturers to optimize the process of the design. The designer is not able to manually reduce the component/semiconductor design in an efficient manner. Hence, this in turn necessitated the fabrication foundries to develop optimized process design kit (PDK) intellectual property (IP) which facilitates reusability of IP and other general or specialized problem building blocks. This also demanded to enhance the number of design files and provide better control over uniformity. Electronic design automation is used extensively to ensure that designers do not violate design rules and this process is called design rule check (DRC) [1]. This process also includes LVS check, XOR checks, and ERC and antenna checks. On further stage, for micro- or nanoscale geometries, more restricted rules have to be included for better yield [2–4]. The verification of the design at the later stage of the layout is very much complicated and has to be efficiently performed using an edge of the envelope EDA software. The objective of the work presented in this paper caters in

- Defining and implementing layout constraints for the design.
- Understanding SVRF from Mentor Graphics.
- Editing the rules for defined layout constraints using SVRF.
- Verifying the results in the Calibre tool.
- The organization of the paper is as follows:

Section 2 deals with the framework of arriving at certain rules by undergoing through study of existing PDK file. Section 3 discusses the implementation where different design rules are described using Calibre tool. Section 4 describes the results obtained on running the modified rule file on the example layout and its discussion and lastly Sect. 5 discusses on conclusion of the work.

2 Framework

Design rules are parameters given by semiconductor manufacturers (foundry) to the designers, in order to ensure the correctness of a mask set during fabrication process [5]. This rule set describes certain restrictions in geometry and connectivity to ensure that the design has sufficient margin to take care of any variability in manufacturing process. This paper discusses the DRC rule writing using design rule file from leading semiconductor foundry called Taiwan Semiconductor Manufacturing Company (TSMC).

TSMC 0.18 μm rule deck consists of some the typical DRC rules like interior rules, exterior rules, enclosure rules and extension rules. Interior rules are responsible for defining the distance of interior facing edge for a single layer or multilayer. Some of the important interior rules in TSMC rule deck are diffusion minimum width, N-well minimum width, poly minimum width, metal minimum width, etc. Likewise, exterior rules define the distance of exterior facing edge of two layers. These rules

are also known as spacing rule. Some of the spacing rules in TSMC are contact minimum spacing, minimum poly spacing, minimum metal spacing and minimum space from N+ diffusion or active. Distance between the inside edge to outside edge is defined by enclosure rule. This includes defining the minimum enclosure of N select, minimum enclosure of contact, minimum enclosure of P+ diffusion or active and minimum enclosure by poly in the referred rule deck. Extension rules define the distance between inside edges to outside edge. Some of the examples are minimum poly extension on diffusion (minimum poly overhang) and minimum diffusion extension on poly. TSMC rule also consists of antenna rules that are commonly expressed as a ratio of wire area over gate area (A_w/A_G) for each metal and cut (via) layer. Antenna rules have not been considered for this work as it increases the complexity. Layout optimization techniques focus on reducing the parasitic effects, power dissipation and interconnect capacitance in the design [6]. Also, these techniques improve the switching power, layout area and delay in the circuit [7]. Designer implements these techniques by minimal length, minimal width routing and other methods in the design phase. Initial emphasis was portrayed on customizing the layout for the circuit under design rather than optimizing it.

The authors in the paper propose seven rules which do not exist in the rule deck. As a case study, inverter circuit is considered. Customization of inverter circuit layout is done by writing backend code in standard rule file, i.e., calibre rule deck. Standard verification rule file (SVRF) manual is referred for syntax and commands [7]. Code in the form of snippets is written from the scratch based on the predefined customization methods. These methods were later implemented on simple inverter layout. Following customization methods are considered for rule writing:

1. Metal rule: Figure 1 shows a pictorial representation of three metal layers running parallel. When external spacing ‘d’ which defines the distance between two parallel running metals is in the range from 0.46 μm to 1.15 μm , the width of the middle metal M1 region should be increased by 1 μm by left and right. This technique reduces the resistance in the circuit.
2. Poly rule: The reactance of the circuit increases if the poly length exceeds the diffusion area. Maximum poly length outside the diffusion is defined in order to overcome this condition. DRC rule is written so that if the poly length is outside the diffusion more than 5 μm then tool throws the error. Figure 2 illustrates the same rule pictorially.

Fig. 1 Metal spacing

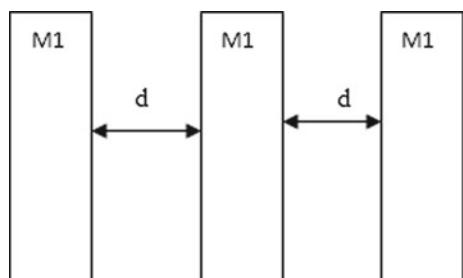
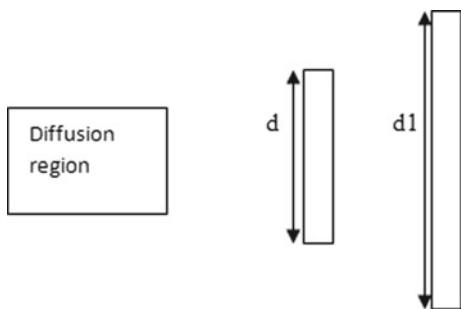
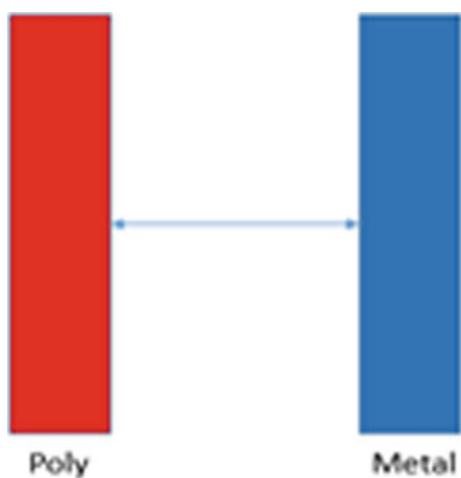


Fig. 2 Poly rule spacing

3. Poly metal rule: In TSMC DRC rule deck, there is no rule specified for maintaining certain minimum distance between poly and metal layers. If metal is immediate layer after the poly and this metal carries a huge current, then there is chance of inducing electrons on the poly which will alter the characteristics of that device. Figure 3 describes the 2D view of the rule mentioned.
4. Contact rules: To build a circuit with minimum RC, extracted view is essential to obtain better post layout simulations. While placing N number of devices in the layout, a check should be performed on the placement of two contacts. The rule deck provides a rule for violating minimum space of contact placement. But a rule should be defined for reducing the resistance of the circuit under design by measuring the distance between the two contacts. If the measured data is greater than predefined value, then an error should pop up to include contacts in between. Figure 4 describes the same.
5. Poly overhang: This rule also discusses on reduction in the resistance of the circuit in layout by connecting the gates of two devices using metal and avoiding the placement of the horizontal polygon. Figure 5 describes the same.

Fig. 3 Poly metal spacing

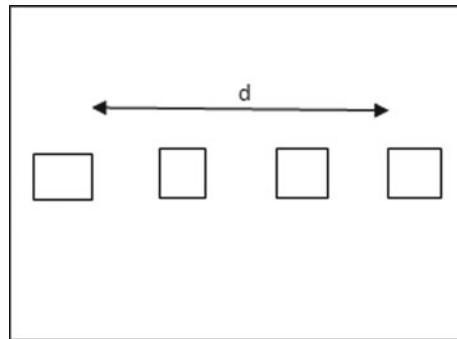


Fig. 4 Contact rule spacing

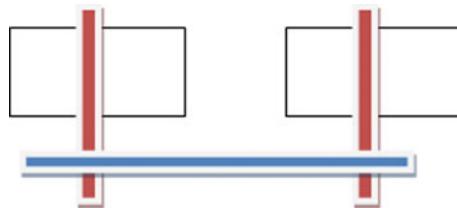


Fig. 5 Replacement of poly by metal

6. Via rule: If two same power lines (VDD/VSS) are running parallel one above the other, add as much as via's possible. This reduces the overall resistance (Fig. 6).
7. Twin well rule: This rule defines the external distance between the N+/P+ diffusion and twin well. Figure 7 describes the block level representation of the same rule.

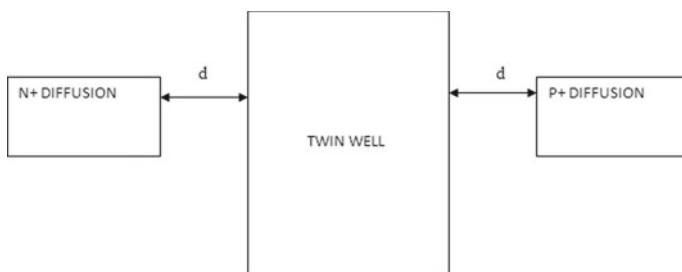


Fig. 6 External distance between the N+/P+ diffusion and twin well

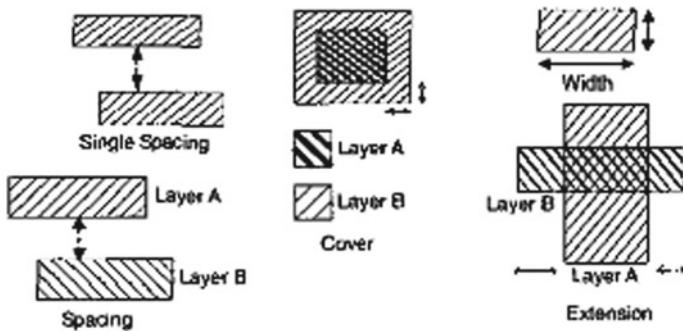


Fig. 7 Sample DRC rule deck

3 Implementation

3.1 Describing Design Rules Using Calibre Tool

The rules specified in the PDK rule deck define the least possible width of different layer tracks, track separation and extension of polygons. Basic Boolean operations such as NAND, AND, NOR and OR are used on layers containing a set of geometric forms which usually are rectangles. A pseudo-layer is created due to the intersection or union of rectangles that is present in combination of operand layers. Figure 7 illustrates certain rules specified in DRC rule deck. There are approximately forty to seventy various design rules where each and every design is applied on millions of polygons. This verification process is carried out repeatedly in a structured manner. The rules specified do not consider the optimization section of the layout. Validation of the design rules requires the description of design constraints in appropriated language.

3.2 Algorithms

Algorithms are written for the different optimized design rules mentioned in Sect. 2.

1. Metal rule: For metal rule, a particular metal layer is defined with the required minimum and maximum dimensions. If any layer has the dimension less than the defined dimension in the layout then that particular metal is made to grow till the required dimensions are met. Following snippet dictates the rule (Fig. 8).

Snippet: M2.S.2.3 {@ GROW METAL IF SPACING IS IN RANGE FROM 0.46 to 1.15X = EXT M2 ≤ 0.46 < 1.15 ABUT < 90 SINGULAR REGION GROW X RIGHT BY 1 LEFT BY 1}

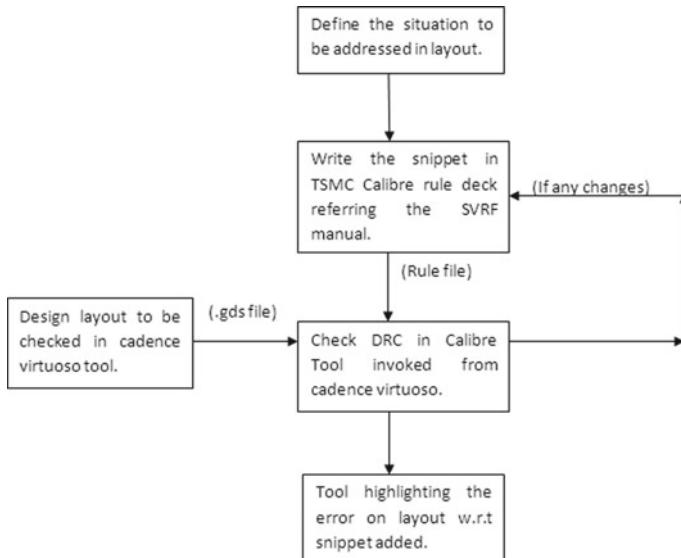


Fig. 8 Algorithm of design rule insertion and checking

2. Poly Rule: For poly rule the region of poly outside the diffusion region is measured and a declaration of maximum poly length is performed and detected in the layout. Following snippet dictates the rule.

Snippet: P0.S.2.3{ @TO DEFINE MAX POLY LENGTH OUTSIDE THE DIFFUSION REGIONX = poly OUTSIDE 0DLENGTH X \geq 5 }

3. Contact rule: The distance between two contacts inside the diffusion region is measured and if the distance between contacts is greater than predefined distance in the layout then an error stating to add multiple contacts in the empty space should be popping up.
4. Poly Overhang: Poly present outside the diffusion region which is connected in horizontal path is detected and a DRC error pops up indicating connection of a metal instead of poly to increase the performance of the circuit under design.
5. Via Rule and Twin Well Rule: To decrease the parasitic resistance via placement in different regions is required. Two power lines running parallel are detected and via are added in that region. For twin well rule the external distance between N+/P+ diffusion region with twin well is measured.

4 Results and Discussions

Cadence Virtuoso tool is used to design the layout for which later customized DRC's are applied. The layout rules are designed for metal and poly. Figures 9 and 10 shows

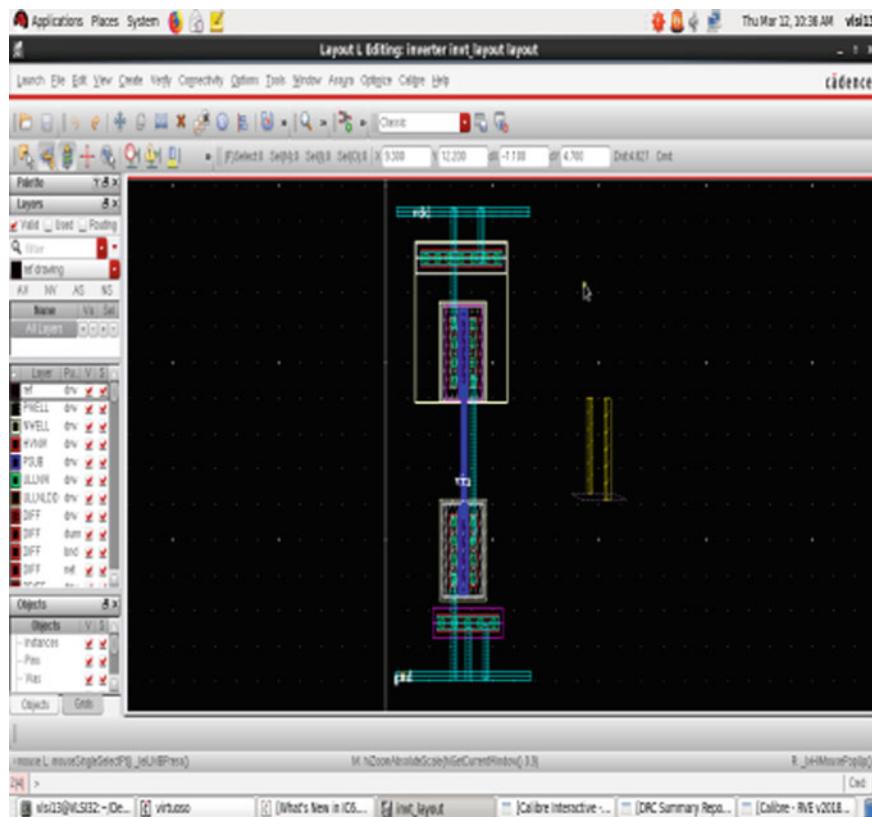


Fig. 9 Cadence virtuoso tool highlighting metal error in layout

the metal error in the layout and calibre for the customized DRC. Figures 11 and 12 highlights the poly error. The DRC code snippets are added to calibre TSMC 0.18 μm rule deck for defined layout constraints in order to customize the layout of design using standard verification rule format (SVRF). All the edited results from layouts with and without the modified techniques are successfully checked and verified in Mentor Graphics Calibre tool. The DRC code snippets are added to calibre TSMC 0.18 μm rule deck for defined layout constraints in order to customize the layout of design using standard verification rule format (SVRF). All the edited results from layouts with and without the modified techniques are successfully checked and verified in Mentor Graphics Calibre tool.

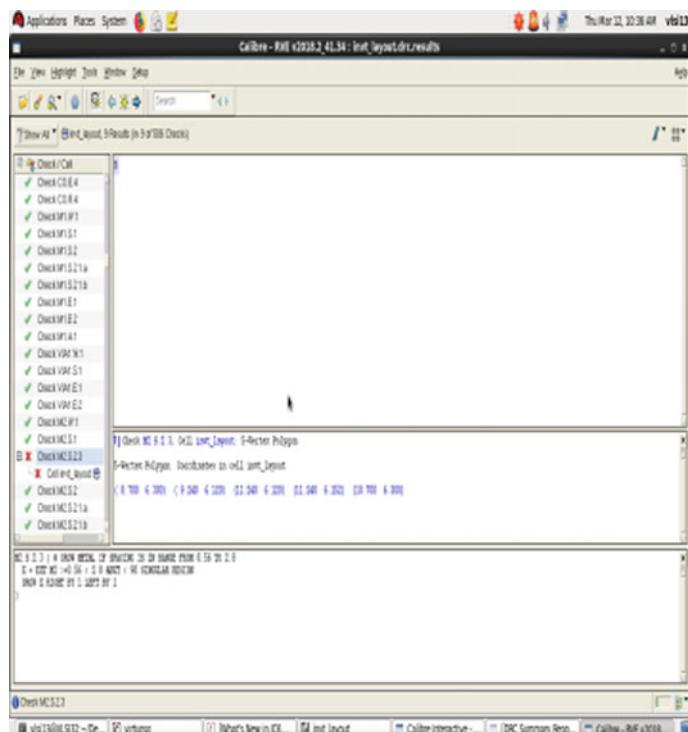


Fig. 10 Calibre tool showing metal error

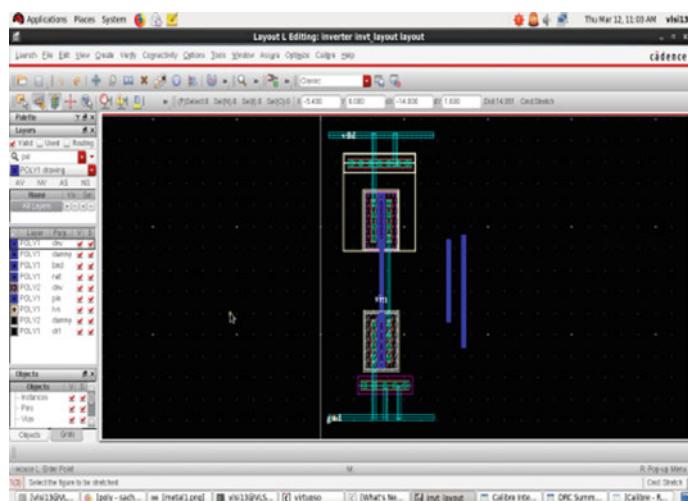


Fig. 11 Cadence virtuoso tool highlighting poly error in layout

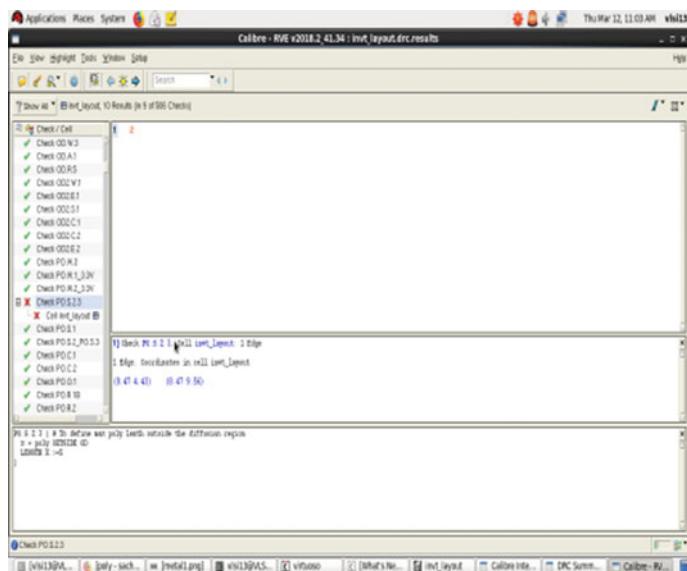


Fig. 12 Calibre showing poly error

5 Conclusions

The authors have an effort to understand rule writing studying the existing calibre rule deck. Few rules which do not exist are implemented and tested on a case study: Inverter. Cadence Virtuoso tool is used to design the layout for which later customized DRC's are applied. The layout rules are designed for metal and poly. The DRC code snippets are added to calibre TSMC 180 nm rule deck for defined layout constraints in order to customize the layout of design using standard verification rule format (SVRF). All the edited results from layouts with and without the modified techniques are successfully checked and verified in Mentor Graphics Calibre tool.

References

1. V. Patel, Senior Engineer (einfochips Ltd) Design Rule Checks (DRC)—A Practical View for 28 nm Technology
2. N. Shanbhag, B. Kiran, Process variability analysis of 14 nm FinFET. IJSERT (2014)
3. L.T. Clarka, V. Vashishtha, L. Shifrenb, A. Gujja, S. Sinhac, B. Clinec, C. Ramamurthy, Gregeric, aArizona State University, Tempe, San Jose, USA ASAP7: A 7-nm finFET predictive process design kit
4. J.E. Stine, J. Chen, I. Castellanos, G. Sundararajan, M. Qayam, P. Kumar, J. Remington, S. Sohoni, Electrical and Computer Engineering Department, Oklahoma State University, Stillwater, USA FreePDK v2.0: Transitioning VLSI education towards nanometer variation-aware designs

5. B.H. Shraddha, S. Manish, V.S. Saroja, N.C. Iyer, M. Prathvish, UVLO system for elapsed time counter using UMC180 Process technology, GCAT-2019
6. K. Sachin, G. Deven, H. Sampath, S. Rohit, A. Shripad, H. Shraddha, V.S. Saroja, K. Sujata, C.I. Nalini, Ultra low power low frequency on chip oscillator for elapsed time counter, VLSID-2019
7. Calibre SVRF (Standard Verification Rule Format) Manual from Mentor Graphics Corporation

Design of Current Mode MOS Logic for Low-Power Digital Applications



Amit Bakshi, S. N. Mishra, and Sandeep Kumar Dash

Abstract Today, the major limitation of designing of high-speed integrated circuits (ICs) with conventional technology is the delay limiting the switching speed of the gates. As evident from already established logic styles like complementary pass transistor logic (CPL), differential cascode voltage swing logic (DCVSL), etc., the circuit exploits the property of noise reduction due to differential inputs. Ideally current mode circuits have constant current gain with no input impedance and finite output impedance. The current gain of the circuit is set to unity as current amplification leads to higher consumption of static power. MOS current mode logic (MCML) is graced with certain advantages which include low level of noise generation, static power dissipation independent of switching activity, low voltage swing, a weak dependence of propagation delay on fan-out load capacitance, lower power dissipation at higher frequencies, etc., out way the few disadvantages such as elaborated design process and increased number of design parameters. In this paper, we present an in-depth study of MCML base approach in which the analysis of low-power applications is performed at the target data rate of 1 Gbps. Work has been done on a standard CMOS technology of 0.18 μm .

Keywords Current mode · High performance · Low power · MCML · Digital IC

A. Bakshi (✉) · S. N. Mishra · S. K. Dash

School of Electronics Engineering, KIIT Deemed to be University, Bhubaneswar, India

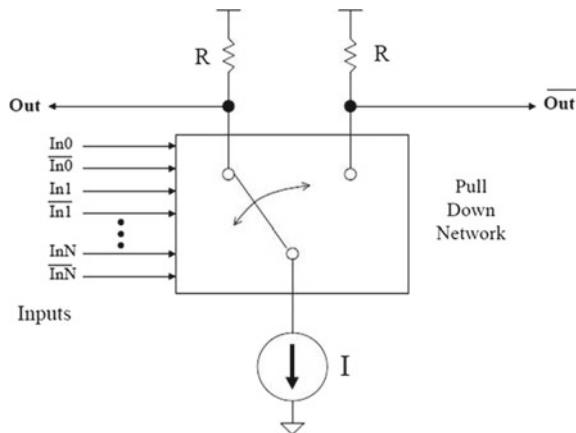
e-mail: amitfet@kiit.ac.in

S. N. Mishra

e-mail: snmishrafet@kiit.ac.in

S. K. Dash

e-mail: sandeepfet@kiit.ac.in

Fig. 1 Basic MCML gate

1 Introduction

In today's world, demands for power efficient device are increasing exponentially. To make battery life more durable, power efficient device design is the need of the hour. Generally, the total power dissipation of a CMOS device consists of static and dynamic power. The static power dissipation occurs due to leakage in MOSFETs. The dynamic power dissipation occurs due to charging and discharging of capacitive load. To reduce the device size and increase the speed of device technology, scaling is done. With the help of technology scaling, a more number of transistors are embedded in device. Due to this, more functionality is added in the device which improves the performance. But due to addition of more transistors, switching activity increases which results in large dynamic power dissipation. Technology scaling enhances leakage current due to which static power dissipation increases [1]. For high-speed applications, MCML is preferred over other techniques [2, 3]. The MCML techniques suffer from large static power dissipation problem [4]. Owing to its numerous advantages, MCML becomes preferred design over other MOS logic styles [5] for digital circuit applications [6, 7]. MCML gives a reduced output voltage swing compared to CMOS circuits thus making it suitable for higher speed applications. The power dissipation using MCML is independent of its frequency of operation [8]. For higher frequency application, MCML design is preferred due to low power dissipation. Constant current in the supply rails of MCML technique reduces the noise as well as crosstalk owing to low voltage swing. In digital logic circuits, signal integrity improves due to differential nature at the inputs of the MCML.

2 MOS Current Mode Logic Operation

The continental structure of MCML block [8] is shown in Fig. 1 which have three most essential components such as pull-up network, pull-down network, and a constant current source. The PDN accepts only differential inputs, i.e., the true and complemented form of all the signals at the inputs. Current ‘ I ’ is originated by the constant current source while the output voltage swing is determined by the load resistance R . The governing principle of the circuit is current steering [2]. Depending on the concept used, PDN directs the current ‘ I ’ to one of the pull-up devices. The voltage drops across the resistor connected to the current source via PDN is equal to $\Delta V = I \times R$. Theoretically, zero current flows through the other branch. The total voltage swing is determined by the magnitude of the current ‘ I ’ steered and the load devices value.

2.1 Derivation of MCML Gate Performance

The basic MCML block is modeled as RC network [3] shown in Fig. 2. The transient analysis of simplified MCML model based on RC network is based on following assumptions:

1. symmetrical output loads
2. the pull-down network is considered to be an ideal switch shifting from left to right at time $t = 0$.

The MCML circuit based on RC network before and after $t = 0$ depends on switch position that can be modeled as shown in Fig. 3a, b, respectively.

A circuit is assumed to have ‘ N ’ identical gates arranged in linear chain, all having similar load capacitance’s denoted by ‘ C ’. So, the total propagation delay for the chain of gates will be proportional to:

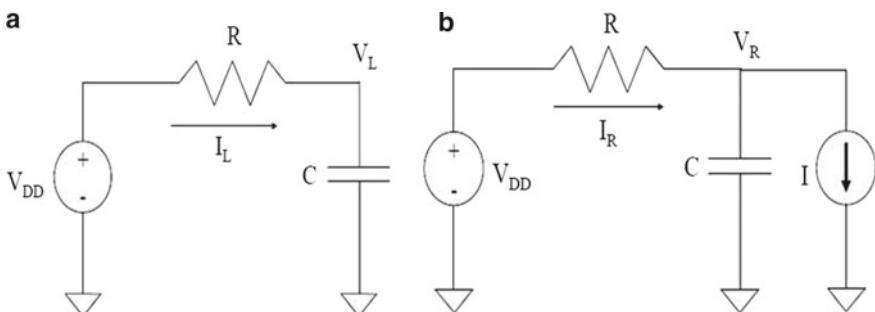
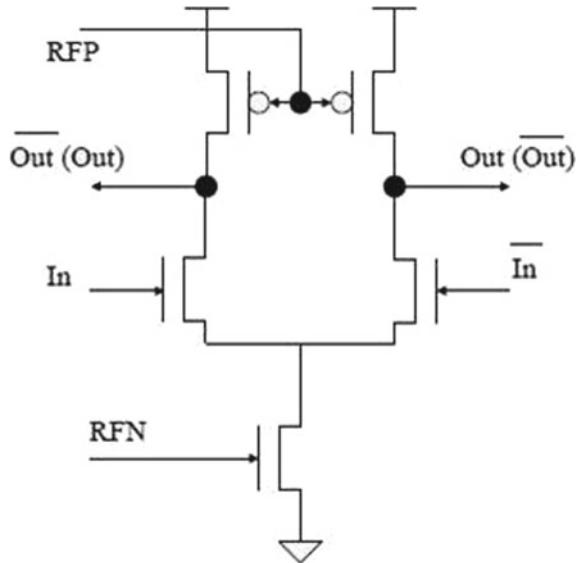


Fig. 2 a, b Equivalent RC model of MCML

Fig. 3 MCML inverter/buffer



$$D_{\text{MCML}} = NRC = \frac{N \times C \times \Delta V}{I} \quad (1)$$

In digital circuits, the power dissipation can be categorized into static and dynamic power components [6]. Thus, in case of MCML under first order consideration, the total power dissipation is a constant. Following this assumption, power, power delay product, and energy delay product can be represented as,

$$P_{\text{MCML}} = N \times I \times V_{dd} \quad (2)$$

$$PD_{\text{MCML}} = NIV_{dd} \times \frac{N \times C \times \Delta V}{I} = N^2 \times C \times \Delta V \times V_{dd} \quad (3)$$

$$ED_{\text{MCML}} = N^2 \times C \times \Delta V \times V_{dd} \times \frac{NC\Delta V}{I} = \frac{N^3 \times C^2 \times V_{dd} \times \Delta V^2}{I} \quad (4)$$

The following equations can be used for comparison with static CMOS logic based on parameters such as power, delay, power delay, energy delay:

$$D_{\text{CMOS}} = \frac{N \times C \times V_{DD}}{\frac{K}{2} \times (V_{DD} - V_T)^\infty} \quad (5)$$

$$P_{\text{CMOS}} = N \times C \times V_{DD}^2 \times \frac{1}{D_{\text{CMOS}}} \quad (6)$$

$$PD_{\text{CMOS}} = N \times C \times V_{DD}^2 \quad (7)$$

$$ED_{\text{CMOS}} = N^2 \times 2 \times \frac{C^2}{K} \times \frac{V_{\text{DD}}^2}{(V_{\text{DD}} - V_T)^\infty} \quad (8)$$

where k and ∞ are parameters dependent on process and transistor size.

2.2 MCML Inverter

The primary design is the MCML inverter/buffer [2] as shown in Fig. 3. The inverter/buffer circuit structure is topologically identical to MCML which employs differential logic to switching the output or input for sense. The load devices in the circuit are implemented using PMOS transistors. They have a fixed gate voltage and to model as resistors, operated in the linear region. It is desirable to make the width of these transistors as small as possible to increase the linearity and to decrease the capacitance. The RFP voltage is controlled by the variable controlled swing shown in Fig. 3.

The PDN is implemented using differential pair of NMOS transistors. The main goal of the pull-down network is to steer the current from the current source to any part of the branch. Theoretically, all current flows through single path and zero current flows through the other path but in reality, some current also flows through the ‘OFF’ path which causes the reduction in the voltage swing. Although the increase in the transistor width increases the voltage gain, it too results in the increase the parasitic capacitance which leads to a trade off between performance and robustness. Thus, in order to preserve the performance, gain must be kept minimum.

A NMOS transistor working in saturation region with a fixed gate voltage act as a current source. Area and robustness form the main trade-off in selecting the size of the current source device. Preference is given for (non-minimum length) transistor to achieve a high output impedance and better current matching ratio. The voltage at the gate of the current source must be kept a few hundred mv from both the power rails if it is determined by a current mirror.

2.3 MCML Tristate Buffer

To remove the influence of the device from the rest of the circuit is the main concept behind high impedance state. Designing tristate gates in CMOS logic is easy [2] since an output of the CMOS gates is connected to any one of the power rails at a time. Therefore, turning off the paths connecting both the supplies to the output is easy in CMOS. In case of current mode logic, the output is constantly connected to the positive supply via load devices which remains always ON.

A switch-based tristate buffer shown in Fig. 4 has two PMOS pass transistors (M6-M7) as reported in [1, 2].

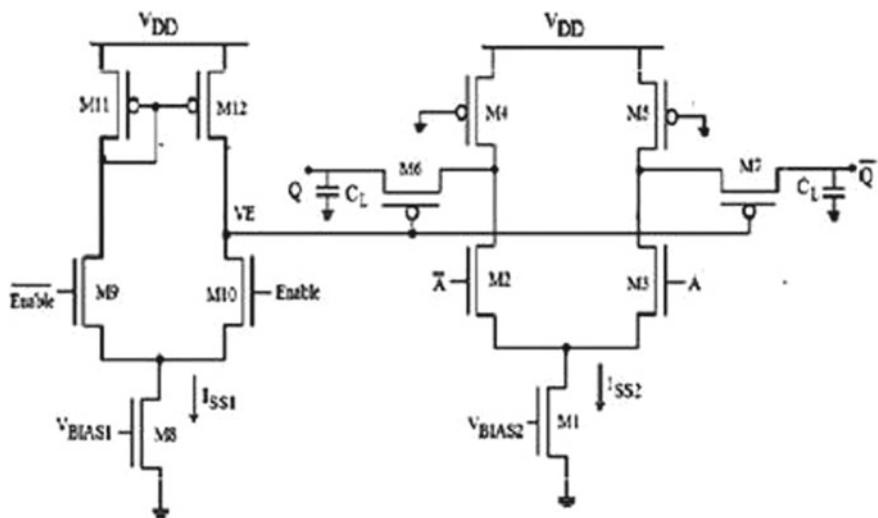


Fig. 4 Switch-based MCML tristate buffer

A low power version of tristate buffer [6] is presented in Fig. 5 which consumes less power as compared to the switch-based tristate buffer by disconnecting the output node from both power supply as well as ground when the buffer is in high impedance state.

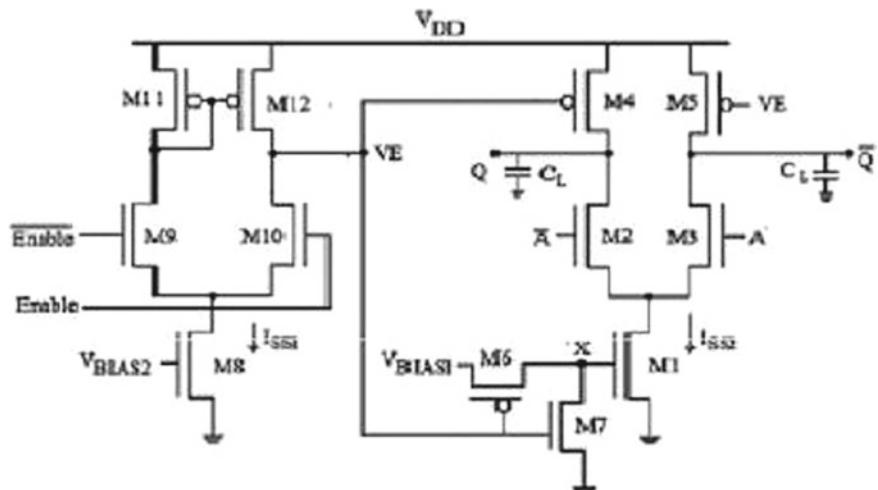


Fig. 5 Low-power MCML tristate buffer

3 Results and Comparison

From Tables 1, 2, and 3, it is observed that the main factors which affect the performance of the gates are their limits on minimum swing at the low current end, but give better propagation delay, good power delay product, and reduced power consumption. With increase in current levels, voltage swing increases in order to attain the gain metric, whereas the propagation delay and the power delay product increase. In order to reduce the capacitance, it is desirable to reduce the width of the transistors but process minimum widths are already achieved.

The performance comparison of the proposed tristate buffers with the existing tristate buffers that are carried out using $0.18 \mu\text{m}$ gpdk CMOS technology parameters. It is noticed that the propagation delay and the power delay product for the switch-based tristate buffer is reduced by fbgv 88.35% and 70.57%, respectively, and that of the low-power tristate buffer reduces by 93.52% and 74.67%, respectively.

Table 1 Simulation results for MCML inverter

I (μA)	ΔV (mV)	V_{DD} (V)	(W_A/L_A)	$(W/L)_{RFP}$	$(W/L)_{RFN}$	T_p (ps)	PDP (fJ)	Static power (pW)
40	300	1.8	0.5/0.18	0.4/0.18	3.0/0.18	63	4.54	19.9
50	400	1.8	0.5/0.18	0.4/0.18	3.0/0.18	69	6.14	20.49
60	500	1.8	0.5/0.18	0.4/0.18	3.0/0.18	71	7.66	21.13

Table 2 Simulation and comparison results for switch-based tristate buffer [8]

Design	I (μA)	ΔV (mV)	V_{DD} (V)	T_p (ps)	Power (μW)	PDP (fJ)
Switch-based tristate buffer [1]	50	400	1.8	481	435	209
Implemented switch-based tristate buffer	50	400	1.8	56	128	7.16

Table 3 Simulation and comparison results for low-power switch-based tristate buffer [8]

Design	I (μA)	ΔV (mV)	V_{DD} (V)	T_p (ps)	Power (μW)	PDP (fJ)
Low-power tristate buffer [1]	50	400	1.8	340	308	104
Implemented low-power tristate buffer	50	400	1.8	22	78	1.71

4 Conclusion

In this paper, a detailed study of MCML operating principle has been presented along with its application to digital circuit design. The paper also addresses the design methodology of high speed, low-power logic gates, and other digital system design blocks by using the contemporary MCML style which has been proved to be better than the conventional CMOS. The static power dissipation of current mode logic remains the main drawback for many applications which stays a future scope of research.

References

1. Radhika, N. Pandey, K. Gupta, M. Gupta, Low power D-latch design using MCML tri-state buffers, *Signal Processing and Integrated Networks (SPIN)*, 2014, pp. 531–534
2. N. Pandey, B. Choudhary, Improved tri-state buffer in MOS current mode logic and its application. *Analog Integr. Circ. Sig. Process* **84**(2), 333–340 (2015)
3. K.P. Sao Pradeep, S. Suresh Kumar, Design and development of high performance MOS current mode logic (MCML) processor for fast and power efficient computing. *Cluster Comput.* (2018)
4. G. Scotti, D. Bellizia, A. Trifiletti, G. Palumbo, Design of low-voltage high-speed CML D-latches in nanometer CMOS technologies. *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.* **25**(12), 3509–3520 (2017)
5. J. Hu, H. Ni, Y. Xia, High speed low power MCML nano meter circuits with near threshold computing. *J. Comput.* **8**(1) (2013)
6. M. Usama, T. Kwasniewski, Design and comparison of CMOS current mode logic latches, in *International symposium on Circuit and Systems, ISCAS'04, 2004*, vol. 4, pp. 353–356
7. S.D. Singh, Analysis and design guidelines for customized logic families in CMOS, in *Proceedings of International Symposium on VLSI Design Test (VDAT)* (2015), pp. 1–2
8. K. Gupta, N. Pandey, M. Gupta, Low-power tri-state buffer in MOS current mode logic. *Int. J. Analog Integr. Circuits Signal Process.* **74**(1) (2013)

Augmented Reality as a Supported Educational Method for Embedded Devices and Technology



Sambit Prasad Kar, Maitrish Ghosh, and Nirmal Kumar Rout

Abstract Technology has always changed and continues to evolve. One of the recent developments in the field of technology is augmented reality (AR). Eighty percent of youth currently own smartphones. Most of them are regular users of smartphones who use these devices to access social media, play games, and communicate with friends and relatives. Meanwhile, far lesser proportions of young adults use telephones to research purposes, do homework, dig knowledge about a subject, etc. The opportunity of combination of smartphones and AR for education is high, though it still needs to be thoroughly explored. AR will provide students with additional knowledge about any issue in a number of ways to make difficult concepts easy to understand. The opportunity to associate reality with digital content has grown slowly, opening up more opportunities for teachers and students. It would quickly impact the traditional learning process in school. AR has the ability to adjust research location and pacing, adding new and innovative forms and methods. Augmented reality technologies will make classrooms more entertaining and more immersive to knowledge. A collective finding of several experiments is that implementations of AR will augment and enhance the learning process, motivation for learning, and performance. The concept is to introduce AR as a mobile app to enable educators, practitioners, and learners to understand critical concept of various embedded hardwares, its main features, and programming. The application generates a virtual label and parts of the product and displays such demonstrations in a step-by-step manner with interactive three-dimensional animations and processes that connect different peripherals with other dedicated devices. This helps the learner to better understand and interact with realistic applications in an immersive way.

S. P. Kar (✉) · M. Ghosh · N. K. Rout

School of Electronics Engineering, Kalinga Institute of Industrial Technology Deemed to be University, Bhubaneswar, Odisha, India

e-mail: spkarfet@kiit.ac.in

M. Ghosh

e-mail: maitrishghosh@gmail.com

N. K. Rout

e-mail: nkrout@kiit.ac.in

Keywords Augmented reality · Embedded hardware · Immersive learning

1 Introduction

Augmented and virtual reality technologies have shown the potential for improving a large variety of fields over the last decade. Virtual reality (VR) technology creates entirely artificial environments through headsets that isolate users from the environment. Increased reality overlays digital interfaces with the physical environment creating both a real and digital world. This combination of physical and virtual information enables AR to improve current simulative functional models further [1, 2]. AR is a real-world, immersive environment, where real-world objects are “increased” by computer-generated perception data, often in several sensory modalities, including visual, auditory, haptic, somatosensory, and olfactory [3]. Although the augmented reality technology and the philosophy are existing from quite some time [4], the recent growths in visual technology and the development of innovative augmented reality applications (ARAs) have attracted user and the tech community. Such technologies are specifically designed combination of software and hardware with AR features in mind and have already been used in many educational settings including natural sciences, chemistry, humanities, and the arts [5]. Augmented reality shows details superimposed in your field of view and can take you into a new world where real and virtual worlds are closely connected. It is not just for the desktop or handheld devices. Augmented reality has been a hot subject for many years in software technology circles, but with the introduction of products like Google Glass, it is receiving renewed interest and popularity. Augmented reality is a technology that uses your device’s camera to work on computer-based vision recognition algorithms to increase sound, video, graphics, and other sensor-based inputs on real-world objects. A simple implementation of the AR implementation is shown in Fig. 1.

AR’s educational use has been examined, and a general finding of many researches is that AR applications can improve the learning process, learning motivation, and efficacy [6]. Despite the promising outcomes, problems remain to be studied such as the teachers’ and students’ production of AR experiences and the formation of interactive content and 3D models [7]. By looking into the potential of AR and rise of small-scale embedded hardware and its application in various fields [8], there is

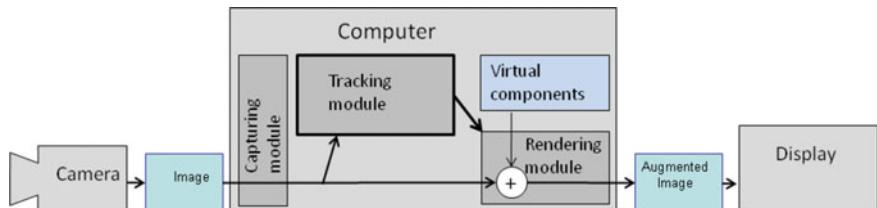


Fig. 1 Block diagram of AR implementation

a need of integration of both to provide an immersive learning experience to the students and the educators.

In this paper, we proposed a method of implementation of AR-based application for educators as well as learners. The proposed application is a novel method of understanding small-scale embedded hardware, its fact sheet, its interfacing ports, and other peripherals. This application also creates a backend connection with a server to fetch the required code for implementing small-scale deployments in the board. The application will act as a virtual educator and assist the learner in learning the fundamentals about the board and its peripherals by self. The paper is organized as follows. Segment 2 details the method and its implementation. The application development part is explained in segment 3. The result analysis and conclusion are depicted in Segment 3 and 4, respectively.

2 Proposed Method and Its Implementation

Augmented reality portrays information in a real-world framework which is right. To do this, the device needs to know where the user is and what the user looks at. Typically, the user experiences the world utilizing a monitor that shows the camera's image along with the details being increased. So the machine is required to calculate the camera's location and orientation in reality. The machine will then render simulated objects in the appropriate location, using a calibrated sensor. The word tracking is a measurement of a camera's relative position (location and orientation) in real time. It is one of the cornerstones of augmented reality [9]. The major types of AR applications available are detailed in Table 1.

Table 1 Types of augmented reality

Category	Type	Characteristics
Trigger-based	Marker-based (paper)	Paper marker activates the process
	Marker-based (object)	Mostly any object activates the process
	Location-based	A GPS signal or particular point on a live map activates the process
	Dynamic augmentation	Meaningful, collaborative augmentation with potential object recognition and/or motion monitoring
	Complex augmentation	Pull internet data and augment based on location marker or object recognition
View-based	Indirect augmentation	Intelligently augmenting the real-world data
	Nonspecific digital augmentation	Augmentation of camera data regardless of location

In our paper, we have used object-based marker as a trigger to augment the embedded hardware and superimposed the information on it and created the application. The process of developing the application is as follows [10, 11]:

- Image acquisition via camera and preprocessing such as denoising, line and edge detection of the marker.
- Affirmative detection of potential markers and discard non-markers.
- Identification and decoding of the valid markers.
- Apply and calculate the orientation of the marker.

3 Application Development

There are several ways to develop an augmented reality application. The approach used in this paper makes use of the unity 3D [12] platform to create the application. The application development flow in unity 3D is as follows.

Modeling of the house: First a three-dimensional (3D) model of the target object (hardware), different pointing objects, and peripheral objects is created using a three-dimensional model creator like blender (Fig. 2).

Importing model to Unity: Once the required 3D model has been generated, then the models are imported to unity for further processing.

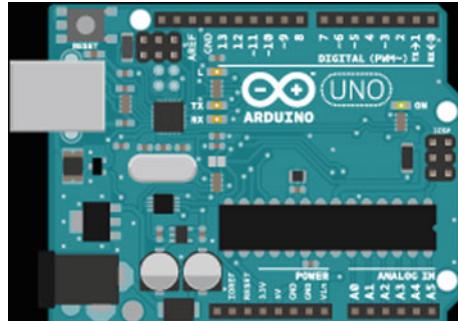
Programming and Scanning: Once the model is imported, then scanning of the 3D object and programming the object for superimposition are done.

Trigger object: The application is tested considering an Arduino Uno [13] embedded board as our marker object. This specific hardware is chosen because of its steep rise in popularity among students, hobbyists, and educators in recent times as a low-cost and flexible microcontroller (Fig. 3).

Fig. 2 3D model created for augmentation



Fig. 3 Arduino Uno embedded microcontroller board (trigger object)



4 Result and Analysis

The application detailed in this paper is based on the unity 3D platform. The blender software used for the production of virtual 3D Objects. Testing during the trial is done on a Xiaomi POCO F1 handset with Vuforia object scanner, and the experimental study is done in a normal room.

Figure 4 represents the welcome screen which got triggered when the object is scanned using the mobile device camera. Once the application detects and recognizes the marker, i.e., the Arduino Uno board. Similarly, Fig. 5 shows the virtual labeling of the board and animated peripherals.

Figures 6 and 7 show a basic implementation of a simple program such as blinking a light-emitting diode (LED) in a stepwise manner with the understanding of the wiring and the necessary code. Once the learner understands the process clearly, then he/she can import the necessary code and mail or deposit the same to his/her email id, else it can also be imported to/from free repositories like GitHub which is shown in Fig. 8.

Fig. 4 View of Introduction page of the application



Fig. 5 Computer-generated 3D labeling of different parts of Arduino using AR

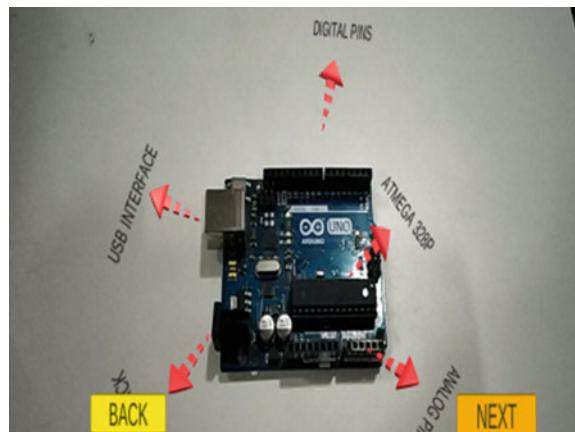


Fig. 6 Basic implementation of tutorial LED blinking using virtual 3D objects such as led and communication cable



Fig. 7 Arduino environment setup instruction using virtual panels

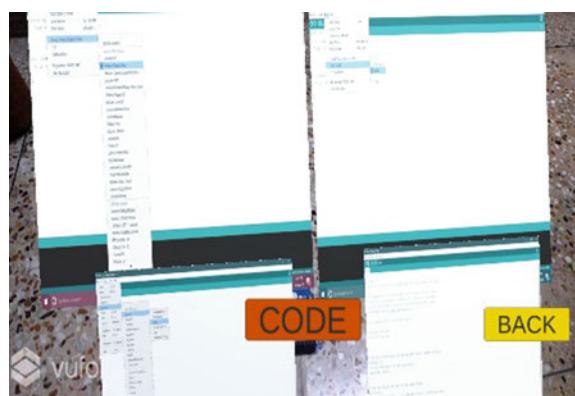


Fig. 8 Fetching of the code GitHub repository and deploying for implementation



```

1 void setup() {
2     // initialize digital pin LED_BUILTIN as an output.
3     pinMode(LED_BUILTIN, OUTPUT);
4 }
5
6 // the loop function runs over and over again forever
7 void loop() {
8     digitalWrite(LED_BUILTIN, HIGH);    // turn the LED on (HIGH is the voltage level)
9     delay(1000);                      // wait for a second
10    digitalWrite(LED_BUILTIN, LOW);   // turn the LED off by making the voltage LOW
11    delay(1000);                      // wait for a second
12 }

```

5 Conclusion

This paper realizes a smartphone-based AR application to create smart data manual for small-scale embedded hardware which assists the developers to understand the device in much lucid and immersive manner. The user interface is made more practical and user-friendly. The motto of our application is to provide an initial assistance to learn about the hardware by self and thus promoting and nurturing self-learning among the learners. There are some relevant products in the market, but they are not created for education or training of these integrated boards. There are also training packages offered to employees in major companies on various devices and appliances where they recommend to use a virtual reality headset such as “Hololens” or “Nreal”. But these are way too costly and difficult to afford. But with our approach, people holding a standard mobile phone with a minimal specification can update it with this AR-based application and use it to learn interactively about various devices and embedded boards. The application developed is in its nascent phase, and in future more augmentation and more detailing of the object are aimed with fluent backend integration. This application can be used to create curiosity among young developers to learn embedded systems without putting much pressure on finding data from data manuals.

References

1. F. Zhou, H.B.L. Duh, M. Billinghurst, Trends in augmented reality tracking, interaction and display: a review of ten years of ISMAR, in *IEEE International Symposium on Mixed and Augmented Reality*, 2008, pp. 15–18
2. Y. Okuda, E.O. Bryson, S. DeMaria Jr. et al., The utility of simulation in medical education: what is the evidence? *Mt. Sinai J. Med.* **76**, 330–343 (2009). <https://doi.org/10.1002/msj.20127>
3. D. Dai, K. Dai, Research on historical and cultural landscape representation based on virtual reality. *J. Hunan Univ. Sci. Technol. (Soc. Sci. Ed.)* **12**, 107–112 (2009)
4. M.L. Heilig, Inventor. Sensorama Simulator (1961)
5. H.-K. Wu, S.W.-Y. Lee, H.-Y. Chang, J.-C. Liang, Current status, opportunities and challenges of augmented reality in education. *Comput. Educ.* **62**, 41–49 (2013). <https://doi.org/10.1016/j.comedu.2012.10.024>

6. J. Bacca, S. Baldiris, R. Fabregat, S. Graf, Kinshuk, Augmented reality trends in education: a systematic review of research and applications. *Educ. Technol. Soc.* **17**, 133–149 (2014)
7. T.C. Huang, C.C. Chen, Y.W. Chou, Animating eco-education: to see, feel, and discover in an augmented reality-based experiential learning environment. *Comput. Educ.* **96**, 7282 (2016)
8. I. Smith, Rise of the (embedded) machines (online article). Available <https://www.abelon.com/news/rise-of-the-embedded-machines.html>
9. A. Edwards-Stewart, T. Hoyt, G. Reger, Classifying different types of augmented reality technology. *Ann. Rev. CyberTherapy Telemed.* **14**, 199–202 (2016)
10. Antonioli, Misty, Corinne Blake, and Kelly Sparks. “Augmented Reality Applications in Education.” *The Journal of Technology Studies* 40, no. 1/2 (2014): 96–107. Accessed August 13, 2020. www.jstor.org/stable/43604312
11. C. Lv, X. Yang, J. Yu, Interaction design in augmented reality on the smartphone, in *2016 9th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)*, Datong, 2016, pp. 1317–1321. <https://doi.org/10.1109/cisp-bmei.2016.7852920>
12. Technologies, U. (n.d.). *Unity*. Unity. <https://unity.com/>
13. *Arduino-Home*. (n.d.). <https://www.arduino.cc/>

Reweighted Zero-Attracting Modified Variable Step-Size Continuous Mixed p -Norm Algorithm for Identification of Sparse System Against Impulsive Noise



Ansuman Patnaik and Sarita Nanda

Abstract The efficiency of traditional adaptive filtering algorithms decreases for sparse system identification when compared to the identification of non-sparse systems. The modified variable step-size continuous mixed p -norm (MVSS-CMPN) algorithm for identifying non-sparse systems was developed in the presence of impulsive noise. In this work, a reweighted zero-attracting MVSS-CMPN algorithm is developed by inducing a sparse penalty function into the MVSS-CMPN algorithm to exploit the sparsity of the system under the effect of impulsive noise. From the simulations, it is found that the presented algorithm achieves a steady state of -25.16 dB for low sparsity in system identification scenario, whereas the MVSS-CMPN algorithm achieves a steady state of -9.96 dB.

Keywords Sparse system identification · Reweighted zero attracting · Continuous mixed p -norm algorithm · Impulsive noise

1 Introduction

A range of practical situations like echo cancellation, underwater acoustics, real-time traffic prediction is affected by the presence of impulsive noise [1]. To counter the consequence of impulsive noise, a sigmoid least mean mixed norm algorithm (SLMMN) [2] and adaptive algorithm based on lower-order norms such as the least mean p -norm (l_p norm) algorithm were introduced [3]. However, the LMP algorithm suffered from stability and steady-state error issues. Motivated from the LMP algorithm, a variable step-size continuous mixed p -norm (VSS-CMPN) algorithm was developed in [4] for robust performance against impulsive noise, thus providing faster convergence. It was also evaluated for the estimation of parameters in power

A. Patnaik · S. Nanda (✉)

School of Electronics Engineering, KIIT Deemed to be University, Bhubaneswar, India
e-mail: snandafet@kiit.ac.in

A. Patnaik

e-mail: ansuman.vssut@gmail.com

signals [5]. However, in the VSS-CMPN, the probability density like function $\lambda(p)$ is considered one, which made the algorithm parameter dependent. In [6], a modified VSS-CMPN (MVSS-CMPN) algorithm was developed in which the probability density like function is updated using a sigmoidal function and thus exhibits faster convergence and low misadjustment.

For different practical applications like acoustic echo cancellers [7], wireless multipath channels [8], the unknown system for identification shows sparsity where the impulse response contains a higher amount of zero coefficients than nonzero coefficients. However, the efficiency of the algorithms mentioned above decreases as they do not consider sparsity while identifying unknown systems. Recently, the sparse LMP algorithms are developed for sparse channel estimation [9]. In [2], zero-attracting (ZA)-SLMMN and reweighted zero-attracting (RZA)-SLMMN algorithm are introduced for the identification of sparse system against impulsive noise. However, when the length of the filter was increased, the performance of both of the algorithms were degraded. This paper presents a reweighted zero-attracting MVSS-CMPN (RZA-MVSS-CMPN) algorithm by inducing a logarithmic sparse penalty into the MVSS-CMPN algorithm to exploit system sparsity in an impulsive noise environment. The organization of the paper is done in the following manner. Section 2 introduces the presented algorithm. Simulation results are discussed in Sect. 3, whereas the conclusion is presented in Sect. 4.

2 RZA-MVSS-CMPN Algorithm

The sparse system identification block diagram is provided in Fig. 1.

Let $x(n)$ be the input to the physical system having impulse response, w with filter coefficients be defined as $w = [w_0, w_1, w_2, \dots, w_{K-1}]$ and K is the filter size. The output of the physical system is corrupted with the system noise $\eta(n)$ uncorrelated to

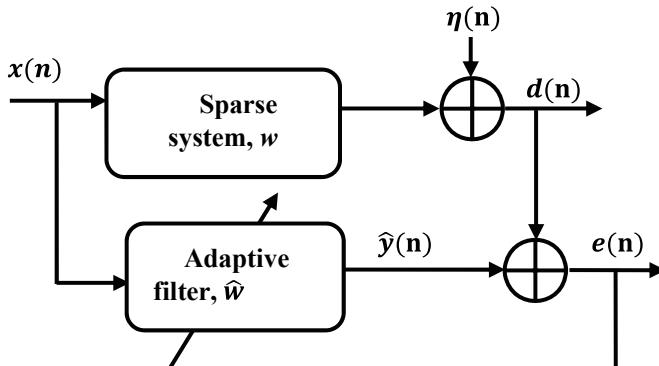


Fig. 1 Sparse system identification block diagram

the input signal. The system noise is a combination of Gaussian noise and impulsive noise. The desired system output $d(n)$ is determined as

$$d(n) = x^T(n)w + \eta(n) \quad (1)$$

where $x(n) = [x(n), x(n-1), x(n-2), \dots, x(n-K+1)]$ represents the signal vector having input signal samples. The physical system w is identified by transmitting the signal $x(n)$ via an adaptive filter \hat{w} . Let the adaptive filter weight coefficients be defined as $\hat{w} = [\hat{w}_0, \hat{w}_1, \hat{w}_2, \dots, \hat{w}_{K-1}]$. The adaptive filter output $\hat{y}(n)$ and the error signal output are determined as

$$\hat{y}(n) = x^T(n)\hat{w}(n) \quad (2)$$

$$e(n) = d(n) - \hat{y}(n) \quad (3)$$

RZA-MVSS-CMPN algorithm's cost function is given by

$$J(n) = \int_1^2 \lambda(n) E\{|e(n)|^p\} dp + \gamma_{RZA} \sum_{i=1}^K \log(1 + \varepsilon_{RZA} |\hat{w}_i(n)|) \quad (4)$$

where γ_{RZA} is the regularization parameter and $\lambda(n)$ is the mixing parameter. The presented algorithm's weight update equation is given by

$$\hat{w}(n+1) = \hat{w}(n) - \mu \nabla_{\hat{w}(n)} J(n) \quad (5)$$

where μ is the step size. Differentiating $J(n)$ with respect to $w(n)$ provides the following equation that represents the presented algorithm's weight update equation

$$\hat{w}(n+1) = \hat{w}(n) + \mu \Upsilon_n \text{sign}(e(n))x(n) - \rho_{RZA} \frac{\text{sgn}(\hat{w}(n))}{1 + \varepsilon_{RZA} |\hat{w}(n)|} \quad (6)$$

where $\rho_{RZA} = \mu \gamma_{RZA} \varepsilon_{RZA}$, $\Upsilon_n = \lambda(n) \int_1^2 p |e(n)|^{p-1} dp$ is the variable step-size dependent on $e(n)$ and is stated as [4]

$$\Upsilon_n = \lambda(n) \frac{\ln(|e(n)|)(2|e(n)| - 1) - |e(n)| + 1}{\ln^2(|e(n)|)} \quad (7)$$

where $\lambda(n)$ is updated using a sigmoid function and is given by [6]

$$\lambda(n) = 1/(1 + \exp(-Q(n))) \quad (8)$$

whereas the variable $Q(n)$ is updated as [6]

$$Q(n+1) = Q(n) + \mu e(n)y(n)\lambda(n) \quad (9)$$

Established on (6), the RZA-VSS-CMPN algorithm's weight update equation is generalized as

$$\hat{w}(n+1) = \hat{w}(n) + \text{Adaptive error update} + \text{Sparse penalty} \quad (10)$$

The RZA-VSS-CMPN algorithm takes into account a logarithm penalty, which is identical to the l_0 norm, the exact measure of sparsity.

3 Simulation Results

The presented sparse algorithm's performance is analyzed for system identification under an impulsive noise environment. The unknown system to be identified and the length of the filter is set as $K = 100$. Different levels of sparsity, i.e., $l = \{1, 4, 8\}$, are considered to evaluate the presented algorithm's performance from high sparsity to low sparsity. By filtering a unit variance, $\sigma_x^2 = 1$, Gaussian white noise through an autoregressive system with 0.8 poles, the correlated (colored) input signal is created. The system noise $\eta(n)$ is a mixture of white Gaussian noise signal with a signal-to-noise ratio (SNR) of 20 dB and Bernoulli–Gaussian (BG) distributed impulsive noise produced as $v(n) = I_m(n)\eta_i(n)$, where $I_m(n)$ is the Bernoulli process and $\eta_i(n)$ is the Gaussian process. $\eta_i(n)$ is a Gaussian random process that has mean zero and variance $\sigma_a^2 = 10^4/12$, and $I_m(n)$ is described with probabilities $P(I_m(n) = 1) = P_i$ and $P(I_m(n) = 0) = 1 - P_i$ with $P_i = 0.01$ [2]. The normalized mean square deviation is the performance criteria used for the presented algorithm's performance evaluation and is given by

$$\text{NMSD}(n) = 10 \log_{10} \frac{\|w - \hat{w}(n)\|_2^2}{\|w\|_2^2} \quad (11)$$

The NMSD is calculated for $n = 15,000$ iterations, and 100 independent trials are averaged for evaluation of results. The presented algorithm's performance is evaluated in comparison with the sparse-based LMP algorithms and SLMMN algorithms for different levels of sparsity. The simulation parameters used in the presented algorithm are $\rho_{\text{RZA}} = 5e-4$, $\varepsilon_{\text{RZA}} = 20$, and $\mu = 0.05$. Figure 2 shows the plot of the sparse impulse response for different values of $l = 1, 4$, i.e., from high sparsity to low sparsity. The simulation results regarding the presented algorithm's performance for different sparsity levels, i.e., $l = 1, l = 4$ and $l = 8$ are provided in Figs. 3 and 4

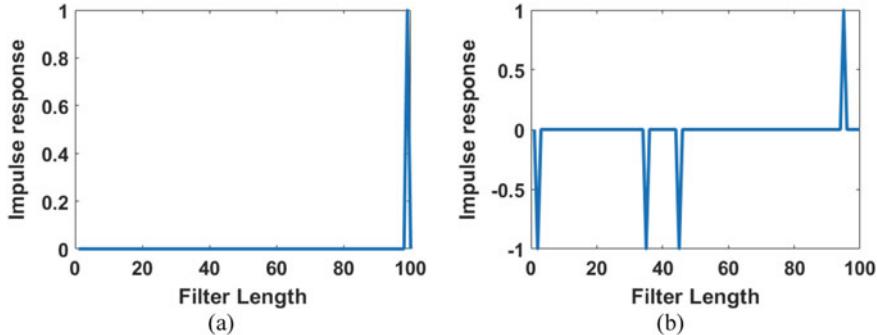


Fig. 2 Impulse response of a system with sparsity **a** $l = 1$ and **b** $l = 4$

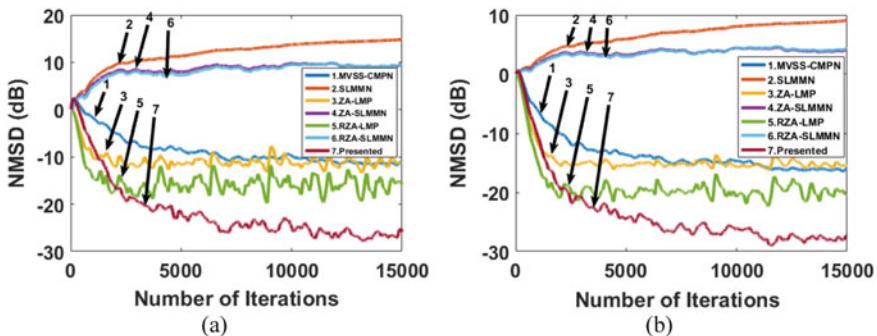


Fig. 3 NMSD comparison of the presented algorithm for a system with sparsity **a** $l = 1$ and **b** $l = 4$ when the system is corrupted with impulsive noise

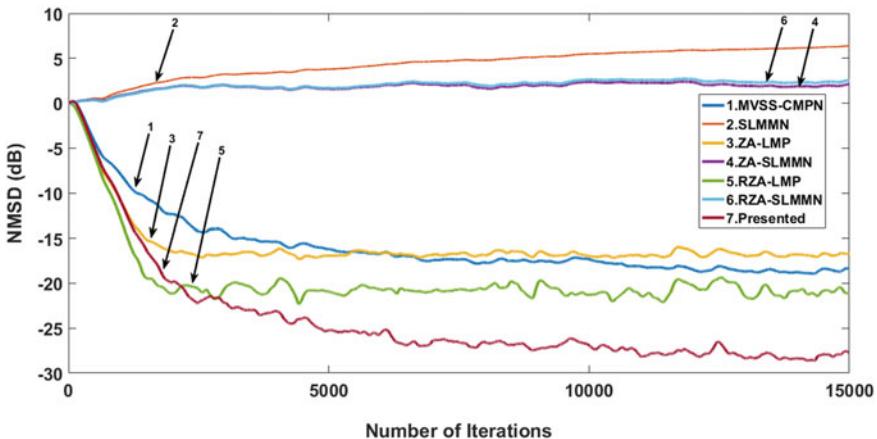


Fig. 4 NMSD comparison of the presented algorithm for a system with sparsity $l = 8$, when the system is corrupted with impulsive noise

Table 1 NMSD values obtained in dB for various algorithms

Algorithms	Sparsity level ($l = 1$) NMSD(dB)	Sparsity level ($l = 4$) NMSD(dB)	Sparsity level ($l = 8$) NMSD(dB)
MVSS-CMPN	-9.96	-14.02	-20.29
SLMMN	15	9.63	6.57
ZA-LMP	-13.14	-14.30	-17.97
ZA-SLMMN	9.69	3.98	2.30
RZA-LMP	-14.37	-17.95	-24.24
RZA-SLMMN	9.95	4.93	3.61
RZA-VSS-CMPN	-25.16	-21.33	-31.48

in the presence of impulsive noise, whereas the steady-state values for the algorithms compared with different sparsity levels are provided in Table 1.

3.1 RZA-MVSS-CMPN Algorithm's Performance for a System with Sparsity $l = 1, 4$

As inferred from Fig. 3, the presented algorithm's performance for a system with sparsity (a) $l = 1$ and (b) $l = 4$ is superior in comparison with the other existing algorithms against impulsive noise in terms of steady-state error. However, the RZA-LMP algorithm exhibits a rapid convergence rate. The MVSS-CMPN algorithm fails to provide lower steady-state error, whereas the SLMMN algorithm and the sparse-based SLMMN algorithms such as ZA-SLMMN and RZA-SLMMN do not converge for a filter of order 100.

3.2 The Presented Algorithm's Performance for a System with Sparsity $l = 8$

Figure 4 demonstrates the presented algorithm's performance for the identification of a sparse system with sparsity $l = 8$. It is observed that the suggested algorithm yields less steady-state error when compared with the other existing sparse-based algorithms such as the RZA-LMP, ZA-SLMMN, and RZA-SLMMN algorithms. However, the RZA-LMP algorithm exhibits a rapid convergence rate. On the other side, the MVSS-CMPN algorithms achieve higher steady-state error for non-sparse system identification.

From Table 1, it is observed that the steady-state values obtained by the presented algorithm is less compared to the other algorithms for the system with different levels of sparsity in the presence of impulsive noise.

4 Conclusion

The MVSS-CMPN algorithm exhibits good performance in the presence of impulsive noise, but it does not promote sparsity. Therefore, we presented a sparse-based MVSS-CMPN algorithm for sparse system identification by inducing a logarithmic sparse penalty into the MVSS-CMPN algorithm to exploit system sparsity. The RZA-MVSS-CMPN algorithm's performance was tested from high sparsity to low sparsity by considering different levels of sparsity. It is shown that the presented algorithm achieves a significant lower steady-state error of -25.16 , -21.33 , -31.48 dBs while identifying sparse systems with different levels (1,4,8) of sparsity in the presence of impulsive noise when compared with other competing algorithms.

References

1. Z. Sheng, J. Zhang, H. Han, Robust shrinkage normalized sign algorithm in an impulsive noise environment. *IEEE Trans. Circ. Syst.: II Expr. Briefs* **64**(1), 91–95 (2016)
2. P. Rakesh, T.K. Kumar, F. Albu, Modified least-mean mixed-norm algorithms for adaptive sparse system identification under impulsive noise environment, in *2019 42nd International Conference on Telecommunications and Signal Processing (TSP)* (IEEE, 2019)
3. M. Shao, C.L. Nikias, Signal processing with fractional lower-order moments: stable processes and their applications. *Proc. IEEE* **81**(7), 986–1010 (1993)
4. H. Zayyani, Continuous mixed p-norm adaptive algorithm for system identification. *IEEE Signal Process. Lett.* **21**(9), 1108–1110 (2014)
5. S. Nanda, T. Chakravorty, P.K. Dash, A new Taylor-LMS adaptive filter for parameter estimation of power signals including distributed generation systems. *Aust. J. Electr. Electron. Eng.* **13**(3), 174–194 (2016)
6. A. Patnaik, S. Nanda, A modified variable step-size continuous mixed p-norm algorithm for system identification, in *2018 International Conference on Applied Electromagnetics, Signal Processing and Communication (AESPC)*, vol. 1 (IEEE, 2018)
7. J. Benesty et al., *Advances in Network and Acoustic Echo Cancellation* (2001)
8. F. Adachi, E. Kudoh, New direction of broadband wireless technology. *Wirel. Commun. Mob. Comput.* **7**(8), 969–983 (2007)
9. W. Ma et al., Sparse least mean p-power algorithms for channel estimation in the presence of impulsive noise. *Sig. Image Video Process.* **10**(3), 503–510 (2016)

A Semantic-Based Information Retrieval System



Alka Ranjan and Soumya Priyadarsini Panda

Abstract Information retrieval (IR) is the process of retrieving relevant documents from a large collection of documents. The applications of these technologies are no longer restricted to only online search engines, instead these are widely being used by organizations to facilitate different organizational information searches. Efficient retrieval of documents from a large collection of organizational data is still a challenging task. In this paper, a semantic-based information retrieval system is presented for institutional student project report retrieval. The model uses a clustering-based technique for fast document retrieval and a semantic-based query processing technique for retrieving the most relevant documents with respect to the user query. A number of tests were conducted to evaluate the performance of the presented model on different random user queries, and the precision and recall measures were determined. The performance of the model is also compared with existing retrieval techniques, and the results obtained show the efficiency of the model in providing relevant documents quickly. The model achieves approximately 91% precision and 90% recall accuracy in the considered domains and data set.

Keywords Information retrieval · K -means clustering · Semantic retrieval · TF-IDF · Domain classification

1 Introduction

Information retrieval (IR) focuses on obtaining required documents from document repositories as per its relevance to the user's query [1]. There is a huge amount of information available over the Web in electronic form, and also, there is a rapid increase in the information sources day by day [2]. Sometimes, it became a key issue

A. Ranjan (✉) · S. P. Panda

Department of CSE, Silicon Institute of Technology, Bhubaneswar, Odisha, India

e-mail: alkarnjn@gmail.com

S. P. Panda

e-mail: soumya.panda@silicon.ac.in

to obtain the most relevant documents that contains the exact answers for the user's query [3]. At many instances, the users were unable to get the required documents at one go and have to reformulate the query to find the suitable documents. Information retrieval is a rapidly growing field and is still continuing as an active area of research as much of the information over the Web or in organizations is expressed in textual form [4]. It has become a key issue to obtain the most relevant data from the collection of available documents. In order to provide the desired level of search accuracy and appropriateness in terms of relevance to user queries, several IR systems have been putted forward since decades [5]. However, understanding the exact intent of the user in the context by analyzing the syntactic and semantic association of words in the entered query and using those information in retrieval process is still a challenging task.

Semantic retrieval of information involves retrieval of documents using the semantic aspect of language processing to ensure more relevant information retrieval [6, 7]. Its applications include retrieval of documents from medical, commercial, educational, or agriculture document repositories [8–10]. As compared to the online retrieval process, retrieval of data from organizational repository receives much attention these days. For various University and Institutes, retrieval of old project reports plays a key role in helping the next batch of students to work on similar areas. Also, from the institute point of view, the idea of student project report retrieval system may be extended as a recommendation system to suggest unique project titles for each new batch of students based on the existing project details and student interests.

Considering the above-discussed issues, this work focuses on development of a semantic-based information retrieval system to retrieve student project reports at institute level. The presented model uses a clustering-based technique for fast document retrieval. It applies a query expansion technique to address the semantic aspect of retrieval for extracting more relevant information. The model also ranks the retrieve documents based on the calculated relevance score. To analyze the performance of the presented model in terms of accuracy, the precision and recall measures were used on a number of user queries covering varied domains. The performance of the model is also compared with the traditional Boolean model (BM) and the fuzzy clustering-based semantic retrieval (FCSR) model [11]. As compared to the existing models, the presented model achieves approximately 91% precision and 90% recall accuracy in the considered domains.

2 Related Work

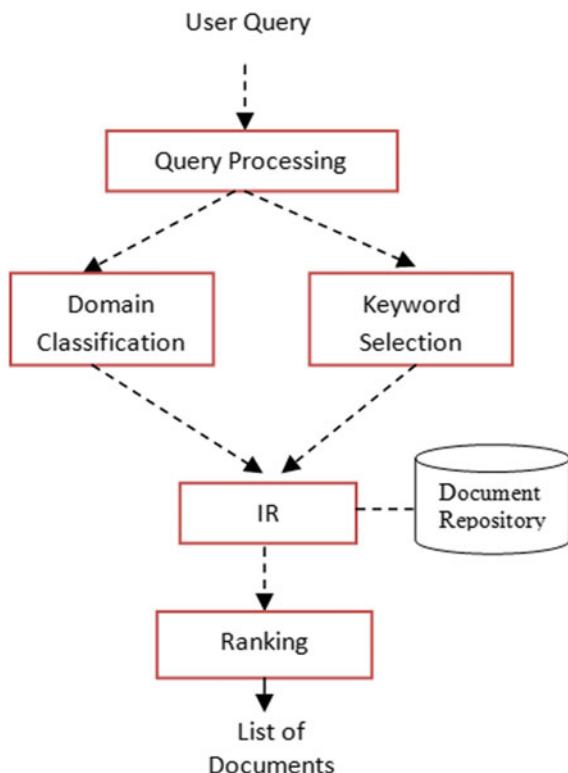
There are a number of researches documented in the field of IR systems; however, semantic-based information retrieval and ranking has no significant research contributions till date on offline repositories. The researchers in [6] presents an intelligent cluster-based model for achieving the IR needs. The retrieval mechanism is also extended to online social networks [7] for enhancing the social information retrieval process [8]. The authors in [9] discusses about a domain-specific IR systems for

retrieving documents. A few IR models focus on identifying the data requirements of people based on their age groups [10]. The authors in [11] discusses about a fuzzy clustering-based semantic IR system to retrieve relevant research articles from a collection of documents using a fuzzy membership value assigned to each document with respect to the match score of keywords in the query. However, semantic-based document retrieval on offline student project report data has no significant research evidence in literature and is the main focus of the presented work.

3 Proposed Model for IR

The architecture of the proposed semantic-based Information retrieval (SIR) system is shown in Fig. 1. The input to the model is any query entered by the user for getting project reports of their interest. The model after processing the query provides a ranked list of project reports related to the query. It uses a document repository which is a collection of project reports. A clustering technique is applied over the reports to create domain-specific clusters which are used for fast information retrieval. The

Fig. 1 Architecture of the model



details of the repository creation and clustering process are discussed below followed by the details of the phases involved in the retrieval process.

3.1 Document Repository Creation and Clustering

For designing the SIR system, a set of 200 project reports were collected and stored in the form of a repository. Each of the project reports were categorized to some pre-defined domain groups by applying clustering technique. For this purpose, documents were represented using vector space model. Each document d is represented as a vector $\{w_1, w_2 \dots w_n\}$, where w_i denotes the associated weight of any term t_i . The term weight value in a document is a measure of the term's significance, and is computed using the term frequency with inverse document frequency (*TF-IDF*). The *TF-IDF* is used to rank the importance of a term in its contextual document corpus. For obtaining the similarity between any documents di and dj , the cosine correlation measure is computed by dividing the dot product of the two vectors by the product of their magnitudes. Using those values as features, the k-means clustering algorithm [6] is applied to create clusters indicating the project domains. The algorithms starts with initial estimates for the K centroids, which can either be randomly generated or randomly selected from the data set. K -means assign data points to the cluster center on the basis of similarity or dissimilarity. Data points, whose features are similar, are aggregated in same cluster as presented below:

Let $X = \{x_1, x_2, x_3 \dots, x_n\}$ be the set of data points and $V = \{v_1, v_2, \dots, v_c\}$ be the set of centers.

1. Choose the number of clusters k
2. Randomly select K cluster center from data point that is centroid
3. Calculate cosine similarity between each document vector.
4. Assign each data point to its closest centroid.
5. Recalculate the centroid of newly formed cluster.
6. Repeat steps 4 and 5 until the centroid position do not change or no data points was reassigned.

3.2 Query Processing

The front-end query processing unit processes the query for removing the stop words and finding the index terms which are used in the retrieval process. The stemming and lemmatization process are applied on the query to obtain the root words to be used in the retrieval process.

3.3 Domain Classification

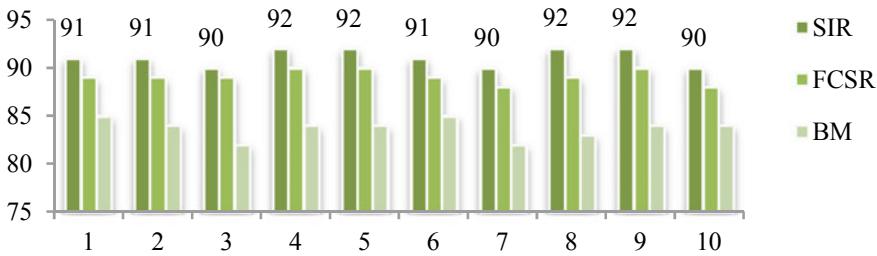
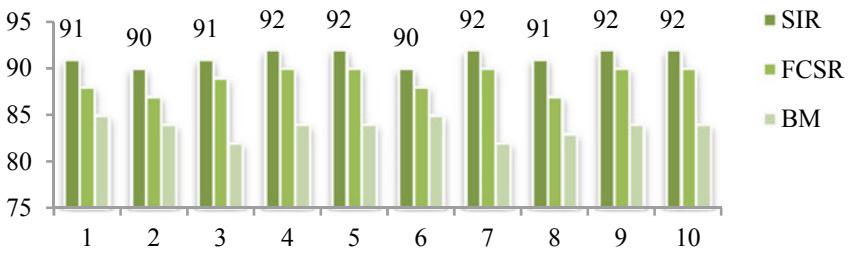
The domain classification phase takes the index terms and classifies the query to its respective domain for faster retrieval of the documents from a large document repository. For this purpose, a data set of 200 queries and their associated domain classes are prepared and are used as training data to implement the classification algorithm for query domain classification. For this purpose, the Naive Bayes classifier is used, and associated class labels are obtained for each of the entered query. This domain label is used for retrieving the documents from the repository which belongs to the same query domain the user is interested in. Therefore, it ensures more meaningful and appropriate contextual retrieval of documents. For this purpose, term-document matrix for each class is generated and used to classify based on the probability values.

3.4 Keyword Selection

The obtained index terms after stop word removal from query processing phase are further enhanced in this phase using the synonyms of the terms to address the semantic aspect and to collect all possible related meanings for document retrieval. For addressing the word relations, WordNet is used which includes all possible senses of a word under various contexts. Therefore, the index term sets obtained from query processing phase are further enhanced by a set of new synonym words to retrieve all possible related documents. The idea of keyword selection phase is to enhance the keyword set to address the semantic aspect of document retrieval. As, in natural languages, the same concept can be expressed using variety of words with similar meaning, including synonyms for the index terms increases the possibility of all possible matches in the document corpus as compared to the traditional retrieval techniques that only matches the keywords in the user query.

3.5 Information Retrieval and Ranking

Based on the classified query domain and the obtained enhanced keyword set, documents are retrieved from the respective domain cluster. The project reports were retrieved from only the cluster whose label matches with the query classified domain label for fast retrieval. Those documents are then ranked based on the number of keyword match scores.

**Fig. 2** Accuracy (Precision %)**Fig. 3** Accuracy (Recall %)

4 Results and Performance Analysis

For implementation of the model, the Python environment and libraries are used. For query pre-processing part, NLTK is used. The domain classification and clustering methods used were evaluated on their accuracy on a number of random queries and project titles, respectively, and the models are able to successfully identify the respective domains in the data set considered. To analyze the performance of the SIR system, the Boolean model (BM) and fuzzy-cluster-based semantic retrieval (FCSR) [11] models are used. The models were tested on a number of random queries, and the performance of the models was measured. For this purpose, the precision and recall measures are used [11]. Due to the use of domain classification technique at query processing part, the model is able to find the related documents belonging to that domain ensuring a high accuracy ratio. Figures 2 and 3 show the accuracy percentage on 10 random queries. The model achieves approximately 91% precision and 90% recall accuracy in the considered domains and data set.

5 Conclusions

In this work, a semantic-based information retrieval system is presented for institutional student project report retrieval. The model uses a clustering-based technique for

fast document retrieval. The presented model is compared with the existing models. The model achieves approximately 91% precision and 90% recall accuracy in the considered domains and data set. The model can be enhanced by considering the word-to-word semantic relationships in query processing to achieve better results.

References

1. C.S. Kumar, R. Santhosh, Effective information retrieval and feature minimization technique for semantic web data. *Comput. Electr. Eng.* **81** (2019) (Article 106518)
2. I. Bounhas, N. Soudani, Y. Slimani, Building a morpho-semantic knowledge graph for Arabic information retrieval. *Inf. Process. Manage.* (2019) (Article 102124)
3. R. Bansal, S. Chawla, Design and development of semantic web-based system for computer science domain-specific information retrieval. *Perspect. Sci.* **8**, 330–333 (2016)
4. M. Pai, M. Chen, H. Chu, Y. Chen, Development of a semantic-based content mapping mechanism for information retrieval. *Expert Syst. Appl.* **40**(7), 2447–2461 (2013)
5. L. Sbattella, R. Tedesco, A novel semantic information retrieval system based on a three-level domain model. *J. Syst. Softw.* **86**(5), 1426–1452 (2013)
6. Y. Djenouri, A. Belhadi, P. Fournier-, J. Lin, Fast and effective cluster-based information retrieval using frequent closed itemsets. *Inf. Sci.* **453**, 154–167 (2018)
7. M. Bouadjenek, H. Hacid, M. Bouzeghoub, Social networks and information retrieval, how are they converging? A survey, a taxonomy and an analysis of social information retrieval approaches and platforms. *Inf. Syst.* **56**, 1–18 (2016)
8. M. Asim, M. Wasim, M. Ghani Khan, N. Mahmood, W. Mahmood, The use of ontology in retrieval: a study on textual, multilingual, and multimedia retrieval. *IEEE Access* **7**, 21662–21686 (2019)
9. T. Gossen, A. Nürnberg, Specifics of information retrieval for young users: a survey. *Inf. Process. Manage.* **49**(4), 739–756 (2013)
10. B. Xu et al., Improve biomedical information retrieval using modified learning to rank methods. *IEEE/ACM Trans. Comput. Biol. Bioinf.* **15**(6), 1797–1809 (2018)
11. D. Mahapatra, C. Maharana, S.P. Panda, J.P. Mohanty, A. Talib, A. Mangaraj, A fuzzy-cluster based semantic information retrieval system, in *Proceeding of the Fourth International Conference on Computing Methodologies and Communication (ICCMC 2020)*, pp. 678–681 (2020)

Multiband Ultrathin Terahertz Metamaterial Absorber for Sensing Application



Shruti and Sasmita Pahadsingh

Abstract A multiband ultrathin metamaterial absorber, namely extended cross pattern with different dimensions and geometrical configurations, has been proposed. The absorption rates of extended cross-shaped absorber have been studied in THz range. The perfect absorption rate of extended cross-shaped absorber was found to be 97.25%, 99.27%, 99.68% and 98.76% at 2.128 THz, 6.088 THz, 9.17 THz and 10.8 THz, respectively. This is suitable for multiband operations, finding its application in sensing. The proposed ultrathin absorber has a thickness of 2.8 μm which is 1/51-th times of the wavelength of the lowest frequency. The impedance of the proposed absorber is calculated and matched to free space impedance in these absorption regions based on impedance matching theory. The proposed absorber has promising applications in sensing, imaging and detection.

Keywords Metamaterial absorber · Dielectric constant · Resonator

1 Introduction

Absorbers are defined as the filters that neither reflect nor transmit the electromagnetic waves in a particular frequency bands. The first Metamaterial Absorber (MA) came into picture by the experimental work of Landy et.al in 2008 [1]. Metamaterials are artificial composite structures that exhibit properties like negative permittivity, permeability and refractive index in certain frequency bands. The negative refractive index of metamaterials has the characteristics of high speed detection and therefore used in biological sensing. The working principle of MA is as follows: as the electromagnetic waves strike the absorber, the wave attenuates and loses energy. The energy loss is due to the conversion from electromagnetic interference (EMI) energy

Shruti (✉) · S. Pahadsingh

School of Electronics engineering, KIIT University, Bhubaneswar 751024, India
e-mail: shruti.fet@kiit.ac.in

S. Pahadsingh

e-mail: spahadsinghfet@kiit.ac.in

to heat energy. The amount of attenuation depends on frequency, permeability and permittivity of the material. The main advantages of metamaterial are compactness, ease of fabrication, lightweight, polarisation independence at broad frequency band and the ability to provide perfect absorption in a wide range of frequencies.

Metamaterials consist of a top layer and a ground metal plane separated by a dielectric substrate which is used to absorb the maximum energy, and a perfect conductor is used as the ground plane to block the transmission. The top layer of metamaterial is designed by periodic unit cells or array whose periodicity is much smaller than the operating wavelengths, in which thickness, size and operational bandwidths are important parameters. So compacting the unit cell and designing multiband metamaterials are the main challenging tasks for researchers nowadays [2]. In recent years, MA has attracted attention in terahertz band as it is very useful in sensing, detecting, bolometer, imaging, etc.

A triple broadband polarisation sensitive MA for terahertz application is designed in [3] consisting of metallic layers separated by GaAs which is ultrathin (0.028λ). The absorption peaks are at 1.71 THz, 3.16 THz and 4.89 THz with the absorption rate of 99.4%, 99.6% and 90.5%, respectively. A six band terahertz absorber is also designed in [4] based on three concentric symmetrical circular split rings and a symmetrical square split ring, and all the absorption peaks are close to 99%. The six absorption peaks were at 0.3 THz, 0.6 THz, 0.816 THz, 1.56 THz, 2.259 THz and 2.52 THz, respectively. A broad dual band switchable graphene-based terahertz MA is designed in [5] with the absorption rate of 80% in the frequency range of 0.473–1.407 THz and 2.273–3.112 THz, respectively.

In this paper, a multiband and ultrathin MA with different dimensions and geometrical configurations has been proposed to study the absorption rate in THz region. The behaviour of electric and magnetic field of the proposed absorber is also studied to gain in-depth knowledge of the absorption peaks. The electric field and magnetic field are produced due to the excitation of the top resonator and current loop between the top resonator and ground plane, respectively.

2 Design of MMA

The unit cell based on extended cross pattern resonator is designed in this paper with periodicity $p = 26 \mu\text{m}$ as shown in Fig. 1a. The proposed absorber consists of three layers in which the ground plate is of thickness $0.3\mu\text{m}$ with conductivity of $4.09 \times 10^7 \text{ S/m}$, substrate used is SiO_2 of thickness $2.8\mu\text{m}$, and dielectric constant of 3.9 and metallic gold resonator of extended cross of thickness $0.1\mu\text{m}$ is designed on the top of the substrate. The thickness of proposed ultrathin MMA is $\lambda/51$ at 2.128 THz (lowest frequency) for multiband operations. The electric field is produced due to the excitation of the top resonator, and magnetic field is produced due to the current loop between the top resonator and ground plane.

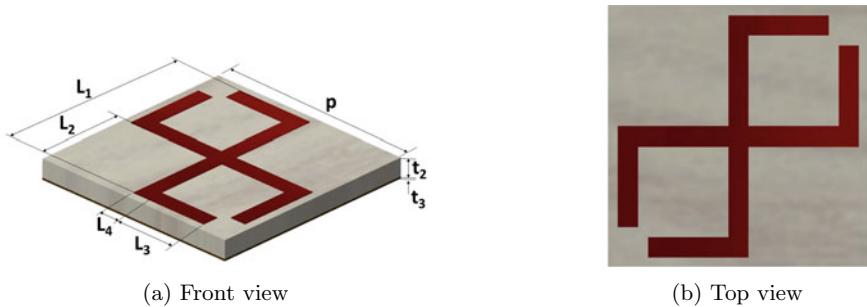


Fig. 1 Schematic view of proposed extended cross pattern ($p = 26 \mu\text{m}$, $t_2 = 2.8 \mu\text{m}$, $t_3 = 0.3 \mu\text{m}$, $L_1 = 24 \mu\text{m}$, $L_2 = 9 \mu\text{m}$, $L_3 = 8 \mu\text{m}$, $L_4 = 2 \mu\text{m}$)

3 Results and Discussion of Simulation Work

The proposed absorber is simulated using ANSYS Electronics Desktop by applying Floquet port and periodic boundary conditions. The absorption rate is calculated by the relation, $A = 1 - R - T$ where T is transmission coefficient and R is reflection coefficient. A metallic ground plane has been used to block the transmitted wave, and thus, the transmission is negligible. The absorption is calculated by using $A = 1 - R$ where R is given by $|S_{11}|^2$. Extended cross pattern exhibits maximum absorption in four frequency bands. The observed absorption peaks are at 2.128 THz, 6.088 THz, 9.17 THz and 10.8 THz with absorption of 97.25%, 99.27%, 99.68% and 98.76%, respectively, as shown in Fig. 2.

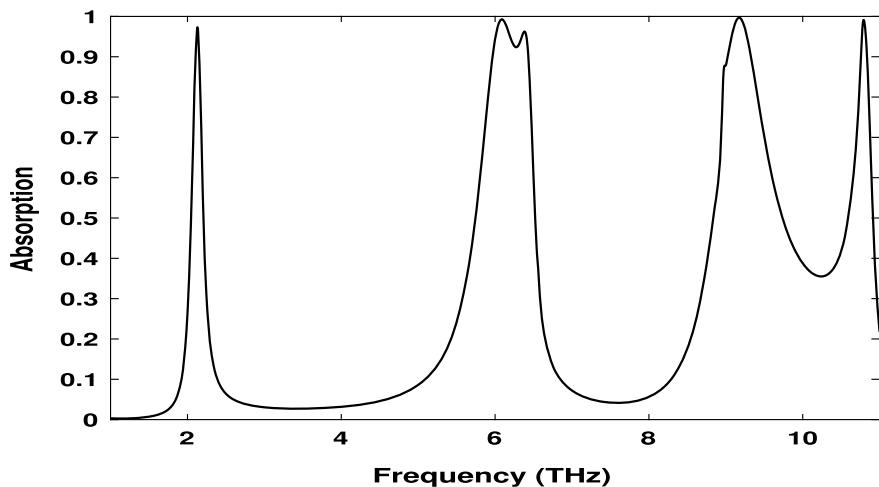


Fig. 2 Simulated absorption spectra of extended cross pattern

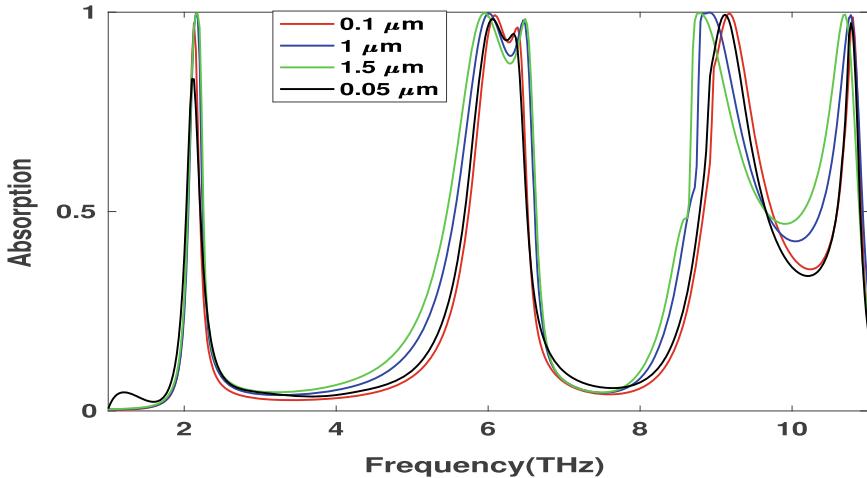


Fig. 3 Absorption spectra with variation in thickness of the gold pattern

The thickness of top gold metallic cross extended pattern has also been varied as shown in Fig. 3. It can be seen that by choosing the thickness of gold pattern greater than $0.1 \mu\text{m}$, a broad band spectra has been obtained, but this is very insignificant. Since the main concern of the current study is to design ultrathin MA and that puts a constraint in selecting the thickness and was thus chosen as $0.1 \mu\text{m}$. When the thickness is reduced to $0.05 \mu\text{m}$, then the results obtained are not very encouraging as at lower frequency values the absorption is observed to be quite low.

Further to study the effect of the thickness of substrate on the absorption characteristics, it was also varied, and the absorption plots were obtained as shown in Fig. 4. Increasing the thickness leads to many random peaks which is not desirable as can be observed for the thickness of $4.0 \mu\text{m}$ where at 8.093 THz there is sharp fall in absorption coefficient. So for the frequency range of $7.63\text{--}9.045 \text{ THz}$, the absorption will be very unpredictable and thus making it unsuitable for application in this range. Further increase in the thickness does not contribute significantly to the desired absorption curve and is against the requirement of an ultrathin MA and hence is not considered for the current study.

Multiband absorption can be explained by impedance matching theory, where the impedance of the structure is made equal to the impedance of free space ($=377\Omega$) for perfect absorption. By impedance matching, maximum power transfer will take place, and there will be no reflection. The real part of ' z ' should be close to one, and imaginary part should be close to '0'. The normalised impedance is calculated by using Eq. 1:

$$Z = \sqrt{\frac{(1 + S_{11})^2 - S_{21}^2}{(1 - S_{11})^2 - S_{21}^2}} = \frac{1 + S_{11}}{1 - S_{11}} \quad (\text{as } S_{21} = 0) \quad (1)$$

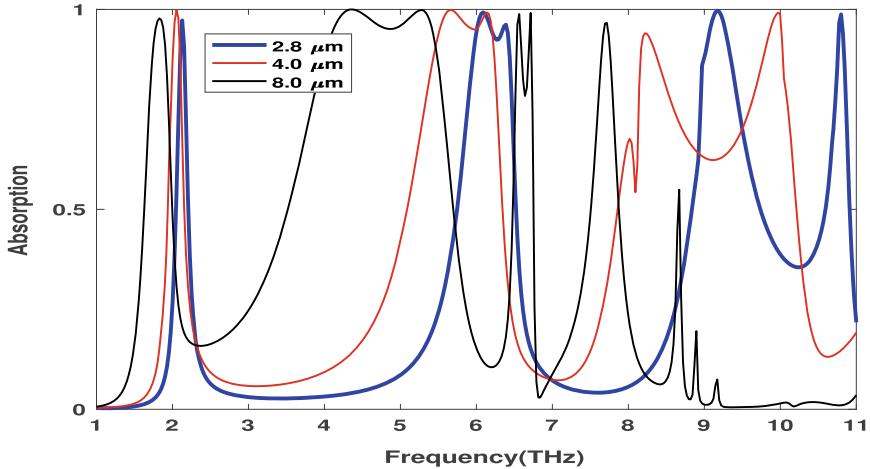


Fig. 4 Absorption spectra with variation in thickness of the substrate pattern

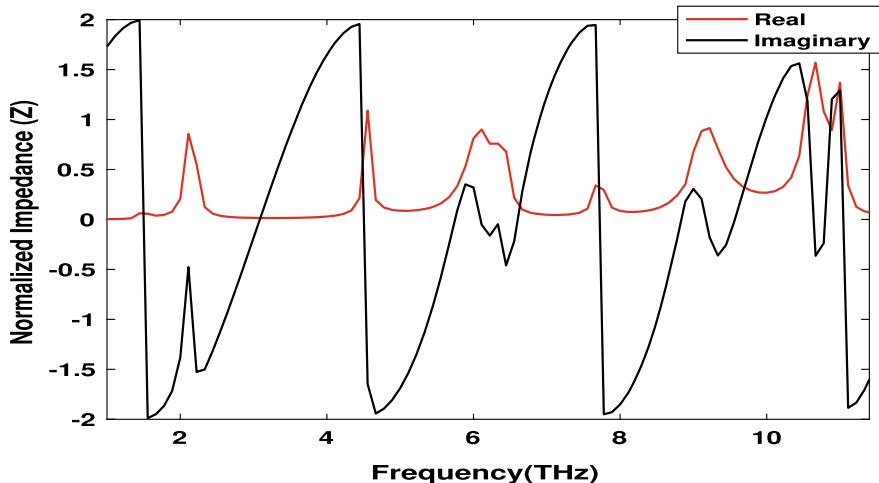


Fig. 5 Normalised impedance of extended cross pattern

The obtained normalised impedance of the proposed extended cross pattern is $0.86-j0.48$, $0.9-j0.05$, $0.92-j0.18$ and $1.08-j0.23$ at 2.128 THz, 6.088 THz, 9.17 THz and 10.8 THz, respectively, as shown in Fig. 5. Therefore, we have received maximum absorption rate at four frequencies.

The behaviour of electric and magnetic field of the proposed extended cross pattern has been studied and shown in Figs. 6 and 7 at 2.128 THz, 6.088 THz, 9.17 THz and 10.8 THz. The first resonance frequency, i.e. at 2.128 THz, the electric field mainly concentrated on the vertical two extended arms of resonator as shown in Fig. 6a. At the second resonance frequency, i.e. at 6.088 THz, the electric field concentrated on

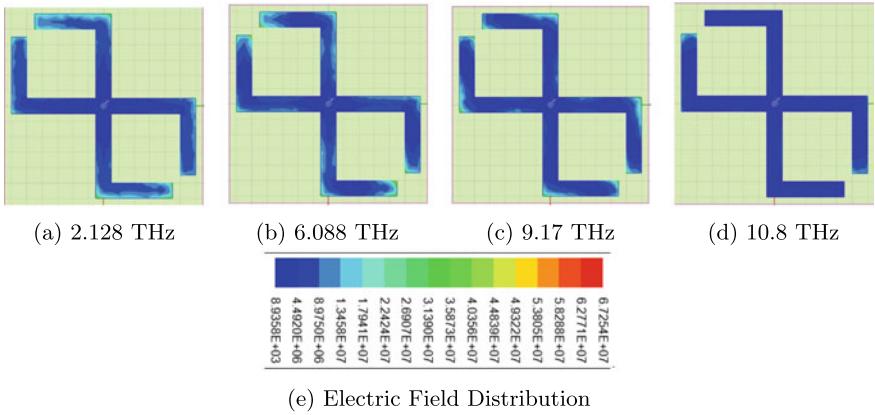


Fig. 6 E-field distribution on top of extended cross resonator

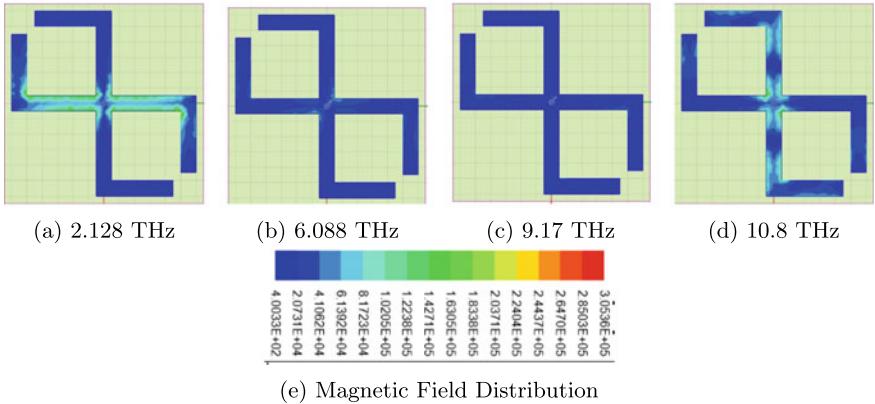


Fig. 7 H-field distribution on top of extended cross resonator

the corner of vertical two extended arms and also little bit on the corner of horizontal two extended arms of the resonator as shown in Figure 6b. At the third resonance frequency, i.e. at 9.17 THz, the electric field is weak but mainly concentrated on the corner of cross resonator as shown in Fig. 6c. At the fourth resonance frequency, i.e. at 10.8 THz, the electric field is very weak as compared to other resonance frequencies as shown in Fig. 6d.

At the first and fourth resonance frequency, i.e. at 2.128 THz and 10.8 THz, respectively, the magnetic field mainly concentrated on the horizontal and vertical arms of cross resonator as shown in Fig. 7a, d, respectively. The second and third resonance frequency, i.e. at 6.088 THz and 9.17 THz, show a weak magnetic field as can be observed from Fig. 7b, c which is consistent with the result shown in Fig. 2. It can be observed from Fig. 6 that the proposed structure is the main contributor resulting in high electric and magnetic field at 2.128 THz and 10.8 THz, respectively.

Table 1 Comparison of present work with previous work

Reference	Pattern	Thickness	Frequency (THz)	Absorption
Limei Qi et.al [2019]	Cross-shaped grooves	95 μm	2.273-3.112	>80% [5]
M. Tianhua et.al [2018]	Circular rings	500 μm	0.9, 1.83, 2.55, 2.92 and 3.71	99.83%, 98.92%, 99.96%, 99.35% and 99% [6]
Z. Haijun et.al [2019]	Cross cave patch	8 μm	0.4-2.2	98% to 99.9%[8]
J. Prince et.al [2018]	Dual ring	4 μm	2.25, 4.18 and 5.2	87% and 89% and 92%[7]
Proposed Work	Extended cross	2.8μm	2.128, 6.088, 9.17 and 10.8	97.25%, 99.27%, 99.68% and 98.76%

However, it is necessary to compare the results with the other recent previous work (given in Table 1). The proposed design shows the properties of multiband structure having absorption peak at around 2.128 THz with absorption rate of 97.25% which is larger than those shown in [5]. In [6], the absorption peak reaches to 99% with a compromise in thickness. In [7], the absorption peak is around 92% at 5.2 THz, but in the proposed design, the peak is around 98.76% at 10.8 THz.

The proposed MA has wide applications such as planner waveguide sensor and planner metamaterial array. Planner waveguide sensor is used to detect harmful gases and chemical analytes. So, the evanescent waves should penetrate to much deeper depths, and this can be increased by symmetry configuration, and since this proposed absorber is symmetrical, so this is appropriate for the application. This proposed absorber is also suitable for planner metamaterial array since sharp resonance is required to achieve high sensitivity in its frequency response.

4 Conclusion

A multiband ultrathin ($\sim \lambda/51$ -th) terahertz metamaterial extended cross-shaped absorber in a unit cell has been designed. The four absorption peaks were found at 2.128 THz, 6.088 THz, 9.17 THz and 10.8 THz with absorption rates of 97.25%, 99.27%, 99.68% and 98.76%, respectively. The results have been compared with some of the designs in the literature, and it is observed that the proposed absorber performs better in terms of thickness and frequencies. The multiband absorption mechanism was also explained by impedance matching theory. To gain an in-depth knowledge of coupling mechanism of the proposed structure, the electric and magnetic field distribution has been analysed at those frequencies. The proposed absorber finds its application in sensing, imaging and detection.

References

1. N.I. Landy, S. Sajuyigbe, J.J. Mock, D.R. Smith, W.J. Padilla, Perfect metamaterial absorber. *Phys. Rev. Lett.* **100**(20), 207402 (2008)
2. M.D. Banadaki, A.A. Heidari, M. Nakhkash, A metamaterial absorber with a new compact unit cell. *IEEE Antennas Wireless Propag. Lett.* **17**(2), 205–208 (2017)
3. P. Jain, A. K. Singh, S. Bansal, K. Prakash, L. Hadda, N. Gupta, A. K. Singh, S. Kumar, N. Sardana, Triple band polarisation sensitive metamaterial absorber for terahertz applications, in *2020 7th International Conference on Signal Processing and Integrated Networks (SPIN)* (IEEE, 2020), pp. 559–562
4. A. Shoieb, J. Nourinia, C. Ghobadi, M. Karamirad, B. Mohammadi, A six-band terahertz absorber based on symmetrical split rings, in *2019 5th Conference on Knowledge Based Engineering and Innovation (KBEI)*, IEEE, pp. 419–423
5. L. Qi, C. Liu, S.M.A. Shah, A broad dual-band switchable graphene-based terahertz metamaterial absorber. *Carbon* **153**, 179–188 (2019)
6. T. Meng, D. Hu, Q. Zhu, Design of a five-band terahertz perfect metamaterial absorber using two resonators. *Optics Commun.* **415**, 151–155 (2018)
7. P. Jain, S. Garg, A. K. Singh, S. Bansal, K. Prakash, N. Gupta, A. K. Singh, N. Sharma, S. Kumar, N. Sardana, Dual band graphene based metamaterial absorber for terahertz applications, in *2018 IEEE 13th Nanotechnology Materials and Devices Conference (NMDC)* (IEEE, 2018), pp. 1–4
8. H. Zou, Y. Cheng, Design of a six-band terahertz metamaterial absorber for temperature sensing application. *Optical Mater.* **88**, 674–679 (2019)

Frequency Estimation Using Adaptive Algorithm



Sananda Kumar

Abstract Frequency estimation is a common problem that is required for a variety of application to provide monitoring, protection and control. Adaptive algorithms generally estimate the unknown system coefficients assuming the frequency is constant. But in real-life scenario, the frequency may deviate from the original value, which will pose problem to estimation techniques. This deviation of frequency from its true value is referred as frequency mismatch (FM). This paper deals with the implementation of discrete Fourier transform (DFT) method in congruence with adaptive least mean square algorithm to estimate the frequency in real-time scenario. The proposed method is applied to estimate signal frequency in presence of Gaussian noise and harmonics and is found to perform well in terms of MSE.

Keywords DFT · LMS · Frequency mismatch

1 Introduction

Frequency estimation application ranges to a variety of applications that include speech and audio signal analysis, radar signal detection, source localization, power delivery and many more. For power system cases, the accuracy of frequency measurement is very important for stability and normal operation of electrical devices. Frequency estimation in case of noise-free signal can be determined using Prony's method [6]. In power system cases, the voltages are considered as pure sinusoidal signal, where the time between two consecutive zero crossings indicates the frequency of the signal [1]. However, the practical signals are always distorted and noisy in nature, contains harmonics up to several orders and may deviate from the true frequency. So, the basic techniques of frequency measurement fail to work. To address these practical signal cases, there has been a lot of development in the field of

S. Kumar (✉)

KIIT Deemed to be University, Odisha, India
e-mail: sanandanit@gmail.com

frequency estimation using notch filter [4, 5], phase-locked loop [2, 3], least square methods and their variants [1, 8, 9].

Of the many techniques, the discrete Fourier transform (DFT) has been fundamental in frequency estimation techniques of power systems and applied to several applications with variations such as modification with interpolation [10], windowing and hybrid techniques [7], etc. However, the interpolation technique requires large number of samples, and the performance becomes off-line.

In this paper, the frequency estimation is proposed using the DFT technique in combination with least mean square method (LMS) that can be implemented in a real-time case with reduced number of samples. The periodic signal that may deviate from its true frequency is communicated after the calculating DFT of the signal, which is treated as training data by the adaptive system for frequency estimation. The performance of the proposed method is demonstrated using simulation of synthetic data in presence of Gaussian noise and harmonics.

2 Mathematical Modeling

A discrete time periodic signal is represented as

$$x(i) = A \cos(2\pi f_0 i \Delta T + \phi) \quad (1)$$

with A , ϕ , f_0 , i , ΔT as amplitude, initial phase angle, true signal frequency, sampled index, sampling rate given by $\frac{1}{f_s}$, respectively, in a noise- and harmonic-free environment. The signal can be expressed in polar form as given below

$$x(i) = \frac{A}{2} [e^{j(2\pi f_0 i \Delta T + \phi)} + e^{-j(2\pi f_0 i \Delta T + \phi)}]. \quad (2)$$

Considering $\frac{A}{2} e^{j\phi}$ as v and similarly $\frac{A}{2} e^{-j\phi}$ as v^* , the equation is rewritten as

$$x(i) = V e^{j(2\pi f_0 i \Delta T)} + V^* e^{-j(2\pi f_0 i \Delta T)}. \quad (3)$$

For any discrete time finite signal, the discrete Fourier transform is calculated as

$$X_{DFT} = \sum_{i=0}^{N-1} x(i) \cdot e^{-j \frac{2\pi i}{N}} \quad (4)$$

for n -th subsequent DFT calculation of N points, it is represented with X_n

$$X_n = \sum_{i=0}^{N-1} x(i+n) \cdot e^{-j \frac{2\pi i}{N}} \quad (5)$$

with $x(i+n)$ defined as $[Ve^{j2\pi f_0(i+n)\Delta T} + V^*e^{-j2\pi f_0(i+n)\Delta T}]$. Putting the above values in terms of Y , the formulation of DFT equation is given as [9]

$$\begin{aligned} Y_n &= \sum_{i=0}^{N-1} (Ve^{j2\pi(f_0+\Delta f)(i+n)\Delta T} + V^*e^{-j2\pi(f_0+\Delta f)(i+n)\Delta T}) \cdot e^{-j\frac{2\pi i}{N}} \\ &= Ve^{\frac{j2\pi(f_0+\Delta f)n}{Nf_0}} \sum_{i=0}^{N-1} e^{\frac{j2\pi\Delta f \cdot i}{Nf_0}} + V^*e^{\frac{-j2\pi(f_0+\Delta f)n}{Nf_0}} \sum_{i=0}^{N-1} e^{\frac{-j2\pi(2f_0+\Delta f) \cdot i}{Nf_0}}. \end{aligned} \quad (6)$$

Denoting the first part of equation (6) as P_n and similarly the second part as Q_n , we have

$$Y_n = P_n + Q_n. \quad (7)$$

Considering $r = e^{\frac{j2\pi(f_0+\Delta f)}{Nf_0}}$ in the equation (6), a relation of n -th first N point DFT can be established with its previous $\{n-1\}^{th}$ N point DFT

$$P_n = r P_{n-1} \text{ and } Q_n = r^{-1} Q_{n-1} \quad (8)$$

The relationship between three consecutive DFT values now can be described using equation below

$$\begin{aligned} Y_n &= P_n + Q_n = r P_{n-1} + r^{-1} Q_{n-1} \\ Y_{n-1} &= P_{n-1} + Q_{n-1} \\ Y_{n-2} &= P_{n-2} + Q_{n-2} = r^{-1} P_{n-1} + r Q_{n-1} \end{aligned} \quad (9)$$

which is further simplified and obtained as

$$\begin{aligned} Y_n + Y_{n-2} &= (r P_{n-1} + r^{-1} Q_{n-1}) + (r^{-1} P_{n-1} + r Q_{n-1}) \\ &= (r + r^{-1}) P_{n-1} + (r + r^{-1}) Q_{n-1} \\ &= (r + r^{-1}) Y_{n-1} \\ &= w Y_{n-1} \end{aligned} \quad (10)$$

The required parameter w is equated as

$$w = \frac{Y_n + Y_{n-2}}{Y_{n-1}} = (r + r^{-1}) = e^{\frac{j2\pi(f_0+\Delta f)}{Nf_0}} + e^{\frac{-j2\pi(f_0+\Delta f)}{Nf_0}} = 2 \cos\left(\frac{2\pi(f_0 + \Delta f)}{Nf_0}\right). \quad (11)$$

The final frequency deviation of Δf can be calculated using equation (11) and is given as

$$(f_0 + \Delta f) = \frac{f_s}{2\pi} \cos^{-1}(w/2) \quad (12)$$

where $f_s = Nf_0$.

Now consider that the deviated signal is being transmitted with its N point DFT as shown in Fig. 1, as initial training phase which will be corrupted by noise while transmission. So, the equation (11) can be written with optimal weight parameter in a linear form given by

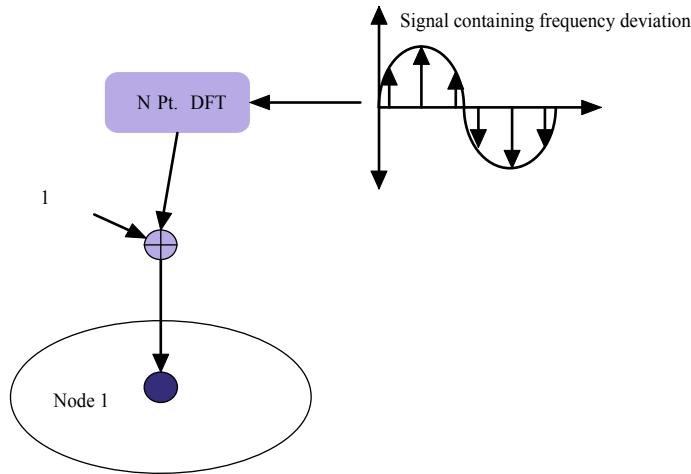


Fig. 1 DFT data signal transmission model

$$(Y_n + Y_{n-2}) = Y_{n-1}w^0 + \eta, \quad (13)$$

where η is the noise part after evaluation of DFT. The above condition holds for the magnitude of the DFT noise η to be very less in magnitude, or else this poses a drifting problem. For least square optimization, the cost function is given by

$$\begin{aligned} J(w) &= E[e^2(n)] = E[e(n).e^*(n)] \\ &= E[(Y_n + Y_{n-2}) - Y_{n-1}w].\{(Y_n + Y_{n-2}) - Y_{n-1}w\}^* \end{aligned} \quad (14)$$

with the error is defined as $e(n) = [(Y_n + Y_{n-2}) - Y_{n-1}w]$ at n -th sample index. The standard least mean square (LMS) algorithm is described by

$$w(n+1) = w(n) - \mu (\nabla_w J(w)) \quad (15)$$

with w as the parameter to be evaluated adaptively in real time. The parameter μ is the step size. The gradient of the cost function is defined as

$$\begin{aligned} \nabla_w J(w) &= -2Y_{n-1}^* (Y_n + Y_{n-2}) + 2\|Y_{n-1}\|^2 w \\ &= -2Y_{n-1}^* \{(Y_n + Y_{n-2}) - Y_{n-1}w\} \end{aligned} \quad (16)$$

So, the final weight update equation is given by

$$w(n+1) = w(n) + \mu Y_{n-1}^* e(n) \quad (17)$$

The w parameter updated using above LMS equation can be used in equation (12), to calculate the final frequency at n -th sample index.

3 Stability in Mean

The equation (17) is subtracted from the optimal weight w^0

$$w^0 - w(n+1) = w^0 - w(n) - \mu Y_{n-1}^* e(n) \quad (18)$$

with the weight deviation $w^0 - w(n+1)$ defined as $\tilde{w}(n+1)$ and using equation (13),

$$\begin{aligned} \tilde{w}(n+1) &= \tilde{w}(n) - \mu Y_{n-1}^* \{Y_{n-1} w^0 + \eta - Y_{n-1} w\} \\ &= \tilde{w}(n) - \mu \|Y_{n-1}\|^2 \tilde{w}(n) - \mu Y_{n-1}^* \eta \end{aligned} \quad (19)$$

Applying the expectation operator to the above equation

$$E[\tilde{w}(n+1)] = (1 - \mu \|Y_{n-1}\|^2) E[\tilde{w}(n)] - \mu E[Y_{n-1}^* \eta] \quad (20)$$

Considering the presence of Gaussian noise only, the last part $E[Y_{n-1}^* \eta]$ can be separated out and written as $E[Y_{n-1}^*] E[\eta]$ with $E[\eta] = 0$. Therefore we get

$$E[\tilde{w}(n+1)] = (1 - \mu \|Y_{n-1}\|^2) E[\tilde{w}(n)] \quad (21)$$

For stability purpose, the following condition must satisfy

$$|(1 - \mu \|Y_{n-1}\|^2)| < 1 \quad (22)$$

which gives the range of step-size μ as

$$\mu < \frac{2}{\|Y_{n-1}\|^2} \quad (23)$$

4 Simulation Result

The simulation is performed in presence of Gaussian noise with 0.01 variance after Fourier transform, and mean square error (MSE) is obtained after implementing 30 independent experiments. The step size is considered to be 0.01, with sampling value $f_s = N \times f_0$ as $N = 60$, and the true frequency is selected to be 50 Hz with maximum deviation from true frequency which is ± 2 . In addition to Gaussian noise, third, fifth and seventh harmonics to the fundamental true frequency are added to the signal with magnitude of 0.01, 0.02 and 0.05, respectively.

Figure 2 describes the input frequency (given in green color) to vary from true frequency with maximum of ± 2 , and the other one (blue line) is the estimated frequency using the algorithm. Similarly, Fig. 3 gives the result of mean square error (MSE) in dB for the first 3000 sample data of Fig. 2, and it shows the steady-state

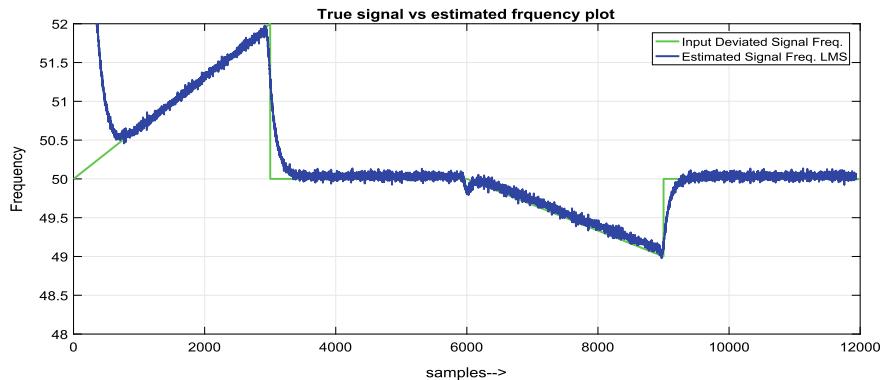


Fig. 2 Deviated frequency vs the estimated frequency

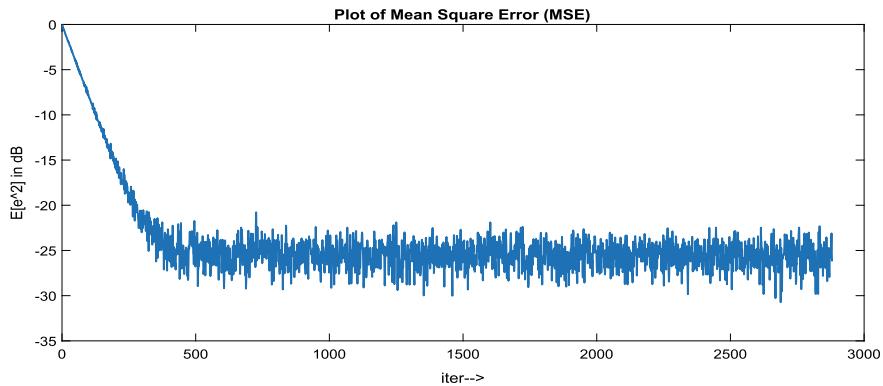


Fig. 3 Plot of mean square error (MSE) in dB for the first 3000 data samples averaged over 30 independent experiments

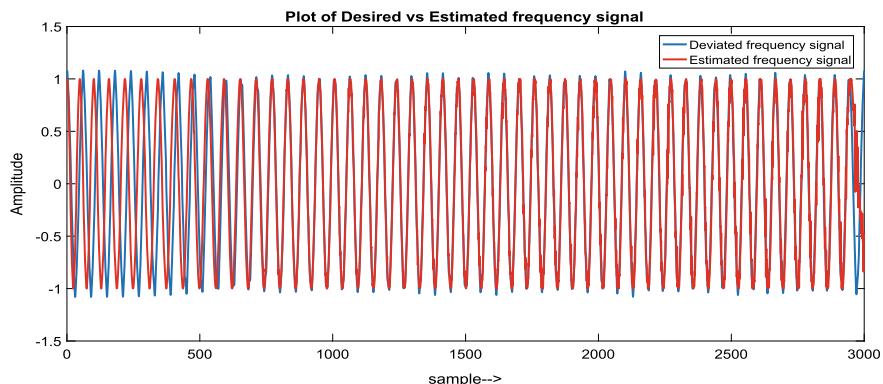


Fig. 4 Plot of frequency deviated cosine wave and plot of estimated cosine wave

performance of the algorithm at -25 dB. Figure 4 depicts the actual cosine signal (blue color) for the first 3000 sample data of Fig. 2, and the other (red) is the estimated cosine signal after implementation of the adaptive algorithm.

5 Conclusion

The proposed algorithm shows a good performance in terms of MSE of -25 dB for signal containing harmonics and in presence of Gaussian noise in the communication environment. The steady state of MSE is attained as much faster stating the algorithm can be implemented in a real-time condition.

References

1. M.M. Begovic, P.M. Djuric, S. Dunlap, A.G. Phadke, Frequency tracking in power networks in the presence of harmonics. *IEEE Trans. Power Deliv.* **8**(2), 480–486 (1993)
2. H. Karimi, M. Karimi-Ghartemani, M.R. Iravani, Estimation of frequency and its rate of change for applications in power systems. *IEEE Trans. Power Deliv.* **19**(2), 472–480 (2004)
3. M. Karimi-Ghartemani, B.T. Ooi, A. Bakhshai, Application of enhanced phase-locked loop system to the computation of synchrophasors. *IEEE Trans. Power Deliv.* **26**(1), 22–32 (2010)
4. C. Li, H. Wang, Distributed frequency estimation over sensor network. *IEEE Sens. J.* **15**(7), 3973–3983 (2015)
5. M. Mojiri, M. Karimi-Ghartemani, A. Bakhshai, Estimation of power system frequency using an adaptive notch filter. *IEEE Trans. Instrumentation Meas.* **56**(6), 2470–2477 (2007)
6. R. Prony, Essai experimental. *J. de l'Ecole Polytechnique* **2** (1795)
7. J. Ren, M. Kezunovic, A hybrid method for power system frequency estimation. *IEEE Trans. Power Deliv.* **27**(3), 1252–1259 (2012)
8. M. Sachdev, M. Giray, A least error squares technique for determining power system frequency. *IEEE Trans. Power Apparatus Syst.* **2**, 437–444 (1985)
9. Y. Xia, Y. He, K. Wang, W. Pei, Z. Blazic, D.P. Mandic, A complex least squares enhanced smart dft technique for power system frequency estimation. *IEEE Trans. Power Deliv.* **32**(3), 1270–1278 (2015)
10. F. Zhang, Z. Geng, W. Yuan, The algorithm of interpolating windowed fft for harmonic analysis of electric power system. *IEEE Trans. Power Deliv.* **16**(2), 160–164 (2001)

An Improved Ridge Regression-Based Extreme Learning Machine for the Prediction of Diabetes



Priya Das and Sarita Nanda

Abstract Diabetes mellitus is a very serious human health problem. Every year the total number of cases is increasing rapidly. The advancement in the machine learning technologies can help in early and accurate detection of the disease. Therefore, an efficient and very fast diabetes prediction model is proposed in this paper using ridge regression extreme learning machine classifier and firefly optimization algorithm for the optimization of the weight vectors. The PIMA Indian Diabetic Database is used for the training and testing of the model. The maximum achieved accuracy, sensitivity and specificity are 93.4%, 97.5% and 85.72%, respectively. The results of the model are compared with two popular methods, support vector machine (SVM) and extreme learning machine (ELM), and it shows that the proposed method outperforms SVM and ELM.

Keywords Diabetes mellitus · Fire fly algorithm · Extreme Learning Machine (ELM) · Ridge Regression ELM

1 Introduction

Diabetes mellitus is a condition in which the production of the required amount of insulin to regulate the blood glucose level in the body decreases [1]. It increases the risk of kidney disease, heart disease, blindness, etc. As there is no cure of it, it is necessary to detect at the early stage to control it. In the last two decades, machine learning applications in medical field are increasing rapidly. Researchers are applying different algorithms to get accurate and most efficient model.

P. Das · S. Nanda (✉)
Kalinga Institute of Industrial Technology (DU), Bhubaneswar, India
e-mail: snandafet@kiit.ac.in

P. Das
e-mail: priya.dasfet@kiit.ac.in

Diabetes classification and diagnosis is a challenging task because of the nonlinear, correlation structured, complex medical data. It also has missing values or outliers. Different methods are being proposed for the classification of diabetes. Hemant et al. [2] use different classifiers (K-means clustering, Naive Bayes, fuzzy ‘C’ mean) for the prediction of diabetes. The maximum accuracy achieved is 76.54%. Genetic algorithms [3] improve the true positive rate. Daho et al. [4] use particle swarm optimization. It shows substantial improvement in accuracy. An hybrid system [5] for diabetes prediction is also proposed by using the machine learning techniques along with deep learning techniques. RDNN, PNN, etc. are also analyzed for this purpose. Artificial neural network [6], tree partitioning adaptive support vector machine [7], decision tree [8] are also proved to be good classifiers for diabetes prediction.

Although many researchers have done good work for the diabetes classification, the maximum accuracy is around 89% till now (as per our literature survey) [9]. The accuracy can be improved further. Extreme learning machine [10] is a very fast classifier. Therefore, a fast and efficient model is proposed in this paper using ridge regression extreme learning machine and firefly optimization technique. Ridge regression ELM has better stability and generalization over ELM. This paper is organized as follows: Section 2 explains the methodologies used. Section 3 elaborates the experimental results, and finally, Sect. 4 presents the conclusions.

2 Extreme Learning Machine

Extreme learning machine (ELM) [11] is a single hidden layer feed-forward network (SLFN). In ELM, X is a $K \times L$ input vector, where K is the total input instances and L is the feature number. i -th input vector X^i can be written as $X^i = [x^{i_1} \ x^{i_2} \ \dots \ x^{i_L}]^T$, and output vector is represented as Y with dimension $K \times 1$. The output matrix of hidden layers is $H^i = [h^{i_1} \ h^{i_2} \ \dots \ h^{i_N}]^T$. β is the hidden layer output weight matrix. The calculated output is $\hat{Y} = H\beta$.

where

$$\beta = H^\perp Y \quad (1)$$

where H^\perp is the *Moore–Penrose* generalized inverse of (H) and it is a non-singular matrix. Therefore, Eq. (1) can be represented as

$$\beta = (H^T H)^{-1} H^T Y \quad (2)$$

2.1 Ridge ELM

ELM is a very fast method, but to reduce the multicollinearity problem and to improve the stability and generalization performance, ridge regression ELM [12] is intro-

duced. In RELM, a constant value is introduced in β .

$$\beta = \left(H^T H + \frac{I}{C} \right)^{-1} H^T Y \quad (3)$$

where I is the identity matrix and C is a regularization parameter.

2.2 Proposed Ridge ELM with Fire Fly Algorithm

The proposed ELM structure is shown in Fig. 1. For both ELM and RELM, the hidden layer output matrix H is calculated using randomly chosen input weight vector. This input weight vector can be optimized using firefly algorithm [12] to achieve the best performance. This algorithm imitates the behavior of the fireflies. This is based upon how fireflies move toward lights. The objective function is represented as

$$f(\theta) = \min \left(\frac{\sum error^2}{K} \right) = \min \left(\frac{\sum (Y - \hat{Y})^2}{K} \right) \quad (4)$$

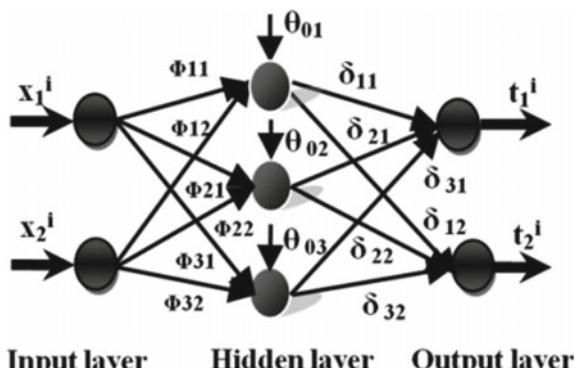
where Y is the actual output and \hat{Y} is the predicted output. The attractiveness α between any two fireflies is given by Eq. (5)

$$\alpha = \alpha_0 e^{-\gamma d_{ij}^2} \quad (5)$$

where α_0 is the initial attractiveness and d_{ij} represents the separation between i -th and j -th fireflies. Here, γ is the absorption coefficient. α can be modified as

$$\alpha = (\alpha_{max} - \alpha_{min})e^{-\gamma d_{ij}^2} + \alpha_{min} \quad (6)$$

Fig. 1 Proposed architecture of ELM [12]



where

$$d_{ij} = \sqrt{\sum_{l=1}^{d_{\text{dime}}} (\theta_{il} - \theta_{jl})^2} \quad (7)$$

θ is the input weight matrix. Initially, θ is random which is updated using Eq. (8).

$$\theta_{i,l} = \theta_{i,l} + \alpha * (\theta_{j,l} - \theta_{i,l}) + \lambda * (\text{rand}_{i,l} - 0.5) \quad (8)$$

λ introduces randomness and its range is between 0 and 1. Using the global optima, Eq. (8) becomes

$$\theta_{i,l} = \theta_{i,l} + \alpha * (\theta_{j,l} - \theta_{i,l}) + \alpha_{\text{best}} * (\theta_{g\text{best},l} - \theta_{i,l}) + \lambda * (\text{rand}_{i,l} - 0.5) \quad (9)$$

where

$$\alpha_{\text{best}} = (\alpha_{\max} - \alpha_{\min}) e^{-\gamma(d_{\text{igbest}})^2} + \alpha_{\min} \quad (10)$$

and

$$d_{\text{igbest}} = \sqrt{\sum_{l=1}^{d_{\text{dime}}} (\theta_{i,l} - \theta_{g\text{best},l})^2} \quad (11)$$

3 Experimental Results

The proposed method uses RELM with firefly optimization algorithm. PIMA dataset [13] is used to validate the model. The dataset consists of 768 (500 non-diabetic and 268 diabetic) instances with 8 features. The dataset is divided into 85% training and 15% testing datasets. The performance of the proposed method is evaluated in terms of accuracy, sensitivity, specificity, MCC, precision, recall and F1 score.

RELM has a tunable parameter C . By taking different values of C , the performance of the model is observed. It has been seen that the best result is obtained when the value of C is 2^{14} . The same dataset is evaluated using ELM and SVM. Table 1 shows the

Table 1 Performance comparison of implemented methods

Method	Accuracy (in %)	Sensitivity (in %)	Specificity (in %)	MCC	Precision	Recall	F1 score
RELM with FF	93.44	97.5	85.71	0.86	0.93	0.98	0.95
ELM	80.65	90.7	57.9	0.52	0.83	0.91	0.87
SVM	71	69.77	73.68	0.40	0.86	0.70	0.77

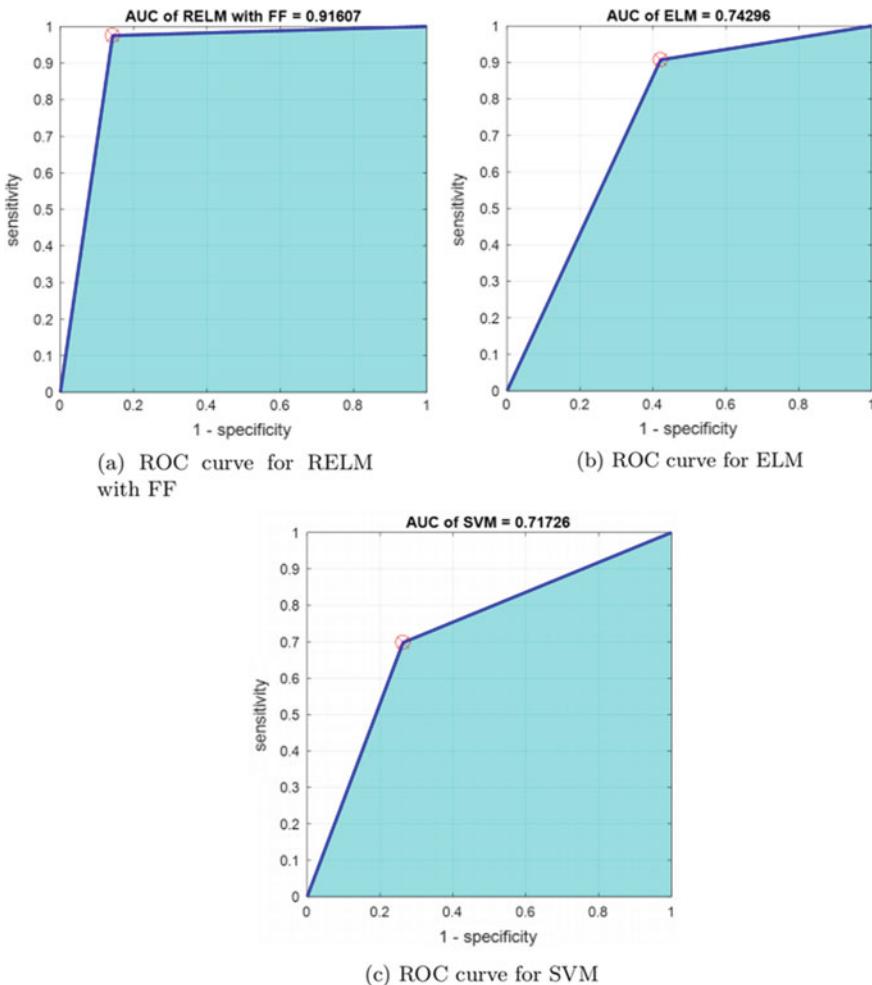


Fig. 2 Comparison of AUC and ROC curves of different methods

performance comparison of RELM with firefly, ELM and SVM. For the proposed model, the maximum accuracy, sensitivity, specificity and precision achieved are 93.44%, 97.5%, 85.71% and 0.93, respectively. The MCC, recall and F1 score are 0.86, 0.98 and 0.95, respectively. For SVM, radial basis kernel is used. The maximum accuracy, sensitivity, specificity and precision achieved for ELM are 80.65%, 90.7%, 57.9% and 0.83, respectively, and those for SVM are 71%, 69.77%, 73.68% and 0.86, respectively. It clearly shows that the proposed model gives better results than ELM and SVM. The receiver operator characteristic (ROC) curves for the implemented three methods are shown in Fig. 2. The area under the curve of the proposed method is much higher than that of SVM and ELM. Figure 3 shows the performance comparison

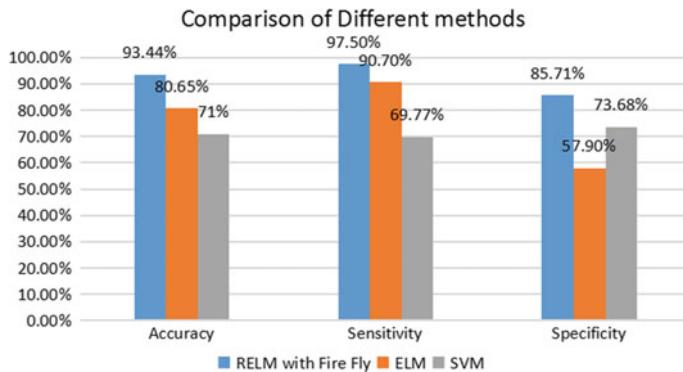


Fig. 3 Comparison of performance measures

Table 2 Performance comparison with existing methods

Paper	Technology Used	Accuracy (in %)
Alić et al. [6] [2017]	ANN	87.29
Syed et al. [7] [2018]	Tree partitioning adaptive support vector machine	89.67
Faruque et al. [8] [2019]	C4.5 Decision tree	73.5
Yahyaoui et al. [5] [2019]	Deep learning	76.81
Kaur et al. [9] [2020]	Linear kernel SVM	89
Proposed Method	RELM optimized by firefly algorithm	93.44
Implemented Method	ELM	80.65
Implemented Method	Radial basis kernel SVM	71

Bold letters are used to show the proposed method and also the methods implemented

of the implemented methods. Table 2 presents the performance comparison of the implemented methods with existing methods. The accuracy obtained from RELM with Firefly algorithm is 93.44% which clearly outperforms the existing methods.

4 Conclusion

The proposed model RELM with firefly optimization algorithm achieves maximum accuracy of 93.44%. The accuracy using ELM and SVM is 80.65% and 71%, respectively. The results obtained from the implemented methods show that the proposed model outperforms SVM and ELM. In future, it would be attempted to reduce the

number of features used for prediction so that the convergence rate becomes faster. This proposed model can be a robust model for the prediction of the chronic diabetes which is commonly found in today's world.

References

1. American Diabetes Association, Diagnosis and classification of diabetes mellitus. *Diabetes Care* **37**(Supplement 1), S81–S90 (2014)
2. P. Hemant, T. Pushpavathi, A novel approach to predict diabetes by cascading clustering and classification, in *2012 Third International Conference on Computing, Communication and Networking Technologies (ICCCNT'12)* (IEEE, 2012, July), pp. 1–7
3. M.F. Hashim, S.Z.M. Hashim, Comparison of clinical and textural approach for diabetic retinopathy grading, in *2012 IEEE International Conference on Control System, Computing and Engineering* (IEEE, 2012, November), pp. 290–295
4. M.E.H. Daho, N. Settouti, M.E.A. Lazouni, M.A. Chikh, Recognition of diabetes disease using a new hybrid learning algorithm for nefclass, in *2013 8th International Workshop on Systems, Signal Processing and their Applications (WoSSPA)* (IEEE, 2013, May), pp. 239–243
5. A. Yahyaoui, A. Jamil, J. Rasheed, M. Yesiltepe, A decision support system for diabetes prediction using machine learning and deep learning techniques, in *2019 1st International Informatics and Software Engineering Conference (UBMYK)* (IEEE, 2019, November), pp. 1–4
6. B. Alić, L. Gurbeta, A. Badnjević Machine learning techniques for classification of diabetes and cardiovascular diseases, in *2017 6th Mediterranean Conference on Embedded Computing (MECO)* (IEEE, 2017, June), pp. 1–4
7. R. Syed, R.K. Gupta, N. Pathik, An advance tree adaptive data classification for the diabetes disease prediction, in *2018 International Conference on Recent Innovations in Electrical, Electronics & Communication Engineering (ICRIECE)* (IEEE, 2018, July), pp. 1793–1798
8. M.F. Faruque, I.H. Sarker, Performance analysis of machine learning techniques to predict diabetes mellitus, in *2019 International Conference on Electrical, Computer and Communication Engineering (ECCE)* (IEEE, 2019, February), pp. 1–4
9. H. Kaur, V. Kumari, Predictive modelling and analytics for diabetes using a machine learning approach, in *Applied Computing and Informatics* (2020)
10. M.K. Pradhan, S. Minz, V.K. Shrivastava, A Kernel-based extreme learning machine framework for classification of hyperspectral images using active learning. *J. Indian Soc. Remote Sens.* **47**(10), 1693–1705 (2019)
11. G.B. Huang, Q.Y. Zhu, C.K. Siew, Extreme learning machine: theory and applications. *Neurocomputing* **70**(1–3), 489–501 (2006)
12. P. Satapathy, S. Dhar, P.K. Dash, A firefly optimized fast extreme learning machine based maximum power point tracking for stability analysis of microgrid with two stage photovoltaic generation system. *J. Renew. Sustain. Energy* **8**(2), 025501 (2016)
13. PIMA Dataset: U. M. L. Repository. <https://archive.ics.uci.edu/ml/index.php>

MIMO-OFDM Outdoor Channel Estimation Using Sparse Momentum Fractional Adaptive Filter



Swetaleena Sahoo, Sarita Nanda, and Harish Kumar Sahoo

Abstract In the next-generation multiple-input multiple-output (MIMO)-based cellular communication system, channel estimation is playing a major role for an efficient signal transmission. To reduce the sparsity of the practical wireless channel, a sparse penalty term is added in the adaptive algorithms to estimate channel parameters which plays a vital role for recovering the original data at the receiver end of MIMO-OFDM system. As the complexity of the MIMO system increases due to the spatial multiplexing, so to reduce the complexity of the system and to achieve fast convergence of the system, a momentum-based fractional-order LMS with sparse penalty term is used in this paper for an efficient channel estimation. A Jake's outdoor channel model is used as the base channel, and the parameters of the channel is estimated through sparse momentum fractional LMS (SmFLMS) for 16- QAM modulated transmitted signal in MIMO-OFDM system. The channel characteristics graph is plotted with respect to signal-to-noise ratio (SNR) values by comparing the proposed method with the mFLMS and other conventional algorithms. The comparative study of mean square error (MSE) is plotted by using MATLAB platform.

Keywords QAM · Sparsity norm · FLMS · MIMO · Channel estimation · MSE

S. Sahoo · S. Nanda (✉)
KIIT Deemed to be University, Bhubaneswar, India
e-mail: snandafet@kiit.ac.in

S. Sahoo
e-mail: swetaleenafet@kiit.ac.in

H. Kumar Sahoo
VSSUT, Burla, Sambalpur, India

1 Introduction

In the growing 5G digital communication environment, it is a challenging task to achieve a high data rate signal transmission with an extremely high quality of service (QoS). So to accomplish high spectral efficiency, multiple-input multiple-output (MIMO) signal transmission is a conventional approach in the fourth and fifth generation wireless digital communications systems. In MIMO system due to presence of multiple transmitting and receiving antennas, the transmission rate increases through the spatial multiplexing. On the other hand, the capacity of the transmitting channel has been increased by using orthogonal frequency division multiplexing (OFDM) techniques with the MIMO system that tends to enhance the complexity of the communication system. The exact estimation of channel impulse response (CIR) [1] is a vital and challenging task, and its accuracy has a significant impression on the overall performance of communication system. The signal processing tools are playing a significant role for various system identification applications. In the recent past, many researchers have proposed various effective adaptive algorithms for the estimation of wireless channel parameters of MIMO-OFDM system which includes modified least mean square (LMS) [2], normalized LMS (NLMS) [3], least mean square forth (LMS/F), recursive least square (RLS) [4], maximum likelihood, minimum mean square error(MMSE) [5], adaptive matching pursuit (NAMP), variable step size LMS/F(VSSLMS/F) [6], block LMS(BLMS), and fast block LMS (FBLMS). Though the stochastic gradient-based algorithms perform well for channel estimation application, these algorithms have some drawbacks. Such as fast block LMS gives a better convergence rather than conventional LMS algorithms, but the computational complexity increases, and in case of MMSE adaptive channel estimation by increasing the nonlinearity of the channel, performance declines. However, keeping these shortcomings into consideration, a new robust algorithm has been proposed based on momentum fractional LMS (mFLMS) for channel estimation. Again to reduce the sparse behavior of the wireless channels, various researchers have introduced different sparse penalty norms [7] in the cost function of the conventional algorithms to deal with the sparse nature of the channel. Inspired by the concept of mFLMS and to reduce the sparsity behavior of the Jake's outdoor channel model, here, a l_p -norm based reweighted zero-attracting (RZA) sparse penalty is introduced in mFLMS (SmFLMS) algorithm which is used for channel estimation in MIMO-OFDM system. The channel estimation process is measured by MSE plot and comparing its value with the conventional algorithms considering QAM signal as input to the transmitter. The content of the paper is organized as follows. Section 2 gives a description of the proposed channel estimation model for Jake's outdoor model. The simulation results for the proposed algorithm and a comparative discussion is given in Sect. 3, and Sect. 4 gives a concluding note.

2 Proposed Channel Estimation Algorithm for Jake's Outdoor Model

In this paper, a MIMO-OFDM system with $N_t \times N_r$ antenna configuration is considered, and Jake's model-based outdoor channel is used for transmission of modulated data stream. The N_t and N_r represent number of transmitting and receiving antennas, respectively. The channel characteristics are estimated through adaptive algorithm which leads to recover original data at the receiver end of the communication system.

The block diagram of the MIMO channel estimation process is shown in Fig. 1. Let a 16-QAM modulated signal is taken as input to the system which is taken as a_i where $i = 1, 2, 3, \dots, N_r$. The transmitted signal s is used as pilot sequence and given input to an OFDM modulator whose operation is performed by inverse Fourier transform (IFFT) and also a cyclic prefix is inserted to the signal and the output the modulator block is denoted by b_i . The next block denotes the outdoor channel which is taken as Jake's outdoor channel [8] whose base characteristic is denoted by H which is a channel matrix with $N_t \times N_r$ dimension. A fading channel can be described using a combination of channel gain and Rayleigh's distribution. The outdoor model is having a time varying impulse response depending upon the number of transmitting and the receiving antenna is denoted by:

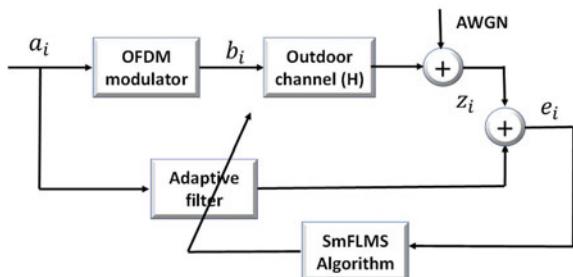
$$H_{t,r} = H_{t,r}^I + j * H_{t,r}^J \quad (1)$$

where $H_{t,r}^I$ the real part with t th transmitting antenna and r th receiving antenna and $H_{t,r}^J$ is the imaginary part of the channel impulse response.

As in this paper, 2×2 MIMO-OFDM is taken, so the channel impulse response is:

$$H = \begin{bmatrix} H_{1,1} & \dots & H_{1,r} \\ \vdots & \ddots & \vdots \\ H_{t,1} & \dots & H_{t,r} \end{bmatrix} \quad (2)$$

Fig. 1 Block diagram of MIMO-OFDM channel estimation model



$$H_{t,r}^I = 2 \sum_{n=1}^{P_0} (\cos \theta_n + \cos w_n t) + \sqrt{2}(\cos \theta_n + \cos w_\theta t) \quad (3)$$

$$H_{t,r}^J = 2 \sum_{n=1}^{P_0} (\sin \theta_n + \cos w_n t) + \sqrt{2}(\sin \theta_n + \cos w_\theta t) \quad (4)$$

where θ_n and θ_P represent the angle of arrival and the initial phase of the n th sinusoid with Doppler's shifted and maximum Doppler frequency denoted by f_θ , respectively, and $P_0 = (\frac{P}{2} - 1)/2$ where $\frac{P}{2}$ is an odd number which represents the number of multipath.

The complex channel impulse response of the Jake's model is represented by:

$$H = E_0 / \sqrt{2P_0 + 1} \{ H_{t,r}^I + j * H_{t,r}^J \} \quad (5)$$

where E_0 represents the average power of fading channel.

An additive Gaussian noise is added to the transmitted signal which is received after added with the channel characteristics. The received signal at the end of the transmitter is denoted by:

$$z_i = \frac{E_0}{\sqrt{2P_0 + 1}} H b_i + \gamma \quad (6)$$

where γ represents the AWGN noise with mean 0 and variance ρ^2 . The signal recovered at the end of the transmitter is passed through the equalizer which is the starting of the receiver, and after getting equalized signal, the original data signal is extracted by using demodulation.

The output of the channel is represented by $z_i = H b_i$ which is taken as the estimated signal where the H value is estimated with the help of various adaptive algorithm by reducing the error signal $e = d - z$ where d is the desired response and z is the estimated signal which leads to recover original signal at the receiver. Here, the value of H is estimated by adapting novel sparse-based adaptive weight updated formula from n th to $n + 1$ th iterations.

According to Chaudhary et al. [9], trade-off has been established between FLMS and mLMS [10] for the power signal parameter estimation application. Applying the mFLMS concept in channel estimation, the weight update equation becomes:

$$H(n+1) = H(n) + v(n+1) \quad (7)$$

$v(n+1)$ is the velocity term which is a combination of momentum term with previously calculated gradient.

$$v(n+1) = \theta v(n) + p(n) \quad (8)$$

$$v(n) = \alpha[H(n) - H(n-1)] \quad (9)$$

$p(n)$ is the gradient part of the FLMS weight update equation, and $\theta \in (0, 1)$ is a controlling parameter which controls the previous gradient part.

$$p(n) = \mu e(n)a_i(n) + \left(\frac{\mu_f}{\Gamma(2-f)} \right) e(n)a_i(n) \odot H^{1-f}(n) \quad (10)$$

where f represents the order of fractional derivative and gamma function represents by:

$$\Gamma(n) = (n-1)! \quad (11)$$

To estimate channel parameter in an efficient way, a sparsity concept is introduced with mFLMS in this paper. The conventional way of reducing the sparsity of the channel is the addition of a sparse penalty term in the weight update equation.

$$H(n+1) = H(n) + \text{adaptive update error} + \text{sparse penalty} \quad (12)$$

The sparse penalty based on $l_1 - norm$ based RZA algorithm is taken in this paper.

The cost function of the RZA LMS is

$$L_{\text{RZA}}(n) = e^2(n) + \rho_{\text{RZA}} \sum_{i=1}^N \log(1 + \theta_{\text{RZA}}|H(n)|) \quad (13)$$

In Eq. (12), ρ_{RZA} is the regularization parameter which reduces the channel sparseness and helps to control square of error.

$$H(n+1) = H(n) - \beta \frac{\partial L_{\text{RZA}}(n)}{\partial H(n)} \quad (14)$$

In place of $H(n)$, the weight update equation of mFLMS equation is added, and the final equation is represented in Eq. (15).

$$\begin{aligned} H(n+1) &= H(n) + \mu e(n)a_i(n) + \left(\frac{\mu_f}{\Gamma(2-f)} \right) e(n)a_i(n) \\ &\odot H^{1-f}(n) + \theta v(n) - \frac{1}{2} * \lambda(n) * \rho_{\text{RZA}} \end{aligned}$$

$$* \theta_{\text{RZA}} * \sum_{i=0}^{N-1} \frac{\text{sgn}(|H(n)|)}{1 + \theta_{\text{RZA}}|H(n)|} \quad (15)$$

$$\begin{aligned} H(n+1) &= H(n) + \mu e(n)a_i(n) + \left(\frac{\mu_f}{\Gamma(2-f)} \right) e(n)a_i(n) \\ &\odot H^{1-f}(n) + \theta v(n) - \varphi_{\text{RZA}} * \lambda(n) * \frac{\text{sgn}(|H(n)|)}{1 + \theta_{\text{RZA}}|H(n)|} \end{aligned} \quad (16)$$

where $\varphi_{\text{RZA}} = \frac{1}{2} * \rho_{\text{RZA}} * \theta_{\text{RZA}}$.

The weight update equation is shown in Eq. 16 which is the proposed RZA sparse-based mFLMS algorithm which is used to estimate the outdoor channel parameters.

3 Simulation and Results

In this paper, the performance evaluation of the channel estimation process is implemented for a 2×2 MIMO-OFDM system characterized by Jake's outdoor channel model whose simulation parameters are described in Table 1 and the simulation is performed by using MATLAB 2018 platform. The channel estimation process is evaluated by using the MSE vs number of iteration, where the MSE value is evaluated by the division of the mean square value of the error with the number of samples taken as input to the MIMO-OFDM system.

Figure 2 shows the plot for MSE vs number of iterations for first transmitting antenna with first receiving antenna. Here, it is shown that the black line is the proposed method which converges nearly about 200th iterations, whereas the other conventional algorithms converge after 500th iterations. However, the MSE plot is taken for 5000 iterations for showing the convergence performance. It is clearly

Table 1 Simulation parameter for MIMO-OFDM system

Parameters	Value
No of transmit antenna (N_t)	2
No of transmit antenna (N_t)	2
Modulation scheme	16-QAM MIMO-OFDM
Channel	Jake's outdoor model
IFFT size	128
Carrier frequency (f_c)	500 MHz
Doppler spectrum f_d	900 kHz
Mobility	100 km/h
SNR	20 dB
Time period (T_s)	$1e-6$ s

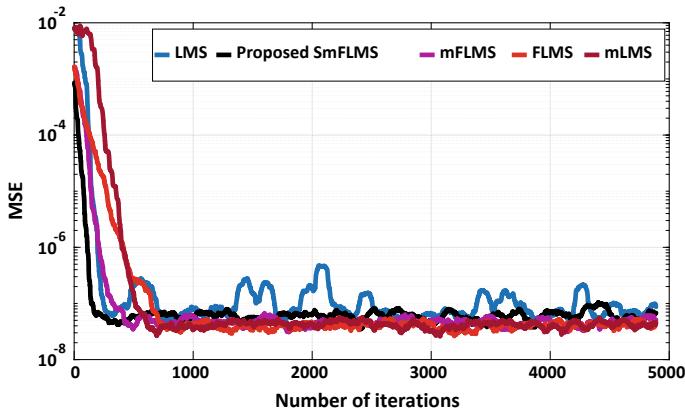


Fig. 2 MSE plot for Tx1 and Rx1

understood that the MSE plot for the proposed method converges faster than conventional algorithms. Figures 3, 4, and 5 show the MSE comparisons plots with respect to two transmitting antennas and two receiving antennas. From the result analysis, it is quite clear that MSE converges for sparse mFLMS is faster in comparison with other algorithms belonging to LMS family in case of outdoor channel estimation.

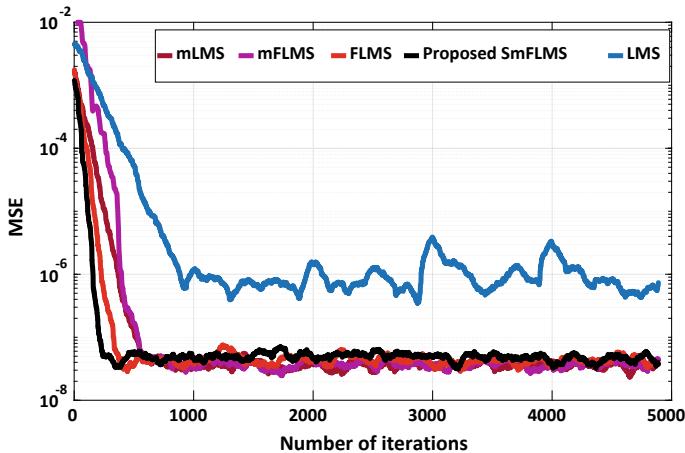


Fig. 3 MSE plot for Tx1 and Rx

Fig. 4 MSE plot for Tx2 and Rx1

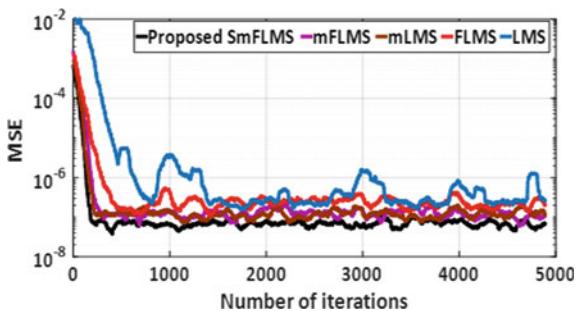
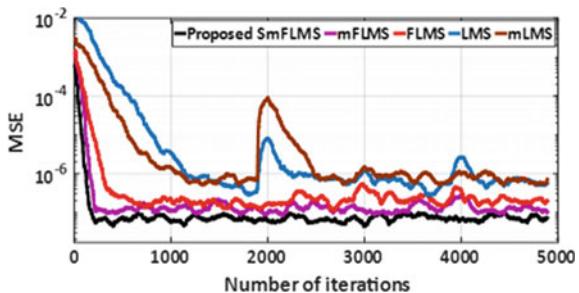


Fig. 5 MSE plot for Tx2 and Rx2



4 Conclusion

The proposed sparse momentum fractional-based adaptive algorithm reveals MSE in the range of 10^{-8} as well as affords a stable convergence property in comparison with the other state-of-the-art approaches. The integration of sparse penalty term with the momentum-based fraction-adaptive algorithm outperforms in the Jake's outdoor channel environment thereby proving its robustness against nonlinearity, ISI, and other channel impairments. The proposed channel estimation approach can also be a potent candidate for practical indoor multipath channels to reconstruct the received bit sequence in presence of fading and user mobility.

References

1. R. Strobel, V. Oksman, Channel estimation. Google Patents (2017)
2. Z. Zhang, L. Xiao, X. Su, J. Zeng, X. Xu, A channel estimation method based on the improved LMS algorithm for MIMO-OFDM systems, in *12th International Symposium on Medical Information and Communication Technology* (ISMICT), pp. 1–5 (2018)
3. P. Saxena, J. Bhalani, Normalized least mean square adaptive filter superimposed on whitening rotation matrix for semi-blind MIMO channel estimation. *Int. J. Comput. Sci. Inf. Secur. (IJCSIS)* **15**(5), 127–131 (2017)
4. D. Sharma, Recursive least square technique for channel estimation for transmit diversity case in MIMO-OFDM. *Int. J. Comput.* **4**(4), 18–24 (2015)

5. A.K. Shrivastava, A comparative analysis of LS and MMSE channel estimation techniques for MIMO-OFDM system. *Int. J. Sci. Res. Dev.* **1**(8), 44–48 (2015)
6. S. Sahoo, S. Nanda, H.K. Sahoo, Adaptive nonlinear equalizer for SISO wireless channels, in *2018 International Conference on Applied Electromagnetics, Signal Processing and Communication* (AESPC), pp. 1–6 (2018)
7. G. Gui, F. Adachi, Stable adaptive sparse filtering algorithms for estimating multiple-input–multiple-output channels. *IET Commun.* **8**(7), 1032–1040 (2014)
8. H.K. Sahoo, B. Mohanty, B. Patnaik, Block and fast block sparse adaptive filtering for outdoor wireless channel estimation and equalization. *Wirel. Pers. Commun.* **98**(3), 3003–3019 (2018)
9. S. Zubair, N.I. Chaudhary, Z.A. Khan, W. Wang, Momentum fractional LMS for power signal parameter estimation. *Sign. Process.* **142**, 441–449 (2018)
10. J.J. Shynk, S. Roy, The LMS algorithm with momentum updating. *IEEE Int. Symp. Circ. Syst.* 2651–2654 (1988)

Real-Time Electromyographic Hand Gesture Signal Classification Using Machine Learning Algorithm Based on Bispectrum Feature



Deepa Nair, R. Rajapriya, and K. Rajeswari

Abstract A real-time hand gesture classification model based on a robust feature extraction and machine learning algorithm is proposed in this paper. The Delsys Trigno Wireless EMG System is used to collect the EMG database. To improve the performance of EMG pattern recognition system, a new feature set based on bispectrum feature extraction is used as it can detect the non-Gaussian and nonlinear characteristics of EMG, unlike conventional feature set. The feature extracted is then fed to the classifiers to classify two hand gestures, hand close and hand open. The performance of the proposed pattern recognition system is measured in terms of their classification accuracies. The classifier algorithms used in this work are set of base classifiers such as Naive Bayes, support vector machine (SVM) and decision tree. Further to achieve increased classification accuracy ensemble classifiers such as random forest, gradient boost and AdaBoost are employed. From the experimental results, it is observed that the AdaBoost performs better with high classification accuracy of 99% as compared to other classifiers such as SVM, decision tree, Naive Bayes, random forest and gradient boost with accuracies of 81%, 75%, 81%, 75% and 75%, respectively. Due to its adaptiveness in correcting the weak learners (base classifiers) in favor of reducing misclassification by the previous classifiers, AdaBoost resulted in higher accuracy compared to other classifiers.

Keywords Surface EMG · Delsys Trigno · Bispectrum · Random forest · Gradient boost · AdaBoost.

1 Introduction

There are nearly 185,000 trauma-related amputations occurring every year. The recovery of hand gesture is one of the major problems for patient suffering from trauma-related disease like vascular disease and brain stroke. To combat with their

D. Nair (✉) · R. Rajapriya · K. Rajeswari
Thiagarajar College of Engineering, Madurai, Tamilnadu, India
e-mail: naairdeepa@gmail.com

daily-life activities, prosthetic limb is used to compensate for the function of upper limb. Recent years have shown so many inventions to make the life of amputees easier, and one such prominent invention is the myoelectric upper limb prostheses control with electromyogram (EMG) [1]. EMG represents bioelectrical signal acquired from the muscles during the muscle contraction and relaxation. It contains rich information about muscular activity occurring during a hand gesture of the person. From EMG, patterns are formed so as to classify the intended hand gesture using pattern recognition (PR). EMG-PR design consists of following processes: (i) EMG data acquisition, (ii) data segmentation, (iii) feature extraction and (iv) gesture classification [1]. Many researchers have begun to study the features of the sEMG signals, such as time-domain features, frequency-domain features and time–frequency features [2, 3]. However, the performances of these features were weak [4], and new features extracted based on EMG’s nonlinear properties are being explored recently [5].

Bispectrum has the capability to detect and characterize the non-Gaussian and nonlinear characteristics of EMG, and therefore, it is utilized for feature extraction, in this paper. This paper focuses on proposing a new feature extraction method that utilizes sEMG—bispectrum for improving hand gesture classification. The strength of this work is that it analyzes the robustness of proposed feature set using various base and ensemble classifiers such as SVM, Naive Bayes, decision tree, random forest, gradient boost and AdaBoost. The work is discussed in the following order: 2. Proposed methods 3. Results and discussion and 4. Conclusion

2 Proposed Methods

The proposed EMG gesture recognition process includes: database acquisition setup, EMG data acquisition protocol, data segmentation, bispectrum based feature extraction and classification.

2.1 Database Description

EMG signal is acquired utilizing the Delsys Trigno Wireless EMG System (shown in Fig. 1a). This wireless electromyogram (w-EMG) measurement equipment is equipped with 3-axis accelerometers and noise-resistant sensors. In Trigno, the EMG signal is transmitted to the receiving base station through a wireless protocol. This helps to minimize the data latency of the EMG sensors. This system can also integrate with other equipment through a variety of interfaces, and the measurement is viewed and can be processed with software as shown in Fig. 1b. Patented parallel-bar technology is used for recording which gives high-quality biomedical signals. This has a multifunctional design 16 possible EMG channels with 3-axis accelerometer in each sensor 37 mm 26 mm × 15 mm size. This can perform well 40 m distance. It has a durable battery lifetime of 8 hours, resolution of 16-bit 2000 Hz sampling rate.

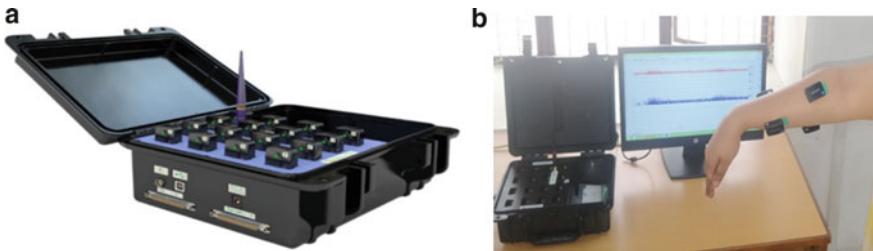


Fig. 1 **a** Shows Delsys Trigno Wireless EMG System and **b** shows setup preparation

2.2 Experiment Protocol

A single subject is recruited to perform two classes of gesture, i.e., hand close and hand open. The recruited person is normally limbed with no neurological or muscular disorder, and thus, the dataset is recorded using Delsys Trigno Wireless EMG system. The sensors are placed on the skin exactly over the center of the muscle belly of the corresponding muscles [6]. To achieve maximum signal strength, the sensors are placed in the line of the muscle fiber's length. The four EMG sensors were placed around the circumference of the right hand and processed by Delsys Trigno Wireless EMG system as shown in Fig. 2a. The sensors were carefully positioned to get accurate measurements for the considered hand gesture. Each of the four sensors are on the palmaris longus, extensor pollicis longus, extensor indicis and extensor digitorum muscles. The sensors are placed with equal spacing, six trials of each gesture are acquired, while each gesture is held for a time period of 5s, with a resting period of 5s between the trials. Finally, a plot is obtained as shown in Fig. 2b, and the EMG signal is stored using the software for later off-line analysis.

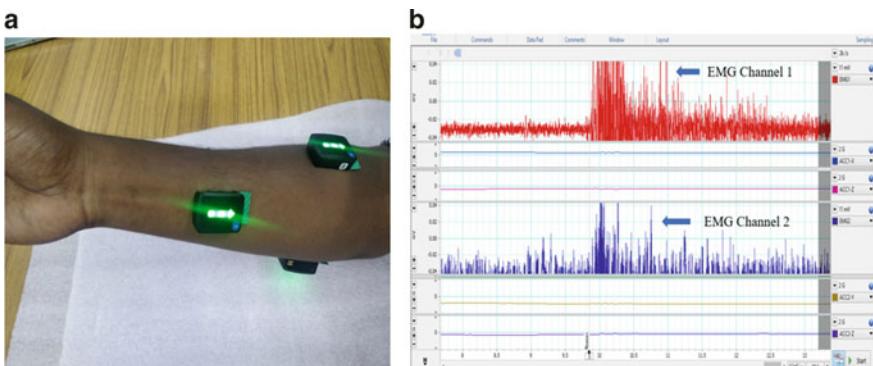


Fig. 2 **a** Shows electrode positions and **b** shows EMG recording

2.3 Data Segmentation

The EMG signal is analyzed offline; therefore, the data is segmented in a length of 256 samples from the raw EMG data using a non-overlapping window. Larger windows acquire more time; hence, it is not suitable to collect and process larger dataset. For each segment, the proposed bispectrum-based classification is performed.

2.4 Bispectrum-Based Feature Extraction

As EMG from neuromuscular system exhibits nonlinearity, bispectrum, higher-order statistics is used for feature extraction in this paper. Bispectrum is the Fourier transform of triple correlation [7]. Algorithm for bispectrum feature extraction is discussed below:

Let $x(n)$, $n = 1, \dots, N$ be the available data

1. Data is segmented into K records of M samples each, i.e., $N = KM$.
2. Let the data in i th segment is given as $\{x^i(n), n = 0, 1, 2, \dots, M - 1\}$, $i = 1, \dots, K$
3. Fourier transform is performed $U_x^i(f)$ for each segment.

$$U_x^i(k) = \sum_{n=0}^{M-1} x^i(n) \exp(-j2\pi nk/M), i = 1, 2, \dots, K, k = 0, 1, \dots, M - 1 \quad (1)$$

4. The third-order spectrum of each segment is obtained as

$$B^i(k_1, k_2) = U_x^i(k_1) U_x^i(k_2) U_x^{i*}(k_1 + k_2), i = 1, 2, \dots, K, k_1 = k_2 = 0, 1, \dots, M - 1 \quad (2)$$

5. Maximum magnitude of each segment is extracted as a feature.

$$Z^i = (\max(B^i(k_1, k_2))) \quad (3)$$

2.5 Classification

Classification is a supervised learning approach; here the machine learns and trains the data input given to it and is used to classify new observation as bi-class or multiclass data. The various classification algorithms used in this work are base classifiers such as Naive Bayes, support vector machine (SVM), decision tree and ensemble classifiers such as random forest, gradient boost and AdaBoost. This feature extraction and these classifiers are implemented in Python 3.

3 Results and Discussion

With the dataset, collected as discussed in Sects. 2.1 and 2.2, the gesture recognition with feature set extracted as discussed in Sect. 2.4 is evaluated by the classification accuracies of various classifiers discussed in Sect. 2.5. The machine learning algorithms are capable of working well across subjects as well as for the same subject in various positions. Machine learning algorithms have wide applications in the design of myoelectric prosthetic devices, and fivefold cross-validation is used in the work with 90% training and 10 % testing data. The classifier selection is based on classification accuracy of algorithm obtained through the confusion matrix. Confusion matrix provides information in finding out the following parameters: accuracy, precision, recall, F1-score and support.

Tables 1 and 2 show the performance measures of base and ensemble classifiers such as precision, recall, F1-score and support of confusion matrix. The obtained outcomes shown in Table 3 show that the bispectrum-based feature extraction provides superior results in terms of average classification accuracy for AdaBoost. Further from the results, it is observed that the proposed feature has better classification performance than other existing traditional features [2, 4, 5] such as mean absolute value (MAV), wavelength (WL), zero crossing (ZC) and slope sign changes (SSC) [3].

Table 1 Performance measures of base classifiers

Base classifiers	Class	Precision	Recall	f1 score	Support	Accuracy (%)
SVM	0	0.80	0.89	0.84	9	81
	1	0.83	0.71	0.77	7	
Decision tree	0	0.73	0.89	0.80	9	75
	1	0.80	0.57	0.67	7	
Naive Bayes	0	0.80	0.89	0.84	9	81
	1	0.83	0.71	0.77	7	

Table 2 Performance measures of ensemble classifiers

Ensemble classifiers	Class	Precision	Recall	F1-score	Support	Accuracy (%)
Random forest	0	0.73	0.89	0.80	9	75
	1	0.80	0.57	0.67	7	
Gradient boost	0	0.69	1.00	0.82	9	75
	1	1.00	0.43	0.60	7	
AdaBoost	0	1.00	1.00	1.00	9	99
	1	1.00	1.00	1.00	7	

Table 3 Average classification accuracy of classifiers

Base classifiers		Ensemble classifiers	
Classifier	Accuracy of the classifier (%)	Classifier	Accuracy of the classifier (%)
SVM	81	Random forest	75
Decision tree	75	Gradient boost	75
Naives Bayes	81	AdaBoost	99

4 Conclusion

The experiments and studies in this work are focused on designing an effective sEMG pattern recognition system to classify various hand gestures with maximum classification accuracy. Bispectrum proves to be a strong tool to detect and characterize non-Gaussianity and nonlinearity in EMG. Bispectrum, the third-order statistics is used for feature extraction in this paper. The feature extracted is then fed to the classifiers, and the performance is measured in terms of their accuracies. Among the classifiers employed, AdaBoost gives highest classification result of 99%, because of its adaptiveness in correcting the subsequent weak learners (base classifiers) in favor of reducing misclassification by previous classifiers.

Acknowledgements This work is part of the research programme supported by AICTE-RPS.

References

1. A. Accogli, L. Grazi, S. Crea, A. Panarese, J. Carpaneto, N. Vitiello, S. Micera, Emg-based detection of user's intentions for human-machine shared control of an assistive upper-limb exoskeleton, in *Wearable Robotics: Challenges and Trends* (Springer, Heidelberg, 2017), pp. 181–185
2. T. Hussain, N. Iqbal, H. F. Maqbool, M. Khan, M. I. Awad, A.A. Dehghani-Sanij, Intent based recognition of walking and ramp activities for amputee using semg based lower limb prostheses, in *Biocybernetics and Biomedical Engineering* (2020)
3. G. Li, J. Li, Z. Ju, Y. Sun, J. Kong, A novel feature extraction method for machine learning based on surface electromyography from healthy brain. *Neural Comput. Appl.* **31**(12), 9013–9022 (2019)
4. A. Phinyomark, F. Quaine, S. Charbonnier, C. Serviere, F. Tarpin-Bernard, Y. Laurillau, Emg feature evaluation for improving myoelectric pattern recognition robustness. *Expert Syst. Appl.* **40**(12), 4832–4840 (2013)
5. N. Nazmi, M.A. Abdul Rahman, S.-I. Yamamoto, S.A. Ahmad, H. Zamzuri, S.A. Mazlan, A review of classification techniques of emg signals during isotonic and isometric contractions. *Sensors* **16**(8), 1304 (2016)
6. J. Hua, G. Li, D. Jiang, H. Zhao, J. Qi, An optimized selection method of channel numbers and electrode layouts for hand motion recognition. *Int. J. Humanoid Robot.* **16**(04), 1941006 (2019)
7. D. Nair, R. Rajapriya, K. Rajeswari, Efficient electromyographic hand gesture signal classification using machine learning algorithm based on bispectrum feature, *Proceedings of International Conference on Instrumentation and Control Engineering (ICECON - 2019)*, Dec 2019

Design of Microwave Filters Using Swarm-Based Optimization Techniques



Piali Chakraborty , Arindam Deb , and Jyoti Ranjan Panda

Abstract The traditional and random drift particle swarm optimization algorithms have been applied to the design of microwave band-pass filter in this paper. The random drift particle swarm optimization has been tuned for best performance through proper choice of control parameter with respect to the defined cost function. The performance of the tuned random drift particle swarm optimization is further compared with that of the conventional particle swarm optimization algorithm. The study on performance comparison has been done with two different swarm sizes of 40 and 30. Random drift particle swarm optimization depicts rapid convergence rate and statistically superior performance compared to the traditional particle swarm optimization algorithm for both the population sizes. Although the thermal coefficient in random, drift particle swarm optimization is varied dynamically over iterations, the drift coefficient values of 1.2 and 1.3 prove to be optimal for population sizes of 40 and 30, respectively. The resultant filter based on the parameters obtained from the best run of the random drift particle swarm optimization shows a better frequency response compared to that from the traditional particle swarm optimization.

Keywords Random drift · Particle swarm optimization · Microwave filter

1 Introduction

Microwave filters are an integral part of wireless communication systems. Design of microwave filter using optimization algorithms has been attempted in the past [1]. A novel optimization technique in this context has been reported in [2]. This paper studies the suitability of random drift particle swarm optimization (RDPSO) in the design of microwave band-pass filter with different control parameter settings and population sizes. Statistical performance comparison of the RDPSO algorithm is also done with conventional particle swarm optimization (PSO). The work reported in

P. Chakraborty · A. Deb () · J. R. Panda

School of Electronics Engineering, Kalinga Institute of Industrial Technology, Deemed to be University, Bhubaneswar 751024, India
e-mail: arindamdeb2004@yahoo.co.in

the paper is expected to help future researchers in judiciously choosing the control parameters of RDPSO algorithm when applying it to filter design problems. The RDPSO algorithm has been applied to power economic dispatch [3] and array pattern synthesis problems [4]. An analysis of the algorithm has also been reported in [5].

The paper is arranged in different sections. Section 1 provides the introduction and objective of the work reported with respect to microwave band-pass filter. Section 2 provides a brief overview of the two variants of PSO algorithms used in this paper. Section 3 explains the filter circuit and the associated cost function for optimization. Section 4 provides the performance comparison between the tuned RDPSO and conventional PSO algorithms with respect to the defined cost function in Sect. 3. The tuning and performance comparison has been done for two different population sizes of 40 and 30. Section 5 concludes the paper.

2 Optimization Algorithms

Application of conventional particle swarm optimization to electromagnetic problems has been reported in [6]. The PSO algorithm works by controlling the movement of bees in a swarm in the solution space of the optimization problem. The coordinates of the bees represent the possible solution for the optimization variables. The position updates of the bees along different dimensions occur based on their velocities in those directions. The velocity update is governed by the present position, local best position, global best position, initial velocity as well as the inertial weight, local and global acceleration coefficients [6]. Based on extensive studies performed by previous researchers on PSO, it is widely accepted that both the local and global acceleration coefficients and the control parameters of the algorithm are to be set at 2.0 for optimal or near optimal performance. The inertial weight varies linearly from 0.9 to 0.4 over the number of iterations or generations.

The RDPSO algorithm [3] works by assuming electron movement in a metal under the effect of an electric field. In this case, the electrons have two velocity components, i.e., the drift velocity component and thermal velocity component. The velocity update equations are given in [3], and the control parameters involved are the drift coefficient (β) and thermal coefficient (α). In our study, the thermal coefficient (α) is varied linearly from 0.9 to 0.3 over the number of generations. The drift coefficient (β) is tuned for optimal performance as explained in Sect. 4.

3 Cost Function Formulation

The general structure of the microwave band-pass filter based on transmission lines considered in this paper is shown in Fig. 1.

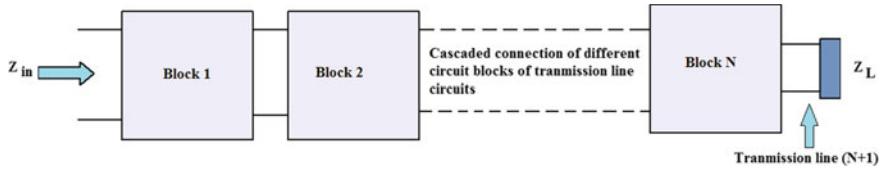


Fig. 1 General representation of the microwave band-pass filter

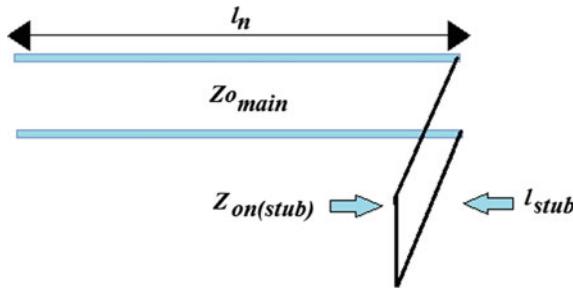


Fig. 2 Transmission line circuit description for each Block_n in microwave band-pass filter

Each ‘ Block_N ’ represents a transmission line circuit with a main line segment having length l_n and characteristic impedance $Z_{\text{O}_{\text{main}}}$, terminated by a short-circuited stub of length l_{stub} and characteristic impedance $Z_{\text{on}(\text{stub})}$.

Ten such blocks as shown in Fig. 2 are considered while optimizing the band-pass filter circuit. There are 23 optimization variables; l_1-l_{10} , lengths of main transmission line segments for different blocks; l_{11} , length of the transmission line connected to load; $Z_{o1}-Z_{o10}$, characteristic impedances of the short-circuited stubs; l_{stub} , length of each stub; $Z_{\text{O}_{\text{main}}}$, characteristic impedance of the main transmission line.

The objective of the design is to achieve band-pass response between 5.15 and 5.35 GHz (indoor wireless communication) and as high attenuation as possible outside this range. The response of the filter can be explained in terms of power loss ratio (PLR) [7] which depends on the reflection coefficient, $\tau(f)$ looking from the input port. The output of the filter is assumed to be terminated by 50Ω . The cost function for the filter is formulated as follows:

$$\text{Cost} = C_1 \frac{1}{N} \sum_{i=1}^N \text{Obj}_i + C_2 (|20 \log_{10}(\tau(f_{L-}))| + |20 \log_{10}(\tau(f_{h+}))|) \quad (1)$$

where f_{L-} and f_{h+} represent two frequencies slightly below 5.15 GHz and above 5.35 GHz, respectively. In this paper, they are considered to be at 4.75 GHz and 5.75 GHz, respectively. The first term in the cost function is aimed to find optimal values for the circuit parameters corresponding to desired pass band and stop band. ‘ N ’ represents the total number of frequency points considering which the cost function is evaluated. The second term in the cost function is incorporated to increase

attenuation outside the pass band. C_1 and C_2 are chosen as 1 and 10, respectively, after repeated numerical experiments. Obj_i is formulated as follows:

$$\text{Obj}_i = |20 \log_{10}(\tau(f_i))| \quad \text{if } f_i \leq 5.15 \text{ GHz or } f_i \geq 5.35 \text{ GHz} \quad (2a)$$

$$\begin{aligned} \text{Obj}_i &= C_3 |-20 \log_{10}(\tau(f_i)) - 40| \quad \text{if } 5.15 \text{ GHz} \leq f_i \leq 5.35 \text{ GHz} \\ &\text{and } 20 \log_{10}(\tau(f_i)) > -40 \end{aligned} \quad (2b)$$

C_3 is chosen to be 10. Both the algorithms are run for 20 independent trials starting from the same initial population. Each trial of an algorithm comprises 100 iterations. The statistical performance parameters of the algorithms are recorded for comparison.

4 Results and Observations

The RDPSO algorithm is tuned for optimal performance corresponding to the cost function described in the previous section. Tuning for the conventional PSO is not done here since many investigative works have previously reported that local and global acceleration coefficients with values of 2.0 give optimal or near optimal algorithm performance. For the RDPSO algorithm, ‘ α ’ is dynamically varied from 0.9 to 0.3, and only ‘ β ’ is tuned by varying it between 1.1 and 2.0 in steps of 0.1 as given in Table 1. The tunings are done for two different population sizes of 40 and 30.

The performance of the RDPSO algorithm is better than the conventional PSO algorithm for all values of β in Table 1, except 2.0 in terms of mean cost value, when the population size is considered to be 40. The best or lowest cost value obtained in case of the RDPSO algorithm with $\beta = 1.2$ is much lower compared to that from the PSO. This parameter value is considered to be optimal for RDPSO with population size of 40.

For population size of 30, the performance of the RDPSO algorithm is better than the conventional PSO algorithm for all values of β except 1.6, 1.8, 1.9 and 2.0 in terms of mean cost value. In this case, the lowest mean and best cost values are obtained for $\beta = 1.3$.

The RDPSO algorithm shows a rapid convergence than the PSO algorithm as shown in Fig. 3.

The parameters of the optimal filters corresponding to the best runs of the PSO and RDPSO algorithm with $\beta = 1.2$, considering population size of 40, are given in Table 2. ‘ λ_o ’ is the wavelength corresponding to frequency of 5.25 GHz.

The frequency response of the microwave band-pass filters with parameters corresponding to the best runs of the PSO and RDPSO algorithm (with $\beta = 1.2$) considering population size of 40 is given in Fig. 4.

The filter designed using PSO has a bandwidth of 240 MHz, whereas the one using RDPSO has a bandwidth of 310 MHz.

Table 1 Performance comparison of optimization algorithms

Optimization algorithms	Population size = 40			Population size = 30		
	Mean cost value	Best cost value	Standard deviation	Mean cost value	Best cost value	Standard deviation
PSO	37.07	29.08	4.87	43.94	33.58	6.95
RDPSO ($\beta = 1.1$)	33.49	19.31	11.12	39.77	18.15	13.18
RDPSO ($\beta = 1.2$)	27.85	11.08	7.94	38.86	23.46	11.55
RDPSO ($\beta = 1.3$)	31.65	16.21	12.28	35.63	11.32	14.47
RDPSO ($\beta = 1.4$)	27.31	22.68	2.27	37.18	13.08	14.67
RDPSO ($\beta = 1.5$)	26.92	16.31	7.41	36.12	22.10	12.32
RDPSO ($\beta = 1.6$)	30.25	19.27	7.23	44.56	18.00	13.66
RDPSO ($\beta = 1.7$)	29.21	12.06	6.03	39.46	22.83	11.95
RDPSO ($\beta = 1.8$)	30.98	20.19	5.1	45.53	18.83	12.94
RDPSO ($\beta = 1.9$)	35.84	25.62	5.35	51.45	29.16	8.54
RDPSO ($\beta = 2.0$)	39.84	24.48	6.43	51.86	36.33	6.71

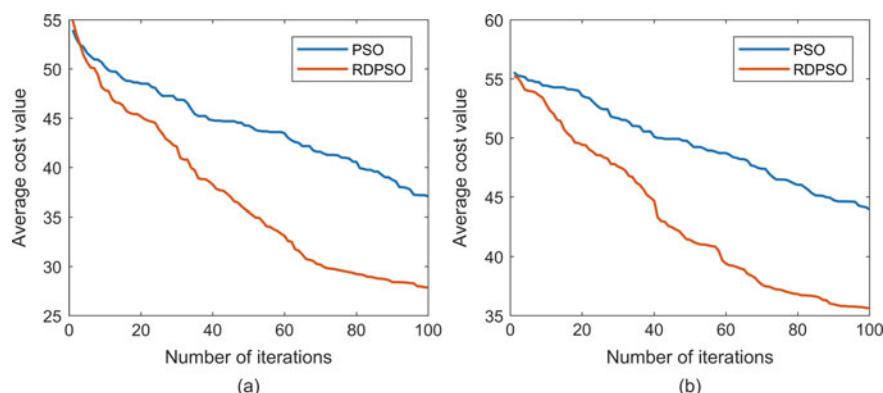
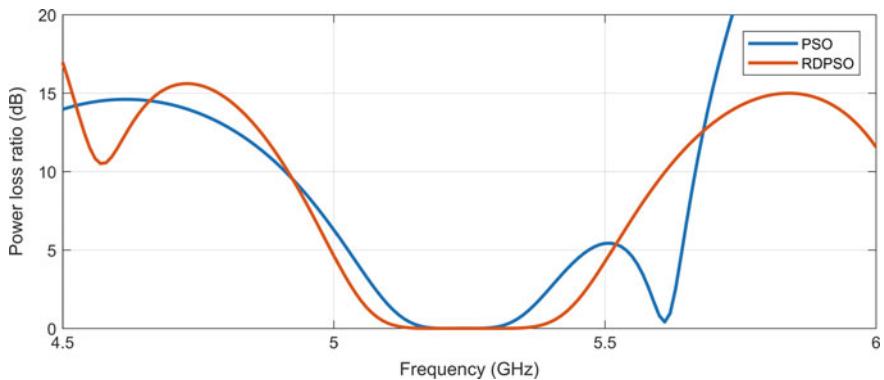
**Fig. 3** Variation of the average cost values with iterations for PSO and RDPSO algorithms with $\beta = 1.2$ and 1.3 for **a** population size = 40, **b** population size = 30, respectively

Table 2 Parameters of the optimal filters

PSO				RDPSO ($\beta = 1.2$)			
Lengths of lines	Fraction of λ_o	Characteristic impedances	Ohms	Lengths of lines	Fraction of λ_o	Characteristic impedances	Ohms
l_1	0.4861	Z_{o1}	73.25	l_1	0.2911	Z_{o1}	69.76
l_2	0.1000	Z_{o2}	20.00	l_2	0.3247	Z_{o2}	58.44
l_3	0.2778	Z_{o3}	90.70	l_3	0.2478	Z_{o3}	66.59
l_4	0.1000	Z_{o4}	78.52	l_4	0.4159	Z_{o4}	88.42
l_5	0.1036	Z_{o5}	20.00	l_5	0.3222	Z_{o5}	83.85
l_6	0.3763	Z_{o6}	47.55	l_6	0.1678	Z_{o6}	92.32
l_7	0.1001	Z_{o7}	92.06	l_7	0.3439	Z_{o7}	36.70
l_8	0.4880	Z_{o8}	24.74	l_8	0.4656	Z_{o8}	40.92
l_9	0.4376	Z_{o9}	32.02	l_9	0.4841	Z_{o9}	60.60
l_{10}	0.1044	Z_{o10}	21.52	l_{10}	0.3861	Z_{o10}	84.95
l_{11}	0.3296	$Z_{o\text{main}}$	64.45	l_{11}	0.2461	$Z_{o\text{main}}$	80.25
l_{stub}	0.3559			l_{stub}	0.1002		

**Fig. 4** Power loss ratio versus frequency for optimal filters obtained from best runs of PSO and RDPSO with $\beta = 1.2$

5 Conclusion

RDPSO algorithm with tuned value of drift coefficient shows better statistical performance and rapid convergence than the traditional PSO algorithm. Future work will focus on the suitability of RDPSO algorithm for the design of planar microwave filters.

References

1. J. Paleček, P. Vestenický, M. Vestenický, J. Spalek, Differential evolution algorithm in optimal design of microstrip filter, in *Aspects of Computational Intelligence: Theory and Applications. Topics in Intelligent Engineering and Informatics*, ed. by L. Madarász, J. Živčák, vol 2 (Springer, Berlin, 2013)
2. F. Feng, W. Na, W. Liu, S. Yan, L. Zhu, J. Ma, Q.-J. Zhang, Multifeature-assisted neuro-transfer function surrogate-based EM optimization exploiting trust-region algorithms for microwave filter design. *IEEE Trans. Microw. Theory Tech.* **68**(2), 531–542 (2020)
3. J. Sun, V. Palade, X.-J. Wu, W. Fang, Z. Wang, Solving the power economic dispatch problem with generator constraints by random drift particle swarm optimization. *IEEE Trans. Ind. Electron.* **10**(1), 222–232 (2014)
4. B. Mishra, A. Deb, Synthesis of antenna arrays with flat-top pattern using conventional and random drift particle swarm optimization algorithms, in *Emerging Trends in Electronic Devices and Computational Techniques*. IEEE, Kolkata, India (2018), pp. 1–4
5. J. Sun, X. Wu, V. Palade, W. Fang, Y. Shi, Random drift particle swarm optimization algorithm: convergence analysis and parameter selection. *Mach. Learn.* **101**, 345–376 (2015)
6. J. Robinson, Y. Rahmat-Samii, Particle swarm optimization in electromagnetics. *IEEE Trans. Antennas Propag.* **52**(2), 397–407 (2004)
7. D.M. Pozar, *Microwave Engineering*, 3rd edn. (Wiley, New York, 2005)

Author Index

A

- Abhishek, D. J., 459
Acharya, Adyasa Priyadarsini, 7
Agarwal, Sumanshu, 127
Akanksha, Kumari, 225
Alamelu, J. V., 187
Amiri, I. S., 1
Anand, Utkarsh, 427
Anne Frank Joe, A., 303
Aryan, Yashaswi, 427
Aseem, K., 443
Atre, Rashmi, 109

B

- Babu, K. S., 467
Bahadure, Nilesh Bhaskarao, 347
Bakshi, Amit, 493
Balaji, M., 65
Balasubramanian, Karthi, 249
Banajyoti, Adyasha, 81
Bandyopadhyay, Ahana, 41, 265
Bansal, Abhishek, 371
Baruah, Satyabrat Malla Bujar, 399
Behera, Manas Chandan, 313
Behera, Seshadri Binaya, 475
Behera, Subhrakanta, 475
Behera, Suchismita, 73
Bhardwaj, Rishu, 339
Bhattacharjee, Debotosh, 99
Bhende, C. N., 81
Biswal, Priyabrata, 7
Biswas, Anindya Kumar, 141
Borah, Shekhar Suman, 23
Bose, Moinak, 179

Budihal, Suneeta V., 257

Burde, Shreyas, 257

C

- Chaki, Aunkit, 295
Chakraborty, Amit, 31
Chakraborty, Anwita, 41, 265
Chakraborty, Piali, 565
Chand, Pawan Kumar, 363
Choudhary, Rishav, 241
Choudhury, Bikram, 281

D

- Darshan, M. N., 117
Das, Biswajit, 399
Das, Debi Prasad, 355
Dasgupta, Mou, 141
Dash, Sandeep Kumar, 493
Dash, Sonali, 195
Das, Niva, 73
Das, Priya, 541
Das, Satyadeep, 157
David, Christina, 215
Deb, Arindam, 565
Dhang, Debraj, 157
Durga Prasad, Ch., 435, 459
Dutta, Koustav, 407

E

- Elamaran, V., 303

G

- Ganesan, P., 303
 Gautham Gopi, S., 117
 George Rajan, Rufina Flora, 233
 Ghatak, Sanjoy, 99
 Ghosh, Maitrish, 501
 Ghosh, Mourina, 23
 Ghosh, Sourodipt, 41, 179, 295
 Guleria, Anchal, 109
 Gundavarapu, Vishal, 65
 Gupta, Ritik, 313

H

- Hegde, Vasudha, 215
 Hiremath, Shraddha B., 451
 Hota, Aditya Kumar, 321

I

- Iyer, Nalini C., 451, 481

J

- Jagadeesh, P., 459
 Jena, U. R., 195
 Jesseca, E. Shirley, 215
 Jhamb, Deepika, 379

K

- Kalivarapu, Puja, 233
 Kanagaraj, G., 89
 Kar, Sambit Prasad, 501
 Kar, T., 163
 Kaur, Amandeep, 389
 Kaur, Manpreet, 389
 Kaur, Preetinder, 389
 Khanra, Soumya, 427
 Khurana, Meenu, 389
 Kotabagi, Sujata S., 451
 Krupa, B. N., 273
 Kudeshia, Ankit, 179, 295
 Kumar, Ayush, 265
 Kumari, Surabhi, 225
 Kumar, Kundan, 127
 Kumar, Rajat, 265
 Kumar, Raunak, 339
 Kumar Sahoo, Harish, 549
 Kumar, Sananda, 533
 Kumar, Sudesh, 371
 Kumar, Tallam Vineeth, 117
 Kundu, Ishita, 41, 265

L

- Leo Joseph, L. M. I., 303

M

- Mahadevaswamy, U. B., 117
 Mahapatra, Sakambhari, 195
 Malik, Varun, 363, 379
 Mallidu, Jayashree, 171
 Mehta, Yagnik, 289
 Mishra, Deepak, 249
 Mishra, S. N., 493
 Mitra, S., 419
 Mittal, Ruchi, 363, 379
 Mittal, Shipra, 15
 Mohanty, Manisha, 281
 Mohanty, M. N., 1
 Mohapatra, Asha Rani, 339
 Mohapatra, Debaniranjan, 31
 Mohapatra, Meryleen, 329
 Mohapatra, Sraddhanjali, 475
 Munda, Archana Kumari, 127
 Mythili, A., 187

N

- Nadgouda, Shreya, 215
 Nair, Deepa, 559
 Nair, Gana C., 249
 Nanda, Sarita, 509, 541, 549
 Narshetty, Pooja, 149
 Nayak, D., 163
 Nibhande, Radhika, 109
 Niranjana Krupa, B., 233

P

- Pahadsingh, Sasmita, 225, 525
 Palai, G., 1
 Pal, Rajarshi, 407
 Panda, Jyoti Ranjan, 565
 Panda, Soumya Priyadarsini, 517
 Pandey, R., 419
 Pandey, Vedika, 127
 Pansare, Safal Deepak, 273
 Patel, Nisarg, 289
 Pathak, Bageshree, 15
 Patnaik, AnsUMAN, 509
 Patnaik, Suprava, 427
 Patnaik, Geeta, 329
 Pavan Nayak, R., 117
 Pawar, Pranjali, 15
 Prabhuram, Malavika, 241
 Prajapati, A., 419

Prasad, D. J. V., 435

Prasad, Rajendra, 407

Premalatha, J., 53

Primya, T., 89

R

Raj, Abhishek, 427

Rajapriya, R., 559

Rajesh Kumar, Burra, 205

Rajesh, P. Srinath, 47

Rajeswari, K., 559

Ramya Sree, Y., 435

Ranjan, Alka, 517

Rath, Kali Charan, 47

Rathod, Jagdish, 289

Rattan, Vikas, 379

Raw, Ram Shringar, 371

Ray, Arun Kumar, 347

Ray, K. B., 163

Reddy, Mekala Harinath, 53

Rosalin, 355

Rout, Deepak Kumar, 339

Rout, Nirmal Kumar, 135, 355, 501

Routray, Sidheswar, 347

Roy, Soumik, 399

S

Sabut, Sukanta, 41, 265

Sahoo, Rudra Swarup, 7

Sahoo, Swetaleena, 549

Sahu, Sudhakar, 157

Samal, Soumya Ranjan, 1

Samantasinghar, Umakanta, 7

Sasipriya, P., 65

Sathish, B. S., 303

Satyaprasad, G. R. K. D., 47

Selva Kumar, S., 443

Sen, Amit Prakash, 135

Sethi, Kabiraj, 321

Shanbhag, Nagaratna, 481

Sharma, Abhay, 53

Sharvani, N. R., 467

Shaw, Ankit Kumar, 31

Sheela, D., 53

Shinde, Akshata, 109

Shinde, Komal, 15

Shirol, Suhas B., 451

Shivalal Patro, B., 313

Shraddha, B. H., 481

Shriram, Revati, 109

Shrivastava, V. K., 419

Shruti, 525

Shukla, Anant, 241

Siddamal, Saroja V., 171, 451, 481

Singh, A. K., 419

Singh, Jaiteg, 363

Singh, Supriya, 225

Singh, Tapas Chhual, 47

Singh, Vedant Raj, 339

Sravan Kumar, K., 435

Srinivasu, N., 459

Subashini, G., 89

Sudharani, Bh., 435

SuhaaS Kiran, D. G., 273

Suresh, Vasagiri, 205

Swain, K. P., 1

Swain, Tanushree, 427

Swapneel, 273

Swati, 225

V

Vani, H. Y., 149

Vankara Jyothi, 205

Vashi, Ronak, 289

Veeramuthu, A., 303

Vrinda, S. M., 215

Y

Yamuna, B., 249