

Shubhakar Kalya
Muralidhar Kulkarni
K. S. Shivaprakasha *Editors*

Advances in VLSI, Signal Processing, Power Electronics, IoT, Communication and Embedded Systems

Select Proceedings of VSPICE 2020

Lecture Notes in Electrical Engineering

Volume 752

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>

Shubhakar Kalya · Muralidhar Kulkarni ·
K. S. Shivaprakasha
Editors

Advances in VLSI, Signal Processing, Power Electronics, IoT, Communication and Embedded Systems

Select Proceedings of VSPICE 2020

 Springer

Editors

Shubhakar Kalya
Engineering Product Development
Singapore University of Technology
and Design
Singapore, Singapore

Muralidhar Kulkarni
Department of Electronics
and Communication Engineering
National Institute of Technology Karnataka
Surathkal, Karnataka, India

K. S. Shivaprakasha
Department of Electronics
and Communication Engineering
N.M.A.M. Institute of Technology
Nitte, Udupi, Karnataka, India

ISSN 1876-1100

ISSN 1876-1119 (electronic)

Lecture Notes in Electrical Engineering

ISBN 978-981-16-0442-3

ISBN 978-981-16-0443-0 (eBook)

<https://doi.org/10.1007/978-981-16-0443-0>

© 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

Contents

Circularly Polarized E-Shaped Patch Antenna for AWS, FMS and MSS Applications	1
Pooja Pandey and Aditya Chinchole	
Bayes' Classifier for Mapping Intermediate and Heterogeneous RS Imagery	9
B. R. Shivakumar and B. G. Nagaraja	
A Study on the Effect of Dimensionality Reduction on Classification Accuracy of Myoelectric Control Systems	25
Praahas Amin and Airani Mohammad Khan	
An Efficient Low Power MEMS-Based Microfluidic Device for the Segregation of Different Blood Components	39
Ranjith B. Gowda, P. Vanishree, and Preeta Sharan	
Design of a Bipedal Robot	55
Chandana N. Aithal, P. Ishwarya, S. Sneha, C. N. Yashvardhan, and K. V. Suresh	
Amalgamation of Neural Network and Genetic Algorithm for Efficient Workload Prediction in Data Center	69
R. Pushpalatha and B. Ramesh	
Pressure Sensor Based on Two-Dimensional Photonic-Crystal Ring Resonator	85
D. L. Lakshmi, Venkateswara Rao Kolli, P. C. Srikanth, D. L. Girijamba, and Indira Bahaddur	
Enhancement of Performance of Round-Trip Time Using Kalman Filtering	99
N. G. Goudru, R. P. Puneeth, and Krishna Prasad N. Rao	
Automatic Modulation Classification Using Cumulants and Ensemble Classifiers	109
M. Venkata Subbarao and P. Samundiswary	

Automation Testing and Validation of Electric Drive System	121
Raqheeba Taneem and Krishnananda Shet	
Dimensionality Reduction Using Principal Component Analysis for Lecture Attendance Management System	135
Ramaprasad Poojary, Mariyam Milofa, and K. Shruthi	
Automatic Modulation Recognition Using Machine Learning Techniques: A Review	145
N. Venkateswara Rao and B. T. Krishna	
Performance Analysis of TCP in Presence of Nonresponsive Traffic in Wireless Networks	155
N. G. Goudru	
Microstrip Patch Antenna Analysis for 5G Millimeter-Wave Communication: A Survey	169
H. V. Pallavi, A. P. Jagadeesh Chandra, and Paramesha	
Kannada Text-to-Speech System using MATLAB	187
K. Sanjana Kamath, K. Raghavendra N. Bhat, Charishma, and Pearl Infancia D'souza	
Design of Dynamic Induction Charging Vehicle for Glimpse of Future: Cutting Down the Need for High-Capacity Batteries and Charging Stations	197
K. Balakrishna and N. G. Sandesh	
Generation of ECG for Heart Block Cases	205
Venkatesh Nayak	
Iterative Thresholding-Based Spectral Subtraction Algorithm for Speech Enhancement	221
Raj Kumar, Manoj Tripathy, and R. S. Anand	
Implementation of Cryptographic Algorithm (One-Time-Pad) with a RISC-V Processor	233
Priyanka Ashok Kurkuri, Saroja V. Siddamal, and Rashmi Kubsad	
Implementation of a Hebbian Learning Algorithm as an Accelerator to a RISC-V-Based Processor	241
Rashmi Kubsad, Saroja V. Siddamal, and Priyanka Ashok Kurkuri	
Design and Implementation of Agricultural Drone for Areca Nut Farms	251
Raju Hajare, C. P. Mallikarjuna Gowda, and M. V. Sanjaya	
Underwater Marine Life and Plastic Waste Detection Using Deep Learning and Raspberry Pi	263
Rahul Hegde, Sanobar Patel, Rosha G. Naik, Sagar N. Nayak, K. S. Shivaprakasha, and Rekha Bhandarkar	

Land Cover Mapping Capability of Chaincluster, K-Means, and ISODATA techniques—A Case Study 273
 Karthik and B. R. Shivakumar

Microcontroller-Based Control Circuit for the Automatic Orientation of Solar Panels in the Direction of Sun 289
 A. Jayashree, Vidya Kudva, and A. G. Ananth

Development of an Automated Plant Classification System Using Deep Learning Approach 303
 K. Ananth Pai, B. R. Apoorva, Daisy Sheetal Mendonca, Durgaprasad S. Hegde, and Roopa B. Hegde

Design and Implementation of Multi-class Logistic Regression for Effective Classification of Low, Medium and High Risk Lung Cancer Problem 317
 Shivaprasad, P. Mahabaleshwara Bhat, and C. Naveena

Voice Controlled IoT Based Grass Cutter Powered by Solar Energy 327
 Mahadevaswamy, Kiran Kumar Humse, K. Chethan, and K. V. Sudheesh

Human Body Measurement Extraction from 2D Images 343
 Sachin S. Bhat, Alaka Ananth, Preema Dsouza, K. Sharanyalaxmi, Shreeraksha, and Tejasvini

Generation of ECG Arrhythmias Using Fourier Analysis 353
 Paresb Praveen, B. Samartha, Vaibhav R. Pai, M. Namith Rao, K. S. Shivaprasad, and Rekha Bhandarkar

A Survey on Methodologies and Database Used for Facial Emotion Recognition 367
 Shashank M. Gowda and H. N. Suresh

Intensity-Based Feature Extraction of Real-Time Transformer Oil Images 379
 C. M. Maheshan and H. Prasanna Kumar

A Comparative Analysis of Statistical Model and Spectral Subtractive Speech Enhancement Algorithms 397
 K. Rohith and Rekha Bhandarkar

Design of Electronic Instrumentation for Isotope Processing 417
 V. Arunprasad

Visual Cryptography Using Hill Cipher and Advanced Hill Cipher Techniques 429
 Jagadeesh Basavaiah, Audre Arlene Anthony, and Chandrashekar Mohan Patil

Analysis of Power Adaptation Techniques Over Beaulieu-Xie Fading Model 445
 Veenu Kansal and Simranjit Singh

About the Editors

Dr. Shubhakar Kalya obtained his Master's degree in Microelectronics from Indian Institute of Science (IISc), Bengaluru, India in 2007 and Ph.D. degree in the field of Nanoelectronics from School of Electrical and Electronics Engineering, Nanyang Technological University (NTU), Singapore in 2012–13. He obtained his Bachelor's Degree in Electronics and Communication Engineering from N.M.A.M. Institute of Technology, Nitte. He is working as a faculty in Engineering Product Development (EPD) Pillar at Singapore University of Technology and Design (SUTD), Singapore, since January 2012. His research work focuses on nanoscale characterization of High- κ gate dielectrics and 2D materials for advanced logic and memory devices, and its reliability analysis. From July 2009 to December 2012, he has worked as a visiting researcher at Institute of Materials and Research Engineering (IMRE), Singapore, and involved in research related to characterization of high- κ gate dielectrics using scanning tunnelling microscopy and atomic force microscopy. During January–June 2017, he was at Massachusetts Institute of Technology (MIT), Cambridge, USA as visiting scientist. He has authored/co-authored more than 30 research publications (journals/ international conferences) and two book chapters. He is also involved in review of various technical papers including IEEE Transactions on Device and Materials reliability, Microelectronic Engineering journals and international conferences related to Nanoelectronics field.

Dr. Muralidhar Kulkarni received his B.E. (Electronics Engineering) degree from University Visvesvaraya College of Engineering, Bangalore University, Bengaluru; M.Tech. (Satellite Communication and Remote Sensing) degree from Indian Institute of Technology, Kharagpur (IIT KGP); and Ph.D. from JMI Central University, New Delhi, in the area of optical communication networks. He has served in various capacities in industry and academics for the last 35 years. He has held the positions of Scientist in Instrumentation Division at the Central Power Research Institute, Bengaluru; Aeronautical Engineer in Avionics group of Design and Development team of Advanced Light Helicopter(ALH) project at Helicopter Design Bureau, Hindustan Aeronautics Limited(HAL), Bengaluru; Lecturer at the Electrical Engineering Department, University Visvesvaraya College of Engineering, Bengaluru; and as an Assistant Professor in Electronics and Communication Engineering (ECE)

Department, Delhi College of Engineering (DCE), Delhi. He has served as Head, Department of Information Technology and Head, Computer Center at DCE, Delhi. Currently, he is a Professor in the Department of Electronics and Communication Engineering (ECE), National Institute of Technology Karnataka (NITK), Surathkal, Karnataka. His teaching and research interests are in the areas of digital communications, information theory and coding, fuzzy digital image processing, optical communication and networks, and OFDM/COFDM techniques. He has published more than 50 papers in reputed international/national journals and conferences. He has also authored five books in the areas of communication.

Dr. K. S. Shivaprakasha received his B.E. (Electronics & Communication Engineering) degree from Bahubali College of Engineering, Visvesvaraya Technological University, Karnataka with IX rank in the University and M.Tech. (Digital Electronics and Communication Systems) degree from Malnad College of Engineering, Visvesvaraya Technological University, Karnataka, with I rank with Gold Medal in the University in 2004 and 2007, respectively. He completed his Ph.D. from National Institute of Technology Karnataka (NITK), Surathkal, Karnataka, in the field of wireless sensor networks in 2015. Currently, he is a Professor in the Department of Electronics and Communication Engineering, N. M. A. M. Institute of Technology, Nitte, Karnataka. His areas of research interest include wireless sensor networks, mobile adhoc networks, information coding theory and cryptography. He has published more than 30 papers in reputed international/national journals and conferences and has co-authored a book on “Information Theory and Coding” for Wiley (India) publications.

Circularly Polarized E-Shaped Patch Antenna for AWS, FMS and MSS Applications



Pooja Pandey and Aditya Chinchole

1 Introduction

Circularly polarized antennas (CP) have high demand due to various advantages like great speed operations, very high tolerance to bad weather conditions and multipath fading resistance. But some designs have large size, very narrow bandwidths and have high complexity in design and fabrication [1].

Microstrip antennas have numerous applications in these days because of the advantages like lightweight, compact size, easiness in integration and reproduction with the circuitry [2, 3]. There is an increasing growth in wireless RF system demand in applications like LAN, point-to-point communications and applications such as AWS, FMS and MSS [4]. Structure of patch is modified to have application specific resonating frequencies and to have higher gain and bandwidth response [5, 6]. Patch geometries such as rectangular, circular and triangular are the most analysed antenna geometries.

Advanced Wireless Service (AWS) is a wireless spectrum band used for mobile voice and data service, messaging and video. Most of the American countries use AWS which includes USA, Canada, Mexico, Chile, Peru and others [7]. AWS has replaced Multipoint Multichannel Distribution System (MMDS) bands. AWS have many versions or blocks like AWS-1, AWS-2, AWS-3 and AWS-4; in this design, AWS-3 and AWS-4 blocks are workable as it covers the required band range of 2180–2200 MHz.

Microwaves are short waves in the upper range of the radio spectrum used in point-to-point communication systems [8]. FMS is nothing but Fixed Microwave Service used in radio communication. Its frequency band ranges from 2160 to 2200 MHz.

P. Pandey · A. Chinchole (✉)
E.T Department, YCCE, Nagpur, India
e-mail: adichin1998@gmail.com

P. Pandey
e-mail: pj17031998@gmail.com

Mobile Satellite Service (MSS) is a radiocommunication service between mobile earth station and space stations [9]. Major LEO satellites operate on the basis of this service.

In this article, E-shaped patch antenna is further modified and analysed to achieve these operations like AWS, FMS and MSS.

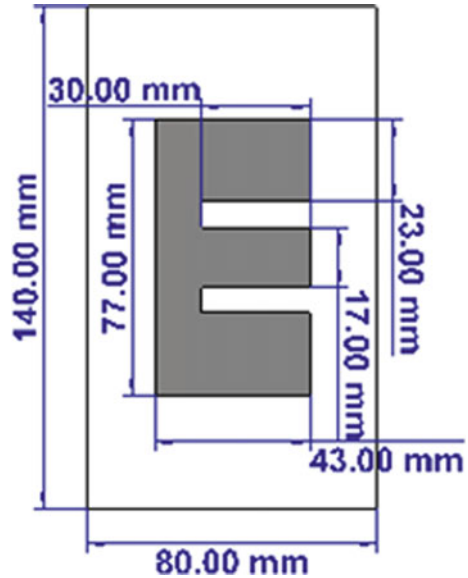
2 Antenna Design

The proposed E-shaped antenna design is depicted in Fig. 1. Substrate used is RT 5880 with dimension of 80×140 mm and thickness of 1.6 mm. The antenna attains circular polarization as Ludwig 3 left and Ludwig 3 right have equal values which is depicted in Fig. 2.

From the above Fig. 1, the E-shaped antenna consists of perfect electric conductor (PEC) as patch and ground and RT duroid 5880 as substrate. Substrate has dimension of 80×140 mm² and thickness of 1.6 mm; on it, a patch of dimension 43×77 mm² is placed. Then, two slots are created of length 13 mm each.

Ludwig 3 left and Ludwig 3 right have same value which proves that circular polarization is achieved.

Fig. 1 E-shaped patch antenna



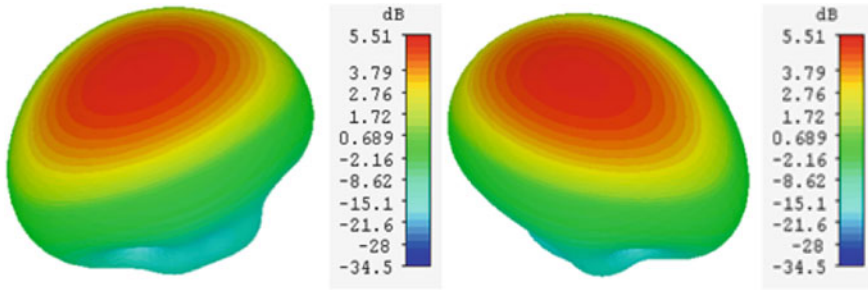


Fig. 2 Ludwig 3 left and Ludwig 3 right

3 Results and Discussion

The proposed antenna was designed on Computer Simulation Technology (CST) environment. CST Studio Suite is simulated electromagnetic tool made by Dassault Systems. It includes many distinct simulation methods like the finite element method (FEM), finite integration technique (FIT), transmission line matrix (TLM), etc.

A. 1D Results

Firstly, S_{11} results or return loss vs frequency curve are depicted in Fig. 3, in which the frequency range of 2.18–2.223 GHz is observed. Moreover, the resonance is at 2.2 GHz with value of 25.5. This bandwidth is useful in many aspects, not only in general communication purpose but also in AWS, FMS and MSS applications.

VSWR is the voltage standing wave ratio which is ideally 1 and can extend up to 2 as per the maximum power transfer theorem. In Fig. 4, VSWR results of the proposed antenna at resonance frequency of 2.2 GHz are depicted. We can observe that the value of around 1.11 has been achieved.

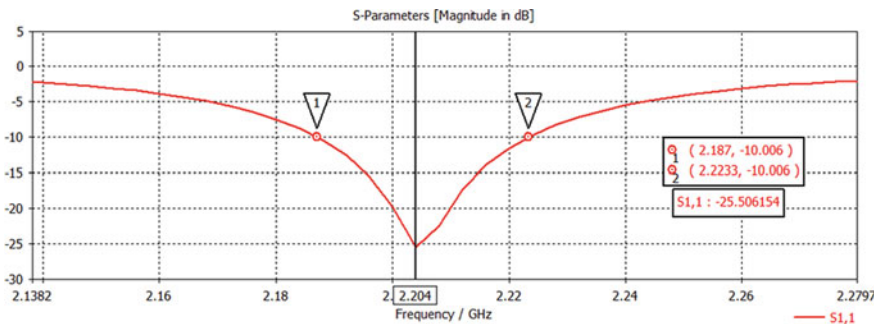


Fig. 3 S-parameter results

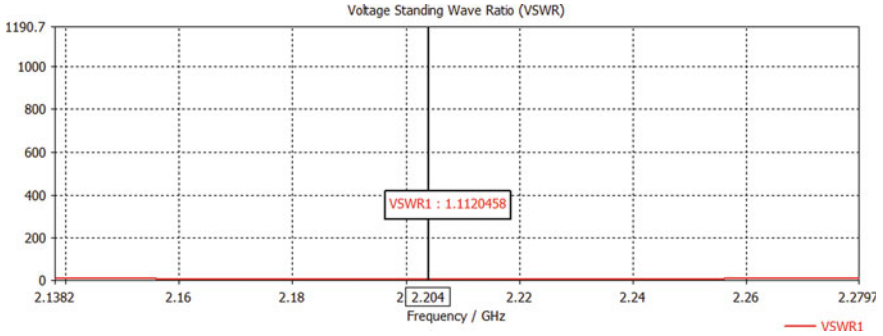


Fig. 4 VSWR at 2.2 GHz frequency

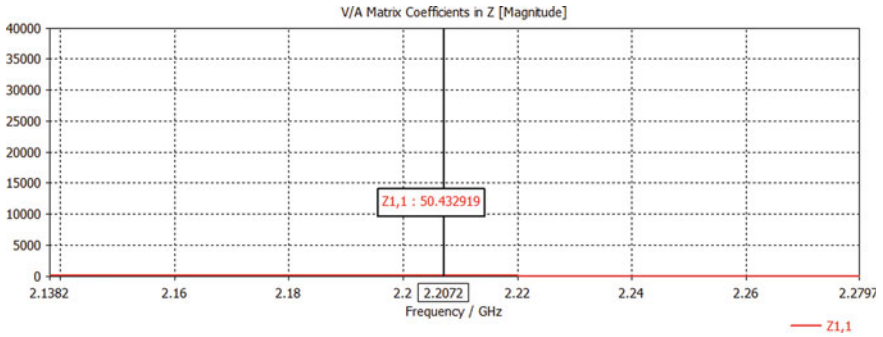


Fig. 5 Impedance measurement

Impedance measurement is shown in Fig. 5, which has impedance matching of around 50Ω , responsible for the maximum power transfer and less signal reflection from load.

Surface current simulation result at resonance frequency of 2.2 GHz is shown in Fig. 6. If we minutely observe the surface current pattern, we can see that the current density is maximum around edges of E-shaped patch. This phenomenon and the observed direction of current waves are related to the circular polarization.

These were the basic 1D results of the proposed antenna design. In the next section, all the 2D and 3D results like radiation pattern, gain and directivity are discussed.

B. 2D-3D results

In the first place, we will talk about gain and directivity results at resonant frequency of 2.2 GHz. As shown in Fig. 7, maximum directivity of 8.68 dBi and gain of 8.52 dB are obtained. From the 3D view of the directivity, we can say that the proposed E-shaped antenna is highly directional and gain of the antenna is also high which leads to the high-power transmission in specific direction and therefore strengthens the transmitted and received signals.

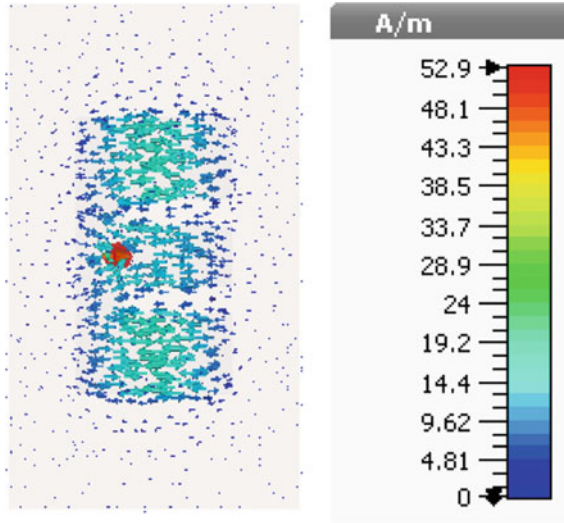


Fig. 6 Surface current simulation

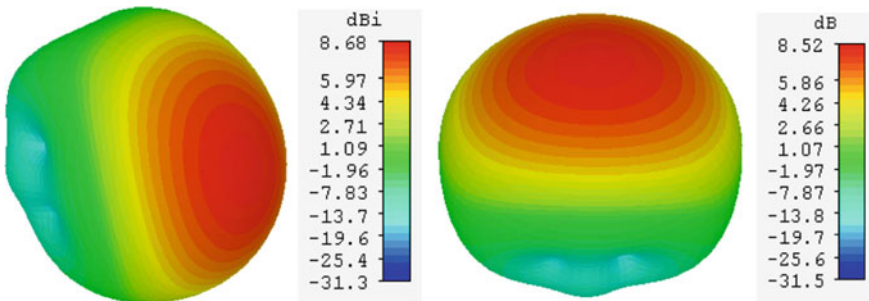


Fig. 7 Directivity and gain at 2.2 GHz

Taking radiation pattern into consideration, we can say that the antenna proposed is omnidirectional with a heart shape pattern shown in Fig. 8. Both constant phi and constant theta radiation patterns are displayed. In Fig. 9, radiation pattern of gain at constant phi and constant theta is portrayed.

4 Conclusion

After the analysis of the proposed E-shaped patch antenna, we can say that it is a highly directional antenna with directivity of 8.68 dBi and also a high gain antenna with gain of 8.52 dB. The proposed antenna also achieves circular polarization and has

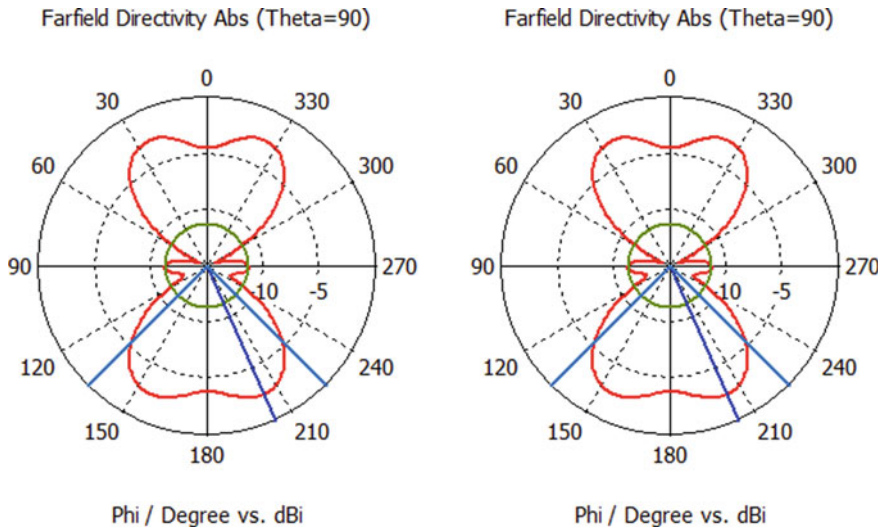


Fig. 8 Directivity radiation patterns

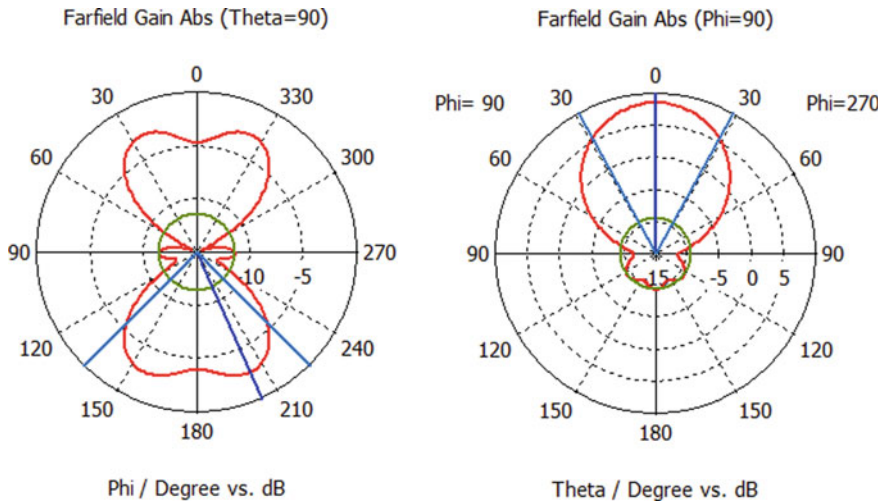


Fig. 9 Gain radiation patterns

efficiency of around 98%. It covers upper L-band frequency range of 2.18–2.2 GHz and has applications like AWS-3, AWS-4, FMS and MSS. Further, E-shaped antennas have higher bandwidth than any regular rectangular patch antenna, so this proposed antenna is light, cheap and easy to assemble than other rectangular antenna for these applications.

References

1. O.G. Kwame, Y. Huang, G. Wen, J. Li, A.E. Ampoma, Broadband circularly polarized square slot antenna with a G-shaped feedline, in *2017 IEEE International Symposium on Antennas and Propagation & USNC/URSI National Radio Science Meeting*, San Diego, CA, pp. 117–118 (2017). 10.1109/APUSNCURSINRSM.2017.8072101
2. K. Carver, J. Mink, Microstrip antenna technology. *IEEE Trans. Antenn. Propag.* **29**(1), 2–4 (1981)
3. R. Garg, P. Bhartia, I.J. Bahl, A. Ittipiboon, *Microstrip Antenna Design Handbook* (Artech House, New York, 2001)
4. S.S. Patel, H.B. Soni, Y.P. Kosta, S.K. Patel, E-shaped patch antenna analysis for multiple applications, in *3rd International Conference on Electronics Computer Technology* (2011)
5. V.K. Tiwari, A. Kimothi, D. Bhatnagar, J.S. Saini, V.K. Saxena, Theoretical and experimental investigation of circular sector microstrip antenna. *Indian J. Radio Space Phys.* **35**, 206–211 (2006)
6. K.L. Wong, W.S. Hsu, Broadband triangular microstrip antenna with V-shaped slot. *Electron. Lett.* **33**(25), 2085–2087 (1997)
7. Advanced Wireless Services page at FCC. *Advanced Wireless Service (AWS)*. About the Federal Communications Commission. FCC.gov. Retrieved 10 Dec 2017
8. <https://www.fcc.gov/wireless/bureau-divisions/broadband-division/point-point-microwave>
9. https://en.wikipedia.org/wiki/Mobile-satellite_service

Bayes' Classifier for Mapping Intermediate and Heterogeneous RS Imagery



B. R. Shivakumar  and B. G. Nagaraja 

1 Introduction

The success of a large number of socio-economic applications such as change detection, creating digital maps, crop analysis and so on depends heavily on the precision with which LULC maps are generated. Therefore, a great deal of research has been conducted on developing algorithms for extracting LULC information from RS imagery. This has resulted in the production of a large variety of classification algorithms ranging from object-based [2, 3], pixel-based [4, 5], sub-pixel-based [6, 7] and scene-based techniques [8, 9]. However, no classification technique has been found to be optimal till date. Also, variation in the complexity of the RS data makes the analysis of classification technique's performance more interesting.

RS data, on the basis of data complexity, can be categorized into three types: homogeneous, intermediate and heterogeneous data. Homogeneous data are characterized by a clear separation between its LULC classes in signal space. Intermediate data exhibits limited amount of overlapping of LULC classes, whereas heterogeneous data sets exhibit large overlapping of the spectral classes making the decision making process during pixel assignment even more complex. When more than two classes overlap spatially, the region covering the overlap constitutes of pixels that have characteristics derived from both the overlapping classes. These pixels are referred to as mixed pixels. When large heterogeneous RS data are considered for the study, mixed pixels are a common finding on the data. Sometimes, a data may consist of more than 50% of the pixels as mixed pixels. In such scenarios, as literature [10, 11]

B. R. Shivakumar (✉)

Department of ECE, NMAM Institute of Technology, Nitte (Visvesvaraya Technological University, Belagavi), Udupi, Karnataka, India
e-mail: shivkumarbr@nitte.edu.in

B. G. Nagaraja

Department of ECE, Jain Institute of Technology, Davanagere, India
e-mail: nagarajbg@gmail.com

has already concluded, hard classifiers are not an optimal choice. This is due to the inherent assumption made by the hard classifiers which is ‘a pixel can only belong to one class’. This leads to the use of soft classifiers.

Soft classifiers, also called fuzzy classifiers, are inherently built on the principle that a pixel can belong to more than one class. This assumption fully defines the mixed pixel scenario. Soft classifiers are constructed using the principle of fuzzy set developed by Zadeh [12, 13]. Since then, a large amount of work has been presented in developing fuzzy-based classifiers [14–16]. Bayes’ classifiers are based on the Bayes’ theorem and are categorized under the probabilistic classifiers. In the literature, they are also referred to as simple Bayes’ and independence Bayes’ [17]. They assume that there exists strong independence between the class features [18]. Bayes’ classifiers are soft supervised classifiers and require training data for performing classification. They are simple models that assign a pixel to a class label that is defined by the analyst. The basic idea behind Bayes’ classifiers is that they assume each feature vector to be completely independent of any other feature vector, given the class variable. Practically, Bayes’ classifiers use a maximum likelihood approach for the estimation of class labels for any feature vector.

The main objective of this study is to characterize the performance of Bayes’ classifier under the severe overlapping of classes on the RS data. This may provide more insight on the principle of working of the Bayes’ classifier as it is tested for two study areas of different characteristics. The classifier is also tested for different input conditions as well as input training sizes. Both the study areas considered exhibit variable levels of LULC class overlapping. Since Bayes’ is a soft classifier, it is interesting to observe its performance under severe class overlapping scenario.

The rest of the paper is structured as follows. In Sect. 2, the RS data considered for the research is assessed and described. The mathematical basis of the Bayes’ classifier is presented in Sect. 3. Section 4 presents and discusses the results obtained during the conduction of the study. Conclusions drawn from the results are discussed in Sect. 5.

2 Study Areas

This study uses two Landsat 8 operational land imager (OLI)/ thermal infrared sensor (TIRS) images. The primary study area encompasses the North Canara District of Karnataka State, India, and is shown in Fig. 1. The secondary study area encompasses the Kumta Taluk, which resides within the primary study area, and is shown in Fig. 2. Both study areas exhibit similar LULC classes with highly similar spectral characteristics. Hence, LULC classes identified on the primary study area are also derived for the secondary study area for uniform evaluation of the results.

Both study areas are characterized by large spatial coverage of thick forests, which are a subset of the famous Western Ghats of the Indian subcontinent. This forest region is further separated as evergreen and deciduous forest classes. Scrub jungles are also profoundly found in large patches throughout the study areas. The

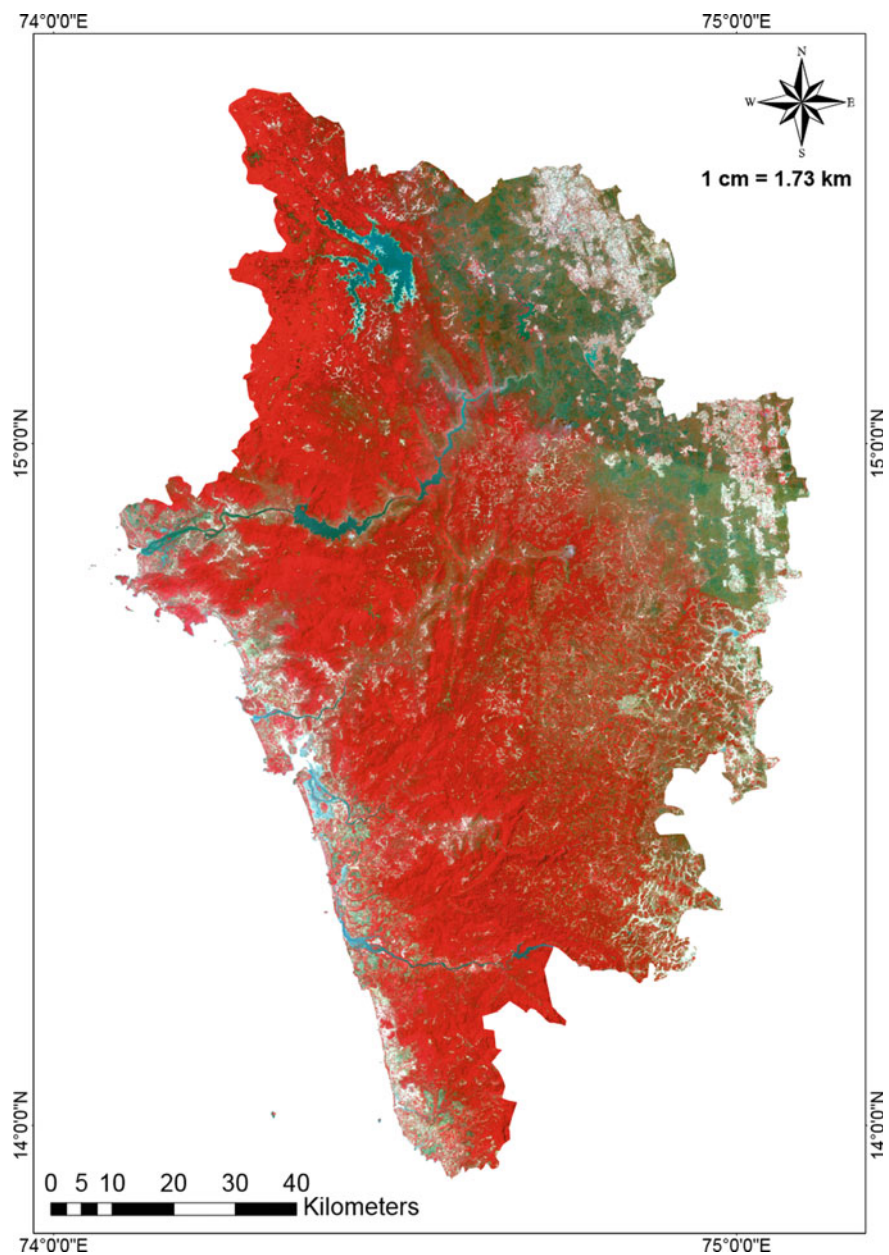


Fig. 1 Map of the primary study area (band combination [5 4 3]). Data courtesy U.S. Geological Survey EarthExplorer [19]

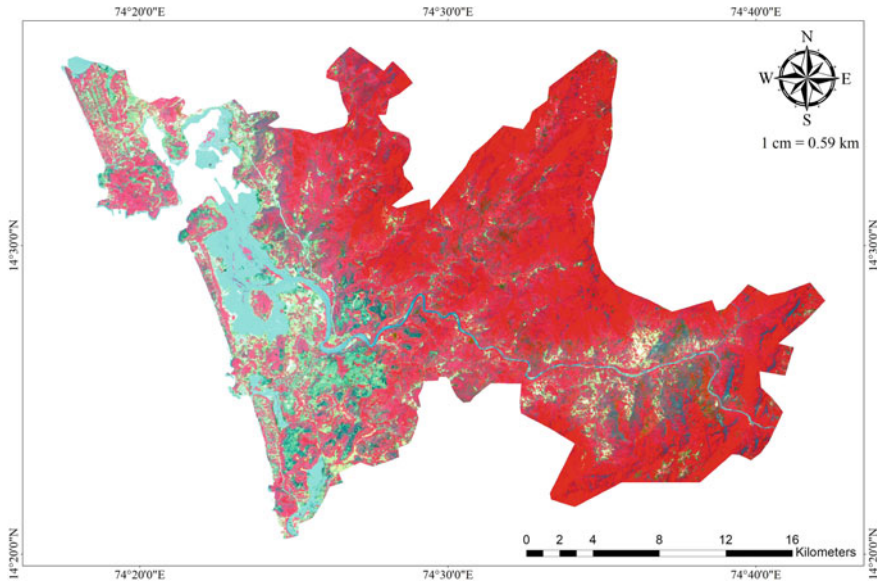


Fig. 2 Map of the secondary study area (band combination [5 4 3]). Data courtesy U.S. Geological Survey EarthExplorer [19]

agricultural fields are categorized as kharif and double crop classes. Kharif are those crops that are bound by the monsoon season. Double Crop includes fields that are harvested twice a year including horticultural plantations. The coastal side of the study areas is very commonly characterized by these horticultural plantations. All the water bodies in the study areas are lumped together to form a single class named water body. Lastly, the urban settlements form the built-up class. Since the study areas are largely covered by thick forests, the spatial proportions of the urban settlement are considerably small in both study areas.

By using an on-screen collection of training data through points, polygons and lines, a total of seven LULC classes were identified on both the study areas. They are evergreen forest (EGF), deciduous forest (DF), water body (WB), kharif (KH), scrub land (SL), double crop (DC) and built-up (BU). These classes are derivatives of Anderson’s LULC system’s level-1 and level-2 classes [1]. The information on signature set size for both study areas is shown in Table 1. It is worth noting that the training data size for primary study area is considerably larger than that for the

Table 1 Signature set information for both study areas

Study area	Training size						
	EGF	DF	SL	WB	KH	DC	BU
Primary	28,192	35,906	25,310	14,927	14,339	6652	2714
Secondary	2059	450	1571	3413	730	561	136

Table 2 Class separability for severely overlapping classes

North Canara			Kumta		
Class pair	JM	TD	Class pair	JM	TD
DF-SL	1.29381150	1.62976945	EGF-DCP	1.46683581	1.95754919
BU-KH	1.45675134	1.60031081	KH-BU	1.90907271	2.00000000
EGF-DCP	1.47137640	1.64988537	SL-DF	1.98816550	2.00000000
KH-SL	1.73930315	1.87348605	SL-BU	1.99885822	2.00000000
BU-SL	1.91547945	1.97042385	EGF-DF	1.99982318	1.99999994
KH-DF	1.94612265	1.99999037	SL-KH	1.99994694	2.00000000
EGF-DF	1.98664047	1.99999998	WB-BU	1.99999962	2.00000000
EGF-SL	1.99332146	1.99998332	EGF-SL	1.99999991	2.00000000

secondary study area. These values were attained by rigorous and repeated testing of the classifier over the study areas. Further, to understand the spectral relatedness between each class pair, class separability between each class pair was measured in terms of Jeffries–Matusita (JM) and transformed divergence (TD) metrics. The most severely overlapping classes on both study areas were identified and are shown in Table 2.

3 Bayes' Classifier

The primary motivation for the use of the Bayesian classifier is the sub-pixel classification of the data. By using the Bayesian classifier, one can estimate the proportions of mixed pixels in the data as well as their relative proportions. In the following subsections, the probabilistic model and different event models used in concurrence with Bayes' classifier are discussed.

A. Probabilistic Model

Naive Bayesian classifier is a conditional probability model. Consider a remotely sensed image consisting of n spectral bands. Each feature vector a of the data is then represented as $a = (a_1, \dots, a_n)$. Bayesian classifier assigns a probability to this feature vector as $p(C_k/a_1, \dots, a_n)$ for each of the K spectral classes on the data [20]. This step is similar to the fuzzification step in fuzzy supervised classification [21]. Probabilities are assigned to the feature vectors based on the Bayesian theorem given by [20],

$$p(C_k/a) = \frac{p(C_k)p(a/C_k)}{p(a)} \quad (1)$$

where $p(C_k/a)$ is the posterior probability of assigning pixel vector a to class label C_k , $p(C_k)$ is the prior probability value of class C_k , $p(a/C_k)$ is the likelihood function

of pixel a belonging to class C_k and $p(a)$ is the pixel evidence given by [20],

$$p(a) = \sum_{i=1}^k p(C_k) p(a/C_k) \quad (2)$$

In Eq. (1), the numerator is the main component that decides the classification labelling. Since the values of the feature vector a_i are known, the denominator is effectively constant.

The numerator of Eq. (1) is the joint probability given as [20],

$$\begin{aligned} p(C_k)p(a/C_k) &= P(C_k, a) = P(C_k, a_1, \dots, a_n) \\ &= p(a_1/a_2, \dots, a_n, C_k)p(a_2, \dots, a_n, C_k) \\ &= p(a_1/a_2, \dots, a_n, C_k)p(a_2/a_3, \dots, a_n, C_k)p(a_3, \dots, a_n, C_k) \dots \end{aligned} \quad (3)$$

$$= p(a_1/a_2, \dots, a_n, C_k), \dots, p(a_{n-1}/a_n, C_k)p(a_n/C_k)p(C_k) \quad (4)$$

Under the assumption that all features of a are mutually independent over the class C_k , the likelihood is given by [20],

$$p(a_i/a_{i+1}, \dots, a_n, C_k) = p(a_i/C_k) \quad (5)$$

Thus, the joint probability defined by Eq. (4) can also be expressed as [20],

$$\begin{aligned} p(C_k/a_1, \dots, a_n) &\propto p(C_k, a_1, \dots, a_n) \\ &= p(C_k)p(a_1/C_k)p(a_2/C_k)p(a_3/C_k) \dots \\ &= p(C_k) \prod_{i=1}^n p(a_i/C_k) \end{aligned} \quad (6)$$

Hence, the conditional distribution of the class variable C_k is [20],

$$p(C_k, a_1, \dots, a_n) = \frac{p(C_k) \prod_{i=1}^n p(a_i/C_k)}{\sum_{i=1}^K p(C_k) p(a_i/C_k)} \quad (7)$$

Using the probability model so defined, Bayesian classifier combines it with a decision rule. The most commonly used decision rule is the maximum a posteriori (MAP) decision rule, which assigns a feature vector a to class label $\hat{b} = C_k$, for some k using the equation [20],

$$\hat{b} = \underbrace{\operatorname{argmax}}_{k \in (1, \dots, K)} p(C_k) \prod_{i=1}^n p(a_i/C_k) \quad (8)$$

B. Parameter Estimation and Event Models

The *a priori* class probabilities can be determined in two ways: i. assuming the classes are equiprobable, and ii. estimating *a priori* class probabilities from the training data set. Estimating parameters for feature distribution require an assumption to be made on the type of data distribution itself or build nonparametric models [22]. The assumption made on the type of data distribution leads to event models of the Bayesian classifier. Depending on the type of input data, there are several distributions. For the continuous type of data, the most commonly occurring distribution is the normal distribution. This leads to the Gaussian Bayes' event model. For discrete data types, the two most popular distributions are multinomial and Bernoulli distributions. These lead to two distinct event models: multinomial Bayes' and Bernoulli Bayes' [23].

i. Gaussian Bayes' Event Model

For continuous data, this type of event model is suitable with the typical assumption that the data related to every class are Gaussian distributed. For example, consider a continuous attribute, a , of the training data. The data will be segmented according to the class, and the mean and variation of the attribute a for each class are calculated. Let μ_k and σ_k^2 denote the mean and Bessel corrected variance of the attribute a for class C_k . Suppose there exists an observation value v . The probability distribution of v for class C_k denoted as $p(a = v/C_k)$ is then given by [17],

$$p(a = v/C_k) = \frac{1}{\sqrt{2\pi\sigma_k^2}} \exp\left[-\frac{1}{2\sigma_k^2}(v - \mu_k)^2\right] \quad (9)$$

ii. Multinomial Bayes' Event Model

In this type of event models, each feature vector represents the frequency of creation of specific events by a multinomial (p_1, \dots, p_n) , where p_i represents the probability that event i occurs. For the multiclass case, K such multinomials are defined, where K is the number of classes on the input data. A feature vector for this model given by $a = (a_1, \dots, a_n)$ represents a histogram where a_i represents the count on the number of times feature a_i was observed in a particular instance. This type of event model is commonly employed in document classification. The likelihood of observing a histogram a is then given by [17],

$$p(a/C_k) = \left(\frac{\sum_i a_i}{\prod_i a_i}\right) \prod_i p_{ki}^{a_i} \quad (10)$$

In this study, Bayesian classifier based on multinomial Bayes' event model is employed for the classification of the remotely sensed data.

iii. Bernoulli Bayes' Event Model

In this type of event model, features are assumed as independent binary variables describing inputs. Such an event model is known to be able to classify document classification tasks, similar to the multinomial case model [24], with the use of binary term occurrence functions, rather than term frequencies. If x_i is a Boolean that expresses the i^{th} term's presence or absence from the vocabulary, the probability of a document given class C_k is given as [24],

$$p(a/C_k) = \prod_{i=1}^n p_{ki}^{a_i} (1 - p_{ki})^{(1-a_i)} \quad (11)$$

where p_{ki} is the probability of class C_k generating the term a_i .

4 Results and Discussions

Both primary and secondary study areas were classified using Bayes' classifier with two input conditions: (i) with *a priori* class probabilities and (ii) without *a priori* class probabilities. The *a priori* class probabilities were extracted from the study areas with a confidence level of 95% with $\pm 4\%$ error margin [21]. The *a priori* class probabilities for the study areas calculated are indicated in Table 3. Table 4 indicates the spatially dominant and spatially subservient classes in each study areas. In the absence of *a priori* class probabilities, equal probability of occurrence was assumed for each class. As each study area has seven LULC classes defined over them, the probability that a class can be assigned to any of the LULC class is equal to $1/7 = 0.14286$. This probability value is used for equal probability case.

For the primary study area, Bayes' classifier produced an overall classification accuracy of 75.50% with equal *a priori* class probabilities. In this case, the classifier extracted water body class with the highest user's accuracy followed by kharif, evergreen and deciduous forest classes. All these classes were extracted with greater than 75% user's accuracy. While user's accuracy only implies the proportion of pixels

Table 3 *A priori* probabilities for the LULC classes

LULC class	North Canara	Kumta
Evergreen forest	0.532	0.590
Deciduous forest	0.192	0.030
Scrub land	0.089	0.100
Double crop	0.089	0.120
Kharif	0.069	0.050
Water body	0.015	0.080
Built-up	0.013	0.030

Table 4 Spatially dominant and subservient classes in the study areas

North Canara		Kumta	
Dominant	Subservient	Dominant	Subservient
Evergreen forest	Double crop	Evergreen forest	Built-up
Deciduous forest	Built-up	Deciduous forest	Double crop
Scrub land		Scrub land	
Kharif		Water body	
Water body		Kharif	

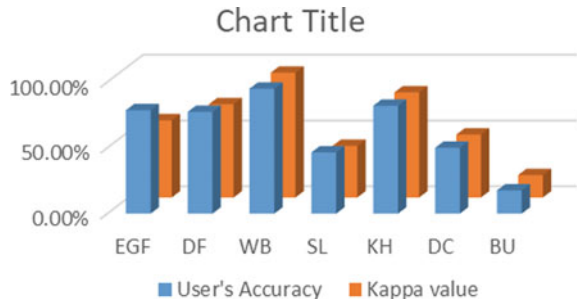
considered for accuracy assessment, using kappa value of the class will enable us to extend the results to the rest of the pixels on the data. In general, the kappa value of a class indicates how accurate is the classification for that class, when interpolated to all other pixels on the data [25]. For example, water body class has a kappa value of 0.9489 indicating that this proportion of all water body pixels on the data is correctly classified on the output map. The classifier further extracted the scrub land and double crop classes with intermediate accuracies, while built-up class was poorly extracted. The reason for such intermediate and poor accuracies for these classes may be derived from Table 2. The first three overlapping class pairs have these classes as their spatial subservient class. A spatial subservient class of a class pair is the one that has smaller spatial covering compared to its class pair. The Bayes' classifier, for this case, has favoured the dominant class of the overlapping class pairs while producing mediocre results for the subservient classes (except for built-up class). Table 5 indicates the accuracy assessment results for this case, and Fig. 3 indicates the variation of user's accuracy and class kappa values for this case.

Table 5 Accuracy assessment results for North Canara study area for equal class probabilities (case 1)

Class name	RT ^a	CT ^b	NC ^c	PA ^d (%)	UA ^e (%)	Kappa
Evergreen forest	529	663	521	98.49	78.58	0.5867
Deciduous forest	244	221	171	70.08	77.38	0.7091
Water body	24	20	19	79.17	95.00	0.9489
Scrub land	132	58	27	20.45	46.55	0.3925
Kharif	115	89	73	63.48	82.02	0.7992
Double crop	46	30	15	32.61	50.00	0.4781
Built-Up	8	17	3	37.50	17.65	0.1704
Totals	1098	1098	829			
Overall Classification Accuracy = 75.50%						
Overall Kappa value = 0.6218						

^aRT: reference totals, ^bCT: classified totals, ^cNC: number correct, ^dPA: producer's accuracy, ^eUA: users accuracy

Fig. 3 Variation of user’s accuracy and Kappa values for classes in case-1



In the presence of *a priori* class probabilities, the overall classification accuracy produced by Bayes’ classifier for North Canara study area is 74.95%. The overall classification accuracy decreased by 0.55%. The individual class performance was highly similar to that produced for the case in which equal *a priori* class probabilities were considered for classification. It is worth noting that the classification performance is not affected by the use of *a priori* class probabilities for the primary study area. The overall Kappa value recorded, for both the cases, when compared with the values provided in the literature [25], the classification performance is termed good. Table 6 indicates the accuracy assessment results for this case, and Fig. 4 indicates the variation of user’s accuracy and class kappa values for this case.

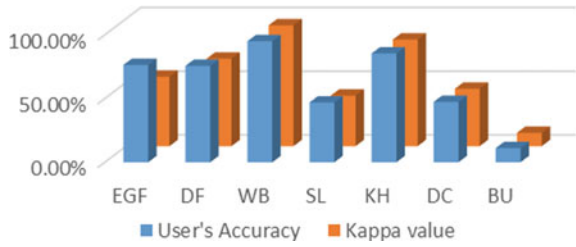
For the secondary study area, the overall classification accuracy when equal class probabilities were considered is 76.90%. The classifier once again extracted water body, evergreen forest, and kharif, with greater than 75% user’s accuracies. Scrub land class was extracted with intermediate user’s accuracy values. These four classes

Table 6 Accuracy assessment results for North Canara study area for *a priori* class probabilities (case 2)

Class name	RT ^a	CT ^b	NC ^c	PA ^d (%)	UA ^e (%)	Kappa
Evergreen forest	529	686	524	99.05	76.38	0.5443
Deciduous forest	244	226	171	70.08	75.66	0.6871
Water body	24	20	19	79.17	95.00	0.9489
Scrub land	132	49	23	17.42	46.94	0.3969
Kharif	115	89	76	66.09	85.39	0.8368
Double crop	46	19	9	19.57	47.37	0.4507
Built-up	8	9	1	12.50	11.11	0.1046
Totals	1098	1098	823			
Overall classification accuracy = 74.95%						
Overall Kappa value = 0.6076						

^aRT: reference totals, ^bCT: classified totals, ^cNC: number correct, ^dPA: producer’s accuracy, ^eUA: users accuracy

Fig. 4 Variation of user's accuracy and Kappa values for classes in case-2



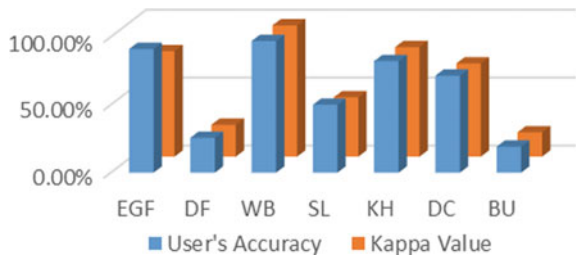
form the dominant classes on the secondary study area. It should be noted that evergreen forest and double crop form the most overlapping class pair on the secondary study area. Yet, the Bayes' classifier extracted the double crop class with an efficient user's accuracy of 71.43%. The other two subservient classes on the study area were extracted rather poorly by the classifier. Table 7 indicates the accuracy assessment results for this case, and Fig. 5 indicates the variation of user's accuracy and class kappa values for this case.

Table 7 Accuracy assessment results for Kumta study area for equal class probabilities (case 3)

Class name	RT ^a	CT ^b	NC ^c	PA ^d (%)	UA ^e (%)	Kappa
Evergreen forest	600	593	540	90.00	91.06	0.77656
Deciduous forest	29	74	19	65.52	25.68	0.234559
Water body	79	66	64	81.01	96.97	0.96710
Scrub land	112	122	61	54.46	50.00	0.436937
Kharif	76	56	46	60.53	82.14	0.806741
Double crop	89	42	30	33.71	71.43	0.686373
Built-up	15	47	9	60.00	19.15	0.179177
Totals	1000	1000	769			
Overall classification accuracy = 76.90%						
Overall Kappa value = 0.6241						

^aRT: reference totals, ^bCT: classified totals, ^cNC: number correct, ^dPA: producer's accuracy, ^eUA: users accuracy

Fig. 5 Variation of user's accuracy and Kappa values for classes in case-3



When a priori class probabilities were used for classifying the secondary study area, the overall classification performance improved to 78.10 % from 76.90%. By careful analysis of the individual class analysis, the use of *a priori* class accuracies has brought a fine balance between the overlapping class pairs by separating them more efficiently. A similar behaviour was observed for maximum likelihood classifier by [26–29], as well as in one of our previous studies [21]. Particularly, the *a priori* class probabilities are helpful in separating the severely overlapping class pairs while not affecting the other classes on the data. This is the reason why some of the classes have highly similar user’s accuracies for each data. By comparing the overall Kappa values with values suggested in the literature [25], the classification performance is termed good. Table 8 indicates the accuracy assessment results for this case, and Fig. 6 indicates the variation of user’s accuracy and class kappa values for this case. Figure 7 indicates the variation of overall classification accuracy and overall Kappa value for the four cases discussed above.

To signify the performance of the Bayes’ classifier in classifying intermediate and heterogeneous RS data, the results obtained are compared with seven other supervised and fuzzy supervised classifiers. The results of these classifiers are shown in Table 9.

Table 8 Accuracy assessment results for Kumta study area for *a priori* class probabilities (case 4)

Class name	RT ^a	CT ^b	NC ^c	PA ^d (%)	UA ^e (%)	Kappa
Evergreen forest	600	613	550	91.67	89.72	0.7430
Deciduous forest	29	56	19	65.52	33.93	0.3195
Water body	79	67	66	83.54	98.51	0.9837
Scrub land	112	133	66	58.93	49.62	0.4327
Kharif	76	54	44	57.89	81.48	0.7995
Double crop	89	34	27	30.34	79.41	0.7740
Built-up	15	43	9	60.00	20.93	0.1972
Totals	1000	1000	781			
Overall classification accuracy = 78.10%						
Overall Kappa value = 0.6366						

^aRT: reference totals, ^bCT: classified totals, ^cNC: number correct, ^dPA: producer’s accuracy, ^eUA: users accuracy

Fig. 6 Variation of user’s accuracy and Kappa values for classes in case-4

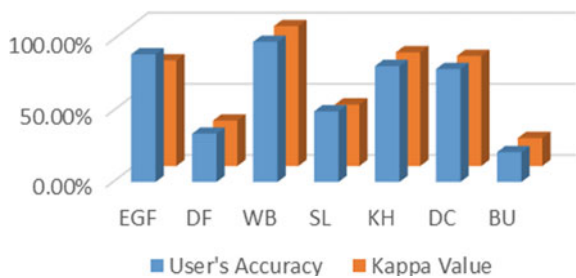


Fig. 7 Variation in accuracy assessment results for the four cases presented

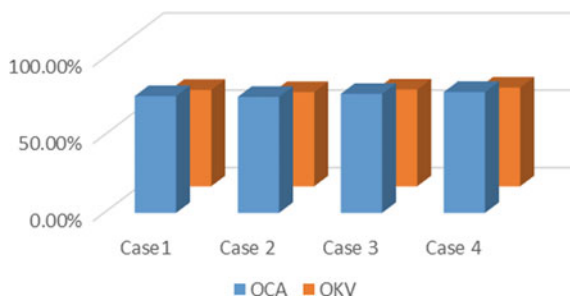


Table 9 Classification accuracy assessment results of some supervised classifiers

Classifier	North Canara		Kumta	
	OCA ^a (%)	OKV ^b	OCA ^a (%)	OKV ^b
Neural network	76.87	0.6501	76.90	0.6174
Multilayer perceptron	69.67	0.5534	78.30	0.6233
Learning vector quantization-1	76.14	0.6327	77.90	0.6049
Learning vector quantization-2	77.78	0.6575	77.40	0.5965
Fuzzy-based maximum likelihood	86.20	0.7870	77.00	0.6069
Fuzzy-based minimum distance	69.58	0.5482	70.40	0.5496
Fuzzy-based Mahalanobis distance	70.58	0.5552	68.80	0.5198

^aOCA: overall classification accuracy, ^bOKV: overall Kappa value

It can be observed that Bayes' classifier's performance is very close to that of neural network and learning vector quantization techniques. Its performance is far superior than that of multilayer perceptron, fuzzy-based minimum distance and Mahalanobis distance classifiers. The only technique that efficiently outperformed the Bayes' classifier is the fuzzy-based maximum likelihood classifier.

5 Conclusion

Soft classifiers have been introduced in the literature to mainly combat the issue of mixed pixels. This paper presented the applicability of Bayes' classifier in handling intermediate and heterogeneous complexity RS data. From the obtained results, it is observed that Bayes' classifier is an excellent choice for the classification of intermediate and heterogeneous study areas which are characterized by severely overlapping LULC classes and mixed pixels. It is illustrated that Bayes' classifier's inherent characteristic of favouring the dominant classes in the study area can be mitigated with the use of *a priori* class probabilities during classification. However, Bayes' classifier has some shortcomings as well. It is very sensitive to the number of training pixels collected for a class. For classes with smaller training sets, the classifier creates

underfitting (Built-Up class being an example for this scenario). The results obtained in this study stand true for the study areas considered. If the study area is changed to a different type or characteristic data, the results may vary in accordance with the LULC overlapping. As a future work, it will be interesting to test Bayes' classifier on hyperspectral RS data.

References

1. J.R. Anderson, *A land use and land cover classification system for use with remote sensor data*, vol. 964 (US Government Printing Office, 1976)
2. S.E. Jozdani, B.A. Johnson, D. Chen, Comparing deep neural networks, ensemble classifiers, and support vector machine algorithms for object-based urban land use/land cover classification. *Remote Sens.* **11**(14), 1713 (2019)
3. S. Liu, Z. Qi, X. Li, A.G.-O. Yeh, Integration of convolutional neural networks and object-based post-classification refinement for land use and land cover mapping with optical and sar data. *Remote Sens.* **11**(6), 690 (2019)
4. M. Wu, X. Zhao, Z. Sun, H. Guo, A hierarchical multiscale super-pixel-based classification method for extracting urban impervious surface using deep residual network from worldview-2 and lidar data. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* **12**(1), 210–222 (2019)
5. C. Zhang, I. Sargent, X. Pan, H. Li, A. Gardiner, J. Hare, P.M. Atkinson, Joint deeplearning for land cover and land use classification. *Remote Sens. Environ.* **221**, 173–187 (2019)
6. X. Xu, J. Li, S. Li, A. Plaza, Subpixel component analysis for hyperspectral imageclassification. *IEEE Trans. Geosci. Remote Sens.* **57**(8), 5564–5579 (2019)
7. P.V. Arun, I. Herrmann, K.M. Budhiraju, A. Karnieli, Convolutional network architectures for super-resolution/sub-pixel mapping of drone-derived images. *Pattern Recogn.* **88**, 431–446 (2019)
8. M. Wang, X. Zhang, X. Niu, F. Wang, X. Zhang, Scene classification of high-resolutionremotely sensed image based on resnet. *J. Geovisualization Spat. Anal.* **3**(2), 16 (2019)
9. H. Huang, K. Xu, Combing triple-part features of convolutional neural networks forscene classification in remote sensing. *Remote Sens.* **11**(14), 1687 (2019)
10. F. Wang, Fuzzy supervised classification of remote sensing images. *IEEE Trans. Geosci. Remote Sens.* **28**(2), 194–201 (1990)
11. F. Wang, Improving remote sensing image analysis through fuzzy information representation, in *Photogrammetric Engineering and Remote Sensing (USA)* (1990)
12. F. Sets, L. Zadeh, *Inform. Control* **8**, 338–353 (1965)
13. L.A. Zadeh, Probability measures of fuzzy events. *J. Math. Anal. Appl.* **23**(2), 421–427 (1968)
14. A. Turčan, E. Ocelíková, L. Madarász, Fuzzy c-means algorithms in remote sensing, in *Proceedings of the 1st Slovakian-Hungarian Joint Symposium on Applied Machine Intelligence (SAMI)*, Herlany, pp. 207–216 (2003)
15. I. Nedeljkovic, Image classification based on fuzzy logic. *Int. Arch. Photogram. Remote Sens. Spat. Inf. Sci.* **34**(30), 3–7 (2004)
16. T. Nakashima, G. Schaefer, Y. Yokota, H. Ishibuchi, A weighted fuzzy classifier and itsapplication to image processing tasks. *Fuzzy Sets Syst.* **158**(3), 284–294 (2007)
17. D.J. Hand, K. Yu, Idiot's bayes—not so stupid after all? *Int. Stat. Rev.* **69**(3), 385–398 (2001)
18. A. McCallum, *Bayesian Network Representation* (2011). [Online]. Available: <https://courses.cs.washington.edu/courses/cse515/11sp/class2-bayesnet.pdf>
19. U. EarthExplorer, Us department of the interior, US Geological Survey. <http://earthexplorer.usgs.gov>
20. M.N. Murty, V.S. Devi, *Pattern Recognition: An Algorithmic Approach* (Springer Science & Business Media, 2011)

21. B.R. Shivakumar, S.V. Rajashekararadhya, An investigation on land cover mapping capability of classical and fuzzy based maximum likelihood classifiers. *Int. J. Eng. Technol.* **7**(2), 939–947 (2018)
22. G.H. John, P. Langley, Estimating continuous distributions in bayesian classifiers, in *Proceedings of the Eleventh Conference on Uncertainty in Artificial Intelligence*, pp. 338–345 (Morgan Kaufmann Publishers Inc., 1995)
23. V. Metsis, I. Androustopoulos, G. Paliouras, Spam filtering with naive bayes-which naive bayes? in *CEAS*, vol. 17, pp. 28–69 (Mountain View, CA, 2006)
24. A. McCallum, K. Nigam et al., A comparison of event models for naive bayes text classification, in *AAAI-98 Workshop on Learning for Text Categorization*, vol. 752, no. 1. Citeseer, pp. 41–48 (1998)
25. J.A. Richards, *Remote Sensing Digital Image Analysis*, vol. 5 (Springer, 2013)
26. D. McIver, M. Friedl, Using prior probabilities in decision-tree classification of remotelysensed data. *Remote Sens. Environ.* **81**(2–3), 253–261 (2002)
27. M. Zheng, Q. Cai, Z. Wang et al., Effect of prior probabilities on maximum likelihood classifier (2005)
28. Z. Mingguo, C. Qianguo, Q. Mingzhou, The effect of prior probabilities in the maximum likelihood classification on individual classes. *Photogram. Eng. Remote Sens.* **75**(9), 1109–1117 (2009)
29. B.R. Shivakumar, S.V. Rajashekararadhya, Investigation on land cover mappingcapability of maximum likelihood classifier: a case study on North Canara, India. *Procedia Comput. Sci.* **143**, 579–586 (2018)

A Study on the Effect of Dimensionality Reduction on Classification Accuracy of Myoelectric Control Systems



Praahas Amin  and Airani Mohammad Khan

1 Introduction

Surface electromyography (SEMG) is a physiological signal recorded from the skin surface electrodes that represents muscle activity such as a hand action. It is a representation of the electric field potential due to the depolarization of the outer muscle fibre membrane [1]. Distinct muscle fibre groups undergo contraction for distinct actions leading to exclusive patterns in the EMG signal corresponding to each action. The electromyography signal is a time series physiological signal and is stochastic. Therefore, the amplitude of the signal cannot be used directly as exclusive action-defining data points. Feature extraction has to be done to identify patterns in the signal for each action. Machine learning techniques can be used to classify actions of the hand based on these feature sets. When many features are used to train a classifier model, it can lead to larger training time. In addition to this, it can also lead to the model becoming too complex in the attempt to include the effect of every feature leading to overfitting. To reduce overfitting, dimensionality reduction can be performed on the feature set. This will reduce the overall number of features and reduce the computational demand for training the algorithm. This will also help in generating a simpler model that may be able to generalize better. This paper is focused on studying the effects of using dimensionality reduction techniques on the accuracy of the classifiers when principal component analysis (PCA), linear discriminant analysis (LDA) and quadratic discriminant analysis (QDA) is used on the time-domain features of SEMG signal

P. Amin (✉) · A. M. Khan
Department of Electronics, Mangalore University, Mangalore, India
e-mail: praahas1234@gmail.com

A. M. Khan
e-mail: asifabc@gmail.com

2 Background

The reviewed literature comprised of studies done on hand action classification systems from SEMG data. The study presented in [2, 3] involved the use of support vector machines (SVM) for identification of hand actions from electromyography data and comparison with the performance of a multi-layer perceptron (MLP) neural networks along with the effects of the electrodes getting displaced. SVM showed a classification performance that matched with other algorithms and was computationally efficient. In [4], a convolutional neural network (CNN)-based feature extraction was proposed. It suggested that a CNN could be an effective feature extractor. They reported that a combination of traditional features with CNN extracted features showed an improvement in accuracy by 4.35, 3.62 and 4.7% for SVM, LDA and K-nearest neighbours (KNN). They also suggest the use of CNN to transfer the raw SEMG into a low-dimensional SEMG feature space. For comparison of performance, they used three classical classifiers SVM, LDA and KNN. In [5], a methodology was proposed to classify the default actions of the myo-armband using machine learning techniques. They evaluated the performance of MLP, KNN, LDA, QDA, decision tree (DT) and Naive Bayes (NB) classifier. They reported achieving overall high accuracy of greater than 90% on single movements with MLP. LDA, KNN, QDA and DT have shown the accuracy of around 80%. NB was shown demonstrated to have poor accuracy. They suggest expanding the feature vector to improve accuracy. The features considered in this study are mean of the absolute value of the signal (MAV), integrated EMG (IEMG), root mean square value of the signal (RMS) and variance of signal (VAR). In [6], a gesture recognition using SEMG in the transient state is proposed. The classification algorithm used was SVM with Gaussian Kernel (SVM-G) and discrete wavelet transform (DWT). The study considered the five default gestures of the myo-armband. A window size of 750 ms was used, and an accuracy of 87.5% was achieved in this study. In [7], an SVM classification model that classified gestures from SEMG data with an accuracy of 89% was proposed. They suggested pre-processing the original data to eliminate the effect of noise by using absolute value function and a Butterworth low-pass filter. In [8], a sensor fusion technique was proposed using the SEMG data from the SEMG sensors and inertial measurement unit data of the myo-armband to classify user-independent hand gestures using MLP. The classifier model was reported to perform with an average accuracy of 78.94%. They demonstrated that it was possible to have a more intuitive human-machine interface that allowed better control of wearable machines. The gestures classified were the default gestures of the myo-armband. In [9], a real-time hand gesture classification system using artificial neural network (ANN) was proposed. They reported an accuracy of 96%. The features considered were time-domain features. The gestures used were the default armband gestures, and an accuracy of $85.08\% \pm 15.21\%$ was reported. They reported a slight delay between the onset of the gesture and the recognition of the gesture attributing it to the use of the sliding window technique. Based on the reviewed literature, the classification algorithms suggested by most of the authors were SVM, ANN, DT and

LDA. Therefore, these were the primary algorithms considered along with logistic regression (LR), random forest (RF), KNN and Gaussian Naïve Bayes (GNB). PCA and LDA were the two dimensionality reduction techniques used in addition to use of CNN. However, deep learning techniques were excluded from this study, and we restricted the study to a comparison of accuracies of classifier algorithms before and after dimensionality reduction.

3 Methodology

In this section, we outline the methodology used for the acquisition of data, training, evaluating and testing the machine learning models. In this study, data is collected from ten consenting volunteers or users of varying age groups and gender. This paper focuses on five hand actions pointer extension (A1) which represents the extension of the pointer finger, middle extension (A2), full grip (A3), half grip (A4) and no action (A5) as shown in Fig. 1. When a person does an action by contracting their muscles, the intention reaches the muscle fibres through the network of neurons in the user's body terminating in fibre contractions. The SEMG signal is then recorded by the acquisition system. The sensors are connected to a computer using a Bluetooth dongle. A software was developed to automate the process of initiating a data recording session and providing instructions to the user to perform the action. The application takes the user's details as input. Prompts are flashed to the users on the screen to initiate an action and relax. Data will be recorded in the form of comma-separated values. This step is repeated for every action required. The unique features characterizing an action are extracted from the raw electromyography data and then labelled. The labelled data is split into training and validation set. The machine learning models are then trained.

A. Data Acquisition and Segmentation

100 instances were collected for each person for each of the five actions. One instance would include a rest window of three seconds followed by performing the action

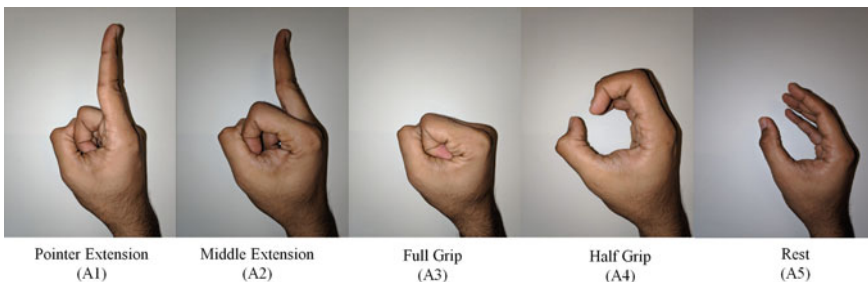


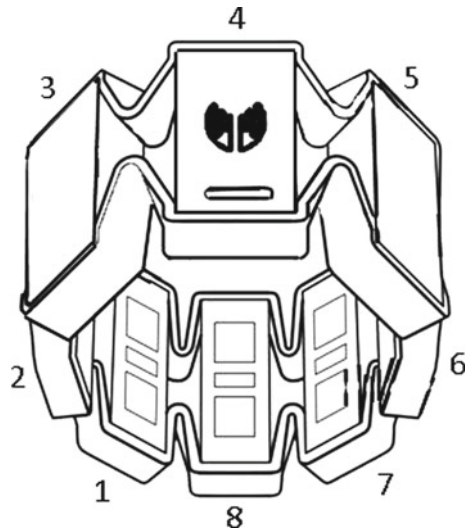
Fig. 1 Actions used in the study

within a window of 2 s followed by a rest window of 3 s. This pattern of action and relaxation is repeated in order to acquire physiological data. Each session is restricted to a batch of 25 actions instances per session to negate the build-up of fatigue. The device used for recording the data is a wearable myo-armband shown in Fig. 2. It has eight surface electromyography sensors which are placed at equal distances and has a sampling frequency of 200 Hz. The muscle data is recorded from ten users. Every person performs the actions being tested in this study. The recorded data contains 100 samples for each of the actions performed myoelectric systems that are meant to be used with real-time systems, and therefore, the data is segmented in analysis of electromyography signals. Segment or window is the number of samples being considered for analysis and feature extraction. Data segmentation can be done using two techniques, i.e., isolated window technique and sliding window technique. In this paper, the technique used for the data collection and feature computation is the isolated window technique. For the training activity, the time instants of the action onset are noted, and then, through the software written in Python language, we initiate the feature extraction process for that window. The window size considered is 8 s over which features are extracted. This process happens across three of the eight channels. For the actions selected in this study, the recordings of sensor 1 was sufficient to accurately classify the actions during offline classification. 20% of the data is used for validation.

B. *Feature Extraction and Classification*

The raw signal data cannot be used to perform hand action classification. In order to identify exclusive parameters of the action contained in the signal, feature extraction is performed. In this study, 19 features were extracted. Integrated signal (IEMG), mean of the absolute value of the signal (MAV) and its variants MAV1 and MAV2

Fig. 2 Myo-armband



[2, 10, 11] are features corresponding to active segment detection and onset indication for a gesture, where x_n is an EMG signal and N is the length of that signal.

$$\text{IEMG} = \sum_{n=1}^N |x_n| \quad (1)$$

$$\text{MAV} = \frac{1}{N} \sum_{n=1}^N |x_n| \quad (2)$$

$$\text{MAV1} = \frac{1}{N} \sum_{n=1}^N w_n |x_n| \quad (3)$$

where $w_n = \begin{cases} 1, & \text{if } 0.25N \leq n \leq 0.75N \\ 0.5, & \text{otherwise} \end{cases}$

$$\text{MAV2} = \frac{1}{N} \sum_{n=1}^N w_n |x_n| \quad (4)$$

where $w_n = \begin{cases} 1, & \text{if } 0.25N \leq n \leq 0.75N \\ \frac{4n}{N} & \text{if } n < 0.25N \\ \frac{4(n-N)}{N}, & \text{otherwise} \end{cases}$

Features corresponding to energy and power of electromyography signal are simple square integral (SSI) and variance (VAR), respectively. Root mean square (RMS) is another feature that corresponds to the power content of the signal [10, 12–15].

$$\text{SSI} = \sum_{n=1}^N |x_n|^2 \quad (5)$$

$$\text{VAR} = \frac{1}{N-1} \sum_{n=1}^N (x_n - \bar{x})^2 \quad (6)$$

$$\text{RMS} = \sqrt{\frac{1}{N-1} \sum_{n=1}^N (x_n - \bar{x})^2} \quad (7)$$

Absolute value of the third, fourth and fifth temporal moment [16]. First and second moments correspond to MAV and VAR.

$$\text{TM}r = \left| \frac{1}{N} \sum_{n=1}^N x_n^r \right| \text{ where } r = 3, 4 \text{ and } 5 \quad (8)$$

The feature corresponding to muscle contraction force is log detector (LOG) feature [10, 17].

$$\text{LOG} = e^{\frac{1}{N} \sum_{n=1}^N \log|x_n|} \quad (9)$$

Features containing frequency information of the electromyography signal are waveform length (WL), average amplitude change (AAC), the absolute standard deviation value difference (DASDV), zero crossing (ZC) and Willison amplitude (WAmplitude). The change in the sign of the slope (SSC) is a feature corresponding to frequency information [2, 10, 11, 14, 18, 19].

$$\text{WL} = \sum_{n=1}^{N-1} |x_{n+1} - x_n| \quad (10)$$

$$\text{AAC} = \frac{1}{N} \sum_{n=1}^{N-1} |x_{n+1} - x_n| \quad (11)$$

$$\text{DASDV} = \sqrt{\frac{1}{N-1} \sum_{n=1}^{N-1} (x_{n+1} - x_n)^2} \quad (12)$$

$$\text{ZC} = \sum_{n=1}^N \text{sgn}(-x_n \times x_{n+1}) \quad (13)$$

where $\text{sgn}(x) = \begin{cases} 1, & \text{if } x \geq \text{threshold} \\ 0, & \text{otherwise} \end{cases}$

$$\text{WAmplitude} = \sum_{n=1}^N f(|x_n - x_{n+1}|) \quad (14)$$

where $f = \begin{cases} 1, & \text{if } x \geq \text{threshold} \\ 0, & \text{otherwise} \end{cases}$

$$\text{SSC} = \sum_{n=2}^{N-1} f((x_n - x_{n-1})(x_n - x_{n+1})) \quad (15)$$

where $f = \begin{cases} 1, & x \geq \text{threshold} \\ 0, & \text{otherwise} \end{cases}$

Percentage of myo pulse (MYOP) is the average value of myo pulse output and first burst amplitude (AFB) which is the first maximum point extracted from the data [18].

$$\text{MYOP} = \frac{1}{N} \sum_{n=1}^N f(|x_n|) \quad (16)$$

where $f = \begin{cases} 1, & \text{if } x_n \geq \text{threshold} \\ 0, & \text{otherwise} \end{cases}$

To classify the gesture using the selected features, the classifiers considered in this study are logistic regression (LR), decision tree (DT), random forest (RF), support vector machine with linear kernel (SVML) support vector machine with radial basis function kernel (SVMR), K-nearest neighbours (KNN), Gaussian Naïve Bayes (GNB) and artificial neural network (ANN). The data acquired from each person is used to train each of the classifier, and the classification accuracy is then checked for each classifier.

C. Dimensionality Reduction

As we increase the number of features, the classification accuracy increases up to a certain limit beyond which accuracy starts to drop. This phenomenon is termed as curse of dimensionality. When there are many features used to identify actions, some of them may be correlated, and therefore, they will be giving redundant information. Too many features may also cause the model to over fit. It may be difficult to perform action classification with sufficient accuracy from a higher dimensional feature space. By dimensionality reduction, we can reduce the dimensions of the feature space by transforming the features into a different perspective. By this technique, the number of features get reduced without losing information. In order to compare the effects of dimensionality reduction on the classification accuracy, the data is first classified without performing dimensionality reduction and then after dimensionality reduction. Dimensionality reduction is performed using PCA, LDA and QDA. The results are then compared. The observations are discussed in the next section.

4 Results and Discussion

The study was focused on observing the effects of dimensionality reduction in hand action classification systems. The accuracy metric is calculated from the confusion matrix for each action based on number of true positives (TP), true negatives (TN), false positives (FP) and false negatives (FN).

$$\text{Accuracy (ACC)} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (17)$$

The accuracy curves of the ANN for the training and validation data without any dimensionality reduction is shown in Fig. 3. The accuracy curves of the ANN for the training and validation data with PCA, LDA and QDA is shown in Figs. 4, 5

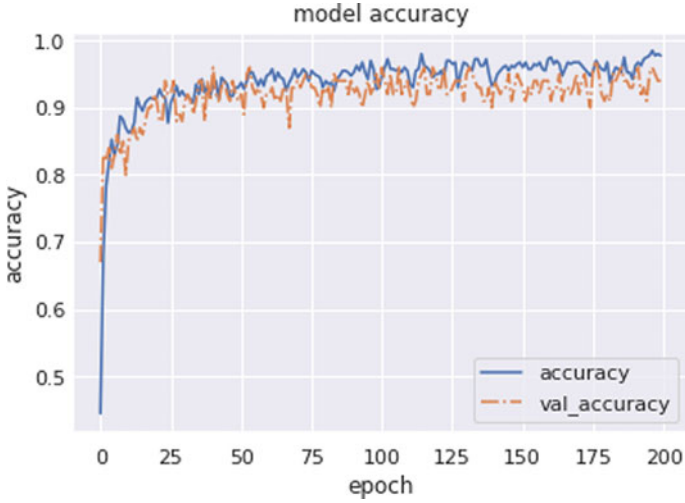


Fig. 3 Accuracy curve (without dimensionality reduction)



Fig. 4 Accuracy curve (with PCA)

and 6, respectively. It is also observed from the accuracy curves that there is no significant reduction in validation accuracy compared to training accuracy in all the cases thereby ruling out any possibility of overfitting.

It can be observed that the performance is consistent without any significant drop in accuracy after performing dimensionality reduction. The confusion matrices for test user 5 has been considered for SVMML and is shown for the cases where

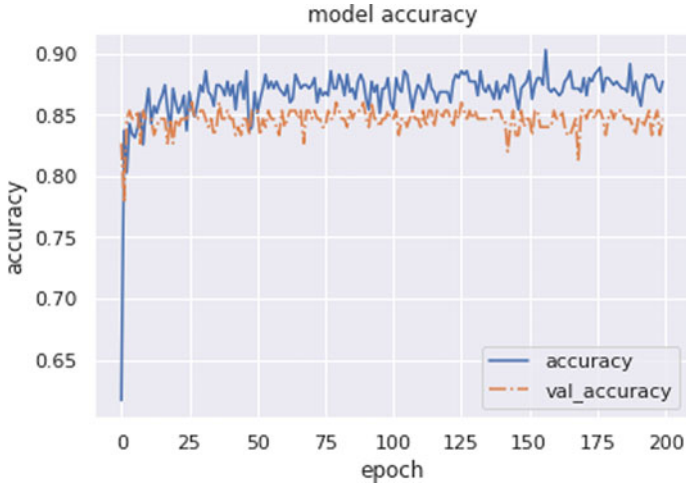


Fig. 5 Accuracy curve (with LDA)

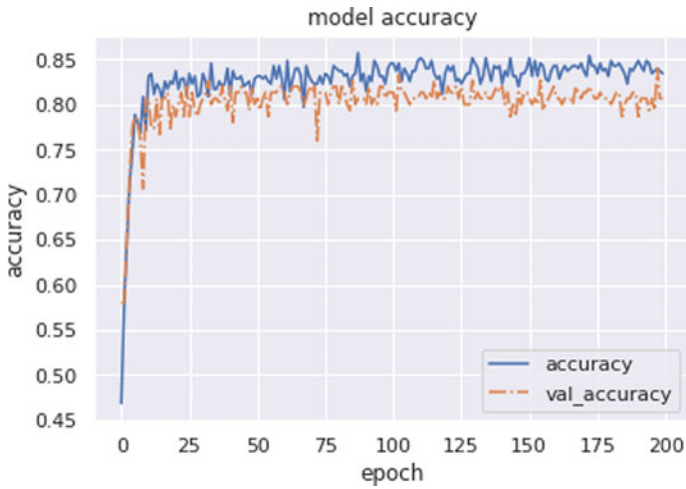


Fig. 6 Accuracy curve (with QDA)

Fig. 7 Confusion matrix for user 5 using SVML with PCA

SVML (PCA)	A1	A2	A3	A4	A5
A1	37	2	0	0	0
A2	1	21	0	1	0
A3	0	0	26	0	0
A4	2	21	2	6	0
A5	0	0	0	0	31

dimensionality reduction has been performed using PCA, LDA and QDA in Figs. 7, 8 and 9, respectively. These can be observed for user 5 in Tables 2, 3, and 4, respectively.

From the confusion matrices, it can be observed that most of the misclassifications have happened when performing cylinder grip action. The cylinder grip action has been misclassified as middle extension. This can be attributed to the fact that the two gestures involve contraction of similar muscle groups.

The classification accuracy of the feature set across all users without dimensionality reduction is given in Table 1. Classification accuracy of all users with

Fig. 8 Confusion matrix for user 5 using SVML with LDA

SVML (LDA)	A1	A2	A3	A4	A5
A1	38	1	0	0	0
A2	1	21	1	0	0
A3	0	1	25	0	0
A4	1	19	0	11	0
A5	0	0	0	0	31

Fig. 9 Confusion matrix for user 5 using SVML with QDA

SVML (QDA)	A1	A2	A3	A4	A5
A1	37	2	0	0	0
A2	1	21	0	1	0
A3	0	0	26	0	0
A4	2	21	2	6	0
A5	0	0	0	0	31

Table 1 Accuracy of classifiers without dimensionality reduction

User	Accuracy (%)							
	LR	DT	RF	SVML	SVMR	KNN	GNB	ANN
1	86	91	96	87	92	85	82	85
2	78	91	91	81	86	82	84	90
3	80	92	96	79	81	80	74	86
4	97	98	99	95	98	90	98	96
5	86	88	90	87	87	87	85	93
6	90	96	93	89	91	86	87	93
7	87	88	86	88	90	90	76	90
8	98	97	98	98	99	96	92	97
9	90	93	98	86	93	89	88	91
10	91	98	99	91	94	90	88	90

dimensionality reduction using PCA, LDA and QDA are given in Tables 2, 3 and 4, respectively. Table 5 shows the average accuracy of all classifiers across all users with and without dimensionality reduction. From the results, we observe that compared to the accuracy of the classifiers without dimensionality reduction, there was no significant drop in accuracy when dimensionality reduction was applied on the feature space of the SEMG data. When ANN was used as a classifier, there was a consistent performance across all users with PCA, LDA and QDA for dimensionality reduction. It was observed that decision tree was prone to overfitting, and the misclassification rate was higher for the testing set. This misclassification rate was reduced by pruning the decision tree. From this observation, it is evident that by performing dimensionality reduction on the original feature space of surface EMG data, the benefits of reduced dimensionality can be achieved without a significant loss of accuracy.

Table 2 Accuracy of classifiers using PCA for Dimensionality Reduction

User	Accuracy (%) PCA							
	LR	DT	RF	SVML	SVMR	KNN	GNB	ANN
1	81	77.3	80.6	79.3	82.6	74	76	85
2	74.6	78.6	76	80	79.3	77.3	75.3	80
3	68.6	70.6	70.6	68.6	71.3	69.3	64.6	71.3
4	86.6	86	90	90.6	93.3	93.3	93.3	92
5	84.6	78.6	82	81	84.6	86	79.3	88
6	89.3	87.3	89.3	89.3	87.3	93.3	91.3	87
7	80	78.6	78.6	82.6	82.6	82.6	75.3	84
8	90.6	90.6	90	91	93	87	86.6	90
9	73.3	77.3	76	70	74.6	80.6	76	83.4
10	84	90.6	93.3	85.3	84.6	87.3	86.6	87.3

Table 3 Accuracy of classifiers using LDA for Dimensionality Reduction

User	Accuracy (%) LDA							
	LR	DT	RF	SVML	SVMR	KNN	GNB	ANN
1	82	81	85	84	86	88	86	84
2	91.3	88	91.3	92.6	92.6	96	91.3	90
3	74.6	76	81.3	75.3	76	82	76	82
4	97.3	96	96	97.3	98	97.3	97.3	98
5	84	86	86.6	84	84	86.6	86.6	86
6	96.6	92.6	96	96	96	98	95.3	93.3
7	79.3	82	85.3	82.6	81.3	85.3	86.6	88
8	95.3	98	96.6	94	95.3	98.6	96	98.6
9	88.6	89.3	91.3	88.6	88.6	91.3	88	91
10	93.3	90.6	89.3	94.6	92.6	94	94.6	90.6

Table 4 Accuracy of classifiers using QDA for Dimensionality Reduction

User	Accuracy (%) QDA							
	LR	DT	RF	SVML	SVMR	KNN	GNB	ANN
1	81.3	77.3	80.6	79.3	82.6	74	76	83.3
2	74.6	78.6	76	80	79.3	77.3	75.3	79
3	68.6	70.6	70.6	68.6	71.3	69.3	64.6	79
4	86.6	86	90	90.6	93.3	93.3	93.3	91.3
5	84.6	78.6	82	81	84.6	86	79.3	88
6	89.3	87.3	89.3	89.3	87.3	93.3	91.3	87.3
7	80	78.6	78.6	82.6	82.6	82.6	75.3	82
8	90.6	90	90	90.6	93.3	87.3	86.6	94
9	73.3	77.3	76	70	74.6	80.6	76	80
10	84	90.6	93.3	85.3	84.6	87.3	87	90

Table 5 Average accuracy of classifiers with and without dimensionality reduction Methods

Method	Accuracy (%)							
	LR	DT	RF	SVML	SVMR	KNN	GNB	ANN
NONE	88.3	93.2	94.6	88.1	91.1	87.5	85.4	91.1
PCA	81.26	81.5	82.6	82.2	83.3	83.0	80.4	84.8
LDA	88.23	87.9	89.8	88.9	89.0	91.7	89.7	90.15
QDA	81.29	81.4	82.6	82.2	83.3	83.1	80.4	85.39

5 Conclusion and Future Work

The hand action classification system would be implemented on an embedded system on which, memory and computing power would be constrained. From the results it is observed that instead of using all 19 time-domain features, dimensionality reduction can be performed on the feature space to reduce computational requirements and training time without a significant loss of accuracy. With artificial neural network as the classifier, the best and consistent accuracy was observed. Future work can be focused on implementing an online action classification system on an embedded development board with sliding window technique.

Acknowledgements The authors would like to thank Department of Electronics, Mangalore University and Flashflow® Technologies (OPC) Private Limited, for all the infrastructure and literature support provided during the research work.

Declaration We have taken permission from competent authorities to use the data as given in the paper. In case of any dispute in the future, we shall be wholly responsible.

References

1. M.Z. Jamal, Signal acquisition using surface EMG and circuit design considerations for robotic prosthesis, in *Computational Intelligence in Electromyography Analysis—A Perspective on Current Applications and Future Challenges* (IntechOpen, 2012)
2. M.A. Oskoei, H. Hu, Support vector machine-based classification scheme for myoelectric control applied to upper limb. *IEEE Trans. Biomed. Eng.* **55**(8), 1956–1965 (2008)
3. L. Hargrove, K. Englehart, B. Hudgins, The effect of electrode displacements on pattern recognition based myoelectric control, in *28th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS'06*, pp. 2203–2206 (2006)
4. H. Chen, Y. Zhang, G. Li, Y. Fang, H. Liu, Surface electromyography feature extraction via convolutional neural network. *Int. J. Mach. Learn. Cybern.* **11**(1), 185–196 (2020)
5. M.L.B. Freitas, J.J.A. Mendes, D.P. Campos, S.L. Stevan, Hand gestures classification using multichannel SEMG armband, in *XXVI Brazilian Congress on Biomedical Engineering, IFMBE Proceedings*, ed. by R. Costa-Felix, J.C. Machado, A.V. Alvarenga (Springer, Singapore, 2019), pp. 239–246
6. J. Yanez, L. Unapanta, M.E. Benalcázar, Short-term hand gesture recognition using electromyography in the transient state, support vector machines, and discrete wavelet transform, in *2019 IEEE Latin American Conference on Computational Intelligence (LA-CCI)* (Guayaquil, Ecuador, 2019), pp. 1–6
7. W. Chen, Z. Zhang, Hand gesture recognition using SEMG signals based on support vector machine, in *Proceedings of the IEEE 8th Joint International Information Technology and Artificial Intelligence Conference*, Chongqing, China, pp. 230–234, May 2019
8. J.G. Colli-Alfaro, A. Ibrahim, A.L. Trejos, Design of user-independent hand gesture recognition using multilayer perceptron networks and sensor fusion techniques, in *Proceedings of the IEEE 16th International Conference on Rehabilitation Robotics (ICORR)*, Toronto, ON, Canada, pp. 1103–1108, July 2019
9. K. Yang, Z. Zhang, Real-time pattern recognition for hand gesture based on ANN and surface EMG, in *Proceedings of the IEEE 8th Joint International Information Technology and Artificial Intelligence Conference*, Chongqing, China, pp. 799–802, May 2019
10. M. Zardoshti-Kermani, B.C. Wheeler, K. Badie, R.M. Hashemi, EMG feature evaluation for movement control of upper extremity prostheses. *IEEE Trans. Rehab. Eng.* **3**(4), 324–333 (1995)
11. B. Hudgins, P. Parker, R.N. Scott, A new strategy for multifunction myoelectric control. *IEEE Trans. Biomed. Eng.* **40**(1), 82–94 (1993)
12. S. Du, M. Vuskovic, Temporal vs. spectral approach to feature extraction from prehensile EMG signals, in *Proceedings of the IEEE International Conference on Information Reuse and Integration-IRI*, Las Vegas, USA, pp. 344–350, (IEEE, 2004)
13. R. Boostani, M. Morad, Evaluation of the forearm EMG signal features for the control of a prosthetic hand. *Physiol. Meas.* **24**(2), 309–319 (2003)
14. K.S. Kim, H.H. Choi, C.S. Moon, C.W. Mun, Comparison of k-nearest neighbor, quadratic discriminant and linear discriminant analysis in classification of electromyogram signals based on the wrist-motion directions. *Curr. Appl. Phys.* **11**(3), 740–745 (2011)
15. S.H. Park, S.P. Lee, EMG pattern recognition based on artificial intelligence techniques. *IEEE Trans. Rehab. Eng.* **6**(4), 400–405 (1998)
16. G.N. Saridis, T.P. Gootee, EMG pattern analysis and classification for a prosthetic arm. *IEEE Trans. Biomed. Eng.* **29**(6), 403–412 (1982)
17. D. Tkach, H. Huang, T.A. Kuiken, Study of stability of time-domain features for electromyographic pattern recognition. *J. Neuroeng. Rehab.* **7**(1), 21 (2010)
18. A. Fougner, *Proportional Myoelectric Control of a Multifunction Upper-Limb Prosthesis*. Ph.D. thesis (2007)
19. L. Philipson, *The Electromyographic Signal Used for Control of Upper Extremity Prostheses and for Quantification of Motor Blockade During Epidural Anaesthesia* (1987)

An Efficient Low Power MEMS-Based Microfluidic Device for the Segregation of Different Blood Components



Ranjith B. Gowda, P. Vanishree, and Preeta Sharan

1 Introduction

The living mechanism is different for all the creatures and animals. The system that helps to have the energy for the individual is mainly by the fluid running in their system. Human blood, which is a major combination of RBC, WBC, platelets and plasma plays a role as fluid of liveliness [1]. Blood has plasma accounting to 55%, 45% of other cells and less than 1% of platelets and WBC in total. Each components of blood have their own functions to carry out. Component RBC plays an essential role in exchanging of oxygen and carbon dioxide between tissues and lungs [2], whereas WBC helps to build the immunity of the system in order to provide strength against infections and diseases. Platelets are concerned in improving the growth factor, promote blood clotting and they are also used as an adjunct to wound recovery [1]. Since they all play a vital role in human living system, a significant imbalance in their quantity can have a serious effect and result in dangerous outcomes like, low platelet count can lead to thrombocytopenia and high-count leads to thrombocythemia [3, 4]. In case of any such occurrence, blood transfusion is one way to restore the required amount of cell count in the patient body and minimize the risk. Blood transfusion method is rarely conducted. There are different kinds of blood cell transfusion based on the particular blood cell component [5]. Above-mentioned facts give a strong reason for sorting of RBC, WBC and platelets as they are the main components of blood. Centrifuge technique is used to separate RBC, platelets and plasma from each other. Platelets transfusion is carried out to reduce the mortality

R. B. Gowda (✉)

Department of Electronics and Communication Engineering, SOE, Dayananda Sagar University, Bangalore, India

Government Polytechnic Sorab, Shimoga, India

P. Vanishree · P. Sharan

Department of Electronics and Communication Engineering, The Oxford College of Engineering, Bangalore, India

rate due to marrow failure [6]. Dielectrophoresis technology can be used for the separation of RBC, WBC and platelets cells efficiently. Microfluidic devices can be modelled to segregate and manipulate blood cells and other biological particles like bacteria. It can be easily organized, as it holds the ups in case of particle size, manufacturing techniques, power requirement, high level of analysis, diagnosis obtained and greater controlling system [7, 8].

In this paper, Sect. 2 explains the theory of DEP, modelling and design are discussed in Sect. 3, simulation result and discussion are given in Sect. 4, finally Sect. 5 concludes.

2 Theory

The polarized particle interacts with the non-uniform electric field in a medium and results in dielectrophoretic force acting on the particle [9]. With the proper DEP design the separation, orientation and trapping of particles can be obtained at a pre-defined point [10]. DEP force is manipulated based on the physical dimension, dielectric properties of particle in the suspended medium. DEP can be brought into action with application of AC or DC voltage. Since DC requires high voltage for the generation of required DEP force, it sets as a drawback for the simulation of the particles. Therefore, alternating voltage can be considered as the suitable method to generate the electric field compatible to drive the particles in a medium [9]. Dielectrophoresis force leads to the formation of dipole at the opposing end of the particles in the presence of electric field depending on their dielectric properties [11], and the particle may experience different amount of forces on them. The force acting on a dipole in an electric field is given as,

$$F = (p \cdot \nabla) E \quad (1)$$

Here, F is force on the particle, ∇E is gradient of EF and p is dipole moment.

Dielectrophoretic responses cannot be quantified directly, though there are many techniques to measure DEP force. This method depends on a proportional response of direction, orientation and strength of the EF force in the model spectrum. Most of techniques use Clausius–Mossotti (CM) factor of the particle. Considering the dimension of particle, spherical particle has the simplest modelling and dielectrophoretic force acting on it is given by [9],

$$F_{\text{DEP}} = 2\pi \epsilon_m f_{\text{cm}} r^3 \text{real}[K] \nabla |E_{\text{rms}}|^2 \quad (2)$$

where, ϵ_m is the medium absolute permittivity, r is the particle radius in the medium, f_{cm} is the Clausius–Mossotti (CM) factor and is given as,

$$f_{cm} = \frac{\epsilon_p - \epsilon_{cm}}{\epsilon_p + 2\epsilon_{cm}} \quad (3)$$

Clausius–Mossotti (CM) factor has a limit of $[-0.5, 1]$. When the CM factor is negative the particle is oriented towards the low electric field, that is, negative-DEP [12]. Since blood cells are responsive to negative force, we can observe a significant response of particle to the negative-DEP force that is produced by the AC electric field. DEP forces can be generated with low voltages. Drag force is one of the forces used along with DEP force in the manipulation of particles and other forces such as electrothermal force, buoyancy force, AC electro-osmotic force and Brownian motion are also considered in the process. AC DEP is the most suitable and advisable for the sorting of particles on a microscale. In such case, the complex permittivity comes into picture, that is, the permittivity in the Clausius–Mossotti factor is modified with complex permittivities. The dipole moment can be modified and expressed as [9],

$$P = 4\pi\epsilon_m f_{cm} r^3 E \quad (4)$$

$$f_{cm} = \frac{\epsilon_p^* - \epsilon_m^*}{\epsilon_p^* + 2\epsilon_m^*} \quad (5)$$

$$\epsilon^* = \epsilon - j \frac{\sigma}{\omega} \quad (6)$$

In the above expression, ϵ_m^* and ϵ_p^* are the complex permittivity of the suspending medium and particle, respectively [13]. Parameter f_{cm} is the CM factor real part and is the frequency at which the force is governed in the electric field. Time averaged DEP force on a particle is given as [9],

$$F_{DEP}(t) = 2\pi\epsilon_m f_{cm} r^3 \nabla |E_{rms}|^2 Re[f_{cm}] \quad (7)$$

where, E_{rms} is the applied AC field RMS value. The DEP force induced on the blood particles can be accurately calculated by considering the shell properties. Since the blood particles have different complex permittivity, as compared to the shell complex permittivity, it is replaced with the complex relative permittivity and the governing equation that include both shell and blood particle complex permittivity is given as [12],

$$\epsilon_{eq}^* = \epsilon_s^* \left(\frac{\left(\frac{r_o}{r_i}\right)^3 + \frac{2(\epsilon_p^* - \epsilon_s^*)}{\epsilon_p^* + 2\epsilon_s^*}}{\left(\frac{r_o}{r_i}\right)^3 - \frac{2(\epsilon_p^* - \epsilon_s^*)}{\epsilon_p^* + 2\epsilon_s^*}} \right) \quad (8)$$

where, r_i and r_o are the inner and outer radii of the shell and the complex permittivity of particle and outer shell is ϵ_p^* and ϵ_s^* , respectively. For single spherical particle, ϵ_{eq}^* represents the equivalent complex permittivity of the particle and shell.

Clausius–Mossotti factor (CM) helps in determining DEP force considering its sign and magnitude, for positive CM factor, the particles are directed towards the electric field and if it is negative, it is directed away from the field. The sign of CM factor can be determined by the medium at a defined frequency. Cross-over frequency is observed when there is a transition taking place between n-DEP and p-DEP at a point as a result of equal complex permittivity of particle and the medium. DEP force acts according to the size, dielectric properties of particle, but it may differ for nano-particles and microparticles. The n-DEP force required for particle separation is calculated and on applying, it can be observed that the particles are affected by n-DEP force and are oriented based on their diameters and hence the separation of particles.

In order to effectively conduct the simulation in microfluidic device, the flow field has a critical role in it. The governing equation for fluid flow in an electric field is associated with the Stokes equation. The flow can be modelled with a parabolic profile, known as Poiseuille flow in case of n-DEP [14] and these values depend on the geometry of the channel along with the parameters of particles. Though the fluid flow is addressed with hydrodynamic lift force associated with particles near the surface [14], it is administered entirely by Stokes and continuity equation, which is expressed as,

$$\nabla u = 0 \quad (9)$$

$$\rho(u \cdot \nabla u)\mu = -\nabla \rho + \mu \nabla^2 u + (\rho - \rho_0)g \quad (10)$$

where, ρ denotes the density, μ denotes viscosity of the liquid and ρ_0 is the fluid density. Fluid flow boundary conditions can be defined accordingly as required for the model design. The velocity at walls and electrodes is found to be zero for the uniform inlet velocity. Particle trajectory is also considered as a result of the dipole moment of particle in the presence of electric field and its strength. The trajectory is computed using integration of velocities of particular particle over time, and these computations require a suitable computational software such as FEM, which enable us to compute and move further with the numerical analysis and design.

3 DEP Modelling

DEP force is produced on applying AC field in the microchannel devices, as it can provide us with no slip boundary condition. The electrodes design, particle size, channel dimensions, inlet and outlet designs, dielectric properties of fluid and particle are associated with the separation of blood cells in the applied non-uniform electric field. The process of sorting the cells based on particle size is carried out by designing suitable electrode with the flow channel. There have been several studies based on the behaviour of particles to the different electrode design. The planar electrodes

are commonly used in models. In order to initialize the particle movement in the flow channel, the electrodes are applied with alternating polarity of voltage. There have been many studies related to this type of electrode design. Paper [15] used this electrodes design type and obtained the separation of RBC, WBC and platelets cells with voltage of 4 V. Paper [12] presented the separation of RBC and platelets cells with applied voltage of 5 V. In this work, the effect of applied voltage, shape and size of the electrode is studied on the particle separation. The dielectric properties and numerical values of RBC, WBC and platelets are taken from the previous studies which is considered as standard values for the particle separation in microfluidic device.

3.1 Modelling Design

As per previous studies, the throughput of the microfluidic model is increased on designing electrodes in the form of an array at the inner side of microchannel with the electric field applied. Inlet and outlet with appropriate angle elevation are considered as an important factor along with size and dielectric properties of particle. There have been comparative studies regarding this, where modelling of microfluidic device is formulated and simulated. Taking previously modelled studies into consideration, particularly, the model designed by Ali [15], as shown in Fig. 1, is used as the reference model for designing the current paper designs. The modifications are made to the electrode design concerning their size, shape and dimensions. The table shown in Fig. 2 shows the numerical values of dielectric properties of blood cells and fluid.

Here, we are intended to use a smaller number of electrodes to obtain higher efficiency with low voltage. In this work, two different model designs are considered. The design has changes in their inlet and outlets, number of electrodes configured and dimensions of the same are compared to each other. Figures 3 and 4 show the designs considered in this work.

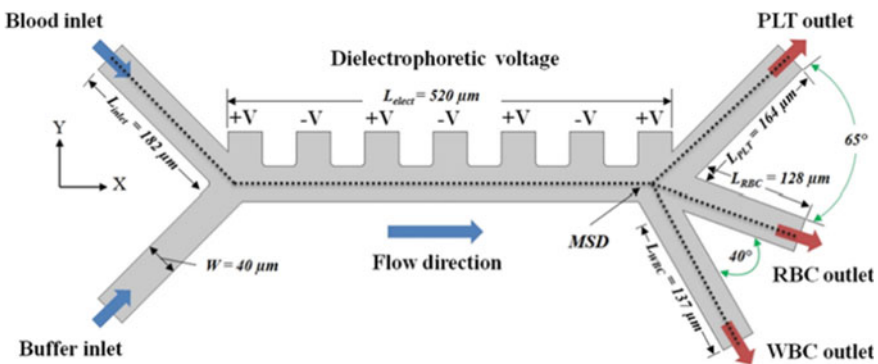


Fig.1 Reference model design

Dielectric Property	RBC	Platelets	WBC	Fluid
Diameter of Particle (μm)	7	2	12	--
Conductivity (S/m)	0.31	0.25	0.65	0.055
Permittivity (ϵ)	59	50	60	80
Shell Conductivity σ_s (S/m)	1×10^{-6}	1×10^{-6}	27.4×10^{-6}	--
Shell Permittivity (ϵ'_s)	4.44	6.8	6	--
Shell Thickness (μm)	0.009	0.008	0.007	--

Fig. 2 Blood cells and fluid dielectric properties

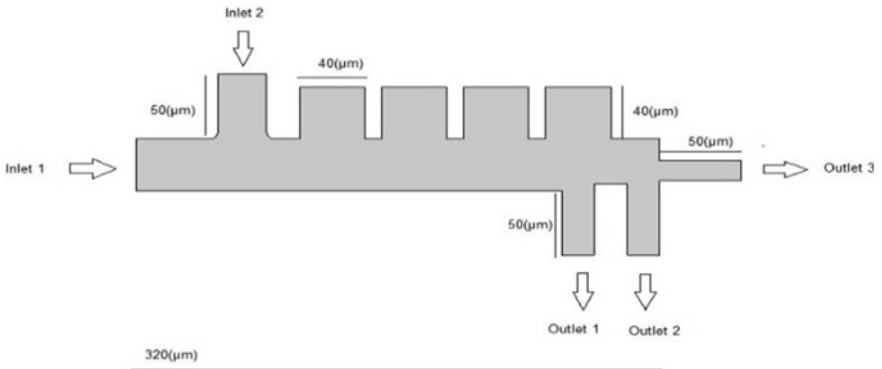


Fig. 3 Model design 1

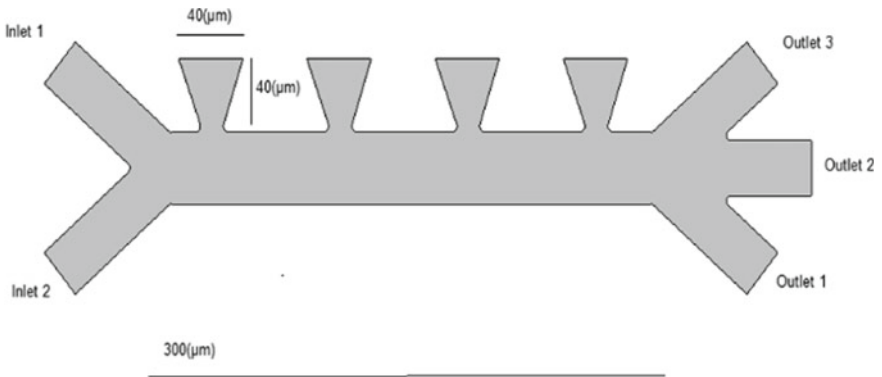


Fig. 4 Model design 2

Figure 3 is modelled with dimensions of channel length = 240 μm , width = 40 μm . Two inlets are designed perpendicular to each other. The upper side of the channel consist of electrode array, which is planar, with dimensions of height = 40 μm and width = 40 μm . Here, the dimensions are considered as same as electrode dimensions used in paper [15] that is in Fig. 1. The dimensions of electrode are kept same, whereas the number of electrodes is reduced to half appropriately and three outlets are designed. The particle deflections are different for applied potential because of their vivid difference in their size and conductivities. The blood fluid is made to flow through inlet2, buffer in inlet1 and exits through outlets designed to get the RBC, WBC and platelets separately.

Figure 4 is modelled with differently shaped electrode than planar electrodes. The main channel length is of 300 μm and width of 40 μm . Inlets1 is angled at 140° to the channel, inlet2 is at mirror angle to inlet1, but in the opposite direction and three outlets are at the other end of the channel. Here, outlet1 is at lower end, outlet2 is in line with the channel and outlet3 at upper end. The flow rate of fluid through inlet1 and inlet2 for both designs is same which is 135 ($\mu\text{m/s}$) and 853 ($\mu\text{m/s}$), respectively. These designed models are simulated for different voltages to accomplish the separation of RBC, WBC and platelets cells with the effect of dielectrophoresis force in the microchannel of the device.

4 Simulation Results

FEM software is used to simulate the above models. Based on potential applied to the electrodes, particle follows different trajectory in the channel. The electrodes are applied with potential of alternating polarity which results in particle motion as it gets exposed to the dielectrophoresis force produced due to non-uniform electric field. The field strength inside the channel also depends on the dimension and shape of the electrode. As per model design 1, as shown in Fig. 3, which uses planar electrodes, it is observed that the high electric field strength at a given point, as compared to electrodes in the model design 2 in Fig. 4. The electrode dimension influences the electric field strength inside the channel. As the field strengths are greater, the particles in the channel experiences more forces and deflected more. Along with the field strengths, particle size will also allow particle to get exposed to different amount of forces on them and trace different trajectory in the medium. The diameter of WBC cells is greater than both RBC and platelet cells, where platelets are said to have smallest size of all three. Considering model design 1, the platelets cells which are smaller in size is deflected less as compared to RBC and WBC cells. In case of RBC and WBC cells, they deflect more from electrodes but there is no much difference between them and they move cordially in the transverse channel. The field strength in the model design 2 is low at the top end of the channel near the electrode, and hence the DEP force is also low. Even in this design, the RBC and WBC are deflected at same rate, whereas platelets are less comparatively. The time duration taken by the particles to reach the outlets after separation is also accounted in this

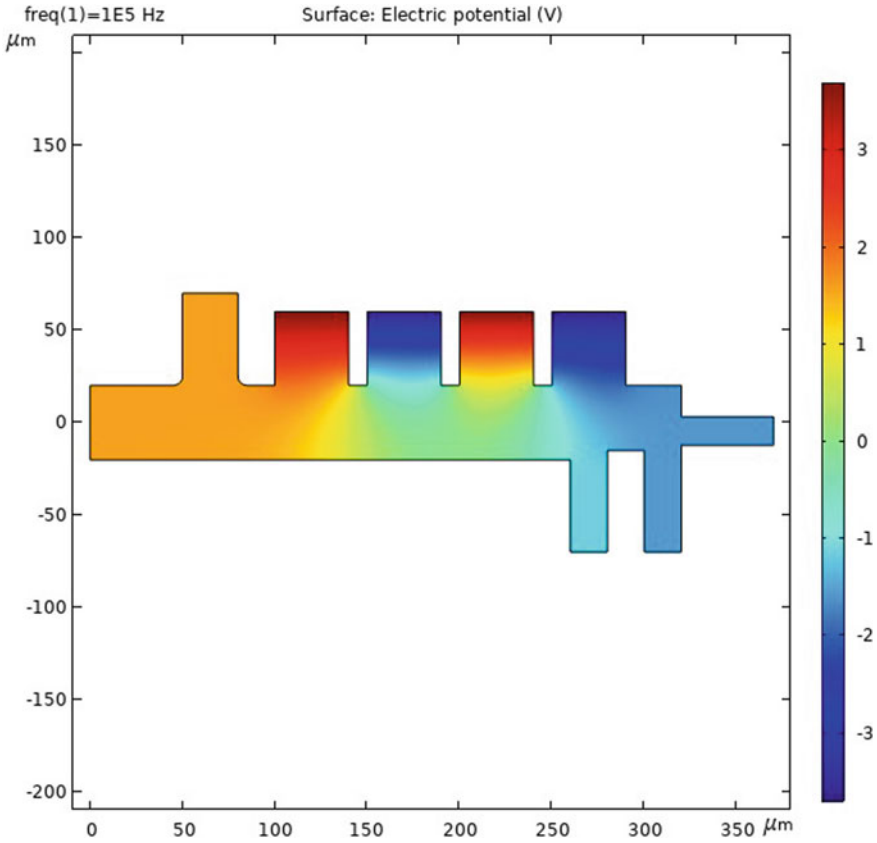


Fig. 5 Model design 1 potential distribution

computation. Model design 1 is observed to have the less time required to separate the particles as it experiences more force for a given voltage and move faster in the channel. Model design 2 requires more time to complete the separation of cells. The electric potential distribution in the model designs is obtained and pictured as given Figs. 5 and 6.

4.1 Effect of Electrode Voltage for Particle Separation

According to the theories, other than electrode design, dimension, outlet and inlet design, the voltage applied can also have an effect on dielectrophoretic force induced. In the paper [15], it is concluded that separation of RBC, WBC and platelets takes place at 4 V of voltage applied to the electrodes. Here, we made an attempt to obtain the separation of particles with different electrode shapes, with reduced electrode

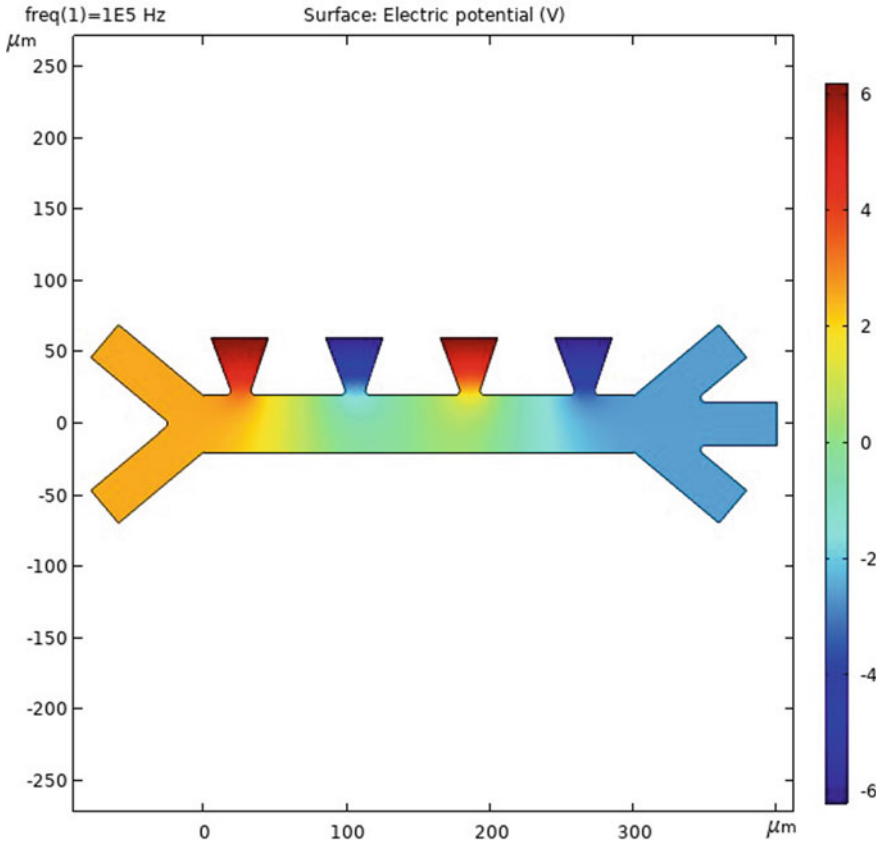


Fig. 6 Model design 2 potential distribution

number and with minimum applied voltage. Applying different voltages resulted in various outcomes. On numerous attempts, particles in model design 1 were found to separate particles at the voltage of 3.7 V and model design 2 was found to give the result at 6.2 V. When these voltages were slightly increased or decreased resulted in separation of RBC and WBC particles in the same outlet and platelets at the other outlet. Hence, separation in these models is possible only at a particular voltage and on applying other voltages, could result in no separation of particles. The below figures show the separation of blood cells at above-mentioned voltages in model design 1 and model design 2.

The electric field distribution inside the microchannel can be estimated by the Poisson equation,

$$E = \frac{\partial V}{\partial L} \tag{11}$$

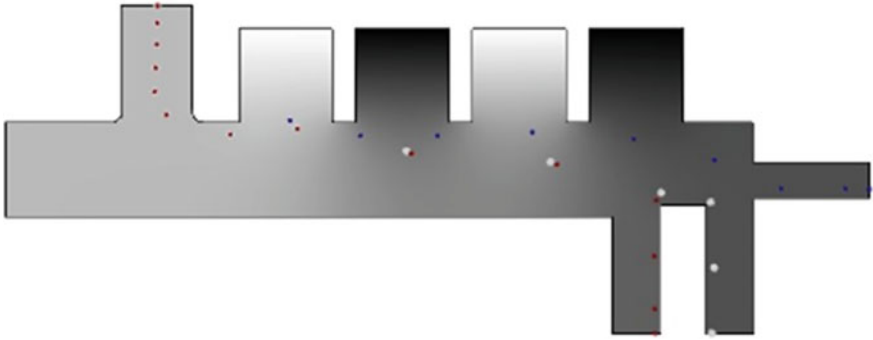


Fig. 7 Separation of RBC, WBC and platelets in model design 1

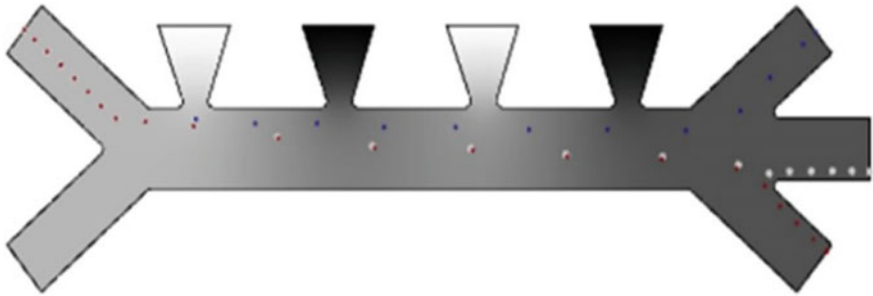


Fig. 8 Separation of RBC, WBC and platelets in model design 2

This is the most commonly used equation to determine the electric field strengths inside the channel of microfluidic device. This equation indicates that the electric field strength is inversely proportional to length and directly proportional to applied voltage. Hence, we can conclude that, changing the physical dimension and shape of the electrodes changes the field strength distribution inside the channel and hence the particle separation. Figure 9 indicates the distribution of electric field strength inside the microfluidic channel for both the model designs shown in Figs. 7 and 8.

4.2 Effect of Changing the Physical Dimensions of Electrodes

From the above results, we can conclude that the shape of electrode and its dimensions influence the field strength inside the channel. In order to provide the particle with required force, the voltage should be applied accordingly and that varies regarding the electrode dimension. This gives us the approach to conduct a further work regarding the effect of change in electrode dimensions. To carry out this, we are intended to

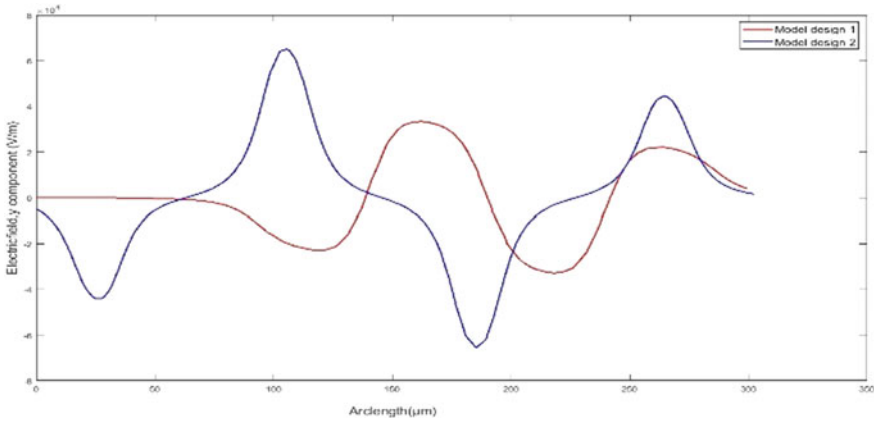


Fig. 9 Electric field distribution versus channel length for designs 1 and 2

do a comparative study considering the reference model [15]. The dimensions of electrode used in this model are of (height = 40 μm , width = 40 μm). Keeping height and width as the main factor of comparison, the simulation was conducted by reducing the height and width of electrode to its half value alternatively. The number of electrodes and length of the channel was also decreased approximately to half of their values compared to reference model. The analysis is done regarding the effect of change in width and height of electrodes.

(1) *Effect of Reducing Height by Keeping Width Same*

The above structures are modelled by reducing height of the electrode to half by keeping same width as of the reference model. On simulating the structures, it is found that structure in Fig. 10 can separate particles at voltage of 2.4 V and structure in Fig. 11 can separate at 4.3 V. From the results, it can be concluded that the voltage required for particle separation is comparatively less than the voltages required for the design as shown in Figs. 3 and 4. Figure 12 indicates the

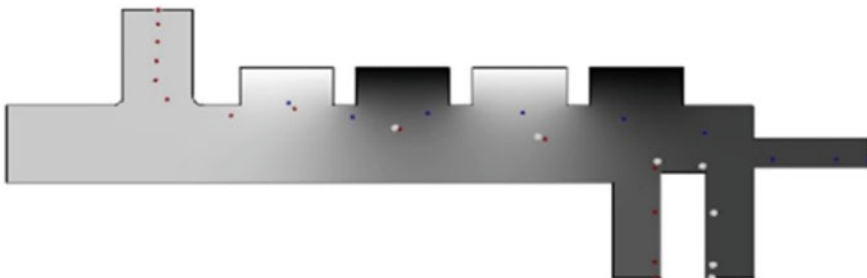


Fig. 10 Particle separation in design 1 by reducing electrode height half without changing width

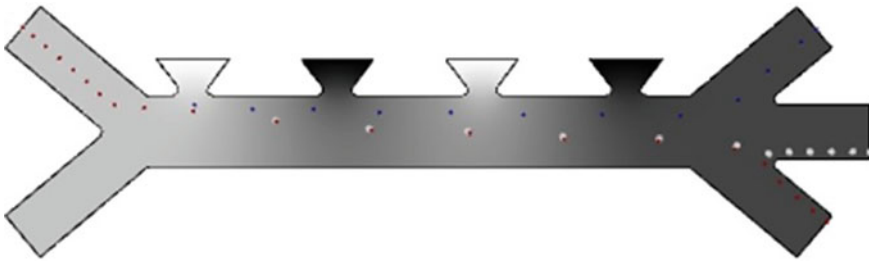


Fig. 11 Particle separation in design 2 by reducing electrode height half without changing width

distribution of electric field strength inside the microfluidic channel for both the model designs shown in Figs. 10 and 11.

(2) *Effect of Reducing Width by Keeping Height Same*

Figures 13 and 14 shows the separation of cells when the width of the electrode is reduced to its half value compared to the reference model. The above designs result in separation of particles at the voltage of 5.6 V and 7.3 V, respectively. From these results, we can conclude that when width of electrode is reduced, significantly the voltage required to drive the particles in the channel will be high compared to the structures shown in Figs. 10 and 11. Figure 15 indicates the distribution of electric field strength inside the microfluidic channel for both the model designs shown in Figs. 13 and 14.

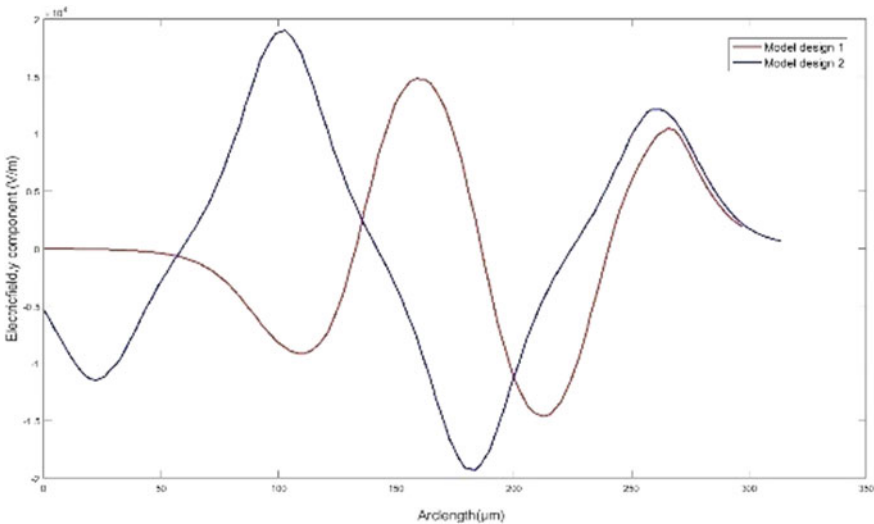


Fig. 12 Electric field distribution versus channel length for design 1 and 2 by reducing height half without changing width

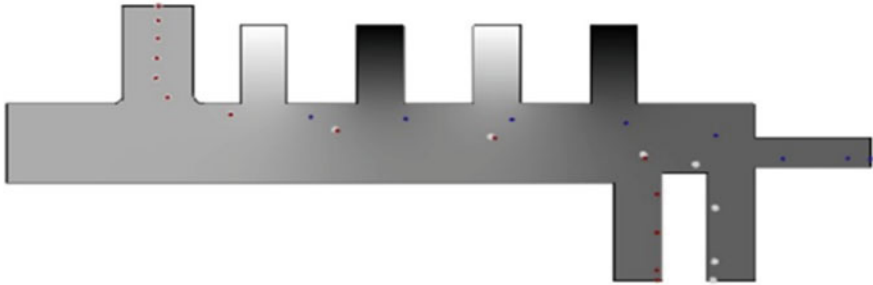


Fig. 13 Particle separation in design 1 by reducing electrode width half without changing height

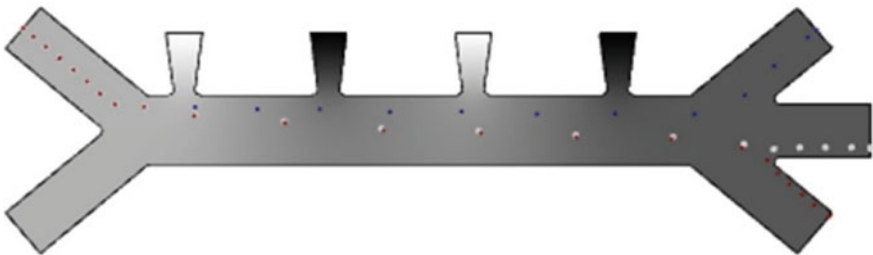


Fig. 14 Particle separation in design 2 by reducing electrode width half without changing height

On considering all the above computations, we can conclude that, when the height of the electrode is less, the possibility of particle separation with the less applied voltage can be obtained.

Based on this analysis, we can model a device for particle separation, with less electrode height to obtain the results with minimum voltage, high efficiency but with an additional electrode without changing channel length. The designs given in the below figures can be considered as the optimized model designs.

Figures 16 and 17 show optimized model designs which have an electrode dimension of height = 10 μm and width = 40 μm . The voltages at which they separate particles is found to be 1.6 V and 2.2 V, respectively. We can observe that, these model designs are comparatively efficient than all the other designs, as it requires less voltage, minimum channel length but with the increase in one extra electrode. Figure 18 indicates the distribution of electric field strength inside the microfluidic channel for both the model designs shown in Figs. 16 and 17.

The table shown in Fig. 19 gives the comparison between the number of electrodes used, voltage required to get the separation of the cells and dimensions of the electrodes. As we can observe that, this comparison table gives clear indication that the electrode physical dimension, shape, channel length and applied voltage are the important parameters that need to be considered while designing the microfluidic device for the segregation of particles.

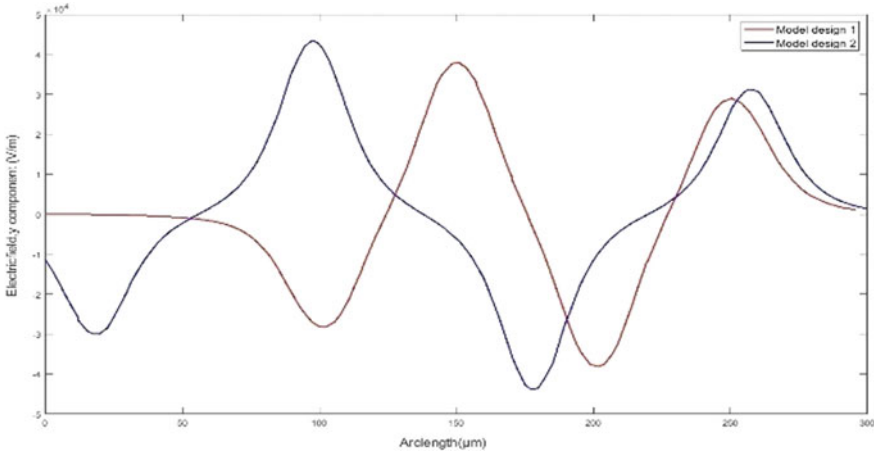


Fig. 15 Electric field distribution versus channel length for design 1 and 2 by reducing width half without changing height

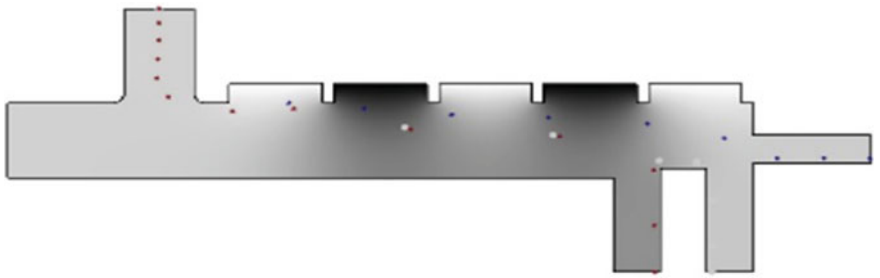


Fig. 16 Particle separation in design 1 with optimized electrode design

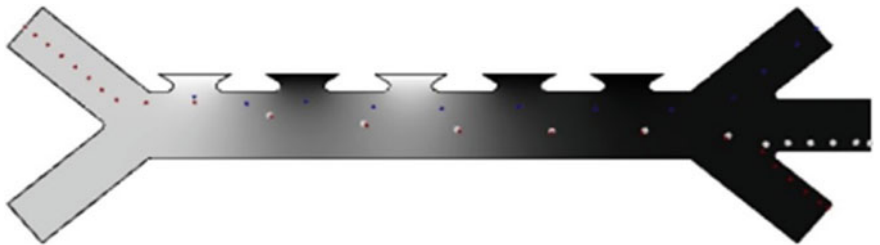


Fig. 17 Particle separation in design 2 with optimized electrode design

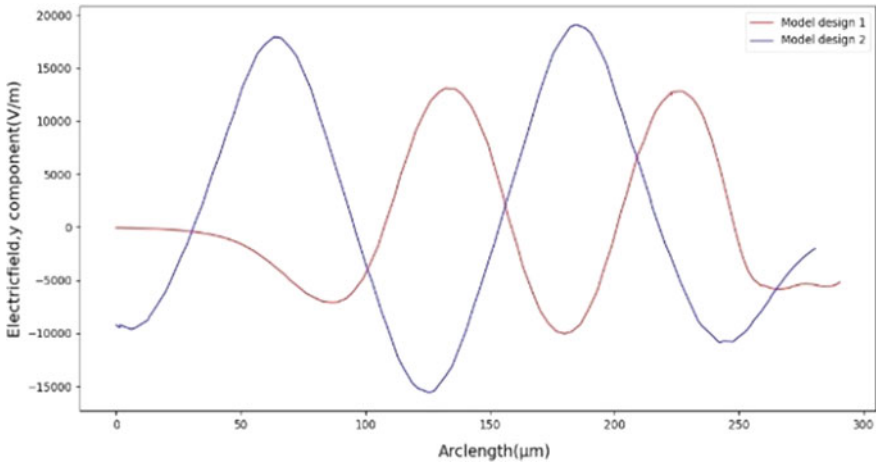


Fig. 18 Electric field distribution versus channel length for optimized design 1 and 2

Parameters	Basic Design		Reduced Height		Reduced Width		Optimized Design	
	Model Design 1	Model Design 2	Model Design 1	Model Design 2	Model Design 1	Model Design 2	Model Design 1	Model Design 2
Height of Electrode (μm)	40	40	20	20	40	40	10	10
Width of electrode (μm)	40	40	40	40	20	20	40	40
Channel length (μm)	240	300	240	300	240	300	240	300
Number of electrodes	4	4	4	4	4	4	5	5
Voltage required for separation (V)	3.7	6.2	2.4	4.3	5.6	7.3	1.6	2.2

Fig. 19 Comparison of model designs

5 Conclusion

In this paper, the separation of RBC, WBC and platelets cells using FEM method is carried out based on model designs. The model designs were simulated and analysed with the help of equations related to the concept used in the process of dielectrophoretic force applied to the microfluidic device, flow rates of the fluid, dielectric properties of fluid and cells. The separation is based on the size of the particles under consideration. The simulation is carried on two different model designs with different electrode dimensions, shape and array length. Along with that, the changes are included in their inlet and outlet positions for the particle separation. The results observed concludes that the separation of cells is accomplished at low voltages when the height of the electrode was reduced with the increased width. The voltage of 1.6 and 2.2 V are the optimized voltage values for the model design

1 and model design 2, respectively, for the segregation of particles with one extra electrode. Since the channel length is kept same, increase in one electrode does not change the device physical dimension. Among both the designs, model design 1 is said to have completed the separation at very low voltage and with same physical device dimension, hence, these models can be used in the applications, where the minimum voltage, low power and less dimension with greater efficiency are the major constraints in the designing of microfluidic devices.

References

1. L. Dean, L. Dean, *Blood Groups and Red Cell Antigens*, vol. 2. NCBI Bethesda, Md, USA, 2005
2. Y. Gong, X. Li, B. Yang, W. Ying, D. Li, Y. Zhang, S. Dai, Y. Cai, J. Wang, F. He et al., Different immunoaffinity fractionation strategies to characterize the human plasma proteome. *J. Proteome Res.* **5**(6), 1379–1387 (2006)
3. N. Bansal, M. Jindal, N.D. Gupta, P. Shukla et al., Clinical guidelines for periodontal management of idiopathic thrombocytopenic purpura: current considerations. *Int. J. Oral Health Sci.* **7**(1), 30 (2017)
4. J.B. Briere, Essential thrombocythemia. *Orphanet J. Rare Dis.* **2**(1), 3 (2007)
5. L.T. Goodnough, M.E. Brecher, M.H. Kanter, J.P. AuBuchon, Transfusion medicine blood transfusion. *N. Engl. J. Med.* **340**(6), 438–447 (1999)
6. P. Tartert, Blood transfusion and infectious complications following colorectal cancer surgery. *Br. J. Surg.* **75**(8), 789–792 (1988)
7. X.J. Li, Y. Zhou, *Microfluidic Devices for Biomedical Applications* (Elsevier, 2013)
8. A. Webster, J. Greenman, S.J. Haswell, Development of microfluidic devices for biomedical and clinical application. *J. Chem. Technol. Biotechnol.* **86**(1), 10–17 (2011)
9. B. Çetin, D. Li, Dielectrophoresis in microfluidics technology. *Electrophoresis* **32**(18), 2410–2427 (2011)
10. T.-S. Leu, Z.-F. Liao, Separating plasma and blood cells by dielectrophoresis in microfluidic chips. *Int. J. Mod. Phys. Conf. Ser.* **19**, 185–189 (2012)
11. J. Cebricos, R. Hoptowitz, S. Jun, Separation of escherichia coli k12 from contaminated tap water using a single-stage, continuous flow dielectrophoresis (dep) device. *LWT* **80**, 185–192 (2017)
12. Y. Zhang, X. Chen, Dielectrophoretic microfluidic device for separation of red blood cells and platelets: a model-based study. *J. Braz. Soc. Mech. Sci. Eng.* **42**(2), 89 (2020)
13. P.R. Gascoyne, J. Vykoukal, Particle separation by dielectrophoresis. *Electrophoresis* **23**(13), 1973 (2002)
14. S. Williams, Dielectrophoretic motion of particles and cells (2008)
15. H. Ali, C.W. Park, Numerical study on the complete blood cell sorting using particle tracing and dielectrophoresis in a microfluidic device. *Korea-Australia Rheol. J.* **28**(4), 327–339 (2016)

Design of a Bipedal Robot



Chandana N. Aithal, P. Ishwarya, S. Sneha, C. N. Yashvardhan,
and K. V. Suresh

1 Introduction

The influence of machines in assisting humans and increasing human capabilities is tremendous. One such class of machines is robots which can sense the environment and perform a specified task with computational intelligence. Robots play a major role in assisting people when they are vulnerable. They also enhance the self-confidence of elderly citizens and specially abled individuals by assisting them in their daily life. Robots with bipedal abilities prove to be much more effective than the usual toddlers.

Bipedal robots have been one of the exciting areas of research in robotics. They can be designed to perform tasks in the same way as that of humans. [1] Many models have been developed on these robots based on the control, design, walking pattern generation, etc. Few notable models developed are passive dynamic walker [2], model based on ZMP [3–5] and model based on the inverted pendulum (IPM) [6–8], etc. Passive dynamic walker, a model developed mainly utilizing the mechanics was able to walk without actuation, it completely relies on the slope of the walking surface and gravitational force for locomotion. Zero moment point-based model utilizes the concept of moment of inertia acting on the robot and the effect of reaction force from the walking surface. In inverted pendulum models, the dynamics of the robot is represented as a single inverted pendulum. In contrast to the passive dynamic walker, ZMP- and IPM-based models are actuation-based models, where high torque servo motors are used for actuation. The concept of zero moment point for balancing of the robot is used in this proposed model.

C. N. Aithal · P. Ishwarya · S. Sneha · C. N. Yashvardhan (✉) · K. V. Suresh
Department of Electronics and Communication, Siddaganga Institute of Technology, Tumkur,
India

K. V. Suresh
e-mail: sureshkvst@sit.ac.in

This paper is organized as follows: Sect. 2 gives the literature review, the system description is discussed in Sects. 3, 4 provides model specification and experimental results and the conclusion is given in Sect. 5.

2 Working Principle of Bipedal Robots

2.1 Bipedal Gait Cycle

The sequence of events occurring during the locomotion of a bipedal robot is called the gait cycle. It starts from lifting one leg and finishes when the same leg reaches the ground.

1. Components of Gait cycle

Stance phase: This is a phase in which the robot's foot stays in contact with the ground. There are five stance phase movements

- Initial contact: Here the heel is the only part of the robot that is making contact with the surface.
- Loading response: The entire weight concentration is transferred to the reference leg here.
- Mid stance: This movement requires body weight balance on the reference foot.
- Terminal stance: Here the toe is still in contact with the floor and the heel is raised.
- Toe off: Here the reference foot's toe is raised off the floor. This results in the beginning of the gait cycle's swing process.

Swing Phase: This is a phase in the gait cycle where the reference foot swings in the air by losing its contact with the ground.

2. Support: As shown in Fig. 1, there are two support phases,

- Single support phase (SSP): Only one foot is in contact with the ground in single support phase. The other foot is in the swing phase.
- Double support phase (DSP): Both feet are in contact with the ground in the double support phase.

2.2 Degrees of Freedom (DOF)

In the context of mechanics, DOFs are unique and specified modes in which a mechanical device or system can move. In human beings, the leg and foot have a total of 30 DOFs.

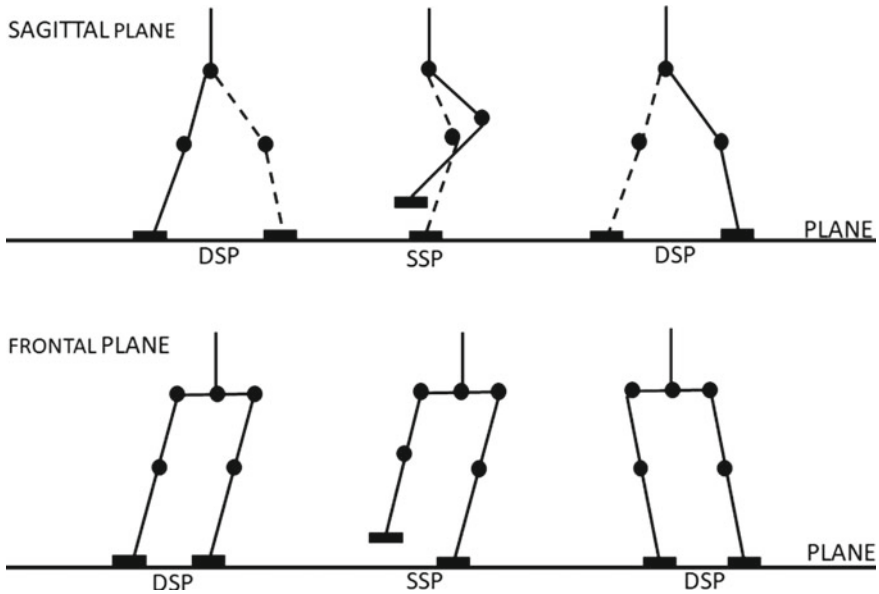


Fig. 1 Double support phase and single support phase of bipedal robot

2.3 Implementation of Degrees of Freedom in Bipedal Robots [9]

- Two degrees of freedom in each leg (1 DOF at hip joint and 1 DOF at ankle) can preserve frontal plane stability but not sagittal plane stability. An extra 1 degree of freedom in each leg at the knee joint would be capable of maintaining equilibrium in the frontal plane.
- As shown in Fig. 2, to have a three-dimensional modification of the supporting surface, it is appropriate to have two additional DOFs in the sagittal plane that totally become 4 DOFs (2 DOFs at the hip joint and 2 DOFs at the ankle).
- We can achieve six DOF by adding a vertical axis at the hip joint perpendicular to the other axis as shown in Fig. 3. It allows the leg to rotate and walk in a circle. Without this, we can only have forward and backward locomotion. Therefore, allowing the robot to change its walking direction is important.

2.4 Bipedal Robots Based on Zero Moment Point

The type of surface and the reaction force exerted by the surface on robots has a big influence on the dynamics of robots. Earth’s gravity, acceleration and deceleration of the robot’s body collectively form the inertial forces acting on the body. It is necessary to balance these forces acting on the robot to have a balanced locomotion.

Fig. 2 Four DOF robot structure

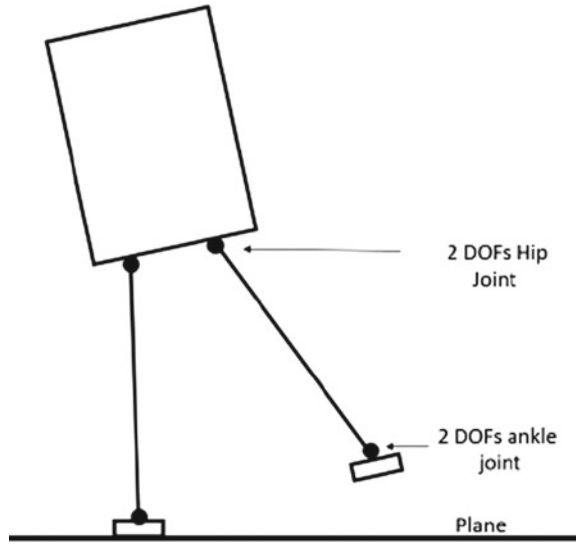
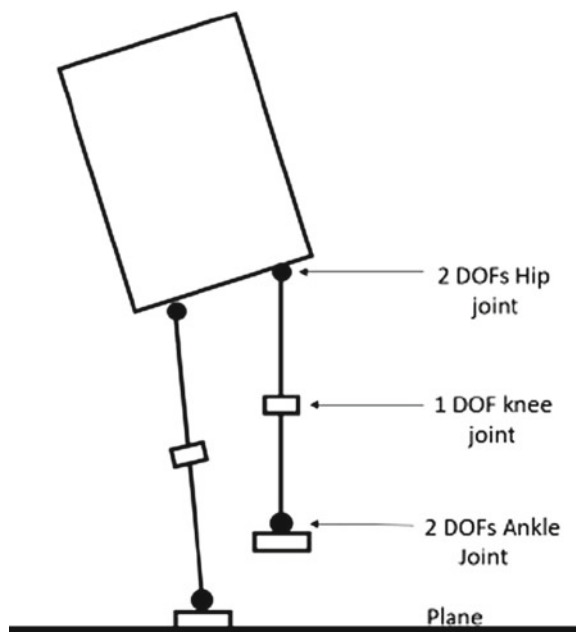


Fig. 3 Six DOF robot structure



If these forces are not balanced, the robot will eventually fall on the surface. Zero moment point is a point on the surface where the total resultant inertial force acting on a body sums up to zero. If this point occurs inside the supporting polygon created by the robot's foot, then the robot will not fall on smooth terrain, i.e. the swing

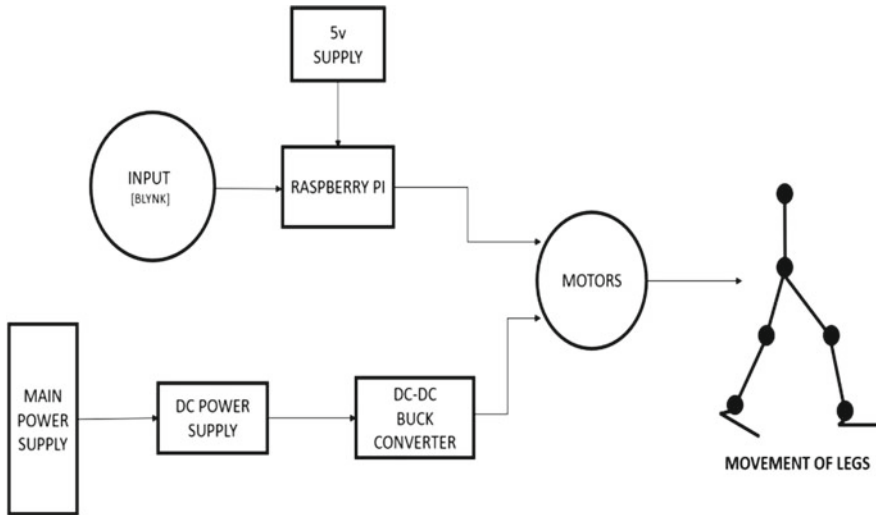


Fig. 4 Proposed block diagram of the system

foot should reside inside the supporting convex polygon once the swing is done. ZMP has been widely used by many bipedal researchers. Although finding ZMP is computationally expensive, the concept of supporting polygon stability can be implemented for achieving basic balanced motion of a bipedal robot.

3 System Description

3.1 System Block Diagram

The block diagram of the system is shown in Fig. 4. It consists of a source of input which is an android app called Blynk, a processor, i.e. Raspberry Pi, main power supply, a DC power supply, a DC-DC buck converter and a set of motors.

3.2 System Hardware

As shown in Fig. 5, the processor used in this proposed model is a Raspberry Pi. It is programmed to control six high torque motors which act as the 6 joints of the two legs, i.e. there are three motors in each leg making it a 6 DOF bipedal structure in total. In each leg, one motor will act as a hip joint, another will act as a knee joint and the third motor will act as an ankle joint. The motors are driven by a 12 V 9 A

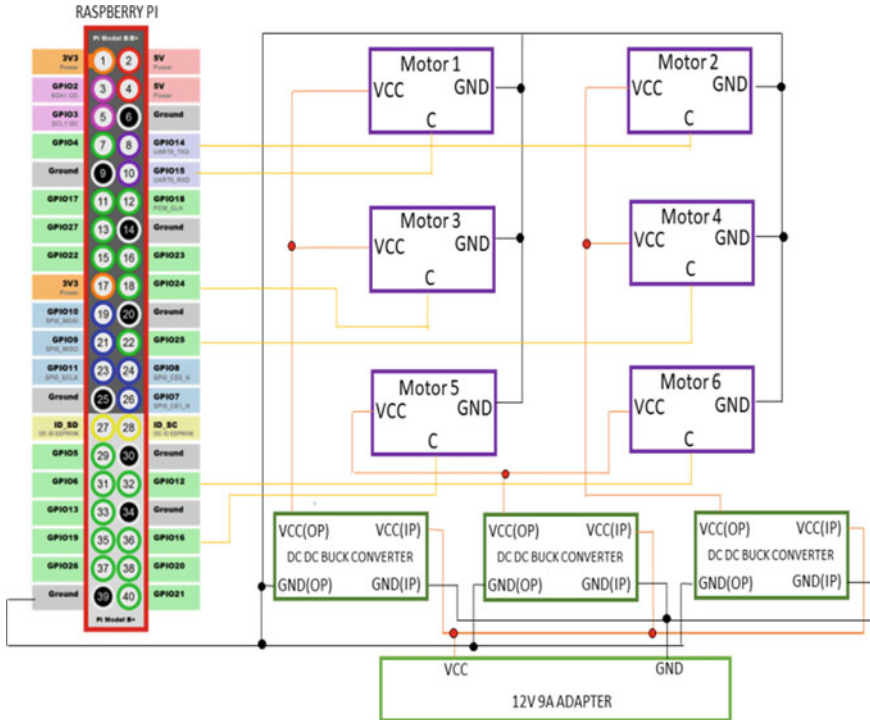


Fig. 5 Circuit diagram of bipedal robot

adapter connected through a DC-DC buck converter which steps down the voltage as required.

The components required to build the structure is listed below:

- Aluminium sheet: The mechanical structure of the bipedal robot is built using aluminium sheet of 3 mm thickness. Aluminium sheets are used in particular because it is strong and also because of its lightweight, the torque required to move the legs will be less.
- Raspberry Pi: Raspberry Pi is a micro-computer with wireless LAN and Bluetooth connectivity. It is a small, powerful and lightweight advanced RISC machine (ARM)-based computer which is capable of performing different operations. It has Quad Core 1.2 GHz Broadcom BCM2837 64bit CPU and 40 general purpose input output pins. It is the heart of a bipedal system which processes the Python code dumbered in it and sends pulse width modulated signal through the general purpose input–output pins to the motors so that the motor rotates to give human-like walking motion to the structure [10].
- Servo Motor: A servo motor is a linear or rotary actuator that provides fast and precise position control for closed loop position control applications. It works on the principle of pulse width modulation (PWM). The duration of pulse applied

to its control pin decides its angle of rotation as shown in Fig. 6. To control the angle of rotation a pulse with minimum pulse width and maximum pulse width is required which repeats itself at a specific rate. Servo motor can turn 90° in either direction from its neutral position. The control pin of the motor should be fed with a signal of frequency 50 Hz because the period at which servo motor expects a pulse is 20 ms, that is the frequency of pulse should be 50 pulses per second or 50 Hz. The duty cycle defines the proportion of ON time to the regular interval or period of time. Angle of rotations between 0 and 180° can be achieved between the duty cycle 2.5–12.5%. The duty cycle (DC) of PWM can be calculated as follows [11].

$$DC = \frac{\text{Angle}}{\frac{180}{12.5-2.5}} + 2.5 \tag{1}$$

$$DC = \frac{\text{Angle}}{18} + 2.5 \tag{2}$$

For 0°

$$DC = \frac{0}{18} + 2.5 = 2.5\% \tag{3}$$

For 90°

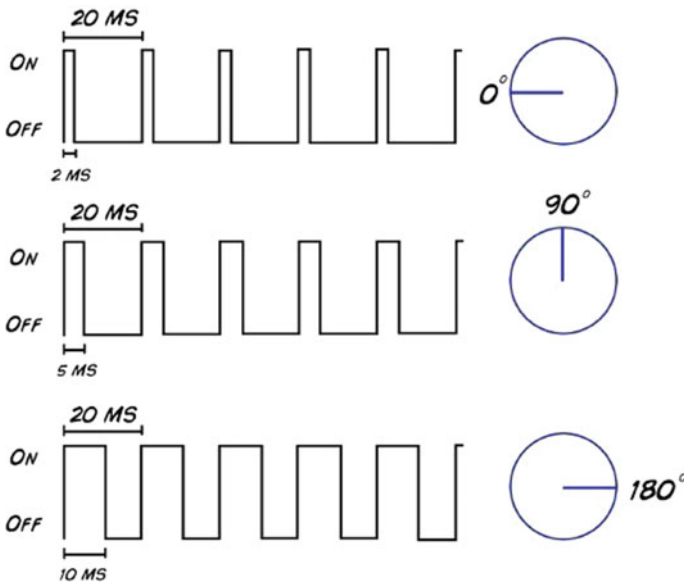


Fig. 6 Timing diagram of control signal for different angle of rotation

$$DC = \frac{90}{18} + 2.5 = 7.5\% \quad (4)$$

For 180°

$$DC = \frac{180}{18} + 2.5 = 12.5\% \quad (5)$$

In the proposed work, two models of servo motors are used based on the torque requirement. They are:

- RKI1206 servo motor: It is a high torque motor with a stall torque of 16kgcm and having a dimension of 40 mm × 20 mm × 40.5 mm. It has a stall current of 1.5 A at 6.0 V and operating voltage from 4.8 to 7.2 V.
- TowerPro MG995 servo motor: It has comparatively less torque of 10kgcm as compared to RKI1206 and has a dimension of 40.7 mm × 19.7 mm × 42.9 mm. It has a stall current of 1.2 A at 6.0 V and operating voltage from 4.8 to 7.2 V.
- DC-DC buck converter: These are used where the DC output voltage must be lower than the DC input voltage in switched mode power supply circuits. LM2596 is used in this system as a buck converter which can drive a 3A load. It has good line and load regulation. Buck converters are available in fixed output voltages of 3.3, 5 and 12 V. An adjustable output version is also available in the market. The LM2596 series operates at a 150 kHz switching frequency, enabling filter components of smaller sizes than would be expected with lower frequency switching regulators.

3.3 Power Requirement

As shown in Fig. 7, two RKI1206 are used which acts as the hip joint of bipedal legs. Since the motors at the hip joint need to lift the weight of the whole leg, high torque motors are used. The remaining four joints, i.e. two knee joints and two angle joints need to lift only a part leg the motor TowerPro MG995 of comparatively less torque is used. In case if the load is maximum on the motors, then the power required is

RKI1206 (2 motors)

- Voltage: 6 V
- Current: 1.5 A + 1.5 A = 3.0 A

TowerPro MG995 (4 motors)

- Voltage: 6 V
- Current: 1.2 A + 1.2 A + 1.2 A + 1.2 A = 4.8 A

Hence, in total, to drive all the motors, the sufficient voltage required is 6 V and 7.8 A (3.0 A + 4.8 A). Therefore, an adapter with standard ratings of 12 V and 9 A is used which is stepped down using a DC-DC buck converter.

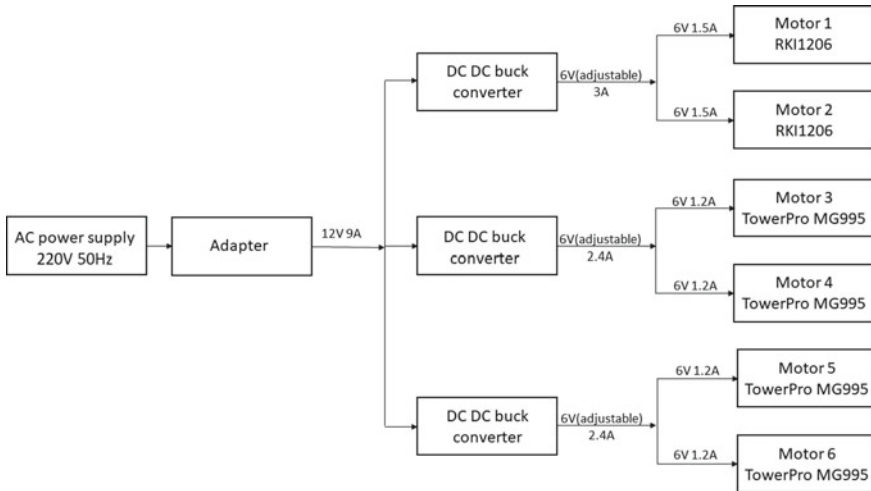


Fig. 7 Power requirement calculation

3.4 System Software

- Blynk: Blynk is a hardware-agnostic IoT platform with mobile applications, private cloud, system management and machine learning [12]. Blynk was designed to control hardware remotely and it can display as well as store the sensor data which makes it highly suitable to implement Internet of things. The platform has three main components:
 - Blynk app—enables users to create a personalized project application using different widgets.
 - Blynk server—It is an open-source network where all the communication between hardware and smartphone takes place with the help of Blynk cloud or a private Blynk server handled locally. It can control various hardware devices easily and can also be launched on a Raspberry Pi.
 - Blynk libraries—It handles all the data exchange and connection routines between the hardware, Blynk cloud and the project application created in the smartphone. It accommodates a library set of popular programming languages like C++, JavaScript, Python, Node.js and others.
- Python: Python 3 is the language used to code in Raspberry Pi. Python is a high level, interpreted and a general-purpose programming language. The usage of significant whitespaces in Python’s design philosophy emphasizes code readability. It becomes very easy for the programmers to write clear and logical code for all kinds of projects due to the language constructs and its object-oriented approach [13].
- Raspbian operating system: Raspbian is an open-source operating system designed for the Raspberry Pi hardware based on the Debian operating system.

The operating system is a user-computer hardware interface that executes all the functions that make the Raspberry Pi run. It comes with over 35,000 packages and pre-compiled applications for simple installation on your Raspberry Pi in zip file form [12].

4 Results

4.1 Models Specifications

The mechanical structure of the bipedal robot is as shown in Fig. 8. The leg height is based on a typical human leg proportion [14]. Naturally in human leg, femur (thigh bone)-to-tibia (shin bone) ratio is approximately between 31:36 and 34:35, i.e. between 0.86 and 0.97 which helps humans for a balanced walk. Taking that into consideration for building the robot, the leg ratio of the structure is 12:13.5, i.e. 0.88 [15].

The length of the robot was determined on the basis of torque calculation, keeping cost reduction as one of the prime factors during the proposed work. Motors 1 and 2 have torque ratings of 16 kg cm, whereas motors 3, 4, 5 and 6 have torque ratings of 10 kg cm. Motors 1 and 2 can lift a maximum weight of 0.8 kg at a distance of 20 cm, similarly motors 3, 4, 5 and 6 can lift up to 1 kg at a distance of 10 cm. So,

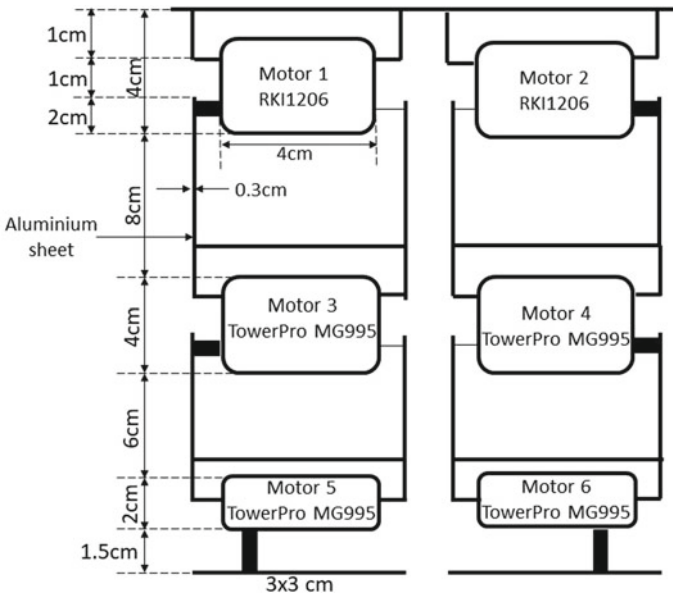


Fig. 8 Mechanical structure

the total length of the robot was made as 25 cm, dividing the robot leg on the basis of leg ratio as mentioned above.

4.2 Simulation Results

Based on the angles obtained after the extensive survey done on different research papers, the angular rotation required by each motor is determined and fixed. The sequence of joint and their respective angles to move forward are:

- (1) Right Ankle, Left Ankle: $0^\circ \rightarrow 18^\circ$
- (2) Left Hip, Left Knee: $0^\circ \rightarrow -60^\circ$
- (3) Right Ankle, Left Ankle: $18^\circ \rightarrow -35^\circ$
- (4) Right Hip, Right Knee: $0^\circ \rightarrow 60^\circ$
- (5) Right Ankle, Left Ankle: $-35^\circ \rightarrow -18^\circ$
- (6) Right Hip, Right Knee: $60^\circ \rightarrow -60^\circ$
- (7) Left Hip, Left Knee: $-60^\circ \rightarrow +60^\circ$

The simulation of the proposed bipedal structure was done on the Tinkercad platform. Snapshots from the simulation are as shown Figs. 9, 10 and 11.

In Fig. 9, all the motors are in the initial position, i.e. both the legs are resting on the ground. In Fig. 10, the angle of rotation by the motors during the stance phase is shown and in Fig. 11, all the motors rotate back to initial position which is the terminal stance of bipedal walking.

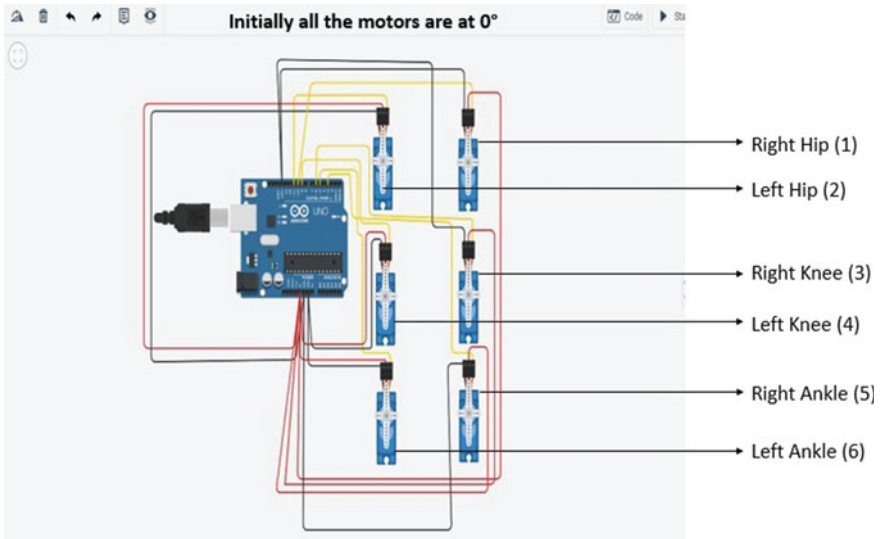


Fig. 9 Initial position

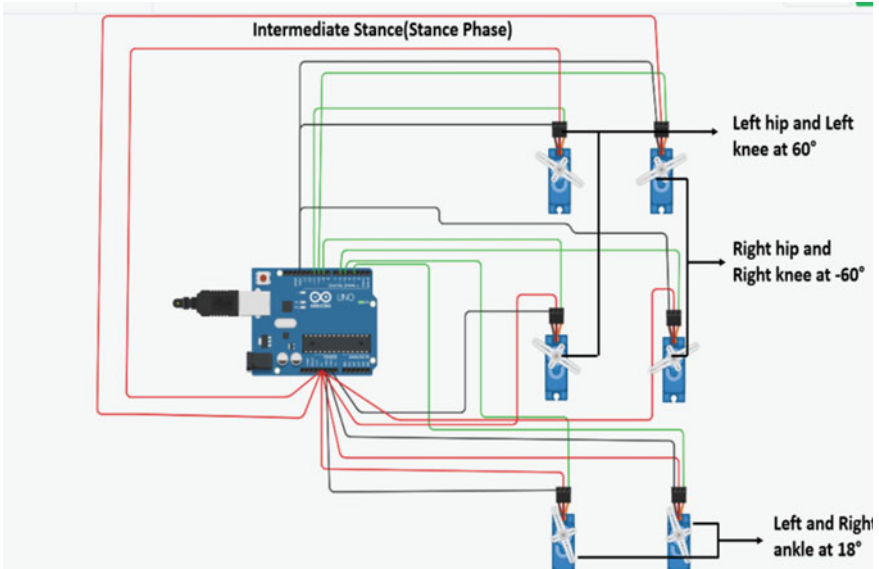


Fig. 10 Stance phase

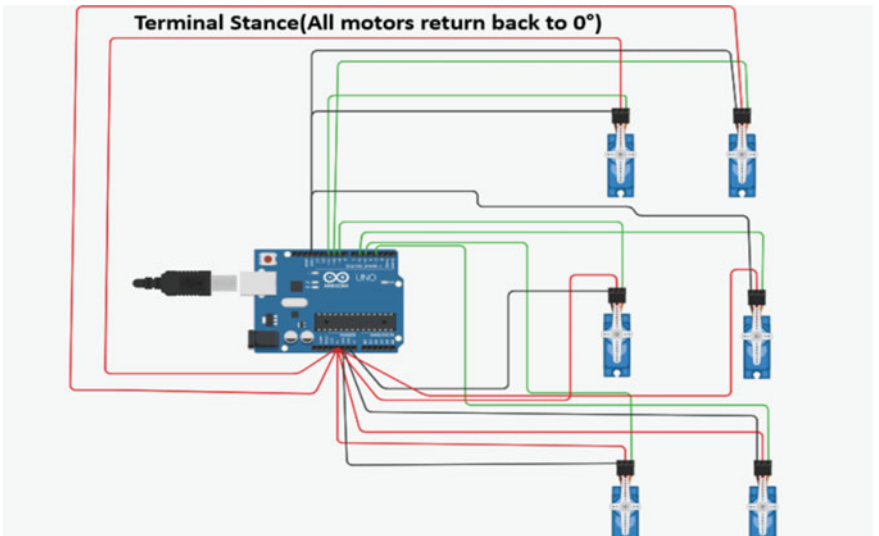


Fig. 11 Terminal stance

5 Conclusion

The model of a bipedal robot having a five link and 6 DOF structure has been developed. To stabilize the robot, ZMP approach has been used. The simulation of the angles being set to the required angles, leading to a balanced motion is done using a free, online 3D modelling platform called Tinkercad. A mobile application has been developed using Blynk. This application allows users to control the robot using simple commands like MOVE and STOP.

References

1. D.N. Nenchev, A. Konno, T. Tsujita, *Humanoid Robots* (Elsevier Inc., 2019)
2. T. McGeer, Passive dynamic walking. *Int. J. Robot. Res.* (1990)
3. M. Vukobratovic, B. Borovac, Zero-moment point- thirty-five years of its life. *Int. J. Human. Robot.* (2004)
4. K. Hirai, M. Hirose, T.T. Yuji Haikawa, The development of honda humanoid robot, in *IEEE International Conference on Robotics and Automation*, 1998
5. E.J. Dijkstraab, E.M.Gutierrez-Farewikabc, Computation of ground reaction force using zero moment point. *J. Biomech.* (2015)
6. J-Y. Kim, I.-W. Park, J.-H. Oh, Walking control algorithm of biped humanoid robot on uneven and inclined floor. *J. Intell. Robot. Syst.* (2007)
7. S. Kajita, F. Kanehiro et al., The 3d linear inverted pendulum mode a simple modeling for a biped walking pattern generation, in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2001
8. H. Bae, Jun-Ho, Biped robot state estimation using compliant inverted pendulum model, in *Robotics and Autonomous Systems*, 2018
9. J. Fred, R. Sias, Y.F. Zheng, How many degrees-of-freedom does a biped need, in *IEEE International Workshop on Intelligent Robots and Systems*, 1990
10. Available <https://www.raspberrypi.org/documentation/> [Online]
11. Available <https://circuitdigest.com/article/servo-motor-basics> [Online]
12. Available <https://docs.blynk.cc/hardware-set-ups-raspberry-pi> [Online]
13. J. Zelle, *Python Programming: An Introduction to Computer Science* (2010)
14. A.S. Kumar, A.G. Krishnan, A. Sridhar, N. Kiruthika, N.K. Prakash. Design and fabrication of bipedal robot. In: *Fifth IEEE International Conference on Computing, Communications and Networking Technologies (ICCCNT)* (2014)
15. Available:<https://royalsocietypublishing.org/doi/full/https://doi.org/10.1098/rsos.170399> [Online]

Amalgamation of Neural Network and Genetic Algorithm for Efficient Workload Prediction in Data Center



R. Pushpalatha and B. Ramesh

1 Introduction

Cloud computing provides every organizations with a plan of action for more effective decision making; expedite revolution, also multitude of other ease. But cloud is further called for attentiveness and problem solving, time observing for management of cost. A present day study on cloud amid IT and Finance executive begin that dearth of tediousness is the sole substantial cloud cost management agony. Respondents mentioned the call for clarified low billing and more precise and predictable budget forecasting. Resource forecasting has become consequential for cost management. We know data centers are spending huge amount for virtual machine migration. The numbers of servers are increasing tremendously. Currently, there are 100 million servers globally where Google and Microsoft own maximum of it [1]. The unnecessary migration can be avoided with workload prediction and turn out the obsolete, idle servers and turn down power consumption. Server consolidation and virtualization are contributing effectively for resource management [1].

The load balancing, reducing the power consumptions by servers, and resource management are few myriad of doing forecasting of resource. Many researchers have proposed various algorithms for prediction, but still reaching 10% accuracy is still challenging. Therefore, it is imperative to accurately predict the workload for resource requirement fulfilling in datacenter. Prediction can be done for long term (1 h) and also for short term (5 min). The accuracy of prediction will be high in long-term prediction.

R. Pushpalatha (✉)

Department of Computer Science Engineering, VTU Center for PG Studies (Visvesvaraya Technological University, Belagavi), Mysuru, India
e-mail: pushpavtu@gmail.com

B. Ramesh

Department of Computer Science Engineering, Malnad College of Engineering, Hassan, India
e-mail: sanchara@gmail.com

Another concern to be considered during prediction is the time spent to predict should not overhead the performance of the system. Also wrong prediction can lead to huge uncertainty for resource allocation which in turn will violate the service level agreement (SLA). The main contribution in this work is, we have attempted to introduce an algorithm called hybrid load prediction algorithm (HLP) which combines neural network (NN) and genetic algorithm (GA) [2, 22]. NN was instructed by updating weights and selecting best inputs using GA and the outcome shows that the proposed HLP algorithm performs best compared to neural network by producing less mean square error.

The other sections are organized below in order. Section 2 discusses about literature survey on prediction. In Sect. 3, we have explained proposed algorithm for prediction. In Sect. 4, we have discussed how experimental set is done and also discussed the results obtained. In Sect. 5, we give conclusion and discuss about future enhancement.

2 Literature Survey

Many of the researchers are working on power optimization techniques and have addressed the pros and cons in the prevailing methods. Research is happening tremendously for optimizing power consumption in data center by performing prediction on server workload.

One of the key elements to optimize the power consumption is by predicting the workload based on resource utilization. So that servers that are underutilized are switched off to save power. Therefore, for this research work, literature survey is done focusing on load prediction. There are several practices for load predictions done, majorly homeostatic prediction and history-based prediction. In homeostatic prediction the imminent workload is predicted by increasing or decreasing the values, which is a mean value of the antecedent workloads.

On the other hand, history-based methods are effortless and admired prediction models.

Models inspect antecedent workloads and extricate patterns to predict the future demands [3]. There are several prediction algorithms exclusively for history-based method which depends on the uniform time intervals, namely auto regression, auto regressive integrated moving average model (ARIMA), exponential smoothing, hidden Markov chain model and many others.

One of the first researches on power management for data center was done by

1. Pinheiro [4]. In his work, authors proposed a method of working for minimizing power expending in a heterogeneous cluster of servers. In this technique, they monitored the load on resources like CPU, disk storage, network interface are made decision on switch off/on servers. The drawback in this paper was the

algorithm runs on master node, which generates a SPF and may lead performance gridlock in huge system and removes, adds solo node at a time, which leads for late operation.

2. Chase [5] has observed the issue of management of systematic energy usage in data center to workout with this issue; they came up with concept called “bid” the resources. The drawback in this paper is the service level agreement is negotiated according to the available resources. However, process deals with management of CPU without considering other resources. The latency of switching the servers on/off is also not considered. The network handles dynamic reconfiguration to alternate the server state from sleep/hibernation mode to active mode when necessary.
3. Elnozahy [6] has come up with issue regarding Web application on energy efficient resource management. Two power-saving techniques are used: handling power of computing nodes on/off and DVFS. The main idea is to estimate the CPU required and find the best number of physical nodes and set the proportional frequent to other nodes. The drawback changeover time for handling the power of node is ignored. Also, usage of resource data is not appropriate and leads to inefficient decision expected to fluctuation workload.
4. Beloglazov [7] has suggested power management by consolidating the virtual machines based on the resource utilization. For VM placement, they have used best fit decreasing algorithm. The limitation of this work is only CPU is considered for workload analysis. The servers have been switched to sleep mode to achieve power optimization. Other limitation is optimization is slow-going expected to compounded computation it wants.
5. Hang-peng [8] came up with flexible resource allocation for energy aware VM deployment in effective cc environment. He introduced online resource allocation method for VM consolidation to improve energy efficiency and decrease SLA violations for data center, taking into the account the usage of power for servers, and he accomplished live migrations, considered length of path for migration in data center. Uniqueness of the project is the author grouped the servers and tried to migrate VM's within the same group by decreasing the number of switches in the path. Here, the switches are switched off when not used.
6. Zhou [9] proposed a strategy called DADTA. In this paper, specific prediction of future workload is done on historical workload data. Based on future workload, over load and under load detection are done before, and later VM migration is done ahead SLA violation. CPU is the only resource considered for workload prediction in his work, VM migration reduces SLAV and servers are shift off to reduce energy consumption. A challenging problem observed in this paper is it will fail to allocate with non- stationary requirement of resource in fluctuating cloud schema.
7. Di and Kondo [10] predicted host load build on Bayesian model with upgraded feature combinations. All the prognosis methods are gauge on traces with 1000 + heterogeneous hosts. Model is compared with state-of-the -art methods build moving average, auto regression. From Google traces, nine features are

extracted to be used as corroboration in Bayes model. The objective of this paper is to predict the load changing for long-term without optimizing the power consumption. Another limitation of this paper is only CPU resource is considered for load prediction.

8. Duy [11] said, the decrease in power consumption can be built by cutting off servers which are not in use. Here, it is designed, implemented and evaluated a Green Scheduling Algorithm integrating NN predictor for optimizing server power consumption. Predicted future load based on historical load. Limitation of this paper is VM is constructing and destructing at each step without paying notice toward VM migration since clients request is supposed to be completely processed in step.
9. Hieu [12] presented a VM combining algorithm with usage prediction (VMCUP) refining the energy efficiency of cloud data center. The unite use of current and predicted CPU utilization metrics allows a well-founded characterization of overloaded and under loaded servers, which will turn off both the load and the power consumption after combining. The limitations observed in this paper are prediction of workload is not done efficiently. CPU is the only parameter considered for prediction. Also, power utilization during VM migrations and unnecessary VM migration is not taken into consideration.
10. Urul [13] initiated a method for CPU usage prediction using dynamic virtual machine allocation and migration method which raise energy efficiency while sustaining agreed quality of service levels in cloud data centers. He suggested a method, called LRAPS, which will estimate short-term CPU utilization of hosts based on their past resources.
11. Kakkar [14] in this research, MBFD algorithm has been analyzed and executed along with neural network as a classification approach for predicting the allocation error. Parameters, namely, allocation strength and error rate are considered for analyzing the whole process. In modified best fit decreasing (MBFD), first step is to category all the VMs in reduction order according to utilization of CPU. Then grant all VMs to respected host which give low increase of power consumption. Limitation of this paper is CPU is the only resource considered to determine the VM for Migration.
12. Zhou [15] proposed VM deployment algorithm (ATEA). Here, servers in data center are classified as little loaded, lightly loaded, heavily loaded and moderate loaded based on server load capacity. The algorithm migrates VM on heavily loaded or little-loaded hosts to lightly loaded hosts, while the VMs on lightly loaded and moderately loaded hosts remain unchanged. They incorporated adaptive method for determining values.
13. Alok Kumar Gupta [16] got patented for a method for lessening the power consumption for a cluster of servers. The method endorse host system power-on servers are overutilized and above the threshold and recommending host system power-off when there is a host system whose utilization is below the threshold. Limitation of this methodology is that VM are migrated without predicting the workloads for the host system and either CPU or Memory is

used as the resource utilized by the host to determine the weather servers are overutilized or underutilized.

14. Meisner [17], he introduced PowerNap concept, the system is getting transition from active state to zero power constant state with respect to load. Limitations of this work are while in nap state, power consumption is truncated, but no processing can occur. When the new work arrives the system wakes and transition back to active state. When work is done, the system will go back to nap mode.
15. Zhou [18] introduced energy-efficient virtual machine consolidation algorithm named PVDE (prediction-based VM deployment algorithm for energy efficiency). They used NN and linear-weighted method to forecast the load of a host and group the hosts in the data center, based on the forecasted host load values. They proposed four different algorithms for VM migration. Limitation of this research is there is huge trade off in optimizing power consumption and SLA. Prediction is not done accurately.
16. Farahnakian [19] introduced dynamic VM consolidation (DVMC) that can decrease energy consumption in data center. Efficiency is achieved by consolidating VMs into the few number of hosts and switching idle hosts to the sleep mode, and they proposed a dynamic VM consolidation approach called utilization prediction-aware best fit decreasing (UP-BFD) that improve the VM placement according to the present and future resource requirements.
17. Xiao et al. [20] gave the concept of “skewness” which measures the unevenness in the multi- dimensional resource utilization of a server. To keep away from overload on system by reducing number of servers to develop allocation of resource. They formed exponentially weighted moving average (EWMA) formula to predict the CPU load on the DNS server used load prediction algorithm that can capture the future resource usages of applications accurately without looking inside the VMs to save energy.
18. Mason et al. [21] instigate a method to forecast in forward the level of CPU consumption of a host. They used neural network for prediction. Several advanced swarm and evolutionary optimization algorithms are implemented to update the weights in the NN to predict host utilization, examples are particle swarm optimization (PSO), differential evolution (DE) and covariance matrix adaptation evolutionary strategy (CMA-ES). CMA-ES is considered to be good with faster convergence.

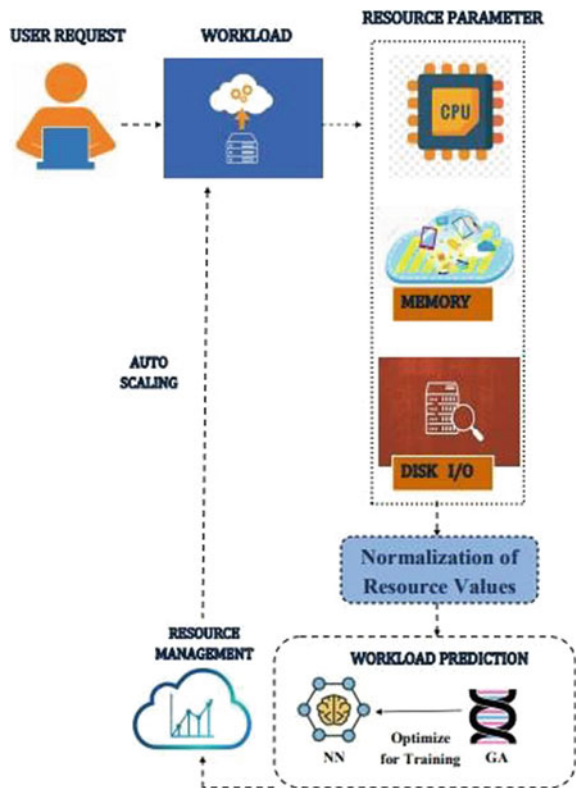
3 Proposed Work

With mighty rise in Internet capacity and its services, the urging for data center and cloud computing technology has turned broad. The boundless urge for cloud-based capacity and processing forces cloud service providers (CSP) to optimize their platforms and facilities. Workload prediction for dynamic workloads is prime issues in this optimization effort. More uncertainty can be seen in the workload coming to

the cloud. In this work, we present a hybrid algorithm which combines neural network and genetic algorithm for load prediction [26]. The parameters considered are CPU, memory and disk storage to upgrade energy efficiency while maintaining the quality of service levels in cloud data centers. Previous researchers have considered CPU as the main resource utilization, ignoring the other resource parameters. The suggested work, called hybrid load prediction algorithm (HLP), tries to predict the resource like CPU, memory and disk storage utilization of servers which will help in turn for load balancing, VM consolidation and in reducing power consumption [23, 24]. We performed simulation experiments using CloudSim simulator which gauge the efficiency and effectiveness of proposed method. The simulation experiments show that proposed method is attainable to put in and predict the workload of the resource based on the resource utilization. The complete flow of our research is shown in the system architecture Fig. 1.

In this research work, objectives are summarized: Analyzing and comparing the existing prediction algorithm such as neural network (NN) with the proposed hybrid load prediction algorithm (HLP).

Fig. 1 System architecture



Proposing hybrid load prediction algorithm (HLPA). We combine GA with NN for better prediction [22]. We choose the best inputs from the noisy dataset by calculating the fitness values of the each inputs and then fed for NN algorithm as input.

Inputs: Workloadlist

Output: Predictedlist

1. Start
2. Initialize the basic parameters. Individuallist,first individualist, second individualist, fitness
3. sval = 0, maxfitindex = 0,maxfit = 0.
4. Generate 10 chromosomes per iteration from the workloadlist.
5. Calculate the fitness value of the chromosomes.
 For i belongs to individual list
 Samplegene = i
 For m belongs to samplegene
 If m > 1,000,000
 Increment fitnessval
 End if
 End for
6. Get the fittest chromosomes in the given list
 For I belongs to individual list
 If maxfit <= fitness value of i Maxfit = fitness value of i Maxfitindex = i
 End if
 End for
7. Add the fittest chromosomes to the firstindividual list.
8. Get the second fittest chromosomes. For I belongs to individualist
 If fitness of I > fitness of maxfit
 Maxfit 2 = maxfit
 Maxfit = i
 Else if fitness of I > fitness of maxfit2
 Maxfit2 = i
 End if
 End for
9. Add the second fittest chromosomes to second individual list
10. Generate the random crossover point
 Crossoverpnt = random(firstchromosomessize)
11. Select firstchromosomesfrom the firstindividuallist(testgene1) and secondindividuallist(testgene2) and add it to samplelist1 and samplelist2

12. For k belongs to crossoverpnt Temp =
samplelist1
Samplelist1 = samplelist2 Samplelist2 = temp
13. Calculate the fitness of first fittestchromosomes
and second fittest chromosomes
For I belongs to list
Samplegene = i
For m belongs to samplegene
If m > 1,000,000
Increment fitness value
End if
End for
End for
14. Creating off Springs. Initialize leastfitindex = 0,
minfitindex = integer.Max_value, minfitindex = 0
For I belongs to individullist
If minfitval > = fitness of i Minfitval = fitness of i
Minfitindex = i
End if
End for
15. Replace the least fit index value with the best fit
index value
If fitness of maxfitindex > fitness of maxfit2
Add maxfit into leastfitindex
Else
Add maxfit2 into leastfitindex
End if
16. Get the newly created offspring and add it to the
finalist. Continue till all the end of input.
17. For I belongs to finalist
Generate two random points p and q within the size
of the genetic algorithm output
18. Get the value in the index of p and q from the
final finalist.
19. Generate random weights between 0 and 1 for the
two inputs *a* and *b*
 $w = \text{math.random}(1-0)$
20. Calculate the summation of the inputs

$$X = (a * w1) + (b * w2)$$

21. Calculate the sigmoid value.

$$\text{Sigmoid} = 1 / (1 + \exp(-x))$$

22. Get the newly predicted value based on the sigmoid value

$$m = (\text{Math. roudom} * 1, 000, 000) * \text{sigmoid}$$

23. Add m to the final list called predictedlist. Endfor.

3.1 Parameters Used

Fitnessval: stores the fitness value of the genes. Maxfitindex: indicates the index of genes with maximum fitness value. Maxfit2: indicates the index of genes with second maximum fitness value.

Workload: an list containing the input values. Genes: lists used to store the genes.

Individualist: used to represent the chromosomes.

Maxfit: stores the maximum fitness values.

Firstindividuallist: contains the genes or values with maximum fitness value.

Second individualist: contains the genes or values with second maximum fitness value.

Testgene: contains first set of values from individual list Testgene 1: contains first set of values from first individual list[maximum fitness].

Testgene2: contains first set of values from second individual list[second max fitness] Crossoverpnt: indicates the crossover point/index Minfitindex: indicates the index of genes with minimum fitness value.

Minfit: stores minimum fitness value.

Finalist: stores the output.

Samplegene: contains one chromosomes from the individual list

Samplelist1: Contains chromosomes from first individual list.

Samplelist2: Contains chromosomes from second individual list.

Predictedlist: contains the predicted values in a list.

3.2 Training of Neural Network with Genetic Algorithm as Optimizers

Adaptation of bio- inspired algorithms is making progress in training and optimizing the neural network. One such algorithm considered is GA to find the fittest among many species. In the growth of animals or plants, a new creature is formed by the crossover of the two healthy parents.

These can be grouped in genes, which posses some piece of data on individual. The genetic information is called the genome and is the genotype of the individual. The solitary product is called phenotype. Similar genotype may give different phenotypes: Twins illustrate this quite well. A genetic algorithm tries to simulate the natural

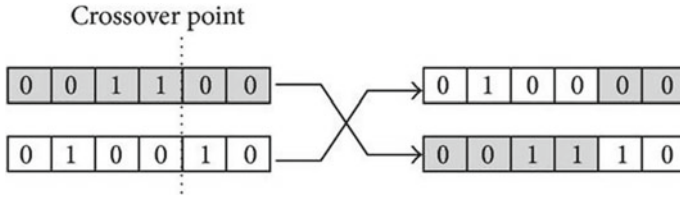


Fig. 2 One-point crossover

evolution process. Its purpose is to optimize a set of parameters. The genetic information is encoded in a bit string of fixed length, called the parameter string or individual [22].

The fitness of the new individuals is stored in the fitness value. Fitness of each species is calculated using fitness function which is mentioned in the algorithm. Each problem has its own fitness solution. For optimization problem, primary function such as sum of a set of calculated parameters related to the problem domain can be used as fitness function.

Crossover, selection and mutation are elementary function is evolutionary algorithms. In this work, we have considered single point crossover and selection process to select the genuine input for neural network. One-point crossover is applied on the inputs to create new springs which possess the trade off genetic material from parents and later creates chromosomes which are alike to parents. Example is shown in the below Fig. 2.

After crossover, the fitness of the off -springs is checked and verified. The chromosomes with highest value are considered as first individual list and chromosomes with second highest value are considered as second individual list. The chromosomes whose value is less than maximum fitness value are considered as weak chromosomes and those chromosomes are replaced by chromosomes with highest fitness value. And output of GA is given as input for NN with inputs with best values [22].

In this ground work, the initial inputs of NN were spot via GA using fitness function. The healthy individuals were selected using fitness function. In the NN, we employed with one input layer, one hidden layer and output layer. Inputs features are CPU, memory and disk s from the Google servers resource usage task table and the output pertains to prediction of resource utilization. We have three inputs, four neurons in one hidden layer and two outputs specifying over utilized or underutilized. The number of neurons selection is done based on trial an error method. For parent selection, Roulette wheel method was not used since it was not used since it was not appropriate for noisy data like Google traces. We extracted best parents from genetic algorithm substitution; we selected of the finest individuals from the latest generation and substituted them with the worst individual belonging to the next generation. In our training, neural network 80% of dataset was used for training and 20% data was considered for testing [24]. Since the dataset consists of noisy data, mini batch of data set is considered of training and testing [27]. Based on these parameters, the neural network has been designed a shown in Fig. 3.

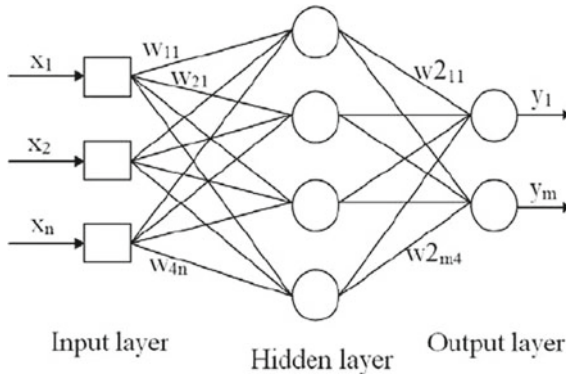


Fig. 3 Structure of neural network

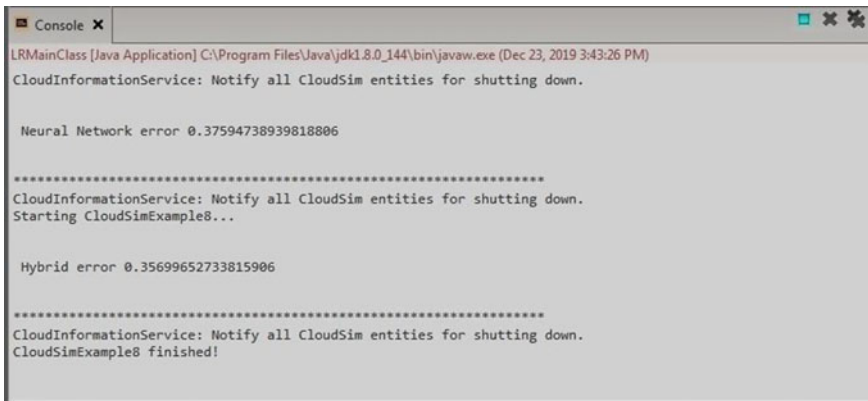


Fig. 4 MSE values for NN and HPLA

In Fig. 3, artificial neural network (ANN) mainly consists of three parts

1. The input layer is $X = (x_1, x_2 \dots x_n)$ and n for number of input nodes.
 2. The hidden layer is the significant function of ANN. Nodes in hidden layers are marked as $H = h_0, h_1, \dots, h_n$ and they are memory cells of the network. Every node in the hidden layer has activation function. If value of activation function is >0.5 , then such nodes are considered to be active. Current messages can be transferred to the next step.
 3. The output layer is Y_1 and Y_m , and gives the predicted values.
- Activation function used in this work is sigmoid function as shown in Eq. (3). The value if sigmoid function is between 0 and 1. The values greater than 0.5 is considered as best neurons.

$$w_1x_1 + w_2x_2 + \dots w_nx_n \quad (1)$$

The variables w and x are weights and inputs, respectively. The value of variable net for the individual neuron should be greater than 0.5 to be considered as fit neuron for prediction. n is the number of inputs from the dataset.

$$\text{net} = \sum_{i=0}^N w_i x_i \quad (2)$$

$$F(\text{net}) = \frac{1}{1 + e^{-\text{net}}} \quad (3)$$

- Since only few inputs are considered from the dataset, mini-batch gradient descent is used. In Eq. (4), y_1 is the actual value and y_2 is the predicted value.

$$\text{MSE} = \frac{1}{n} + \sum_{i=0}^n (y_1 - y_2)^2 \quad (4)$$

- Weights are updated using Eq. (5). Where w_{new} is the new updated value and w_{old} is previous weights value. η is the learning rate with the value 0.3 based on survey. L is the loss value which is the difference between actual and predicted value.

$$w_{\text{new}} = w_{\text{old}} - \frac{\sigma L}{\sigma w_{\text{old}}} \quad (5)$$

4 Experimental Setup and Results

4.1 Experimental Setup

The experiment was conducted using cloudsim Simulator. We simulated a data center consisting of fourteen heterogeneous servers interconnected through a three-tier fat-tree network topology [24], which is the most widely adopted topologies within DC: 7 Dual Core servers & 7 Quad Core servers. Seven servers have 2 cores with 1.5 GM of memory and storage of 2 GB. Other set of 7 servers which has four core processor with 2 GB of memory, x86 system architecture; 64bit processing is used with Linux operating system. The virtual machine monitor Xen is a type-1 hypervisor, which permits multiple OS to execute on similar hardware simultaneously. Micro-instance

virtual machines with 512 MB memory and 250 MIPS are created. Totally, 500 VM's are used for experiment [25].

4.2 Results

There are abundant methods for prediction, in which NN plays prominent role [28]. In the following, we compare the performance of HLP algorithm with the existing neural network algorithm for the Google workload trace collected on 11 March. Figure 5a, b shows the relationship between actual and predicted values of neural network and hybrid load prediction algorithm (HLP). The prediction efficiency is achieved in proposed HLP algorithm compared to NN. These two algorithms are well trained and tested to achieve high efficiency. Figure 4 shows the mean square error of 0.356 for proposed HLP and 0.375 for neural network. Genetic algorithm is used as optimizers in selection of inputs for NN from noisy data set. Fitness function is used to select the inputs for NN. Fitness value is calculated after every crossover. The prediction is done for every 5 min of intervals.

5 Conclusion and Future Enhancement

There are many physical methods to reduce power consumption. Technically, the proposed HLP algorithm shows that by combining genetic algorithm with NN can improve prediction efficiency. Predicting the resource utilization which will benefit in reducing power consumption, load balancing, resource allocation will be shown in my next paper. The results show that proposed HLP algorithm is more efficient than existing NN by selecting the best inputs from the noisy dataset. In future, the NN can be trained with other bio-inspired optimizer like Rider NN, Stochastic fruitfly algorithm, etc., by considering other data set.

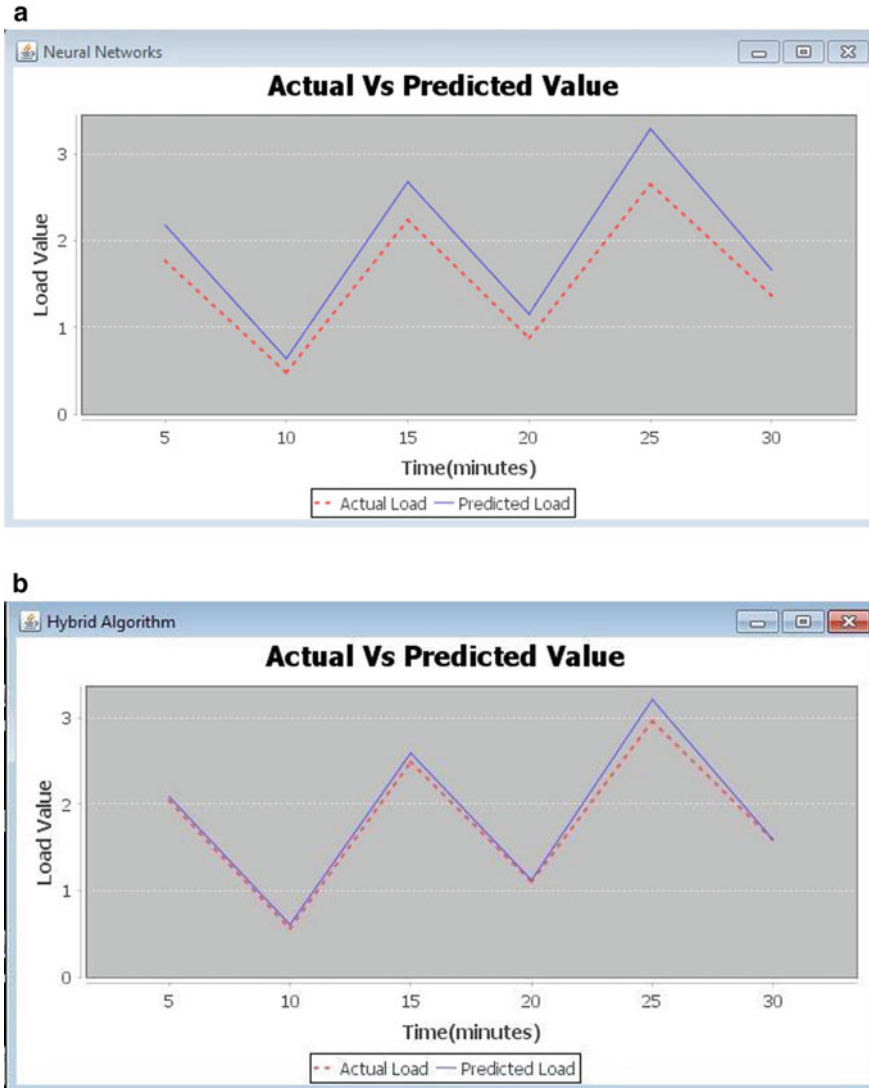


Fig. 5 a Forecast using neural network. b Forecast using HPLA

References

1. R. Pushpalatha, B. Ramesh, Pre-eminent strategy for effective utilization of power in data center. *IJRTE* 8(5), 1442–1447 (2020)
2. Y. Zhangang, C. Yano, K.W. Eric Cheng, Genetic algorithm-based RBF neural network load forecasting model. *IEEE* (2007)
3. H. Xiao, Z. Hu, K. Li, Multi-objective VM consolidation based on thresholds and ant colony system in cloud computing. *IEEE Trans. Content Min.* 7, 53441–53453 (2019)

4. E. Pinheiro, R. Bianchini, E.V. Carrera, T. Heath, Load balancing and unbalancing for power and performance in cluster-based systems, in *Proceedings of the Workshop on Compilers and Operating Systems for Low Power*, vol. 1, pp. 182–195 (2001)
5. J.S. Chase, D.C. Anderson, P.N. Thakar, A.M. Vahdat, R.P. Doyle, Managing energy and server resources in hosting centers, in *Proceedings of the 18th ACM Symposium on Operating Systems Principles*, vol. 35 (ACM, New York, NY, USA, 2001)
6. America's Data Centers Are Wasting Huge Amounts of Energy, White Center data center-IB-Final-826 published on August 2014
7. A. Beloglazov, J. Abawajy, R. Buyya, Energy-aware resource allocation heuristics for efficient management of data centers for cloud computing. *Future Gener. Comput. Syst. J.* **28**, 755–768 (2011). (Elsevier)
8. H.-P. Jiang, W.-M. Chen, Self-adaptive resource allocation for energy-aware virtual machine placement in dynamic computing cloud. *J. Netw. Comput. Appl.* **120**, 119–129 (2018). (Elsevier)
9. H. Zhou, Q. Li, K.-K.R. Choo, H. Zhu, DADTA: A novel adaptive strategy for energy and performance efficient virtual machine consolidation. *J. Parallel Distrib. Comput.* **121**, 15–26 (2018). (Elsevier)
10. S. Di, D. Kondo, W. Cirne, Google hostload prediction based on Bayesian model with optimized feature combination. *J. Parallel Distrib. Comput.* **74**, 1820–1832 (2014). (Elsevier)
11. T.V.T. Duy, Y. Sato, Y. Inoguchi, Performance evaluation of a green scheduling algorithm for energy savings in cloud computing. *JAIST, IEEE*, 1–8 (2010)
12. N.T. Hieu, M. Di Francesco, A. Yla-Jaaski, Virtual machine consolidation with usage prediction for energy-efficient cloud data centers, in *IEEE 8th International Conference on Cloud Computing* (2015), pp. 750–757
13. G. Urul, Energy efficient dynamic virtual machine allocation with CPU usage prediction in cloud datacenters, Thesis submitted to the Graduate School Of Engineering And Science, January 2018
14. D. Kakkar, Heuristic of VM allocation to reduce migration and energy at cloud server, A Thesis Presented to the Faculty of California State Polytechnic University, Pomona, 2017
15. Z. Zhou, Z. Hu, K. Li, *Virtual machine placement algorithm for both energy-awareness and SLA violation reduction in cloud data centers*, vol. 2016 (Hindawi Publishing Corporation Scientific Programming, 10 Mar 2016)
16. A.K. Gupta, M. Ji, T.P. Mann, T. Mobashir, U. Rencuzogullari G.. Shanmuganthan, L. Wang, A. Marie, A.M. Holler, Reducing Power Consumption In a Server Cluster, United States Patent, 10 Sept 2009
17. D. Meisner, B.T. Gold, T.F. Wenisch, PowerNap: eliminating server idle power, in *ASPLOS'09* (Washington, DC, USA, March 7–11, 2009)
18. Z. Zhou, Z.-g. Hu, J.-y. Yu, J. Abawajy, M. Chowdhury, *Energy-efficient virtual machine consolidation algorithm in cloud data centers*, vol. 24, pp. 2331–2341 (Central South University Press, Springer-Verlag GmbH Germany, 2017)
19. F. Farahnakian, T. Pahikkala, P. Liljeberg, J. Plosila H. Tenhunen, Utilization prediction aware VM consolidation approach for green cloud computing, in *IEEE 8th International Conference on Cloud Computing* (2015), pp. 381–388
20. Z. Xiao, W. Song, Q. Chen, Dynamic resource allocation using virtual machines for cloud computing environment. *IEEE Trans. Parallel Distrib. Syst.* **24**, 1107–1117 (2013)
21. K. Mason, M. Duggan, E. Barrett, J. Duggan, E. Howley, Predicting host CPU utilization in the cloud using evolutionary neural networks. *Future Gener. Comput. Syst. J.* **86**, 162–173 (Elsevier, 2018)
22. Z. Arabasadi, R. Alizadehsani, M. Roshanzamir, H. Moosaei, A.A. Yarifard, *Computer aided decision making for heart disease detection using hybrid neural network Genetic algorithm*, vol. 141 (Elsevier, 2017), pp 19–26
23. S. Ismaeel, R. Karim, A. Miri, Proactive dynamic virtual-machine consolidation for energy conservation in cloud data centres. *J. Cloud Comput. Adv. Syst. Appl.* (2018)

24. Y. Mao, D. Ren, X. Chen, *Adaptive load balancing algorithm based on prediction model in cloud computing* (ICCC Wuhan, China, 2013), pp. 165–170
25. A. Beloglazov, R. Buyya, Optimal online deterministic algorithms and adaptive heuristics for energy and performance efficient dynamic consolidation of virtual machines in cloud data centers. *Concurr. Comput. Wiley Inter Sci.* **24** (2012)
26. J. Kumar, A.K. Singh, Workload prediction in cloud using artificial neural network and adaptive differential evolution, vol. 81, pp. 41–52. (2018). (Elsevier)
27. A. Beloglazov, R. Buyya, Y.C. Lee, A. Zomaya, A taxonomy and survey of energy efficient data centers and cloud computing systems. *Adv. Comput. Sci.* **82**, 47–111 (2011). (Elsevier)
28. M. Chandini, R. Pushpalatha, B. Ramesh, A brief study on prediction of load in cloud environment. *Int. Adv. Res. Commun.* **5**, 157–162 (2016)

Pressure Sensor Based on Two-Dimensional Photonic-Crystal Ring Resonator



D. L. Lakshmi, Venkateswara Rao Kolli, P. C. Srikanth, D. L. Girijamba, and Indira Bahaddur

1 Introduction

Photonic-crystals (PCs) are the periodically structured electromagnetic media. It is called so because the materials having different refractive index are arranged in periodic manner. *Photonic-crystals* possess *photonic-band-gap (PBG)*, the range of frequency where light cannot propagate through the structure [2]. The wavelengths of a wave allowed to flow through the structure are modes, and group of such modes forms a band. Laser propagates through the crystal for most of the frequencies. PBG will not allow light propagate through the crystal for a band of frequencies. This is called as PBG property [2–4].

There are some existing pressure sensors with limitations like low quality factor. There is always scope for the pressure sensor designed with good quality factor. Photonic-crystal ring resonator which can work as a pressure sensor detects the signal and the damage of structures if its sensitivity, stability, and quality factor are improved.

D. L. Lakshmi (✉)

Department of E&C Engineering, BGS Institute of Technology, BG Nagara, India
e-mail: lakshmidl.ec@gmail.com

V. R. Kolli · P. C. Srikanth · I. Bahaddur

Department of E&C Engineering, Malnad College of Engineering, Hassan, India
e-mail: venkukolli@gmail.com

I. Bahaddur

e-mail: indirabahaddur@gmail.com

D. L. Girijamba

Department of E&C Engineering, Vidyavardhaka College of Engineering, Mysuru, India
e-mail: girijamba@vvc.ac.in

2 Photonic-Crystal for Pressure Sensing

Photonic-crystal ring resonator designed for pressure sensing (Fig. 1) in which the waveguide and hexagonal ring is formed over a silicon substrate which is having the refractive index of 3.48 and holes are created having the refractive index 1, radius 110 nm with lattice constant 410 nm.

The fabrication type of *photonic-crystal* employed here is the “holes in slab” or hole-type photonic-crystal slab configuration. This configuration is preferred over the “rods in air” as it is easy to fabricate holes on slab. Photonic-crystal ring resonator has four ports, and they are.

- Input port
- Transmission port
- Forward loop

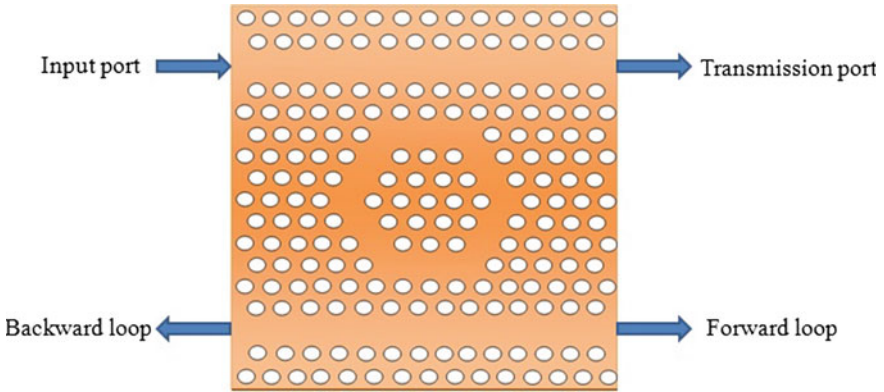


Fig. 1 Schematic diagram of photonic-crystal for pressure sensing

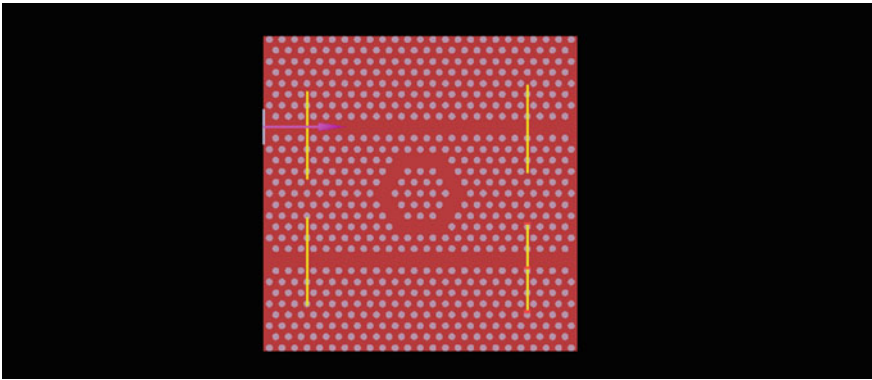


Fig. 2 Photonic-crystal ring resonator designed using Lumerical FDTD software tool

- Backward loop

The material on which ring resonator formed is silicon. We can create ring resonator by defining the proper dimensions for silicon, creating holes on silicon and removing holes of two rows which act as a linear waveguide. Create a hexagonal ring at the center of the slab, and it should be between the two waveguides (Fig. 2).

Steps to create photonic-crystal ring resonator

- A *rectangular shape* is considered as a silicon slab which should be given the dimensions as $10 * 10 \mu\text{m}$ to form a square silicon slab.
- Create holes in the silicon slab with hole radius 110 nm and lattice constant 410 nm to form a hexagonal lattice structure.
- Create two line defects that act as straight waveguides. It is created by removing holes in one complete row or by not creating holes.
- Create a hexagonal ring at the center of the silicon slab between the waveguides. It is also created by removing *holes* in hexagonal shape.
- Set a source at the input port which should be of the type mode source that injects guided source into the simulation region.
- Set monitors at different ports to observe electric field, magnetic field and pointing vector and *normalized-transmission* as a function of space and wavelength or frequency.
- The design parameters chosen to create a photonic-crystal ring resonator are as given below:
 - Silicon slab of width $10 * 10 \mu\text{m}$ and thickness 220 nm.
 - Holes of radius 110 nm and thickness 220 nm.
 - Lattice constant 410 nm.
 - Holes in slab configuration.
 - Hexagonal lattice
 - Operating wavelength ranges between 1.5 and 1.65 μm .
 - Frequency domain power monitors to monitor the simulation.

The device works on photo-elastic effect. The RI of the photonic-crystal depends on applied pressure. The change in RI leads to a shift in resonant wavelength. The applied pressure and RI are related by the equations [5, 6] given below.

$$n = n_0 - (C_1 + 2C_2)\sigma \quad (1)$$

$$C_1 = n_0^3 \left(\frac{p_{11} - 2V p_{12}}{2E} \right) \quad (2)$$

$$C_2 = n_0^3 \left(\frac{p_{12} - V(p_{11} + p_{12})}{2E} \right) \quad (3)$$

where E, V are Young’s modulus and Poisson’s ratios of silicon. P_{11} and P_{12} are the strain constants.

3 Simulation Results

Sensors designed with photonic crystal shows good quality factor and sensitivity. When there is no pressure on structure, RI of the material remains as earlier. When pressure is applied on sensor, RI will be increased and will be added to n_0 for the next trial.

The transmission characteristics with wavelength for applied pressures is observed at Monitor-2 (Transmission port) (Figs. 3, 4, 5, 6, 7, 8, 9 and 10).

Table of simulation results.

The table summarizes the simulation for different pressure values. Applied pressure on photonic-crystal ring resonator promotes change in refractive index and this in turn causes shift in wavelength.

Equations 1, 2 and 3 give the relation between the applied pressure and refractive index. For the different pressure values like 0 Pa, 500 kPa, 1 Mpa, 500 MPa, 1 GPa, and 10 GPa, the refractive index and resonant wavelength values are observed as in the Table 1.

High Quality factor:

The high quality factor analysis object produces the following plots (Fig. 11).

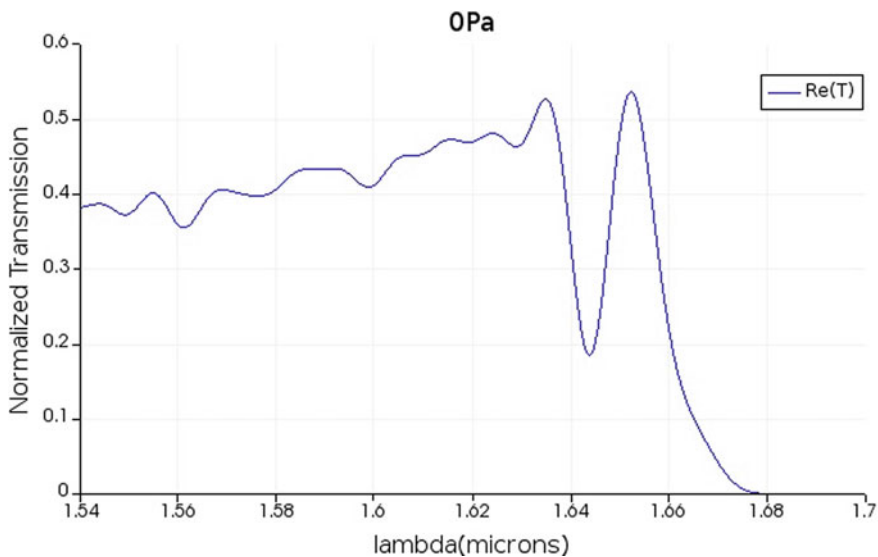


Fig. 3 Transmission characteristics at no pressure

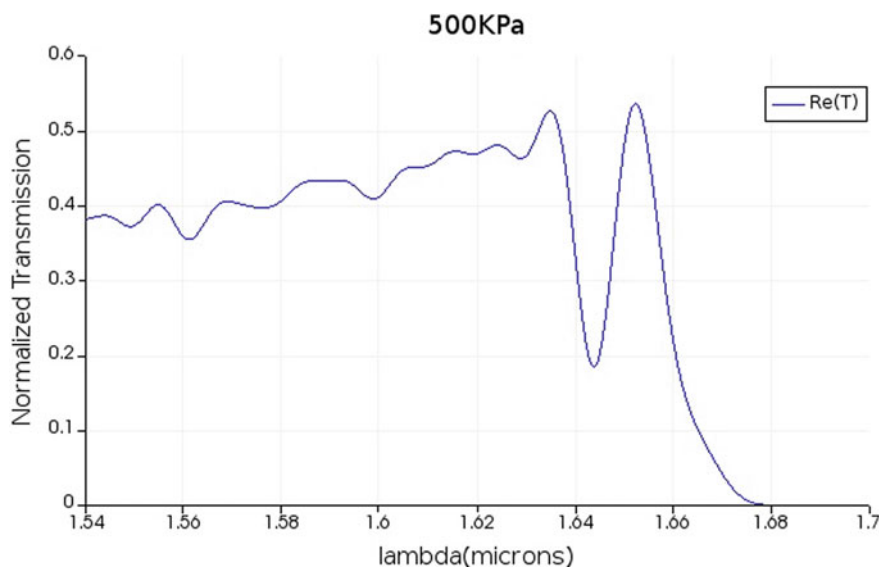


Fig. 4 Transmission characteristics at 500 KPa

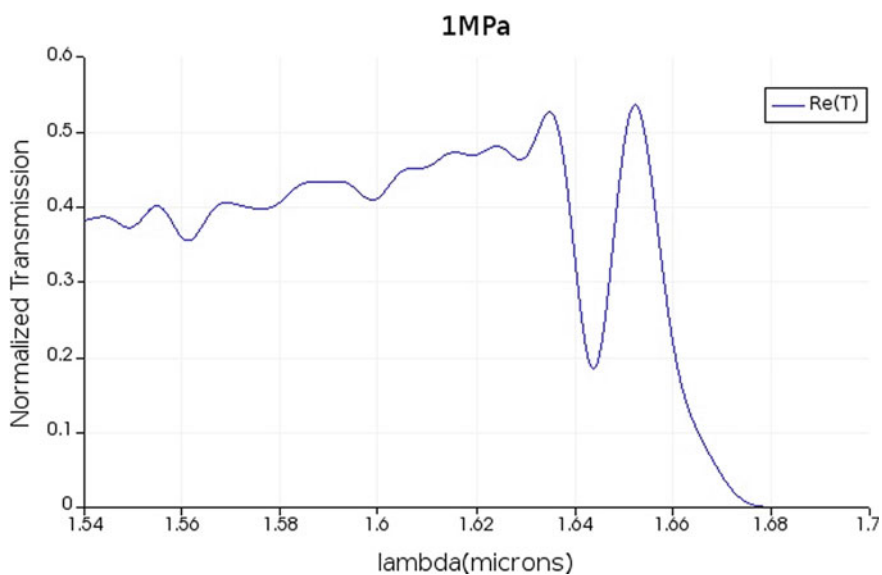


Fig. 5 Transmission characteristics at 1 MPa pressure

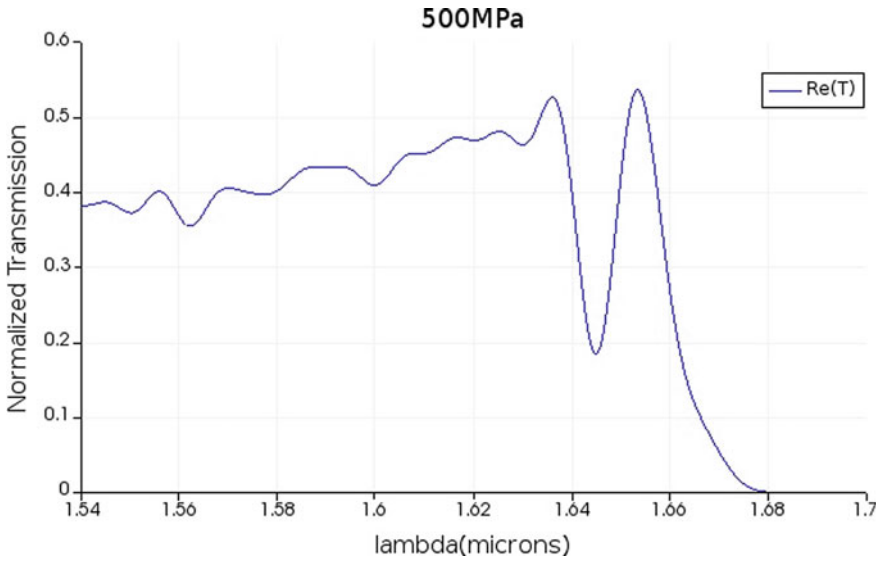


Fig. 6 Transmission characteristics at 500 MPa pressure

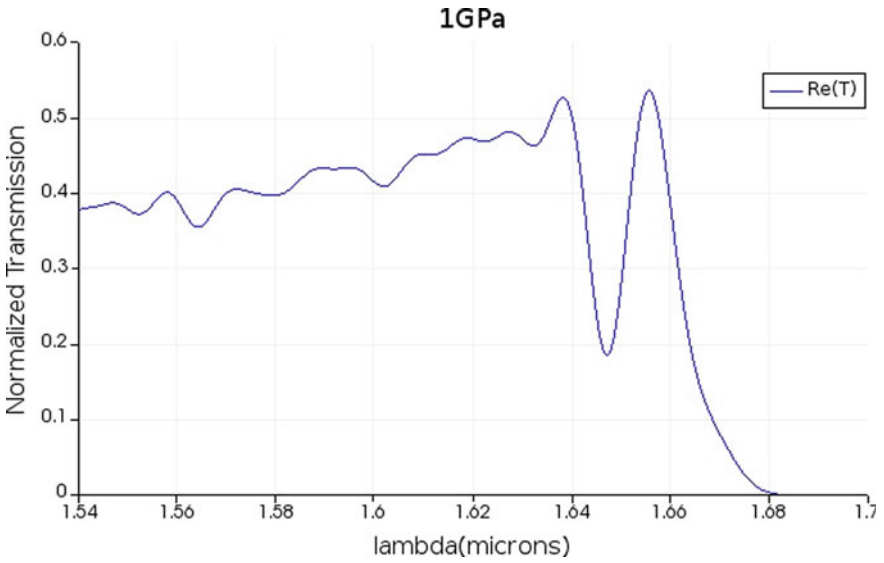


Fig. 7 Transmission characteristics at 1 GPa pressure

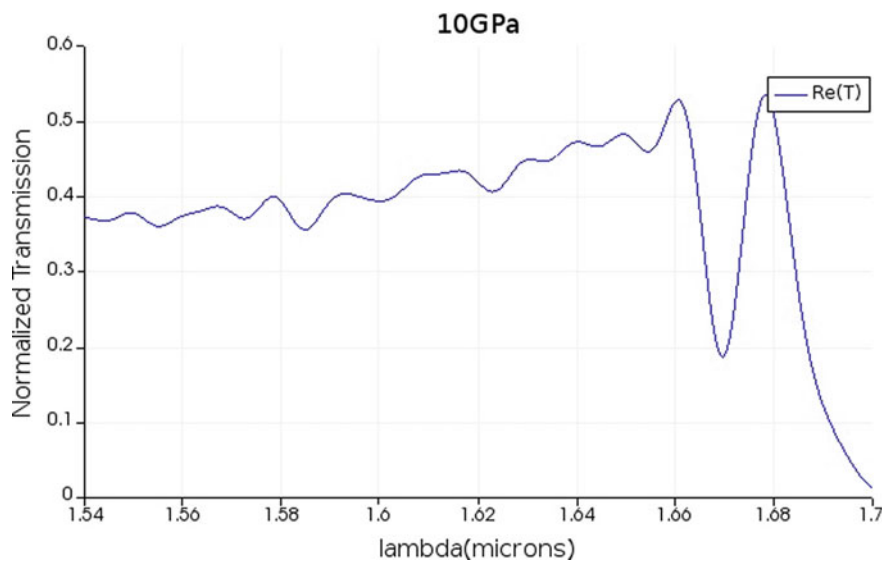


Fig. 8 Transmission characteristics at 10 GPa pressure

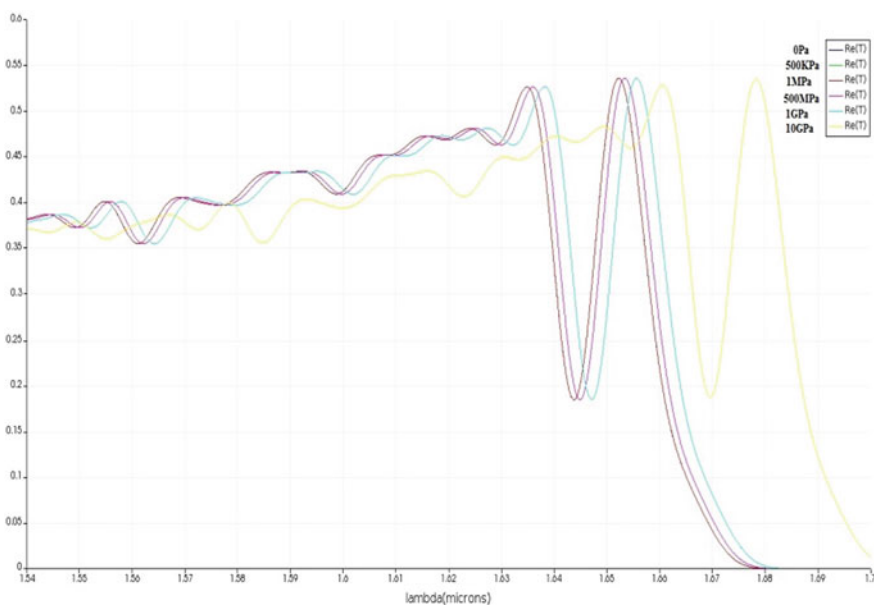


Fig. 9 Wavelength shift for different pressures

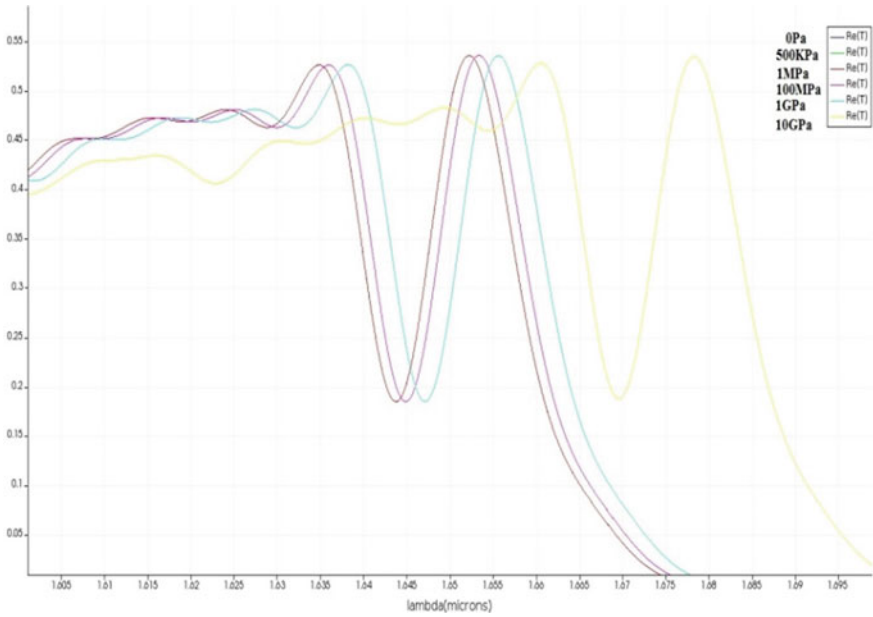


Fig. 10 Zoomed view of wavelength shift for different pressures

Table 1 Simulation results for different pressures

Pressure	RI	Resonant wavelength (μm)
0 Pa	3.0	1.65
500 kPa	3.000002128	1.65218
1 MPa	3.000006384	1.65219
500 MPa	3.002134424	1.65331
1 GPa	3.006399567	1.65553
10 GPa	3.049233044	1.66051

Figure 12 shows the decay of the EM components with respect to time (Fig. 12).

Figure 12 shows the decay of the sum of squared EM components on a logarithmic scale. The slope of these lines is used in the Q calculations [7].

In the plot spectrum of resonances (Fig. 13), each resonant peak appears in different colors [7] (Fig. 14).

The resonant peaks are isolated by using spectrum and the Gaussian filters [7] (Fig. 15).

Low Quality factor.

The low Q-analysis group finds the Q-factor of each resonance frequencies. [7].

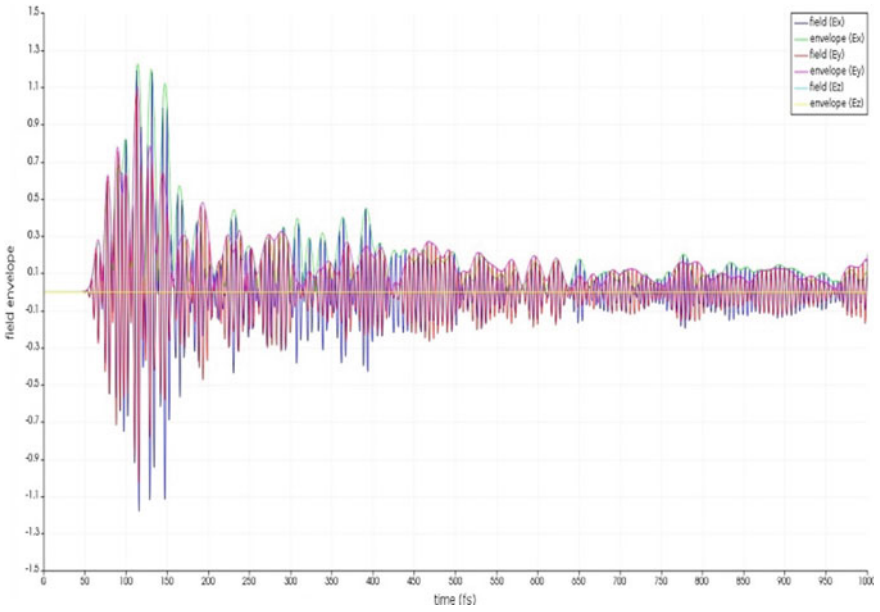


Fig. 11 Field envelope

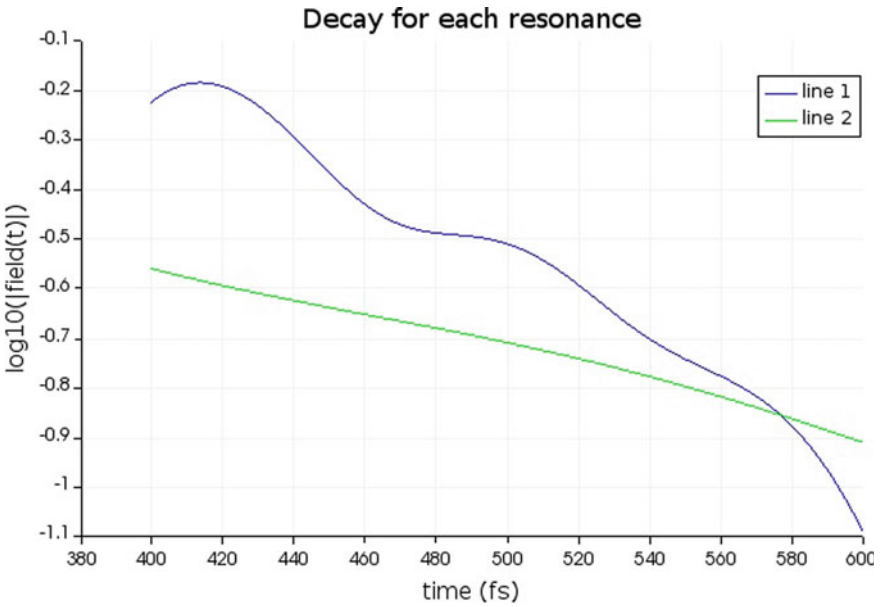


Fig. 12 Decay for each resonance

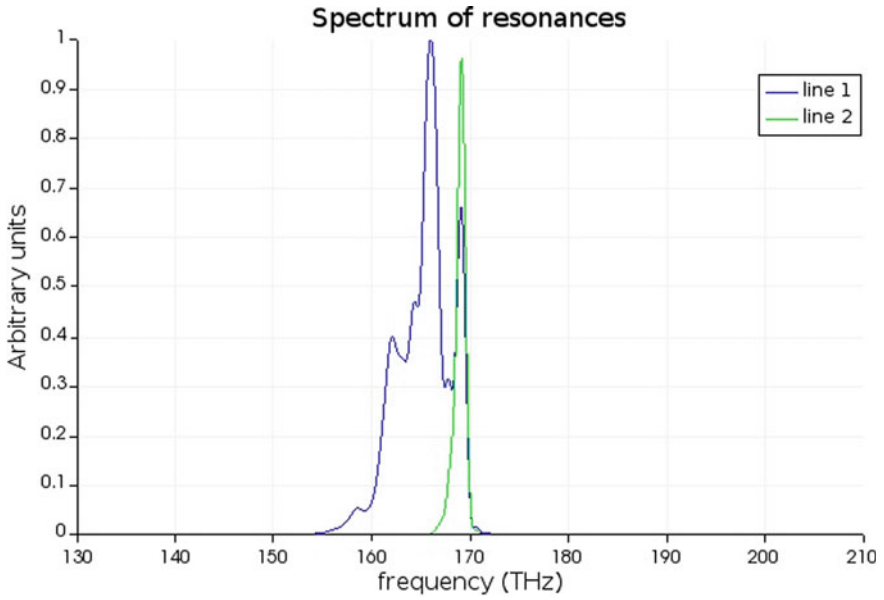


Fig. 13 Spectrum of resonances

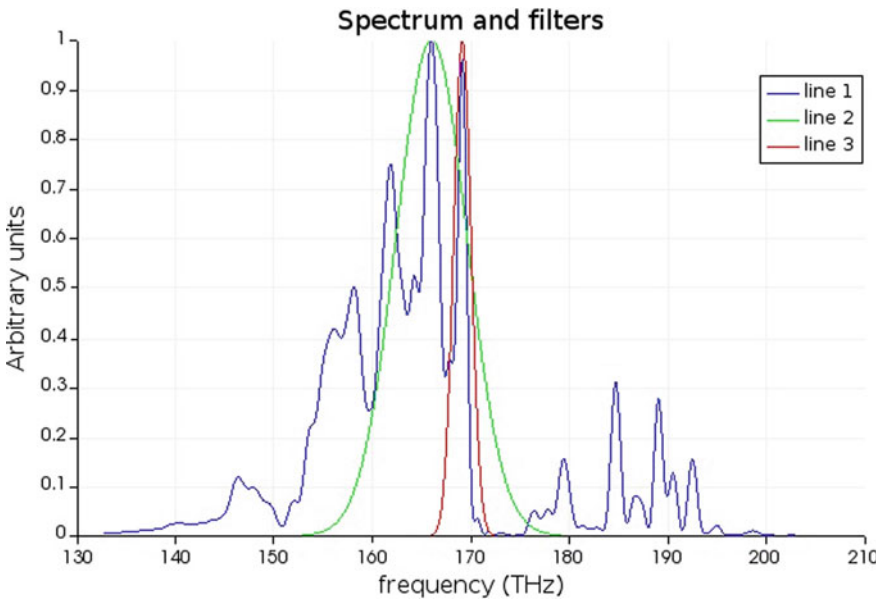


Fig. 14 Spectrum and filters

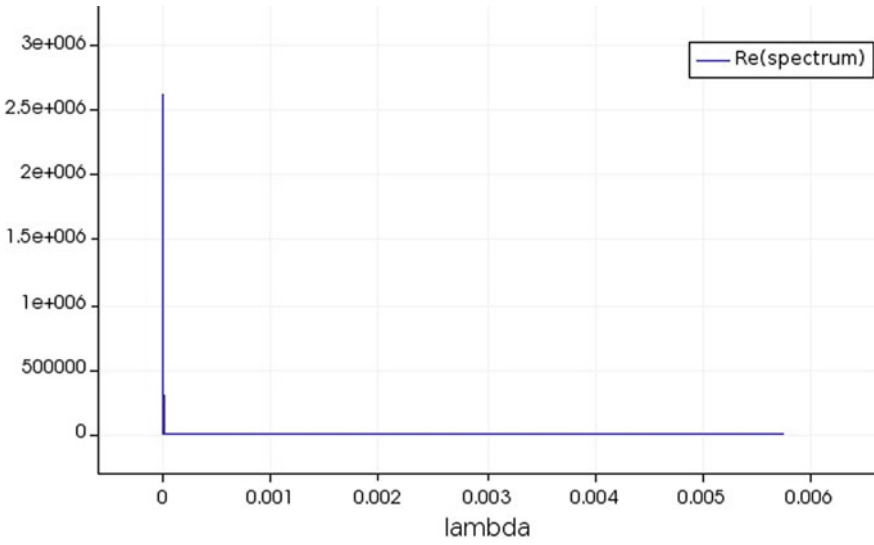


Fig. 15 Quality analysis spectrum versus spectrum

The Q-analysis generates two plots. In the plot field component (Hz), the field decays with respect to time. The spectrum plot shows the location and relative amplitude of the resonance peaks [7] (Fig. 16, 17 and 18).

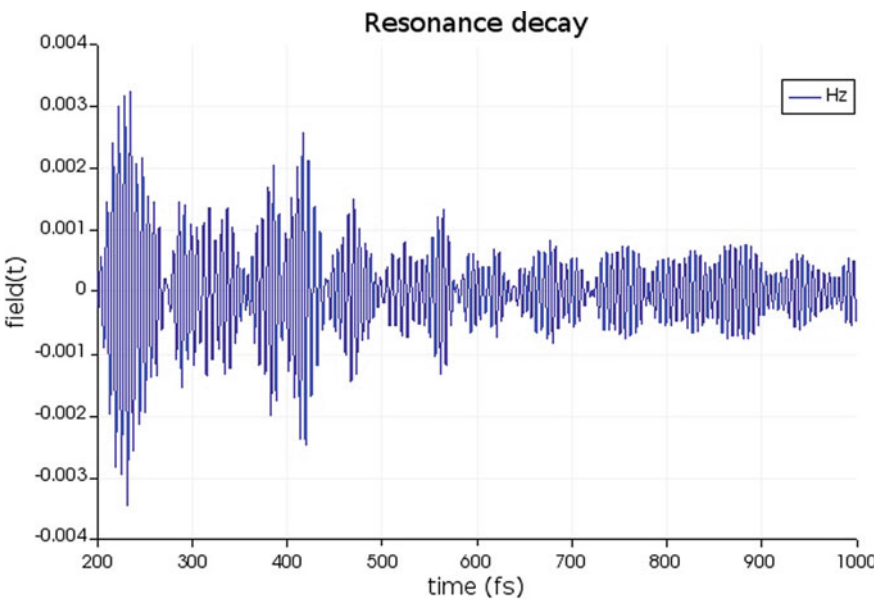


Fig. 16 Resonance decay

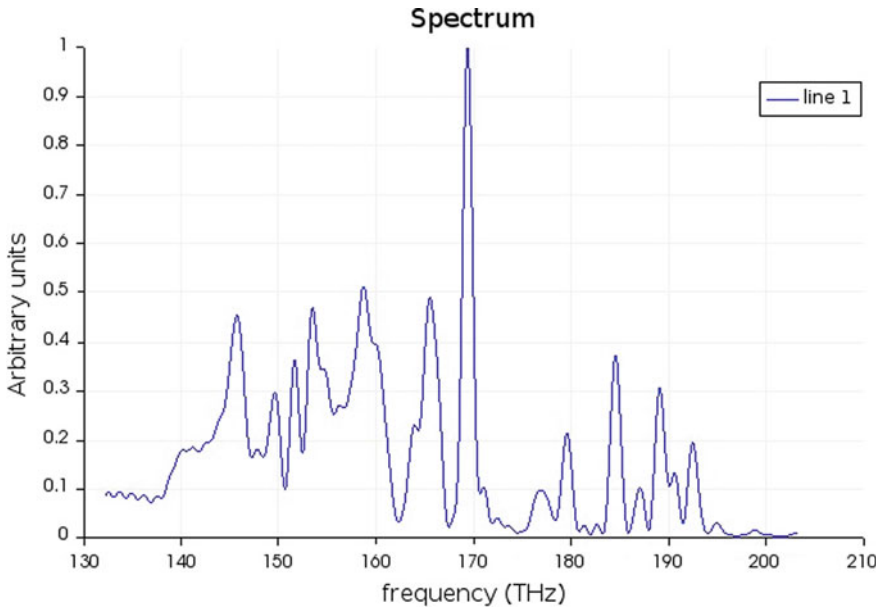


Fig. 17 Quality factor Spectrum versus frequency

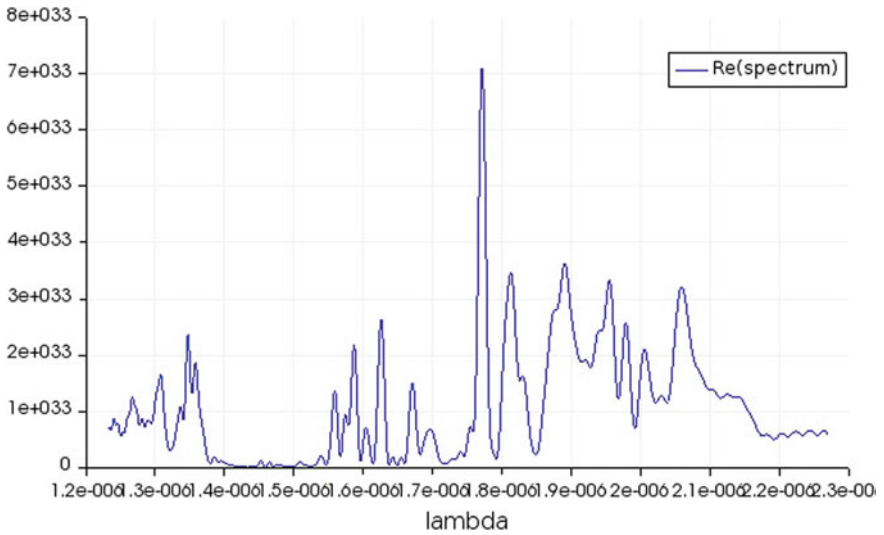


Fig. 18 Quality factor spectrum versus wavelength

4 Conclusion

In this paper, a 2-D photonic-crystal which can work as pressure sensor has been demonstrated. The sensor is formed by creating the ring resonator on silicon slab with holes in slab type configuration. *Normalized-transmission* at input port and transmission port is observed. Resonant wavelength is observed for each pressure and compared with different pressures. The pressure is varied from 0 Pa to 1 GPa. For change in refractive index on applying pressure on the material, there is shift in the wavelength, and hence, it acts as a pressure sensor. Analysis of wavelength shift is done with the equations which relate applied pressure and refractive index. The wavelength shift for the pressure values from 0 to 1GPa is 1.65 to 1.66051. Simulation and analysis is done using finite difference time domain (FDTD) method.

References

1. B. Troia, A. Paolicelli, F. De Leonardis, V.M.N. Passaro, Photonic-crystals for optical sensing: a review. in *Advances in Photonic Crystals* (Chapter 11), ed. by V.M.N. Passaro, IntechOpen Ltd, 2013
2. S.G. Johnson, J.D. Joannopoulos, Introduction to photonic-crystals: Bloch's theorem, band diagrams, and gaps (but no defects). (MIT 3rd February 2003)
3. J.D. Joannopoulos., S.G. Johnson, J.N. Winn, R.D. Meade, *Photonic Crystals: Molding the Flow of Light*, 2nd edn. (Princeton University Press, Princeton, 2008)
4. R.D. Meade, A.M. Rappe, K.D. Brommer, J.D. Joannopoulos, Existence of a photonic band gap in two dimensions. *Appl. Phys. Lett.* **61**(4), 495 (1992)
5. T.A.O, Shangbin, D. Chen, J. Wang, J. Qiao, Y. Duan, A high sensitivity pressure sensor based on two-dimensional photonic-crystal. *Photonic Sens.* **6**, 137–142 (2016)
6. K.V. Shanthi, S. Robinson, Two-dimensional photonic crystal based sensor for pressure sensing. *Photonic Sens.* **3**(3), 248–253 (2014)
7. https://kb.lumerical.com/en/diffractive_optics_cavity_q_calculation.html

Enhancement of Performance of Round-Trip Time Using Kalman Filtering



N. G. Goudru , R. P. Puneeth, and Krishna Prasad N. Rao

1 Introduction

Congestion control is one of the important key components for seamless flow of data over the Internet. The success of dominant transport protocol TCP Reno and its versions are due to handling efficiently the congestion control mechanism. One of the factors that contribute for the performance of throughput and bandwidth utilization of individual users is the round-trip delay. One of the ways of verifies that the network system performing fine is by calculating the total path delay (RTT). A fair approach to monitor RTT is to monitor the time taken for a TCP packet to be acknowledged. There are several factors affecting RTT delay such as clock time, the geographical distance, the number of hops, the queue length at the bottleneck like and so on. A simple model-based formula extensively used for RTT estimation in most of the research work [11–13] is influenced with queuing delay and the propagation delay and is given by $R(t) = \frac{q(t)}{C_d} + T_p$.

In wired networks, the transmission delay is negligible. The queue $q(t)$ is the number of packets queuing at the router. In each node, the RTT delay from current node to the destination is estimated. Many researchers recognize that the network state variables such as queuing delay and RTT delay are essential for efficient congestion control.

N. G. Goudru (✉)

Department of Information Science and Engineering, Nitte Meenakshi Institute of Technology (affiliated to Visvesvaraya Technological University), Bengaluru, India
e-mail: goudru.ng@nmit.ac.in

R. P. Puneeth · K. P. N. Rao

Department of Computer Science and Engineering, NMAM Institute of Technology (affiliated to Visvesvaraya Technological University), Nitte, Udupi, India
e-mail: puneeth.rp@nitte.edu.in

K. P. N. Rao

e-mail: krishnaprasad@nitte.edu.in

2 Literature Review

Heterogeneous communication network with heterogeneous traffic of time varying traffic load poses complex and challenging problem in providing guaranteed QoS. RTT is an important parameter for efficient end-to-end congestion control through which guaranteed QoS can be achieved. The authors in [1] developed a model for latency in data transfer for a single TCP connection considering the delay contribution due to connection establishment (three-way handshake) and the latency due to transfer a given amount of data. During data transfer, all the operations in TCP like packet loss, request for packet retransmission and delayed acknowledgement are considered. The researchers in [2] propose a new model that combines Kalman filter with the RTT model for RTT estimation at data link layer level using network state information. The disadvantage of this method is that it can detect long lived TCP flows but filter out the short lived TCP flows as noise which is incorrect. In reality, short lived TCP flows are also contributing for network traffic. The authors in [3] demonstrate the distribution of RTT values for one way that is from sender to the receiver or receiver to a sender for a TCP traffic. It is because the path traversed by a data packet to reach the destination is totally different from the path followed by an acknowledgement packet to reach the source. In the research work, the authors use RTT values to validate the flow models and to study the effect of RTT on sender window dynamics. The researchers in [4] tried to analyse worst case round-trip time (WCRTT), predicted by applying extreme value theory for the first time. Two extreme values called block maxima and threshold exceedances were applied for WCRTT prediction more reliable. The investigators in [5] studied RTT estimation suitable for congestion control applications. A Kalman filter is combined with CUSUM change detection is proposed in the work. The RTT algorithm works fine for real-time data but treats short TCP signals as noise which is an absurd. The authors in [6] propose a congestion control algorithm for real-time communication using Kalman filter for the estimation of end-to-end one way delay variation from source to destination. This results in maintaining queue delays low. But the idea of using TCP and RTT in real-time communication is advisable only when there are sufficient network resources. The researchers in [7] develop a new algorithm for RTT estimation between two hosts. The hosts are communicating the border routers using TCP protocol where bulk data is being transmitted. The algorithm is used to find RTT in large-scale high-speed networks, but does not help a service provider to provide guaranteed QoS to his clients. The investigators in [8] developed an expert framework using machine learning. The proposed approach is capable of tuning very quickly for the variation that can happen in RTT by reducing the packet losses. The authors in [9] develop an algorithm for RTT estimation for a TCP application in high bandwidth delay. The results obtained are compared with passive or normal bandwidth delay networks. The disadvantage of the algorithm is about not filling its bandwidth \times delay product that differs practically. The researchers in [10] develop an algorithm combining the TCP connection establishment using 3-way handshake and transfer of data packet based on slow start phase of TCP. The algorithm results are validated

by using NLANR traces for both short and long time scales. The research survey reveals that no adequate work has been carried out by the researchers on RTT delay estimation from the TCP perspective. In this research work, a novel discussion has been made to improve the RTT estimation using stochastic model by eliminating the noise through the application of Kalman filtering.

3 System Model for Round-Trip Time

The modified model [11–13] representing round-trip time for a single hop wired networks after introducing one-dimensional Kalman filtering is

$$R(t) = \frac{q(t)}{C_d} + T_p + \text{bias} \quad (1)$$

where T_p is the delay due to propagation in the wired medium, $q(t)/C_d$ model the delay due to the formation of queue in the bottleneck link. The term bias describes the introduction of noise into the signals during transmission, and it acts as a correction factor added to the RTT value. The noise induced affects the true RTT value. A stochastic model representing the behaviour of packets queue at the ingress point of a router is

$$\frac{\partial q(t)}{\partial t} = \frac{w(t)}{R(t)} - C_d. \quad (2)$$

where the term $w(t)/R(t)$ describes the increase in queue length by TCP flow of the sender, $C_d = q(t)/R(t)$ represents decrease in queue length because of packet servicing by the router. It is also called as delay due to packet departure from the router. Model describing the TCP window dynamics is given by

$$\frac{\partial w}{\partial t} = \frac{1}{R(t)} - \frac{w(t)w(t - R(t))}{2R(t - R(t))} p(t - R(t)). \quad (3)$$

TCP window works based on the principle of additive increase and multiplicative decrease congestion control. On the right hand side of Eq. (3), the first term describes additive increase of the sender window size. The window increases linearly until the sender notices congestion. The second term describes multiplicative decrease of the window. Model representing the probability of packet dropping due to random early detection (RED) is

$$P(t) = \begin{cases} 0, & q(t) \in [0, t_{\min}] \\ \frac{q(t) - t_{\min}}{W_{\max} - t_{\min}} P_{\max}, & q(t) \in [t_{\min}, W_{\max}] \\ 1, & q(t) \geq W_{\max} \end{cases} \quad (4)$$

All incoming packets are received by the router when the buffer is between $[0, t_{\min}]$, starts dropping the packets with a provability value when the buffer is between $[t_{\min}, w_{\max}]$ and all incoming packets are dropped when the buffer overflows when queue length is greater than w_{\max} .

A. Estimation of Bias Term

- (i) Capture the RTT values from the model called RTT estimation. Determine the variance. Find the interval of variance which contains lower bound and upper bound. $E_0 =$ lower bound.
- (ii) Capture the RTT values from the Internet using Wireshark, called RTT measurement. Determine the variance. Find the interval of variance which contains lower bound and upper bound. $E_{\text{mea}} =$ upper bound.
- (iii) Determine initial Kalman Gain (K_0).

$$K_0 = \frac{E_0}{E_0 + E_{\text{mea}}}$$

- (iv) $\text{bias}_t = \text{bias}_{t-1} + K_t(Y_t - \text{bias}_{t-1})$
where Y_t is RTT measurement.
- (v) Kalman Gain, $K_t = \frac{E_{\text{est}}}{E_{\text{est}} + E_{\text{mea}}}$
- (vi) $E_{\text{estt}} = (1 - K_t) * E_{\text{est}} - 1$

4 Simulation and Performance Analysis

Simulation was conducted by coding in MATLAB R2014a. The network model is as shown in Fig. 1. R_1 and R_2 are the routers. The network control parameters used to carry out the simulation are $C_d = 400$ packets, $P_{\max} = 0.5$, $t_{\min} = 200$ packets, $W_{\max} = 300$ packets, $t_p = 0.001$ s and packet size = 1000 bytes.

The traces of the graph of Fig. 2 give variance estimated to understand the range of possible error in estimated RTT using the stochastic model given by Eq. (1) without bias term. Total number of RTT samples taken is 500. The possible error range is $[0.4088 \text{ s} - 0.4230 \text{ s}]$ which ascertains the model error. The graph of Fig. 3 shows that traces of RTT measurement using the stochastic model given by Eq. (1) without

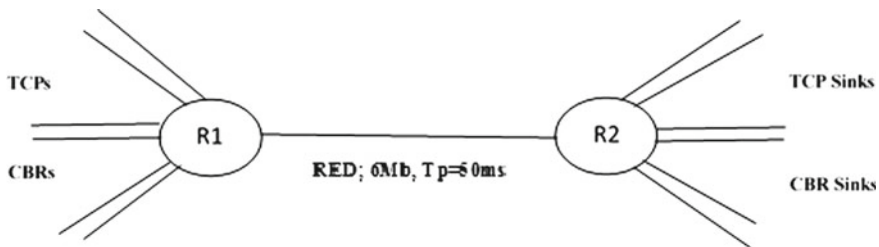


Fig. 1 Network model

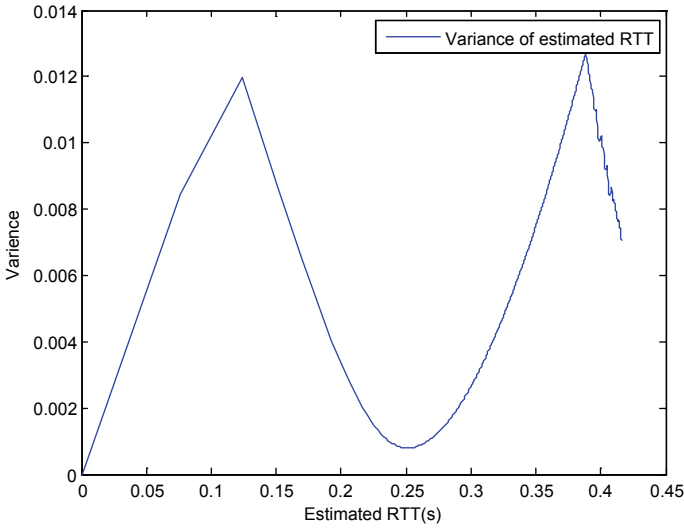


Fig. 2 Variance with respect to estimated RTT

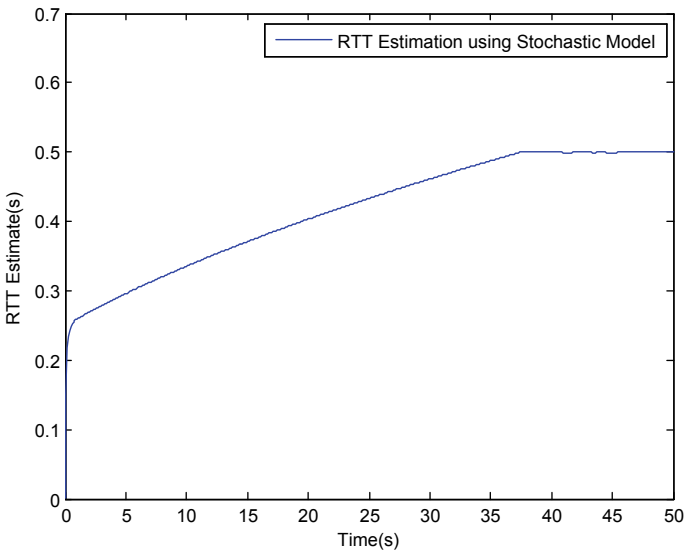


Fig. 3 RTT estimation by model without bias

bias term. The RTT estimation fluctuates over the interval of 0.218219 s–0.500197 s, having an average value of 0.41589 s. It is practically observed that end-to-end performance of TCP is higher when the RTT value is lower. If the new RTT delay is much less than the current RTT delay time, the system ignores it for some time. If

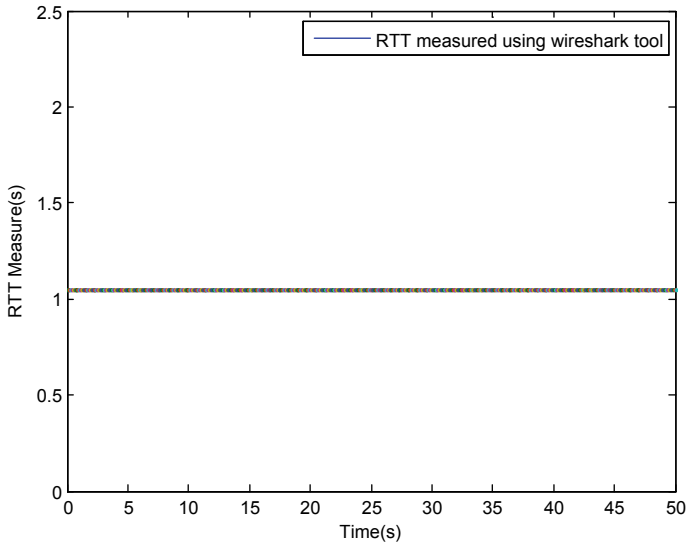


Fig. 4 Measured RTT by Wireshark

the same new RTT value continues for some more time, then the system drops the current value and adds the new RTT value. An effective RTT value can also be used as a metric for load balancing the network traffic. The traces in Fig. 4 represent the distribution of measured RTT values captured on-line using Wireshark tool for the same network parameters. RTT measurement values vary over the range of 0.2527 s–2.9042 s, with an average value of 0.6402 s. The graph of Fig. 5 demonstrates the distribution of Kalman gain. Kalman gain is the ratio of error in estimated RTT to the sum of error in estimated RTT plus error in measured RTT. Kalman gain varies over the range 0.4445–0.496 with an average value of 0.491. Kalman gain value is small when error in the measurement is relatively large compared to error in the estimate. Kalman gain is large when error in the estimate is relatively large compared to error in the measurement. The traces of the graph of Fig. 6 show the estimated RTT distribution obtained from the modified stochastic model after adding the bias term. The modified estimation model RTT value varies over the range of 0.4169 s–1.9533 s, with an average value of 0.73002 s. The graph of Fig. 7 demonstrates performance of queue along the bottleneck link. Queue length varies between 86 to 200 packets with an average value of 166 packets. When the queue length is large, RTT value in the network increases when queue is small, and RTT value decreases. RTT value directly depends on queue length. Our research work uses RED routers built based on random early detection algorithm. Model representing RED is presented in Eq. (4). The model uses two threshold values called minimum threshold value and maximum threshold value. In the experiment, minimum threshold values are 200 packets, and maximum threshold value is 300 packets. When the incoming packets are below 200 packets, RTT value is about 0.2 s. As the incoming packets increase, queue formation

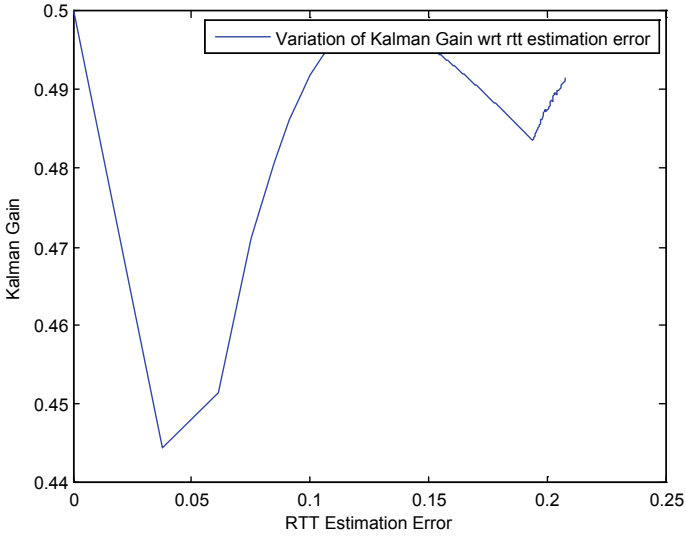


Fig. 5 Kalman gain with respect to estimated RTT

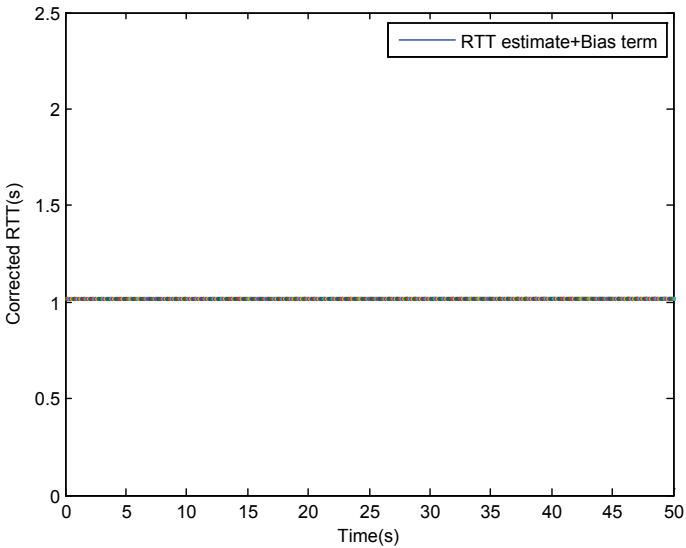


Fig. 6 RTT estimation by modified model with bias

starts, and RTT value increases. When the incoming packet exceeds the maximum threshold value called burst flow, the router discards the packets. The traces of the graph of Fig. 8 show the dynamics of sender TCP window. It varies over the range of 100 packets to 200 packets having an average value of 167 packets. A TCP flow is

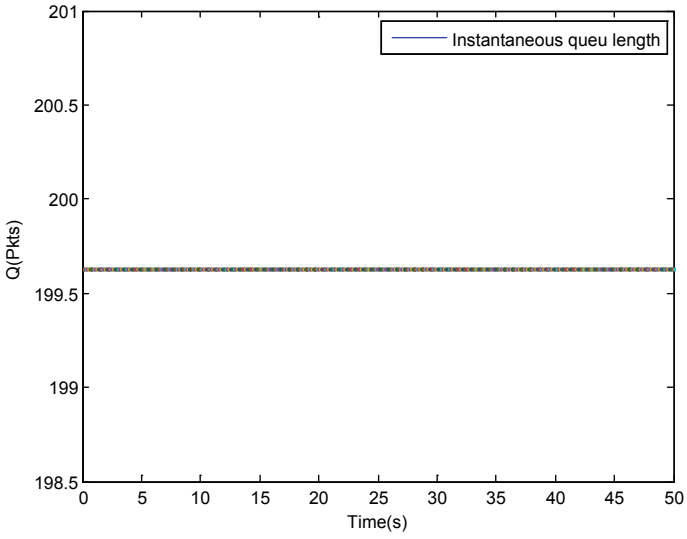


Fig. 7 Performance of queue at bottleneck link

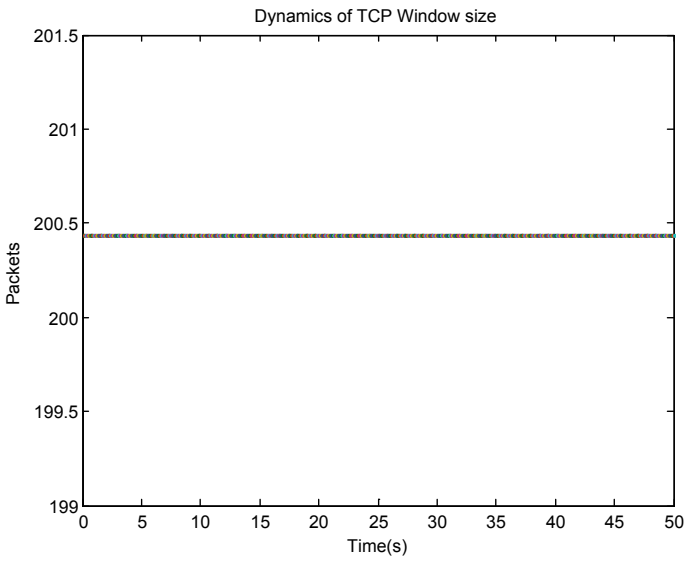


Fig. 8 Performance of sender window size

inversely proportional to its RTT distribution. TCP window dynamics given in model (3) is a rate-based model where rate defines an average number of bits transferred per second.

5 Conclusion

In this research work, an attempt is made to enhance the performance of RTT estimation model used extensively by the researchers for network traffic analysis. The error term in the model given in Eq. (1) is called bias term. The bias term eliminates the noise that introduced into the signals during data transmission over the Internet. The bias term is calculated by the application of Kalman filtering. The error in RTT estimation using model (1) without bias term (Fig. 3) has been reduced, and the values are smoothed after adding the bias term (3) (Fig. 6). The accuracy of the performance of the modified model is validated by comparing the graphs given in Figs. 3 and 6. The estimated RTT values are nearly converging with the measured RTT values. The accuracy in the modified RTT model in turn influences the performance of queue length and the sender window dynamics. This can be observed by comparing the graphs shown in Figs. 7 and 8. The queue length gets smooth and converges to a value of 199.6 packets. Similarly, window size converges to a value of 200.4 packets. The source sending rate matches with the queue length minimizing the congestion and packet loss.

References

1. N. Cardwell, S. Savage, T. Anderson, *Modeling Latency TCP* (Department of Computer science and Engineering, University of Washington, Seattle, WA 98195, USA, 2000)
2. K. Jacobsson, H. Hjalmarsson, N. Moller, K. Henrik Johansson, *Estimation of RTT and Bandwidth for Congestion Control Applications in Communication Networks* (Department of Signals, Sensors and Systems, KTH, SE-100 44 Stockholm, Sweden, 2005)
3. S. Shakkottai, R. Srikant, *The RTT Distribution of TCP Flows in the Internet and Its Impact on TCP-Based Flow Control* (University of Illinois, Illinois, USA, 2005).
4. Z. Zhu, A.M.K Cheng, *Worst Case Round Trip Time Prediction and Statistical Analysis Using Extreme Values Theory* (Real-time system laboratory, Department of Computer Science, University of Houston, Texas, 77204, USA, 2005)
5. K. Jacobsson, H. Hjalmarsson, N. Moller, Karl Henrik Johansson, *Round-Trip Time Estimation in Communication Networks Using Adaptive Kalman Filtering* (Department of Signals, Sensors and Systems, KTH, SE-100 44 Stockholm, Sweden, 2005)
6. G. Carlucci, L. De Cicco, S. Holmer, S. Mascolo, *Congestion Control for Web Real-Time Communication* (IEEE/ACM Trans. Netw, 2017).
7. Q. Su, J. Gong, X. Hu, RTT estimation with sampled flow data. *IEICE Trans. Commun.* **E98-B**(9) (2015)
8. B.A.A. Nunes, K. Veenstra, W. Ballenthin, S. Lukin, K. Obraczka, A Machine learning approach to end-to-end RTT estimation and its application to TCP. 978-1-4577-0638-7 /11/\$26.00 ©2011 IEEE

9. I. Prieto, M. Izal, E. Magaña, D. Morato, A simple passive method to estimate RTT in high bandwidth-delay networks, in *The Seventh International Conference on Evolving Internet* (Public University of Navarre, Navarre, Spain, 2015)
10. H. Jiang, C. Dovrolis, *Passive Estimation of TCP Round-Trip Times* (University of Delaware, Computer and Information Sciences, 2011).
11. Z. Zhao, S. Darbha, A.L. Narasimha Reddy, A method for estimating the proportion of nonresponsive traffic at a router. *IEEE/ACM Trans. Netw.* **12**(4) (2004)
12. C.V. Hollot, Y. Liu, V. Misra, D. Towsley, Unresponsive flows and AQM performance. *IEEE INFOCOM* (2003)
13. N. G. Goudru, B.P. Vijayakumar, Evaluation of TCP performance using NDG loss predictor in wireless networks. <https://doi.org/10.1109/ADCOM.2015.10>

Automatic Modulation Classification Using Cumulants and Ensemble Classifiers



M. Venkata Subbarao and P. Samundiswary

1 Introduction

Security is the key issue in many communication applications especially for military applications. In contrast to this, the authorities of boarder security may wish to monitor unlicensed transmitters for jamming their signals [1]. The necessary action of doing so is to identify or recognize the modulation class of that intercepted signal. Such type of actions also arises in several other applications such as interference management, signal authorization, verification, and selection of appropriate demodulation techniques in electronic combat, threat analysis, and so on. Modulation recognition is also useful to recognize the suspicious transmitter in the near geographical site and to generate jamming signals to stop communication between the suspicious users. In recent civilian applications, a greater number of modulation formats can be employed by a transmitter to manage the data rate, to reduce the individual bandwidths of every user, and to assure the integrity of the message [2]. However, the group of modulation formats is known both to transmitter and receiver. The choice of the modulation format is adaptive and may not be known at the receiving end. Therefore, an AMC mechanism is required for the receiving end to recognize modulation format of the received signal and to select the proper demodulation approach in order to recover the original message. Moreover, in civilian applications, several techniques developed to reduce overhead of reference signals required for channel estimation have motivated the research in blind and semi-blind MIMO techniques [3].

M. Venkata Subbarao (✉)

Department of ECE, Shri Vishnu Engineering College for Women, Bhimavaram, Andhra Pradesh, India

e-mail: mandava.decs@gmail.com

P. Samundiswary

Department of EE, Pondicherry University, Kalapet, India

e-mail: samundiswary_pdy@yahoo.com

Blind techniques are also expected to play a role in software defined radio and cognitive radio. Configuration information required by a software defined radio system is transmitted as overhead to the data. Intelligent receivers capable of extracting this information blindly may improve transmission efficiency through reductions in overhead, i.e., automatic modulation classification eliminates the need for supplementary information on the modulation type [4].

AMC techniques are mainly classified into decision theoretic (DT) or maximum likelihood and feature-based (FB) or pattern recognition approaches. Some of the AMC approaches are tabulated in Table 1. From the literature, it is observed that many authors considered a basic modulation classes as a dataset. So that the accuracy achieved is about 90%. This paper performs AMC of higher order modulation classes.

Table 1 AMC techniques

Ref. No.	Approach type	Features used	Modulation type	Accuracy (%)
[5]	DT	Ratio of variance of envelope to square of mean	AM, FM, SSB, DSB	80–95
[6]	DT	Likelihood ratio of Phase	BPSK, QPSK	100
[7]	DT	Instantaneous amplitude and phase	AM, DSB, SSB, VSB, LSB, USB, FM	91–100
[8]	FB	Instantaneous amplitude, phase, freq, spectrum symmetry	ASK2, ASK4, PSK2, PSK4, FSK2, FSK4	97
[9]	DT	I-Q data	16QAM, 32QAM, 64 QAM	100
[10]	FB	Combination of spectral and higher order cumulant	QAM16, QAM64, ASK2, ASK4, PSK2, PSK4, FSK2, FSK4	98
[11]	FB	Power and cyclic spectral features	QPSK, PSK8, QAM8, 16QAM CW, FSK2, FSK4, FSK8, ASK4, BPSK	94.9
[12]	DT	Moments	PSK4, PSK8, QAM16 ASK4, ASK8, FSK4, BPSK,	89.76
[13]	FB	Moments of wavelet	ASK4, ASK8, PSK8, QAM16,32, 64, FSK4, BPSK, PSK4	98
[14]	FB	Energy of signal, zero crossing rate, variance	ASK4, ASK8, PSK8, QAM16, 32, 64 FSK4, BPSK, PSK4	91
[15]	FB	Two stage classification	BPSK, QPSK, QAM16 & 64	70–100
[16]	DT	Cumulants	16 QAM, 64 QAM, BPSK, QPSK & 8 PSK	50–90

Organization paper is as follows: Sect. 1 describes about the scenario of existing wireless communication systems and their current challenges. Also, an elaborate study on existing different AMC techniques is presented. Section 2 deals with the extraction of different higher order statistical features and framework of PR approach. Section 3 presents the development of PRC for AMC algorithms using ensemble classifiers. Section 4 deals with the performance analysis of proposed ensemble-based PRCs under non-ideal channel conditions. Section 5 depicts the important surmises of the research work.

2 Framework

The functional block diagram representation of proposed PR classification model is shown in Fig. 1. In order to classify the modulation classes, a set of features are extracted from each class. The set of features are considered in the proposed approach are cumulants, and they are extracted from the moments.

The received signal $y(n)$ by a communication receiver is

$$y(n) = Me^{i(2\pi nTf_o + \theta_n)} \sum_{l=-\infty}^{\infty} x(l)h((n-l+\epsilon)T) \quad (1)$$

where T is duration of the symbol, M is the magnitude, θ_n is phase shifts and ϵ is time shifts caused by the channel. $h(n)$ is channel impulse response, and $x(l)$ is input binary data.

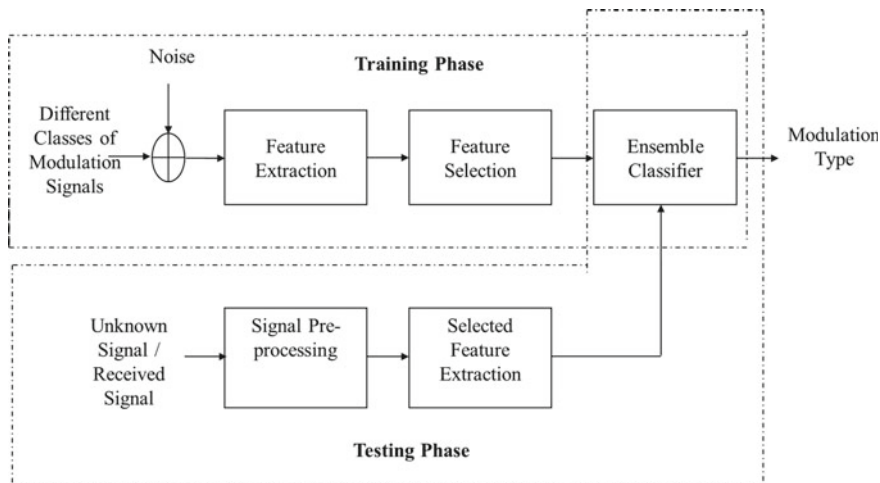


Fig. 1 Proposed classification model

The moments of signal $y(n)$ are

$$M_{ab} = E[y(n)^{a-b} y^*(n)^b] \tag{2}$$

Here, a and b are integers. $y(n)$ and $y^*(n)$ are received signal and *its* conjugate.

Multi-Order moments (for $a=2, 4, 6$ and 8) are derived from Eq. 2. To train the classifier a set of 11 cumulants of order 2, 4, 6 and 8 are derived from the moments [17].

3 Ensemble Classifiers

The performance of PR classifiers such as DT, KNNs, and SVMs is varied across different data sets because of their strategy in classification. Prediction of best model for all data sets is not at all possible for all the time. Keeping this in mind, the proposed classifiers are built with a combination of various classes of classifiers to achieve better accuracy than individuals. It works on a principle of majority vote, i.e., initially individual classifiers identified the class of unknown signal and finally ensemble classifiers identify the majority vote, and gives the class label for unknown signal. An ensemble classifier contains a set of independently trained classifiers whose predictions are pooled when recognizing new object. The block diagram representation of an proposed classifier is shown in Fig. 2 [17].

In this paper, a variety of classifiers are used for AMC. These are boosted trees, bagged trees, subspace KNN, RUSBoosted trees, and discriminant KNN. Boosted trees and bagged are constructed from decision trees, and these are slower in speed.

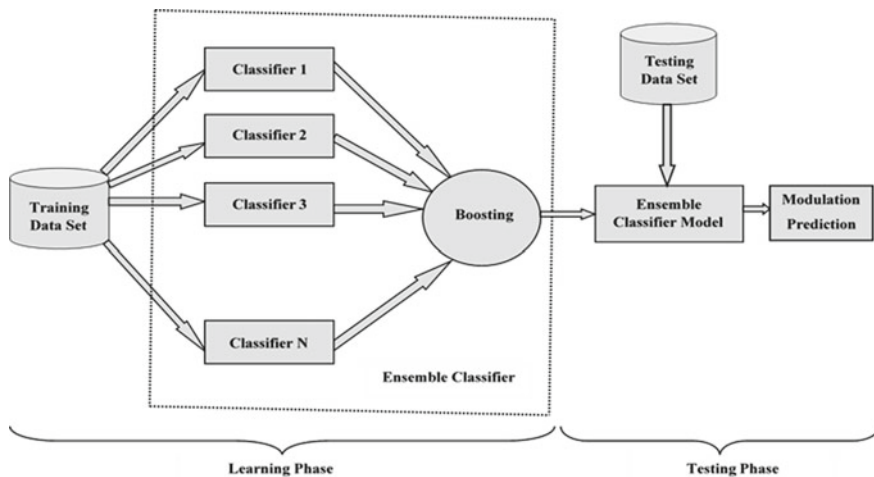


Fig. 2 Ensemble classification model

Whereas, discriminant KNN and subspace KNN both used nearest neighbors and discriminant analysis. Based on random undersampling, weak learners are boosted in RUSBoosted trees.

4 Simulation Results

In this section, the performance analysis of proposed ensemble-based PRCs is carried out under SNR of 0 dB to 20 dB with modulation classes of M-ary QAM ($M = 4, 16$ and 64) and M-ary PSK ($M = 2, 4$ and 8). The simulation parameters are tabulated in shown in Table 2.

The performance of five ensemble classifiers is tested with different SNR conditions. Table 2 represents the confusion matrix for different ensemble classifiers using multi-order cumulants at different SNRs. Diagonal elements in the confusion matrix denote the true classification rates, and off diagonal elements represent misclassification rates. From Table 3, it is proved that the proposed ensemble classifiers are able to distinguish the different modulating classes at low SNR than that of existing classifiers.

The performance accuracy of boosted tree classifier for different modulation classes with different SNR values is shown in Fig. 3. The average modulation classification accuracy of boosted tree from 0 dB to 20 dB is 81.7%, 95%, 98.3%, 98.9%, and 100%, respectively.

The classification performance of proposed bagged tree classifier with 90% training is shown in Fig. 4. At 0 dB, the average classification accuracy of bagged tree classifier is 81.7%. The classification performance of proposed subspace discriminant classifier with 90% training is shown in Fig. 5. At 0 dB, the average classification accuracy of subspace discriminant classifier is 77.8%.

The classification performance of proposed subspace KNN classifier with 90% training is shown in Fig. 6. At 0 dB, the accuracy of subspace KNN is 72.2% because the classifier is unable to distinguish higher order modulation classes at lower SNR values.

Table 2 Simulation parameters

Parameters	Description
Modulation types	MQAM ($M = 4, 16$ and 64), MPSK ($M = 2, 4, 8$)
Channels	Fading and Gaussian
SNR	0–20 dB
Data set	6000 * 11 (11 features, 1000 of each class)
Data set for training	50-90%
Data set for testing	10-50%

Table 3 Confusion matrix of proposed ensemble classifiers with 90% training

Classifier	True class	SNR																																							
		0					5					10					15					20																			
		0	1	2	3	4	0	1	2	3	4	0	1	2	3	4	0	1	2	3	4	0	1	2	3	4	0	1	2	3	4										
Boosted trees	BPSK	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	QPSK	0	97	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	8PSK	0	0	84	0	3	13	0	0	97	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	4QAM	0	0	0	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	16QAM	0	0	33	0	30	37	0	0	0	0	0	0	87	13	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	64QAM	0	3	10	0	7	80	0	0	0	0	0	0	10	90	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Bagged trees	BPSK	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	QPSK	0	97	0	0	3	0	0	0	97	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	8PSK	0	0	83	0	7	10	0	0	0	97	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	4QAM	0	0	0	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	16QAM	0	0	30	0	50	20	0	0	0	0	0	0	90	10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	64QAM	0	3	10	0	27	60	0	0	0	0	0	0	10	90	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Subspace discriminant	BPSK	97	3	0	0	0	0	0	0	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	QPSK	0	93	0	0	0	7	0	93	0	0	0	7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	8PSK	0	0	80	0	17	3	0	0	0	97	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	4QAM	0	0	0	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	16QAM	0	0	47	0	30	23	0	0	0	0	0	0	90	10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	64QAM	0	6	7	0	20	67	0	0	0	0	0	0	17	83	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Subspace KNN	BPSK	97	3	0	0	0	0	0	0	97	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	QPSK	13	77	4	0	3	3	0	90	0	0	0	10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	8PSK	0	3	63	4	23	7	0	0	0	90	0	10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	4QAM	0	0	0	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	16QAM	3	0	27	0	40	30	0	0	0	4	0	73	23	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
	64QAM	0	13	0	0	27	57	3	0	4	0	23	70	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

(continued)

Table 3 (continued)

Classifier	True class	SNR																																							
		0					5					10					15					20																			
		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
RUSBoosted trees	BPSK	100	0	0	0	0	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	QPSK	0	93	0	0	0	7	0	97	0	0	3	0	100	0	0	0	0	0	0	0	0	100	0	0	0	0	0	0	0	0	0	100	0	0	0	0	0	0	0	0
	8PSK	0	0	73	0	14	13	0	0	97	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	4QAM	0	0	0	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	16QAM	0	0	23	0	40	37	0	0	0	0	0	87	13	0	0	0	3	0	90	7	0	0	0	0	0	0	0	0	0	0	0	97	3	0	0	0	0	0	0	0
	64QAM	0	7	10	0	3	80	0	0	0	0	0	10	90	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	97	0	0	0	0	0	0	0

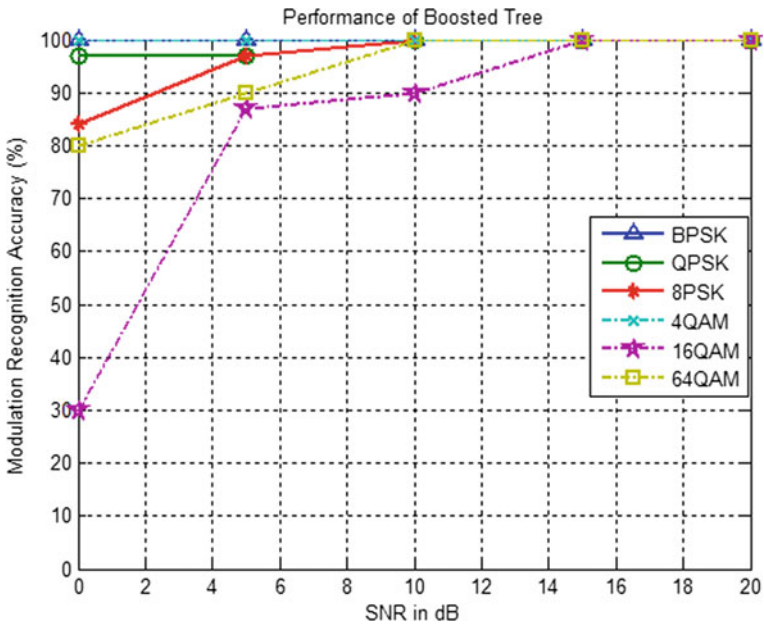


Fig. 3 Performance of boosted tree classifier

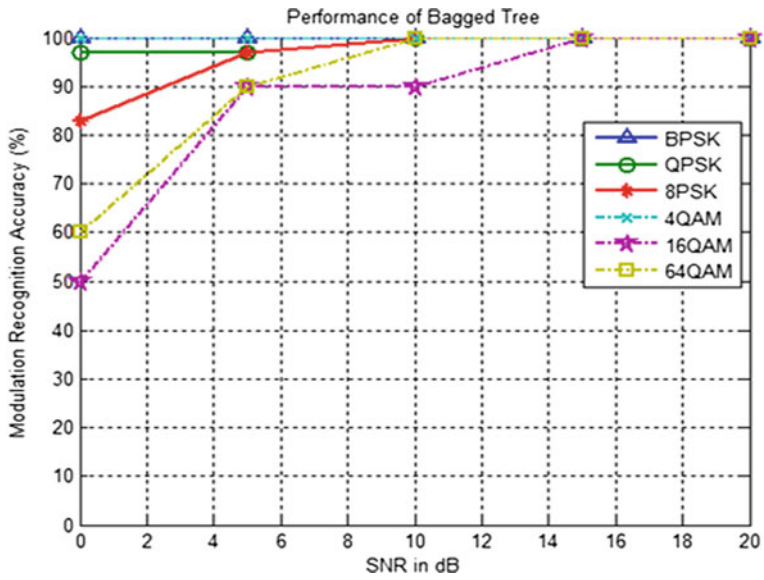


Fig. 4 Performance of bagged tree classifier

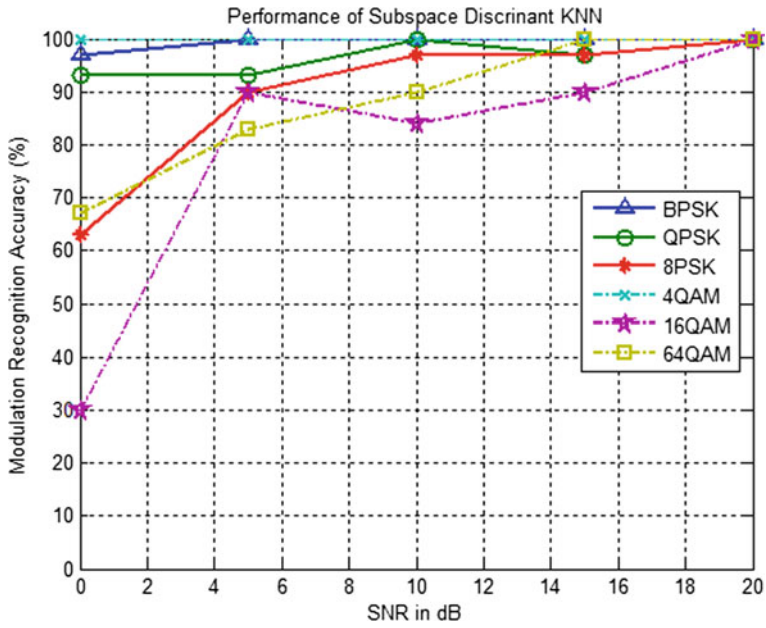


Fig. 5 Performance of subspace discriminant KNN classifier

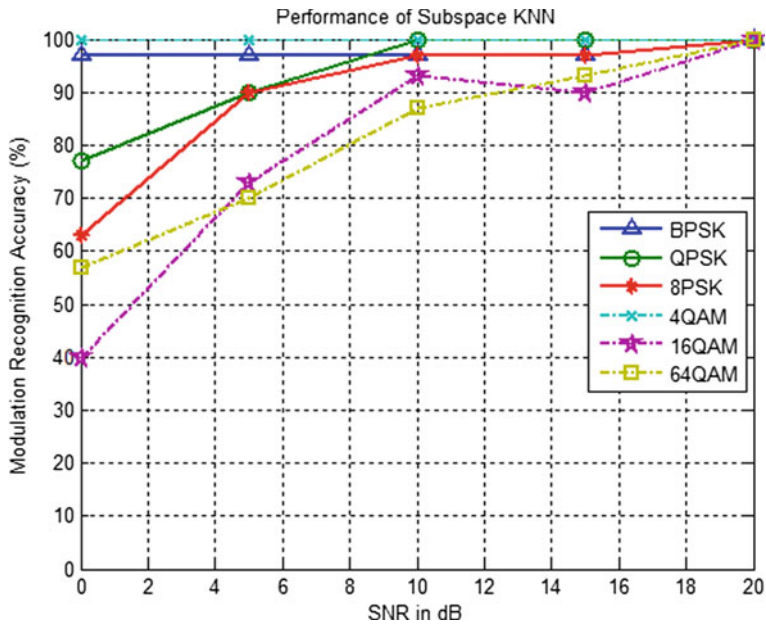


Fig. 6 Performance of subspace KNN classifier

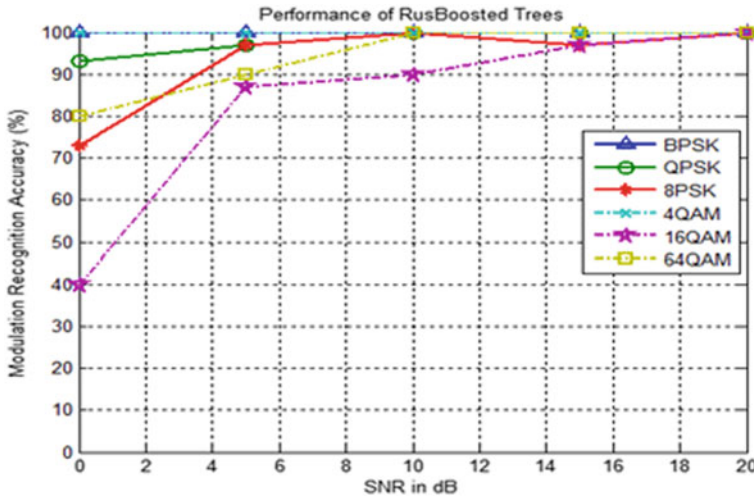


Fig. 7 Performance of RUSBoosted tree classifier

The classification performance of proposed RUSBoosted tree classifier with 90% training is shown in Fig. 7. At 0 dB, the average classification accuracy of RUSBoosted tree is 81.1%. From simulations, bagged, boosted, and RUSBoosted tree classifiers have higher classification accuracy than others.

At different training rates, the performance of ensemble classifiers is tabulated in Table 4. It is showed that the accuracy of classifiers is higher even at high testing rates. Among all the classifiers, bagged tree classifier has the highest accuracy.

5 Simulation Results

In this paper, a variety of ensemble classifiers is developed for AMC. Thereafter, performance of proposed ensemble classifiers is analyzed under non-ideal channel conditions with various values of SNR along with different cases of training. Further, performance of proposed ensemble classifiers is compared with the existing approaches to prove the efficiency of the proposed classifiers in modulation recognition. From the simulation results, it is proved that even with more classes the proposed ensemble classifiers achieve more classification accuracy even at lower SNRs.

Table 4 Performance at different training rates

Training %	Classifier	Classification accuracy (%)				
		0 dB	5 dB	10 dB	15 dB	20 dB
80	Boosted Trees	81.2	94.6	99.2	99.5	100
	Bagged Trees	81.2	95.2	99.3	99.3	100
	Subspace Discriminant KNNs	77.5	93.4	96.8	97.5	100
	Subspace KNNs	71.6	86.2	96.0	96.9	100
	RUSBoosted trees	80.8	94.7	98.8	98.7	100
70	Boosted trees	80.4	94.1	98.8	99.2	100
	Bagged trees	80.3	94.8	99.1	99.3	100
	Subspace discriminant KNNs	76.2	93.1	96.2	97.4	100
	Subspace KNNs	71.1	85.4	95.8	96.4	100
	RUSBoosted trees	79.9	94.2	98.3	98.2	100
60	Boosted trees	80.1	93.8	98.3	99.1	100
	Bagged trees	80.1	94.1	98.7	99.5	100
	Subspace discriminant KNNs	75.7	92.5	95.6	97.7	100
	Subspace KNNs	70.9	85.4	95.2	96.5	100
	RUSBoosted trees	79.8	93.9	97.9	98.6	100
50	Boosted trees	79.8	93.3	98.1	98.9	100
	Bagged trees	79.8	93.7	98.2	99.3	100
	Subspace discriminant KNNs	75.4	92.1	95.6	97.2	100
	Subspace KNNs	70.1	85.3	95.1	96.2	100
	RUSBoosted trees	79.2	93.7	96.8	98.6	100

References

1. M.Venkata Subbarao, P.Samundiswary, An intelligent cognitive radio receiver for future trend wireless applications. *Int. J. Comput. Sci. Inf. Secur. (IJCSIS)*, **14**, 7–12 (2016)
2. M. Venkata Subbarao, P. Samundiswary, *Spectrum Sensing in Cognitive Radio Networks Using Time–Frequency Analysis and Modulation Recognition*, vol. 471, pp. 827–837 (Springer Lecture Notes in Electrical Engineering, 2018)
3. M.R. Bahloul, M.Z. Yusoff, A.-H. Abdel-Aty, M. Naufal, M. Saad, M. Al-Jemeli, Modulation Classification for MIMO Systems: State of the Art and Research Directions, vol. 89, pp. 497–505 (Chaos, Solitons & Fractals, 2016). <https://doi.org/10.1016/j.chaos.2016.02.029>
4. X. Lin, Y.A. Eldemerdash, O.A. Dobre, S. Zhang C. Li, Modulation classification using received signal’s amplitude distribution for coherent receivers, *IEEE Photonics Technol. Lett.* **29**(21), 1872–1875. 1 Nov 2017. <https://doi.org/10.1109/lpt.2017.2754501>
5. Y. Chan, L. Gadbois, P. Yansouni, Identification of the modulation type of a signal, in *ICASSP’85. IEEE International Conference on Acoustics, Speech, and Signal Processing*, Tampa, FL, USA, pp. 838–841 (1985). <https://doi.org/10.1109/icassp.1985.1168289>
6. C.-Y. Huan, A. Polydoros, Likelihood methods for MPSK modulation classification. *IEEE Trans. Commun.* **43**(2/3/4), 1493–1504 (February/March/April 1995). <https://doi.org/10.1109/26.380199>

7. A.K. Nandi, E.E. Azzouz, Algorithms for automatic modulation recognition of communication signals. *IEEE Trans. Communi.* **46**(4), 431–436 (1998). <https://doi.org/10.1109/26.664294>
8. G. Arulampalam, V. Ramakonar, A. Bouzerdoum, D. Habibi, Classification of digital modulation schemes using neural networks, in *ISSPA'99. Proceedings of the Fifth International Symposium on Signal Processing and its Applications (IEEE Cat. No.99EX359)*, vol. 2, pp. 649–652 (Brisbane, Queensland, Australia, 1999). <https://doi.org/10.1109/isspa.1999.815756>
9. W. Wei, J.M. Mendel, Maximum-likelihood classification for digital amplitude-phase modulations. *IEEE Trans. Commun.* **48**(2), 189–193 (2000). <https://doi.org/10.1109/26.823550>
10. M.L.D. Wong, A.K. Nandi, Automatic digital modulation recognition using artificial neural network and genetic algorithm. *Signal Process.* **84**(2), 351–365 (2004). <https://doi.org/10.1016/j.sigpro.2003.10.019>
11. Y.-x. Wu, L.-d. Ge, F.-f. Liu, Comprehensive features based digital modulation identification using a neural tree network, in *Proceedings of 2005 International Conference on Communications, Circuits and Systems*, pp. 752 (Hong Kong, China, 2005). <https://doi.org/10.1109/iccascas.2005.1495219>
12. A. Ebrahimzadeh, A. Ranjbar, Intelligent digital signal-type identification. *Eng. Appli. Artif. Intell.* **21**(4), 569–577 (2008). <https://doi.org/10.1016/j.engappai.2007.06.003>
13. K. Hassan, I. Dayoub, W. Hamouda, M. Berbineau, Automatic modulation recognition using wavelet transform and neural network, in *2009 9th International Conference on Intelligent Transport Systems Telecommunications*, (ITST), pp. 234–238 (Lille, 2009). <https://doi.org/10.1109/itst.2009.5399351>
14. A. Ebrahimzadeh, H. Azimi, S.A. Mirbozorgi, Digital communication signals identification using an efficient recognizer. **44**(8), 1475–1481 (2011). ISSN 0263-2241. <https://doi.org/10.1016/j.measurement.2011.05.019>
15. M.W. Aslam, Z. Zhu, A.K. Nandi, Automatic modulation classification using combination of genetic programming and KNN. *IEEE Trans. Wirel. Commun.* **11**(8), 2742–2750 (2012). <https://doi.org/10.1109/TWC.2012.060412.110460>
16. M.S. Muhlhaus, M. Oner, O.A. Dobre, F.K. Jondral, A low complexity modulation classification algorithm for MIMO systems. *IEEE Commun. Lett.* **17**(10), 1881–1884 (2013). <https://doi.org/10.1109/LCOMM.2013.091113.130975>
17. M. Venkata Subbarao, P. Samundiswary, Performance analysis of modulation recognition in multipath fading channels using pattern recognition classifiers. *Wireless Pers. Commun.* **115**, 129–151 (2020). <https://doi.org/10.1007/s11277-020-07564-z>

Automation Testing and Validation of Electric Drive System



Raqheeba Taneem and Krishnananda Shet

1 Introduction

The use of electronic frameworks in automotive industry is constantly growing, even at quicker pace. Contrast to the past models of vehicle, every single novel model has more capacities, intended to improve security, connectivity, drivability and also solace. Growing vehicle capacities imply increasing count of electronic control modules with progressively complex programming. Because of the multifaceted nature and size of the control programming, it is required to utilize a development procedure which will improve the general quality of the firmware, increment in development effectiveness and wiping out systematic firmware bugs.

The standard programming development procedure utilized in the automotive industries is the V-cycle as shown in Fig. 1.

V-cycle is the conventional software development process utilized in the automotive domain [1]. The V-model parts the software development process into a few stages that incorporates prerequisite definitions, programming configuration, plant model design, testing and validation processes.

The initial step of the process involves the analysis of functional requirements. Next, based on these requirements, model development is done in which plant models are developed using Simulink or Scilab tools. A series of tests called model-in-loop tests are carried on the developed model to test the system functionality. Further, the ECUs are programmed in C language or other suitable scripts. Also, C code is extracted from the model and compiled to check if the required functionality is met, using the software-in-loop tests. Further, a number of ECUs are integrated and tested

R. Taneem (✉) · K. Shet
Department of E&CE, NMAM Institute of Technology, Nitte (Visvesvaraya Technological University, Belagavi), Udupi, India
e-mail: raqheebataneem33@gmail.com

K. Shet
e-mail: konikrishna@nitte.edu.in

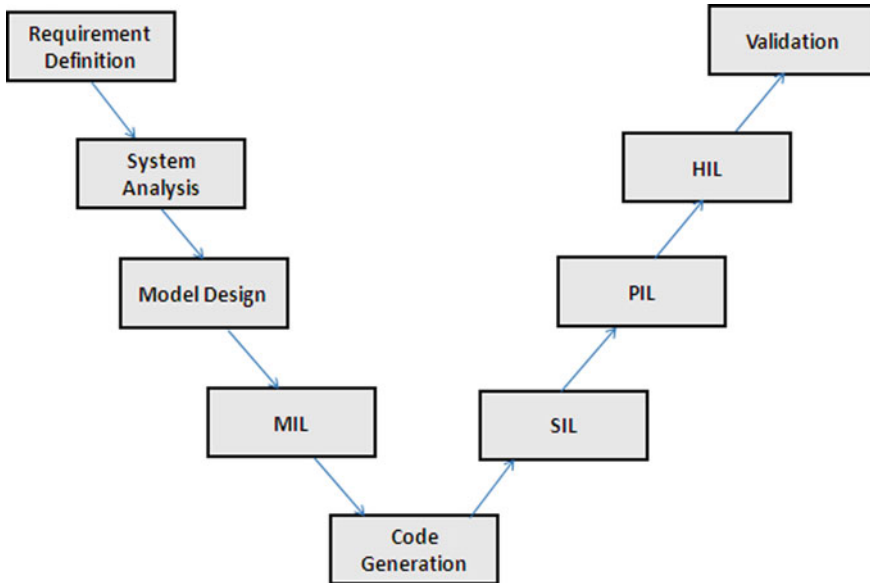


Fig. 1 V-cycle process for automotive testing and validation

using hardware-in-loop concept. In the HIL setup, a simulator is used that mimics a real vehicle. The simulator provides all the necessary electrical connections for ECUs like power supply, battery, sensors, actuators, etc. Thus, the HIL setup mimics the real vehicle behavior in an indoor environment. This validation process ensures the correctness of the developed model and verifies that the functional requirements are met.

2 Literature Review

The following is the analysis of the work carried out by several researchers.

The work proposed by Ili, Velibor, SrđanPopi and Milan Kovači [2] depicts methods for secure and automated data flow of ECUs testing procedure. In the explained testing method, it is prescribed to utilize the computerized testing conditions that create definite reports about each fragment of tested ECU. The automation testing process decreases human errors done while manual testing. The automation testing is done with the help of programming languages. The test results are obtained in automatically generated test reports in html, txt formats.

Ramaswamy, Deepa, et al. [3] utilized two sorts of simulation model designs to test ECU. The principal kind is known as Open-Loop Test Platform (OLTP) models that mimics just the low-level input–output functionalities in the framework. The simulation of the physical plant is excluded from the OLTP. The subsequent kind,

called Closed-Loop Test Platform (CLTP) models entire physical I/O functionalities along with plant I/O processing. Both testing designs discovered broad use in the controller improvement process.

Li, Bo, et al. [4] assembled a test stage dependent on LabVIEW to examine vehicle control unit (VCU) and battery management system (BMS), and control methodologies of fabricated test stage are tested by various experiments. Utilizing equipment items, information obtaining, CAN communication, signal modeling, HIL setup is established. In this framework, real-time vehicle model and battery model are created by Simulink, and with load simulation units, a vehicle simulation condition is built up. It is demonstrated that VCU and BMS run well on the HIL test stage and meet their test prerequisites, and the combined diagnosing capability toward VCU and BMS is confirmed.

Vo-Duy, Thanh, et al. [5] proposed a signal HIL model of electric vehicles that incorporates driving framework and vehicle model running continuously in dSPACE-DS1103 control card. All the necessary conditions of electric vehicle are observed in MATLAB/Simulink platform. The model is tested and approved by running necessary test cases.

Haupt, Hagen, et al. [6] proposed HIL test for BMS. Testing with HIL test systems has become a standard methodology in ECU advancement. Creating HIL test systems for battery management frameworks caused new difficulties for HIL gadgets and real-time models. dSPACE has built up a battery cell voltage imitating board that recreates cell voltages with high accuracy and furthermore has adequate current sink and source ability to permit latent and dynamic cell balancing on the HIL test system. The new multi-cell battery models reproduce the fundamental static and dynamic impacts of the cells and are real-time proficient.

Bringmann, Eckard, et al. [7] proposed TPT test approach for Model-based Testing of Automotive Systems. Model-based improvement caused an extreme change in automotive framework advancement. The need to test sooner, to test on different incorporation levels, under real-time requirements, with practical multi-faceted nature, gives interdisciplinary exchangeability, encourage consistent taking care of signals and numerous other aspects set elevated standards for testing systems. TPT is one test approach that encourages the plan, execution, evaluation and report generation of experiments for automotive model-based frameworks. TPT experiments are compact and reusable on various test stages, for example, MIL, SIL or HIL. TPT reinforces responsive tests and can be executed progressively.

Shruthi and NazMufeeda [8] presented the use of CANoe tool for testing of communication between brake ECU and electronic stability program. Manual testing of various ECUs in a vehicle becomes a tedious job; hence, Vector CANoe tool is used for automated testing of ECU communication. Since this testing is done prior to the component production, rate of failure and manufacturing cost are narrowed. By using Vector CANoe setup, ECUs can be tested effortlessly and productively. CANoe tool gives precise estimation and test outcomes.

Stahl, Konrad, et al. [9] presents HIL testing approach for research in dynamic vehicle security frameworks. The framework comprises a PC simulation of the vehicle and the vehicle environmental factors. A steering system with a real-time

ECU and sensor equipment is hardwired in the close control loop. Steer controlling by both driver and actuator is taken into the simulation environment, and then stability, performance and safety of vehicle are tested and validated.

Nibert, Jonathan, et al. [10] presented multi-stage model development approach for vehicle testing. In this approach, a model is passed through the diverse advancement phases of MIL, SIL and HIL without keeping up various variants of the full model. Rather, just lightweight variants containing only the interface layers should be separated. Hence, keeping up variants is made simpler since just a little subsection of the model has to be changed between variants. Also, it guarantees probity of the control logic since the practical layer can be decoupled from the interface layer, and hence, it is equivalent in all variations of the model, regardless of the equipment conditions identified with interfacing.

Ciceo, Sebastian, et al. [11] proposed model-based structure and testing technique concentrating on the electric vehicle driveability perspective. The procedure is classified into two steps. The initial step is the MIL co-simulation along with electric drive which is modeled in MATLAB/Simulink. The second step is HIL simulation in which physical and plant model electric drive is integrated in real time. The co-simulation approach is favored because of the upsides of joining control designing given by MATLAB/Simulink and framework building given by LMS Imagine Lab Amesim in development stage, and the simplicity of coordinating both the control and framework models in the HIL testbench for verification reason.

Closed-Loop Test Platform models all the physical plant behavior in addition to the I/O processing; hence, hardware-in-loop concept is preferred as it allows tests of controllers creating a real vehicle behavior within an indoor environment. Hardware-in-loop simulation allows to check how controller responds in real time.

3 Methodology

Figure 2 shows various components used in HIL simulation.

The following is the description of various steps involved in HIL simulation process:

- (1) System analysis and understanding:
Understanding the electric drive system specifications and writing test specifications.
- (2) Test implementation and test script generation:
 - Test cases are written based on system specification understanding in Excel sheet. It includes the testing steps and states of signals that are expected in outputs.
 - From Excel sheet, test scripts are generated using test automation scripting (VBScript), such that they are understandable by the test execution software.
- (3) Creation of Diagnostic Binary Caesar (DBC) files:

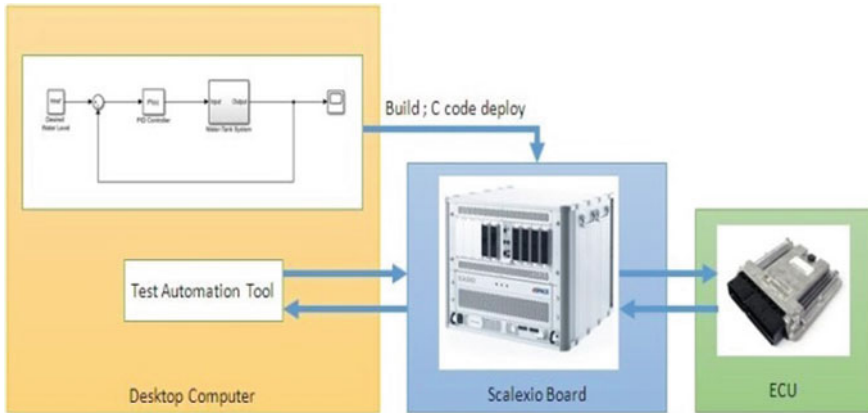


Fig. 2 HIL simulation components

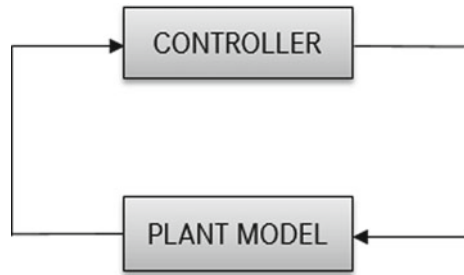
A database is used to maintain the information related to CAN bus communication. CANdb++ is a data organization tool that is utilized to create and modify CAN communication databases for various systems of electric drive. CAN database involves objects such as networks, control units, network nodes, signals and messages. DBC is created consisting of different nodes. Signals sent and received by the nodes are listed. Each signal is given name, definition, length, order, value type and min/max values. The signals are mapped into relevant messages. Messages include message name, message ID, type, transmitters and receivers. Then, the messages are mapped for different network nodes. For each network node, there is Tx message group and Rx message group. The messages are segregated and mapped accordingly. The network consists of different ECUs involved in the particular CAN system.

(4) Plant Modeling:

Simulink is incorporated with MATLAB as a collective setup for plant modeling, analysis and simulation. Models are created based on block diagrams through drag and drop.

Figure 3 shows a closed-loop system block diagram.

- Controller is the ECU that is to be tested.
- A real-time environment for testing the ECU is obtained by designing the plant model based on system requirements.
- Rest bus simulation is done when real ECU is not present.
- When real ECUs are present, mapping of appropriate CAN signals is done using RTICANMM blocks.
- When part of model is present, its logic is analyzed using requirement definition and system specification. Data type matching and value limit assignment are done. A part of plant model logic like state flow for key position, gear selection, etc., is developed.

Fig. 3 Closed-loop system**(5) HIL Setup:**

- Understanding of the input/output (I/O) pins and CAN communication channels for ECU, referring wiring harness schematics which gives pin configuration of ECU.
- Allocating appropriate dSPACE cards for the ECU pins referring the SCALEXIO documentation and dSPACE tools pin details.
- Setting up breakout board with phoenix connectors to perform wiring for integration of ECU side to HIL side. Breakout board simplifies the wiring between control hardware and HIL system. It provides easy access to all interface signals like control feedback signals, analog/digital I/O signals, etc.
- ECU–CAN communication verification is done by connecting CAN-H, CAN-L, power supply, Gnd lines of ECU and SCALEXIO and check if signals from ECU are transmitted to node created in SCALEXIO side.

(6) Integration of logical plant model and I/O model using ConfigurationDesk: ConfigurationDesk is a graphical design and execution tool suitable for controlling huge HIL real-time applications based on dSPACE real-time components such as SCALEXIO and to implement plant models and input–output code on real-time hardware. ConfigurationDesk gives a distinctly organized outline of external components like ECUs, dSPACE channels and simulated models.

(7) Automation testing:

- VB test scripts are imported into test automation tool. These scripts contain various vehicle behaviors to be tested. Also, the SDF file obtained from the build process is put in Provetech.
- Vehicle behavior to be tested is observed in the workpage available in the tool. This shows up the CAN signal and the model signal values whose values change according to the test conditions.
- Manipulation blocks are also available to modify the model signal values according to the requirement.
- Faults can be induced manually via diagnostics to check if the behavior is as expected.

- Test cases are then checked for pass and failure conditions. Failed test cases have to be debugged for syntax and logical errors to get a successful testing environment.
- Test reports are generated for all the executed scripts. These reports will be in a prescribed format, i.e., MDF or BLF format.
- These files are then imported on the CANoe tool for vehicle log analysis. Vehicle log can be analyzed online or offline. Online analysis is done in parallel when scripts are running, whereas offline log analysis is done once the reports are generated.

(8) Test Report and Vehicle log analysis:

CANoe is a vigorous tool for ECU and bus simulation, vehicle log analysis and test automation.

Steps in CANoe analysis:

1. Select CAN configuration in vector CANoe.
2. In simulation window, required DBC files are selected.

Figure 4 shows the Vector CANoe simulation setup window, in which the required CAN DBC files are selected.

3. Online simulation is done during execution based on test script commands.
4. For offline simulation, in the measurement setup, offline mode is selected. BLF files are loaded.

Figure 5 shows the Vector CANoe measurement setup window, in which offline or online mode for analysis is selected.

5. In graphic window, vehicle log analysis is done by adding appropriate signals for evaluation, and it is run, and then, analysis is done to check if the vehicle behavior is as per the requirements.

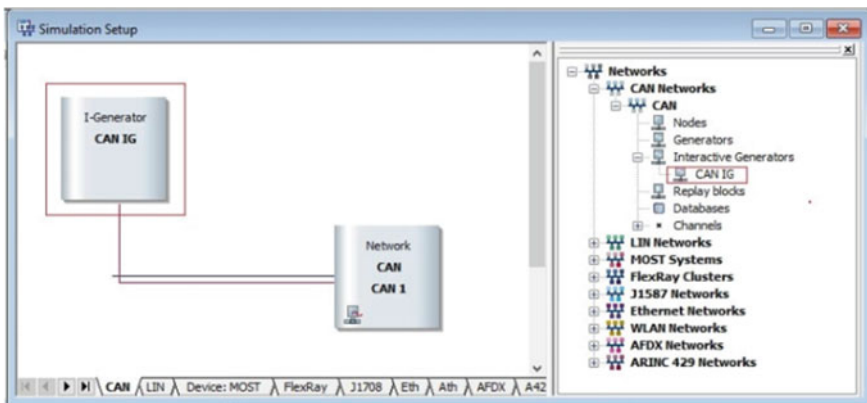


Fig. 4 Simulation setup window in Vector CANoe

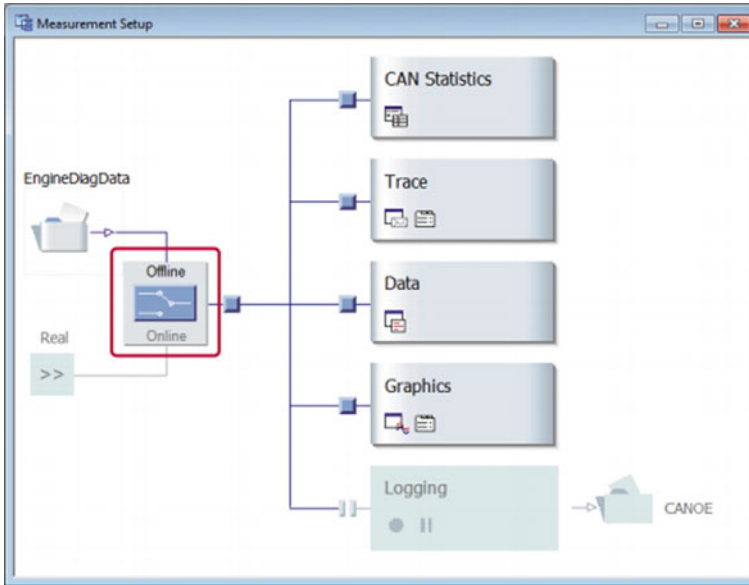


Fig. 5 Measurement setup window in Vector CANoe

4 Results and Discussions

Based on system analysis and understanding, test cases are written, and test scripts are generated to test the behavior of the system. Some examples for test cases are high voltage activated when vehicle is cranked, high voltage activated when battery charger is connected, high voltage deactivated when vehicle ignition switch is turned off, and depending on torque demand of powertrain controller, motor rotates, etc.

CAN DBC files are created that consist of the CAN messages and signals relevant to a particular system as seen in Fig. 6.

The CAN DBC file forms a database that contains the signals and messages that are relevant to a particular system or ECU. Each signal is given name, definition, length, order, value type and min/max values. The CAN messages include message names, message ID, type, transmitters and receivers.

After ECU-CAN communication setup, setting RTICANMM block and loading of CAN DBC files, it is checked if the ECU transmits the messages. The CAN messages received at the receiver node are displayed as seen in Fig. 7.

The information like message bit size, description of the message and sender ECU of the message can be obtained from Fig. 7.

SDF files are generated after successful build process in ConfigurationDesk by adding Simulink model and hardware functional blocks. Then, the test scripts along with SDF files are imported to Provetech and executed. After debugging and clearing of errors and warnings in the test scripts, a pop-up message 'Syntax error is correct' is displayed as seen in Fig. 8.

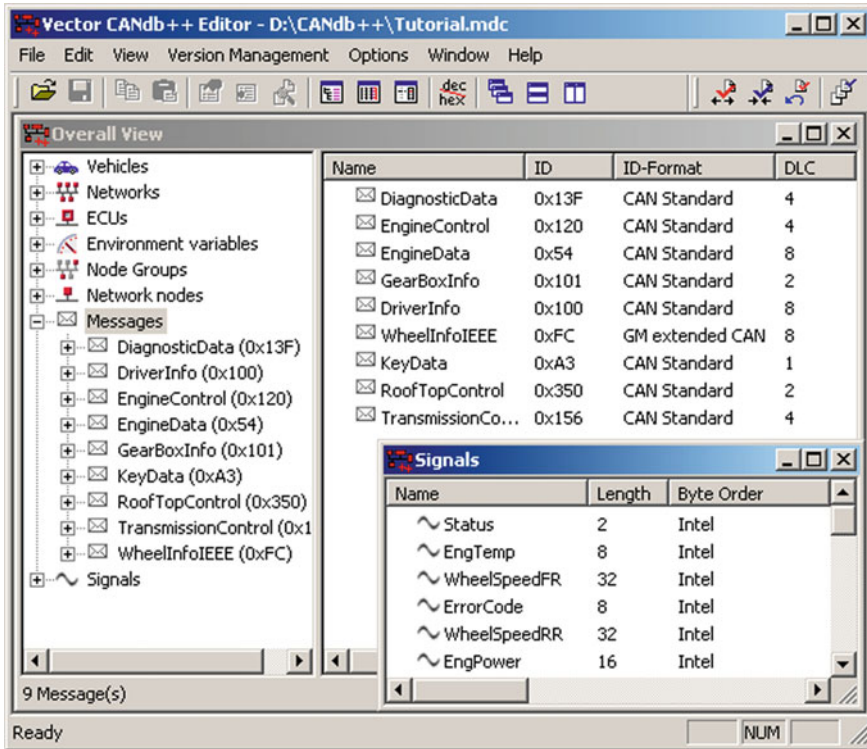


Fig. 6 CAN signals and messages window

After checking syntax errors, test reports are generated as shown in Fig. 9. These reports describe if any other errors are present, time taken for execution and the test results (pass or fail status).

Signal waveforms with resultant values obtained in graphic window of Vector CANoe are shown in Fig. 10. These signals are analyzed. A measurement cursor is used to read the values of the signals over a time duration. The obtained signal values are compared with the expected signal values, and the results are validated. Thus, it is checked if behavior of signals meets system requirements.

5 Conclusion

This paper highlights the significance of testing, verification and validation of electric vehicle to deliver its high performance and safety.

System specifications containing the functionalities of the electric drive system to be validated are analyzed. Based on its understanding, test implementation and test cases are written. Then, test scripts are generated using test automation scripting.

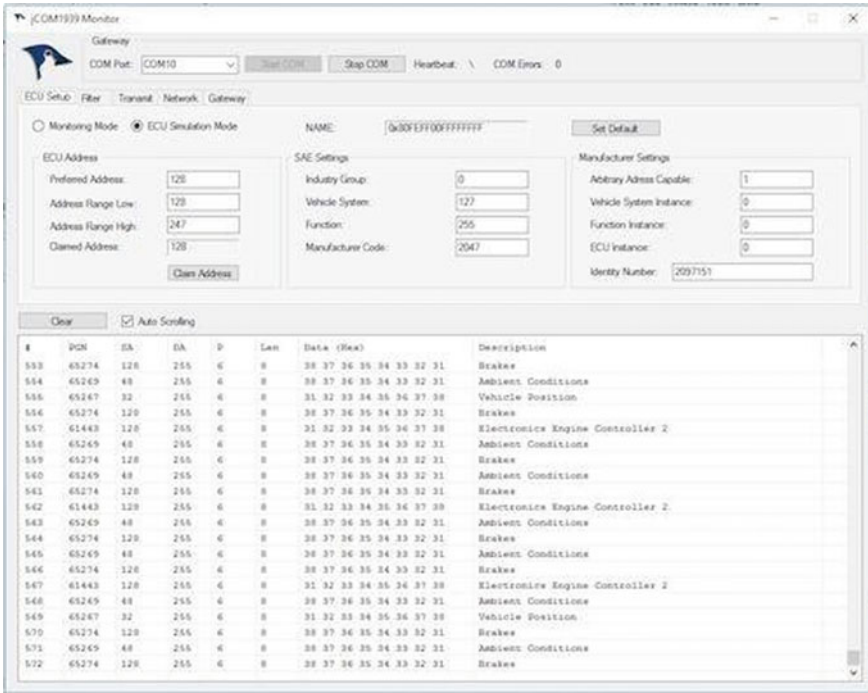


Fig. 7 Received CAN messages display

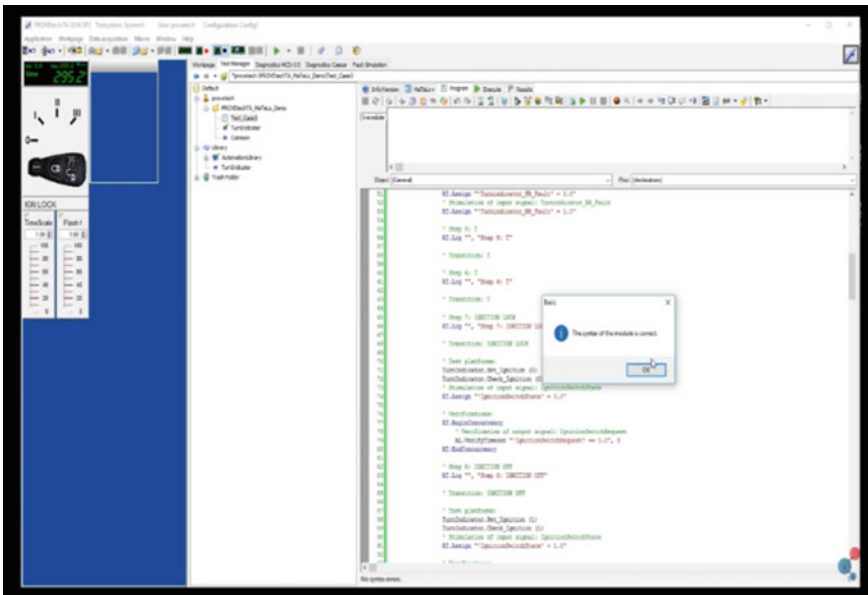


Fig. 8 Syntax check display

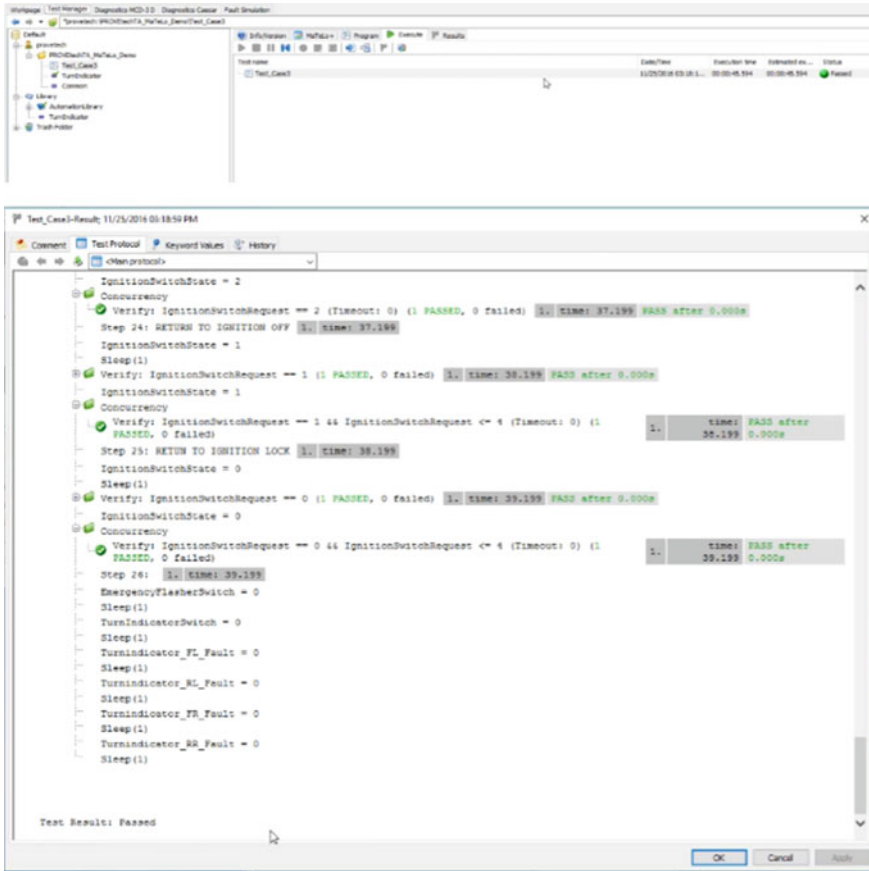


Fig. 9 Test report pass/fail status

CAN DBC files are created for different systems in Vector CANdb++. Simulink is used for plant modeling, and correctness of model is verified. A closed-loop model is used. In the HIL commissioning side, classification of dSPACE boards and selection of proper channels to meet electrical requirements are done. Then, the SCALEXIO and ECU sides are integrated through breakout box. Basic testing of ECU-CAN communication is done in order to verify if the ECU is sending messages. These messages are interpreted using corresponding CAN DBC files. Further logical plant model and I/O model are integrated using ConfigurationDesk. ConfigurationDesk gives a distinctly organized outline of external components like ECUs, dSPACE channels and simulated models. Once the signal chain is developed, checked and resolving of conflicts is done, it is given for build process. Using the SDF files generated from build process, the test scripts are executed and test reports are generated with pass or fail status in the test automation tool. Vector CANoe tool is used

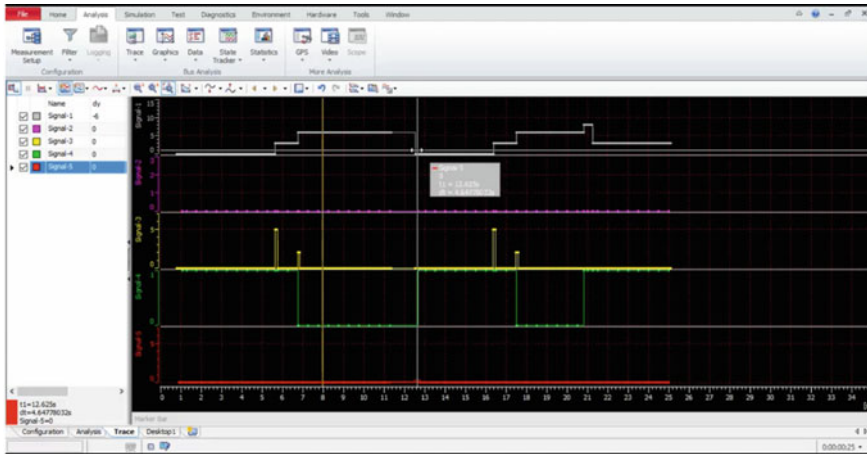


Fig. 10 Signal waveforms in graphics window

to analyze the logs obtained after the test script execution. It is validated that the obtained results meet the system requirements.

References

1. V. Kharchenko, et al., V-models of safety critical system life cycle: Classification and application, in *2018 IEEE 9th International Conference on Dependable Systems, Services and Technologies (DESSERT)* (IEEE, 2018)
2. V. Ili, S. Popi, M. Kovači, Data flow in automated testing of the complex automotive electronic control units, in *2016 Zooming Innovation in Consumer Electronics International Conference (ZINC)* (IEEE, 2016)
3. D. Ramaswamy, et al., A case study in hardware-in-the-loop testing: development of an ECU for a hybrid electric vehicle. No. 2004-01-0303. SAE Technical Paper (2004)
4. B. Li, et al., Study on HIL system of electric vehicle controller based on ni. *IOP Conf. Ser. Mater. Sci. Eng.* **382** (2018)
5. T. Vo-Duy, M.C. Ta, A signal hardware-in-the-loop model for electric vehicles. *Robomech. J.* **3**(1), 1–11 (2016)
6. H. Haupt, M. Plöger, J. Bracker, Hardware-in-the-Loop test of battery management systems. *IFAC Proc. Vol.* **46**(21), 658–664 (2013)
7. E. Bringmann, A. Krämer, Model-based testing of automotive systems, in *2008 1st international conference on software testing, verification, and validation* (IEEE, 2008)
8. T.S. Shruthi, K.H. NazMufeeda, Using VT system for automated testing of ECU. *Int. Organ. Sci. Res. J. Comput. Eng.* **18**(3) (2016)
9. K. Stahl, K.-D. Leimbach, Vehicle dynamics simulation by using hardware in the loop techniques, in *17th International IEEE Conference on Intelligent Transportation Systems (ITSC)* (IEEE, 2014)

10. J. Nibert, M.E. Herniter, Z. Chambers, Model-based system design for MIL, SIL, and HIL. *World Electr. Veh. J.* **5**(4), 1121–1130 (2012)
11. S. Ciceo, et al., Model-based design and testing for electric vehicle driveability analysis, in *2016 IEEE 16th International Conference on Environment and Electrical Engineering (EEEIC)* (IEEE, 2016)

Dimensionality Reduction Using Principal Component Analysis for Lecture Attendance Management System



Ramaprasad Poojary, Mariyam Milofa, and K. Shruthi

1 Introduction

Security and authentication play vital role in a society. Face recognition is one of the well-known techniques for security and authentication [1]. In a university, student attendance record is important to control student discipline. There are many ways of attendance recording mechanisms which are being used in practice such as manual system, radio-frequency identification (RFID)-based method and biometric technologies such as iris recognition, fingerprint recognition, face recognition and so on. Each of these technologies has advantages and disadvantages. Face recognition-based attendance management system is also one of the most advanced methods of biometric technology. Success rate of this technology is high. Because of biometric technology becoming simpler to use, many institutions started using this method for attendance marking of students [2, 7]. Extensive study has been done in the literature on human faces compared to other objects. Reason behind this study is due to the remarkable face recognition capability of the human visual system and also due to number of important applications of face recognition technology [3–5]. Principal component analysis (PCA) is a dimensionality reduction technique which has been used prominently in industries and universities for person face identification.

Principal component analysis (PCA) has been developed to make the computation simpler. Dimensional reduction of linear projection that maximizes the scatter of all

R. Poojary (✉) · M. Milofa
School of Engineering and IT, Manipal Academy of Higher Education, DIAC, Dubai Campus,
Dubai, UAE

e-mail: rama.poojary@manipaldubai.com

M. Milofa
e-mail: mmilofa11@gmail.com

K. Shruthi
Indian Institute of Technology, Dharwad, India
e-mail: shruthikarunakar84@gmail.com

projected samples is chosen by PCA to represent a high-dimensional training set into a lower dimensional subspace [6]. Work presented by Paul Viola and Michael Jones with a title “Fast and robust method for face recognition” was found to be 15 times faster than any method during the time of release and said to give 95% accuracy at about 17 fps [8, 9].

Organization of this paper: Sect. 2 presents the proposed lecture attendance management system. Section 3 validates the proposed work by using statistical parameters such as reduction rate and recognition accuracy. Results are verified by considering different datasets. Section 4 concludes the work.

2 Related Work

Shubhobrata et al. [10] proposed a study on automatic attendance management system based on face recognition developed using ubiquitous components. Zainab Hussain Arif et al. [11] presented a detailed critical review on technology, application domain and main findings used by different researchers in the field of automated attendance management system. Shreyak et al. [12] proposed a study on student attendance management system based on face recognition. This work uses eigenface values, principal component analysis (PCA) and convolutional neural network (CNN) for its implementation. Kumar Neela Ashish et al. [13] presented a smart attendance marking system based on face recognition. In this study, Viola–Jones algorithm is used for face detection, and PCA is used for feature selection.

It is clear from the literature review that most of the attendance management system is based on face recognition method and demands large memory when the number of face values increases. Hence, the proposed study focuses on reducing number of eigenvectors without affecting recognition accuracy.

3 Lecture Attendance Management System

The purpose of this system is to update the student attendance register automatically using face recognition technique [14]. This system is implemented in two phases as mentioned in Sect. 1 of the paper. In the first phase, student registration and two face datasets are created. In phase 2, a test face is detected and recognized, and attendance register is updated.

3.1 Student Registration and Generation of Training Face Database

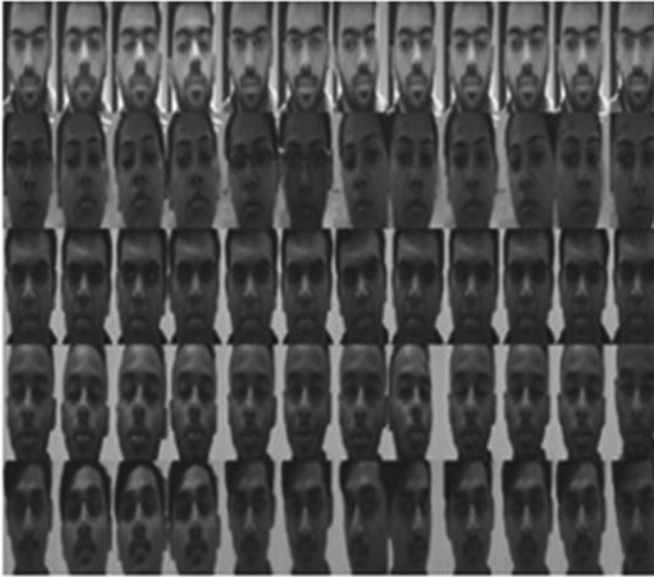
- Student Registration: Student registration is a step carried out for the first time when a new student enrolls for a course. During student registration, student name and university seat number are recorded in the attendance register as shown in Fig. 1.
- Image Capture: After successful registration, each student will be given proper instructions so as to obtain face images with different facial expressions. A total of 15 images is captured using a web camera.
- Face Detection: Detection of face part from a given image is a crucial and essential preprocessing task in face recognition. From preliminary tests, it was concluded that this step increases efficiency by a large margin. Viola–Jones algorithm is a well-known face detector technique. Viola–Jones algorithm makes use of cascade classifier for face detection. Face parts of all the 15 images of a student are obtained by using Viola–Jones algorithm.

Viola–Jones Algorithm: It uses Haar-like features that is scalar product between an image and Haar-like templates [9].

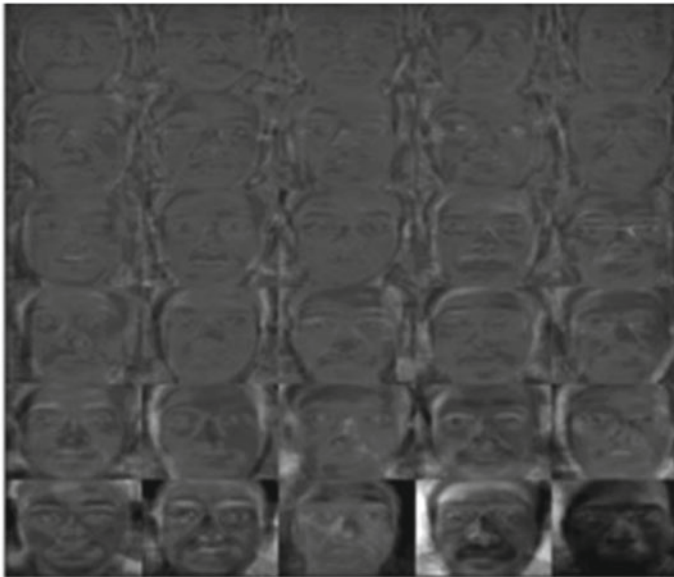
- Face Database: After face detection, images in the training dataset are resized to a lower dimension of 160×120 and stored as gray-level images. Sample face images from dataset 1 are shown in Fig. 1, and sample eigenfaces for the same are as shown in Fig. 2.

Fig. 1 Screenshot showing student registration





(a)



(b)

Fig. 2 Sample face images and eigenfaces are shown in **a** and **b**, respectively (Written informed consent was obtained from the students for the publication of images included in the manuscript)

	SUN	MON	TUE	WED	THURS	FRI	SAT	SUN	MON	TUE	WED	THURS	FRI
1	01-Jan-17	02-Jan-17	03-Jan-17	04-Jan-17	05-Jan-17	06-Jan-17	07-Jan-17	08-Jan-17	09-Jan-17	10-Jan-17	11-Jan-17	12-Jan-17	13-Jan-17
2	ID	NAME											
3	1001	Ram											
4	1002	Aadya											
5	1003	Shubha											
6	1004	Tesla											
7	1005	Sherlock											

Fig. 3 Sample attendance register showing student attendance data

3.2 Face Recognition and Attendance Update

- Test Face Detection: Once an image is captured, the test face is extracted using Viola–Jones algorithm, resized to 160 × 120 resolution and stored as gray-level image [10].
- Face Recognition using PCA: PCA-based face recognition is well discussed in most of the literatures [15–18]. Principal component analysis (PCA) is one of the widely used feature extraction methods which generates feature vectors that are a linear combination of initial features [19, 20]. Advantage of PCA also lies in the reduction rate which is discussed in Keerthi Vasan et al. (2016)
- Attendance Update: Once the face recognition is complete, attendance register is updated as per the result. Sample of attendance register used for face recognition-based attendance system is as shown in Fig. 3.

4 Results and Discussions

The experiments were carried out on databases captured during the registration phase.

4.1 Test Using Student Database 1

A total 750 facial images of students were used for creating student database 1. Facial images were captured with 15 different facial variations for 50 students. It has been

Table 1 Recognition accuracy with PCA + ED when student database 1 is used

No. of principal components (K)	Reduction rate (RR) K/D	Recognition accuracy RA (%)
500	0.67	99.2
400	0.53	99.15
300	0.4	99
200	0.27	98.2
100	0.13	93
70	0.09	87
60	0.08	81
50	0.067	70
30	0.04	55
20	0.027	38

Grey shaded region indicates that, even with less than 50% of Principal Components (K), high accuracy of 98.2% is achieved

ensured that lighting conditions are kept normal and constant throughout the test. All the face images are cropped to a dimension of 160×120 and converted into gray scale. In this case, the original dimension of eigenvector is 749. By applying PCA, the number of eigenvectors used for recognition can be reduced, resulting in much faster computation. For determining the recognition accuracy of PCA with Euclidean distance (ED) as the distance classifier, a test was carried out. Results showed over 99% recognition accuracy (RA) when more than 60% of the feature vectors were being used for distance classification. The test also showed that ideal value for K (feature vector of principal components) as highlighted in Table 1 was 200 which gave an accuracy of 98.2%. Thus, it can be seen that even when half the number of eigenvectors is used for computation, the accuracy does not vary as much.

4.2 Reduction Rate (RR) for PCA

RR is a measure compression ratio in the case of PCA. If M is the dimensional space and K is the dimensional subspace, then RR can be calculated using Eq. (1).

$$RR = \frac{K_{ideal}}{M} \quad (1)$$

Table 1 also shows the RR rate of the algorithm which is 0.267 at K ideal.

4.3 Test 2 Using Student Database 1

This test checks the robustness of the algorithm against occlusion. When test images are captured in a classroom environment with N number of students, the chances of getting occluded images are high. Hence, this test becomes a vital part of the proposed work. The student database 1 created as per the previous section has been used for this test (Fig. 4).

Table 2 shows the reduction rate (RR) and recognition accuracy (RA) for different values of K. It has been observed from the figure that RA is highly influenced by occlusion. In the figure, it is seen that there is a drastic decrease in RA when less than 50% of the eigenvectors only are used (highlighted in the figure).

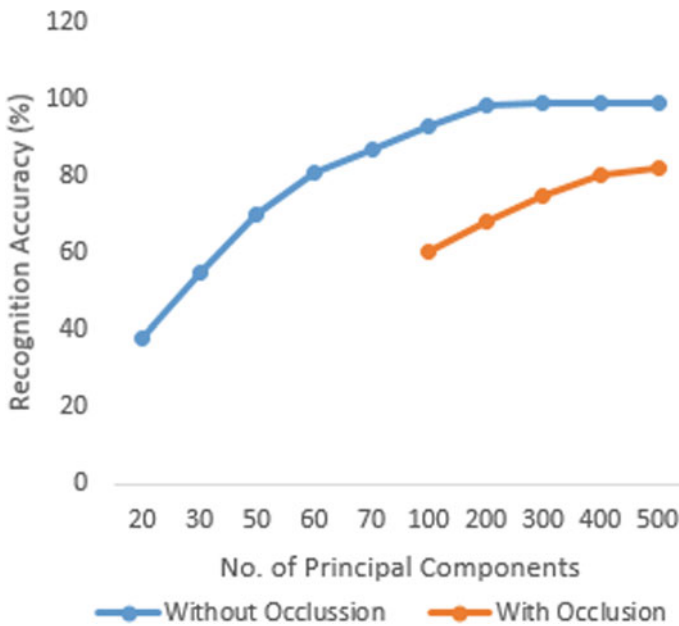


Fig. 4 Graph showing relationship between K and RA with and without occlusion

Table 2 PCA-ED for $D = 749$ with occlusion using student database 1

No. of principal components (K)	Reduction rate (RR) K/D	Recognition accuracy RA (%)
500	0.67	82.1
400	0.53	80
300	0.4	75
200	0.27	68
100	0.13	60

5 Conclusion

Face recognition based on PCA is incorporated in this attendance management system. The impact of the Viola–Jones algorithm in face detection has been highlighted in the paper. From the experimental results, it has been found that recognition accuracy of the proposed algorithm is 98.2% for the database without occlusion and 75% with occlusion for K ideal. Results also show that ideal value for RR for database 1 without occlusion is 0.267, and with occlusion, it is 0.4. Results also show that varying lighting conditions will degrade the performance of the algorithm.

References

1. Using principal component analysis and neural networks. *Int. J. Comput. Netw. Wirel. Mob. Commun. (IJCNWMC)* **3**(1), 245–256 (2013). ISSN 2250-1568 © TJPRC Pvt. Ltd
2. K. Jaikumar, M. Santhosh Kumar, S. Rajkumar, A. Sakthivel, Fingerprint based student attendance system with Sms alert to parents. *Int. J. Res. Eng. Technol. (IJRET)* **4**(2), 293–297 (2015)
3. D.R. Wijaya, I. Asror, Integrated and efficient attendance management system based on RFID. *J. Theor. Appl. Inform. Technol.* **73**(1), 149–154 (2015)
4. Wagh, R. Thakare, J. Chaudhari, Attendance system based on face recognition using eigen face and PCA algorithms, in *2015 International Conference on Green Computing and Internet of Things* (2015)
5. J. Wright, A.Y. Yang, A. Ganesh, S. Shankar Sastry, Y. Ma, Robust face recognition via sparse representation. *IEEE Trans. Pattern Anal. Mach. Intell.* **31**(2), 1–18 (2009)
6. K. Keerthi Vasani, B. Surendiran, Dimensionality reduction using principal component analysis for network intrusion detection. *Persp. Sci.* **8**, 510–512 (2016)
7. G. Lakshmi Priya, M. Pandimadevi, G. Ramu Priya, P. Ramya, Implementation of attendance management system using SMART-FR. *Int. J. Adv. Res. Comput. Commun. Eng.* **3**(11), 8532–8536 (2014)
8. M.K. Dabhi, B.K. Pancholi, Face detection system based on Viola–Jones algorithm. *Int. J. Sci. Res. (IJSR)* **5**(4) (2016)
9. Y.-Q. Wang, An analysis of the Viola–Jones face detection algorithm. *Image Process. Line* **4**, 128–148 (2014)
10. S. Bhattacharya, G.S. Nainala, P. Das, A. Routray, Smart attendance monitoring system (SAMS): a face recognition based attendance system for classroom environment, in *IEEE 18th International Conference on Advanced Learning Technologies (ICALT)*, Mumbai (2018), pp. 358–360
11. Z.H. Arif, N.S. Ali, N.A. Zakaria, M.N. Al-Mhiqani, Attendance management system for educational sector: critical review. *Int. J. Comput. Sci. Mob. Comput.* **7**(8), 60–66 (2018)
12. S. Sawhney, K. Kacker, S. Jain, S.N. Singh, R. Garg, Real-time smart attendance system using face recognition techniques, in *9th International Conference on Cloud Computing, Data Science & Engineering (Confluence)*, Noida (2019), pp. 522–525
13. K.N. Ashish, P. Swarnalatha, C. Prudhvi, N.J. Krishna, K.K. Sandeep, Smart attendance marking system using facial recognition. *Res. J. Sci. Technol.* **11**(2), 101–108 (2019)
14. E. Rekha, P. Ramaprasad, An efficient automated management system based on eigen face recognition, in *International Conference on Cloud Computing, Data Science & Engineering*, 12th–13th January 2017, Amity University Uttar Pradesh, Noida, India
15. M.J.R. Mulla, R.P. Patil, S.K. Shah, Facial image based security system using PCA, in *2015 IEEE Access, International Conference on Information Processing (ICIP)*

16. P. Mehta, P. Tomar, An efficient attendance management system based on face recognition using matlab and raspberry pi 2. *Int. J. Eng. Technol. Sci. Res.* **3**(5) (2016)
17. P. Singh, A. Sharma, Face recognition using principal component analysis in MATLAB. *Int. J. Sci. Res. Comput. Sci. Eng.* **3**(1) (2015)
18. S. Suhas, B. Kurhe Ajay, P. Khanle, Face recognition using principal component analysis and linear discriminant anlysis on holistic approach in facial images database. *IOSR J. Eng.* **2**(12), 15–23 (2012)
19. S.M. Bah, F. Ming, An improved face recognition algorithm and its application in attendance management system. *Elsevier Array* **5**, 100014 (2020)
20. O. Godswill1, O. Osas, O. Anderson, I. Oseikhuemen, O. Etse, Automated student attendance management system using face recognition. *Int. J. Edu. Res. Inform. Sci.* **5**, 31–37

Automatic Modulation Recognition Using Machine Learning Techniques: A Review



N. Venkateswara Rao and B. T. Krishna

1 Introduction

AMC is the method of finding the modulation format of an intercepted signal with no prior information about its features. The main advantage of AMC is it does not require any supplement information in communication to inform the receiver about the changes in transmitted signal modulation type [47]. The concept of AMC is originated from military applications. In this scenario, authenticated users should be given access to the spectrum, whereas unauthorized users should be barred from access. In this real-time environment, a receiver is required that is capable (intelligent) of discriminating authorized users from unauthorized users [1]. In [2], spectrum sensing is applied for civilian purpose. The terms SDR and CR were enhanced by Mitola in 1991 and 1998, respectively [3]. SDR has the capability of handling a greater number of modulation classes at a time. Due to recent advancements in SDR, a lot of importance was given to AMC. To utilize the spectrum efficiently, the idea of CR was introduced [4]. CR is a software radio which helps the user to get access to the spectrum by situational presence and knowledge [5]. CR's will sense the spectrum extending the underutilized, licensed frequency bands to unlicensed (secondary) users [6]. Spectrum sensing (SS) is also a major step to implement next-generation adaptive radio system [7]. SS can be performed in centralized or distributed approaches [8]. In centralized approach, a master entity is required to gather all information from subsequent CR nodes through the common channel. In distributed approach, CR nodes will exchange information through local channels.

Among SS methods, cooperative sensing is more accurate than non-cooperative schemes [9]. In cooperative SS, a centralized SS system will search for the vacant

N. V. Rao (✉) · B. T. Krishna
Electronics and Communication Engineering, JNTU Kakinada, Andhra Pradesh, India
e-mail: venkatesh0621@gmail.com

B. T. Krishna
e-mail: tkbattula@gmail.com

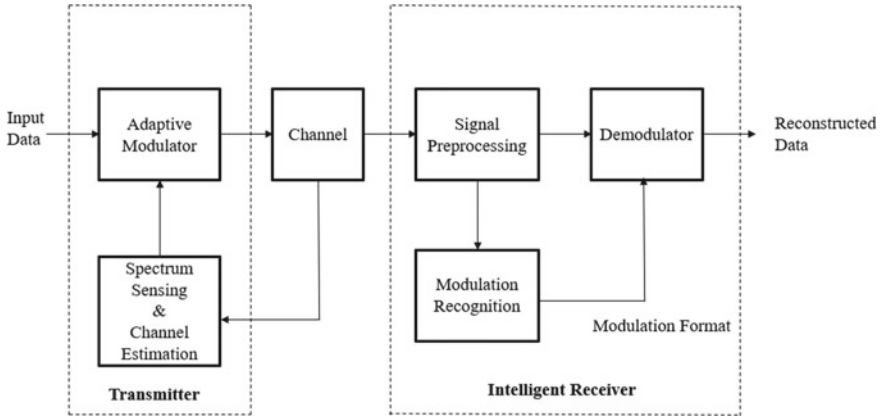


Fig. 1 Block diagram of adaptive modulation system [11]

bands in the spectrum. It allows secondary users (SU) to access licensed spectrum. For SU's, the channel is not fixed, they will send the data, which is available at that instant of time. Because of this, SU transmitters operate according to channel conditions.

The adaptive modulation system block diagram is shown in Fig. 1. Modulation recognition is a process applied between signal detection and signal demodulation [10].

The rest of the paper is organized as follows. Section 2 discusses likelihood-based (LB) approach. Machine learning (ML) approach is discussed in Sect. 3. Section 4 concludes the article.

2 Likelihood-Based [LB] Approaches

In LB approach, classification is done by comparing the likelihood ratio over a pre-determined threshold value [12]. The LB approach is applied for AMC which is a hypothesis problem. Probability density function (PDF) of the observed signal waveforms will contain all information related to classification [13]. Based on the model chosen, LB-AMC approaches are categorized as generalized likelihood ratio test (GLRT) [14], average likelihood ratio test (ALRT) [15–17], quasi ALRT [18], hybrid likelihood ratio test (HLRT) [20] and quasi HLRT [19].

ALRT represents unspecified parameters as random variables and computes the likelihood function by averaging them. By applying maximum likelihood estimations on unknown parameters, GLRT calculates PDF. In [14–20], LB approach is used for automatic modulation recognition (AMR). Quasi ALRT recognizes frequency shift keying (FSK) and phase shift keying (PSK) signals with good accuracy, but they fail in recognizing modulated quadrature amplitude modulation (QAM) signals. Even

though GLRT offers many advantages, it degrades its performance when the input is nested constellations. HLRT has high complexity due to more unknown parameters. Complexity in quasi HLRT is low compared to HLRT.

LB approach requires perfect knowledge of the channel model and their parameters for classification. Additionally, LB techniques have faced difficulty in determining the appropriate solutions, especially when unknown parameters are more. This approach is computationally complex.

3 ML Approaches

The feature extraction system approach comprises two phases feature extraction and classifier. Various researchers used spectral features, cyclostationary features, statistical features, constellation features and transform features in their work.

Mostly, all modulated signals are represented as in-phase (I) and quadrature (Q) components. Higher-order statistical (HOS) features such as moments, ratios and cumulants are the statistical features. These features eliminate the background noise occurred in channel [12]. These features will be helpful in discriminating amplitude shift keying (ASK), PSK and QAM. In [21–23], different orders of cyclic cumulants and statistical moments are used.

In [10], spectral features are used for signal classification. Amplitude, phase and frequency are used as spectral parameters. In [24], transform features are used in their work. Constellation features will be helpful while classifying PSK and QAM signals. In [25], constellation feature is used for recognizing M-QAM.

In AMC, after feature extraction, the next step is to identify the type of modulation format. Several approaches, such as K-nearest neighbor (KNN), decision tree (DT), support vector machine (SVM), ensemble classifiers, artificial neural networks (ANNs) and combination of genetic algorithm (GA) are used for classification.

3.1 DT

DT uses a tree-like model for decision making [12]. Decisions like propagating from one stage to another rely on threshold values, which accounts for overall performance. A multi-class problem can also be handled by DT. DT takes a considerable amount of time during design [26]. Based on number of trees, DTs are categorized into medium tree (MT), coarse tree (CT) and fine tree (FT) [11].

3.2 ANN

ANN is commonly used as classifier in AMC. ANNs find importance in pattern recognition and ML [27]. It comprises of three main layers: (1) input layer, (2) hidden layer and (3) output layer. Based on the outcome of weights, the nodes will propagate from one layer to another. Weights play a key role in mapping [28].

3.3 SVM

It is a supervised ML approach used for AMC. SVM is preferred over ANN due to training time and overfitting [29]. In [30], kernel is used for feature selection. Radial basis function network (RBFN) and multi-layer perceptron are used as kernels.

3.4 KNN

It is the easier of all ML algorithms. In KNN, classification is done by majority vote of its K-nearest associate members. If the object is having similarities with its associated members, then that object is assigned to that group [2]. Depending on k value and distance, KNN is classified into medium KNN, fine KNN, cosine KNN, coarse KNN, weighted KNN and cubic KNN [46].

These approaches can be applied in computer vision, prediction and semantic analysis [31]. In [32], SVM is applied for breast cancer detection and diagnosis. In [33], SVM algorithm is applied for underwriting process. In [34], KNN algorithm is applied for object detection.

Generalized block diagram of modulation classifier is shown in Fig. 2. It is divided into two phases: (1) training and (2) testing. During training phase, different types of modulated signals are given as inputs. For these inputs, noises like white Gaussian noise, correlated non-Gaussian noise and additive white Gaussian noise (AWGN) are added. During testing phase, unknown modulated signals are tested for classification.

Table 1 describes the approaches used by different researchers in selecting features, classifier and modulation set for AMC using ML.

All the machine learning classifiers will perform well when data set is small. However, their performance is degraded when large amount of data is given [45].

Some of the classifiers will fail under low SNR values. ANN tends to struggle with the computational complexity while handling image data. To overcome these limitations, deep learning came into existence.

In [11], authors used cumulants as the feature and selected 11 features for training the classifier. They used.

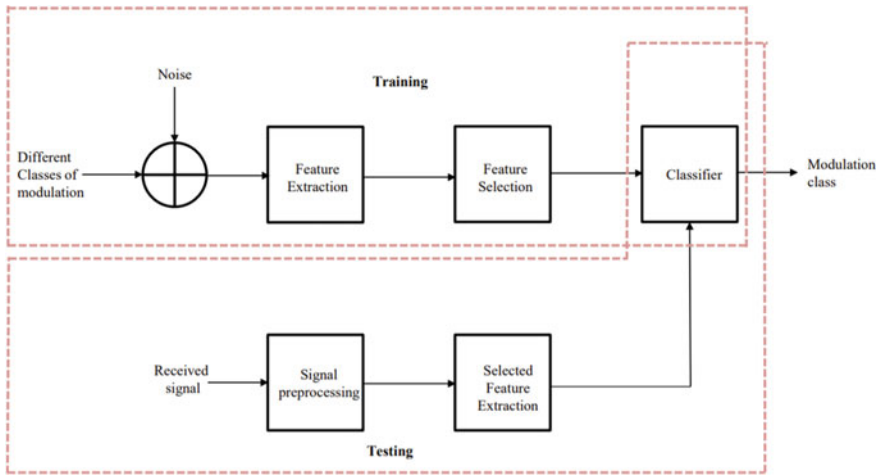


Fig. 2 Generalized block diagram of modulation classifier [11]

Some of the classifiers will fail under low SNR values. ANN tends to struggle with the computational complexity while handling image data. To overcome these limitations, deep learning came into existence.

In [11], authors used cumulants as the feature and selected 11 features for training the classifier. They used different classifiers like DT (FT, MT and CT), KNN (fine, medium, coarse, cubic, cosine, weighted), SVM (linear, quadratic and cubic) and ensemble classifiers. They have chosen QPSK, BPSK, 4 QAM, 8 PSK, 64 QAM, 16 QAM as modulated signals. Several multipath fading conditions are considered at different SNR values ranging from 0 to 20 dB. They found out that ensemble classifiers and SVM achieved higher accuracy under fading and AWGN conditions. Quadratic SVM got classification accuracy (%) of 97.2 even at low SNR (5 dB).

In [35], authors proposed Gabor technique to select features. In the next stage, cuckoo search algorithm (CSA) is used to optimize the features. For decision making, they used extreme learning machine (ELM). They have chosen nine modulation classes (QPSK, 16 PSK, 64 PSK, BFSK, QFSK, 16 FSK, QAM, 16 QAM, 64 QAM). AWGN and Rayleigh fading channels are considered with SNR values ranging from -10 to 5 dB. Their proposed algorithm has given accuracy of 95 at -5 dB.

In [12], author used moments as a feature. Medium KNN, fine KNN, cosine KNN, coarse KNN, weighted KNN and Cubic KNN are used as classifiers. 2 PSK, 4 PSK, 8 PSK, 64 QAM 4 QAM, 16 QAM signals are considered as modulation class for classification. Performance is carried under non-ideal channel conditions with SNR ranging from 0 to 20 dB. They found that out of all KNNs, medium KNN and cubic KNN classifiers performed well. Classifiers got 98.3% accuracy at 10 dB and approached 100% at 20 dB.

In [36], moments are considered as features. They used a hybrid model (intersection of LB and feature-based classifiers). They carried their investigations by

Table 1 Summary of ML classifiers

Ref. No.	Features	Modulation class	Classifier	Category	Channel conditions	SNR	Accuracy	Year
[38]	Cumulants moments	Estimators (CSI, time, frequency and phase offset)	SVM	ML	AWGN channel SNR (-10 to 15 dB)	0 dB 15 dB	90 100	2019
[39, 46]	Multiple cumulants	2PSK, 4 PSK, 8 PSK, QAM, 16 QAM, 64 QAM	DT	ML	AWGN channel with SNR (0-20) dB	15 dB 20 dB	98.9 100	2018 2019
[40]	Statistical features	4ASK, 2ASK, 2PSK, 2FSK, FM, MSK, USB, LSB	SVM and DT	ML	AWGN channel SNR (0-20 dB)	4 dB 20 dB	78 100	2018
[41]	I, Q, SNR	BPSK and FSK	SVM, DT, KNN, ensemble boosted trees	ML	AWGN channel SNR (10-30 dB)	12 dB 27 dB	87.5 97.5	2016
[2]	Cumulants	BPSK, QPSK, QAM16 and QAM64	Genetic programming (GP)+KNN	ML	AWGN channel SNR (5-20 dB)	5 dB $N = 4096$ 15 dB $N = 4096$	95 100	2012
[42]	Spectral features, moments, HOC	2 ASK, 4 ASK, 2 FSK, 4 FSK, 2 PSK, 4 PSK, 16 QAM, QAM64	RBFN	ANN	AWGN channel with SNR (3-27 dB)	6 dB 18 dB	85 95	2009
[43]	Cyclic spectrum	OFDM/BPSK, QPSK, ASK5, ASK7, PSK5, PSK7	Independent component analysis (ICA)—SVM	ML	AWGN channel with SNR (0-20 dB)	0 dB 10 dB	80 100	2009
[44, 45]	Statistical features	AM, DSB, LSB, VSB, FM, PSK2, PSK4, ASK2, ASK4, FSK2, FSK4	Decision-theoretic approach and ANN algorithm	ANN	Band-limited Gaussian noise SNR (15-20 dB)	15 dB Decision-theoretic approach 15 dB ANN algorithm	94 96	1998

Table 2 Merits and demerits of different approaches

Approach	Merits	Demerits
LB [23–29]	<ul style="list-style-type: none"> • Optimized performance with known parameters • Easy to understand 	<ul style="list-style-type: none"> • It gives in accurate results when unknown parameters are more • Computationally complex
ML [2, 11, 12, 35–44]	<ul style="list-style-type: none"> • Good classification accuracy when data set is small • Training time will be less 	<ul style="list-style-type: none"> • It gives inaccurate results when data set is large • It cannot take image as features

comparing their proposed algorithm with different approaches. 64 QAM, BPSK, QPSK, 8 PSK and 16 QAM are chosen as modulation classes. These signals are carried in different conditions like AWGN channel, normalized Rayleigh fading channel with different sequence lengths. They concluded that hybrid model outperformed remaining classifiers. Proposed classifier is achieved classification accuracy of 100% at 18 dB.

In [37], spectral features and zero-crossing sequence shape (ZCSS) features are considered to extract features. Spectral features are used to differentiate analog and digital modulations. On the other side, ZCSS features are used to differentiate phase and frequency signals. Adaptive binary neural network approach is used as classifier. ASK, AM, QAM, PSK and FSK are used as modulated signals. They used signal level of $[-55$ to $10]$ dBm for observations and detected error rate of 45% at -55 dBm and 0% error rate at 10 dBm.

Table 2 shows the merits and demerits of different approaches which are discussed above.

In ML approach, different features are selected depending on the modulation technique preferred (digital and analog). ML approach does not give satisfactory results when huge data is given. This paves the path for deep learning in AMC.

4 Conclusion

This article presents an in-depth study of AMR on LB and ML approaches. LB approaches provided better accuracy, but poor handling of more unknown parameters made them to be rarely used. ML approaches used different classifiers (KNN, SVM, DT) shown good results on classification accuracy. ML used different ways to generate the features like cumulants, moments and statistical features. It achieved promising results in different operating conditions with different SNRs making them highly recommended in realistic environment. ML approaches cannot handle big data sets resulting in performance degradation.

Acknowledgements I would like to thank Management, Principal and Head of the Department ECE, Malla Reddy Institute of Technology and Science, Hyderabad for providing research facilities to complete my work.

References

1. P. Ghasemzadeh, S. Banerjee, M. Hempel, H. Sharif, Accuracy analysis of feature-based automatic modulation classification With Blind Modulation Detection, in *International Conference on Computing, Networking and Communications (ICNC)* (2019), pp. 1000–1004
2. M.W. Aslam, Z. Zhu, A.K. Nandi, Automatic modulation classification using combination of genetic programming and KNN. *IEEE Trans. Wirel. Commun.* **11**, 1–9 (2012)
3. Qing Zhao, B.M. Sadler, A survey of dynamic spectrum access. *IEEE Sig. Process. Mag.* **24**(3), 79–89 (2007)
4. J. Mitola, G.Q. Maguire, Cognitive radio: making software radios more personal. *IEEE Pers. Commun.* **6**(4), 13–18 (1999)
5. J. Neel, J. Reed, R.M. Buehrer, L. Dasilva, R. Gilles, A. Mackenzie, Analysis and design of cognitive radio networks and distributed radio resource management algorithms (2006)
6. T. Ikuma, M. Naraghi-Pour, A comparison of three classes of spectrum sensing techniques, in *IEEE GLOBECOM 2008—2008 IEEE Global Telecommunications Conference* (2008), pp. 179–182
7. H. Celebi, I. Guvenc, S. Gezici, H. Arslan, Cognitive-radio systems for spectrum, location, and environmental awareness. *IEEE Antennas Propag. Mag.* **52**(4), 41–61 (2010)
8. T.-W. Wu, Y.-E. Lin, H.-Y. Hsieh, Modeling and comparison of primary user detection techniques in cognitive radio networks, in *IEEE GLOBECOM 2008—2008 IEEE Global Telecommunications Conference* (2008), pp. 1–5
9. C. Lee, W. Wolf, Multiple access-inspired cooperative spectrum sensing for cognitive radio, in *MILCOM 2007—IEEE Military Communications Conference* (2007), pp. 1–6
10. C. Yuanzeng, Z. Hailong, W. Yu, Research on modulation recognition of the communication signal based on statistical model, in *2011 Third International Conference on Measuring Technology and Mechatronics*, vol. 3 (2011), pp. 46–50
11. M. Venkata Subbarao, P. Samundiswary, Performance analysis of modulation recognition in multipath fading channels using pattern recognition classifiers. *Wirel. Pers. Commun.* (2020). <https://doi.org/10.1007/s11277-020-07564-z>
12. L. Jung Hwan, K. Jaekyum, K. Byeoungdo, Y. Dongweon, C. Jun Won, Robust automatic modulation classification technique for fading channels via deep neural network, in *Information Theory, Probability and Statistics* (2017), pp. 1–11
13. O.A. Dobre, A. Abdi, Y. Bar-Ness, W. Su, Survey of automatic modulation classification techniques: classical approaches and new trends. *IET Commun.* **1**(2), 137 (2007)
14. P. Panagiotou, A. Anastopoulos, A. Polydoros, Likelihood ratio tests for modulation classification, in *MILCOM 2000 Proceedings. 21st Century Military Communications. Architectures and Technologies for Information Superiority*, vol. 2 (2000), pp. 670–674
15. J.A. Sills, Maximum-likelihood modulation classification for PSK/QAM, in *MILCOM 1999. IEEE Military Communications. Conference Proceedings*, vol. 1 (1999), pp. 217–220
16. B.F. Beidas, C.L. Weber, Asynchronous classification of MFSK signals using the higher order correlation domain. *IEEE Trans. Commun.* **46**(4), 480–493 (1998)
17. B.F. Beidas, C.L. Weber, Higher-order correlation-based classification of asynchronous MFSK signals, in *Proceedings of MILCOM '96 IEEE Military Communications Conference*, vol. 3 (n.d.), pp. 1003–1009
18. B.F. Beidas, C.L. Weber, Higher-order correlation-based approach to modulation classification of digitally frequency-modulated signals. *IEEE J. Sel. Areas Commun.* **13**(1), 89–101 (1995)
19. A. Abdi, O.A. Dobre, R. Choudhry, Y. Bar-Ness, W. Su, Modulation classification in fading channels using antenna arrays, in *IEEE MILCOM 2004. Military Communications Conference*, vol. 1 (2004), pp. 211–217
20. L. Hong, K.C. Ho, Antenna array likelihood modulation classifier for BPSK and QPSK signals, in *MILCOM 2002. Proceedings*, vol. 1 (2002), pp. 647–651
21. P. Marchand, C. Le Martret, J.-L. Lacoume, Classification of linear modulations by a combination of different orders cyclic cumulants, in *Proceedings of the IEEE Signal Processing Workshop on Higher-Order Statistics* (1997), pp. 47–51

22. O.A. Dobre, Y. Bar-Ness, W. Su, Higher-order cyclic cumulants for high order modulation classification, in *IEEE Military Communications Conference*, vol. 1 (2003), pp. 112–117
23. S.S. Soliman, S.-Z. Hsue, Signal classification using statistical moments. *IEEE Trans. Commun.* **40**(5), 908–916 (1992)
24. M. Venkata Subbarao, P. Samundiswary, An intelligent cognitive radio receiver for future trend wireless applications. *Int. J. Comput. Sci. Inform. Sec.* **14**, 7–16 (2016)
25. C. Zhendong, J. Weining, X. Changbo, L. Min, Modulation recognition based on constellation diagram for M-QAM signals, in *2013 IEEE 11th International Conference on Electronic Measurement & Instruments*, vol. 1 (2013), pp. 70–74
26. P. Cutno, C.-H. Cheng, A software-defined radio based automatic modulation classifier, in *2017 Wireless Telecommunications Symposium (WTS)* (2017), pp. 1–6
27. U. Satija, M.S. Manikandan, B. Ramkumar, Performance study of cyclostationary based digital modulation classification schemes, in *2014 9th International Conference on Industrial and Information Systems (ICIIS)* (2014), pp. 1–5
28. S. Athira, Rohit Mohan, Prabaharan Poornachandran, Soman Kp, Automatic modulation classification using convolutional neural network. *Int. J. Comput. Technol. Appl.* **9**, 7733–7742 (2016)
29. Z. Zhu, A. Nandi, Automatic modulation classification: principles, algorithms and applications (2015)
30. M.H. Valipour, M.M. Homayounpour, M.A. Mehralian, Automatic digital modulation recognition in presence of noise using SVM and PSO, in *6th International Symposium on Telecommunications (IST)* (2012), pp. 378–382
31. P.P. Shinde, S. Shah, A review of machine learning and deep learning applications, in *2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA)* (2018), pp. 1–6
32. D. Bazazeh, R. Shubair, Comparative study of machine learning algorithms for breast cancer detection and diagnosis, in *2016 5th International Conference on Electronic Devices, Systems and Applications (ICEDSA)* (2016), pp. 1–4
33. Y. Tan, G.-J. Zhang, The application of machine learning algorithm in underwriting process, in *2005 International Conference on Machine Learning and Cybernetics* (2005), pp. 3523–3527
34. S. Bozkurt, G. Elibol, S. Gunal, U. Yayan, A comparative study on machine learning algorithms for indoor positioning, in *2015 International Symposium on Innovations in Intelligent Systems and Applications (INISTA)* (2015), pp. 1–8
35. S.I.H. Shah, S. Alam, S.A. Ghauri, A. Hussain, F.A. Ansari, A novel hybrid cuckoo search-extreme learning machine approach for modulation classification. *IEEE Access* **7**, 90525–90537 (2019)
36. M. Abu-Romoh, A. Aboutaleb, Z. Rezki, Automatic modulation classification using moments and likelihood maximization. *IEEE Commun. Lett.* **22**(5), 938–941 (2018)
37. Y. Eteffagh, M.H. Moghaddam, S. Eghbalian, An adaptive neural network approach for automatic modulation recognition, in *2017 51st Annual Conference on Information Sciences and Systems* (2017), pp. 1–5
38. P. Ghasemzadeh, S. Banerjee, M. Hempel, H. Sharif, Accuracy analysis of feature-based automatic modulation classification with blind modulation detection, in *2019 International Conference on Computing, Networking and Communications (ICNC)* (2019), pp. 1000–1004
39. M. Venkata Subbarao, P. Samundiswary, Automatic modulation recognition in cognitive radio receivers using multi-order cumulants and decision trees. *Int. J. Rec. Technol. Eng. (IJRTE)* **7**, 61–69 (2018)
40. X. Zhang, T. Ge, Z. Chen, Automatic modulation recognition of communication signals based on instantaneous statistical characteristics and SVM classifier, in *2018 IEEE Asia-Pacific Conference on Antennas and Propagation (APCAP)* (2018), pp. 344–346
41. I. Valieva, M. Bjorkman, J. Akerberg, M. Ekstrom, I. Voitenko, Multiple machine learning algorithms comparison for modulation type classification for efficient cognitive radio, in *MILCOM 2019—2019 IEEE Military Communications Conference (MILCOM)* (2019), pp. 318–323

42. A. Ebrahimzadeh, Automatic modulation recognition using RBFNN and efficient features in fading channels, in *2009 First International Conference on Networked Digital Technologies* (2009), pp. 485–488
43. D. Boutte, B. Santhanam, A feature weighted hybrid ICA-SVM approach to automatic modulation recognition, in *2009 IEEE 13th Digital Signal Processing Workshop and 5th IEEE Signal Processing Education Workshop* (2009), pp. 399–403
44. A.K. Nandi, E.E. Azzouz, Algorithms for automatic modulation recognition of communication signals. *IEEE Trans. Commun.* **46**(4), 431–436 (1998)
45. S. Wei, S. Zou, F. Liao, W. Lang, W. Wu, Automatic modulation recognition using neural architecture search, in *2019 International Conference on High Performance Big Data and Intelligent Systems (HPBD&IS)* (2019), pp. 151–156
46. M. Venkata Subbarao, P. Samundiswary, K-nearest neighbors based automatic modulation classifier for next generation adaptive radio systems. *Int. J. Sec. Appl. (IJSIA), NADIA* **13**(4), 41–50 (2019)
47. P. Ghasemzadeh, S. Banerjee, M. Hempel, H. Sharif, A new framework for automatic modulation classification using deep belief networks, in *IEEE International Conference on Communications Workshops (ICC Workshops)* (2020)

Performance Analysis of TCP in Presence of Nonresponsive Traffic in Wireless Networks



N. G. Goudru 

1 Introduction

Because of technological advancement, wireless network communication achieved an exponential growth in its use and popularity in recent years. Wireless technology evolved very much faster than any other technologies. This evolution in technology made simple and easy for its users. The data communication between a sender and a receiver uses OSI-layered architecture. Transmission control protocol (TCP) and user datagram protocol (UDP) are important protocols in the transport layer. Many survey reports reveal that TCP carry about 85% of the total network traffic load. Because of the popularity in the use of real-time and multimedia applications over Internet, the role of UDP has been considered very important in the traffic management. TCP provides reliable end-to-end service. The reliability is achieved through acknowledgement from receiver to sender. The process of retransmission introduces delay which is the main reason for not using TCP in real-time and multimedia applications, so these applications use UDP. TCP is responsive to network, whereas UDP is nonresponsive to network traffic. In this paper, we consider UDP to transmit the data at a constant bit rate (CBR) into the network. This results in congestion and loss of data. In this research work, we tried to throw some light by analyzing the performance of TCP, in presence of nonresponsive traffic (UDP).

N. G. Goudru (✉)

Department of Information Science and Engineering, Nitte Meenakshi Institute of Technology (affiliated to Visvesvaraya Technological University), Bengaluru, India

e-mail: goudru.ng@nmit.ac.in

2 Literature Review

The researchers in [1] discuss the ways of improving TCP performance over wireless network by adopting changes in three protocols. In end-to-end protocol, sender knows the wireless link. The data link layer protocol is for providing local reliability. Split connection protocol divides the end-to-end connection into two paths. Sender to base station is connected by wire and base station to receiver is connected by wireless. The researchers tried to reach good performance avoiding the division of end-to-end connection at the base station. They applied selective repeat acknowledgement and explicit loss notification concepts for improving the TCP performance. The authors in [2] enhance TCP Reno to TCP VenO by refining additive increase and multiplicative decrease parts of TCP Reno algorithm. The performance of TCP VenO is much higher than TCP Reno. It is capable detecting more accurately the packet loss due to high bit error and hence increasing the throughput performance considerably. The researchers in [3] propose a new variant of TCP called TCP jersey. TCP-Jersey is performing better in identifying the packet loss reasons like due to congestion or high bit error in the medium. The researchers use two algorithms called algorithm for estimation of available bandwidth and algorithm for congestion warning. They enhance the performance of these two algorithms for faster understanding of the packet losses due to congestion or wireless transmission. The authors in [4] discuss the problems of bandwidth estimation algorithm applied at source side of TCP. The authors try for reviewing, analyzing and estimation of accuracy in the proposed algorithm, and they proposed a new bandwidth estimation algorithm known as time interval-based bandwidth estimation technique that involves modification at sender-side TCP. The researchers in [5] present the estimation of proportion of incoming traffic at a router. When the observed queue length is not matching with predicted queue length, then the error is because of non-responsive traffic. The authors in [6] extended the TCP model as a function of network utilization factor for steady-state TCP throughput. The authors tried to increase performance of TCP during fast retransmission and fast recovery in timeout periods. The researchers in [7] tried to enhance the performance of TCP in 3G wireless networks by reducing the channel delay. The authors develop window regulator algorithm in which the receiver's acknowledgment packet contains the information on the wireless channel condition. It improves the selective repeat acknowledgement module of TCP. It minimizes the packet losses by correcting the misinterpretation that loss of packet is because of transmission error and not because of congestion. The authors in [8] discuss on traffic characteristics comprising of voice, video and Skype. They apply transport layer protocol, service type and network conditions for analyzing the traffic characteristics and classification of traffic type. The researchers in [9] discuss on video quality of Skype video calls. They analyze the performance of Skype by changing rate of packet loss, propagation delay, and bandwidth. They conclude that Skype works well during small amount of packet loss, and for propagation delay change. Skype utilizes the available bandwidth efficiently. The authors demonstrate that Skype video calls TCP friendly and respond for congestion. The authors in [10] tried to improve TCP performance in wireless

networks using NDG loss predictor. Using the loss predictor, the sender can understand the packet loss is due to transmission or congestion in wireless networks. The authors develop an explicit condition for loss control parameter P_{\max} . The authors also discuss the stability analysis of the network system. The researchers in [11] demonstrate that real-time communication needs congestion control as well as minimizing queuing delay. TCP is not advisable for real-time web communication because of its retransmission policy. They introduced a controller at the receiver which maintains the queue size. When there is high queue formation, inform the sender to reduce the sending rate. The literature review reveals that less work has been made by the researchers to understand the proportion of effect of non-responsive traffic on TCP in wireless networks.

3 System Model and Performance Analysis

A stochastic model representing dynamics of TCP window in wireless networks is,

$$\frac{\partial w}{\partial t} = \frac{1}{R(t)} - \frac{w(t)w(t-R(t))}{2R(t-R(t))}p(t-R(t)) + \beta \frac{w(t-R(t))}{R(t-R(t))} \quad (1)$$

On the right side of Eq. (1), first term defines additive increase, second term defines multiplicative decrease and third term controls the transmission loss in wireless network. A stochastic model representing queue dynamics at the bottleneck link in the absence of UDP flows is,

$$\frac{\partial q(t)}{\partial t} = \frac{Nw(t)}{R(t)} - C_d \quad (2)$$

On right side of Eq. (2), first term indicates the rate of incoming packets and second term $C_d = q(t)/R(t)$ indicates the service rate, and N denotes number of active TCP flows. The packet drop probability is given by,

$$P(t) = \begin{cases} 0, & q(t) \in [0, t_{\min}] \\ \frac{q(t) - t_{\min}}{t_{\max} - t_{\min}} P_{\max}, & q(t) \in [t_{\min}, t_{\max}] \\ 1, & q(t) \geq t_{\max} \end{cases} \quad (3)$$

t_{\min} and t_{\max} are the threshold values of the AQM router buffer. When the queue length is $[0, t_{\min}]$, all the incoming packets are stored in the buffer; during $[t_{\min}, t_{\max}]$ randomly some packets are dropped to reduce the sending rate, when size of queue exceeds the threshold value and buffer overflows; thus, all incoming packets are dropped by the router. The round-trip delay calculation model is given by,

$$R(t) = \frac{q(t)}{C} + T_p \quad (4)$$

T_p is the propagation delay, whose value is fixed at 5 ms for 1000 km distance. The non-responsive flow (UDP) is assumed to be constant bit rate (CBR). A mathematical model representing UDP flow is given by,

$$X_u = C_t \quad (5)$$

where C is a constant bit rate value which is pumping into the network by UDP application. The queue dynamics in presence of TCP and UDP applications is given by,

$$\frac{\partial q(t)}{\partial t} = \frac{(Nw(t) + X_u)}{R(t)} - C_d \quad (6)$$

3.1 Wireless Loss-Predictor: Normalized Delay Gradient (NDG)

$$f_{\text{NDG}} = \frac{(\text{RTT}_i - \text{RTT}_{i-1}) (W_i + W_{i-1})}{(\text{RTT}_i + \text{RTT}_{i-1}) (W_i - W_{i-1})} \quad (7)$$

Using f_{NDG} , is packet loss was because of transmission or congestion can be predicted. NDG has lower value when traffic load is less. NDG has higher value when traffic load is high. Based on that information whether to increase or decrease the sender window size.

- (i) If $f_{\text{NDG}} > 0$, the next packet loss is because of congestion.
- (ii) if $f_{\text{NDG}} \leq 0$, the next packet loss is because of wireless transmission error.

Thus,

If ($f_{\text{NDG}} > 0$) % congestion loss.

$\beta = 0$

end

If ($f_{\text{NDG}} \leq 0$) % Wireless transmission error.

$\beta = 0.1$

end

The network model used for simulation is as given in Fig. 1. R_1 and R_2 are the RED routers communicating using wired network. Router R_2 is a part of base station which provides wireless service. The simulation was done using MATLAB programming. The initialization of network parameters is $C_d = 2600$ packets, $P_{\text{max}} = 0.05$, $t_{\text{min}} = 200$ packets, $t_{\text{max}} = 500$ packets, $t_p = 0.05$ s, and packet size = 1000 bytes, $N =$

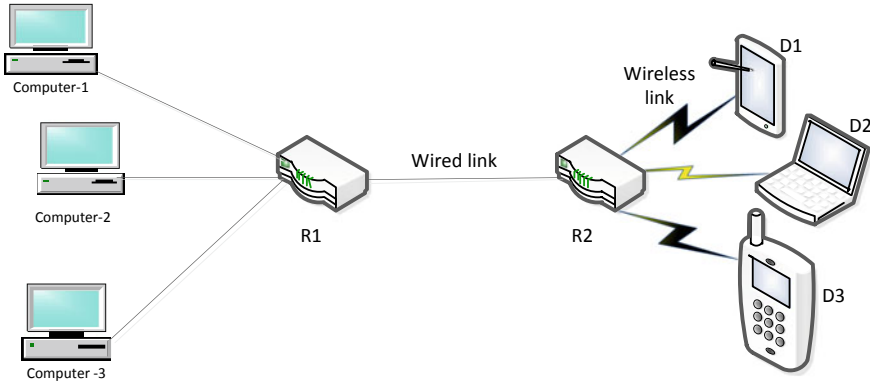


Fig. 1 Network model

10, $C = 5$ packets. The experiment was conducted in two parts—(i) in absence of non-responsive (UDP) flows and (ii) in presence of non-responsive (UDP) flows.

The distribution of round-trip time delay in the absence of UDP flows is as shown in Fig. 2. RTT value changes over the interval [0.0186, 0.2424 s] having an average RTT value of 0.2299 s. RTT is a function queue length. When queue length increases at the bottleneck link, RTT value also increases and when length decreases, RTT value queue also decreases. That is the reason the graph appears as a seesaw shape. The increase and decrease properties of queue length are due to responsiveness of TCP. The performance of RTT in presence UDP flows is shown in Fig. 3. The RTT

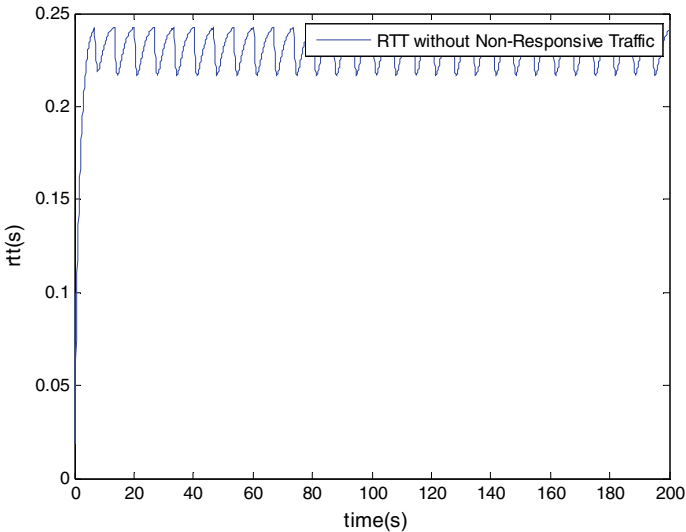


Fig. 2 Variation RTT in absence of UDP

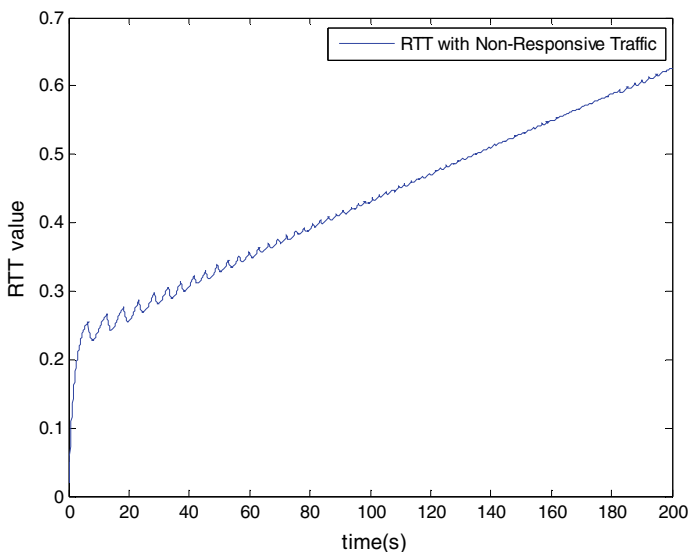


Fig. 3 Variation RTT in presence of UDP

value changes over the interval $[0.0184, 0.6261 \text{ s}]$ having an average RTT value of 0.4276 s . The graph of the curve linearly increases with time because traffic is the combination of TCP and UDP flows. The seesaw path of the graph shrinks because TCP is responsive to network and UDP is non-responsive to network. The average RTT value changes from 0.2299 to 0.4276 s , which is almost double. The queue dynamics of TCP flows is demonstrated in Fig. 4. Over the bottleneck link, queue size changes over the interval $[20, 500 \text{ packets}]$ having an average queue length of 467 packets. Queue is a function of TCP sender window. When sender TCP notices the congestion, decrease the sending rate results in decrease of queue size and increase in resources without the indication of congestion, and increase the flow rate. The performance of queue in presence of UDP flow is shown in Fig. 5. The queue length changes over the interval of $[22, 500 \text{ packets}]$ having an average queue of 482 packets. The average queue size increases from 467 to 482 packets in presence of UDP. UDP transmits 5 packets of size 1000 bytes per second continuously into the network, and UDP is non-responsive to network. Random early detection (RED) algorithm monitors the congestion level at the ingress point of the router. The traces of probability of packet dropping are shown in Fig. 6. There is no packet drop till buffer queue size attains a minimum threshold value of 200 packets. RED starts dropping the packets randomly after 200 packets noticing the formation of congestion. When the buffer queue size reaches maximum threshold value of 500 packets, all the incoming packets are dropped and the probability value becomes 1 . The average probability value is 0.06 . The traces of probability of packet dropping by RED in presence of UDP flows are shown in Fig. 7. The average probability value is 0.09 . The graph of Figs. 6 and 7 shows that the number of packets dropped is

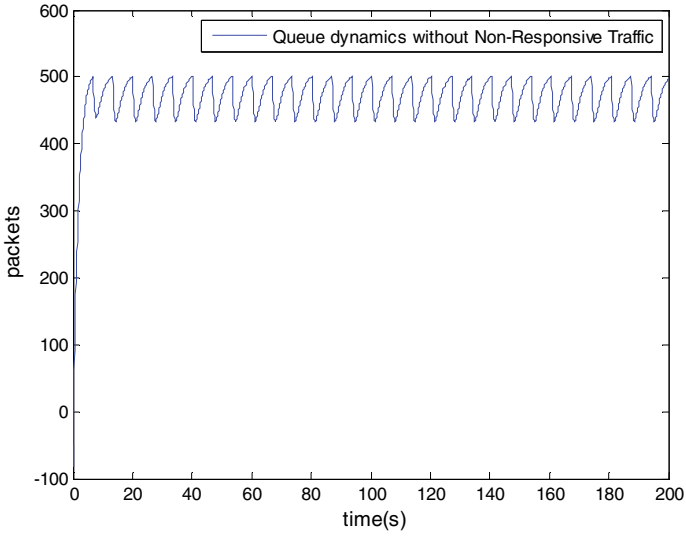


Fig. 4 Queue dynamics in absence of UDP

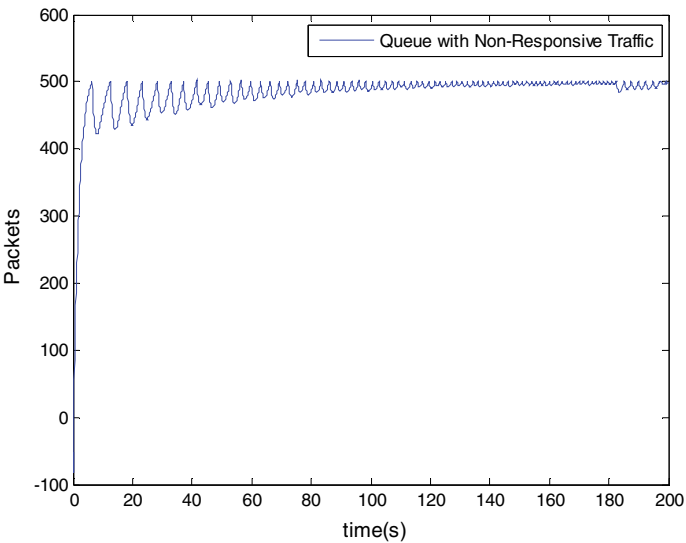


Fig. 5 Queue dynamics in presence of UDP

larger in Fig. 7 because of average probability value is 0.09. To understand the right direction of sender window adjustment, the concept of normalized delay gradient is applied. The performance of NDG is demonstrated in Figs. 8 and 9. Based on NDG value, the packet that enters into the network has lost because of transmission or

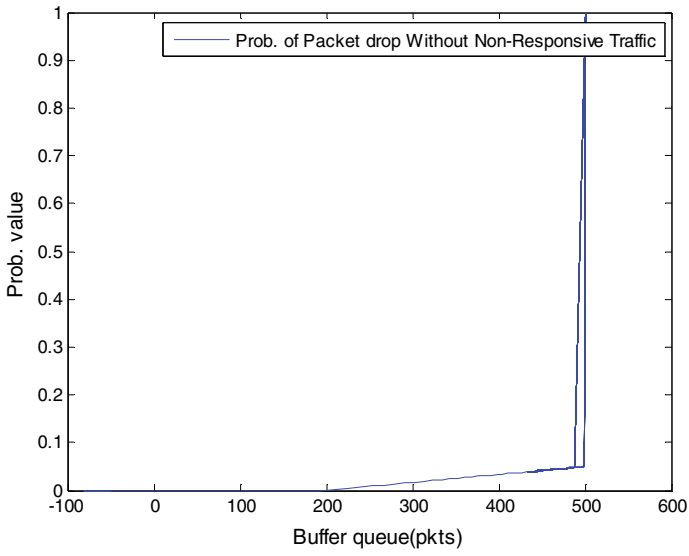


Fig. 6 Packet dropping probability of RED in absence of UDP

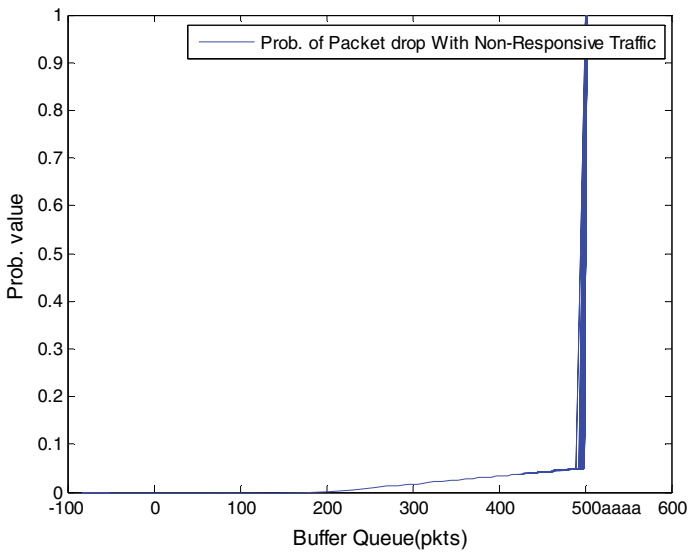


Fig. 7 Packet dropping probability of RED in presence of UDP

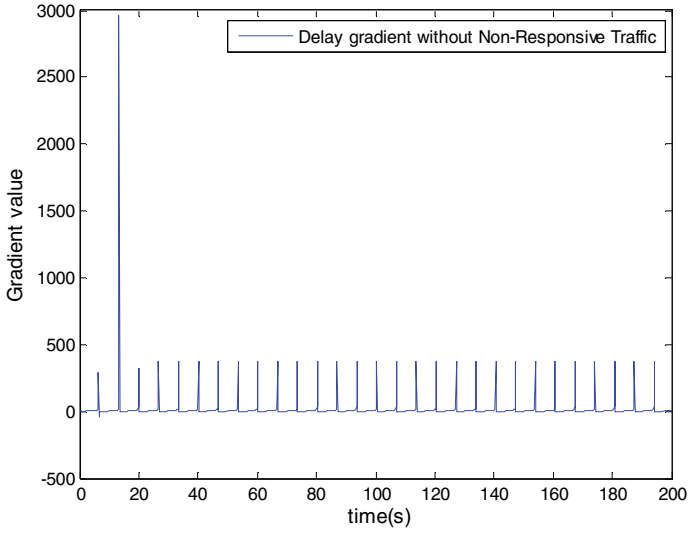


Fig. 8 f_{NDG} in absence of UDP

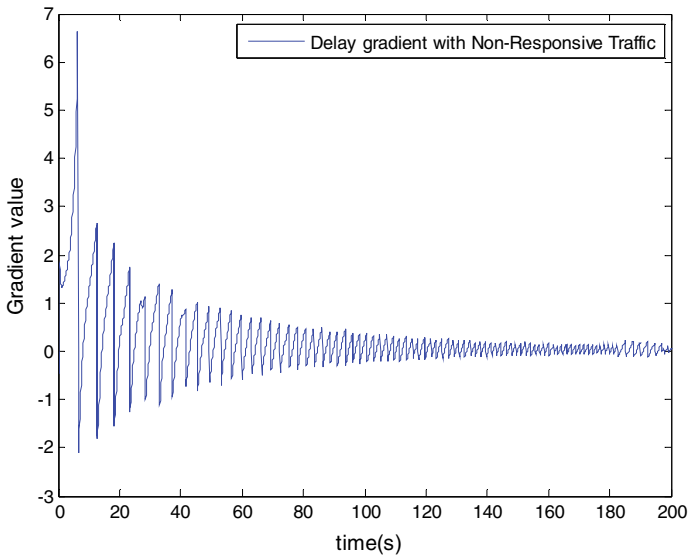


Fig. 9 f_{NDG} in presence of UDP

congestion can be predicted. Beta is a parameter used to invoke the algorithm suitably for congestion losses and transmission losses over wireless link. The variation of beta values is shown in Fig. 10. The traces of Fig. 11 demonstrate the CRB flow. CRB flow is non-responsive to network. The linearly increasing graph shows that UDP

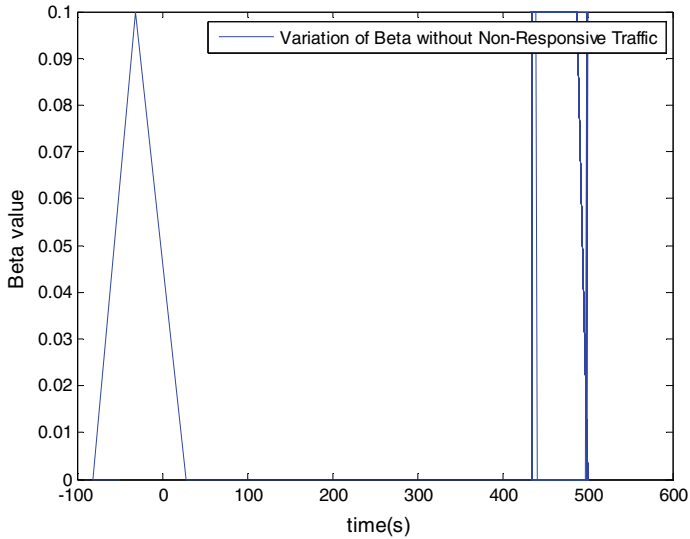


Fig. 10 Distribution of beta value

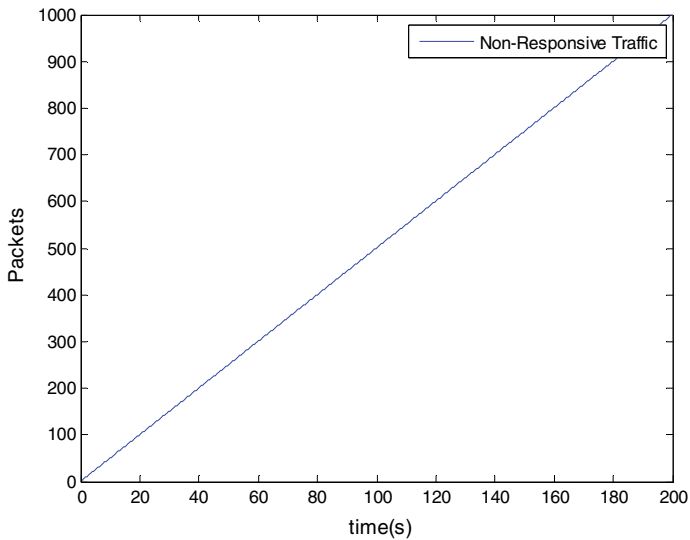


Fig. 11 Transmission of CBR flow into the network

is continuously pumping the data into the network at a fixed rate. Figures 12 and 13 show the performance sender window for TCP flows and for combined flows of TCP and UDP, respectively. The average window size is 58 packets and 63 packets,

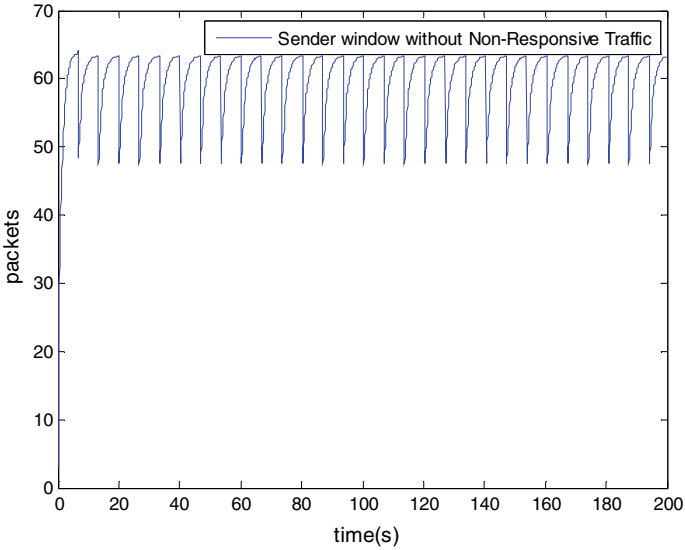


Fig. 12 Sender window dynamics in absence of UDP

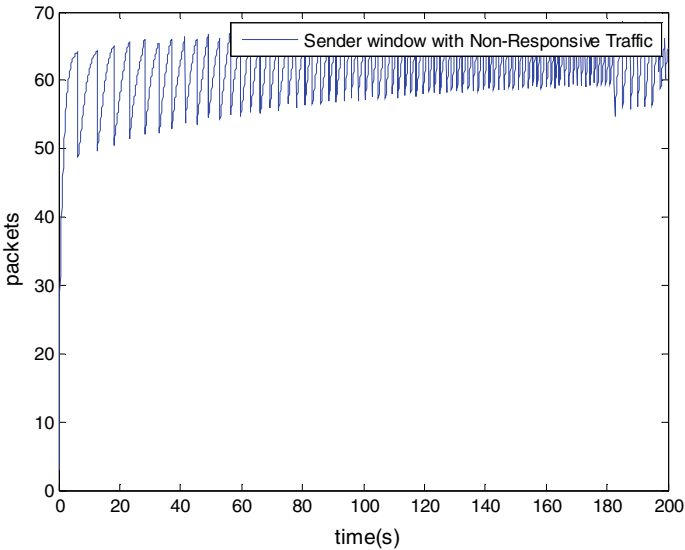


Fig. 13 Sender window dynamics in presence of UDP

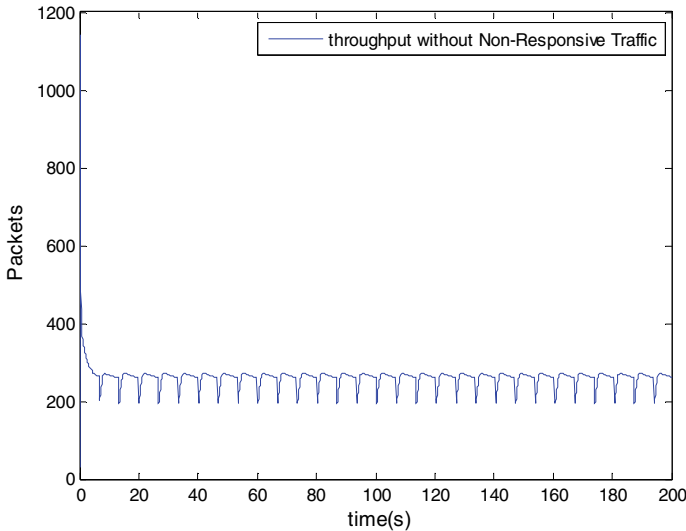


Fig. 14 Performance of throughput in absence of UDP

respectively. The performance of window depends on RTT value, probability of packet drop and beta. Window changes are very frequent in Fig. 13 because CBR flows add more packets into the network causing congestion which results in slower the sending rate. The performance of throughputs in the absence of UDP flow and in presence of UDP flow is shown in Figs. 14 and 15, respectively. Throughputs for TCP flows reduce by 40% in presence of UDP; even CBR flow rate is 5 packets per sec. It is because UDP captures more network resources.

4 Conclusion

The research work deals with performance analysis of TCP flows in presence UDP flows. TCP is responsive networks. Sender TCP increases or decreases its sending rate understanding the congestion status. UDP is non-responsive to network. In this work, CBR flow is a UDP flow which continuously sending the data into network as determined in the beginning till all the data gets exhausted. This characteristic of UDP has a negative effect on TCP flows. We notice the degradation in the performance of TCP. In this experiment, a small data rate of 5 packets per second of UDP flow decreases the throughput performance of TCP by 40% which is a matter of great concern. UDP captures large amount of network resources, and TCP gets deprived of network resources.

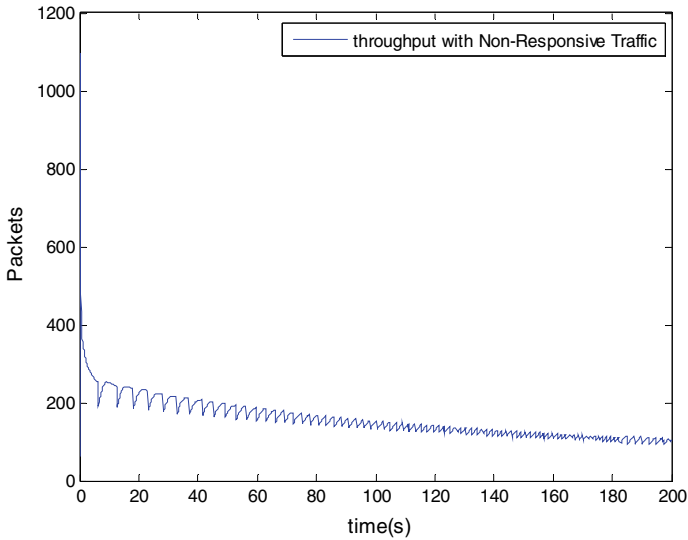


Fig. 15 Performance of throughput in presence of UDP

References

1. H. Balakrishna, V.N. Padmanabhan, S. Sheshan, R.H. Katz, A comparison of mechanisms for improving TCP performance over wireless links, in *Proceedings of ACM SIGCOMM 96*, Stanford, CA, August 1996
2. C.P. Fu, S.C. Liew, TCP VenO: TCP enhancement for transmission over wireless access networks areas. *IEEE J. Sel. Commun.* **21**(2) (2003)
3. K. Xu, Y. Tian, N. Ansari, TCP-Jersey for wireless IP communications. *IEEE J. Sel. Commun.* **22**(4) (2004)
4. A. Capone, L. Fratta, F. Martignon, Bandwidth estimation schemes for TCP over wireless networks. *IEEE Trans. Mob. Comput.* **3**(2) (2004)
5. Z. Zhao, S. Darbha, A.L. Narasimha Reddy, A method for estimating the proportion of nonresponsive traffic at a router. *IEEE/ACM Trans. Netw.* **12**(4) (2004)
6. A. Jayanathan, H. Sirisena, V. Garg, Analytical model of TCP with enhanced recovery mechanism for wireless environment. 1-4244-0353-7/07(c) *IEEE Proce. ICC 2007*
7. M.C. Chan, R. Ramjee, Improving TCP/IP performance over third generation Wireless Networks. *IEEE Trans. Mob. Comput.* **7**(4) (2008)
8. D. Bonfiglio, M. Mellia, M. Meo, D. Rossi, Detailed analysis of Skype traffic. *IEEE Trans. Multimed.* **11**(1) (2009)
9. X. Zhang, Y. Xu, H. Hu, Y. Liu, Z. Guo, Y. Wang, Profiling skype video calls: rate control and video quality. Institute of Computer Science and Technology, Peking University, Beijing 100871 (2015)
10. N.G. Goudru, B.P. Vijayakumar, Performance analysis and enhancement of TCP using normalised delay gradient in wireless networks. *Elixire Netw. Eng.* **90**, 37473–37489 (2016)
11. G. Carlucci, L. De Cicco, S. Holmer, S. Mascolo, Congestion control for web real-time communication. *IEEE/ACM Trans. Netw.* (2017)

Microstrip Patch Antenna Analysis for 5G Millimeter-Wave Communication: A Survey



H. V. Pallavi, A. P. Jagadeesh Chandra, and Paramesha

1 Introduction

For any wireless application, we need antenna. A microstrip patch antenna is an important device in the communication systems. The rectangular microstrip patch antennas are largely used in wireless communication systems because of their low cost, small-profile, small-size, lightweight, and they are especially preferred for narrow bandwidth and bandwidth enhancement. Though different geometrical shapes and dimensions can be used in the patch antenna design, the rectangular and circular patches are the ones most commonly used in many of the applications. The antenna dimension becomes very small at millimeter-wave frequency. At 60 GHz, the wavelength is already 5 mm, when fabricated it provides a still smaller size. Important wireless applications lie in the band 900 MHz to 5.8 GHz. The millimeter-wave (mm wave) mobile communication works on 28 and 38 GHz frequency. In order to satisfy the requirements of mobile data, millimetre-wave communication has got a great importance. Millimetre wave is that section of electromagnetic spectrum of wavelength 10–1 mm, and it translates to a frequency scale of 300 to 30 GHz.

Some of the popular applications of millimetre wave are Bluetooth, WI-FI, wireless LAN. The data rate offered by Bluetooth is in kbps, WI-FI offers a data rate

H. V. Pallavi (✉)

Department of Electronics and Communication Engineering, Government Engineering College, Hassan, Karnataka, India

e-mail: bpallavi178@gmail.com

A. P. J. Chandra

Department of Electronics and Communication Engineering, Adichunchanagiri Institute of Technology, Chikamagalur, India

e-mail: apjchandra@gmail.com

Paramesha

Department of Electronics and Communication Engineering, Central University of Karnataka, Kalaburagi, India

e-mail: parameshap91@gmail.com

of Mbps and a few tens of Mbps. The wireless LAN provides a data rate of Gbps. High-resolution RADAR or high-resolution imaging is possible using millimeter-wave frequency range. Although there are many advantages of millimetre wave, they are still not popular because of many problems or the challenges in millimeter wave. Some of them have already been solved. The main goal of 5G will be aimed at enhancing the capacity of the network with a good coverage at a low cost. The entire spectrum of the millimeter wave is divided into several sections. The Ka band frequency ranges between 26.5 and 40 GHz [1, 2]. Several researches show that the bandwidth enhancement can be done by using various techniques like slotting the patch, using thick substrates and also the substrates with low effective permittivity, by incorporating a greater number of resonances and so on. An array configuration may be used to improve the gain, as an array antenna can easily fit for millimeter-wave application [3].

Organization: An Introduction to microstrip patch antenna, a 5G millimeter-wave bands with the basic design of a patch antenna are discussed in Sects. 1 and 2, respectively. In Sect. 3, various patch antenna designs for 5G applications are highlighted. In Sect. 4, a comparative analysis of different types of patch antennas as well as array antennas is tabulated separately in terms of their performance parameters, and conclusions are contained in the Sect. 5.

2 Microstrip Patch Antenna Design

The microstrip antenna is a metallic patch that is placed over a dielectric substrate on a ground plane as shown in Fig. 1a. The thickness ' t ' of patch should be such that it is very much smaller than λ_0 . [10]. The length ' L ' of the rectangular patch should normally be $\lambda_0/3 < L < \lambda_0/2$. Length takes care of frequency. The patch width ' W ' should be less than λ_0 . [11]. Ground plane and the patch are detached by a thick

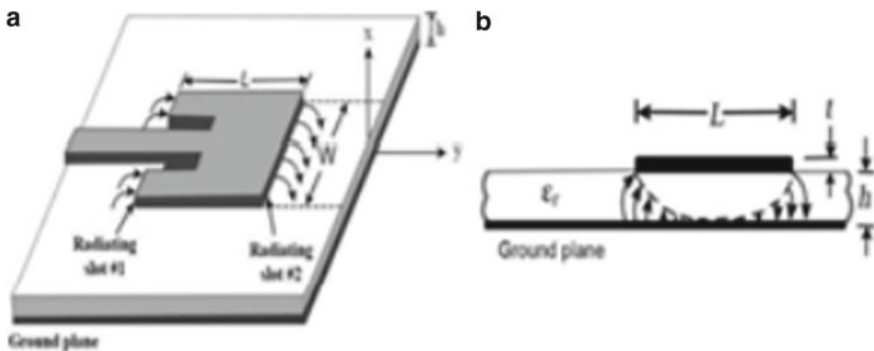


Fig. 1 a Microstrip antenna and b side view [12]

substrate material as shown in Fig. 1b [12]. The antenna dimensions are calculated from Eqs. 1–6.

The width of patch can be calculated using the equation:

$$W = \frac{C}{2f_0\sqrt{\frac{\epsilon_r+1}{2}}} \quad (1)$$

C indicates the speed of light, which is given by equation $C = 3*10^{11}$ mm/s, f_0 indicates resonant frequency, ϵ_r indicates dielectric constant.

The substrate height is calculated by the equation:

$$h = \frac{0.0606\lambda}{\sqrt{\epsilon_r}} \quad (2)$$

The effective dielectric constant is given by the equation:

$$\epsilon_{re} = \frac{\epsilon_r + 1}{2} + \frac{\epsilon_r - 1}{2} \left[1 + 12 \frac{h}{w} \right]^{-\frac{1}{2}} \quad (3)$$

The path length ΔL due to the fringing fields is extended on its both, as defined in Eq. (4):

$$\Delta L = 0.412h \frac{(\epsilon_r + 0.3)\left(\frac{w}{h} + 0.264\right)}{(\epsilon_{re} - 0.258)\left(\frac{w}{h} + 0.8\right)} \quad (4)$$

Effective length of antenna is given in Eq. (5); from this the original length of the rectangular microstrip patch antenna can be found

$$L_{\text{eff}} = \frac{C}{2f_0\sqrt{\epsilon_{re}}} \quad (5)$$

The length and the width of the substrate are given as:

$$\begin{aligned} Lg &= Lp + 6h \\ Wg &= Wp + 6h \end{aligned} \quad (6)$$

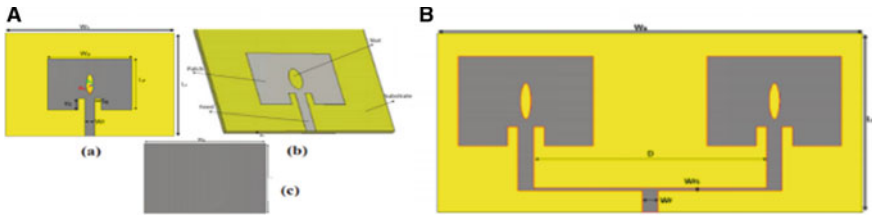


Fig. 2 a Single antenna and b 2×1 array antenna [4]

3 Literature Review

3.1 Dual-Band Rectangular Patch Antenna Array

Firstly, a single radiating element as shown in Fig. 2a resonating at the frequencies 3.43 and 5.4 GHz, having bandwidths 400 and 250 MHz, respectively, is designed using FR-4 Epoxy substrate. An elliptical slot is introduced to enhance the bandwidth. VSWR is 1.29 and 1.07, respectively, at 3.43 and 5.4 GHz. The S_{11} at 3.43 GHz is -18.27 dB, gain is 6.85 dB. At 5.4 GHz S_{11} is -29.75 dB, gain is 6.44 dB. Then an array concept is used as shown in Fig. 2b to improve the gain.

The 1×2 proposed antenna has gain 8.77 dB, S_{11} 1.06, while a center series-fed technique is employed to excite antenna. Simulation is done using CST microwave studio [4].

3.2 Dual-Band Circular Patch Antenna Array

In this case, the 28 and 45 GHz are the resonating frequencies, having bandwidths 1.3 and 1 GHz, respectively, as shown in Fig. 3a with an elliptical slot. Efficiency is found to be 85.6 and 95.3% at these two frequencies. S_{11} is -40 and -14 dB, gain is 7.6 and 7.21 dB. It has a $6 \text{ mm}^3 \times 6 \text{ mm}^3 \times 0.578 \text{ mm}^3$ size which is compact. In order to improve the gain, an array antenna of 1×4 elements is implemented for mobile communication as shown in Fig. 3b, with 28, 34 and 45 GHz frequencies and gain being 13.5 dB with the radiation efficiency being 98.75%. The array is excited using center series-fed technique [3].

3.3 Virtual Education and for Other 5G Applications

A microstrip patch antenna has been designed here for a good virtual education quality and for other 5G applications at a resonant frequency of 26 GHz. The ϵ_r is 2.2 with dielectric loss tangent 0.0010. The simulation with analysis is done using

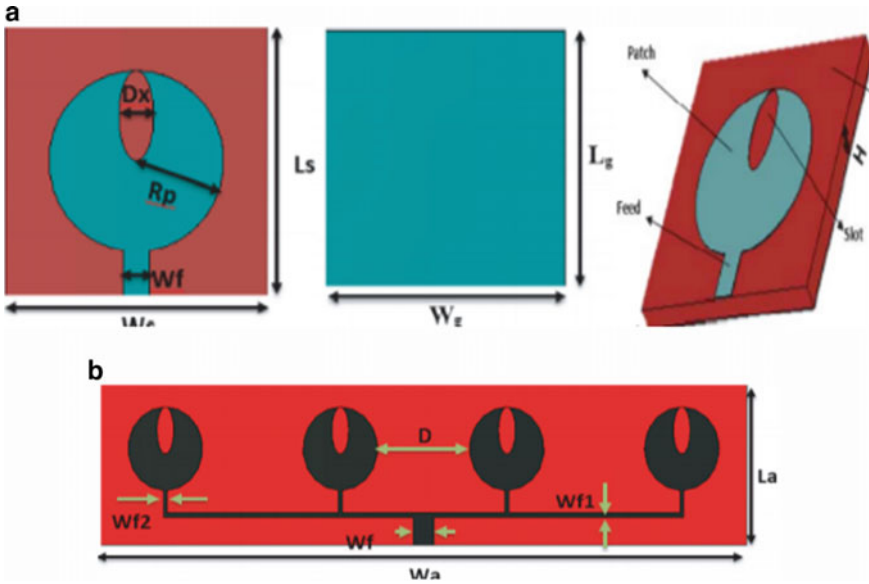


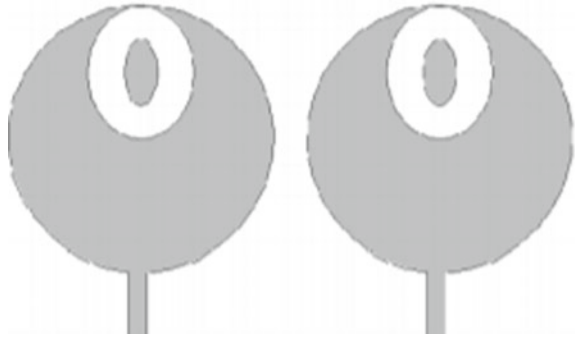
Fig. 3 a Single antenna. b 1×4 array antenna [3]

FEKO software. The S_{11} is -33.4 dB. A bandwidth of 3.56 GHz is achieved, and a VSWR less than 2 is obtained. A gain 10 dB with a radiation efficiency of 99.5% has been achieved which was benefitted during on-going lockdown around the world. The gap identified here is that the proposed antenna has return loss and bandwidth a little small that can be improved [5].

3.4 Doughnut Slot 5G MIMO Antenna

The author has designed a circular patch with two resonant frequencies at 28 and 45.54 GHz as shown in Fig. 4. The material used for the design is Rogers RT-5880 with ϵ_r 2.2, loss tangent is found to be 0.0009 for the given material, and the height of the dielectric material is found to be 0.508 mm. An elliptical doughnut slot is done onto the circular patch. The antenna dimension is 6 mm \times 6 mm. The bandwidths obtained are 1.14 and 45.54 GHz at the above two frequencies, respectively. The gain is found to be 5.94 and 4.89 dB, respectively. Later, a 2×2 MIMO antenna is designed to enhance the gain, radiation and efficiency, etc. The return loss now is found to be -45.22 and -15.32 dB at the two frequencies, respectively. The gain is also progressed now to 7.67 dB. The ECC technique is used measure the correlation between array elements [6].

Fig. 4 Doughnut slot 5G MIMO antenna [6]



3.5 Directional Slotted Single-Element Antenna

A compact-sized antenna is designed here which to operates at 28 GHz using FR4 substrate. The patch is modified by slotting the radiator to achieve specific requirements. After simulation, it is seen that the antenna size is reduced. The radiation efficiency is 86.73%. The impedance matching is good with the VSWR less than 2. The ϵ_r is also found more. There is an improvement in the radiation characteristic compared to the conventional microstrip antenna. The improvement is the impedance bandwidth, and the gain is enhanced by 3.86% and 11.63%, respectively; the pattern is more directive with HPBW decreased by 8.34% at $\phi = 0^\circ$ and at $\phi = 90^\circ$ at 13.42%, compared to HPBW of a conventional antenna [7].

3.6 Linear Array Antenna

This paper presents a 16×1 array element in a linear fashion as shown in Fig. 5c, operating at 28 GHz. The antenna element is made of Teflon-based RT/duroid 5880 with ϵ_r 2.2 and thickness 0.381 mm.

There is a gap associated with the feed element as shown in Fig. 5a. In between the adjacent elements to reduce the coupling, a slot is used which is the DGS in the array as shown in Fig. 5b. The maximum gain here is 17.4 dBi, and bandwidth is greater than 2 GHz [8].

3.7 A Square Patch Antenna

A square patch antenna with improved return loss is proposed here at 30 GHz which has a diamond-shaped slot. At first a conventional inset-fed square patch is designed with quarter wavelength impedance matching line. In order to achieve a

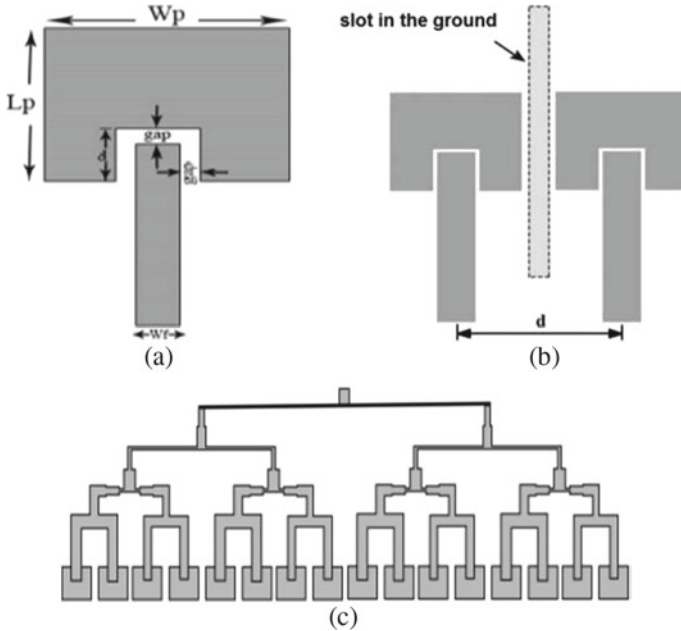


Fig. 5 a Single element with gap-coupled feeding, b two adjacent antennas with DGS, c 16×1 linear antenna array [8]

low return loss and high bandwidth, the diamond-shaped slot is added. This increases the bandwidth to 46 MHz and a return loss to 10.751 dB as shown in Fig. 6a, b [9].

3.8 Massive MIMO for 5G

Antenna type used here is 10×10 elements massive multiple in multiple out patches.

This antenna exhibits a return loss of -14.29 dB, a VSWR of 1.43, a 450 MHz bandwidth is achieved with a gain of 27.4 dB. Here a single element is first designed according to the specifications, and then, an array of 100 elements is designed to meet the 5G requirements. The Rogers RO4003C material is used here as it produces the highest bandwidth. Figure 7a, b illustrate a single antenna and its array [10].

3.9 Multiband Patch Antenna

A multiple in multiple out technology with 4-antenna linear array is used in this paper with the frequencies 28, 37, 41 and 74 GHz. A resonant cavity model is used here. Introduction of slits and slots helps in multiband operations. Inset-fed technology is

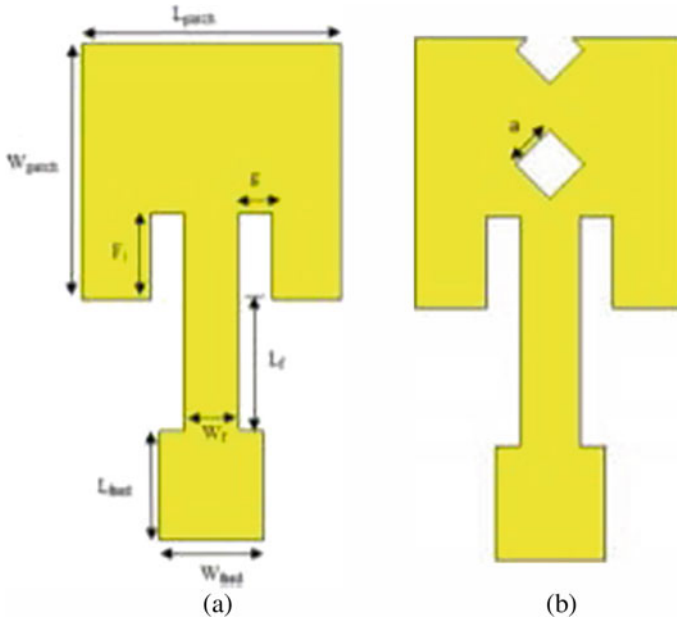


Fig. 6 **a** Square inset-fed microstrip patch antenna and **b** square inset-fed microstrip patch antenna with diamond-shaped slots [9]

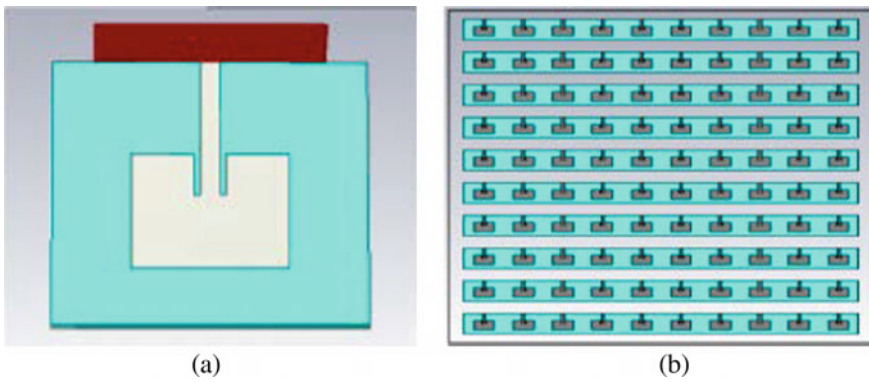
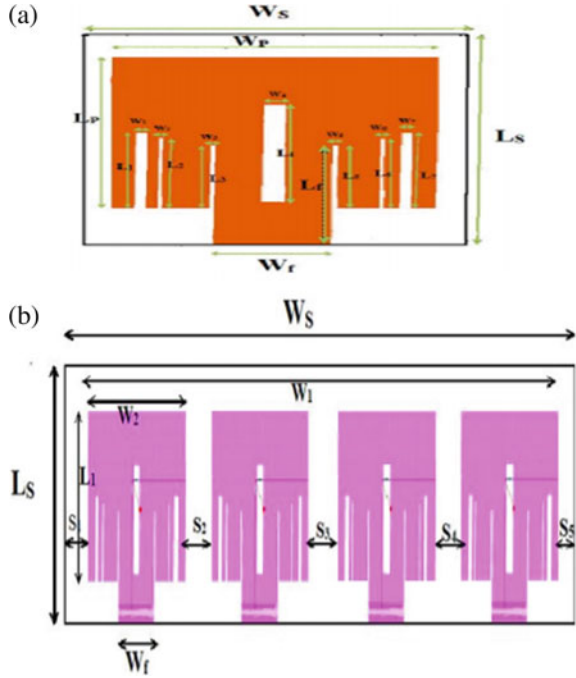


Fig. 7 **a** Single antenna and **b** array antenna 100 elements (10×10) [10]

achieved to match the impedance between the patch and the feed line as shown in Fig. 8a, b for a single antenna and an array, respectively [11].

Fig. 8 **a** Single antenna and **b** MIMO antenna



3.10 Single-Band Antenna at 60 GHz

In this design, for a single-band antenna, the E and H slots are made on the patch as shown in Fig. 9 with the 50Ω microstrip line feed being used. Simulation is done using electromagnetic simulator. Rogers RT5880 material is used as a substrate with ϵ_r 2.2, loss tangent 0.0009 and height 1.6 mm. The return loss at 60 GHz is -40.99 dB. Due to the H and E slots, there is an improvement in the bandwidth. A bandwidth of 4.028 GHz is achieved here at a frequency 59.93 GHz [12] (Fig. 10).

3.11 A Dual-Band Patch Antenna

The feeding structure in this proposed work for 28 and 39 GHz is a combination of two strips, one close to feeding port and another strip located near the slot. A coplanar waveguide (CPW) is used as feed to match the port and 50Ω resistor. An array of 8 elements is also constructed, which covers a beam width from -30 to 30° in YOZ-plane. A beam-forming requirement for the 5G wireless communications can be fulfilled here [13].

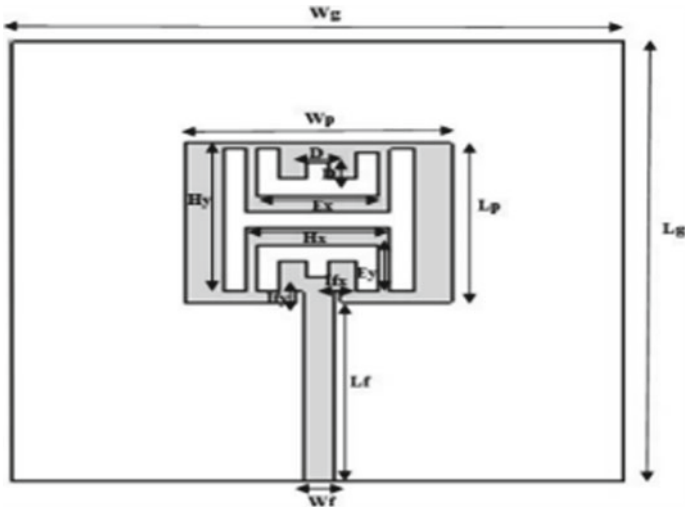


Fig. 9 Proposed single-band antenna [12]

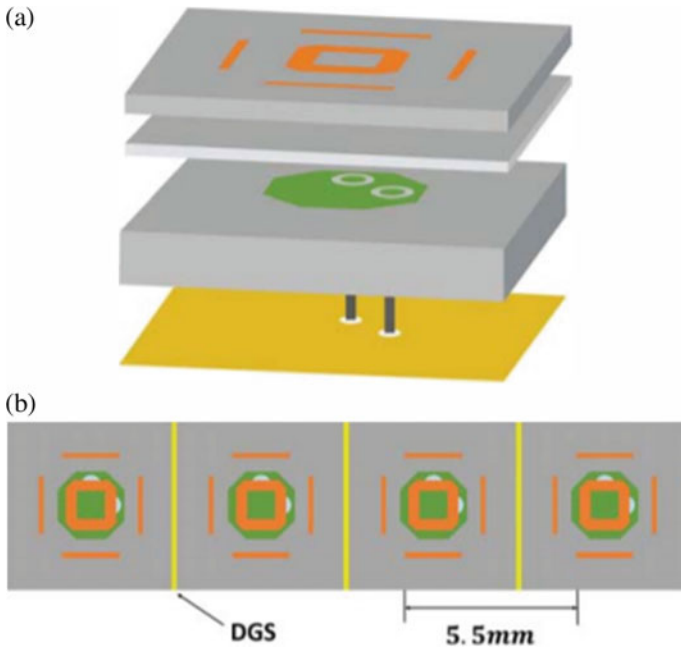


Fig. 10 a Stacked patch antenna and b its array [14]

3.12 Dual-Band Dual-Polarized Patch Antenna Array

A stacked patch at 26 and 36 GHz frequency bands has been designed here with a small patch on the top whose function is to generate the higher resonant frequency. The bottom patch helps to obtain a lower resonant frequency. The feeding technique used is a capacitive feed. Rogers RT4350 ($\epsilon=3.52$) is the material used to construct the two-layer dielectric substrate with industrial standard thickness. The Rogers RO4450F with 0.1 mm thickness acts as an adhesive layer to hold the upper and lower dielectric materials. The designed 4-element antenna has 5.5 mm spacing. The array has a good impedance matching and less than 10 dB return loss compared to an earlier antenna array which used which used a defected ground structure (DGS) technology [14].

3.13 Collinear Array Concept for Millimeter-Wave Spectrum

Here in this design, the author has used a collinear array concept. A good impedance matching is obtained between the feed line and array by using an edge feed mechanism. Four frequencies here are 24.7, 27.7, 35.4 and 40.3 GHz. Polarization is linear. This antenna provides a good efficiency, a return loss and a large gain in the entire given spectrum. It has a wider beam width in the E-plane and the H-plane [15].

4 Comparison of Various Patch Antenna Designs at Millimeter-Wave Frequency

Different configurations of microstrip patch antennas have been designed for millimeter-wave communication systems. A comparative analysis of the patch antennas is done and shown in Tables 1 and 2 based upon the microstrip antenna configuration for a single and an array element, respectively. Various antenna parameters are highlighted in the tables, and a comparison is done with respect to the size, return loss, bandwidth, gain, VSWR, etc.

4.1 Microstrip Patch Antenna Arrays

For 5G applications, the antenna requires a higher gain and directive beam that can be controlled in specific direction. It may be difficult to accomplish such high gain by single antenna. However, if there is an array of antenna, a MIMO, beam steering techniques, which is used for radar applications can be done. Several small antennas can be gathered together in an array antenna to obtain such high increased gain

Table 1 Comparative analysis of patch antennas for 5G millimeter-wave frequency

MPAs configuration	Fc (GHz)	Material for design	Technique applied	€r	Size of the antenna (mm)	BW (GHz)	Gain (dBi)	S ₁₁ (dB)	VSWR	Ref
Slotted microstrip patch	59.93	Rogers RT5880	Microstrip line	2.2	8 × 8	4.028	5.42	-40.99	Less than 2	[16]
Single-band microstrip square patch	28	air	NA	1	4.8 × 4.4	4.3	10.3	-57	1	[17]
Microstrip antenna with two rectangular slots	29.74	RT5880	NA	2.2	3.32 × 4.456	2.26	7.1-7.65 dBi	-21.8	NA	[18]
Double F-slot microstrip patch antenna	58.10	silicon substrate	Inset feed	11.9	0.9843 × 0.62	NA	5.99	-32.5	NA	[19]
Pin-fed 3-notch circular patch antenna	60	RT5880	NA	2.2	5 × 5 × 0.1	NA	70839	-27	NA	[20]
Dual-band mmw microstrip antenna	28/39	Thick dielectric substrate	Coplanar wave guide feed	4.4	4 × 9.45 × 0.26	9.3/4	2.3/4.8	NA	NA	[13]
Multiband microstrip patch antenna	37/54	RT5880	Inset feed	2.2	7.2 × 5 × 0.787	5.5/8.67	5.5/6	-25.8/-27.8	NA	[21]
Compressed microstrip patch antenna	28.3/50.3	RT5880	Inset feed	2.2	11 × 8 × 0.5	NA	2.6	21/31	1.58/0.5	[22]

(continued)

Table 1 (continued)

MPAs configuration	F _c (GHz)	Material for design	Technique applied	€r	Size of the antenna (mm)	BW (GHz)	Gain (dBi)	S ₁₁ (dB)	VSWR	Ref
Microstrip patch antenna	27.91	RT5880	Inset feed	2.2	4.105 × 3.362	582 MHz	6.69	-12.59	1.77	[23]
Single-band microstrip patch antenna	38	FR-4	Microstrip line	4.4	NA	1.021	low	-24.35	1	[24]
A wide-band patch antenna	28	RT5880	Coaxial feed	2.2	5 × 5.26 × 0.254	2.11	5.2	-20.03	1.22	[25]
Rectangular microstrip patch antenna	27.954	RT5880	Inset feed	2.2	6.285 × 7.235 × 0.5	900 MHz	6.63	-13.48	1.5376	[26]
Rectangular microstrip patch antenna	38.14	Taconic RF-60(tm)	Coaxial—probe feed	6.15	5 × 4 × 0.64	4.06	4	-31	1.10	[27]
Rectangular microstrip patch antenna	27.4	RT5880	Inset feed	2.2	6.3 × 7.1 × 0.5	NA	6.85	-22.2	1.34	[28]
Microstrip patch antenna	26	RT5880	Corporate feed	2.2	6 × 6 × 0.65	3.56	10	-33.4	1.03	[5]

Table 2 Comparative analysis of patch antenna array for 5G millimeter-wave frequency

MPAs configuration	F_c (GHz)	Material for design	Technique applied	ϵ_r	Size of the antenna (mm)	BW (GHz)	Gain (dBi)	S_{11} (dB)	VSWR	References
Massive MIMO patch antenna with 10×10 array	28	Rogers RO4003C	Inset feed from the edge of the patch	3.38	2.78×3.62	450 MHz	12.4	-14.29	1.43	[29]
MIMO 4-antenna linear array	28/37/41/74	FR4	Inset feed	4.4	2.767×2.8	1.98/0.64/0.23/4.95	-22/-37/-10/-19	NA	Less than 2	[30]
MMW microstrip patch antenna array	38.6/47.7/54	Rogers RT5880	Tapered line feeding	2.2	$6 \times 6.25 \times 0.578$	3.5/2.5/1.3	12.2/11.6/12.1	-13.6/-22.5/-18	NA	[31]
Linearly polarized 4×16 array antenna	26.2	Rogers RO4003C	Grounded coplanar wave guide	3.55	NA	2	Drops by 4 dB when scanning to -60°	Less than -10 dB	NA	[16]
Dual-band dual-polarized array	28/39	RT4350	Capacitive feed	3.52	$3.5 \times 3.5 \times 1.03$	11.9/12.1	6.5/5.8	Less than -10 dB	NA	[32]
2×4 dual-band mm-w antenna array	28/38	NA	Wave guide/SMA feed port	NA	6.8×6.3	1.189/1.42	15.8/13.9	-19.75/-19.55	Less than 1.5	[33]

(continued)

Table 2 (continued)

MPAs configuration	F_c (GHz)	Material for design	Technique applied	ϵ_r	Size of the antenna (mm)	BW (GHz)	Gain (dBi)	S_{11} (dB)	VSWR	References
Microstrip antenna array of $2 \times 1/4 \times 1$	29.8	RO3006	Microstrip line	6.5	$7 \times 14.5 \times 0.13/9 \times 27 \times 0.13$	337.3 MHz/301.4 MHz	9.4/11.7	-25.82/-19.61	0.8888/1.8230	[34]
Microstrip antenna array of $2 \times 1/4 \times 1$	28	FR-4	Inset feed	4.35	2.303×3.215	NA	9.52/11.2	-3.77/-21.44	1.624/1.650	[35]
Multi-MMW antenna array	24.7/27.7/35.4/40.3	RT5880	Edge feed	2.2		1	18.8 at 26.6	-39.3 at 35.4 GHz	1 at 35.4	[36]
4 \times 1 element circular phase array of rectangular patch antenna	24.81/33	RT5880	Inset feed	2.2	$25 \times 20 \times 1.6$	24.8-33 GHz	10.14	-18.64	1.85	[37]
4-element corporate series-fed antenna array	27.39	Roger RT Duroid 6002	Corporate series feed	2.94	3.89×2.93	NA	12.12	-29	1.0375	[38]
Circular microstrip patch antenna and its array of 1×4 elements	28	RT5880	Centre series fed technique	2.2	$6 \times 6 \times 0.578$	NA	8.9	Less than -10 dB	Less than or equal to 2	[39]

directive pattern that can be electronically checked in a specific direction. It becomes easier to create lot more antennas in millimeter wave using massive MIMO.

5 Conclusion

In this survey, many patch antenna designs are discussed for millimeter-wave frequency. Out of which the antenna design for 28 GHz frequency is suitable for 5G mobile phones. Antenna array can be introduced to improve overall performance of the antenna. Multiple antenna methods can be used, and MIMO can be employed to increase the capacity, reliability, etc. Gain, return loss, side lobe, beam steering are affected by the proximity of other antennas in the array. By changing the orientation of the antennas, engineers can optimize these properties. One of the challenges in millimeter-wave frequency is that the antenna dimension becomes very small. Secure data link is possible with millimeter wave by sending a pencil like beam. In millimeter-wave frequency, the antenna coverage also becomes small. Various other challenges in millimeter-wave frequency are atmospheric attenuation, rain attenuation, long-distance communication, weather conditions, fading, Doppler shift, shadowing effect, etc. Some of them have already been solved, and others are in process.

References

1. M. Ali, A. Sebak, Directive antenna for future 5G mobile wireless communications, in *32nd URSI GASS, Montreal*, 16–26 August 2017
2. T. Rappaport et al., Millimeter wave mobile communications for 5G cellular: it will work! Proc IEEE **I**(10), 335–349 (2013)
3. M.I. Khattak, A. Sohail, U. Khan, Z. Barki, G. Witjaksono, Elliptical slot circular patch antenna array with dual band behaviour for future 5G mobile communication networks. Prog. Electromagn. Res. **89**, 133–147 (2019)
4. M. Hanaoui, M. Rifi, Elliptical slot rectangular patch antenna array with dual band behaviour for future 5G wireless communication networks, in *IEEE 7th Mediterranean Congress of Telecommunications (CMT)* (2019), pp. 1–4
5. J. Colaco, R. Lohani, Design and implementation of microstrip patch antenna for 5G applications, *IEEE 5th International Conference on Communication and Electronics Systems (ICCES)* (2020), pp. 682–685
6. G.M. Amrutha, T. Sudha, Millimeter wave doughnut slot MIMO antenna for 5G applications, in *IEEE Region 10 Conference (TENCON)* (2019), pp. 1220–1224
7. A.F. Kaeib, N.M. Shebani, A.R. Zarek, Design and analysis of a slotted microstrip antenna for 5G communication networks at 28 GHz, in *19th IEEE International Conference on Sciences and Techniques of Automatic Control and Computer Engineering (STA)* (2019), pp. 648–653
8. Saada, M. Abu, T. Skaik, R. Alhalabi, Design of efficient microstrip linear antenna array for 5G communications systems, in *IEEE International Conference on Promising Electronic Technologies (ICPET)* (2017), pp. 43–47

9. B.G. Hakanoglu, M. Turkmen, An inset fed square microstrip patch antenna to improve the return loss characteristics for 5G applications, in *XXXIInd IEEE General Assembly and Scientific Symposium of the International Union of Radio Science (URSI GASS)* (2017), pp. 1–4
10. S. Choudhury, Effect of dielectric permittivity and height on a microstrip fed rectangular patch antenna. *Int. J. Electron. Commun. Technol.* 1–2 (2014)
11. D.M. Pozar, D. Schaubert, *The Analysis and Design of Microstrip Antennas and Arrays* (1992)
12. W.F. Richards, Y.T. Lo, D.D. Harrison, An improved theory for microstrip antennas and applications. *IEEE Tran. A&P* **29**, 137–145 (1979)
13. C.-Z. Han, G.-L. Huang, T. Yuan, A dual-band millimeter-wave antenna for 5G mobile applications, in *IEEE International Symposium on Antennas and Propagation and USNC-URSI Radio Science Meeting* (2019), pp. 1083–1084
14. Y. He, M. Rao, Y. Liu, G. Jing, M. Xi, L. Zhao, 28/39-GHz dual-band dual-polarized millimeter wave stacked patch antenna array for 5G applications, in *IEEE International Workshop on Antenna Technology (iWAT)* (2020), pp. 1–4
15. S. Pandey, L. Garia, Design of millimeter-wave spectrum microstrip patch antenna array for 5G wireless systems, in *Women Institute of Technology IEEE Conference on Electrical and Computer Engineering (WITCON ECE)* (2019), pp. 168–171
16. J. Saini, S.K. Agarwal, Design a single band microstrip patch antenna at 60 GHz millimeter wave for 5G application, in *IEEE International conference on Computer, Communications and Electronics (Comptelix)* (2017), pp. 227–230
17. M. Faisal, A. Gafur, S.Z. Rashid, Md.O.F. Shawon, K. Hasan, Md.B. Billah, Return loss and gain improvement for 5G wireless communication based on single band microstrip square patch antenna, in *1st IEEE International Conference on Advances in Science, Engineering and Robotics Technology (ICASERT)* (2019), pp. 1–5
18. Z. Li, R. Jian, Y. Chen, T. Chen. A novel design for millimeter wave microstrip antenna with bandwidth enhancement, in *IEEE Student Conference on Research and Development (SCoReD)* (2018), pp. 1–4
19. F. Ahmad, B. Tlili, Design and analysis of millimeter wave double F slot patch antenna for future 5G wireless communication, in *IEEE International Conference on Electrical and Computing Technologies and Applications (ICECTA)*, pp. 1–4 (2017)
20. Z. Lodro, M.Y. Lodro, M.M. Lodro, Millimeter wave pin-fed three notch circular patch antenna for 5G communication, in *2nd IEEE International Conference on Computing, Mathematics and Engineering Technologies (iCoMET)* (2019), pp. 1–3
21. Z. Lodro, N. Shah, E. Mahar, S.B. Tirmizi, M. Lodro, mmWave novel multiband microstrip patch antenna design for 5G communication, in *2nd IEEE International Conference on Computing, Mathematics and Engineering Technologies (iCoMET)* (2019), pp. 1–4
22. T. Kiran, N. Mounisha, Ch. Mythily, D. Akhil, T.V.B. Phani Kumar, Design of microstrip patch antenna for 5G applications. *IOSR J. Electron. Commun. Eng. (IOSR-JECE)* **13**, 14–17 (2018)
23. D. Mungur, S. Duraikannan, Microstrip patch antenna at 28 GHz for 5G applications. *J. Sci. Technol. Eng. Manag. Adv. Res. Innov.* **1**(1) (2018)
24. C. Şeker, T. Ozturk, M.T. Güneşer, A single band antenna design for future millimeter wave wireless communication at 38 GHz. *Eur. J. Eng. Formal Sci.* **2**(2), 35–39 (2018)
25. V.G. Prachi, S. Vijay, A novel design of compact 28 GHz printed wideband antenna for 5G applications. *Int. J. Innov. Technol. Explor. Eng. (IJITEE)* **9**(3) (2020)
26. Md.J. Ahammed, R. Sambasiva Nayak, N. Mohammed, Simulation of rectangular microstrip patch antenna for range of ka band frequency. *J. Adv. Res. Technol. Manag. Sci.* **1**(3) (2019)
27. A.M. Jajere, Millimeter wave patch antenna design antenna for future 5G applications. *Int. J. Eng. Res. Technol. (IJERT)* **6**(02), 289–291 (2017)
28. A. Sivabalan, S. Pavithra, R. Selvarani, K.M. Vinita, Design of microstrip patch antenna for 5G. *Int. J. Control Autom.* **13**(2), 546–552 (2020)
29. T. Yuwono, M. Ismail, I. Hajar, Design of massive MIMO for 5G 28 GHz, in *2nd IEEE International Conference on Computer Applications and Information Security (ICCAIS)* (2019), pp. 1–4

30. P.M. Sunthari, R. Veeramani, Multiband microstrip patch antenna for 5G wireless applications using MIMO techniques, in *First IEEE International Conference on Recent Advances in Aerospace Engineering (ICRAAE)* (2017), pp. 1–5
31. D. Imran, M.M. Farooqi, M.I. Khattak, Z. Ullah, M.I. Khan, M.A. Khattak, H. Dar, Millimeter wave microstrip patch antenna for 5G mobile communication. *IEEE Int.*
32. Y. He, M. Rao, Y. Liu, G. Jing, M. Xi, L. Zhao, 28/39-GHz dual-band dual-polarized millimeter wave stacked patch antenna array for 5G applications, in *IEEE International Workshop on Antenna Technology (iWAT)* (2020), pp. 1–4. *Conference on Engineering and Emerging Technologies (ICEET)* (2018), pp. 1–6
33. Y. Rahayu, L. Fitria, Y. Hakiki, A. Kurniawan. A new 2×4 array design of dual-band millimeter-wave antenna for 5G applications, in *IEEE International Workshop on Antenna Technology (iWAT)* (2018), pp. 1–4
34. B. Aghoutane, N. Meskini, M. Elghzaoui, H. El Faylali, Millimeter-wave microstrip antenna array design for future 5 g cellular applications, in *IEEE International Conference on Electronics, Control, Optimization and Computer Science (ICECOCS)* (2018), pp. 1–4
35. E.L. Mashade, M. Bakry, E.A. Hegazy, Design and analysis of 28 GHz rectangular microstrip patch array antenna. *WSEAS Trans. Commun.* **17**, 1–9 (2018)
36. S. Pandey, L. Garia, Design of millimeter-wave spectrum microstrip patch antenna array for 5G wireless systems, in *IEEE Women Institute of Technology Conference on Electrical and Computer Engineering (WITCON ECE)* (2019), pp. 168–171
37. Md.F. Shah, A.D. Singh, Design and analysis of microstrip patch antenna arrays for millimeter wave wireless communication. *Int. J. Eng. Adv. Technol. (IJEAT)* **9**(3) (2020)
38. G.V. Raviteja, A 4-element corporate series feed millimeter-wave microstrip antenna array for 5G applications. *Int. J. Electron. Commun. Eng.* **13**(11), 674–684 (2019)
39. G.U. Kumar, V.S. Sanish, Millimeter wave circular microstrip patch antenna for 5G applications. *Int. J. Res. Sci. Innov. (IJRSI)* **VI**(V) (2019)

Kannada Text-to-Speech System using MATLAB



**K. Sanjana Kamath, K. Raghavendra N. Bhat, Charishma,
and Pearl Infancia D'souza**

1 Introduction

Kannada, being one among the Dravidian languages, is famous in one of the South Indian states, Karnataka [1]. In the present situation, there might be few people who are unable to read Kannada text due to various problems such as blindness, illiteracy, or immigration of people who are a novice to the language. In such situations, the text-to-speech synthesizer (TTS) system can be utilized to solve these problems.

The TTS system is designed especially for the Kannada language. There have been several efforts made in commercially developing TTS converters for various languages. However, very little effort has been put into developing a TTS converter for Kannada [2].

The TTS converter converts the linguistic texts to associated speech in the form of an audio. This system helps blind people to read Kannada documents without any difficulty. This paper explains the development of the TTS system using MATLAB. MATLAB has various in-built functions which advances the accuracy of the system.

The first step in the system is text recognition, which is carried out by optical character recognition (OCR) using the concept of discrete wavelet transform (DWT). OCR is used to extract the text from the required document or image by recognizing the character accurately. The paper explains different phases involved to build an

K. S. Kamath (✉) · K. R. N. Bhat · Charishma · P. I. D'souza
Department of Electronics and Communication, NMAM Institute of Technology, Nitte
(Visvesvaraya Technological University, Belagavi), Udupi, Karnataka, India
e-mail: ksanjanakamath@gmail.com

K. R. N. Bhat
e-mail: bhatraghav98@gmail.com

Charishma
e-mail: shettycharishma@nitte.edu.in

P. I. D'souza
e-mail: pearlidsouza@gmail.com

efficient as well as robust OCR system: pre-processing, feature extraction, and classification [3]. In the stage of pre-processing, the input image is converted into an appropriate processed image used for feature extraction. Feature extraction is the most significant stage in the OCR system. It extracts a suitable feature from the text as structural features. The classification phase deals with recognizing Kannada characters from the desired document. Once the character is recognized, it carries out TTS conversion where the recognized characters are read out loud.

The TTS conversion is carried out by using a few MATLAB functions to interface the recognized characters to any speech synthesis API. Therefore, Microsoft Win32 Speech API (SAPI) is used which is already integrated into MATLAB for good quality of audio.

The TTS system considered here was developed with the help of MATLAB functions which helps us to read basic Kannada characters that have decent voice quality from the desired document or image.

2 Prior Works

In today's world, there are people who are blind, illiterate or novice to the Kannada language. These people find it hard to read any document or text and might have to be dependent on someone else to read it. Hence, an approach was made by using previous related work to make an efficient and robust Kannada TTS system.

Text-to-speech systems have been developed for numerous languages, but only a few approaches have been made for the Kannada language. The separation of Kannada text into the smallest units of the word and concatenating sound segments extracted from pre-recorded speech to produce speech output [2]. It was developed using MATLAB 2010. First, segmenting is done where the entire document is separated into its consequent parts, i.e., lines, words, and letters. Second, concatenation where the segmented letters are concatenated back to form Akshara (syllable). The TTS system helps people to save time in this hectic world. It mainly helps visually weakened people to read the Kannada text and elderly people suffering from reduced sight.

The OCR system that is developed for handwritten Kannada characters comprises pre-processing, feature extraction, and classification [3]. In pre-processing, techniques that fit right to convert an input image to a suitable form are used for feature extraction. Pattern recognition and image processing lead to a great demand for the OCR system comprising handwritten characters. An OCR system is developed to recognize any given Kannada character from the document. Unique features are extracted from the characters as structural features and wavelet transform does its work of extracting global features. An artificial neural network classifier is used for accurate identification of the Kannada characters. This method is tested on 4800 images of handwritten Kannada characters and gained an average accuracy of 91.00%.

An attempt of using neural network approaches have been made for the OCR of the Kannada characters [4]. The fact that not many efforts have been made in developing such approaches for south Indian languages like Kannada has been emphasized. The process of carrying out OCR is classified into two. The first classification involves two neural classifiers, each for character classification as well as subscript classification. The second classification involves about 34 smaller neural classifiers for the consonant group of the Kannada language. These neural classifiers are denoted as group classifiers. The group classifier is carefully chosen by the first classifier and handles around 8 Kannada characters which result in recognizing the characters.

In [5], how to develop an accurate method to recognize characters has been described. The difficulty due to the similarities between the two characters can affect the OCR process. Therefore, two methods have been developed: template matching and feature extraction. In feature extraction, certain features are taken out in a specified manner in order to recognize the character using discrete wavelet transform (DWT). In template matching, a template of each character is created and matched with the feature extraction output.

Several processes included in OCR are explained in [6]. The procedure includes two steps mainly, training and testing which further break into sub-steps. Training includes steps like pre-processing, feature extraction, and model estimation, whereas testing includes steps like pre-processing, feature extraction, and classification. Further, the steps under pre-processing include binarization, morphological operators, and segmentation (the most important characteristic as it allows the system to extract features from every individual character). Also, a MATLAB-based graphical user interface (GUI) has been implemented in the training process of the model.

TTSBOX is a toolbox designed to explain TTS synthesis. It accomplishes the blend of Genglish (for “Generic English”), a non-existent language achieved by replacing English words with generic words. Genglish vocabulary is very minimal, and it contains no numerals, no acronyms, no intricate names, and no strange words. Every letter in the word is replaced by synthesized symbols. TTSBOX contains a Genglish synthesizer for segmentation. WAVESURFER tool is used to check the correctness of segmentation. Here, the simple concatenation is used, which takes out selected diaphones from the speech body and gathers it into synthetic speech [7].

TTS has been stated as the ability of the system to speak the recognized character out aloud. Developing the Kannada speech template is a task in itself as there is no readily available database for the Kannada language. Hence, a template is created by concatenating pre-existing speech for each character in the language. The speech application programming interface (SAPI) has been used for speech recognition. The SAPI used has been technologically advanced by Microsoft that is also used for speech synthesis [8].

Fig. 1 List of Kannada letters

ಅ ಆ ಇ ಈ ಉ ಊ ಋ ೠ ಎ ಏ ಐ ಒ ಓ ಔ ಂ ಃ

(a): Swaragalu and Yogavaahakagalu

ಕ ಖ ಗ ಘ ಙ

ಚ ಛ ಜ ಝ ಞ

ಟ ಠ ಡ ಢ ಴

ತ ಥ ದ ಧ ನ

ಪ ಫ ಬ ಭ ಮ

(b): Vargeeya Vyanjanagalu

ಯ ರ ಲ ವ ಶ ಷ ಸ ಹ ಳ

(c): Avargeeya Vyanjanagalu

3 Kannada Language

The Kannada language has been evolved from the Kadamba script (fifth century). Currently, the language is written using the Kannada script. Kannada uses a non-cursive script which is written horizontally that is left-to-right form [2]. The term non-cursive indicates that words in Kannada do not have joined characters, but a small gap between the characters of each word. There is no concept of case sensitivity in Kannada as there are not any upper-case and lower-case present.

The language has 49 letters which are divided into three parts namely: *swaragalu*, *vyanjanagalu*, and *yogavaahakagalu*. *Swaragalu* consists of 13 letters which can be assumed to be vowels. *Vyanjanagalu* consists of 34 letters which can be assumed to be consonants. *Yogavaahakagalu* is of two types: *anusvara* ಂಃ and *visarga* ಃಃ [9]. *Vyanjanagalu* is further classified as *Vargeeya Vyanjana* and *Avargeeya Vyanjana*. These *Vyanjanagalu* are reliant on vowels.

The list of Kannada letters are shown in Fig. 1. Figure 1a consists of Swaragalu and Yogavaahakagalu. Figure 1b consists of Vargeeya Vyanjanagalu and Fig. 1c consists of Avargeeya Vyanjanagalu [2].

On combining the Vyanjanagalu with Swaragalu, form *Kagunitha* for Kannada alphabet can be formed. Every symbol that is written in Kannada produces a syllable which is known as *Akshara*, whereas every symbol in English produces a phoneme.

4 Optical Character Recognition

Optical character recognition is an electronic translation of typed, handwritten or printed characters into machine encoded characters [10]. This could either be from a document that is scanned, photograph of the document, etc. The aim of OCR is

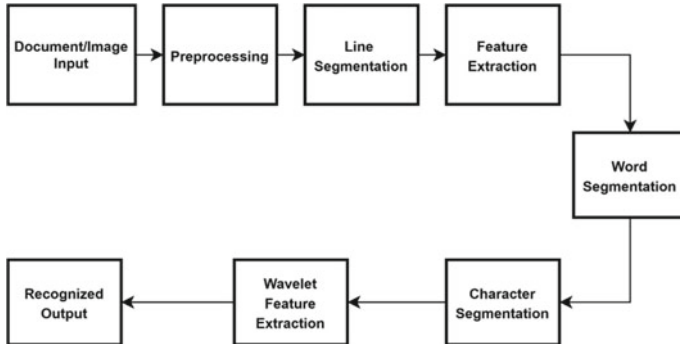


Fig. 2 Block diagram of OCR system

to categorize optical outlines that repeatedly occur in a digital image equivalent to alphanumeric or any other characters [6]. Computer systems stimulated with such OCR systems can advance the speed of operation and reduce the possible human errors. Identification of the printed characters is challenging since there may be an alternative of the same character depending on the person, the resolution of the scanner, or different kinds of noises. Also, the difference in the font sizes makes the recognition task difficult. Hence, methods like pre-processing, feature extraction, and recognition which work besides the OCR system to make it more robust are required to be used [3].

The block diagram of OCR system is shown in Fig. 2 which shows various methods involved in recognizing the character which is efficient and robust:

A. Pre-processing

Pre-processing is a very important step that alters an input image into a form which is suitable for improved feature extraction. The input image is exposed to numerous pre-processing stages such as binarization, removal of noise, normalization, and thinning. The resulting image from the process of scanning may contain some noise. Depending on the resolution of the camera used to scan, the characters may be broken or smeared. Hence, pre-processing can be done to smooth the digitized characters. In addition to smoothing, pre-processing also involves normalization of the characters.

B. Line or Word Segmentation

Segmentation involves the isolation of characters and words. First, the input image is segmented line by line to check for characters, once it recognizes some character it goes on for word by word segmentation, then the last part includes a letter by letter segmentation. In the end, only letters are extracted from the image by ignoring the background. For the ease of segmentation, the input image is standardized to a resolution of 42×24 pixels.

C. Feature Extraction

Feature extraction is used for capturing the characteristics of the recognized symbols, and it is achieved by matching the matrix containing the input characters with a set of model characters that represent a class. This can be carried out by extraction of two types of features, namely structural features and global transformation features. The topological as well as geometrical properties of the character are a part of structural features. However, the global transformation feature used here is the wavelet transformation.

D. Wavelet Transform

Wavelet transformation is used in extracting the global features. Wavelet transforms are mathematical functions that decompose data into a set of basic functions called wavelets. Wavelets, by dilations and shifting, are achieved from the original prototype called the mother wavelet. Wavelet transform decodes the time-amplitude signal into a time-frequency signal that is compressed as wavelet coefficients. In general, the wavelet transform is computed using the Eq. (1) and its inverse is shown in Eq. (2):

$$f(a, b) = \int_{-\infty}^{\infty} f(x) \Psi_{(a,b)}^*(x) dx \quad (1)$$

$$f(x) = \iint f(a, b) \Psi_{(a,b)} db da \quad (2)$$

Here, $f(a, b)$ is a function, the scaling and shifting parameters are a and b , respectively. $*$ is the symbol used to represent complex conjugate and Ψ is a function that is chosen subjectively. Wavelet transform is separated into continuous wavelet transform (CWT) and discrete wavelet transform (DWT). Wavelets that are not orthogonal in nature are used for CWT, whereas, in DWT, the wavelets used are orthogonal. In the proposed system, DWT is utilized as it decomposes into a set of wavelets that are orthogonal to translation and scaling. However, DWT also tolerates good localization in both time as well as spatial frequency domain. DWT even allows good identification of data that is applicable to human perception and is easy to implement which in turn shrinks the complexity and the computation time required. The equations used for high-pass wavelet coefficient (Hi) and low-pass wavelet coefficient (Li) are shown in Eqs. (3) and (4), respectively:

$$H_i = \sum_{m=0}^{k-1} x_{2i-m} \cdot s_m(z) \quad (3)$$

$$L_i = \sum_{m=0}^{k-1} x_{2i-m} \cdot t_m(z) \quad (4)$$

Here, wavelet filters are assumed to be $s_m(z)$ and $t_m(z)$ and the length of the filter is 'k'.

5 Text-To-Speech System

Text-to-speech synthesis is a complex combination of various processes like signal processing, language processing, and computer science that can be implemented using the software. One can create synthesized speech by concatenating fragments of recorded speech that are stored in a database [2]. The block diagram of the text-to-speech system is shown in Fig. 3.

A text-to-speech engine is made up of two parts: a front-end and a back-end. The front-end consists of two major tasks. In the first task, a process called text normalization takes place where raw text that consists of symbols like numbers and abbreviations are converted into the corresponding words that are written out. Phonetic transcriptions are assigned to each word in the front-end process; the text is then divided and symbolized into prosody units, like phrases, clauses, and sentences. This process where phonetic transcriptions are allotted to words is known as text-to-phoneme conversion. Phonetic transcriptions and information form the symbolic linguistic representation that is the output given by the front-end. The back-end alters the symbolic linguistic representation into an audio format that is accomplished using a synthesizer [11].

The tool used for TTS synthesis is called as 'TTSBOX'. TTSBOX carries out the synthesis of Genglish (Generic English). Genglish is assumed to have no abbreviations. Although the vocabulary usage of Genglish is simpler than English, it sustains a lot of its phonetic and prosody complexity. This feature makes it supportive in converting it to other languages (in this case, Kannada) by assigning a word which acts like a phoneme of a Kannada character which has been carried out to every basic character of Kannada. The TTS synthesizes speech from TXT string and speaks it as output. The format of the audio used is mono, 16 bit, and 16 k Hz by default. There is a function to select any specific voices from the list of available voices. The pace of speech can be changed between the ranges -10 (slowest) to 10 (highest). All these functions require a speech application program interface (API). In this case, Microsoft Win32 Speech API (SAPI) is used which is already integrated into MATLAB.

Fig. 3 Block diagram of text-to-speech system



A template is created by concatenation of previously existing speech for each character of the Kannada language. This template is correlated with the recognized output and the template with the maximum correlation is given to the speech API which is Microsoft Win32 SAPI. This SAPI helps the system to read the recognized character out aloud.

6 Results

An image of Kannada text is given as an input to the TTS system which will first carry out the OCR, where the binarization, removal of noise, and conversion of image to gray takes place after which lines and words are segmented. Finally, the feature extraction is carried out and the character is recognized using template matching. This recognized output is correlated with the speech template and is given to the SAPI which helps in reading out the recognized character.

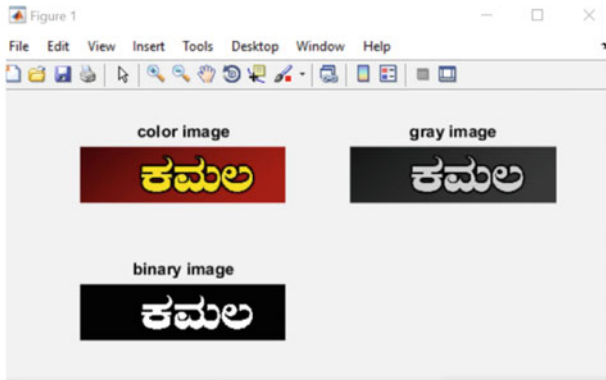
Figure 4a shows the pre-processing output carried out using MATLAB for the word 'Kamala'. Figure 4b shows the line and word segmentation output for the same word. Figure 4c shows the wavelet transform output used to recognize each character which in this case is 'Ka', 'Ma', and 'La'.

The recognized image is then correlated with the template where each character is assigned with a particular phoneme. The one with the maximum correlation is selected and given to the Microsoft Win32 SAPI and the system reads out the recognized characters, which in this case is 'Kamala'.

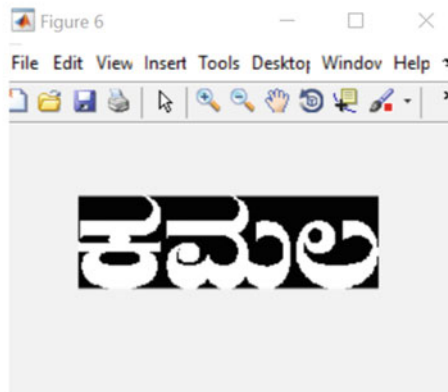
7 Conclusion and Future Scope

Shortly, speech recognition is most likely to become a vital element in communication among humans and machines. The effectiveness of the TTS system is based on how exactly segmentation units are formed and after synthesizing different words how meticulously the concatenation joint distortion is reduced. The different segmentation techniques are implemented for syllable formation. The general processing, phoneme recognition, and text recognition are the three phases of speech recognition carried out in MATLAB. The speaker-independent speech recognition systems were efficaciously trained to identify speech inputs that were interfaced with Microsoft Win32 Speech API with the help of MATLAB and speak it out as output.

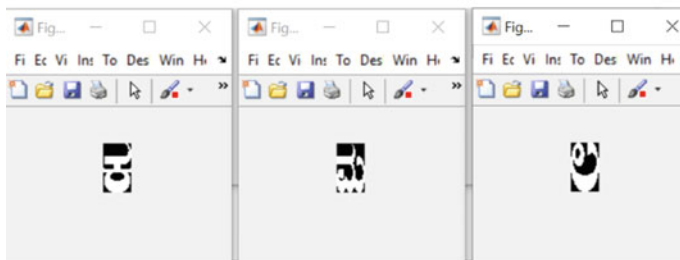
Crucial research challenges such as robustness of the speech, various critical pronunciations, an extension of vocabulary, and precision in techniques used for extracting abstracts from characters may have to be faced, especially for languages like Kannada. In languages like Kannada, "Ottakshara" plays a major role to represent complicated words. Hence, the improvisation of the system by reading out the words containing ottakshara would lead to increase the efficiency of the system. The acquired consequences can be enhanced by adjusting the system with larger



(a): Pre-processing output for the word 'Kamala'



(b): Line and word segmentation output for the word 'Kamala'



(c): Wavelet transform output recognising the characters 'Ka', 'Ma', and 'La'

Fig. 4 Output of the OCR part

databases. Recognizing live speech would be difficult in the following stage, which would demand additional resources, bigger databases, and wide vocabularies to yield accurate character recognition outcomes. The application field for speech synthesis is becoming broader all the time which brings more reserves into research and development areas. Combining text-to-speech system of different languages and making a single reader will give the user the choice to read different languages without being dependent on someone else, yielding a marketable product.

References

1. M.S. Dhananjaya, B. Niranjana Krupa, R. Sushma, Kannada text to speech conversion: a novel approach, in *International Conference on Electrical, Electronics, Communication, Computer and Optimization Techniques (ICEECCOT)* (IEEE, 2016), pp. 168–172
2. A. Joshi, D. Chabbi, M. Suman, S. Kulkarni, Text to speech system for Kannada language, in *International Conference on Communications and Signal Processing (ICCSPP)* (IEEE, 2015), pp. 1901–1904
3. S. Pasha, M.C. Padma, Handwritten Kannada character recognition using wavelet transform and structural features, in *International Conference on Emerging Research in Electronics, Computer Science and Technology (ICERECT)* (IEEE, 2015), pp. 346–351
4. K. Kalaivani, R. Praveena, V. Anjalipriya, R. Srimeena, Real time implementation of image recognition and text to speech conversion. *Int. J. Adv. Res. Technol* **2**, 171–175 (2014)
5. B. Dash, S. Pradhan, D. Rana, Odia offline character recognition using DWT features. *IOSR J. Electron. Commun. Eng.* 31–37 (2016)
6. J. Hansen, *A Matlab Project in Optical Character Recognition (OCR)*, vol. 6. (DSP Lab, University of Rhode Island, 2002)
7. T. Dutoit, M. Cernak, TTSBOX: a MATLAB toolbox for teaching text-to-speech synthesis, in *Proceedings. (ICASSP'05). IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 5 (IEEE, 2005), pp. v–537
8. R.S. Kunte, R.D.S. Samuel, An OCR system for printed Kannada text using two-stage multi-network classification approach employing wavelet features, in *International Conference on Computational Intelligence and Multimedia Applications (ICCIMA 2007)*, vol. 2 (IEEE, 2007), pp. 349–353
9. C. Naveena, V.N. Manjunath Aradhya, Handwritten character segmentation for kannada scripts, in *World Congress on Information and Communication Technologies* (IEEE, 2012), pp. 144–149
10. A.E. Manjunath, B. Sharath, Implementing Kannada optical character recognition on the android operating system for Kannada sign boards. *Int. J. Adv. Res. Comput. Commun. Eng. (IJARCCE)* **2**(1) (2013)
11. P. Malik, R. Dixit, Handwritten character recognition using wavelet transform and Hopfield network, in *International Conference on Machine Intelligence and Research Advancement* (IEEE, 2013), pp. 125–129

Design of Dynamic Induction Charging Vehicle for Glimpse of Future: Cutting Down the Need for High-Capacity Batteries and Charging Stations



K. Balakrishna and N. G. Sandesh

1 Introduction

The decarbonation of transportation relies largely on electromobility for the present and future decade in India. Future markets predict that electrical vehicles will reach a dominant share compared to present running vehicles on road by 2030, taking into consideration of road transportation modes [1]. Even though electric vehicle technology is rapidly developing, but charging the battery is a challenging task to adopt this technology for a routine lifecycle. The recent electric vehicle technology allows it to get fully charged within 30 min from the previous 8 h through a cable, pantograph or wireless [2].

The advent of charging enables an array of opportunities and markets. Which will advantageous from autonomous taxis to cargo transport without halt, innovating transport methods are possible. Wireless charging is distinguished into two types such as stationary charging and dynamic charging. In stationary charging, electrical vehicles are charged wirelessly through charging pads when they are parked similar to mobile devices [3]. But dynamic charging electrical vehicles specify the technology that can simultaneously supply power and perform charging, while vehicle is driving or in motion.

Presently, investments are being sought to develop dynamic wireless charging from many companies research and development wing to improve its usage and efficiency to reduce carbon dioxide (CO₂) emission. Dynamic remote charging of electric vehicles [4] could turn into a favored technique since it would empower power trade between the vehicle and the grid while the vehicle is moving pervasively. Mobile energy disseminators (MED) are a new concept [4] that can facilitate EVs to extend

K. Balakrishna (✉) · N. G. Sandesh
Department of ECE, Maharaja Institute of Technology Mysore, Belavadi, Karnataka, India
e-mail: balakrishnak_ece@mitmysore.in

N. G. Sandesh
e-mail: sandeshhg_ece@mitmysore.in

their range in a typical urban scenario. Our proposed method exploits inter-vehicle (IVC) communications in order to eco-route electric vehicles taking advantage of the existence of MEDs. Combining modern communications between vehicles and state of the art technologies on energy transfer, vehicles can broaden their movement time without the requirement for huge batteries or incredibly expensive framework [5].

But in a typical electric car, it needs to be charged only in the charging stations or else in the parking lots but in our proposed system the car will be charged within the vehicle when the vehicle is in the motion this will change the problem of the vehicle because there will be no problem of recharging the vehicles in the parking lots or in the charging stations and this is the major advantage of our proposed system.

2 Literature Survey

The literature survey is carried out as a part of the research work. It has provided review of the past research about the electric vehicles kind of a battery they used in it and the way they used to charge the vehicle. The previous exploration exertion will appropriately manual for legitimize the degree and bearing of the current exertion.

Mazlan et al. [6] carried an experimental study on the effect of alternator speed to the car system. In this context, an attempt has been made to charge the battery as fast as possible. The system that they have developed is capable of charging a 12 V, 42 Ah battery within 2 h, if the speed of the alternator rotor is maintained around 2000 rpm. To achieve this they have used different pulley ratios. The pulley ratio that they have maintained to charge the battery within two hours is 1:6 ratios and it produced a desired result.

Koumartzakis et al. [7] worked on designing and developing of a prototype electric vehicle chassis. The paper describes about the designing of the chassis to build a prototype of an electric vehicle and the mechanisms of breaking system, steering system as well as the drive shaft. The paper describes about the exact measurement of the chassis that they have designed and the steering system working and the details of the transmission system.

Udaeta et al. [8] proposed a model for energy efficient looking into the electric mobility in electric vehicles analysis. This paper describes about the evolution of electric vehicle has taken place where the power consumption is dramatically decreased and the performance of the vehicle is increased and higher energy densities battery are required with electric vehicles, which restricts them due to the volumes and the masses associated with it and also another limitation of it is that the least availability of the charging stations.

Suhas et al. [9] designed an electric vehicle battery with self-charging capacity for its own propulsion and compared its performance. The proposed work manages the manufacture and testing of a battery electric vehicle with a self-charging framework. An endeavor has been made to create a self-charging battery electric vehicle which uses the rotational energy of wheels to charge the batteries, in this way presenting a

framework which makes the vehicle contamination free. After the drive they tested the distance that can be achieved so that the capacity of the battery can be found.

The main objectives of the project are

1. To reduce the pollution caused by ICs and to make the environment pollution free.
2. To make the car using the proposed method, so that no need for the car to charge in the parking lot or in the charging stations.
3. A low cost and highly effective method of producing the energy almost at the rate of it is depleting and to make the vehicle moving without problems of energy depletion.

3 Proposed Work

The system mainly includes electric motors, boost converter, battery packs, smart switching system, alternators and microcontroller as shown in Fig. 1. The motor gives the mechanical movement required to move the car. The battery packs act as a power source to the motor. Boost converter is used to boost the voltage from 24 to 48 V [10]. The alternators are used to charge the batteries while the vehicle is in motion. Smart switching system is used to switch the battery packs to the motor and alternators based on the voltage levels of the batteries and the controller is used to check the voltage levels and excite the relay. The hardware components are the major

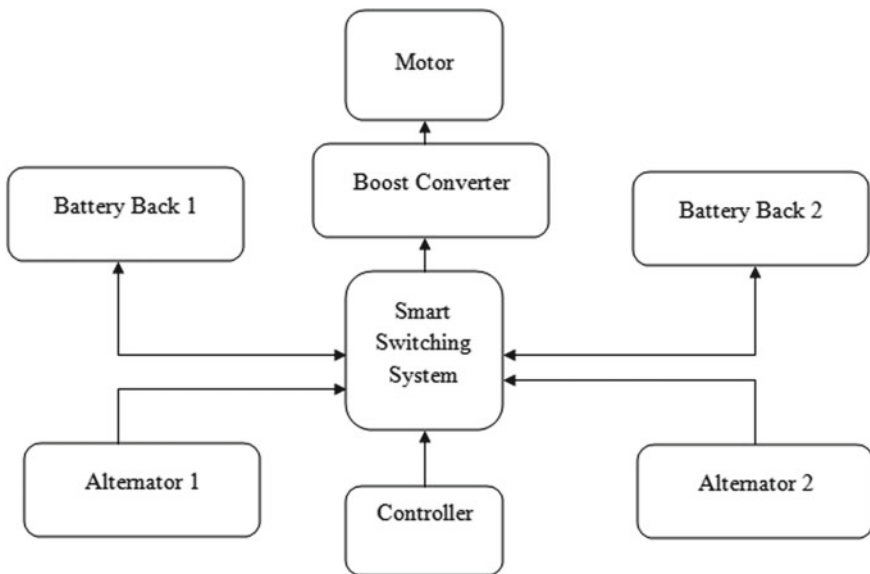


Fig. 1 Block diagram of self-charging electric vehicles

contributor for the proposed model when it is compared with the software. Hardware components used are Arduino Nano, brushless DC motor, speed controller, alternator, batteries and relays and software used are Arduino IDE 1.8.5.

4 Experimentation

As shown in Fig. 2, alternator acts as a power source for the electrical vehicles. The alternator is another form of mechanical energy created by the engine through combustion of fossil fuel; here, we have used the battery to rotate the motor to rotate the alternator to generate electricity. The induction coil transfers the electric energy from one coil to another coil, i.e., wound-up coil through an electromagnetic field [11]. Every time, vehicles in motion receive electricity for a fraction of time and it will continue while in motion due to the rotor and stator windings. The synchronous Table 1 shows the relation between frequencies and rotation per minute (RPM) for selected poles.

Battery contains an electrochemical cells with external connection to motor through a relays, and it converts the mechanical energy into chemical energy [12]. The voltaic cell depends on a unique piezoelectric film that moves lithium particles

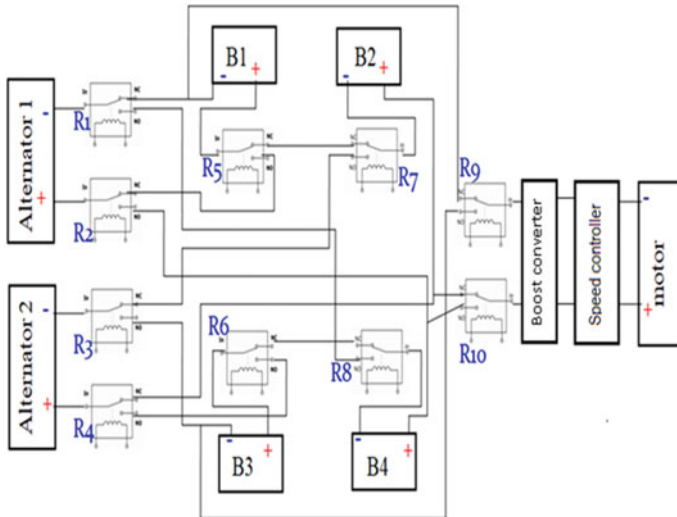


Fig. 2 Circuit diagram of smart switching circuit

Table 1 Synchronous speed

Poles	2	4	6	8	10	12
RPM for 50 Hz	3000	1500	1000	750	600	500

from one side of the cell to the next when the layer is exposed to mechanical power. By driving the lithium particles through the spellbound layer utilizing the piezoelectric potential, the vehicle ready to store synthetic energy straightforwardly utilizing an electrochemical cycle. The relay is connected for the isolation switching function for with standing high power density and circuit protection [13].

The DC–DC boost converter is necessary to interface with fuel cell and battery, and it stores the DC input temporarily either magnetic field (inductors or transformers) or electric field (capacitors) components. Here, to achieve the required voltage regulation, two control loops are used as shown in Fig. 3. The control mode operates on the knowledge of inductor current and using RST controllers. The electric vehicle designed here works with a DC voltage of varying between 96 and 190 V [14].

The construction and working of DC motor are shown in Fig. 4. The DC bias of 12 V is given to DC motor, which is limited by resistor and fed to timer IC; here, diode

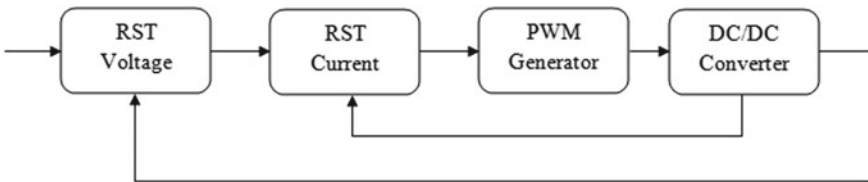


Fig. 3 Control mode operation for DC–DC converter

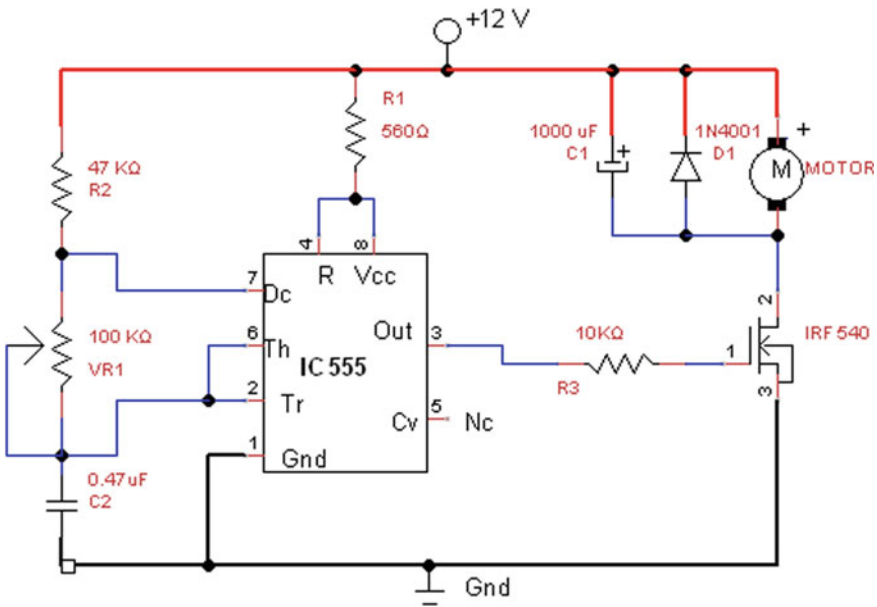


Fig. 4 DC motor speed controller

protects the motor from back electromotive force [15]. The timer IC produces a timing pulse at pin 3 relies upon VR1 and C2 components, and this timing pulse controls the MOSFET consequently the engine associated through MOSFET is controlled, and the beat yield from IC 555 has the control of DC engine speed, by changing the pulse span we can control the speed of DC engine [16].

Brushless DC motor has permanent magnets on the outer side called stator and spinning armature inside called rotator. A brushless DC engine is basically flipped back to front, taking out the requirement for brushes to flip the electromagnetic field [17, 18]. The efficiency of the motor is about 90% with 360° rotation of motor.

5 Result

The self-charging electric car should reduce the use of fuel vehicles so that which makes the vehicle pollution less and environment pollution free. Self-charging electric vehicle is designed and developed and it is in the working condition capable of carrying or taking a load up to 500 kg by the motor. The sufficient amount of energy has been generated using the alternator and it is redirected to the battery. The smart switching action is also performing well by switching when the voltage level of the one battery pack is depleted and it has been reached to the threshold level so at that time it has been performing the switching action. The snapshot of the working model is shown in Fig. 5.



Fig. 5 Snapshot of the working model

6 Conclusion

Electric vehicle should make the environment pollution free because it does not release any toxic contents and it does not require fuels because it is using the motor to run which requires electricity and that power is supplied through the battery. It significantly reduces the use of fuel vehicles and it is pollution-free journey. The smart switching system which has been installed enables the user to carry out the journey uninterrupted. Because of its capability of self-charging, it does not require to charge in the parking lots, and in the charging station, it will charge while it is in the movement.

Acknowledgements We would like to show our gratitude to the Maharaja Institute of Technology Mysore and we thank teaching and non-teaching staff of department of electronics and communication engineering. Also, thanks to our parents and friends who all are directly or indirectly supported for this research.

References

1. K.Ç. Bayindir, M.A. Gözükküçük, A. Teke, A comprehensive overview of hybrid electric vehicle: powertrain configurations, powertrain control techniques and electronic control units. *Energy Convers. Manage.* **52**(2), 1305–1313 (2011)
2. S.Y. Choi, B.W. Gu, S.Y. Jeong, C.T. Rim, Advances in wireless power transfer systems for roadway-powered electric vehicles. *IEEE J. Emerg. Sel. Top. Power Electron.* **3**(1), 18–36 (2015). <https://doi.org/10.1109/JESTPE.2014.2343674>
3. K. Balakrishna, WSN-based information dissemination for optimizing irrigation through prescriptive farming. *IJAEIS* **11**(4), 41–54 (2020). <https://doi.org/10.4018/ijaeis.2020100103>
4. S. Moschoyiannis et al., Dynamic wireless charging of electric vehicles on the move with mobile energy disseminators. *Int. J. Adv. Comput. Sci. Appl. (IJACSA)* **6**(6), 239–251 (2015)
5. Y.M. Nie, M. Ghamami, A corridor-centric approach to planning electric vehicle charging infrastructure. *Transp. Res. Part B: Methodol.* **57**, 172–190 (2013)
6. R.K. Mazlan et al., Experimental study on the effect of alternator speed to the car charging system, in *MATEC Web of Conferences*, vol. 90 (EDP Sciences, 2017)
7. G. Koumartzakis, P. Spanoudakis, N.C. Tsourveloudis, Design and development of a prototype electric vehicle's chassis, in *7th BETA CAE International Conference* (2017)
8. H.F. Camilo et al., Assessment of photovoltaic distributed generation—issues of grid connected systems through the consumer side applied to a case study of Brazil. *Renew. Sustain. Energy Rev.* **71**, 712–719 (2017)
9. V. Suhas et al., Performance of a battery electric vehicle with self charging capacity for its own propulsion. *Int. Res. J. Eng. Technol. (IRJET)* **2**(03) (2015)
10. N. Kumar et al., Physical design and modeling of 24v/48v dc–dc boost converter for solar PV application by using SIMSCAPE library in MATLAB. *Int. J. Appl. Control Electr. Electron. Eng. (IJACEEE)* **2**(2) (Wireilla Publication, Australia, 2014)
11. R.J. Forsyth et al., The underlying Parker spiral structure in the Ulysses magnetic field observations, 1990–1994. *J. Geophys. Res. Space Phys.* **101**(A1), 395–403 (1996)
12. H. Ganter et al., Relay support device for an electric motor, in particular for an electrically commutated DC motor. U.S. Patent No. 6,873,072, 29 Mar 2005
13. K. Balakrishna, M. Rao, Tomato plant leaves disease classification using KNN and PNN. *IJCVIP* **9**(1), 51–63 (2019). <https://doi.org/10.4018/ijcvip.2019010104>

14. L. Maheswari et al., A unique control strategy to improve the life cycle of the battery and to reduce the thermal runaway for electric vehicle applications. *J. Therm. Anal. Calorim.* 1–13 (2020)
15. J. Shao, An improved microcontroller-based sensorless brushless DC (BLDC) motor drive for automotive applications. *IEEE Trans. Ind. Appl.* **42**(5), 1216–1221 (2006)
16. A. Yasuo, T. Tamotu, Variable pulse system for controlling dc motor speed by variation of supplied current. U.S. Patent No. 3,409,814, 5 Nov 1968
17. K. Balakrishna, Fusion approach-based horticulture plant diseases identification using image processing, in *Applications of Advanced Machine Intelligence in Computer Vision and Object Recognition: Emerging Research and Opportunities*, eds. by S. Chakraborty, K. Mali (IGI Global, Hershey, PA, 2020), pp. 119–132. <https://doi.org/10.4018/978-1-7998-2736-8.ch005>
18. P. Sreekala, A. Sivasubramanian, Speed control of brushless DC motor with PI and fuzzy logic controller using resonantpole inverter, in *ISGT2011-India* (IEEE, 2011)

Generation of ECG for Heart Block Cases



Venkatesh Nayak

1 Introduction

An electrocardiogram, short for ECG or EKG, is a tool used to measure the electrical activities of the heart. This is measured by attaching a pair of electrodes on the surface of the chest and on various other body parts, for accuracy. Small differences in the potential as a consequence of cardiac muscle depolarization (contraction) and immediate re-polarization (relaxation) constitutes for one complete heartbeat. Every such heartbeat is measured and plotted on the graph.

Real-time ECG signals are contaminated with various noises and artefacts, resulting in a signal which would pose difficulty in analyzing the output. Hence, there is a requirement for a signal to be processed, in order to interpret an accurate result from the graph which enables to infer required diagnostic information. Albeit, to achieve such outcomes, a reliable signal processing is required, which consumes time and ample efforts to come to a viable solution. Therefore, there is a need for standardized reference. Additionally, accuracy is of not a major concern, if it comes down to analyzing the data for academic studies. With this notion, various mathematical models have been developed in the past. One of such most widely used models is Dynamical model based on three coupled ordinary differential equations [1]. Similarly, generating

Artificial ECG based on geometrical features of a real ECG
Signal [2, 3]. And creation of a synthetic ECG signal
Model based on the Fourier analysis [4].

V. Nayak (✉)

Department of Electrical and Electronics Engineering, NMAM Institute of Technology, Nitte
(Visvesvaraya Technological University, Belagavi), Udupi, India
e-mail: venkatesh17nayak@gmail.com

Whereas, this paper throws a light on generating artificial ECG signals using geometrical features and Fourier series approximation, with the goal to simulate several heart blocks. Furthermore, this study is one half of the entire project undertaken to generate and detect various abnormal heartbeats.

This paper is organized as follows: Sect. 2 gives an insight into the typical ECG heartbeat characteristic, Sect. 3 briefs the mathematical model with the flowchart. Section 4 presents the results obtained using the proposed model and validation, and Sect. 5 concludes the paper.

2 Features of ECG

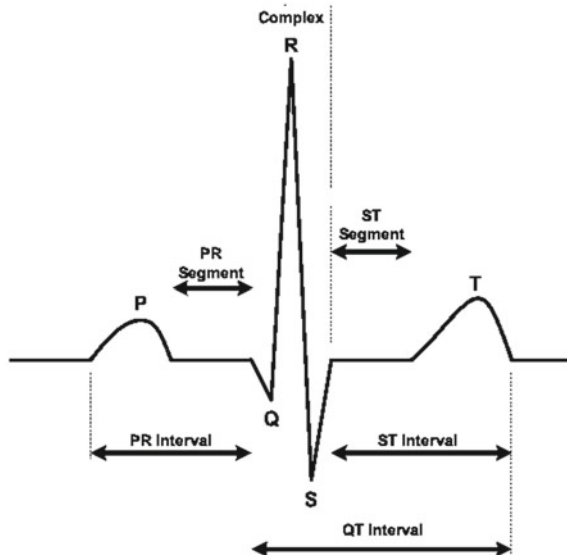
An ECG signal is a continuous response of each heartbeat. Each beat consists of a series of several deflection of the current, away from the baseline or isoelectric line. A typical heartbeat consists of five waves, namely P, Q, R, S, T waves (Fig. 1).

The ECG interpretation starts with a depolarization of the left and right atrium which is responsible for the current deflecting above the baseline accounting for the P wave [5].

The QRS complex are a result of depolarization of the ventricles. Q, R and S waves together constitute for QRS complex. The negative deflection corresponds to the Q wave and the next instantaneous increase in the deflection corresponds to the R wave and the consecutive deflection results in an S wave.

The distance between P and the QRS complex is the PR segment [5, 6].

Fig. 1 PQRST waveforms and the intervals of an ECG signal



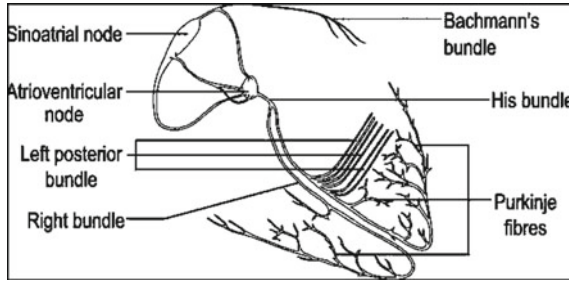


Fig. 2 Conduction system of the heart [7]. Adapted from Wikimedia Commons by Madhero88 (original files); Conduction system of the heart without the Heart.png CC BY-SA 3.0, <https://upload.wikimedia.org/wikipedia/commons/a/ae/ConductionsystemoftheheartwithouttheHeart-en.svg>

QRS complex is followed by a ST segment, similar to PR segment. ST segment is the distance between QRS complex and the initiation of the T wave. The T wave is a result of the rapid deflection of the current due to the ventricular re-polarization. Unlike P wave, T wave is slightly asymmetric and with a steeper downward slope [5].

This constitutes an entire heartbeat. Such beats recorded for t time interval are plotted using the electrocardiogram. The above discussed parameters must be present in order to call it a normal heart beat. Typically, the resting heart rate is between 60 and 100 BPM (beats per minute). If the heart rate is above or below than the mentioned BPM, implies that heart condition is abnormal, conditions such as Bradycardia or Tachycardia, respectively [5, 6] (Fig. 2).

The initiation of the electrical signal, or the P wave, here, occurs in the Sinoatrial node or Sinus node or SA node. Subsequently followed by the Atrioventricular node or AV node. This node contains a secondary pacemaker, at certain time acts as the backup for the SA node in case it fails to initiate. The signal then conducts into the Bundle of His and through the Bundle Branches, right and left. Therefore, all the blockages are characterized and classified based on the location of the blockage.

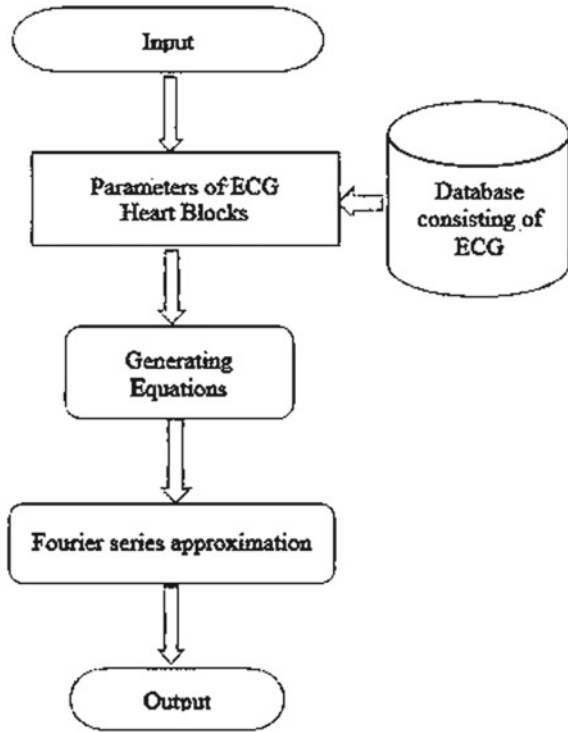
3 Methodology

The proposed model uses geometric features of each of the major waves as discussed in the previous section and approximates each waveform using the Fourier series approximation. Therefore resulting in a signal which is free of noise and artefacts but rather an ideal waveform with a good approximation and consuming relatively less computation time (Fig. 3).

A. Flowchart

Therefore, the necessary parameters required to generate an ECG signal are the intervals of the P Q R S T waves and PR interval and QT interval.

Fig. 3 Flowchart of the process



B. *Mathematical Model*

Having analysed the ECG pattern from the previous section, the heartbeat initiates with an isoelectric line for a short period of time which hovers around the zero line, also known as baseline. Hence taken zero.

- Modelling P and T waves

Although T is slightly asymmetric, the P and T waves can be approximated to the same mathematical function. It can be simply defined by a parabolic function. A typical parabolic equation is given by,

$$f(x) = -x^2 [4]$$

This conic can be precisely defined as follows

$$f(x) = -a(x - t)^2 + A \tag{1}$$

Here, a defines the width of the section. t represents the position and A is the amplitude. For a given amplitude and interval(s), the P and T wave can be represented as,

$$f(x) = ai(x - ti)^2 + A \quad (i = p \text{ or } t) \tag{2}$$

And ai is defined as

$$ai = \frac{Ai}{\left(\frac{ti}{2}\right)^2} \tag{3}$$

The equations generated using geometrical features are further decomposed into sums of sines and cosines using the Fourier series and are approximated [4, 8].

- Modelling QRS Complex

QRS consists of three major waves, they are Q, R and S waves. These three waves can be particularly represented using a triangular function.

A rising or falling edge can be defined as follows,

$$f(x) = mx + b.$$

where m is the slope of the line and b represents the intercept of the wave.

A triangular function can be deduced from the intersection of raising and the falling edge slopes. Generally, it can be defined as,

$$f(x) = \begin{cases} Mx + B, & x < 0 \\ -Mx + B, & x \geq 0 \end{cases} \tag{4}$$

Here, slope M and intercept B can be more precisely defined as follows,

$$M = 2 \times Ai/ti \tag{5}$$

$$B = Ai \times ti \quad (i = p \text{ or } t) \tag{6}$$

Ai is amplitude and ti is the interval for the Q, R and S waves individually.

The QRS complex computed using the above equations is approximated using the Fourier series [4, 8] (Figs. 4 and 5).

Fig. 4 Fourier approximation of P or T wave

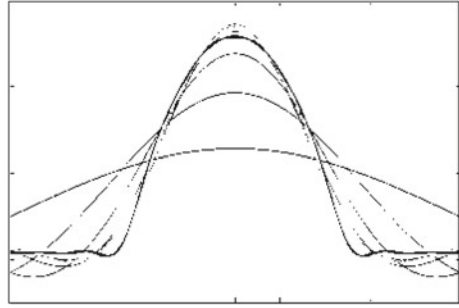
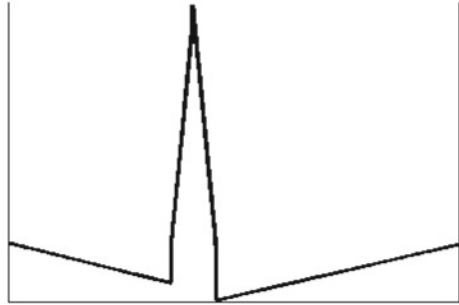


Fig. 5 Fourier approximation of QRS



C. Fourier Series

The ECG signals are assumed to be periodic, the Fourier series is used to represent the function in terms of its fundamental sums of Sines and Cosines.

The coefficient of Fourier series is given by,

$$A_0 = \frac{1}{L} \int_{-L}^L f(x) dx \tag{7}$$

$$A_n = \frac{1}{L} \int_{-L}^L f(x) \cos\left(\frac{n\pi x}{L}\right) dx \tag{8}$$

$$B_n = \frac{1}{L} \int_{-L}^L f(x) \sin\left(\frac{n\pi x}{L}\right) dx \tag{9}$$

$$f(x = i) = \frac{A_0}{2} + \sum_{n=0}^n A_n \times \cos\left(\frac{n\pi x}{L}\right) + \sum_{n=0}^n B_n \sin\left(\frac{n\pi x}{L}\right) \tag{10}$$

i = p, q, r, s, t waves *L* is the period of the function.

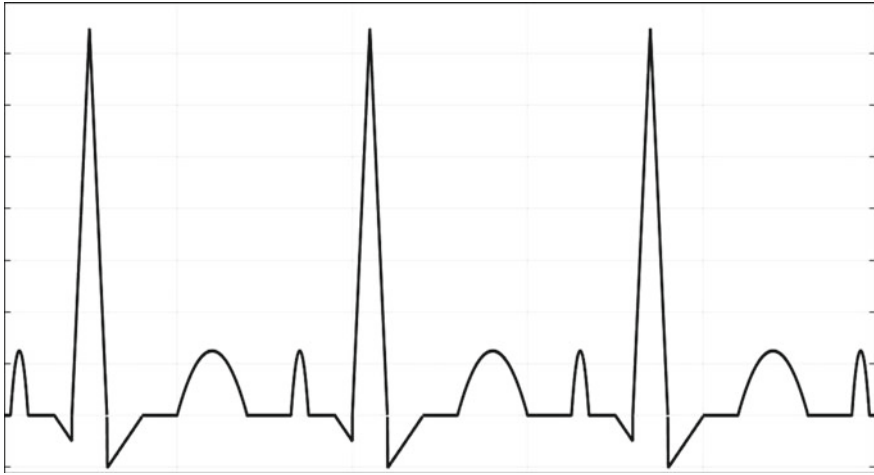


Fig. 6 Fourier approximation of PQRST waves, ECG signal output after the fourier analysis using random parameters

Since the generation is only considered with the real components, therefore only real component of the Fourier series is denoted (Fig. 6).

4 Result

This section describes and validates the generated signal using the proposed model. Using the proposed model, various heart blocks are simulated. The simulation of the various heart arrhythmias described below were carried out using the MATLAB software. The heart blocks taken in consideration are Sinoatrial Blocks (SA), Atrioventricular Blocks (AV) and the Bundle Branch Blocks (L/R BB).

The resulting output obtained is being validated with the real-time ECG strip. The resulting output obtained is a very close approximate to that of a real ECG signal.

The section below describes the signal in consideration and the important features considered while modelling the ECG signal.

A. Sinoatrial Node (SA)

- SA Block

SA block or the Sinus block conducts the P waves normally until a P wave is missed and a complete electric cardiac cycle of heartbeat fails to conduct itself [9].

Due to the absence of the P wave, it leads to the absence or the missing of the atrial depolarization (contraction) which, therefore, leads to a dropped beat; however, a moment later is followed with a normal sinus rhythm.

Features considered while modelling were,

The absence of the P wave, resulting in one complete dropped beat. Followed by a normal sinus rhythm (Figs. 7 and 8).

- Sinus Arrest

In a Sinus Arrest, SA node fails to even produce and transmit the P wave electrical impulses. Therefore, no more cardiac activity, until another electrical focus takes over. As a result, the pause will be longer and the conducted beat would be not Sinus.

Whereas in a Sinus Block dropped beat is followed by an immediate P wave, here, it continuous to be paused until any automaticity focus kicks in. That is, heart will begin its own pacemaker irrespective of conduction of the P wave [9] (Figs. 9 and 10).

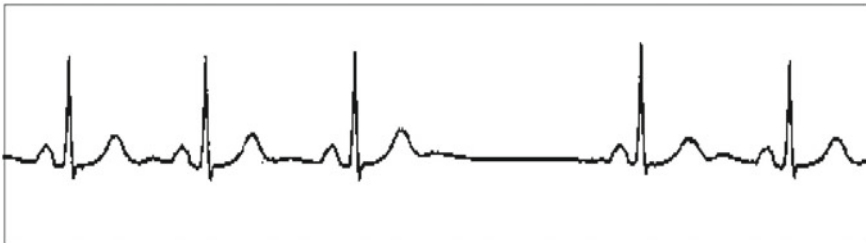


Fig. 7 Real-time sinus block ECG strip [10]. Adapted from life in the fastlane, by Ed Burns, Sep 20, 2020, from <https://litfl.com/wp-content/uploads/2018/08/ECG-Type-II-sino-atrial-exit-block.jpg> licensed under CC BY 4.0

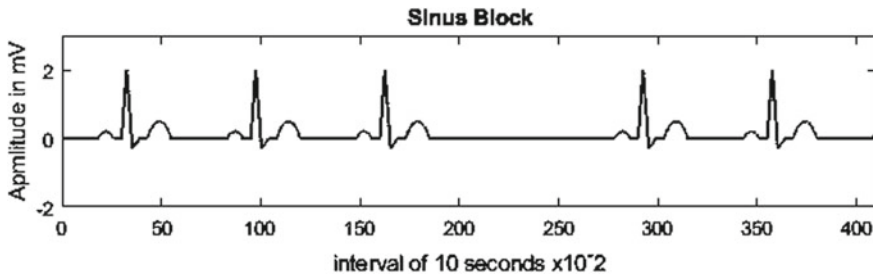


Fig. 8 Generated sinus block, with the dropped beat followed by a normal sinus rhythm



Fig. 9 Real-time sinus arrest ECG strip [11]. Adapted from ddxof, Sep 20, 2020, from <https://ddxof.com/ecg-guide-part-ii/sinus-arrest/> under CC BY 4.0

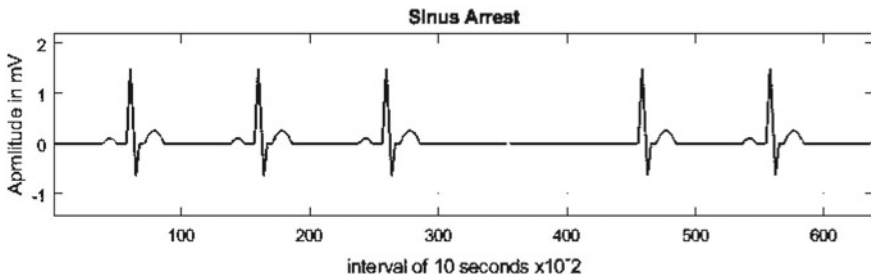


Fig. 10 Generated sinus arrest. Complete drop beat followed by a missing P wave

B. *Atrioventricular Node (AV)*

- First Degree AV Block

First degree blocks are often easy to detect and are caused due to the delay in the action potential in the AV node from the Atria to the Ventricles [9, 12].

Feature considered while modelling,

The delay caused is more than the normal and is greater than 200 ms, therefore resulting in a prolonged PR interval. However, QRS complexes and T waves remain unaffected (Figs. 11 and 12).

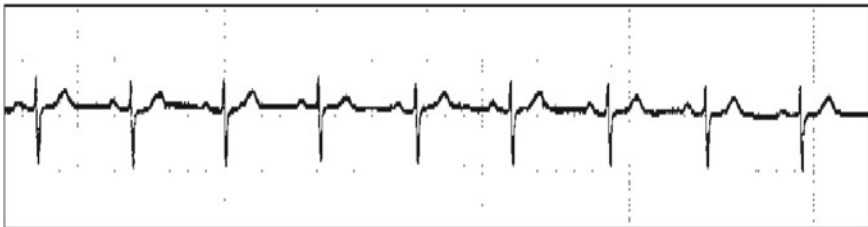


Fig. 11 Real-time first degree AV block ECG strip [13]. Adapted from Wikimedia Commons, by Michael Rosengarten BEng, MD McGill, Sep 2020 from https://upload.wikimedia.org/wikipedia/commons/e/5/21_%28CardioNetworks_ECGpedia%29.jpg licensed under CC BY 3.0

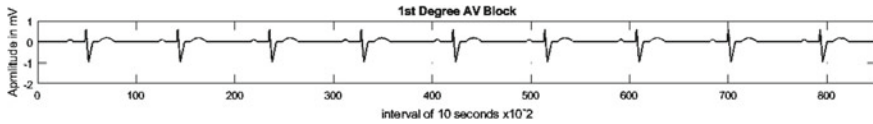


Fig. 12 Generated first degree AV block. With a prolonged PR interval about (350 ms)

- Second Degree AV Block (Type I/II)

In a second degree AV block, a block occurs when each successive impulse starting from the SA node is delayed. These successive delay adds up during each beat and continues until the impulse fails to conduct the ventricles or is very minuscule to note the difference.

There are two types, the second degree Type I AV block or Wenckebach and Mobitz Type II Second degree AV Blocks.

Features were considered to model Type I block. Since the SA node initiates normally, the P wave is of normal size and duration.

Albeit, the PR interval progressively increases until it appears without the QRS complex. It repeats itself after each dropped beats [9, 14] (Figs. 13 and 14).

Features considered while modelling the Type II AV Block are in contrast to Type I AV block. Since this occurs below, the AV node within the His Bundle and the right or left bundle branches [9, 12] there is no prolonging of the PR interval prior and after the non-conducted P wave or dropped QRS complex (Figs. 15 and 16).

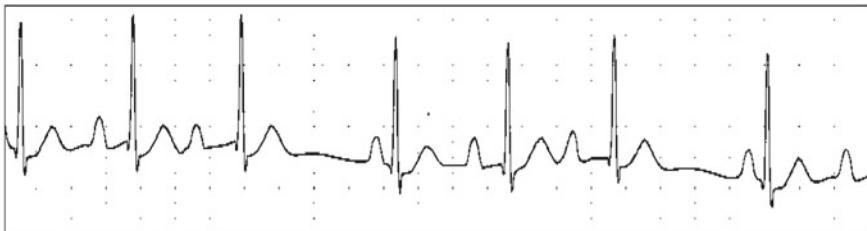


Fig. 13 Real-time second degree AV block type I ECG strip [15]. Adapted from life in the fastlane, by Ed Burns, Sep 20, 2020, from <https://litfl.com/wp-content/uploads/2018/08/Wenckebach-AV-block-with-Wenckebach-SA-block.jpg> licensed under CC BY 4.0

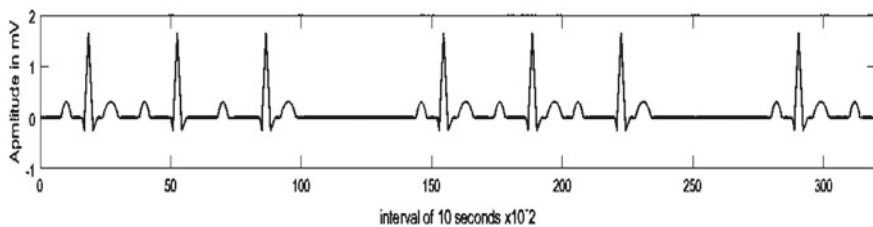


Fig. 14 Generated second degree AV block type I, with the prolonged PR interval until a drop beat

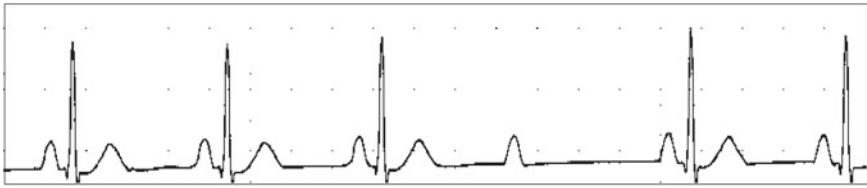


Fig. 15 Real-time second degree AV block type II ECG strip [16]. Adapted from life in the fastlane, by Ed Burns, Sep 20, 2020, from <https://litfl.com/wp-content/uploads/2018/08/ECG-Mobitz-II-Hay-AV-Block-1.jpg> licensed under CC BY 4.0

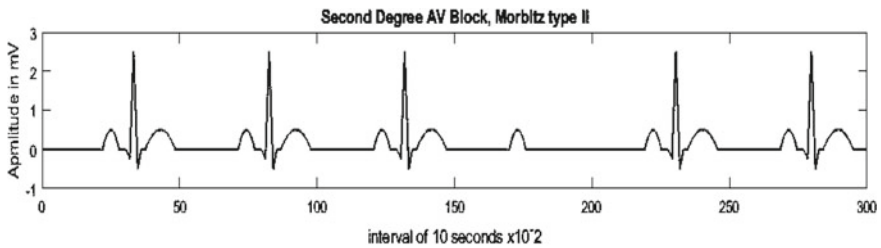


Fig. 16 Generated second degree AV block type II, with no prolonged PR interval unlike the type I, however, followed by a dropped beat

The QT interval and T wave are under normal limits, and in both the types, RR remains constant.

- Third Degree AV Block

Third degree AV block is a complete heart block, which occurs when any action potential from the atria to the ventricle completely fails to conduct from the AV node. This type of pattern on the EEG is called a complete Heart Block or Complete Atrioventricular (AV) Dissociation. Here, the atria conducts normally but ventricles conduct their own impulses separately [9].

Features to be considered—Due to the regular impulses of AV node, the P wave starts off normally. However, those impulses do not conduct through AV node due to a complete block at the AV node. P–P intervals are regular and QRS are regular; however, they occur at different rates independent of one another. Therefore, it has no association when P wave is followed by QRS complex resulting in a prolonged PR interval. There are two possibilities: complete AV block with junctional focus and ventricular focus. Junctional focus has a narrow QRS complex and is usually around 40–60 BPM, whereas ventricular focus has wider QRS complex and is usually 20–40 BPM (Figs. 17 and 18).

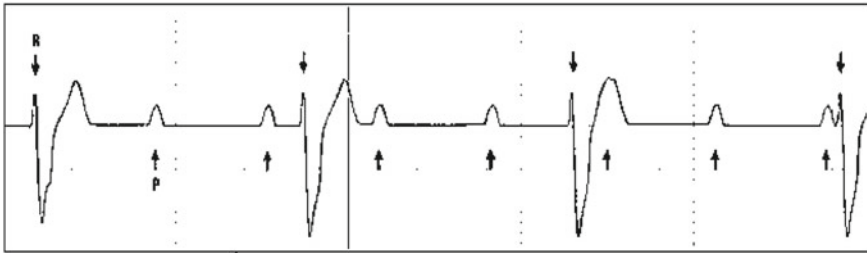


Fig. 17 Real-time third degree AV block ECG strip (ventricular focus) [17]. Adapted from life in the fastlane, by John Larkin, Sep 20, 2020, from <https://litfl.com/wp-content/uploads/2018/08/3rd-degree-heart-block.jpg> licensed under CC BY 4.0

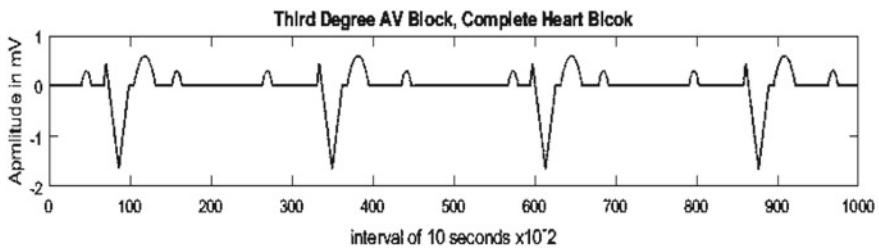


Fig. 18 Generated third degree AV block (ventricular focus), With wider QRS complex and slowed heartbeat

C. Bundle Branch (LBB/RBB)

• Bundle Branch Blocks

Right Bundle Branch Block or RBBB occurs when the electric conduction, specifically in the His Purkinje system is interrupted. Resulting in a widened QRS complex [18].

Features considered while modelling the Branch block were widened QRS complex greater than or equal to 120 ms.

T waves tend to be discordant to the terminal QRS vector [18] and Therefore results in an inverted T Wave (Figs. 19 and 20).

Left Bundle Branch Block or LBBB occurs when the conduction in the left bundle is delayed. This is results in delayed depolarization of the left ventricle and resulting in a widened QRS complex [9].

Features considered while modelling the LBBB. Widened QRS complex, greater than or equal to 120 ms. A narrow Q wave maybe present or mostly there is the absence of Q waves. T waves usually are opposite to the QRS deflection (Figs. 21 and 22).

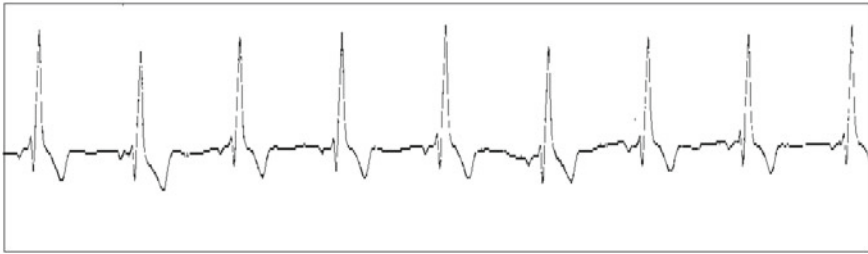


Fig. 19 Real-time right bundle branch block ECG strip [19]. Adapted from Wikimedia Commons, by Steven Fruitsmaak, Sep 20, 2020, https://upload.wikimedia.org/wikipedia/commons/9/96/Bifascicular_block_ECG.png licensed under CC by 3.0

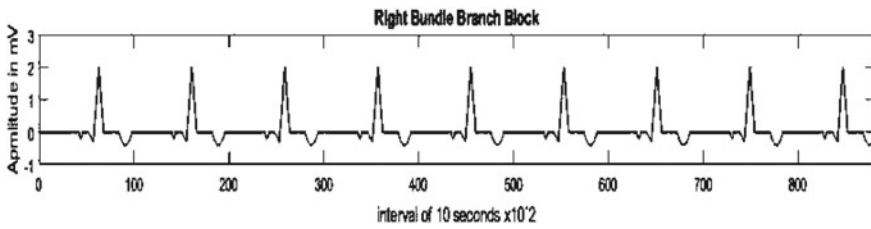


Fig. 20 Generated right bundle branch block. With the widened QRS complex and an inverted T wave

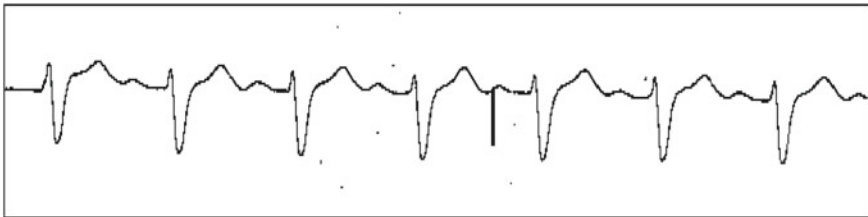


Fig. 21 Real-time left bundle branch block ECG strip [20]. Adapted from Wikimedia Commons, Sep 20, 2020, from https://upload.wikimedia.org/wikipedia/commons/5/5b/ECG_008_a.jpg licensed under CC BY 3.0

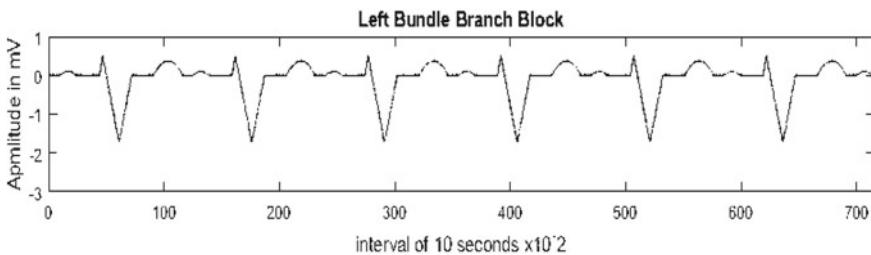


Fig. 22 Generated left bundle branch block with a wide QRS complex and T wave opposite to the QRS deflection

5 Conclusion

Various arrhythmias are generated using the proposed model elaborated in Sect. 3 using MATLAB as a computation tool. The heart blocks including, but not limited to, Sinus Block, Sinus Arrest, First Degree AV Block, Second Degree AV Block Type I, Type 2, Right and Left Bundle Branch Blocks have been generated and validated with the real ECG signal and resulting in an analogous output to that in comparison.

Further, this project can be extended as a base, in developing machine learning models to detect the generated heartbeat to that of the real ECG signal, and therefore enabling the generated model as a benchmark to compare with the real ECG signal for greater accuracy of diagnosis.

Acknowledgements This work is a part of the project supported by KSHEMA, Nitte Deemed to be University, Mangaluru, under the notion of a social cause aiming at giving an illustrative insight into electrocardiogram for the purpose of academic use.

References

1. P.E. Mcsharry, G.D. Clifford, L. Tarassenko, L.A. Smith, A dynamical model for generating synthetic electrocardiogram signals. *IEEE Trans. Biomed. Eng.* **50**(3) (March 2003)
2. P. Dolinský, I. Andras, L. Michaeli, D. Grimaldi, Model for generating simple synthetic ECG signals. *Acta Electrotechnica et Informatica*. **18**, 3–8 (2018). <https://doi.org/10.15546/aeii-2018-0019>
3. J. Kubíček, M. Penhaker, R. Kahankova, Design of a synthetic ECG signal based on the Fourier series, in *Proceedings of the 2014 International Conference on Advances in Computing, Communications and Informatics, ICACCI* (2014), pp. 1881–1885. <https://doi.org/10.1109/icaaci.2014.6968312>
4. P. Kovács, ECG signal generator based on geometrical features. *Annales Universitatis Scientiarum Budapestinensis de Rolando Eötvös Nominatae Sectio Computatorica* **37**, 247–260 (2012)
5. *ECG Interpretation Made Incredibly Easy!*, 5th edn. (Wolters Kluwer—Lippincott Williams & Wilkins)
6. E.J. Berbari, *Principles of Electrocardiography, Biomedical Engineering Handbook*, 2nd edn.
7. Madher088, Conduction system of the heart. Wikimedia Commons (20 Sep 2020). Retrieved from <https://commons.wikimedia.org/w/index.php?curid=29922595>
8. R. Karthik, ECG simulation using MTAB, in *Principle of Fourier Series* (College of Engineering, Guindy, Anna University, Chennai)
9. D. Dubin, *Rapid Interpretation of EKG's: An Interactive Course* (Cover Publishing Company, 2000)
10. Ed Burns, Sino atrial exit block. Life in the fastlane (20 Sep 2020). Retrieved from <https://litfl.com/wp-content/uploads/2018/08/ECG-Type-II-sino-atrial-exit-block.jpg>
11. Ddxof, Sinus arrest (20 Sep 2020). Retrieved from <https://ddxof.com/ecg-guide-part-ii/sinus-arrest/>
12. O.S. Narula, B.J. Scherlag, P. Samet, R.P. Javier, Atrioventricular block: localization and classification by His bundle recordings. *Am. J. Med.* **50**(2), 146–165 (1971)
13. Michael Rosengarten, First degree AV block. Wikimedia Commons (20 Sep 2020). Retrieved from https://upload.wikimedia.org/wikipedia/commons/e/e5/21_%28CardioNetworks_ECGpedia%29.jpg

14. E.C.G. Challenge, identifying second-degree atrioventricular block (type I). *J. Cardiovasc. Nurs.* (2007)
15. Ed Burns, (20 Sep 2020) from <https://litfl.com/av-block-2nd-degree-mobitz-i-wenckebach-phenomenon/>
16. Ed Burns, 2nd degree AV block type II. Wikimdia Commons (20 Sep 2020). Retrived from <https://litfl.com/wp-content/uploads/2018/08/ECG-Mobitz-II-Hay-AV-Block-1.jpg>
17. J. Larkin, 3rd degree, complete heart block. *Life In The Fastlane* (20 Sep 2020). Retrived from <https://litfl.com/wp-content/uploads/2018/08/3rd-degree-heart-block.jpg>
18. W.T. Harkness, M. Hicks, Right bundle branch block (RBBB) [Updated 2020 Aug 13], in *StatPearls [Internet]. Treasure Island (FL)* (StatPearls Publishing; Jan 2020). Available from: <https://www.ncbi.nlm.nih.gov/books/NBK507872/>
19. Steven Fruitsmaak, Right bundle branch block (20 Sep 2020). Retrived from https://commons.wikimedia.org/wiki/File:Bifascicular_block_ECG.png
20. Left bundle branch block (20 Sep 2020). Retrived from https://upload.wikimedia.org/wikipedia/commons/5/5b/ECG_008_a.jpg

Iterative Thresholding-Based Spectral Subtraction Algorithm for Speech Enhancement



Raj Kumar, Manoj Tripathy, and R. S. Anand

1 Introduction

Speech enhancement (SE) techniques find many applications such as automatic speech recognition (ASR) systems, speaker recognition, online conferencing, and voice-controlled devices for noise suppression and intelligibility improvement.

Spectral subtraction (SS) algorithm introduced by Boll [1] is still the most preferred speech enhancement algorithm because of simple and reliable design, which makes it the best choice for real-time application, e.g., online conferencing or audio calling. SS algorithm is easily implementable at moderate computing platforms, e.g., DSP processor [2] or FPGA [3], which makes this algorithm a better choice for hearing aids or speech assistive devices. In the SS algorithm, the noise component is subtracted from the spectrum of noisy speech. SS is still an active research area of speech enhancement, used in combination with other techniques like deep recurrent neural network [4], least mean square adaptive filter [5], statistical models [6], deep neural network [7, 8], orthogonal matching pursuit [9], etc.

If clean speech $x(n)$ is corrupted by additive noise $d(n)$, then noisy speech $y(n)$ is given by 1.

$$y(n) = x(n) + d(n) \quad (1)$$

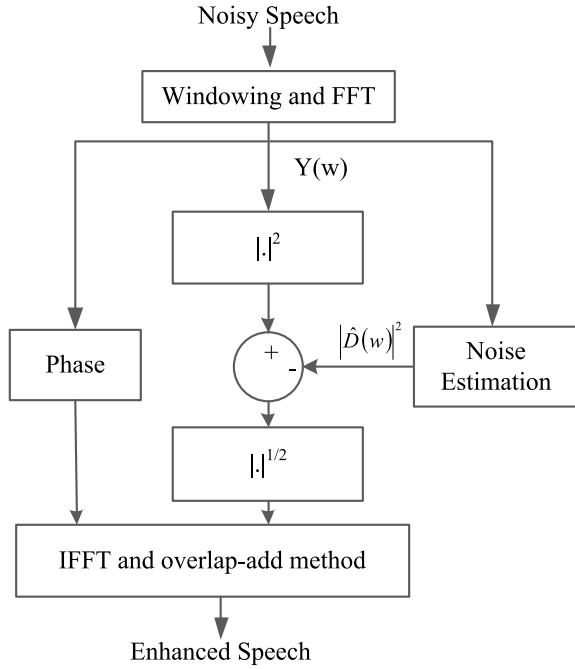
The approximate relationship between spectral power of clean speech, noise, and noisy speech is given by 2.

R. Kumar (✉) · M. Tripathy · R. S. Anand
Electrical Engineering Department, Indian Institute of Technology Roorkee, Roorkee, India
e-mail: rkumar17@ee.iitr.ac.in

M. Tripathy
e-mail: manoj.tripathy@ee.iitr.ac.in

R. S. Anand
e-mail: r.anand@ee.iitr.ac.in

Fig. 1 General spectral subtraction algorithm



$$|Y(w)|^2 = |X(w)|^2 + |D(w)|^2 \tag{2}$$

where $Y(w)$ represents noisy spectrum, $X(w)$ represents clean speech spectrum, and $D(w)$ represents noise spectrum. A general methodology used in the SS algorithm is shown through the block diagram in Fig. 1. The term $\hat{D}(w)$ denotes estimated noise spectrum, which is estimated during silence intervals, i.e., when speech is absent.

In the simplest form, SS can be formulated as shown in 3 to get the estimated clean speech spectrum $\hat{X}(w)$.

$$|\hat{X}(w)|^2 = |Y(w)|^2 - \alpha |\hat{D}(w)|^2 \tag{3}$$

Here, $\alpha (\alpha \geq 1)$ ensures subtraction result non-negative value. However, this results in musical noise and removes residual noise in noise only region, which is a must for speech’s naturalness. To remove the mentioned problem, Berouti et al. [10] have proposed the parametric spectral power subtraction algorithm, as described in 4.

$$|\hat{X}(w)|^2 = \begin{cases} |Y(w)|^2 - \alpha |\hat{D}(w)|^2 & \text{if } |Y(w)|^2 \geq (\alpha + \beta) |\hat{D}(w)|^2 \\ \beta |\hat{D}(w)|^2 & \text{else} \end{cases} \tag{4}$$

Here, $\beta(0 < \beta \ll 1)$ avoids isolated peaks in spectrum to suppress musical noise. The noisy phase, $\angle Y(w)$ is used in the final step because phase information is almost unchanged except at very low SNR [11] as shown in 5.

$$\hat{X}(w) = \left| \hat{X}(w) \right| e^{\angle Y(w)} \tag{5}$$

SS performance degrades at low SNR conditions [12]. Several modified SS algorithms have been proposed like multi-band SS [13], reduced delay convolution, adaptive averaging SS [14], and geometric SS [15] to deal with limitations of the SS algorithm.

From Fig. 1, it is clear that the performance of the SS algorithm greatly depends on noise estimation; hence, better the noise estimation, better will be performance [16, 17]. Other enhancement techniques also require prior noise information, e.g., the probability distribution of noise is assumed to be known in Wiener filter [18] and MMSE algorithm [19] or it assumes that noise and speech have independent spectral feature as in the case of subspace approach [20].

Martin and Cohen [21] has introduced the improved minima controlled recursive averaging (IMCRA) algorithm, which performs better than all methods mentioned earlier. It estimates noise even in speech-dominant frames and updates noise power recursively. It uses minima tracking of smoothed periodogram for noise estimation.

The algorithm presented in the paper use speech characterize to use as a clue to locate noise-dominant frame. Recently developed compressive sensing (CS) [22, 23] exploits signal characteristic. Equation 6 represents a compressible signal $x \in R^N$ in sparsifying basis $\psi = [\psi_1, \psi_2, \dots, \psi_N]$ while in 7, x_T represent reconstructed speech by keeping T largest coefficients of X where $T \ll N$.

$$x = \sum_{i=1}^N X(i)\psi(i) \tag{6}$$

$$x_T = \sum_{i=1}^T X_T(i)\psi_T(i) \tag{7}$$

The signal $x(n)$ will be sparse if reconstruction error (RE) as shown in 8 is exactly or nearly zero.

$$\text{R.E.} = \sum_{i=1}^N (x(i) - x_T(i))^2 \tag{8}$$

A CS-based signal recovery problem from noisy measurement y is shown in 9 which is referred as l_0 minimization problem.

$$x^{\text{opt}} = \arg \min_x \frac{1}{2} \|y - Ax\|_2^2 + \lambda \|x\|_0 \tag{9}$$

Here, $\|x\|_0$ represents number of non-negative elements in x . Above problem is basically a search problem to recover sparse vector x^{opt} which become exhaustive if dimension of x is large. The alternate solution method is to solve 9 by l_1 minimization as shown in 10.

$$x^{\text{opt}} = \arg \min_x \frac{1}{2} \|y - Ax\|_2^2 + \lambda \|x\|_1 \quad (10)$$

where l_1 -norm is defined as:

$$\|x\|_1 = \sum_i |x(i)| \quad (11)$$

It has been proved that if measurement is sufficient and A satisfy coherence property, then l_1 minimization problem will give same solution as l_0 minimization problem [24]. In the time domain, speech is not a sparse signal, but it shows some sparsity level in some other domain like wavelets, discrete Fourier, or discrete cosine transform domain [25]. The major advantage of CS-based SE is that it relies on signal characteristics rather than noise characteristics; hence, performance does not change whether the noise is stationary or non-stationary, which makes it useful for a real-world scenario. In paper [26], authors have described all available methods to solve the recovery problem defined in 9 and 10.

2 Iterative Soft Thresholding

Iterative thresholding is a technique commonly used to recover the signal from degraded or under-sampled signal [27, 28] based on the fact that the signal is sparse in some sparsifying domain. For solving l_1 -regularized least square problem in 10, an algorithm, called split augmented Lagrangian shrinkage algorithm (SALSA) [29] based on the iterative thresholding, has been utilized in the proposed algorithm.

The proposed algorithm for spectral denoising of time domain noisy signal y_{fr} is as follow:

$$\begin{aligned}
 &\mathbf{Initialize} : y_{\text{fr}}, X = |A(y_{\text{fr}})|, \lambda, \mu > 0, d = 0 \\
 &\mathbf{Repeat} \\
 &\quad v = \text{soft}(X + d, \lambda/\mu) - d \\
 &\quad d = 1/\mu A(y_{\text{fr}} - A^T(v)) \\
 &\quad X = d + v \\
 &\mathbf{End} \quad (12)
 \end{aligned}$$

where A and A^T represent Fourier and inverse Fourier transform, respectively, such that $AA^T = I$, I is identity matrix. The function $\text{soft}(x, \tau)$ in 12, attenuates input value x above threshold τ while values lower than threshold is made zero as shown in 13.

$$\text{soft}(x, \tau) = \begin{cases} x - \tau & \text{if } \tau < x \\ 0 & \text{if } -\tau < x < \tau \\ x + \tau & \text{if } x < -\tau \end{cases} \quad (13)$$

While X in 12 represents the spectral feature of a frame after thresholding. If a particular frame contains a speech, then after thresholding, larger magnitude peaks will be obtained. Thus, if $|X|^2 > P_{\min}$, then that frame is a speech-dominant frame; otherwise, it is noise-dominant frame. Where P_{\min} is the minimum power corresponding to residual noise. If a frame is a noise dominant, then noise power is updated recursively using 14.

$$\hat{D}_{\text{fr}} = a\hat{D}_{\text{fr}-1} + (1 - a)|A(y_{\text{fr}})|^2, 0 < a < 1 \quad (14)$$

where a is smoothing coefficient, \hat{D}_{fr} is estimated noise for current frame and $\hat{D}_{\text{fr}-1}$ is estimated noise previously. In the final step, estimated noise will be subtracted from noisy spectrum using 4.

The proposed algorithm is similar to the minima tracking of periodogram as in IMCRA; instead, it uses thresholding of spectra based on the sparsity of speech.

3 Experiment

A. Experiment Setup

In the experiment, NOIZEUS [30] data has been used, having 30 sentences spoken by three male and three female speakers. All speech samples are sampled at 8 kHz. Noises used in the experiment are babble, airport, exhibition, and car environment. Noises have been added to clean speech with SNR from -10 to 10 dB. The Hamming window has been used for windowing with 50% overlapping in all experiments.

B. Performance Evaluation Measures

For evaluation purposes, frequency weighted segmental SNR, fwSNR [31] is used to assess the gain in quality. Perceptual evaluation of speech quality (PESQ) [32] has been used to measure the overall quality of the enhanced speech. Short-time objective intelligibility (STOI) [33] measures speech intelligibility by computing the average correlation between the clean and enhanced speech temporal envelope in multiple frames and bands.

Table 1 Effect of smoothing coefficient on fwSNR performance of proposed algorithm in various noise environment

Smoothing coefficient (a)						
Noise type	0.4	0.5	0.6	0.7	0.8	0.9
Exhibition	4.8385	4.8544	4.8473	4.8540	4.8344	4.7881
Babble	4.8125	4.8114	4.8078	4.8062	4.7973	4.8039
Airport	4.7888	4.7944	4.8023	4.8052	4.8019	4.8086
Car	4.8939	4.9099	4.9098	4.9075	4.8901	4.8602

C. Effect of Smoothing Coefficient and Frame Length

Table 1 shows the effect of variation of smoothing coefficient, a on fwSNR gain under various noise condition at 0 dB with frame length of 80 ms and $P_{\min} = 5 \times 10^{-2}$. The results shown in Table 1 indicate that, in most of the cases, the proposed algorithm performs good when $a \in [0.5, 0.8]$ for the chosen database.

Results in Table 2 show the effect of varying frame size on the speech quality of the enhanced speech using the proposed algorithm for various noises under different SNR levels. The value of smoothing coefficient, a is 0.7 for whole experiment with $P_{\min} = 5 \times 10^{-5}$. The bold numerals represent the highest value in the row. As it is clear, fwSNR of enhanced speech increases as frame size increases from 20 to 80 ms in almost all cases. When frame size increases beyond 80 ms, there is a drop in intelligibility performance in terms of STOI though quality improves. It happens because a larger frame contains both noise and voice activity components, which cause the removal of voice components from the subsequent frame, resulting drop in intelligibility.

D. Comparison of Proposed Algorithm

In this experiment, the proposed method with a frame size of 80 ms (adopted from the previous experiment) is compared with the statistical approach IMCRA [21] algorithms. All parameters in the proposed algorithm have kept constant except λ , which represents weightage given to l_1 regularization. For higher noise, λ is kept high and vice versa.

Figures 2 and 3 show the performance comparison of the proposed algorithm with the IMCRA algorithm in term of fwSNR and PESQ for quality gain. In Fig. 2, fwSNR of enhanced speech using the proposed algorithm is higher under all noisy conditions at SNR level from -10 to 10 dB. In terms of PESQ, the proposed algorithm performs better in all noise conditions except car noise at SNR level more than 0 dB (Fig. 3, bottom-right).

Figure 4 shows the intelligibility performance comparison of the proposed algorithm with the IMCRA algorithm in terms of STOI. It concludes that the STOI of enhanced speech using the proposed algorithm is higher under all noisy conditions at all SNR levels. The intelligibility improved significantly at negative SNR.

Table 2 Effect of frame length on fwSNR performance of proposed algorithm in various noise environment

Noise type	Input SNR (dB)	Frame length		
		20 ms	40 ms	80 ms
Exhibition	-10	2.1037	2.4046	2.4913
	-5	2.9003	3.2261	3.3196
	0	4.4816	4.6089	4.7172
	5	6.7669	6.7235	6.6216
	10	8.7186	8.9080	8.9671
Babble	-10	2.1121	2.1258	2.1618
	-5	2.9921	3.1548	3.2000
	0	4.3492	4.6991	4.8108
	5	6.5673	6.7860	6.9251
	10	8.7706	9.3968	9.5200
Airport	-10	1.8858	1.9671	2.0662
	-5	2.9102	3.0546	3.1396
	0	4.4212	4.6955	4.7928
	5	6.3888	6.8418	6.9828
	10	8.8742	9.4703	9.6426
Car	-10	1.6770	1.9211	2.0912
	-5	2.6459	2.9867	3.1784
	0	4.1308	4.6079	4.8078
	5	6.5275	6.6981	6.8895
	10	9.2413	9.4334	9.3804

4 Conclusion and Future Work

The proposed algorithm extracts non-speech or noise dominated frames effectively in various non-stationary noises like babble, exhibition, airport, and car noise compared to the statistical approaches. The proposed algorithm’s performance does not depend on noise characteristics; hence, performance remains the same whether the noise is stationary or non-stationary.

The proposed method uses a larger smoothing coefficient ($0.5 < a < 0.8$) for noise update, i.e., higher weightage to current frame noise estimate compared to the previous frame; hence, it adapt to a large variation in noise power.

In this paper, the Fourier transform has been used as a sparsifying domain for speech. Various other transforms, e.g., wavelet transform, will be explored in which speech is more sparse than Fourier transforms in short time, thus, showing better performance in shorter frame time. A shorter frame time will reduce delay for real-time applications.

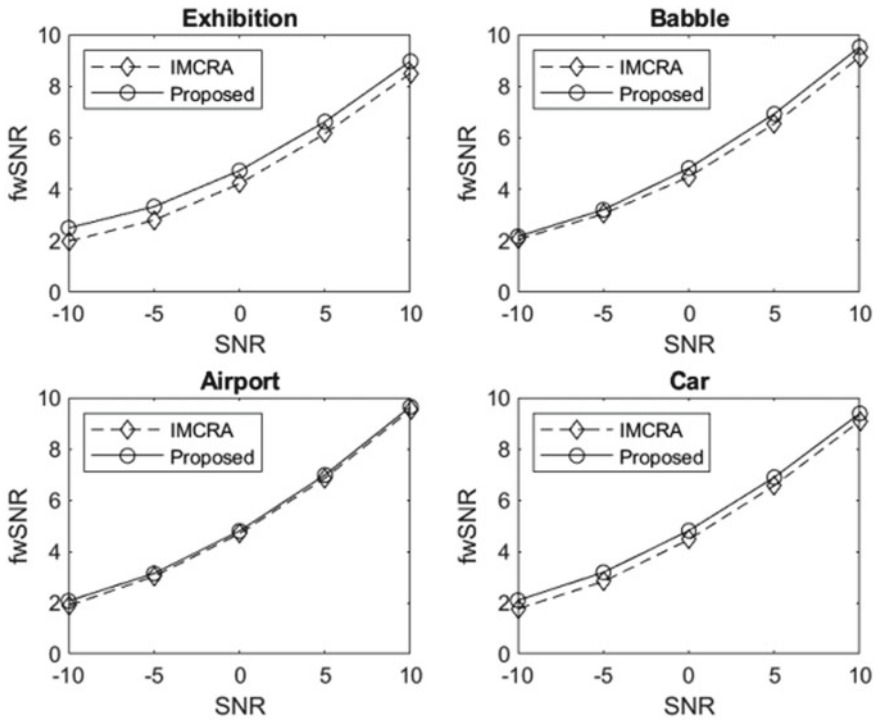


Fig. 2 fwSNR improvement performance

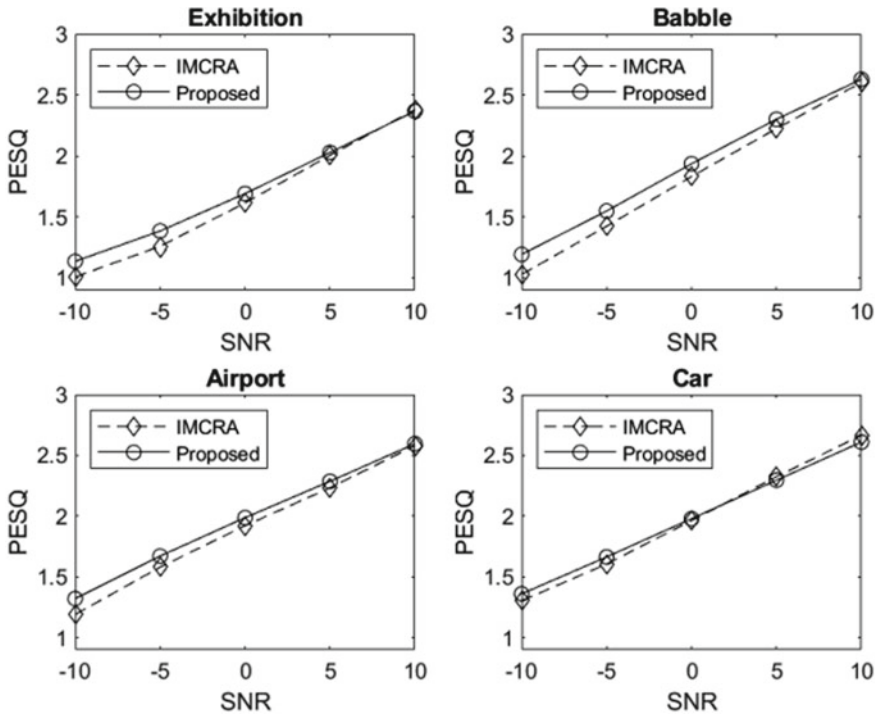


Fig. 3 PESQ improvement performance

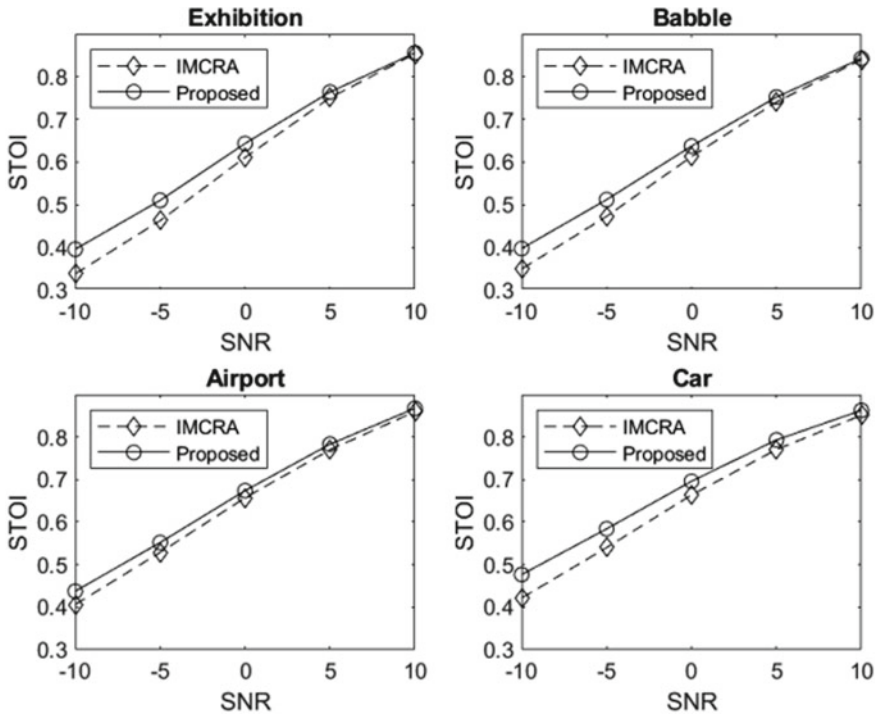


Fig. 4 Intelligibility improvement performance

References

1. S.F. Boll, Suppression of acoustic noise in speech using spectral subtraction. *IEEE Trans. Acoust.* **27**(2), 113–120 (1979). <https://doi.org/10.1109/TASSP.1979.1163209>
2. U. Purushotham, K. Suresh, Implementation of spectral subtraction using sub-band filtering in DSP C6748 processor for enhancing speech signal, in *Advances in Intelligent Systems and Computing* (Springer, Singapore, 2018), pp. 259–267
3. M. Bahoura, FPGA implementation of multi-band spectral subtraction method for speech enhancement, in *Midwest Symposium on Circuits Systems*, vol. 2017-Augus (2017), pp. 1442–1445. <https://doi.org/10.1109/mwscas.2017.8053204>
4. M. Keshavarzi, T. Goehring, R.E. Turner, B.C.J. Moore, Comparison of effects on subjective intelligibility and quality of speech in babble for two algorithms: a deep recurrent neural network and spectral subtraction. *J. Acoust. Soc. Am.* **145**(3), 1493–1503 (2019). <https://doi.org/10.1121/1.5094765>
5. D. Cao, Z. Chen, X. Gao, Research on noise reduction algorithm based on combination of LMS filter and spectral subtraction. *J. Inf. Process. Syst.* **15**(4), 748–764 (2019). <https://doi.org/10.3745/JIPS.04.0123>
6. V.R. Balaji, S. Maheswaran, M. Rajesh Babu, M. Kowsigan, E. Prabhu, K. Venkatachalam, Combining statistical models using modified spectral subtraction method for embedded system. *Microprocess. Microsyst.* **73**, 102957 (2020). <https://doi.org/10.1016/j.micpro.2019.102957>
7. T.K. Dash, S.S. Solanki, Speech intelligibility based enhancement system using modified deep neural network and adaptive multiband spectral subtraction. *Wirel. Pers. Commun.* **111**(2), 1073–1087 (2020). <https://doi.org/10.1007/s11277-019-06902-0>

8. Q. Zhou, Research on English speech enhancement algorithm based on improved spectral subtraction and deep neural network. *Int. J. Innov. Comput. Inf. Control* **16**(5), 1711–1723 (2020). <https://doi.org/10.24507/ijicic.16.05.1711>
9. H. Haneche, B. Boudraa, A. Ouahabi, A new way to enhance speech signal based on compressed sensing. *Meas. J. Int. Meas. Confed.* **151**, 107117 (2020). <https://doi.org/10.1016/j.measurement.2019.107117>
10. M. Berouti, R. Schwartz, J. Makhoul, Enhancement of speech corrupted by acoustic noise, in *IEEE International Conference on Acoustics, Speech, and Signal Processing*, 1 (1979), pp. 208–211. <https://doi.org/10.1109/icassp.1979.1170788>
11. Z. Chen, Y. Liu, G. Wang, S. Wang, W. Geng, Multiband spectral subtraction speech enhancement algorithm with phase spectrum compensation, in *Proceedings of the 2019 IEEE 4th Advanced Information Technology, Electronic and Automation Control Conference IAEAC*, vol. 20 (2019), pp. 2681–2685. <https://doi.org/10.1109/iaeac47372.2019.8997837>
12. T.K. Dash, S.S. Solanki, Comparative study of speech enhancement algorithms and their effect on speech intelligibility, in *Proceedings of the 2nd International Conference on Communication and Electronics Systems ICCES*, 2017, vol. 2018-Janua (2018), pp. 270–276. <https://doi.org/10.1109/cesys.2017.8321280>
13. S. Kamath, P. Loizou, A multi-band spectral subtraction method for enhancing speech corrupted by colored noise, in *ICASSP, IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 4, no. 2 (2002), p. 4164. <https://doi.org/10.1109/icassp.2002.5745591>
14. H. Gustafsson, S.E. Nordholm, I. Claesson, Spectral subtraction using reduced delay convolution and adaptive averaging. *IEEE Trans. Speech Audio Process.* **9**(8), 799–807 (2001). <https://doi.org/10.1109/89.966083>
15. Y. Lu, P.C. Loizou, A geometric approach to spectral subtraction. *Speech Commun.* **50**(6), 453–466 (2008). <https://doi.org/10.1016/j.specom.2008.01.003>
16. R. Dahlan, D. Krisnandi, A. Ramdan, H.F. Pardede, Unbiased noise estimator for Q-spectral subtraction based speech enhancement, in *Proceedings of the International Conference on Radar, Antenna, Microwave, Electronics and Telecommunications ICRAMET*, no. 2 (2019), pp. 65–68. <https://doi.org/10.1109/icramet47453.2019.8980396>
17. K. Ozawa, M. Morise, S. Sakamoto, K. Watanabe, Sound source separation by spectral subtraction based on instantaneous estimation of noise spectrum, in *6th International Conference on Systems Informatics, ICSAI*, vol. 1 (2019), pp. 1137–1142. <https://doi.org/10.1109/icsai48974.2019.9010477>
18. I. Almajai, B. Milner, J. Darch, S. Vaseghi, Visually-derived Wiener filters for speech enhancement, in *ICASSP, IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 4 (2007), pp. 2–5. <https://doi.org/10.1109/icassp.2007.366980>
19. Y. Ephraim, D. Malah, Speech enhancement using a minimum-mean-square error short-time spectral amplitude estimator. *IEEE Trans. Audio, Speech Lang. Process.* **32**(6), 1109–1121 (1984)
20. Y. Ephraim, H.L. Van Trees, A signal subspace approach for speech enhancement. *IEEE Trans. Speech Audio Process.* **3**(4), 251–266 (1995). <https://doi.org/10.1109/89.397090>
21. R. Martin, I. Cohen, Single-channel speech presence probability estimation and noise tracking, in *Audio Source Separation and Speech Enhancement* (Wiley, 2018), pp. 97–99
22. D.L. Donoho, Compressed sensing. *IEEE Trans. Inf. Theory* **52**(4), 1289–1306 (2006). <https://doi.org/10.1109/TIT.2006.871582>
23. R.G. Baraniuk, E. Candes, M. Elad, Y. Ma, Applications of sparse representation and compressive sensing. *Proc. IEEE* **98**(6), 906–909 (2010). <https://doi.org/10.1109/JPROC.2010.2047424>
24. M. Elad, *Sparse and redundant representations, from theory to applications in signal and image processing* (Springer, New York, 2010)
25. D. Wu, W.P. Zhu, M.N.S. Swamy, On sparsity issues in compressive sensing based speech enhancement, in *ISCAS 2012 IEEE International Symposium on Circuits and Systems* (2012), pp. 285–288. <https://doi.org/10.1109/iscas.2012.6271907>

26. M. Rani, S.B. Dhok, R.B. Deshmukh, A systematic review of compressive sensing: concepts, implementations and applications. *IEEE Access* **6**, 4875–4894 (2018). <https://doi.org/10.1109/ACCESS.2018.2793851>
27. I. Daubechies, M. Defrise, C. De Mol, An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. *Commun. Pure Appl. Math.* **57**(11), 1413–1457 (2004). <https://doi.org/10.1002/cpa.20042>
28. A. Beck, M. Teboulle, A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM J. Imaging Sci.* **2**(1), 183–202 (2009). <https://doi.org/10.1137/080716542>
29. M.V. Afonso, J.M. Bioucas-Dias, M.A.T. Figueiredo, Fast image recovery using variable splitting and constrained optimization. *IEEE Trans. Image Process.* **19**(9), 2345–2356 (2010). <https://doi.org/10.1109/TIP.2010.2047910>
30. Y. Hu, P.C. Loizou, Subjective comparison and evaluation of speech enhancement algorithms. *Speech Commun.* **49**(7–8), 588–601 (2007). <https://doi.org/10.1016/j.specom.2006.12.006>
31. J.M. Tribolet, P. Noll, B.J. McDermott, R.E. Crochiere, A study of complexity and quality of speech waveform coders, in *IEEE International Conference on Acoustics, Speech, and Signal Processing* (1978), pp. 586–590. <https://doi.org/10.1109/icassp.1978.1170567>
32. Perceptual evaluation of speech quality (PESQ), and objective method for end-to-end speech quality assessment of narrowband telephone networks and speech codecs, ITU, ITU-T Recomm. (2000), p. 862. [Online]. Available: <https://www.itu.int/rec/T-REC-P.862-200102-I/en>
33. C.H. Taal, R.C. Hendriks, R. Heusdens, J. Jensen, A short-time objective intelligibility measure for time-frequency weighted noisy speech, in *IEEE International Conference on Acoustics, Speech and Signal Processing* (2010), pp. 4214–4217. <https://doi.org/10.1109/icassp.2010.5495701>

Implementation of Cryptographic Algorithm (One-Time-Pad) with a RISC-V Processor



Priyanka Ashok Kurkuri, Saroja V. Siddamal, and Rashmi Kubsad

1 Introduction

Accelerators to processors are used to improve the execution of specified algorithm or may be complex computation by providing high amount of concurrency, providing data paths for temporary variables used for the computation and also by reducing the cost of instruction control in the common fetch–decode–execute cycle. To provide such a concurrency in performing cryptographic functions to the processor, One-Time-Pad a cryptographic algorithm that is known to be “non-crackable” is implemented as an accelerator to the Rocket chip processor. The cryptographic algorithm is a symmetric and stream cipher that can encrypt and decrypt a stream of plaintext(data to be encrypted). The processor used here is the Rocket chip processor, which is general SoC, generated using the RISC-V ISA. As the generator of the processor and ISA both are open source, it makes it possible to add additional functionalities in the future. As of now, there are various applications of cryptographic algorithms in day-to-day life, especially to store the data in a remote database, end-to-end encryption in communication, transfer of data. So, the applications of having an hardware accelerator would increase the computation speed, in performing the algorithm that is used extensively.

2 Literature Survey

The authors in paper [1] present a hardware accelerator called Falcon, for cryptography applications on energy limited devices. The architecture is a general engine rather than a specific fixed set of ciphers. It accelerates ciphers such as Cha-Cha,

P. Ashok Kurkuri · S. V. Siddamal (✉) · R. Kubsad
School of Electronics and Communication, K.L.E. Technological University, Hubli, India
e-mail: sarojavs@kletech.ac.in

AES, and RAS. The authors claim the energy consumption is reduced by 5–60× as compared to software. The authors also claim it provides the best security maintaining the same deployment time. The authors in paper [2] have developed Phalanx which is a parallel processor and accelerator form shared memory clusters. Clusters are connected with larger bandwidth. In paper [3], the authors have proposed tightly coupled accelerators which are integrated into RISC-V. The authors have implemented it in ASIC. They claim for ASIC implementation the energy consumption is reduced by 8.8. The authors in paper [4] have presented DSP accelerator for RISC-V. The accelerator is suitable for digital FIR filtering, bidimensional convolution, and pattern matching. The accelerator consists of 64-bit processing elements. It is implemented in a 28 nm FDSOI library of standard cells.

2.1 RISC-V SoC

For the design of hardware accelerator a 32-bit RV32I, RISC-V core with integer functionality is used. RISC-V ISA [5] is an open-source ISA by University of California, Berkeley. For the SoC, Berkeley's Rocket chip open source is opted. Rocket Chip Generator [6] was designed by the Berkeley Architecture Research (BAR) at University of California Berkeley (UCB). It is an open-source system-on-chip generator that generates general-purpose, open RISC-V ISA-based processor cores. It produces synthesizable RTL for customized system-on-chip. CHISEL an open-source hardware construction language is used to implement Rocket chip generator. The parameterization feature of Chisel is leveraged to devise a collection of sophisticated generators cores, caches, and interconnects which can be configured depending on need and hence produces customized integrated system-on-chip.

2.2 Rocket Core

Rocket core is a component of Rocket Chip Generator [6] as shown in Fig. 1, originally developed by UCB and currently supported by SiFive. It is capable of producing cores that configured have different size/type/level of caches, number of on-chip buses or any different configuration parameters. The core can either be in-order Rocket core or a out-of-order BOOM core.

Rocket has a classical five stage in-order scalar pipelined processor. It can implement 32-bit base integer RISC-V ISA and the 64-bit RISC-V ISA. The core has a branch predictor which is configurable. It also has a Memory Management Unit (MMU) with optional paged virtual memory support. The Rocket Custom Coprocessor Interface (RoCC) [7] is an interface between the Rocket Core and the attached custom coprocessors; it enables decoupled communication between the two. The RoCC interface provides set of signals that are generally essential for accelerators. The RoCC provides two types of interface: default RoCC interface and extended

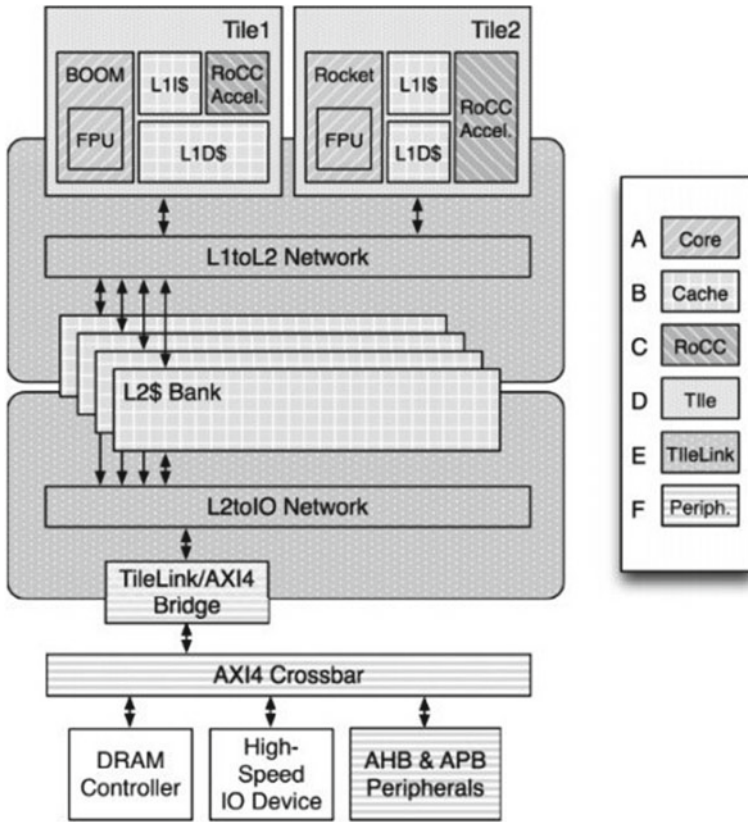


Fig. 1 Rocket chip generator

RoCC interface. The default RoCC interface has three subgroups: control, register, and Memory mode signals. The core subgroup signals ensure coordination between the core and the accelerator.

3 Contribution

This paper makes the following three contributions:

- Study of RISC-V ISA-based processor core
- Study of Rocket chip core and RoCC interface with Rocket Chip, i.e., coordination between core and the accelerator.
- Simulation of hardware accelerator and demonstration of advantage of accelerator in terms of speed.

4 Accelerator Design

The One-Time-Pad is a “non-crackable” cipher. For encryption, a stream of random keys generated is combined with plain text (see Fig. 2). Here, the input to the accelerator is sent via the rs1 field of the R-type instruction register and the encryption on this data is done using the LFSR.

4.1 Accelerator Hardware Architecture

The accelerator communicates with the Rocket chip via the RoCC link. The RoCC link uses R-type Instruction register for the communication. The inputs to the accelerator are given in the rs1 data register, so that the accelerator can perform encryption on it. Figure 3 shows how the Rocket chip and accelerator communicate via RoCC link using R-type register.

Figure 4 shows the architecture of the accelerator performing One-Time-pad algorithm. Figure 5 shows the architecture of accelerator connecting to Rocket chip via ROCC link. Accelerator in the implementation takes inputs as a vector (collection in CHISEL) of 64-bit unsigned integer. Accelerator has an LFSR method that produces the “pseudo-random” numbers as the key used to encrypt and decrypt the data. Output from the accelerator is also a vector of 64-bit unsigned integers. The input to the accelerator is plaintext that is the data to be encrypted. The key is randomly generated using an LFSR. This LFSR generates pseudo-random numbers which is of the same length as the plaintext. Here, the plaintext is of 64-bits, and the key is also 64-bits. One-Time-Pad algorithm implemented as an accelerator is connected to the Rocket core via RoCC interface that communicates through an R-type instruction register as shown in the figure above. Here, the input to the accelerator is sent via the rs1 field of the R-type instruction register and the encryption on this data is done using the LFSR.

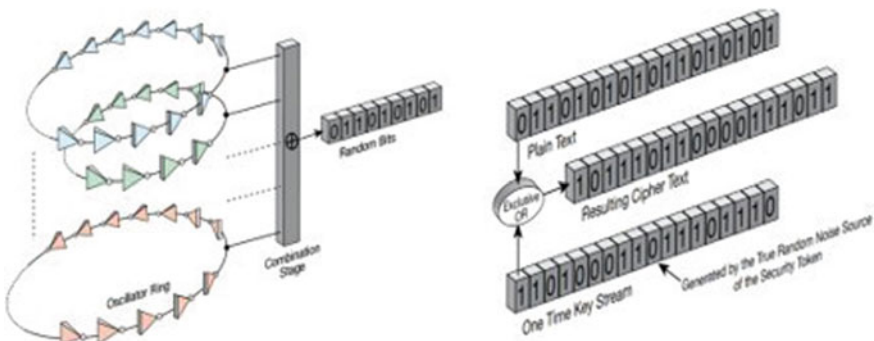


Fig. 2 One--time-pad process

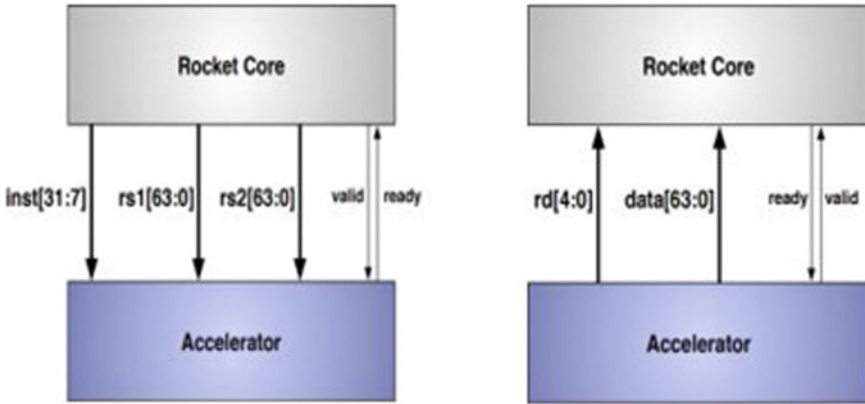


Fig. 3 Communication between rocket chip and accelerator via ROCC

4.2 Accelerator Hardware Implementation

When the data is to be encrypted, One-Time-Pad simply generates the key using LFSR and then performs an XOR (exclusive OR) operation between plaintext and the key.

$$\text{Encryption} = \text{plaintext XOR key}$$

When the data is to be decrypted, the encrypted text is again XOR'ed with the same key to get back the original plaintext.

$$\text{Decryption} = \text{Encryption XOR key}$$

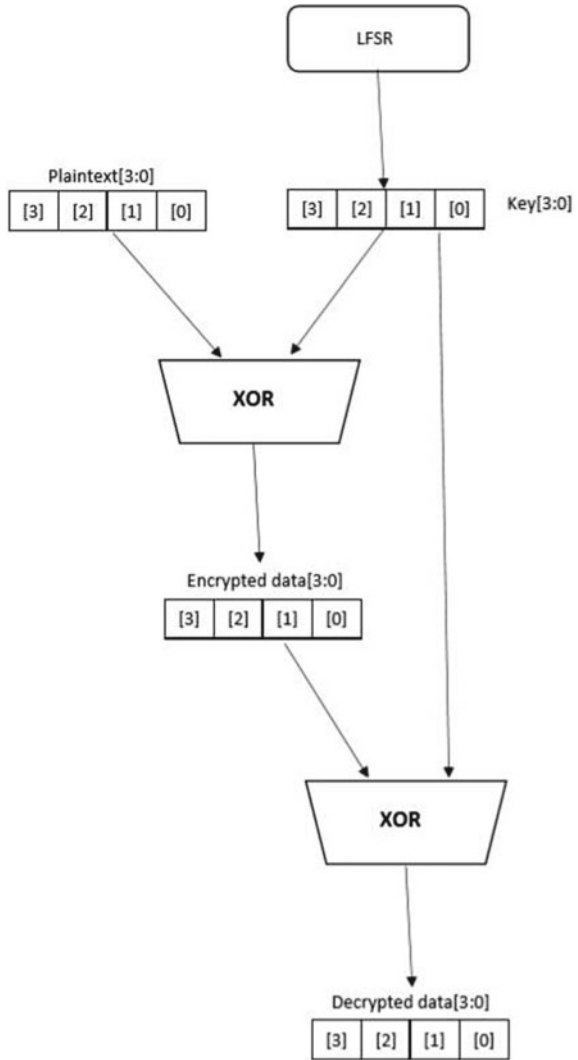
The algorithm is implemented in Constructing Hardware. In Scala Embedded Language (CHISEL), as an accelerator to the Rocket chip processor. The accelerator is connected to the Rocket chip processor by an RoCC link.

The accelerator is implemented such that it uses CHISEL principles of inheritance and composition to connect to the various modules of RoCC link and Rocket chip. By doing so, the accelerator gets the inputs via the RoCC link and outputs are given out via RoCC link itself.

5 Experimental Results

As the rocket chip, RISC-V and accelerator are built using CHISEL, and wrapper is eliminated. The accelerator comprises of multiply, add units. The total area is 131,150 μm^2 . The simulation is done using Scala Build Tool. The clock frequency of rocket

Fig. 4 Architecture of the accelerator performing one-time-pad algorithm



chip, i.e., 1 GHz, is considered as target operating frequency. The functionality of RISC-V and hardware accelerator are tested at this frequency. The implementation results of accelerator are as shown in Table 1.

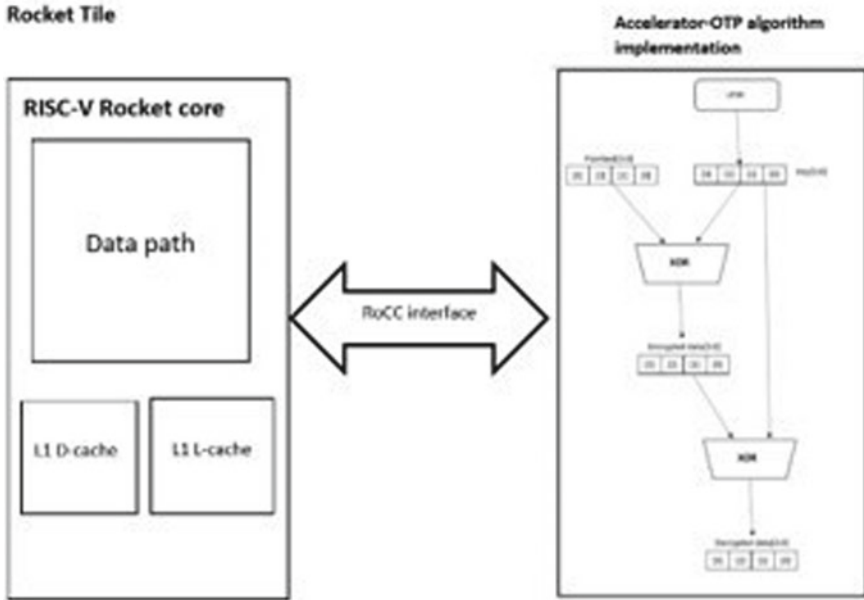


Fig. 5 Architecture of accelerator hardware implementation

Table 1 Implementation results of hardware accelerator

Area [μm^2]	f_{max}	Application
131,150	1 GHz	Encryption and decryption

6 Conclusion

The authors have presented an accelerator for RISC-V ecosystem suitable for cryptographic applications. The present implementation of the accelerator is limited to a simple stream and symmetric cipher; it can be further developed into a complex cryptographic algorithm implementation, which will open a lot many doors in hardware implementation and applications of cryptographic algorithms. By implementing specialized hardware for complex algorithms and using them as accelerators to processors (GPUs) in an SoC environment, we are enabling higher efficiency in computation when compared to software enabled algorithm implementation with a GPU.

Acknowledgements We are extremely grateful to K.L.E. Technological University for providing us with all the support, resources needed to accomplish this research. We would also like to heart full thank SiFive for helping throughout the process of research and helping us to achieve the desired result.

References

1. K. Kinningham et al., Falcon—A flexible architecture for accelerating cryptography, in *2019 IEEE 16th International Conference on Mobile Ad Hoc and Sensor Systems (MASS)* (2019), pp. 136–144
2. J. Gray, GRVI phalanx: a massively parallel RISC-V FPGA accelerator, in *International Workshop on Overlay Architectures for FPGAs (OLAF2016)* (Monterey, CA, USA, Feb 21, 2016), pp. 23–24
3. T. Fritzmann, G. Sig, J. Sepúlveda, RISQ-V: tightly coupled RISC-V accelerators for post-quantum cryptography, in Cryptology ePrint Archive, Report 2020/446 (2020)
4. L. Calicchia et al., Digital signal processing accelerator for RISC-V, in *26th IEEE International Conference on Electronics Circuits and Systems IEEE*. <https://doi.org/10.1109/ICECS46596.2019.8964670>
5. The RISC-V Instruction Set Manual Volume I: Unprivileged ISA, CS Division, EECS Department, University of California, Berkeley
6. K. Asanovic et al., The rocket chip generator, in Technical Report No. UCB/EECS-2016–17 Electrical Engineering and Computer Sciences University of California at Berkeley. <https://www.eecs.berkeley.edu/Pubs/TechRpts/2016/EECS-2016-17.html>
7. A. Waterman, Design and programming of a coprocessor for a RISC-V Architecture, in Electrical Engineering and Computer Sciences University of California at Berkeley, Technical Report No. UCB/EECS-2016-1. <https://www.eecs.berkeley.edu/Pubs/TechRpts/2016/EECS-2016-1.html>

Implementation of a Hebbian Learning Algorithm as an Accelerator to a RISC-V-Based Processor



Rashmi Kubsad, Saroja V. Siddamal, and Priyanka Ashok Kurkuri

1 Introduction

Accelerators to processors are used to improve the execution of specified algorithm or may be complex computation by providing high amount of concurrency, providing data-paths for temporary variables used for the computation and also by reducing the cost of instruction control in the common fetch–decode–execute cycle. To provide such a concurrency in performing cryptographic functions to the processor, Hebbian, a unsupervised neural network algorithm, is implemented as an accelerator to the Rocket chip processor. The processor used here, is the Rocket chip processor, which is general SoC, generated using the RISC-V ISA. As the generator of the processor and ISA both are open-source, it makes it possible to add additional functionalities in the future. The optimized weights are obtained using the Hebb network. Realization of logic gates as accelerator using Hebb network results in low power consumption.

2 Literature Survey

The authors in paper [1] present a hardware accelerator called Falcon, for cryptography applications on energy limited devices. The architecture is a general engine rather than a specific fixed set of ciphers. It accelerators ciphers such as ChaCha, AES, RAS, The authors claim that the energy consumption is reduced by 5–60× as compared to software. The authors also claim that it provides the best security maintaining the same deployment time. The authors in paper [2] have developed Phalanx which is a parallel processor and accelerator form shared memory clusters. Clusters are connected with larger bandwidth. In paper [3], the authors have proposed tightly

R. Kubsad · S. V. Siddamal (✉) · P. Ashok Kurkuri
School of Electronics and Communication, K.L.E. Technological University, Hubli, India
e-mail: sarojavs@kletech.ac.in

coupled accelerators which are integrated into RISC-V. The authors have implemented it in ASIC. They claim for ASIC implementation the energy consumption is reduced by 8.8. The authors in paper [4] have presented DSP accelerator for RISC-V. The accelerator is suitable for digital FIR filtering, bi-dimensional convolution, and pattern matching. The accelerator consists of 64-bit processing elements. It is implemented in a 28 nm FDSOI library of standard cells.

2.1 RISC-V SoC

For the design of hardware accelerator, a 32-bit RV32I, RISC-V core with integer functionality is used. RISC-V ISA [5] is an open-source ISA by University of California, Berkeley. For the SoC, Berkeley's Rocket chip open source is opted. Rocket chip generator [2] was designed by the Berkeley Architecture Research (BAR) at University of California Berkeley (UCB). It is an open-source system-on-chip generator that generates general-purpose, open RISC-V ISA-based processor cores. It produces synthesizable RTL for customized system-on-chip. CHISEL an open-source hardware construction language used to implement Rocket chip generator. The parameterization feature of Chisel is leveraged to devise a collection of sophisticated generators cores, caches, and interconnects which can be configured depending on need and hence produces customized integrated system-on-chip.

2.2 Rocket Core

Rocket core is a component of Rocket chip generator [6] as shown in Fig. 1, originally developed by UCB and currently supported by SiFive. It is capable of producing cores configured have different size/type/level of caches, number of on-chip buses or any different configuration parameters. The core can either be in order Rocket core or a out-of-order BOOM core. Rocket has a classical five stages in order scalar pipelined processor. It can implement 32-bit base integer RISC-V ISA and the 64-bit RISC-V ISA. The core has a branch predictor which is configurable. It also has a Memory Management Unit (MMU) with optional paged virtual memory support. The rocket custom coprocessor interface (RoCC) [7] is an interface between the Rocket Core and the attached custom coprocessors; it enables decoupled communication between the two. The RoCC interface provides set of signals that are generally essential for accelerators. The RoCC provides two types of interface: default RoCC interface and extended RoCC interface. The default RoCC interface has three subgroups: control, register, and memory mode signals. The core subgroup signals ensure co-ordination between the core and the accelerator (Fig. 2).

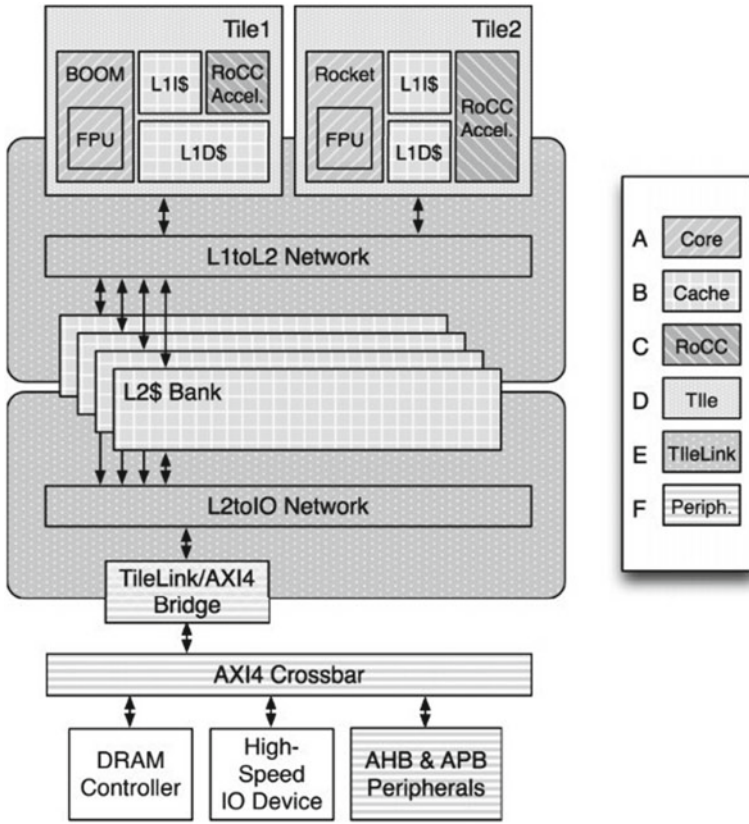
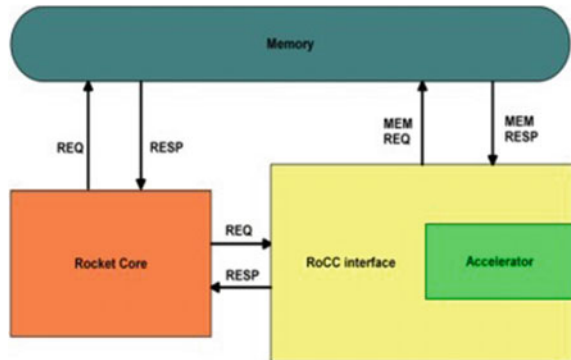


Fig. 1 Rocket chip generator [6]

Fig. 2 RoCC interface to RISC V and accelerator



3 Contribution

This paper makes the following three contributions:

- Study of RISC-V ISA-based Processor core (Sect. 2.1)
- Study of Rocket chip core and RoCC interface with Rocket chip (Sect. 2.1) i.e. co-ordination between core and the accelerator
- Simulation of hardware accelerator and demonstration the design of hardware accelerator (Sects. 4 and 5).

3.1 Accelerator Design

An unsupervised Hebbian learning is designed as an accelerator. According to Hebb's rule, the weights will increase proportional to the product of input and output. If two neurons are interconnected, then the weights associated with these neurons can be increased by changes in the synaptic gap. This learning algorithm is used design of logic gates. Hebb's principle can be used to update the weights between learning nodes such that the weight better represents the relationship between the nodes.

$$w_{ij}[n + 1] = w_{ij}[n] + \eta x_i[n]x_j[n] \quad (1)$$

where η is the learning rate and $x_i[n]x_j[n]$ are the output of the i th and j th element at time step n .

Hebbian learning algorithm is used to implement AND logic by adjusting the weights to the inputs and the bias.

3.2 Hebbian Learning Algorithm

The algorithm to find weight between neurons using Hebbian learning algorithm is as follows:

- Initially weights are set to zero i.e. $w_1 = 0$ and $w_2 = 0$.
- Consider x_1 and x_2 be return-to-zero type of inputs arrays.
- The activation function for input is set as an identity function.
- Let y be the target output array.
- The activation function for output is set to target array.
- The updated weights are calculated using the equation:

$$w_1(\text{new}) = w_1(\text{old}) + (x_1(i) * y(i)) \quad (2)$$

$$w_2(\text{new}) = w_2(\text{old}) + (x_2(i) * y(i)) \quad (3)$$

- The updated bias is calculated using the equation:

$$b(\text{new}) = b(\text{old}) + y(i) \tag{4}$$

- The net value is calculated for each pair of input using the equation:

$$\text{net}(i) = (x_1(i) * w_1) + (x_2(i) * w_2) + b(i) \tag{5}$$

To implement the algorithm, the architecture is designed as shown in Fig. 4

4 Accelerator Hardware Implementation

The accelerator communicates with the Rocket chip via the RoCC link. The RoCC link uses R-type instruction register for the communication. The inputs to the accelerator are given in the rs1 data register, so that the accelerator can perform Hebbian learning algorithm on it.

Figure 3 shows the architecture of accelerator connecting to Rocket chip via ROCC. The input arrays x_1 , x_2 , and y are sent to the accelerator using rs1, rs2, and function section of the R-type register. After the accelerator computes the output, the result is sent back to rocket chip using output (rd) section of the R-type register. The

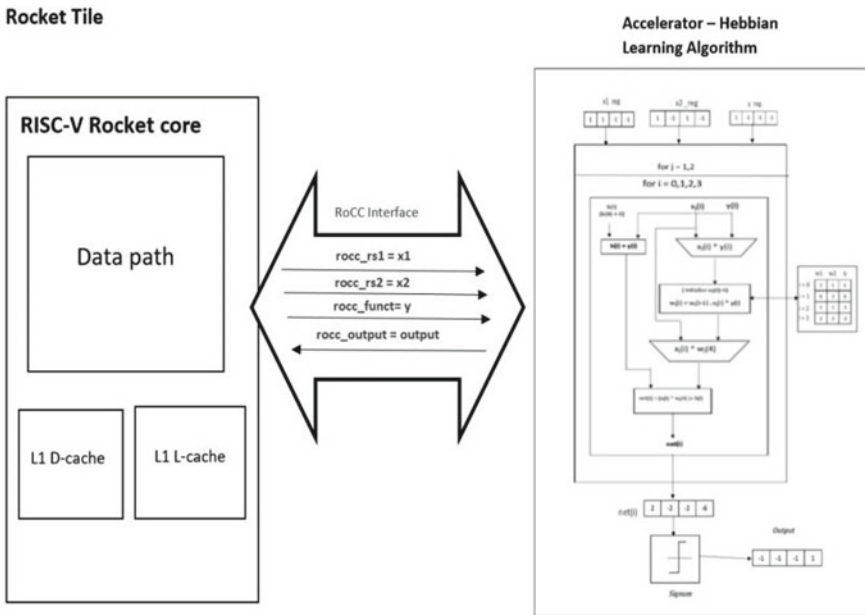


Fig. 3 Architecture of accelerator interfaced with rocket chip through RoCC interface

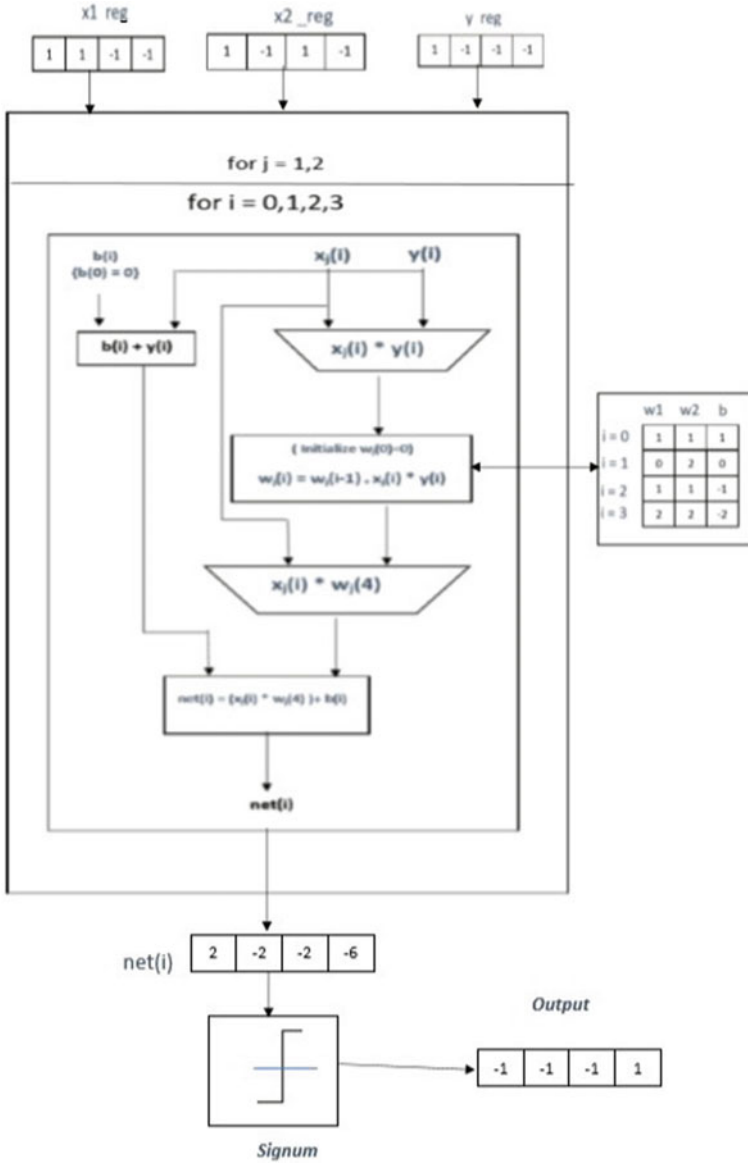


Fig. 4 Architecture of accelerator

algorithm is implemented in CHISEL as an accelerator to the Rocket chip processor. The accelerator is connected to the Rocket chip processor by a RoCC link. The accelerator is implemented such that it uses CHISEL principles of inheritance and

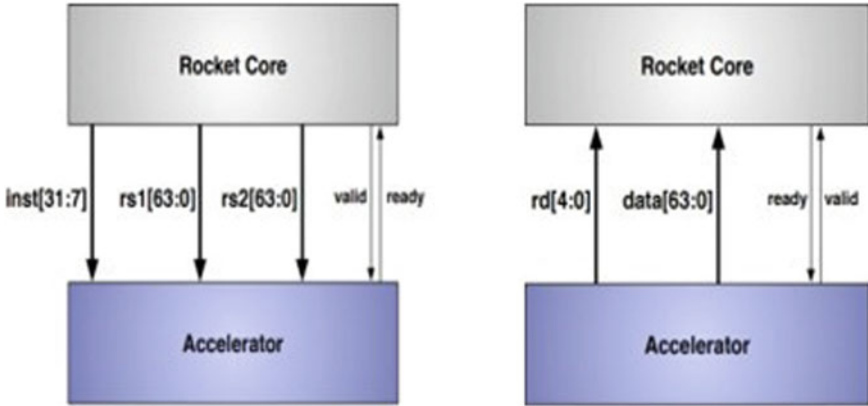


Fig. 5 Communication between rocket-chip and accelerator via RoCC

composition to connect to the various modules of RoCC link and rocket chip (Figs. 4 and 5).

5 Experimental Results

The experimental results carried on Rocket chip SoC are presented in this section. As the rocket chip, RISC-V and accelerator are built using CHISEL, and wrapper is eliminated. The simulation is done using Scala Build Tool. The clock frequency of rocket chip, i.e. 1 GHz is considered as target operating frequency. The functionality of RISC-V and hardware accelerator is tested at this frequency. The implementation results of accelerator are shown in Table 1.

The updated weights are calculated using Eq. (3) as shown in Table 2. The updated bias is calculated using Eq. (4) as shown in Table 3 and last set of values of the weights, and the output is calculated in order to check if the weights obtained are optimal. If the net value for a input set is greater than 0, output is set to 1. Alternatively, if the net value for a particular input set is less than 1, then output is set to -1. This optimized calculation results to low power consumption.

Table 1 2-bit RZ input with corresponding output implemented and logic

Input	Input	Bias	Target
x_1	x_2	b	y
1	1	1	1
1	-1	1	-1
-1	1	1	-1
-1	-1	1	-1

Table 2 Update weight and bias for each pair of input

Input	Input	Bias	Target	Weight for x_1	Weight for x_2	Updates bias
x_1	x_2	b	y	w_1	w_2	b
1	1	1	1	1	1	1
1	-1	1	-1	0	2	0
-1	1	1	-1	1	1	-1
-1	-1	1	-1	2	2	-2

Table 3 Ouput value for each pair of inputs

Input	Input	Bias	Target	Weight for x_1	Weight for x_2	Updates bias	Net output	Output
x_1	x_2	b	y	w_1	w_2	b	Net	output
1	1	1	1	1	1	1	2	1
1	-1	1	-1	0	2	0	-2	-1
-1	1	1	-1	1	1	-1	-2	-1
-1	-1	1	-1	2	2	-2	-6	-1

6 Conclusion

The authors have presented an accelerator for RISC-V ecosystem suitable for neural network applications. The rocket chip, RISC-V, and accelerator are built in CHISEL and simulated in Scala Build tool at frequency of 1Ghz. The hardware accelerator is designed for application such as neural network design of logic gates. The results show an optimized calculation which reduces power consumption.

References

1. K. Kinningham et al., Falcon—a flexible architecture for accelerating cryptography, in *2019 IEEE 16th International Conference on Mobile Ad Hoc and Sensor Systems (MASS)*, pp. 136–144 (2019)
2. J. Gray, RVI phalanx: a massively parallel RISC-V FPGA accelerator, in *International Workshop on Overlay Architectures for FPGAs (OLAF2016)* (Monterey, CA, USA, Feb 21, 2016), pp. 23–24
3. T. Fritzmann, G. Sig, J. Sepúlveda, RISQ-V: tightly coupled RISC-V accelerators for post-quantum cryptography, in *Cryptology ePrint Archive*, Report 2020/446 (2020)
4. L. Calicchia et al., Digital signal processing accelerator for RISC-V, in *26th IEEE International Conference on Electronics Circuits and Systems IEEE*. <https://doi.org/10.1109/ICECS46596.2019.8964670>
5. The RISC-V Instruction Set Manual Volume I: Unprivileged ISA, CS Division, EECS Department, University of California, Berkeley
6. K. Asanovic et al., The rocket chip generator, in Technical Report No. UCB/EECS-2016–17 Electrical Engineering and Computer Sciences University of California at Berkeley. <http://www.eecs.berkeley.edu/Pubs/TechRpts/2016/EECS-2016-17.html>

7. A. Waterman, Design and programming of a coprocessor for a RISC-V Architecture, in Electrical Engineering and Computer Sciences University of California at Berkeley, Technical Report No. UCB/EECS-2016-1. <https://www.eecs.berkeley.edu/Pubs/TechRpts/2016/EECS-2016-1.html>

Design and Implementation of Agricultural Drone for Areca Nut Farms



Raju Hajare, C. P. Mallikarjuna Gowda, and M. V. Sanjaya

1 Introduction

Scientists and engineers are striving hard for better solutions in order to resolve societal problems. So there is always a need of better and more practical approach which is easier to interact with human beings. There have been many efforts made in this direction in resolving the issues and guaranteeing the operational functions and with less efforts.

To accomplish this, unmanned aerial vehicles (UAVs) were developed to perform actions that were difficult for human beings to execute.

Over the last couple of years, UAVs have received a lot of attention due to their huge potential applications particularly in military and civilian field and industry as well. Some of the application areas are photography, agriculture and surveillance. Among all the UAVs, quad rotors have become most popular due to their rotor configuration which simplifies their analysis and control considerably.

With the advancements in new trending technologies like machine learning and artificial intelligence, a vast range of smart embedded and IOT applications are being developed by technological giants like Google and Amazon. These technologies combined with features like object detection, automatic routing and tracking using GPS, the potential applications are vast.

Through literature review, it has been understood that the Facebook has UAV built on solar power to enable internet access in remote zonal areas. A new prototype has

R. Hajare (✉) · C. P. Mallikarjuna Gowda
Department of Electronics and Telecommunication Engineering, BMS Institute of Technology and Management, Bengaluru, India
e-mail: rajuhajare@bmsit.in

C. P. Mallikarjuna Gowda
e-mail: cpmallikarjunagowda@bmsit.in

M. V. Sanjaya
Department of Electronics Engineering, IIIT, Bengaluru, India

been launched by Amazon, which is capable of delivering small order packages to the height of 120 m and then flying horizontally for up to 24 km [1–4].

In developing economies like India, the country's growth is majorly dependent on the agricultural sector. Becoming self-reliant in meeting the basic needs is the primary objective of every nation. To feed a nation of 1.3 billion people, it is necessary to give up the conventional agricultural methods and adopt new scientific ones employing automation. Thus, an engineer also shares the same amount of responsibility as that of a farmer in the agricultural sector. Monitoring large farms is a key issue faced by most farmers in India. This includes irrigation, spraying of pesticide and herbicide, and reaping the crop at the right time [5–9]. Till recently, manual labor was employed for this purpose. Some advancements have been taken place in the area and when survey is carried out through the research papers and it is understood that some new way of doing the things are reported as seen in Fig. 1. Here developed a rover that can climb tree to spray the Bordeaux mixture onto Areca-nut fruits. The product could do the intended work but it did not quite serve the purpose of reducing the tediousness. Researchers have built a wonder climber as shown in Fig. 1, but the product could

Fig. 1 Tree climber [10]





Fig. 2 Concept of proposed system

do the intended work and serve the purpose, firstly, it was quite difficult to control the machine from below.

Secondly, it is required to be removed from one tree and then clamped to another every time. Considering the number of trees in one farm, this still remained a tedious process.

Instead of using such a machine, a multi-rotor drone for this purpose proves to be a better solution of the problem. This is majorly due to the fact that drones can easily maneuver from one tree to another unlike the climbing machine.

Thus automating these tasks is of great scope in the modern world. To automate the above discussed system development, the following concept is proposed as in Fig. 2.

To initiate the work on drone, it is important to develop a skeleton and layout of the drone as shown in Fig. 3.

2 Proposed System

The following block diagram shows how the above described system can be implemented.

This block diagram shown in Fig. 4 represents the design of our complete system. The operator uses a radio transmitter to control the quad copter dynamics and navigates it manually. PWM signals are received by the radio receiver and it is forwarded

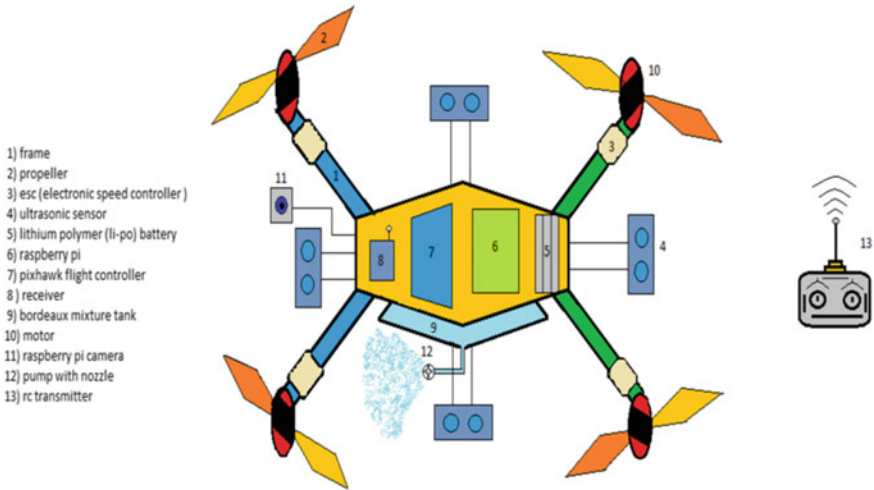


Fig. 3 Drone layout diagram

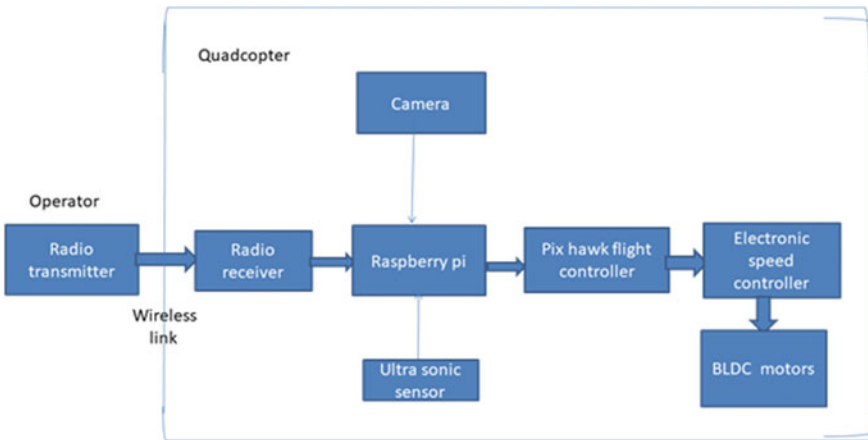


Fig. 4 Block diagram

to the Raspberry Pi. The Raspberry Pi detects the presence of the obstacle if any and commands the PixHawk to take the necessary action overriding the operator’s input. If no obstacle is detected, the signal received from the transmitter is forwarded to the PixHawk flight controller. The flight controller sends the speed information of each motor to its corresponding ESC thus performing the necessary action. The Raspberry Pi sends the appropriate roll and pitch triggers to the PixHawk controller via a Mavlink connection established. These commands are sent as RC override commands.

The design specification for the implemented drone as follows:

S. No.	Parameters	Value/capacity
1	Full dimension ($L \times W \times H$)	220 × 600 × 250
2	Weight	3.5 kg
3	Power source	Battery power
4	Pump discharge 1 min^{-1}	1.75
5	No of nozzles	4
6	Tank capacity	500 ml

3 Methodology Adopted

- (a) Development of a quadcopter drone with a payload capacity of about 500 g.
 - Design of the motor propeller combination.
 - Assembling the basic quad copter drone structure with all necessary components on it.
 - Flashing the ArduPilot firmware on the PixHawk flight controller.
 - Setting up the PixHawk flight controller and mounting it on the frame.
 - Flight test of the quad copter system.
- (b) Implementation of the circular trajectory tracing feature into the existing flight controller.
 - Analysis of the ArduPilot software architecture.
 - Analysis of the various possible design methods to implement circular trajectory tracing feature.
 - Implementation of the circular mode.
 - Interfacing Raspberry Pi and the PixHawk controller.
 - Testing the system.
- (c) Implementation of the spray mechanism.
 - Design of the spray mechanism.
 - Analysis and testing the working of a relay.
 - Assignment of one channel for the spray activation switch.
 - Interfacing PixHawk with relay.
 - Testing the system and mounting it on the quadcopter frame.
- (d) Implementation of the live video transmission system.

Analyses of the various possible methods implement a live video transmission system

 - Interfacing Raspberry Pi camera with Raspberry Pi.
 - Installation of GStreamer on a windows platform.
 - Testing and performance analysis of the system.

4 Structure of the Arducopter System

ArduCopter is an open-source multi-copter unmanned aerial vehicle platform community based on the Arduino platform. It is an advanced autopilot system for multi-copters, helicopters, and other rotor vehicles. ArduCopter supports multiple hardware platforms among which Pixhawk and APM are popular. We have employed a Pixhawk 2.4.8 hardware unit.

The structure of Ardupilot shown in Fig. 5 has been divided into main five parts.

- Hardware abstraction layer (AP_HAL)
- Shared libraries
- Vehicle code
- Tool directives and
- External support (i.e., maylink drone kit).

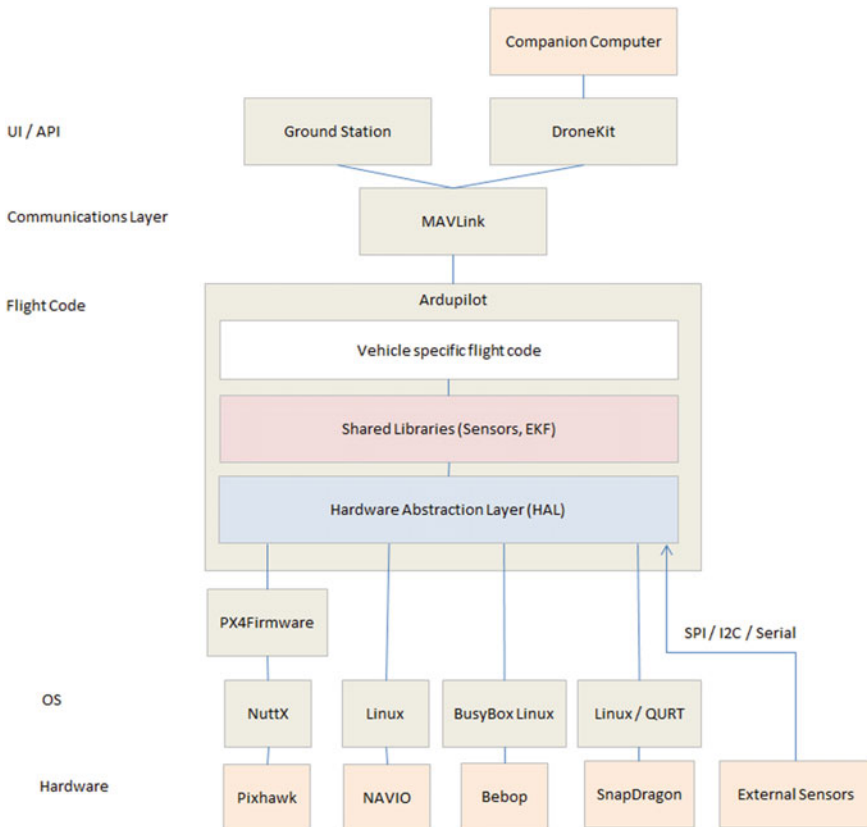


Fig. 5 Architecture of an Arducopter-based system [11]

5 Software Architecture

The Ardupilot software is designed in a modular approach for easier readability and understanding. The following schematic shown in Fig. 6 shows a high-level architecture of the complete software.

The blue box at left side justifies how sensor driver runs in the background thread. The collected data from the sensor is converted into standard form and then held within buffers within driver.

The main thread of the vehicles code runs at the frequency of 400Hertz for copter. It is able access the latest data available in driver’s front end. To calculate the latest altitude estimate, the AHRS/EKF would pull the latest accelerometer, gyro and compass information from the sensor at driver’s front end. With the help of I2C or SPI, drivers must run in the background thread so that the high rate communication with sensor

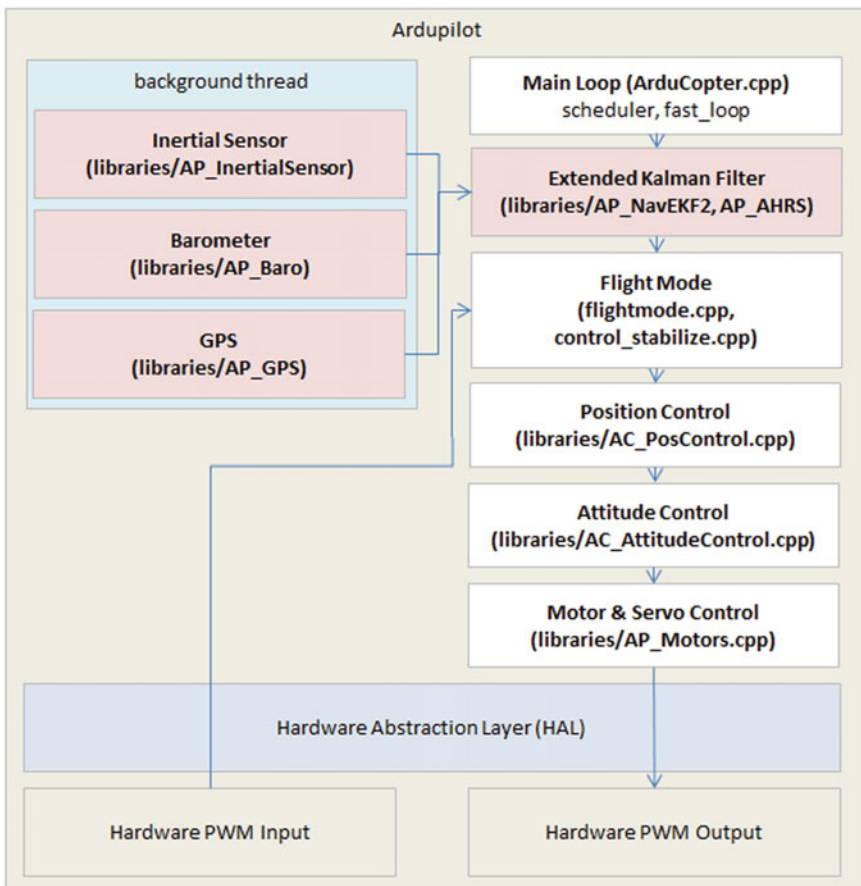


Fig. 6 Software architecture [11]

does not affect the performance in the main loop, but for driver with serial interface, it is safe to run in the main loop thread because the underlying serial driver itself collects data in the background includes a buffer.

6 Flight Modes in Arducopter

Arducopter has about twenty different flight modes. The vehicle behaves very differently in each of these modes. These modes support different types of stabilization and autopilot features. These modes can be switched on and off via a dedicated channel from the transmitter during flight.

Few of the commonly used modes are:

- **Stabilize mode:** This mode allows the drone to be controlled manually. It performs self-leveling of the roll and pitch axis.
- **Loiter mode of operation:** Here, it automatically tries to maintain the present location, altitude and heading. Here, pilot may fly the copter in mentioned loiter mode as if it is working in manual flight mode. When sticks get released, vehicle will slow down to stop and hold position. It uses GPS as well.
- **Altitude hold mode:** Here, copter itself maintains a consistant altitude while allowing roll, pitch, yaw to be controlled normally.
- **Circle mode:** Circle mode allows the UAV to orbit appoint located circle radius in front of the vehicle with its nose pointed at the center autonomously.

7 Design of the Trajectory Tracing Algorithm

We have proposed an economical approach to design a trajectory tracing algorithm. This approach uses companion computer that gives navigation commands to flight controller. The diagram shown in Fig. 7 gives the basis for designing circular motion.

The drone is altitude locked and the spray mechanism is activated using a switch. The distance sensor is mounted in the front along with a Gimbal arrangement. The drone is set to constant YAW and ROLL rate that achieves a circular motion around the tree. The real-time data from the ultrasonic sensors controls the pitch for the drone and maintains the drone at a constant distance from the tree.

8 Live Video Streaming from Raspberry PI

In order to establish a video streaming from Raspberry Pi camera to the PC, one should be able to connect these 2 systems where one will be acting as a server and client. Here, Raspberry Pi will be acting as a server which will stream the data that are received in the PC. Since the video transmission is done wirelessly, the IP address of

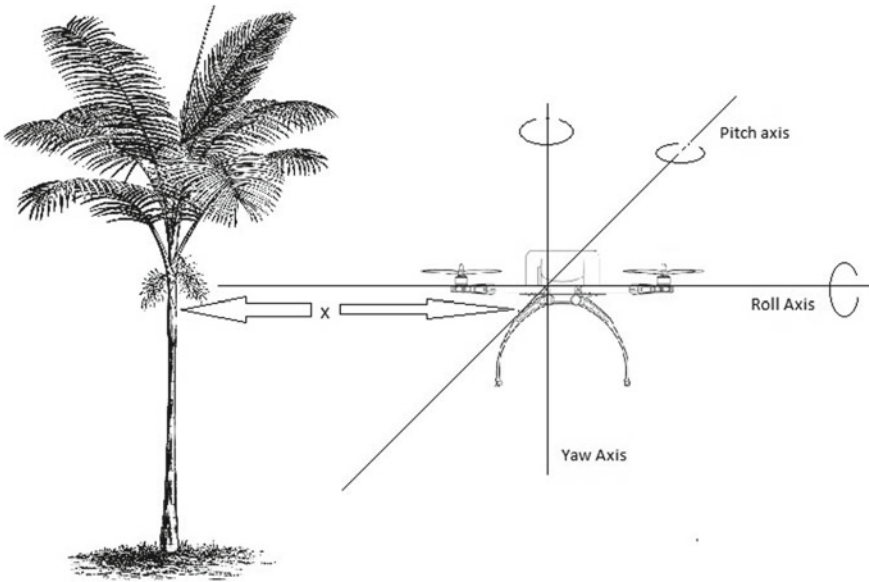


Fig. 7 Circular motion [3, 12]

pi is noted and modified. Through the Wi-Fi, all these devices are connected where an ad hoc network is formed [13, 14].

Basically, there are three methods to stream a video from Raspberry Pi to PC,

- i. RTP/RTSP flow with VLC.
- ii. GStreamer.
- iii. Sending a raw h.264 flow through net cat to a named pipe.

The first method uses the VLC command tool for as CVLC in the program on raspberry side and streaming RTSP over RTP on port 8554. The scripts on raspberry side and vlc on the windows side are launched. The disadvantage here is the latency is not really better so we can go for second method.

Among this, second method is followed in our case which uses GStreamer for streaming the live video from Raspberry Pi. This method is chosen due to the reason that first method has glitches more than 5 ms and is very difficult to minimize this [15,16].

9 Results and Analysis

From last one decade, there have been number of examples of applications of UAVs or drones in monitoring of environment and precision farming, etc. However, there are still some crucial limitations related to UAVs including high initial costs, platform

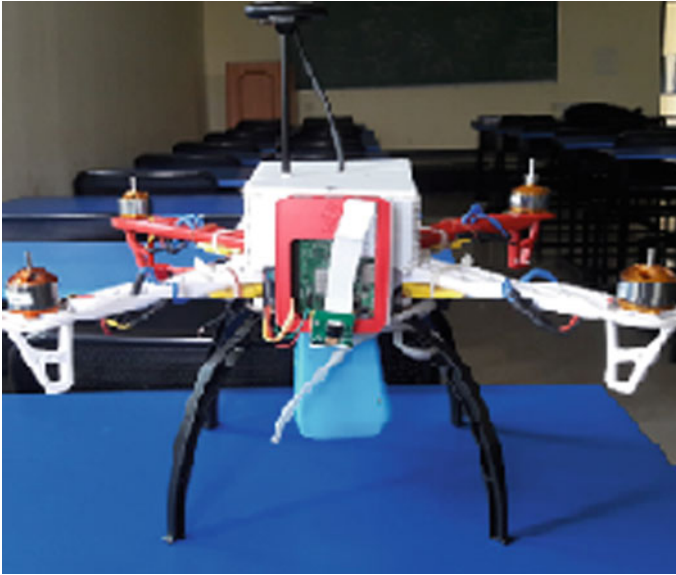


Fig. 8 Drone with a payload capacity of 500 g

reliability, sensor capability, and lack of standard procedures to process large volume of data. Keeping this point of objective to help farmers to reduce their work in the fields also to reduce human risk that they may fall. A quad copter can achieve vertical flight in a stable manner and are used to monitor. The cost has been reduced by the technological advancements and at the same time, system performance has improved. The implemented drone with payload capacity is as shown in Fig. 8 and drone working is demonstrated in Fig. 9.

10 Conclusion

The user-friendly agricultural drone has been implemented successfully for helping farmers. The developed semi-autonomous system is a fungicide spraying drone for agricultural purpose particularly for areca nuts during monsoon seasons to avoid the fungus attack, which helps the farmer to spray the fungicide onto every tree so that it reduces his work which can evenly spray all over his farm. The only work of the farmers will be only to take the drone up to 50 feet by using a RC transmitter after that it will autonomously spray the fungicide all over the farm. We have a Raspberry Pi camera attached which captures a live video transmission which helps the farmers to increase the visualize that helps him to judge whether the part of areca nut is sprayed or not, so that fungicide is saved and fully sprayed.



Fig. 9 Demonstration of a drone working

Firstly, our proposed drone is capable of maneuvering in a much stabilized manner and maintaining the altitude lock at 50 ft. Next the live video transmission is captured from Raspberry Pi camera. Once the altitude is locked to our requirement, the circular trajectory motion is activated which will make the drone to move around a tree and simultaneously spraying system is activated for spraying the fungicide. Presently, we have achieved to lift the drone with the payload capacity of our system is 500 g.

Acknowledgements Author acknowledges the Department of Ocean Engineering at IISc, Bangalore, for funding support to carry out the project.

References

1. P. Tripicchio, Towards smart farming and sustainable agriculture with drones, in *International Conference on Intelligent Environments (IE)* (2015)
2. R.F. Maia, I. Netto, Precision agriculture using remote monitoring systems in Brazil, in *IEEE Global Humanitarian Technology Conference (GHTC)* (2017)
3. D. Murugan, IEEE, Development of an adaptive approach for precision agriculture monitoring with drone and satellite data. *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*

4. F.N. Murrieta-Rico, Resolution improvement of accelerometers measurement for drones in agricultural applications
5. P. Tripicchio, M. Satler, Towards smart farming and sustainable agriculture with drones, in *2015 International Conference on Intelligent Environments*
6. M. Reinecke, T. Prinsloo, The influence of drone monitoring on crop health and harvest size, in *1st International Conference on Next Generation Computing Applications (NextComp)* (2017)
7. H.N. Saha, N. Saha, Recent trends in implementation of Internet of Things—A review, in *IEEE 7th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON)* (2016)
8. Zhou et al., Mapping paddy rice planting area in rice-wetland coex-istent areas through analysis of Landsat 8 OLI and MODIS images. *Int. J. Appl. Earth Obs. Geoinf.* **46**, 1–12 (2016)
9. Y. Shao, R.S. Lunetta, B. Wheeler, J.S. Iiames, J.B. Campbell, An evaluation of time-series smoothing algorithms for land-cover classifications using MODIS-NDVI multi-temporal data. *Remote Sens. Environ.* **174**, 258–265 (2016)
10. A.G. Deepak, Development of an adaptive approach for precision agriculture monitoring with drone and satellite data. *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.* **10**(12) (2017)
11. H. Shakhatareh, A. Sawalmeh, Unmanned aerial vehicles: a survey on civil applications and key research challenges. *IEEE Dig. Explore* **7**, 48572–48634 (2019)
12. S. Agarwal, D. Singh, An adaptive statistical approach for non-destructive underline crack detection of ceramic tiles using millimeter wave imaging radar for industrial application. *IEEE Sensors J.* **15**(12), 7036–7044 (2015)
13. P. Patel, Agricultural drones finally cleared for take-off news. *IEEE Spectrum* **53**(11), 13–14 (2016)
14. T. Pobkrut, T. Eamsa-Ard, T. Kerdcharoen, Sensor drone for aerial odor mapping for agriculture and security services
15. D. Murugan, A. Garg, T. Ahmed, D. Singh, Fusion of drone and satellite data for precision agriculture monitoring, in *Proceedings of 2016 IEEE 11th International Conference on Indian Information System*, Dec 2016
16. M. Podhradsky, C. Coopmans, N. Hoffer, Improving communication security of open source UAVs: encrypting radio control link, in *International Conference on Unmanned Aircraft Systems (ICUAS)* (IEEE, 2017), pp. 1153–1159

Underwater Marine Life and Plastic Waste Detection Using Deep Learning and Raspberry Pi



Rahul Hegde, Sanobar Patel, Rosha G. Naik, Sagar N. Nayak, K. S. Shivaprakasha, and Rekha Bhandarkar

1 Introduction

Machine learning for object detection has been used since a long time. Object detection is an application of computer vision. In this method, a target object is to be detected by means of a computer algorithm, which is the first phase of object detection. The object is then localized and the class of the object is determined. Several algorithms have been developed to give a better result in training an object detector. Each algorithm has its own advantages and disadvantages. In this paper, the single shot detector (SSD) MobileNet quantized V2 model is used for the training of object detector to detect marine life and underwater plastic waste.

Waste management has been a major issue for mankind since ages. In the current times clearing out accumulated waste is a major challenge. Clearing of underwater waste raises the difficulty of the situation. Detection of plastic waste in an underwater environment is difficult due to challenges such as reduced visibility, the water current, and the nature of the plastic waste. Computer vision has aided in this process as it provides algorithms for effective image processing and machine learning.

The rest of the paper is organized as follows: In Sect. 2, a brief description of the literature survey is presented. Section 3 presents a brief note on SSD is provided. A description to the system architecture of the project is given in Sect. 4 that outlines the major components used in making of the system. The results and tabulation of the performance of the object detector are presented in Sect. 5. Section 6 gives the concluding remarks of the paper. Section 6 provides areas for future work.

R. Hegde (✉) · S. Patel · R. G. Naik · S. N. Nayak · K. S. Shivaprakasha · R. Bhandarkar
Department of Electronics and Communication Engineering, NMAM Institute of Technology,
Nitte (Visvesvaraya Technological University, Belagavi), Udupi, India

K. S. Shivaprakasha
e-mail: shivaprakasha.ks@nitte.edu.in

R. Bhandarkar
e-mail: rekhabhandarkar@nitte.edu.in

2 Literature Survey

Deep learning algorithms such as You Only Look Once v3 (YOLOv3), Faster Region—Convolution Neural Network (Faster R-CNN), Viola Jones, and SSD were surveyed for their efficiency and robustness for the project. The use of YOLOv3 for underwater object detection was outlined in [1]. YOLOv3 is a robust algorithm for object detection in real-time embedded system. Its implementation in real-time object detection is discussed in [2–4]. YOLOv3 had a limitation in detecting small objects. Faster R-CNN [5, 6] overcame this disadvantage in this field but lacked in speed at which it detected objects in a real-time environment. SSD algorithm [7] is used in real-time object detection as it had high accuracy but lacked in its ability of identifying similar objects. It was faster than Faster R-CNN in its detection speed but slower than YOLOv3 algorithm. SSD algorithm was decided as Raspberry Pi was the target device. Raspberry Pi is a resource constrained device in terms of power and processing ability when it came to deep learning.

Tutorials on deep learning and object detection by Rosebrock [8, 9] and Edge Electronics [10, 11] provided a basis to understand machine learning. Tutorials on basics of machine learning and deep learning were obtained from ‘Codebasics’ [12] and deep learning with applications using Python by [13] also provided useful information on basics of machine learning.

A quantized model of a SSD MobileNet neural network [14, 15] was used in the training process. In general, SSD models were pre-equipped with VGG-16 network. MobileNet, on the other hand, proved to be useful when deployed in resource constrained devices. The process of model quantization reduced the model size and made it fit to be deployed in resource constrained devices [16].

3 Single Shot Detector

SSD here uses MobileNet as its base network for image classification. Convolutional feature layers that gradually reduce in size are added to its end so that it is able to predict object at different scales by an aspect ratio [1]. The convolution filters are applied to feature maps to predict the confidence score of a particular type of object and to obtain the offset coordinates of the bounding boxes. This enabled feature extraction at multiple scales and progressively decreased the size of the input to each successive layer [2].

MobileNet is primarily used as a classification layer during the training process of images for object detection. It uses depthwise separable convolutions. Depthwise separable convolution uses a combination of depthwise convolution and 1×1 pointwise convolution. In MobileNetV2 model, it uses three layers, first one being 1×1 convolution, second one being 3×3 depthwise convolution, and last being 1×1 convolution layer. First and second use rectified linear unit 6 (ReLU6) activation function. The function of the depthwise convolution is to apply a single filter to

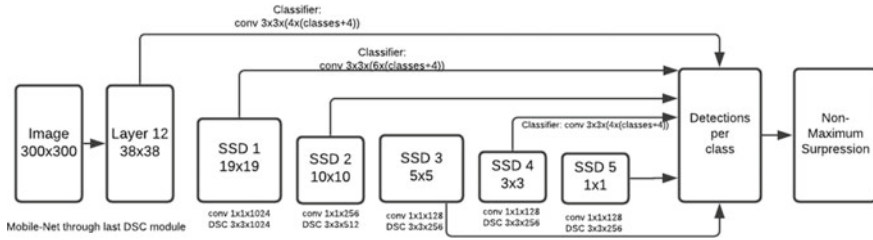


Fig. 1 Layers of SSD architecture with MobileNet for image classification

a single input channel. The pointwise convolution then takes over to combine the output of the depthwise convolution. In the first layer of SSD MobileNet, there are a total of $38 \times 38 \times 4$ predictions done with four predictions done per cell. It utilizes convolution 3×3 layers for object detection purposes as it detects both the relative location and the score of objects. It has five more layers called as extra feature layers added (in Fig. 1, SSD 1, SSD 2, SSD 3, SSD 4, SSD 5 represent extra feature layers). It uses six predictions per cell contrary to the four predictions per cell of the first layer for the next three layers and four predictions per cell for the next three convolution filter layers. Each layer applies a combination of depthwise convolution and pointwise convolution. SSD model uses multibox technique to obtain the bounding boxes. Non-maximum suppression is the technique used in order to filter out the large number of bounding boxes created during the detection process and to keep the highest predicted scores per class [17]. The model here is also quantized. The quantization process removes unwanted connection in the layers to improve accuracy and also reduce the weight of the trained model. The quantization used is unit8 quantization.

Figure 1 shows SSD MobileNet [14] layers. It replaces the Visual Geometry Group 16 (VGG16) neural network used for image classification. MobileNet is faster in detection speed and smaller in size when compared to VGG net [18].

The learning rate during the training process for the model was taken as 0.004. The batch size was taken as 24. It was taken to accelerate the rate of learning and reduce memory consumption of the graphical processing unit (GPU). Weighted smooth l1 loss function is used for localization loss and weighted sigmoid was the loss function used for classification loss. Weighted smooth l1 loss functions are used to steady the gradients. The box predictor uses convolution parameters such as l2_regularizer in the regularizer category. L2_regularizer balances the weights in the model during training. The activation function used is ReLU6. It is used as it has a good convergence as it prevent gradient from vanishing (vanishing gradient problem), thereby enabling faster learning rate. The feature extractor also uses l2_regularizer. Root mean square prop (RMSprop) optimizer was used as the optimizer during the training process. RMSprop optimizer balances between learning rate and the gradient descent.

4 General System Description

The main function of the system is to perform object detection of plastic bottles, plastic bags, bottle caps, and mackerel in an underwater environment. The image is captured by means of Raspberry Pi camera. The captured image is then processed by a Raspberry Pi. The model and labels files generated from the training process to detect the desired class of object in the image are stored in the Raspberry Pi. The object detection script is used to detect the target object in the image and draw bounding boxes around it. The image is captured at regular intervals of 2 s. Figure 2 gives the basic block diagram of the project. Figure 3 shows the general flow of the project.

4.1 Raspberry Pi Camera

The Raspberry Pi camera is a 5 mega pixel camera. It is specifically suited for use with a Raspberry Pi. It captures images of the surrounding underwater environment and sends the images to the Raspberry Pi for the detection of the target object in the image.

4.2 Raspberry Pi 3B+

Raspberry Pi 3B+ looks after the processing and control tasks. It receives the image from the Pi camera and passes it through the object detection algorithm. The object detection algorithm detects the target object and draws bounding boxes around the detected object. The power requirements of the Raspberry Pi are 5 V, 2 A.

Fig. 2 General block diagram of the system

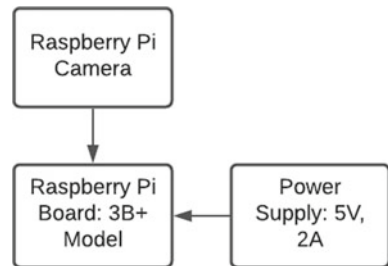
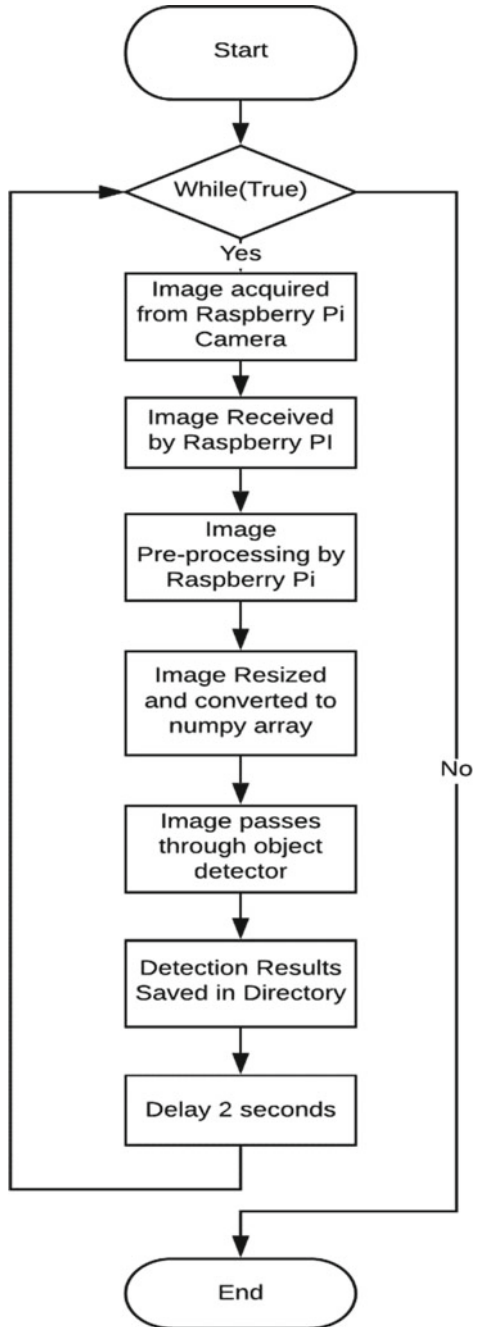


Fig. 3 General process flow diagram



4.3 Dataset Collection

The model was trained from images obtained from Google. The object detector had to detect four classes of images, namely plastic bottles, plastic bags, plastic caps, and mackerel. The images were resized to a dimension of 300×300 and then were augmented [19] to obtain the images for the training and test dataset. The dimensions of the image were 300×300 as it would reduce the training time for the object detector as well as it is a standard size for SSD models. The object detector was trained using SSD MobileNet V2 quantized model. The number of images collected for each class is as follows.

- Plastic bags: 2500
- Plastic bottles: 2500
- Plastic caps: 2500
- Mackerel: 2500

80% images were taken as training dataset and 20% were taken for validation dataset. Training of the model was done on Google Colab. In Google Colab, NVidia's Tesla GPU was used for training the model.

4.4 Implementation

- Raspberry Pi camera is used to capture the images from the surrounding environment.
- The captured image is sent to the Raspberry Pi for processing and detection of the target objects.
- The image once acquired by the Raspberry Pi is taken through the process of object detection. Here, the image is preprocessed which involves the resizing of the image, conversion into a numpy array. The images are resized to fit the dimensions on which the images were trained. Numpy arrays are an efficient way to handle and process image data.
- The image is then passed through the trained model to detect the object.
- Once the image has completed the process of object detection, the bounding boxes, classes, and scores are applied to the image and it is stored in a directory.
- The acquisition of image takes place at every 2 s.

5 Results and Analysis

5.1 Testing of Object Detection

- For testing the object detector, a simulation-based approach was followed.

- The images for testing in Fig. 4 were taken from Google images.
- The results in the tabulation were calculated from 2000 training images. It consisted of 500 images each of plastic bottle, plastic bag, plastic caps, and mackerel.
- Table 1 lists the particulars of the results of the object detector such true positive, true negatives, false positives, and false negatives for training images. The

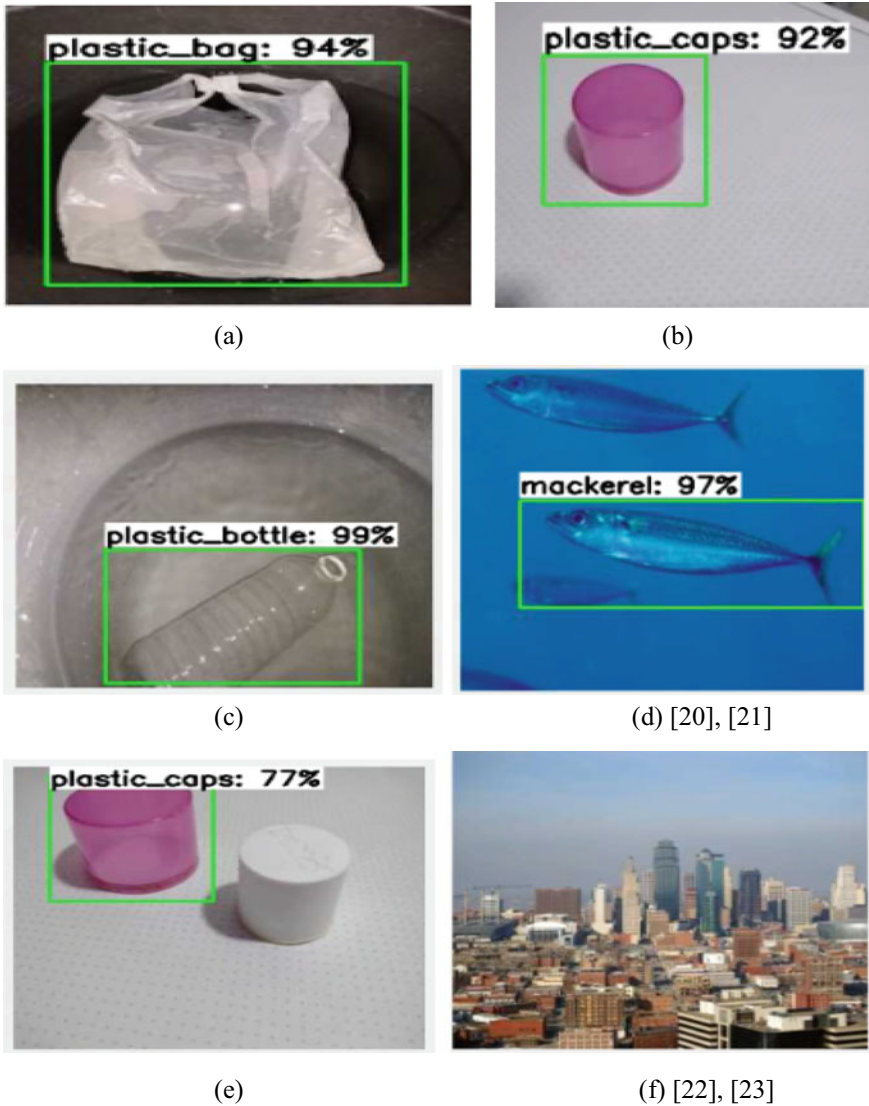


Fig. 4 Object detection results

Table 1 Results of training images for object detection

Class	True positive	True negative	False positive	False negative
Plastic bags	562	28	4	26
Plastic bottles	433	28	6	139
Plastic caps	518	24	21	471
Mackerel	547	28	3	104

Table 2 Parameters of object detection for training image dataset

Parameter (in %)	Plastic bags (in %)	Plastic bottles (in %)	Plastic caps (in %)	Mackerel (in %)
Accuracy	81.94	76.07	52.41	84.31
Misclassification rate	18.05	23.92	47.58	15.68
True positive rate	81.68	75.69	52.37	84.02
False positive rate	12.50	17.64	46.67	9.67
True negative rate	87.5	82.35	53.33	90.32
False negative rate	18.31	24.30	47.62	15.97
Precision	99.29	98.63	96.10	99.45
Recall	81.68	75.69	52.37	84.02
F1 score	85.89	85.65	84.88	85.95

threshold for recording the object detection results was kept at 0.5. Any object having a score less than 0.5 would not be detected. The detection results can vary based on the threshold score.

- Table 2 contains the results of true positive, true negatives, false positives, false negatives rates, the precision, accuracy, misclassification rate, recall, precision, and F1 score for training images.

Figure 4a–c shows the detection results of plastic bottle, cap, bags, and 4d [20, 21] shows the results of mackerel. Fig. 4e shows ambiguity is multiple detection of bottle caps and Fig. 4f [22, 23] shows true negative detection where none of the classes are falsely detected. The object detector was trained based on transfer learning approach [24, 25].

6 Conclusion

The simulation-based testing of the object detector was done by feeding 2000 images. The images were the ones used for training. The images were taken from Google. The object detector detected objects with an average score ranging from 70 to 90%. Detection of multiple bottle caps produced ambiguous results. The object detector

detects single objects with a high score. The detection results are summarized in Tables 1 and 2. Plastic bags, plastic bottles and mackerel have a precision in the range of 95–98% stating that the objects have a higher chance to be detected out of the total true cases. The $F1$ score of plastic bottle, plastic bags, plastic caps and mackerel are at range of 80–85%.

References

1. W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, A.C. Berg, Ssd: single shot multibox detector, in *European Conference on Computer Vision (ECCV)*, 2016, pp. 21–37. Understanding Real-Time SSD Multi Box Object Detection in Deep Learning [Online]. Available <https://towardsdatascience.com/understanding-ssd-multibox-real-time-object-detection-in-deep-learning-495ef744fab>. Accessed on 12.10.2019
2. B. Benjdira, T. Khursheed, A. Koubaa, A. Ammar, K. Ouni, Car detection using unmanned aerial vehicles: comparison between faster R-CNN and YOLOv3, in *2019 1st International Conference on Unmanned Vehicle Systems-Oman (UVS)* (Muscat, Oman, 2019), pp. 1–61; S. Jacobs, C.P. Bean, Fine particles, thin films and exchange anisotropy, in *Magnetism*, vol. III, ed. by G.T. Rado, H. Suhl (Academic, New York, 1963), pp. 271–350
3. Z. Lu, J. Lu, Q. Ge, T. Zhan, Multi-object detection method based on YOLO and ResNet hybrid networks, in *2019 IEEE 4th International Conference on Advanced Robotics and Mechatronics (ICARM)* (Toyonaka, Japan, 2019), pp. 827–832
4. R. Girshick, *Fast R-CNN*. [arXiv:1504.08083](https://arxiv.org/abs/1504.08083) (2015)
5. M. Manana, C. Tu, P.A. Owolawi, Preprocessed faster RCNN for vehicle detection, in *2018 International Conference on Intelligent and Innovative Computing Applications (ICONIC)* (PlaineMagnien, 2018), pp. 1–4
6. W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, A.C. Berg, Ssd: single shot multibox detector, in *European Conference on Computer Vision (ECCV)*, 2016, pp. 21–37
7. Image Classification with Keras and Deep Learning [Online]. Available <https://www.pyimagesearch.com/2017/12/11/image-classification-with-keras-and-deep-learning/>. Accessed on 16.01.2020
8. Turning any CNN image classifier into an object detector with Keras, TensorFlow, and OpenCV [Online]. Available <https://www.pyimagesearch.com/2020/06/22/turning-any-cnn-image-classifier-into-an-object-detector-with-keras-tensorflow-and-opencv/>. Accessed on 15.01.2020
9. How to run tensorflow lite for object detection on Raspberry Pi [Online]. Available <https://www.youtube.com/watch?v=aimSGOAU18Y>. Accessed on 10.01.2020
10. How To Train an Object Detection Classifier Using TensorFlow (GPU) on Windows 10 [Online]. Available <https://www.youtube.com/watch?v=Rgpfk6eYxJA&t=1167s>. Accessed on 10.01.2020
11. Deep Learning with Python, Tensorflow and Keras Neural Network for Image Classification [Online]. Available <https://www.youtube.com/watch?v=qEyEijUDOCA>. Accessed on 02.05.2020
12. N.K. Manaswi, N.K. Manaswi, S. John. *Deep Learning with Applications Using Python* (Apress, 2018)
13. Accelerated Training and Inference with the Tensorflow Object Detection API [Online]. Available <https://ai.googleblog.com/2018/07/accelerated-training-and-inference-with.html>. Accessed on 22.06.2020
14. A.G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, H. Adam, *Mobilenets: Efficient Convolutional Neural Networks for Mobile Vision Applications*. [arXiv:1704.04861](https://arxiv.org/abs/1704.04861) (2017)

15. B. Jacob, S. Kligys, B. Chen, M. Zhu, M. Tang, A. Howard, H. Adam, D. Kalenichenko, Quantization and training of neural networks for efficient integer-arithmetic-only inference, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 2704—2713
16. SSD object detection: single shot multibox detector for real-time processing [Online]. Available https://medium.com/@jonathan_hui/ssd-object-detection-single-shot-multibox-detector-for-real-time-processing-9bd8deac0e06. Accessed on 06.09.2020
17. Compressing Deep Neural Nets [Online]. Available <https://machinethink.net/blog/compressing-deep-neural-nets/>. Accessed on 16.06.2020
18. Keras ImageDataGenerator and Data Augmentation [Online]. Available <https://www.pyimagesearch.com/2019/07/08/keras-imagedatagenerator-and-data-augmentation/>. Accessed on 10.04.2020
19. Izuzuki [Online]. Available <https://www.izuzuki.com/>, CC BY-SA 3.0, <https://commons.wikimedia.org/w/index.php?curid=32245503>. Accessed on 01.12.2020
20. Izuzuki [Online]. Available CC BY-SA 3.0, <https://creativecommons.org/licenses/by-sa/3.0>, via Wikimedia Commons. Accessed on 01.12.2020
21. Leprecub at en.wikipedia, CC BY 3.0 [Online]. Available <https://commons.wikimedia.org/w/index.php?curid=6656059>. Accessed on 01.12.2020
22. Leprecub at en.wikipedia, CC BY 3.0 [Online]. Available <https://creativecommons.org/licenses/by/3.0>, via Wikimedia Commons
23. Tensorflow/ Models [Online]. Available <https://github.com/tensorflow/models>, Accessed on 01.12.2020
24. Tensorflow 1 Detection Model Zoo [Online]. Available https://github.com/tensorflow/models/blob/master/research/object_detection/g3doc/tf1_detection_zoo.md. Accessed on 01.12.2020
25. TensorFlow-Lite-Object-Detection-on-Android-and-Raspberry-Pi [Online]. Available <https://github.com/EdjeElectronics/TensorFlow-Lite-Object-Detection-on-Android-and-Raspberry-Pi>. Accessed on 01.12.2020

Land Cover Mapping Capability of Chaincluster, K-Means, and ISODATA techniques—A Case Study



Karthik and B. R. Shivakumar

1 Introduction

Remote sensing (RS) is the process of observing and monitoring the objects or areas of interest on the surface of the Earth from a remote distance by placing sensors on aircraft or geostationary satellites [1–3]. The field of remote sensing has been experiencing steady growth for the last five decades due to the advent of space technology. The current technology provides an option to assess RS data over a variety of spatial, spectral, and temporal resolutions. Recently, the public also has access to several advanced sensors, such as hyperspectral sensors, which can scan the Earth's surface using more than 240 spectral bands at extremely high spatial resolution. However, for large-scale mapping and change detection studies, medium resolution data are shown to be sufficient [4, 5] and hence, in this paper, we use Sentinel-2A sensor data.

One of the most commonly used techniques for extracting information from RS data is image classification. Classification of RS data can be broadly categorized based on training information into supervised and unsupervised classification. Supervised classifiers, as their name implies, are performed under the supervision of the analyst. Supervised classification techniques are commonly employed in two stages: (i) feature extraction and (ii) classification [1, 6, 7]. In the feature extraction step, the analyst collects the samples of spectral classes available on the data. This implies the analyst should be well aware of the study area's terrain and flora. This information is then used as the training data during the classification stage. One of the difficulties in performing supervised classification is the collection of training data. If the analyst fails to define and record data samples for an available spectral class, the efficiency of that spectral class is compromised and depends only on the effectiveness of the

Karthik (✉) · B. R. Shivakumar

Department of ECE, NMAM Institute of Technology, Nitte (Visvesvaraya Technological University, Belagavi), Udipi, Karnataka, India

classification algorithm [1, 2]. Supervised classifiers are further categorized based on whether the algorithm uses the spectral patterns of objects to classify as spectral classifiers [8, 9].

Unsupervised classifiers are commonly called as spectral identifiers and are basically employed when no *a priori* information about spectral classes is available. They are essentially computer-automated classification techniques in which the analyst is required to feed the system only a bare minimal input [3, 10]. In these classification techniques, numerical operations are conducted on the data to produce natural pixel groupings. This process is also termed as clustering [9]. The clustering technique is one of the most effective methods of partitioning the data in signal space as well as information extraction. Since these classifiers are completely based on software analysis of the data, they help detect new spectral classes that may not have been recorded previously, hence the name spectral identifiers. Field visits can provide insights into the terrain characteristics of the data when no information is available on the spectral features of the study area.

The unsupervised classification produces a thematic map with m spectral clusters which the analyst then merges to create a meaningful land use land cover (LULC) class. Prior knowledge about the terrain characteristics is required to merge spectral clusters to create a LULC classes. The LULC class may require several spectral clusters to be combined since, depending on the clusters formed, a class can constitute more than one cluster. This study has a twofold objective; (i) analyze the classification performance of chaincluster, iterative self-organizing data analysis technique (ISODATA), and k-means algorithms using Sentinel-2 RS data. (ii) Identify the spatially and spectrally dominant LULC classes on the study area by performing an accuracy assessment of the classified thematic maps.

The rest of the paper is organized as follows. Section 2 describes the data products used and the study area considered for experimentation. The unsupervised classification techniques used in this study are discussed in Sect. 3. In Sect. 4, we present the methodology followed in this study. The results obtained during this study are presented and discussed in Sect. 5. Lastly, the conclusions drawn from the results obtained are presented in Sect. 6.

2 Data Products and Study Area

2.1 Sentinel-2 Data

Sentinel-2 is an Earth observation mission from the Copernicus Program. The sensors placed on the satellite capture information at very high spatial resolution (10–60 m) over land and coastal waters. The mission is a constellation with twin satellites, Sentinel-2A and Sentinel-2B. The mission supports a broad range of services and applications such as agricultural monitoring, emergency management, land cover classification, and water quality. The Sentinel-2A carries a multispectral imager

(MSI) sensor which provides data at 13 spectral bands ranging from 10 to 60 m pixel size. The spectral bands are; Coastal Aerosol (443 nm), Blue (490 nm), Green (560 nm), Red (665 nm), Visible and Near Infrared-1 (VNIR-1) (705 nm), VNIR-2 (740 nm), VNIR-3 (783 nm), VNIR-4 (842 nm), VNIR-5 (865 nm), Short Wave Infrared-1 (SWIR-1) (940 nm), SWIR-2 (1375 nm), SWIR-3 (1610 nm), and SWIR-4 (2190 nm) [11]. The spatial resolution of bands Blue, Green, Red, and VNIR-4 is 10 m. Spectral bands VNIR-1, VNIR-2, VNIR-3, VNIR-5, SWIR-3, and SWIR-4 have a spatial resolution of 20 m, and Coastal Aerosol, SWIR-1, and SWIR-2 have 60 m pixel size.

2.2 Study Area

The study area of the Udipi Taluk boundary lies in the south-western part of Karnataka state, India, and is shown in Fig. 1. It is identified as belonging to the coastal agro-climatic zone by the Government of India [12]. The primary reason for the selection of this area is its heterogeneous nature. The study area experiences a blend of several LULC classes such as vegetation, water body, and agricultural fields. The data was acquired on January 12, 2020 which is the presummer season and is free from clouds. All 13 bands were used for creating the composite input image. The spatial resolution of the image is 10 m, the projection type of the data was set to geographic (Lat/Lon), and the spheroid and datum type was set to WGS-84.

3 Unsupervised Classification

In this section, we present a detailed discussion of the unsupervised classification techniques employed in this study.

3.1 Chaincluster

The chaincluster technique is one of the primitive unsupervised clustering techniques. It is basically a two-pass clustering technique. The operation of the algorithm under each pass is discussed below.

First pass: In the first pass, the algorithm strategically groups pixels in the spectral space. Each group of pixels so created is called a cluster. A mean vector is then constructed and associated with each cluster. The data is evaluated sequentially pixelwise from left to right as if in a chain. After one line is processed the next line is considered. This process continues until the number of pixels evaluated is greater than some prespecified value N . When the number of pixels processed becomes equal to N , the algorithm stops evaluation and assesses the nature of clusters produced till

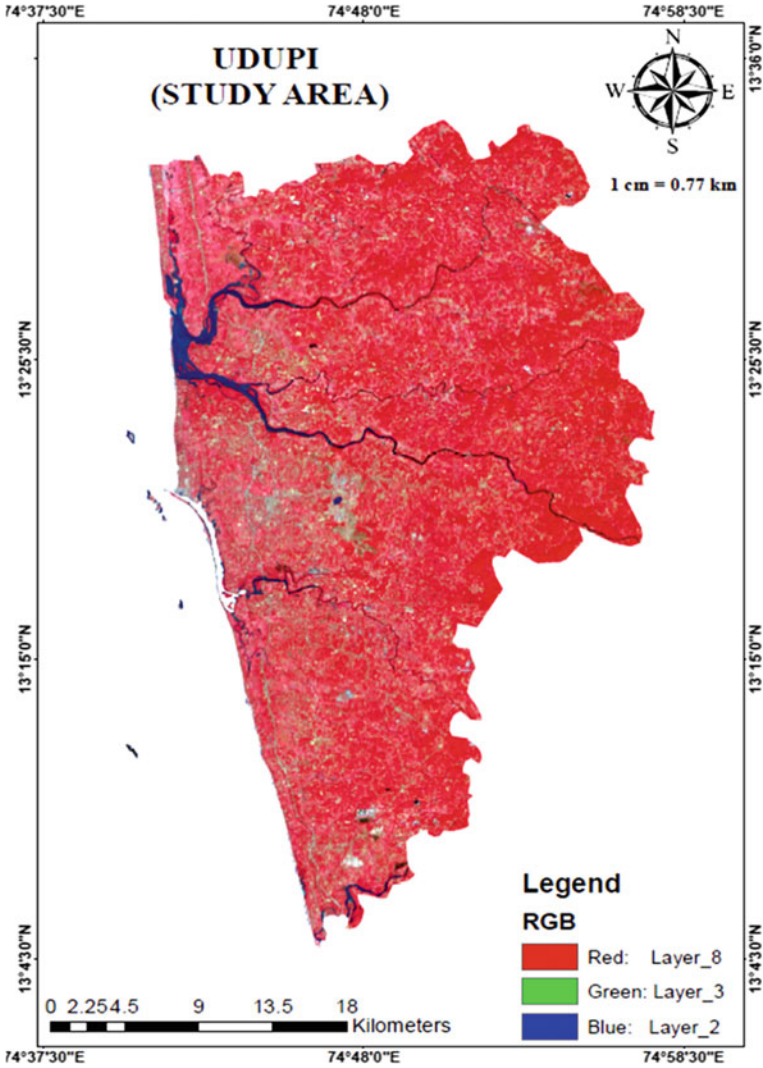


Fig. 1 Study area of Udupi Taluk boundary (Band combination [8 3 2])

that instant. If the spectral space at that instant has two clusters that are separated by a distance less than some value C , they are merged to create a single cluster [1]. A new cluster mean is created for this merged cluster, which is the weighted average of the two original cluster means, and the weight is the sum of the two individual weights. This process continues until the spectral space is modified such that no two clusters are separated by a distance C . The classifier then continues to evaluate the next pixel in line. This whole process is reiterated until the entire multispectral data is evaluated.

Second pass: In the second pass, the minimum distance-to-means classification technique is used for assigning all the pixels to one of the clusters created in Pass 1. The minimum distance-to-means algorithm does this by calculating Euclidean distance between the pixel under consideration and mean clusters of all clusters in spectral space. The pixel is then assigned to the cluster that creates the smallest Euclidean distance between itself and the mean vector of that cluster. This decision rule is applied to every pixel in the spectral space until all of the pixels are assigned to one of the clusters [1].

3.2 *K-means*

K-means is a clustering algorithm falling under the vector quantization technique and originates from the signal processing domain. It has been employed in data mining studies due to its exploratory nature [13–15]. By principle, the k-means technique behaves similarly to chaincluster technique by partitioning the spectral space and assigning pixels to the nearest cluster means. However, it differs in its operation in dividing the spectral space and has control over the total number of clusters to be created. The k-means clustering is performed in two steps.

- Step 1 This step is also called as the initialization step. The k-means algorithm specifically uses Forgy and random partitioning techniques for initialization [16]. The Forgy initialization selects k randomly selected pixels from the RS data as initial cluster means. On the other hand, the random partitioning initialization technique randomly assigns pixels to a cluster and updates the cluster means after each pixel assignment. This variant of the k-means technique may therefore be considered as a variant of the expectation–maximization technique with the assignment step referred to as expectation and the update step as maximization.
- Step 2 This step is further carried out in two steps. In the first step, the clustering technique assigns the next pixel in line to the nearest cluster. The nearest cluster is found by calculating the Euclidean distance between the pixel and the mean vector of the cluster. In the second step, after a new pixel is assigned to one of the clusters, the corresponding cluster’s mean vector is updated to incorporate the new pixel. These steps are repeated until every pixel in the image is assigned to one of the clusters.

The running time of the algorithm is generally $O(nkdi)$ where n is the number of d -dimensional vectors, k is the number of clusters, and i is the number of iterations needed until convergence [14]. The k-means algorithm partitions the spectral space in such a way that the squared error between the empirical cluster means and pixels in a cluster is minimal. Let μ_k be the mean vector of the cluster c_k . The squared distance between μ_k and the points in the cluster c_k is given by [3]

$$J(c_k) = \sum_{x_i \in c_k} x_i - \mu_k^2 \quad (1)$$

The goal of k-means is to minimize the sum of the squared error over all k clusters. i.e., [3]

$$J(C) = \sum_{k=1}^K \sum_{x_i \in c_k} x_i - \mu_k^2 \quad (2)$$

K-means clustering measures the separation between pixels and the cluster centers by calculating the Euclidean distance between them [17].

3.3 ISODATA

ISODATA is a self-organizing algorithm which is a reformation of the k-means clustering algorithm. At first, ISODATA arbitrarily assigns all its C_{\max} clusters along the n-dimensional vector which is running in feature space between very specific points. The standard deviation (σ_k) and mean (μ_k) of each input image band characterize a region in the multidimensional feature space. The initial C_{\max} mean vector location will vary to effectively divide the feature space.

First iteration: The initial C_{\max} mean vectors pass through the database with the upper left corner of the matrix as the initializing point. Each test pixel is then compared with the mean vectors of every cluster and depending upon the calculated distance the pixel is allocated to the nearest cluster [1]. ISODATA also uses Euclidean distance to measure the distance between the test pixel and the cluster's mean vector which is given by

$$L_i = \sqrt{x - m_i} \quad \text{for } i = 1, \dots, C_{\max} \quad (3)$$

where x is the test pixel vector which is then assigned to C_i if it produces the smallest L_i for x , i.e., test pixel x is assigned to cluster C_i if $L_i > L_j, \forall j = 1, \dots, C_{\max}$ and $i \neq j$.

At the end of the first iteration, a thematic map that consists of C_{\max} cluster is created.

Second to M th iteration: After the first iteration, for each cluster, based on the spectral locations for each test pixel, new mean vectors are calculated. This process is then repeated by comparing the test pixels with new cluster means and assigning them to the cluster for which it produces the smallest Euclidean distance. The process is iterated until convergence is achieved. Convergence is achieved when there are no minimum changes in class assignments between two consecutive iterations or when the maximum number of iterations completed [1]. The final output of the

ISODATA technique is an image with C_{\max} clusters in it. The recoding operation can be performed by the analyst to create a classified image.

4 Methodology

In this section, the methodology employed in this study is presented. The overall methodology is divided into three phases: preprocessing, classification, and post-processing. The dispute about these phases is as follows.

4.1 Preprocessing

Sentinel-2A data of Udupi Taluk boundary was acquired from the EarthExplorer Web site [18]. The spectral bands of the data are originally at a different spatial resolution as discussed in Sect. 2. A composite image with spectral bands that are at various spatial resolutions cannot be produced. Hence, the spectral bands of higher spatial resolution than 10 m are then resampled using ARCGIS 10.5[®] RS and GIS applications to get 10 m spatial resolution. The spectral bands are then layer-stacked to produce 10 m spatial resolution multispectral imagery. This data is then used in classification.

4.2 Classification

As already discussed in Sect. 3, the study classifies the considered sample area using three unsupervised classifiers: chaincluster, k-means, and ISODATA. The role of analysts is limited, as the analysis emphasizes unsupervised classification algorithms. However, details such as the estimated number of clusters, convergence, and the maximum number of merges per iteration, needs to be given by the analyst. Table 1 displays the values of these input parameters used for classification.

4.3 Post-classification Processing

On the output cluster maps, two post-classification processing operations are carried out to construct meaningful classes as well as to quantify the results. These operations are described as follows.

Recoding: A land cover on the surface of the Earth can be scattered with different proportions and boundary shapes around the study areas. The clustering methods thus produce distinct clusters, while they essentially reflect a single land cover. These

Table 1 Parameter configuration used for conducting classifications

Parameter	Chaincluster	K-means	ISODATA
No. of clusters	36	36	36
Convergence threshold	–	0.950	0.950
Max. no. of iterations	–	10	10
Merges per iteration	–	–	1
Std. dev. threshold to split the cluster	–	–	5.0
Euclidean distance threshold to merge clusters	–	–	4.0
Radius distance	2500	–	–
Merging threshold	100	–	–
Number of pixels between merging	10,000	–	–

clusters will appear as they are on the output thematic maps. Since land cover maps ought to reflect the actual land cover on the Earth's surface, it is important to merge clusters that represent a class of land cover. This mechanism is called recoding.

Accuracy assessment: The thematic maps produced in the classification technique only provide qualitative information of the LULC classes in the study area. To quantify the results and assess the effectiveness of the overall process, the spatial accuracy assessment method is employed on the output thematic maps. An independent set of randomly selected pixels from the classified image are selected and verified for their correctness. This study aims to obtain 85% classification accuracy at 95% confidence with a $\pm 4\%$ error margin. The minimum number of pixels required to satisfy this constraint is given by [2];

$$N = \frac{4P(1 - P)}{e^2} \quad (4)$$

where N is the number of pixels required for the accuracy assessment and P is the expected map accuracy and e is the acceptable error margin.

For the expected classification accuracy of 85%, the value of N is calculated to be 319. Several accuracy assessment parameters were considered for the accuracy assessment of thematic maps produced by the classifiers. It is not within the scope of this paper to study to discuss the accuracy assessment and its parameters. We suggest the readers refer [2] for more information on accuracy assessment.

5 Results and Discussions

To identify the LULC classes extracted in the study areas, the outcomes of the clustering techniques were subjected to a stringent recoding technique. We have used other sources, such as Google Maps [19] and Wikimapia maps [20] to understand

the terrain characteristics of the study area under consideration. We find five of Anderson's level-1 classes in the study area by close review of all three classification output maps; they are: vegetation, water body, built-up, barren land, and agricultural fields. Of the five classes identified, three classes form the spatially dominant classes. They are vegetation, agricultural fields, and barren land. The other two classes water body and built-up form the spatially subservient classes.

Further, based on the spectral characteristics of the LULC classes, they can be categorized as spectrally dominant and spectrally subservient classes. Spectrally dominant classes are those which have a clear distinction in their spectral characteristics from others and are easy to separate in spectral space. On the other hand, spectrally subservient classes are not spectrally independent and normally overlap each other in spectral space. The spectrally dominant classes in the study area are vegetation and water body. The spectrally subservient classes are agricultural fields, barren land, and built-up. Since the study area is categorized under the coastal agro-climatic zone, vegetation forms the primary land cover type. It should be noted that this study only concentrates on the level-1 classes of Anderson's classification system.

Figure 2 indicates the output thematic map of the Udupi Taluk study area produced by the chaincluster technique after recoding, and Table 2 indicates the accuracy assessment results of the same. The chaincluster technique successfully identified and separated all the spatially dominant classes. With a user's accuracy of 84.91%, the vegetation class was the most efficiently extracted class. Out of the 212 vegetation class pixels, 180 pixels were correctly classified. Further, the algorithm also extracted the spectrally dominant classes very efficiently. The classifier extracted agricultural fields and barren land with mediocre user's accuracies. The classifier, however, entirely failed to extract the built-up class. It should be remembered that in the study area, the built-up class is both spatially and spectrally subservient. The overall classification accuracy and overall Kappa value are 69.28% and 0.4212 respectively.

Figure 3 indicates the output thematic map of the Udupi Taluk study area produced by the k-means clustering technique after recoding, and Table 3 indicates the accuracy assessment results for the same. K-means classifier extracted all LULC classes on the data successfully. Once again, vegetation and water body classes were extracted excellently with greater than 85% class accuracies. Agricultural fields and barren land classes were extracted with below average accuracies. It must be noted that the k-means algorithm extracted built-up class excellently with 87.50% of the user's accuracy. The overall classification accuracy and overall Kappa value produced by k-means clustering are 69.28% and 0.3987, respectively. The k-means technique extracted the spectrally dominant classes excellently while the spatially subservient classes were poorly extracted.

Figure 4 indicates the output thematic map of the Udupi Taluk study area produced by the ISODATA classification technique after recoding, and Table 4 indicates the accuracy assessment results for the same. The ISODATA technique, similar to the k-means technique, also successfully extracted all LULC classes from the results. Once again, the vegetation class is the most successfully extracted category. With 88.89% of

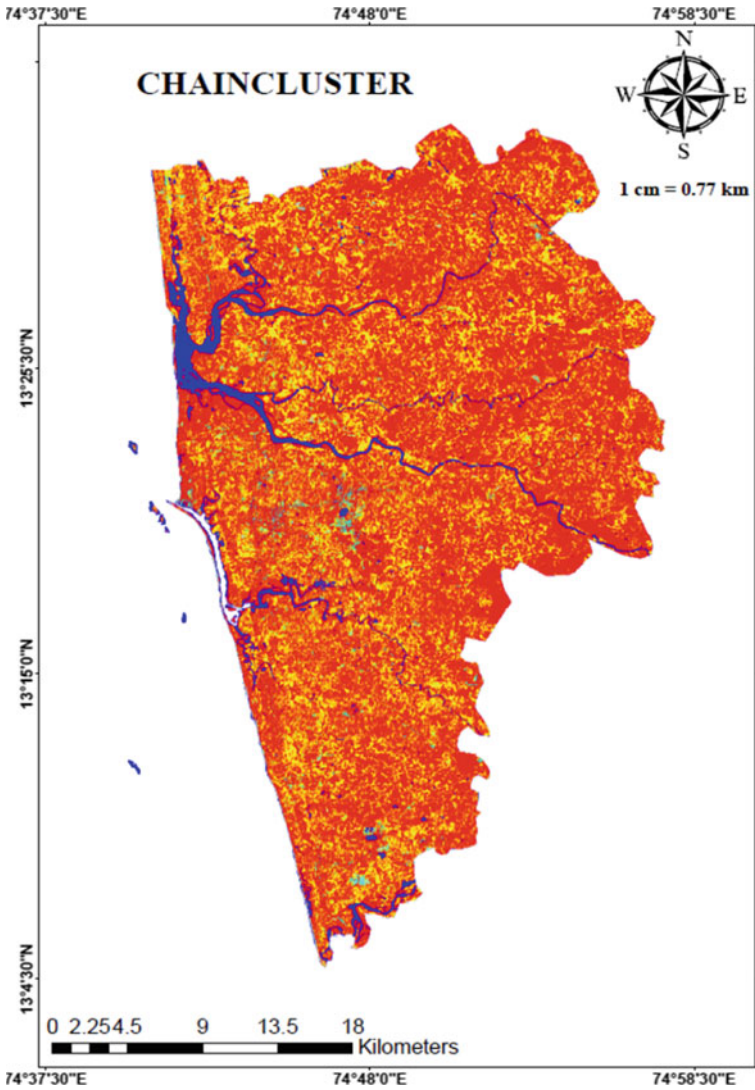


Fig. 2 Output thematic map produced by chaincluster technique (after recoding)

the user's accuracy, water body class was also retrieved excellently. The other three classes were extracted with average accuracies. The spectrally dominant classes were excellently identified by the ISODATA clustering technique. For ISODATA clustering, the overall classification accuracy and overall Kappa value are 75.86% and 0.5668, respectively.

Table 2 Accuracy assessment results of chaincluster technique

Class name	RT ^a	CT ^b	NC ^c	PA ^d (%)	UA ^e (%)	K_{hat} ^f
Vegetation	207	212	180	86.96	84.91	0.5701
Agricultural field	40	91	31	77.50	34.07	0.2461
Barren land	41	2	1	2.44	50.00	0.4263
Water body	10	14	9	90.00	64.29	0.6313
Built-up	21	0	0	–	–	0.0000
Totals	319	319	221	–	–	–
Overall classification accuracy = 69.28%						
Overall Kappa statistics = 0.4212						

^aRT Reference totals^bCT Classified totals^cNC Number correct^dPA Producers accuracy^eUA Users accuracy^f K_{hat} : Kappa value

6 Conclusions

Unsupervised classifiers have been consistently used in remote sensing studies and are particularly of a large advantage when the terrain information is not available. In this paper, we have analyzed three unsupervised classifiers in their ability to both identify and create meaningful LULC classes in the study area. All three classifiers have excellently extracted the spectrally and spatially dominant classes in the study area. However, a great difference is observed when it comes to extracting spectrally as well as spatially subservient classes. Particularly, the chaincluster technique failed to identify built-up class in the study area, even though it produced the same amount of overall classification accuracy as the k-means clustering technique. However, it should be noted that the overall Kappa value produced by the chaincluster technique is better than that produced by k-means, which indicates that the output thematic map produced by chaincluster is more efficient than that produced by the k-means technique. ISODATA clustering produced considerably better classification performance than both chaincluster and k-means techniques. The Kappa statistic for ISODATA is better by 0.14 compared to other two techniques and this is a great improvement in the overall thematic map quality. Hence, we conclude that the ISODATA classification technique is more suitable for identifying LULC classes than chaincluster and k-means techniques. Also, as per our second objective, the spectrally dominant and subservient classes were identified successfully. It would be interesting to test the ability of these classifiers applied to very high-resolution RS data such as hyperspectral imagery.

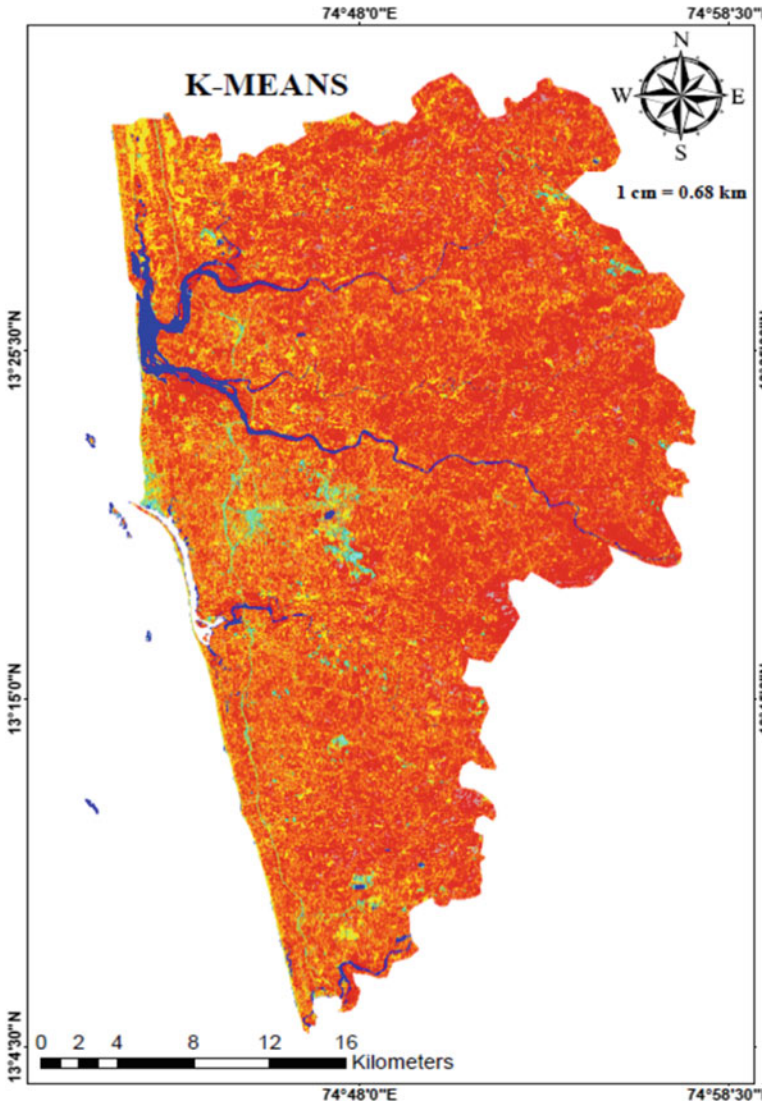


Fig. 3 Output thematic map produced by k-means technique (after recoding)

Table 3 Accuracy assessment results of k-means technique

Class name	RT ^a	CT ^b	NC ^c	PA ^d (%)	UA ^e (%)	K_{hat} ^f
Vegetation	220	209	178	80.91	85.17	0.5221
Agricultural field	39	82	24	61.54	29.27	0.1942
Barren land	35	11	4	11.43	36.36	0.2852
Water body	11	9	8	72.73	88.89	0.8849
Built-up	14	8	7	50.00	87.50	0.8693
Totals	319	319	221			

Overall classification accuracy = 69.28%

Overall Kappa statistics = 0.3987

^aRT Reference totals

^bCT Classified totals

^cNC Number correct

^dPA Producers accuracy

^eUA Users accuracy

^f K_{hat} Kappa value

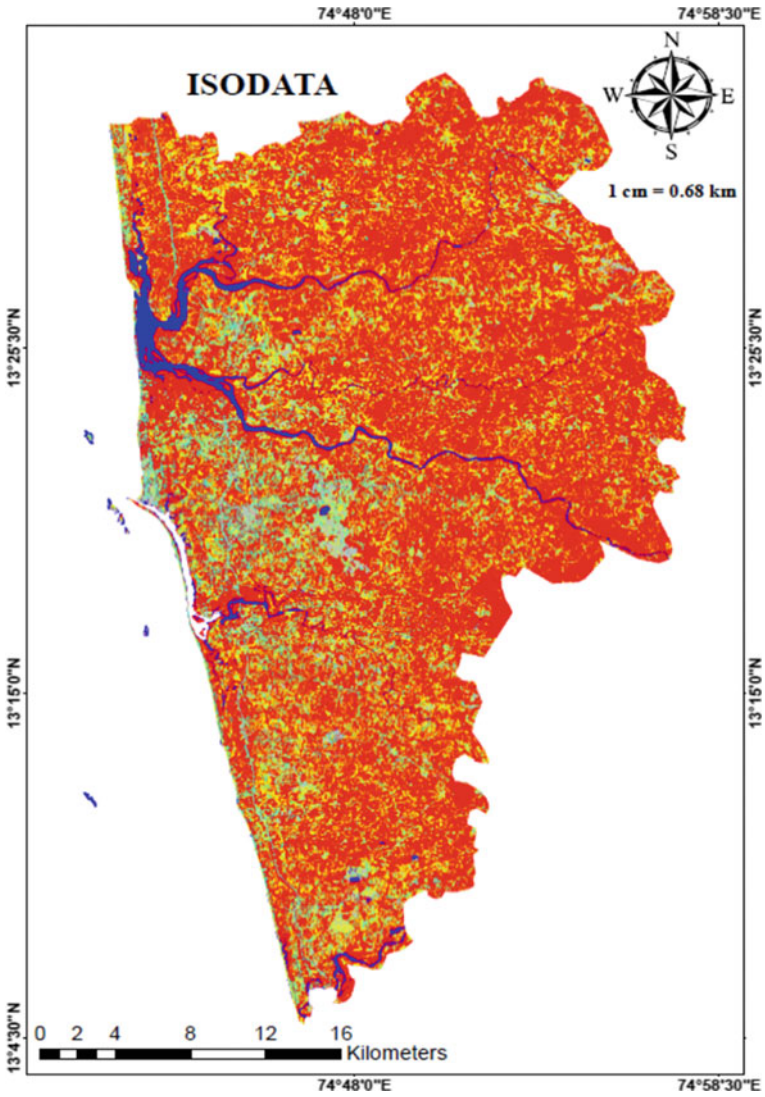


Fig. 4 Output thematic map produced by ISODATA technique (after recoding)

Table 4 Accuracy assessment results of ISODATA technique

Class name	RT ^a	CT ^b	NC ^c	PA ^d (%)	UA ^e (%)	K_{hat} ^f
Vegetation	207	197	185	89.37	93.91	0.8265
Water body	10	9	8	80.00	88.89	0.8853
Built-up	21	18	8	38.10	44.44	0.4053
Barren land	41	19	11	26.83	57.89	0.5168
Agricultural field	40	76	30	75.00	39.47	0.3080
Totals	319	319	242			
Overall classification accuracy = 75.86%						
Overall Kappa statistics = 0.5668						

^aRT Reference totals^bCT Classified totals^cNC Number correct^dPA Producers accuracy^eUA Users accuracy^f K_{hat} Kappa value

References

1. J.R. Jensen, *Introductory Digital Image Processing: A Remote Sensing Perspective*, 4th edn. (Pearson, 2016)
2. J.A. Richards, *Remote Sensing Digital Image Analysis: An Introduction*, 5th edn. 2013
3. B.R. Shivakumar, Study and analysis of pixel-based classification of remotely sensed data using different classifiers, in Thesis submitted to Visvesvaraya Technological University, 2020
4. M.A. Alrababah, M.N. Alhamad, Land use/cover classification of arid and semi-arid Mediterranean landscapes using Landsat ETM. *Int. J. Remote Sens.* **27**(13), 2703–2718 (2006). <https://doi.org/10.1080/01431160500522700>
5. R. Manandhar, I.O.A. Odehi, T. Ancevt, Improving the accuracy of land use and land cover classification of landsat data using post-classification enhancement. *Remote Sens.* **1**(3), 330–344 (2009). <https://doi.org/10.3390/rs1030330>
6. B.R. Shivakumar, S.V. Rajashekararadhya, Investigation on land cover mapping capability of maximum likelihood classifier: a case study on North Canara, India. *Proc. Comput. Sci.* **143**(January), 579–586 (2018). <https://doi.org/10.1016/j.procs.2018.10.434>
7. B.R. Shivakumar, S.V. Rajashekararadhya, An investigation on land cover mapping capability of classical and fuzzy based maximum likelihood classifiers. *Int. J. Eng. Technol.* **7**(2), 939 (2018). <https://doi.org/10.14419/ijet.v7i2.10743>
8. B.R. Shivakumar, S.V. Rajashekararadhya, Spectral similarity for evaluating classification performance of traditional classifiers, in *Proceedings of 2017 International Conference on Wireless Communication Signal Process. Networking, WiSPNET 2017*, vol. 2018-Janua, pp. 1999–2004 (2018), <https://doi.org/10.1109/WiSPNET.2017.8300111>
9. D. Lu, Q. Weng, A survey of image classification methods and techniques for improving classification performance. *Int. J. Remote Sens.* **28**(5), 823–870 (2007). <https://doi.org/10.1080/01431160600746456>
10. B.R. Shivakumar, S.V. Rajashekararadhya, Classification of landsat 8 imagery using Kohonen's self organizing maps and learning vector quantization. *Lect. Notes Electr. Eng.* **614**, 445–462 (2020). https://doi.org/10.1007/978-981-15-0626-0_35
11. J. Li, D.P. Roy, A global analysis of Sentinel-2a, Sentinel-2b and Landsat-8 data revisit intervals and implications for terrestrial monitoring. *Remote Sens.* **9**(9) (2017). <https://doi.org/10.3390/rs9090902>

12. T.V. Ramachandra, G. Kamakshi, B.V. Shruthi, Bioresource status in Karnataka. *Renew. Sustain. Energy Rev.* **8**(1), 1–47 (2004). <https://doi.org/10.1016/j.rser.2003.09.001>
13. D. Aloise, A. Deshpande, P. Hansen, P. Papat, NP-hardness of Euclidean sum-of-squares clustering. *Mach. Learn.* **75**(2), 245–248 (2009). <https://doi.org/10.1007/s10994-009-5103-0>
14. J.A. Hartigan, M.A. Wong, Algorithm AS 136: A K-means clustering algorithm. *Appl. Stat.* **28**(1), 100 (1979). <https://doi.org/10.2307/2346830>
15. P. Drineas et al., Clustering large graphs via the singular value. *Mach. Learn.* **56**(1–3), 9–33 (2004)
16. G. Hamerly, C. Elkan, Alternatives to the k-means algorithm that find better clusterings. *Int. Conf. Inf. Knowl. Manage. Proc.*, 600–607 (2002). <https://doi.org/10.1145/584792.584890>
17. S. Phillips, Reducing the computation time of the isodata and K-means unsupervised classification algorithms. *Int. Geosci. Remote Sens. Symp.* **3**(C), 1627–1629 (2002). <https://doi.org/10.1109/igarss.2002.1026202>
18. Earth Explorer. <https://earthexplorer.usgs.gov/>. Accessed 25 June 2020
19. Google Maps. <https://www.google.com/maps/>. Accessed 15 July 2020
20. Wikimapia. <https://www.wikimapia.org/>. Accessed 18 July 2020

Microcontroller-Based Control Circuit for the Automatic Orientation of Solar Panels in the Direction of Sun



A. Jayashree, Vidya Kudva, and A. G. Ananth

1 Introduction

The energy obtained from the sun is immense in quantity and is inexhaustible. The planet receives around 172 PW of solar power. This is many thousands of times of our present day requirement. It is very important to properly harness this energy. Sun is capable of providing energy both to present and to future generations. With the disquieting rate of depletion of key conventional sources of energy such as oil, coal, and natural gas, investing in renewable energy sources such as solar energy has become a necessity. Solar panels convert the sunlight into electrical energy which can power electrical loads. Solar panels operate at less than 40% efficiency if they are stationary. As a result, one has to use more number of solar panels to meet energy demands or a single system with large output. Increased exposure to sun and hence maximum power generation from the solar panels can be achieved through tracking the sun and adjusting the orientation of solar panels. The tracking mechanism moves and orients the solar array in such a way as to align it for optimum output power. Solar tracking is a system which increases power output of solar panels from 30 to 60% compared to stationary systems [1]. In this paper, a system for tracking the sun and adjusting the orientation of solar panels is described.

A. Jayashree (✉) · V. Kudva · A. G. Ananth
Department of E&C Engineering, NMAM Institute of Technology, Nitte (Visvesvaraya Technological University, Belagavi), Udupi, India

V. Kudva
e-mail: vidyakudva@nitte.edu.in

2 Literature Survey

Oltu et al. [2] proposed a low-cost method for tracking solar energy utilizing differential method and a special microdetector. The movement of solar panel was restricted from 0 to 180°.

Stjepanović et al. [3] proposed a PIC16F877 microcontroller-based solar tracking system, in which sequence of light sensor values was read, compared, and rotated the motor in the direction of light sensor that provided the maximum value.

Li et al. [4] developed a system based on solar geometry and monthly horizontal radiation. They provided a mathematical technique to evaluate daily collectible radiation on fixed two-axis and horizontal single-axis monitored solar panels. Results showed that the annual solar gain on a horizontal single-axis tracked solar panel and the orientation of the horizontal tracking axis was related.

Zolkapli et al. [5] designed and developed a high-efficiency dual-axis solar tracking system by utilizing Arduino Uno platform. Five LDRs were used for capturing full light source in hardware implementation. The servo motor was used to shift the solar panel to the optimum position of light source sensed by LDRs. The coding was performed using programming language C. At different time intervals, the efficiency of the system was tested and compared with static solar panel.

Mamun et al. [6] designed and implemented a multidirectional solar tracking system. This model could move in both horizontal and vertical axis and also to a point making combination of both horizontal and vertical axis.

Veligorskyi et al. [7] proposed a two-axis active solar tracker system. For the proposed system, the efficiency was calculated for three particular days of year summer solstice, equinox solstice, and winter solstice. The weather conditions were neglected in this experiment.

Zhu et al. [8] developed a photoelectric tracking system which consisted of signal detection circuit for photoelectric sensor, analog-to-digital converter, microcomputer, stepper motor, stepper motor driver, and a computer. The program for the system was based on windows platform. The main program and subroutine for the photoelectric sensor were developed using Keil software.

Chaïb et al. [9] designed and implemented a robot manipulator and a robot controller for positioning and orientation of solar panel. MATLAB software was used for software simulations. The robot manipulator had two degrees of freedom.

Petrusev et al. [10] developed and tested the acrylic concentrator and tracker for the solar panels. Electric drive parameters for the solar tracker were selected based on weight and size of the solar panel. The use of acrylic concentrators in conjunction with the solar tracker helped solar power plant costs to be reduced by 2–5 times.

Mustafa et al. [11] designed a precise automatic dual-axis solar tracking device using Arduino Uno, DC motors, and LDR. This tracking system considered both azimuth angle and altitude angle. The results revealed that this tracking system was efficient and reliable than the fixed solar panels.

Sharma and Tayal [12] proposed a solar tracking system that tracks the sun movement throughout the day. The system design was based on 8051 microcontroller, DC

motor, and light sensor. The output of the proposed solar tracking system was found to be 32.17% higher than the static flat solar panel.

Mao et al. [13] designed a solar tracking system using ARM processor, photoelectric sensor module, and a motor. ARM processor was implemented in solar trajectory tracking mode. The photoelectric sensor tracked the sun and the motor is used to orient the solar panel. The tracking accuracy was within 0.15° .

Jasim et al. [14] developed a low-cost portable solar panel (PSP) prototype motivated by origami folding mechanism, tilt-and-swing dual-axis solar tracker. To save space, the foldable solar panels were mounted in a motorized, compact, and holding box. This kind of design made the system portable. Further, the active dual-axis solar tracking system consisted of Arduino microcontrollers, LDRs, motor shields, and stepper motors.

Chow and Abiera [15] developed a prototype which uses a solar tracker and data logger that increases the capabilities of a solar panel. The solar panel efficiency was theoretically boosted by aligning it perpendicular to the sun rays. This was implemented by constructing a mounting for the solar panel which was controlled by servo motors. The programming for the collection of data from the sensors and the suitable motor actions were performed using MATLAB programming software. The data logging capabilities and a graphical user interface (GUI) was developed to support the tracker function.

Jallal et al. [16] presented a novel machine-learning model DNN-RODDPSO to enhance the real-time prediction accuracy of the hourly energy provided by four dual-axis solar trackers. This approach combined a modern iteration of the deep neural network (DNN) with the current version of the particle swarm optimization (PSO) algorithm mentioned as a randomly occurring delayed particle swarm optimization algorithm (RODPSO).

Kang et al. [17] proposed a smart solar photovoltaic blinds (SPB) system. This is a window blind device with a photovoltaic module that can control the sunlight entering the room and simultaneously reduces the energy used in the building by generating electricity itself.

Zhang et al. [18] implemented a dual-axis hybrid tracking system using global positioning system (GPS) for geological location, photodiodes for closed-loop tracking and orientation algorithm for open-loop tracking. An initializing calibration was suggested to correct the errors caused due to the mismatch of photodiodes.

Zheng et al. [19] proposed an open-loop control strategy based on the sun tracking control strategy with z variation, and it depends on the predefined apparent motion of the sun. The three-extensible-rod (TER) solar tracker has a GPS (NEO-6M) module to determine the position coordinates and the current date and time at the installation site.

Argatu et al. [20] stated that the most important parameter to be considered is the solar panel efficiency, which is the ratio of the solar energy radiated on the solar panel and the energy produced by the solar panel. The study concluded that a solar panel fixed on a mobile system has greater efficiency than the static solar panel.

Fernandez-Ahumada et al. [21] proposed backtracking to reduce the effects of shading. This strategy involved deviating the direction of the solar trackers from the

solar position to avoid shading between the collectors when necessary. The results revealed that the maximum irradiance backtracking (MIBT) improves the energy collection by 1.31% in comparison to astronomical tracking with no shading (ATNS).

Jamroen et al. [22] designed and developed an automatic dual-axis tracking system. The tracking system was accurately adjusting the photovoltaic (PV) module through primary and secondary axes to follow the sun trajectory using the digital logic design of LDR. This tracker had a pseudoazimuthal system for good stability of movement mechanism. Simple configuration of LDRs were employed and installed to minimize tracking errors caused by the complicated orientation.

Wang et al. [23] proposed one-direction-automatic-rotation (1-DAR), two-direction-automatic-rotation (2-DAR), and three-direction automatic-rotation (3-DAR) perpendicular sun tracking shading panel arrays. The proposed array of shading panels combined with the automatic rotation control system, simultaneously ensured sun-glare prevention and energy saving.

Many techniques have been used to develop solar tracking system. One among them is use of LDRs for sensing maximum light source. Sensed signal is used to drive the servo motors which in turn orient the solar panel in the direction of maximum light source. Different microcontrollers namely PIC16F877 microcontroller, Arduino Uno controller, ARM processor, 8051 microcontroller were used for generating the error signal. Error signal is produced by the disorientation of the solar panels with respect to the light source. This signal is fed to a servo motor to correct the solar panel orientation to generate maximum signal from the light source. Different programming software such as PICBASIC PRO, C, Keil, graphical user interface (GUI), and MATLAB were used for developing the main program for the automatic tracking system to achieve maximum capturing of sunlight by the solar panels. The tracking accuracy of 0.15° has been achieved and the output of the solar panel tracking system was 32% higher compared to static flat solar panel system. In the proposed system, STM32F103C8T6 microcontroller is used for generating error signal and DC gear motor is used to reorient the solar panel.

3 Methodology

Overview of the system for tracking the sun and adjusting the orientation of solar panels is shown in Fig. 1.

The tracker control system contains a microcontroller board, a control program, a power supply board, motor drivers, solar panel, and LDRs. The system uses LDRs to sense the solar energy. Microcontroller is programmed in such a way that output of LDR which receives maximum amount of sunlight will direct the DC gear motors which in turn revolve the solar panels in a direction such that they receive maximum sun light.

Implementation of the system is divided into two phases, simulation phase and hardware implementation phase, of the simulated system.

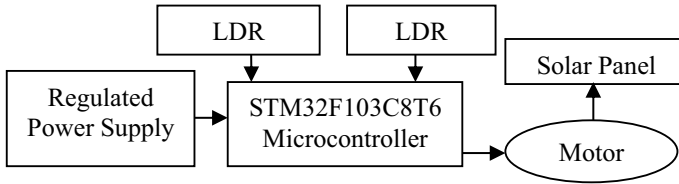


Fig. 1 Overview of the system for tracking the sun and adjusting the orientation of solar panel

3.1 Simulation Phase

The simulation is performed using Proteus 8 professional software. We used ATmega 2560 microcontroller, LDRs, and servo motors. The ATmega 2560 acts as a great development platform for projects with 8-bit microcontrollers. It has 70 I/O pins, 256 KB of flash memory, 8 KB of SRAM, and 4 KB of EEPROM. In Arduino IDE software, the ATmega 2560 is programmed using the C language. The Proteus microcontroller simulation works by either applying a hex file or a debug file to the microcontroller part on the schematic. It is then co-simulated with any attached analog and digital electronics device. This allows its use in areas such as motor control, temperature control, and user interface design across a broad range of project prototyping. The simulated circuit diagram of the tracker control system is shown in Fig. 2.

Four LDRs are mounted on the edges of solar panel and are separated by 90° with respect to each other. These LDRs sense the light and convert them into voltage. The voltages are then fed as an analog signal to ADC channel of the ATmega 2560 microcontroller, which converts the analog signal to digital signal. Two DC geared motors are used to move the solar panel in both horizontal and vertical axis. If one of the LDRs receives more light intensity than the others, the voltage difference is transmitted as an analog signal to the ADC interface of ATmega 2560. The ATmega 2560 microcontroller analyses this signal and generates an error signal which drives the DC geared motors to align the tracker to a point where equal light is being illuminated on pair of LDRs. If the illumination of sun on all four LDRs are equal, the analog voltage signals obtained at the microcontroller's ADC channel will have identical values and the microcontroller does not produce any error signal to drive the DC geared motors and hence the solar panel does not reorient. The intensity of the LDRs and also the difference between them is displayed in the Liquid Crystal Display (LCD). Metal Oxide Semiconductor Field Effect Transistor (MOSFET) is used to isolate the panel and the battery during the measurement of voltage across them. The panel and battery voltages are displayed in the LCD.

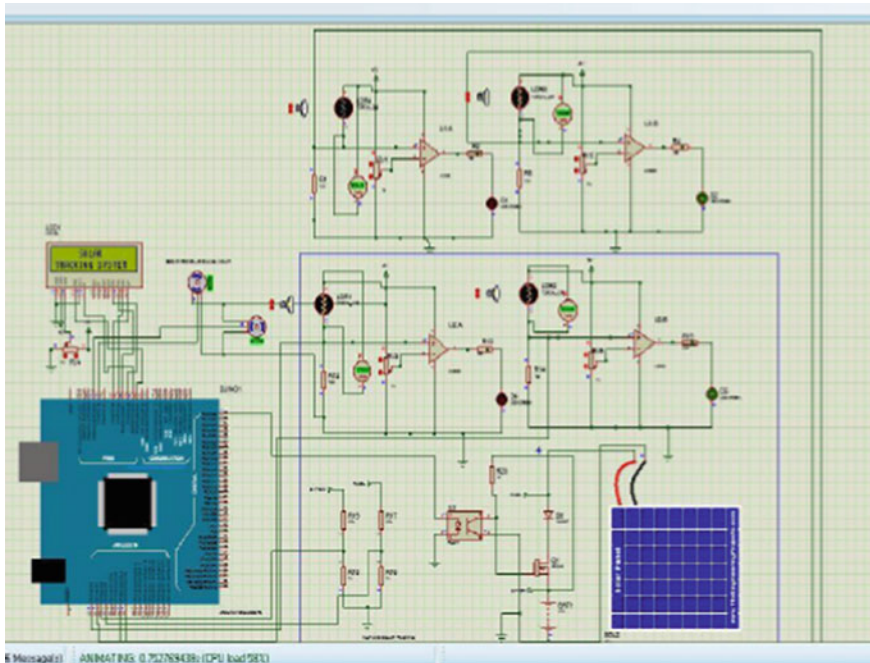


Fig. 2 Circuit diagram of the tracker control system

3.2 Hardware Implementation Phase

For hardware implementation, STM32F103C8T6 microcontroller is used. In deep sleep mode, it consumes the current up to nanoamps. It has ARM cortex M3 32-bit architecture and operates at 72 MHz frequency. It has 10 analog channels, 64 KB flash memory, and 20 KB SRAM. These are the major factors considered to select STM32F103C8T6 microcontroller. The STM32F103C8T6 is programmed using the C language in Arduino IDE software. The system uses LDRs to sense the intensity of solar energy. The block diagram of the hardware implementation is shown in Fig. 3.

The experimental setup of the tracking system is shown in Fig. 4.

STM32F103C8T6 microcontroller board is interfaced with two LDRs, namely LDR1 and LDR2. These LDRs are placed on the two opposite edges of the solar panel. The LDRs sense the intensity of the solar radiation, which is an analog signal. This analog output of LDRs is fed to the PA0 and PA1 pins, which are the analog pins of the microcontroller. These analog pins are then interfaced with the 12-bit ADC present in the microcontroller to convert the analog intensity values to digital. The microcontroller is programmed to perform binary subtraction of these two intensity values. This intensity difference between two LDRs is differential intensity. The DC geared motor driven by L293N motor driver, rotates and reorients the solar panel until this differential intensity becomes zero. Some motors can be directly

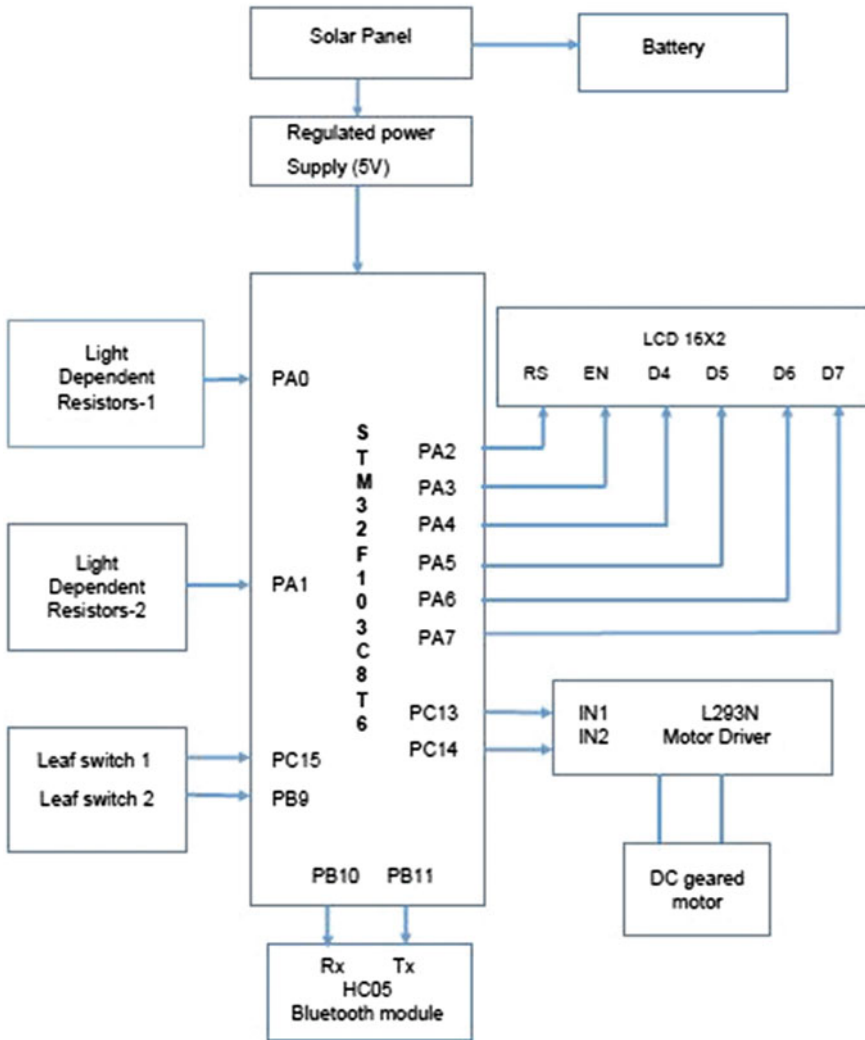


Fig. 3 Block diagram of the hardware implementation

controlled by connecting the power supply to them and make it work but if these motors are controlled by microcontrollers then we need motor drivers. If there is no differential intensity, then solar panel does not orient. The differential intensity is also displayed in the LCD. This motor driver requires 12 V supply, whereas LDRs, LCD, and microcontroller require 5 V supply, LM7805 voltage regulator is used for this purpose. The leaf switches are used to limit the rotation of solar panel in both the directions. The panel voltage and battery voltage are also displayed in the LCD.

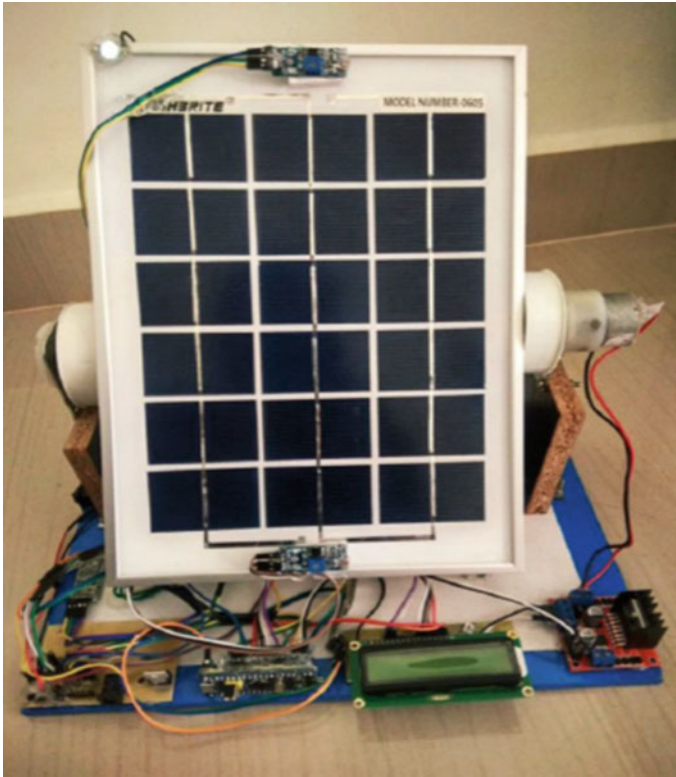


Fig. 4 Experimental setup of the solar tracking system

The solar panel voltage is stored in a battery and the entire system is operated by the power which is stored in battery.

Experimental procedure: The solar panel orientation and tracking system was kept horizontal. Light sources used for this experiment was mobile flashlight of Redmi Note 4. The flashlight is of 50 lm (lumens), i.e., 0.5555555556 W. Two light sources were used and fixed parallel to the solar panel and was focused to LDRs. Due to the use of mobile flashlight distance between LDRs and light source was limited to 1m. With a distance more than 1m, the LDRs did not sense the light. When the intensity of light was same on both LDRs the orientation did not take place since there is no differential intensity. Then the solar panel was tilted to different angles like 10°, 20°, 30°, etc., keeping the light source constant. The differential intensity required to reorient the solar panel to the direction of maximum solar energy was noted for corresponding angles. The solar panel was tilted in both clockwise and anticlockwise directions, but for the same angle, the differential intensity noted to reorient the solar panel that slightly differed from one another due to backlash of servo motor.

4 Results and Discussions

4.1 Result of Simulation Phase

In the simulation of the solar panel orientation, two LDRs namely LDR1 and LDR2 were placed on the edges of the solar panel. Intensity at LDR1 was kept constant and intensity of the source at LDR2 is varied. The voltage output at LDR2 is measured using a voltmeter and differential voltage and differential intensity displayed in the LCD are noted. The same process is repeated but by keeping the intensity of source at LDR2 constant and varying the intensity of source at LDR1. The readings of these two experimentations are listed in Tables 1 and 2, respectively.

It can be observed from Tables 1 and 2 that differential voltage produced is proportional to the differential intensity. As this differential voltage is used to drive the motor which in turn rotates the solar panels, the angular rotation of the solar panel is proportional to the differential voltage. A plot of differential voltage versus differential intensity is shown in Fig. 5.

From Fig. 5, it can be observed that the differential voltage increases as the differential light intensity falling on the LDRs increases corresponding to the increasing angle of rotation of the solar panel. The differential voltage increases linearly with differential intensity between 123 and 804 W/m² which corresponds to differential voltage 0.6–3.93 V. The minimum differential intensity required to trigger the servo motor is 0.6 V, which corresponds to 123 W/m². The maximum differential intensity above which the differential voltage saturates is 804 W/m². The operating differential voltage range of the servo system designed for orientation of the solar panel is between 0.6 and 3.93 V. The servo system can be conveniently operated in the linear region. The same results are observed for the other two LDRs namely LDR3 and LDR4.

Table 1 Differential intensity and differential voltages when the intensity at LDR1 is kept constant and at LDR2 is varied

Voltage at LDR 1 (V)	Intensity at LDR 1 (W/m ²)	Voltage at LDR 2 (V)	Intensity at LDR 2 (W/m ²)	Differential intensity (W/m ²)	Differential voltage (V)
0.24	975	0.24	975	0	0
0.24	975	0.46	930	45	-0.24
0.24	975	0.84	852	123	-0.6
0.24	975	1.68	681	294	-1.43
0.24	975	2.51	510	465	-2.27
0.24	975	3.34	340	635	-3.1
0.24	975	4.17	171	804	-3.93
0.24	975	4.55	94	881	-4.3
0.24	975	4.76	50	925	-4.51
0.24	975	4.95	12	963	-4.7

Table 2 Differential intensity and differential voltages when the intensity at LDR2 is kept constant and at LDR1 is varied

Voltage at LDR 1 (V)	Intensity at LDR 1 (W/m ²)	Voltage at LDR 2 (V)	Intensity at LDR 2 (W/m ²)	Differential intensity (W/m ²)	Differential voltage (V)
0.24	975	0.24	975	0	0
0.46	930	0.24	975	-45	0.24
0.84	852	0.24	975	-123	0.6
1.68	681	0.24	975	-294	1.43
2.51	510	0.24	975	-465	2.27
3.34	340	0.24	975	-635	3.1
4.17	171	0.24	975	-804	3.93
4.55	94	0.24	975	-881	4.3
4.76	50	0.24	975	-925	4.51
4.95	12	0.24	975	-963	4.7

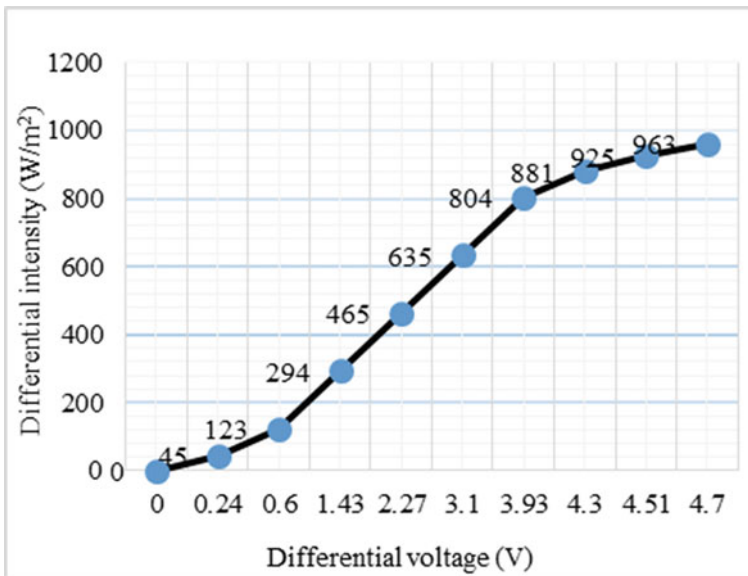


Fig. 5 Differential voltage versus differential intensity plot

4.2 Results of Hardware Implementation Phase

In the hardware implementation of the system, the light intensity of Redmi Note 4 mobile flash light at both the LDRs are adjusted equal such that the differential intensity is zero and the tilt angle of the solar panel is observed to be zero. Then the

Table 3 Intensity difference and its corresponding tilt angle during coarse tuning of the system for both clockwise and anticlockwise orientation of the solar panel

Clockwise		Anticlockwise	
Differential intensity (W/m ²)	Tilt angle (°)	Differential intensity (W/m ²)	Tilt angle (°)
0	0	0	0
194	10	196	10
493	20	495	20
568	30	569	30
850	40	852	40
1053	50	1057	50

solar panel was tilted to different angles like 10°, 20°, 30°, etc., keeping the intensity of flashlight constant. The intensity is power transferred per square meter of solar panel, where the power is derived from the analog output of the LDR. The differential intensity is the difference of intensity between two LDRs. The differential intensity was noted for corresponding angles. The solar panel was tilted in both clockwise and anticlockwise. Results of this experimentation are listed in Tables 3 and 4.

It is observed that the minimum differential intensity between the two LDRs namely LDR1 and LDR2 has to be 165 W/m² and the corresponding tilt angle is found to be 5°. Tables 3 and 4 list the differential intensity and its corresponding tilt angle of the solar panel during coarse and fine tuning of the system for clockwise and anticlockwise rotation of the solar panel. It is observed that the small intensity difference (less than 0.5%) seen for the same tilt angle of the solar panel between clockwise and anticlockwise orientation can be attributed to the backlash of the motor driver system. Plots of differential intensity versus tilt angle plot for coarse and fine tuning of the system for clockwise and anticlockwise rotation of solar panel are shown in Figs. 6 and 7.

Table 4 Intensity difference and its corresponding tilt angle during fine tuning of the system for both clockwise and anticlockwise orientation of the solar panel

Clockwise		Anticlockwise	
Differential intensity (W/m ²)	Tilt angle (°)	Differential intensity (W/m ²)	Tilt angle (°)
0	0	0	0
165	5	168	5
183	7	184	7
194	10	197	10
288	15	286	15
493	20	495	20
510	23	513	23
527	25	529	25

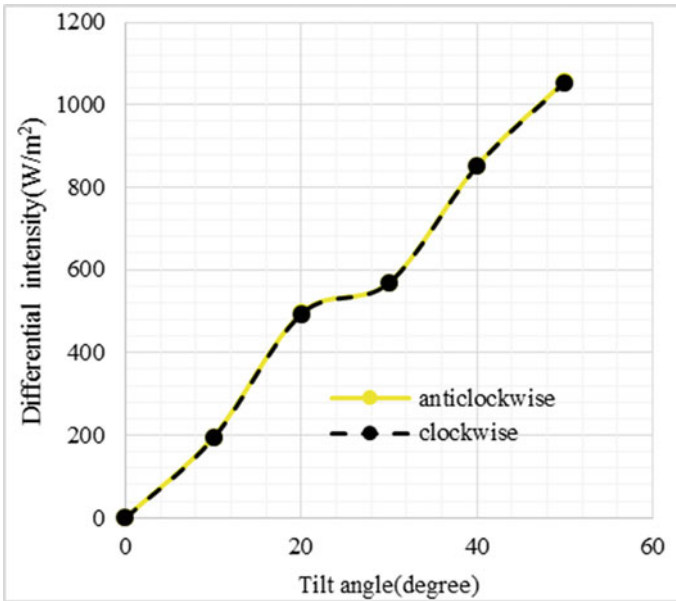


Fig. 6 Differential intensity versus tilt angle plot for the coarse tuning of the system for clockwise and anticlockwise rotation of solar panel

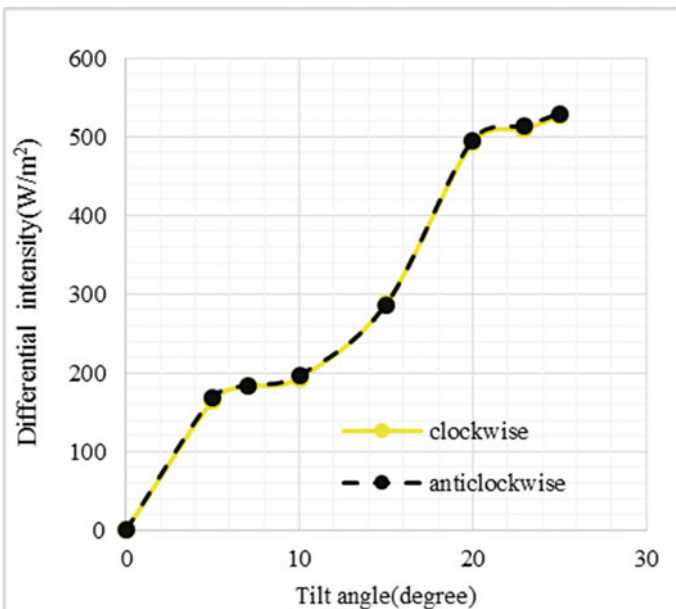


Fig. 7 Differential intensity versus tilt angle plot for fine tuning of the system for clockwise and anticlockwise rotation of solar panel

It is observed that tilt angle variation increases as the differential intensity between the two LDRs increases. The tilt angle shows linear variation with differential intensity, further it may be noted from the Fig. 6 that there is no significant difference between the behavior of tilt angle with differential intensity for clockwise and anticlockwise rotation of the solar panel. Similarly Fig. 7 shows the tilt angle versus differential intensity plot for fine tuning of the system for clockwise and anticlockwise rotation of solar panel. It is observed that for minimum differential intensity of 165 W/m^2 , the tilt angle of solar panel will be 5° . The graph clearly suggests that the tilt angle variation of the solar panel even up to 5° can be accurately corrected for deriving maximum power from solar panel. If the intensity on both the LDRs are same, then the differential intensity is zero and the tilt angle of solar panel is also zero. Hence, there is no reorientation of the solar panel takes place.

5 Conclusion

A system for tracking the sun and adjusting the orientation of solar panels was developed. The focus of the system was to design the controlling part which generates a control signal proportional to the differential intensity to drive the DC-gear motor. The experimental results indicated that differential voltage which is used as the control signal to drive the motor is proportional differential intensity. The proposed system can be used in non-critical domestic and industrial applications for generation of electricity.

References

1. F.T. Fernandes, F.A. Farret, C.R. De Nardin, J.A. Pires, M. Brondani, A. Longo, Economical feasibility of photovoltaic array power increase by alternative structures inclusion. *Int. J. Emerg. Electr. Power Syst.* **19**, 30 (2018)
2. O. Oltu, P.L. Milea, M. Dragulescu, E. Franti, Solar panel energetic efficiency optimization method, based on an specific detector and orientation microsystem, in *IEEE International Semiconductor Conference*, 2007, pp. 127–130
3. A. Stjepanović, S. Stjepanović, F. Softić, Z. Bundalo, Microcontroller based solar tracking system, in *IEEE 9th International Conference on Telecommunication in Modern Satellite, Cable, and Broadcasting Services*, 2009, pp. 518–521
4. G. Li, R. Tang, H. Zhong, Optical performance of horizontal single-axis tracked solar panels. *Energy Proc.* **16**, 1744–1752 (2012)
5. M. Zolkapli, S.A.M. Al-Junid, Z. Othman, A. Manut, M.A. Mohd Zulkiffi, High-efficiency dual-axis solar tracking development using Arduino, in *IEEE International Conference on Technology, Informatics, Management, Engineering and Environment*, 2013, pp. 43–47
6. N. Mamun, N.H. Samrat, N. Mohamma, M.J. Islam, A.A. Mamun, M.H.R. Prince, R. Adib, M. Iftakherahmed, Multi-directional solar tracker using low cost photo sensor matrix, in *IEEE International Conference on Informatics, Electronics & Vision (ICIEV)*, 2014, pp. 1–5

7. O. Veligorskiy, R. Kosenko, S. Stepenko, High-efficiency solar tracker development and effectiveness estimation, in *IEEE International Conference on Intelligent Energy and Power Systems (IEPS)*, 2014, pp. 153–158
8. Z. Zhu, Y. Xu, G. Shao, An automatic sun tracking system for solar energy based optical fiber lighting, in *IEEE 4th Asia-Pacific Conference on Antennas and Propagation (APCAP)*, 2015, pp. 69–71
9. A. Chaïb, M. Kesraoui, E. Kechadi, PV panel positioning using a robot manipulator, in *IEEE 3rd International Renewable and Sustainable Energy Conference (IRSEC)*, 2015, pp. 1–6
10. A.S. Petrusev, V.M. Rulevskiy, Y.Z. Sarsikeyev, D.Y. Lyapunov, Solar tracker with active orientation, in *IEEE 2nd International Conference on Industrial Engineering, Applications and Manufacturing (ICIEAM)*, 2016, pp. 1–4
11. F. Mustafa, A.S. Al-Ammri, F.F. Ahmad, Direct and indirect sensing two-axis solar tracking system, in *IEEE 8th International Renewable Energy Congress (IREC)*, 2017, pp. 1–4
12. V. Sharma, V.K. Tayal, Hardware implementation of sun tracking solar panel using 8051 micro-controller, in *IEEE 6th International Conference on Reliability, Infocom Technologies and Optimization (ICRITO)*, 2017, pp. 483–486
13. K. Mao, F. Liuand, I.R. Ji, Design of ARM-based solar tracking system, in *IEEE 37th Chinese Control Conference (CCC)*, 2018, pp. 7394–7398
14. B. Jasim, P. Taheri, An origami-based portable solar panel system, in *IEEE 9th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON)*, 2018, pp. 199–203
15. L.S. Chow, M. Abiera, Optimization of solar panel with solar tracking and data logging, in *IEEE student Conference on Research and Development*, 2018, pp. 15–19
16. M.A. Jallal, S. Chabaa, A. Zeroual, A novel deep neural network based on randomly occurring distributed delayed PSO algorithm for monitoring the energy produced by four dual-axis solar trackers. *Renew. Energy* **149**, 1182–1196 (2019)
17. H. Kang, T. Hong, S. Jung, M. Lee, Techno-economic performance analysis of the smart solar photovoltaic blinds considering the photovoltaic panel type and the solar tracking method. *Energy Build.* **193**, 1–14 (2019)
18. J. Zhang, Z. Yin, P. Jin, Error analysis and auto correction of hybrid solar tracking system using photo sensors and orientation algorithm. *Energy* **182**, 585–593 (2019)
19. T. Zheng, F. Zheng, X. Rui, X. Ji, K. Niu, A novel ultra-light dish system based on a three-extensible-rod solar tracker. *Sol. Energy* **193**, 335–359 (2019)
20. F. Argatu, F. Tarcau, V. Argatu, S. Grigorescu, B. Enache, L. Constantinescu, Building and designing a dual axis solar tracker for educational purposes, in *11th International Conference on Electronics, Computers and Artificial Intelligence (ECAI)*, 2019, pp. 1–4
21. L.M. Fernández-Ahumada, J. Ramírez-Faz, R. López-Luque, M. Varo-Martínez, I. M. Moreno-García, F.C. de la Torre, A novel backtracking approach for two-axis solar PV tracking plants. *Renew. Energy* **145**, 1214–1221 (2020)
22. C. Jamroen, P. Komkum, S. Kohsri, W. Himananto, S. Panupintu, S. Unkat, A low-cost dual-axis solar tracking system based on digital logic design: Design and implementation. *Sustain. Energy Technol. Assessments* **37**, 100618 (2020)
23. R. Wang, G. Li, L. Xu, Y. Wang, C. Peng, Integration of sun-tracking shading panels into window system towards maximum energy saving and non-glare day lighting. *Appl. Energy* **260**, 114304 (2020)

Development of an Automated Plant Classification System Using Deep Learning Approach



K. Ananth Pai, B. R. Apoorva, Daisy Sheetal Mendonca, Durgaprasad S. Hegde, and Roopa B. Hegde

1 Introduction

Plants play an important role in maintaining balance in the environment. Plants are vital for human existence as they rely upon plants for energy, medicine, and many other things. With the introduction of new species, there has been an increase in the number of unknown plants. On the other hand, some of the known species with medicinal values are on the verge of extinction without human knowledge. The plant classification is necessary for the identification of the alien plant species to understand their importance and avoid their invasion of the ecosystem of an area. It is also required to protect plant species and slow down their extinction. Manual identification of plants is a tedious task and requires experts in this field. Hence, the researchers must come up with different approaches for the identification and classification of plant species.

Artificial intelligence (AI) is an emerging field that makes it possible to automatically classify plant species without human intervention. The term machine learning (ML) refers to an algorithm and mathematical model designed to perform a specific task, but for the increase in the amount of data, the efficiency of ML algorithms is not pleasing. Hence, in such a scenario, deep learning (DL) is suitable. Convolutional neural network (CNN) [47] is a commonly used deep network. CNNs are gaining popularity because of their ability to process huge amounts of data. Also, segmentation and feature engineering are performed by the network.

The recent advances in ML techniques have led several research groups to use this in plant recognition. Every single leaf possesses a unique feature, which was cleverly

K. Ananth Pai (✉) · B. R. Apoorva · D. S. Mendonca · D. S. Hegde · R. B. Hegde
NMAM Institute of Technology, Nitte (Visvesvaraya Technological University, Belagavi), Udupi
574110, Karnataka, India

R. B. Hegde
e-mail: roopabhegde@nitte.edu.in

exploited by several research groups for the classification of plants. Various techniques were used for plant identification and feature extraction of leaves. Several research groups used artificial neural network (ANN) [1–5], probabilistic neural network (PNN) [6–10] for classification of plants using leaf images and found that the accuracy of plant leaf classification to be in the range of 84.42%–97.33% and 82.04%–93%, respectively. A comparison between state vector machine–binary decision tree (SVM-BDT), probabilistic neural network (PNN), and Fourier moment’s classifiers was carried out [11] for classification of leaf images and their findings show that SVM-BTD with the accuracy of 96% outperformed PNN and Fourier moment’s classifiers. Support vector machine (SVM) was used for classification of leaves by many research groups [12–17] and the accuracy of classification was found to be the range of 90–95.6%. Features were extracted using histogram of oriented gradients (HOG) and local binary pattern (LBP) to obtain the classification accuracy in the desired range. The use of the k-nearest neighbor classifier was experimented with by several researchers for leaf identification and obtained the classification accuracy was around 83.5–94.3% [18–22]. Multilayered perceptron (MLP) was used for leaf classification [23, 24] and accuracy was observed to be in the range of 85.65–87.8%. A combination of gray-level co-occurrence matrix (GLCM) and principal component analysis (PCA) was used for feature extraction [25] and an accuracy of 98% was observed. Morphological features like shape, color, and texture were used for leaf classification [26–30]. Many research groups used contour-based leaf recognition [31, 32].

Recent advancements in deep learning techniques made several researchers to build convolutional neural network (CNN) for leaf identification with accuracy ranging from 90–97.7% [33–38]. A comparative study of existing deep networks such as GoogleNet, MobileNet, ResNet, Inception-V3, VGGNet, AlexNet, and their modified forms for leaf classification was carried out by many research groups and the accuracy was observed to be in the range of 76.87–95.54% [39–41]. Amuthalingeswaran et al. [42] designed a deep network called medicinal neural network (MNN) for the purpose of leaf identification which reported an accuracy of 85.15%.

ANN with backpropagation was used for leaf image identification [43, 44] with the accuracy obtained ranging from 96 to 97.2% for different standard datasets. Snakes technique and cellular neural network were used for plant identification [45], and growing convolutional neural network (GCNN) was employed as the classifiers to identify the plant leaves which yielded an accuracy of 88% [46]. In this paper, we propose an automated system for the classification of plants using leaf images. A CNN consisting of fourteen layers is designed for the identification of plants.

2 Methodology

The block diagram of the proposed system is shown in Fig. 1. A camera module operated using a Raspberry Pi clicks the leaf images in real time, and it is fed to



Fig. 1 Proposed plant classification system

the classifying system. We used CNN for the classification of leaf images. CNN processes the leaf images and then classifies it to a particular class. The result of classification is displayed to the user using a Raspberry Pi screen.

The model was designed and trained in Anaconda IDE using the Python framework. Since the memory bandwidth of GPU is greater than that of CPU and can process multiple computations simultaneously, it is used for training the CNN model. The specifications of the system used are as follows.

Intel Core i7-8750H CPU operating at 2.2 GHz with 8 GB RAM and NVIDIA GTX 1050Ti GPU with 4 GB memory.

Python-based libraries such as OpenCV, Numpy, Keras, and Tensorflow were installed in Raspberry Pi and stored in the memory card. A 5 MP camera module, button, and the Raspberry Pi screen which also form the hardware of the system were tested and configured separately with the help of these libraries. The hardware components were then integrated with the Raspberry Pi that contains the trained model. As soon as the button is pressed, the camera module captures the leaf image which is given as the input to the classifying system. The model then processes this image to predict the plant name and displays the same along with a sample image of the identified plant on the Raspberry Pi screen.

2.1 Data Collection

Images required for this study were acquired under varied lighting conditions, different viewpoints, and various color backgrounds. The dataset consists of 5958 leaf images of six classes namely acacia, curry, guava, ivy gourd, jackfruit, and neem. This dataset was split into a training set, testing set, and validation set in the ratio 81:10:9. Details of the dataset are provided in Table 1. Sample images of each type are shown in Fig. 2. It is evident from the figure that, the dataset consists of leaf images of different shapes and sizes.

CNN requires a large number of training data. Therefore, data augmentation was performed on the training set to increase the training images. The variations were induced in terms of image width, image height, zoomed images, horizontal flip, vertical flip, shear, and rotation of images [50].

Table 1 Dataset details

Plant species	Number of images in the training set	Number of images in the testing set	Number of images in the validation set	Total number of images
Acacia	846	114	84	1044
Curry leaves	809	107	91	1007
Guava	752	99	87	938
Ivy gourd	819	98	92	1009
Jackfruit	779	93	93	965
Neem	820	85	90	995

**Fig. 2** Sample images of the dataset

2.2 Proposed Method

A CNN with fourteen layers was designed for plant classification using leaf images. It has six convolutional layers, three pooling layers, four fully connected layers, and an output layer which is as shown in Fig. 3. The acquired images were resized from $3500 \times 3500 \times 3$ to $150 \times 150 \times 3$ which are given as input to the CNN.

The images were resized in such a way that no features are lost and also ensuring that the image quality is retained.

There are six convolutional layers that form layers 1, 2, 4, 5, 7, and 8 of the designed network having 64, 128, 64, 128, 128, and 256 filters, respectively. Each of these filters has a kernel size of 3×3 and moves with a stride length of 1×1 , because of which at the output of every convolutional layer the resolution is reduced

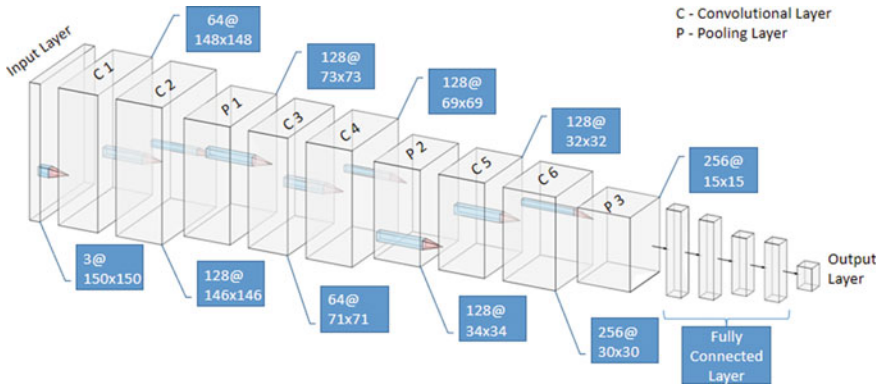


Fig. 3 Proposed CNN architecture [52]

by 2×2 . For example, the input to convolutional layer 1 is of size 150×150 which is reduced to 148×148 at the output. The number of layers, number of filters, stride size, kernel size, and other hyperparameters are selected empirically after conducting several experimentation.

Every convolutional layer is followed by the ReLU activation function. ReLU is used so that the model can learn faster and perform better because it overcomes the vanishing gradient problem. The equation of ReLU is as given in Eq. (1). Zero paddings are added to the frame of the image in layer 1 to allow for more space for the kernel to cover the image. The addition of padding to an image processed by a CNN allows for a more accurate analysis of images.

$$f(x) = \max(0, x) \tag{1}$$

Three pooling layers form layers 3, 6, and 9 of the designed CNN. Each pooling layer has a pool size of 2×2 and uses max pooling [51]. Since the kernel size is 2×2 , the output image resolution will be half the input image resolution. For example, in layer 3, the resolution of the input image is 146×146 and that of the output image is 73×73 . This will enrich the image features and reduces the computational complexity in subsequent steps. The output from the last pooling layer, i.e., layer 9 is flattened to form the fully connected layer. There are four fully connected layers. The first and second fully connected layers have 1000 and 500 neurons, respectively. The third and fourth fully connected layers have 500 neurons each. Dropout function is used on these fully connected layers to overcome the problem of over-fitting. The probability of the number of neurons being dropped during each epoch for these fully connected layers is 0.4, 0.3, 0.3, and 0.2, respectively. The last output layer uses softmax as the activation function. The number of neurons in this layer is six which is equal to the number of leaf classes. The softmax function is applied at the output layer as it turns logic into probabilities which are as given in Eq. 2.

$$f(Xi) = \frac{\exp(Xi)}{\sum_j \exp(Xj)} \quad (2)$$

The exponential function is applied to each element Xi of input vector X and then is divided by the sum of all exponential values of the input vector represented by Xj .

The CNN was trained for 251 epochs with 2000 steps per epoch. The model uses a learning rate of 0.0001. The loss function used is categorical cross-entropy, and the performance metric considered here is accuracy. Adam [48] was the optimizer used because it provides optimized algorithm that can deal with sparse gradients on noisy problems by utilizing the best features of the AdaGrad and RMSProp algorithms. After the training process, CNN was tested for an unseen dataset which was also collected along with the main dataset and kept aside for the testing purpose.

3 Results

The performance of the model is determined by accuracy, specificity, and sensitivity. Specificity can be calculated as the ratio of true negative (TN) to the sum of TN and false positive (FP) as shown in Eq. (3). Sensitivity can be calculated as the ratio of true positive (TP) to the sum of TP and false negative (FN) as in Eq. (4).

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}} \quad (3)$$

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (4)$$

Color images of size $150 \times 150 \times 3$ were fed to the built model and the following results were obtained. The CNN reported a training accuracy of 97.38% with a loss of 0.0884, validation accuracy of 99.25% with a loss of 0.0475, and testing accuracy of 98.3%. Out of 596 predictions, the system was able to predict 586 images correctly. The values of specificity and sensitivity for each class of plant leaf are listed in Table 2. It can be observed from the table that, less sensitivity is obtained for curry leaves. This could be due to its appearance which is similar to neem leaves. Though there is

Table 2 Performance measures for color images

Plant name	Specificity	Sensitivity
Acacia	1	1
Curry leaves	1	0.9252
Guava	0.9899	0.9898
Ivy gourd	1	1
Jackfruit	0.9940	1
Neem	0.9960	0.9882

a small difference at the edges of the curry leaves and neem leaves, this feature for some of the leaves was not highlighted due to the resizing of images.

To test the designed network and to possibly reduce the computation, we performed the classification of leaf images considering the grayscale representation of the leaf images of size $150 \times 150 \times 1$. We obtained a training accuracy of 96.66% with a loss of 0.1103, validation accuracy of 90.16% with a loss of 0.5226, and testing accuracy of 91.94%. Out of 596 predictions, the system was able to predict 548 images correctly. The values of specificity and sensitivity for each class of plants are as listed in Table 3. It is clear from the table that, performance measure values are less compared to that of color images. This is because the color feature is also one of the important features for leaf classification. We obtained a testing accuracy of 98.3% for color images and 91.94% for grayscale images. It is true that using color images will significantly increase the number of computations compared to that of grayscale images. But it is observed that there is an increase in the classification accuracy when color images were used as input instead of grayscale images.

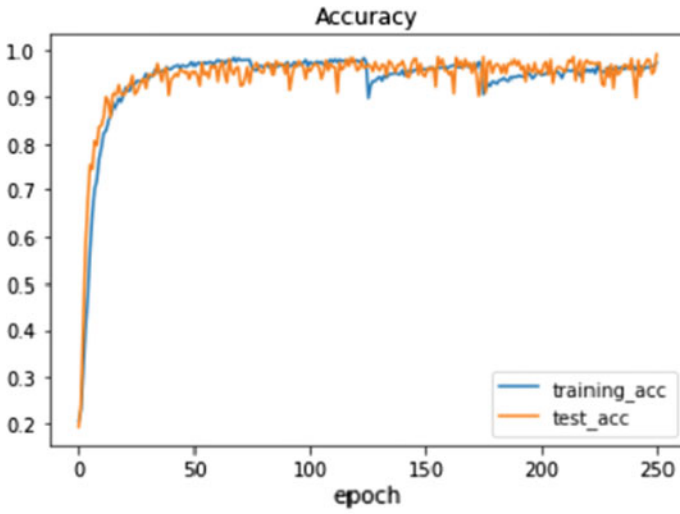
Plots for training and testing accuracy for color and grayscale images are shown in Fig. 4a, 4b, respectively. It can be observed from the plots that, testing accuracy in the case of grayscale images is deviating from that of training accuracy (Fig. 4b), whereas the accuracy of training and testing is not deviating much in the case of color images.

To study the impact of the size of input images, we carried out this experiment considering different sizes of input images [49]. Apart from images of size 150×150 , CNN was also trained for downsized images of 100×100 and 50×50 and compared the results. The performance of the CNN for different sized input images are listed in Table 4.

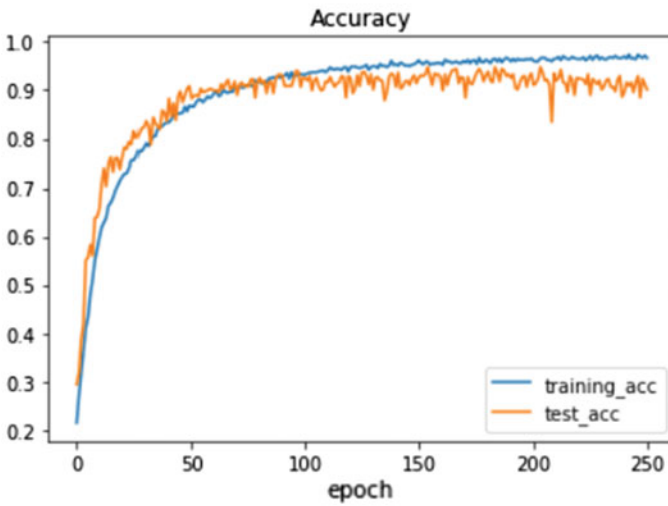
We can infer from the table that accuracy decreases and loss increases when the resolution of the images is decreased for both color and grayscale images. This is because the features of the leaves are lost by reducing the image size. Hence, the size of the input image plays a vital role in classifying leaf images. Therefore, careful selection of image size is required for accurate classification. Smaller image size results in loss of information whereas larger size image increases the computation time and complexity. Hence, the selection of input image size requires several experiments. The designed model with the chosen Raspberry Pi model supported the image size $150 \times 150 \times 3$. This limited the input image size in our experiments.

Table 3 Performance measures for grayscale images

Plant name	Specificity	Sensitivity
Acacia	0.9875	0.9736
Curry	0.9979	0.8317
Guava	0.9919	0.8224
Ivy gourd	0.9839	0.9693
Jackfruit	0.9781	0.9462
Neem	0.9647	1



(a)



(b)

Fig. 4 Plot of training and testing accuracy, **a** color images and **b** grayscale images

Increasing the size of the input image may improve accuracy. This needs a selection of Raspberry Pi model which supports the requirements.

Table 4 Accuracy and loss for different input image sizes

Input image size	Training accuracy (%)	Training loss	Validation accuracy (%)	Validation loss	Testing accuracy (%)
150 × 150 × 3	97.30	0.0884	99.25	0.0475	98.30
100 × 100 × 3	97.67	0.0843	96.08	0.1126	96.47
50 × 50 × 3	96.10	0.1254	87.70	0.6159	88.25
150 × 150 × 1	96.66	0.1103	90.13	0.5226	91.94
100 × 100 × 1	91.91	0.2535	84.35	0.5289	88.00
50 × 50 × 1	76.62	0.6188	94.30	0.8039	75.67

4 Observations

Several research groups have tried to develop an automated plant identification system using ML and DL. It can be observed from the state-of-the-art methods that, CNN outperforms many traditional ML approaches. This is because segmentation of the region of interest and feature extraction is taken care of by the network itself. CNN model and deep learning frameworks can be retrained on a custom dataset which makes it more flexible than traditional ML and CV approaches which tend to be more domain-specific. Therefore, deep learning could be used for the automated classification of plants based on the availability of resources.

Classification of plants can be achieved using leaves, stem, roots, flowers, fruits, seeds, etc. Since leaves are available inadequate and in every season, it is feasible to classify plants based on their leaves. Also, the shape and size of the leaves are different for different plants. Hence, it is convenient to use leaf images for the classification of plants.

In this work, we proposed a novel method for the classification of plants that could be used in real-time applications. In the proposed method, we plucked the leaves for designing and testing the model. But this can be improved by designing a model without plucking the leaves. Plucking the leaves of rare species may harm the plant. Hence, there is a need for improving the model which can process the leaves with complex background. The current model was trained with images having a single leaf. This model could be further modified to recognize the plant in an image having more than one leaf. Further, the number of classes in the dataset can be increased, and also the number of images in each of the classes can be increased to train the model more accurately. But this may increase the computation complexity.

The overall accuracy of the proposed method was compared with few of the existing methods and the comparison is given in Table 5. Different techniques have been implemented to identify and classify leaves but it can be observed from Table 5 that the proposed method has higher accuracy than the existing methods. Though the

Table 5 Comparison of proposed method with existing methods

Papers	Method	Accuracy (%)
Proposed method	CNN	98.3
Fu et al. [1]	ANN	97.33
Prasad et al. [13]	SVM	95.6
Lee et al. [37]	CNN	97.7
Şekeroğlu et al. [44]	ANN	97.2
Zhao et al. [49]	GCNN	88

difference is not huge, with few changes in the dataset, the accuracy can be improved significantly. The proposed method not only has high classification accuracy but also is robust since it classifies the plant leaf image with any background, as the dataset used to train the model had the images of different background. Also, the use of color images over grayscale helped in improving the accuracy significantly since more features were captured by the model.

5 Conclusion

In this paper, an automated plant identification system was designed using CNN. Leaf images of six different plants were used in developing the CNN with fourteen layers. Training and testing of CNN were carried out by considering both color as well as grayscale images separately of various resolution. The specificity and sensitivity were obtained in the range of 0.9899–1 and 0.9252–1, respectively, with an accuracy of 98.3% for color images with the image resolution of 150×150 . It was found that the proposed system gave better classification accuracy when trained on color images rather than grayscale images. Also, it was observed that the image resolution was directly proportional to the classification accuracy. Such a system can classify the leaves and hence can be used to automatically identify the plants. This will reduce the burden on botanists and also helps in the identification of medicinal plants. This system can be used in speeding up the plant identification process. Also, the system can be trained to identify new species and the species in the verge of extinction.

References

1. H. Fu, Z. Chi, Combined thresholding and neural network approach for vein pattern extraction from leaf images. *IEE Proc. Vision Image Signal Process.* **153**(6), 881–892 (2006)
2. W. Qing-feng et al, Feature extraction and automatic recognition of plant leaf using artificial neural network (2007)
3. V. Satti, A. Satya, S. Sharma, An automatic leaf recognition system for plant identification using machine vision technology. *Int. J. Eng. Sci. Technol. (IJEST)* **5**, 874–879 (2013)

4. H. Fu, Z. Chi, D. Feng, J. Song, Machine learning techniques for ontology-based leaf classification, in *ICARCV 2004 8th Control, Automation, Robotics and Vision Conference*, vol. 1, Kunming, China, 2004, pp. 681–686
5. H. Fu, Z. Chi, A two-stage approach for leaf vein extraction, vol. 1, pp. 208–211. <https://doi.org/10.1109/ICNNSP.2003.1279248.2004>
6. K. Mahdikhanlou, H. Ebrahimzhad, Plant leaf classification using centroid distance and axis of least inertia method, in *2014 22nd Iranian Conference on Electrical Engineering (ICEE)* (2014)
7. D. Sandya, Y. Herdiyeni, MedLeaf: mobile application for medicinal plant identification based on leaf image. *Int. J. Adv. Sci. Eng. Inf. Technol.* **3** (2013). <https://doi.org/10.18517/ijaseit.3.2.287>
8. C. Uluturk, A. Ugur, Recognition of leaves based on morphological features derived from two half-regions, in *2012 International Symposium on Innovations in Intelligent Systems and Applications*, Trabzon, 2012, pp. 1–4
9. S.G. Wu et al., A leaf recognition algorithm for plant classification using probabilistic neural network, in *IEEE International Symposium on Signal Processing & Information Technology*, 2007, pp. 11–16
10. X. Gu, J.-X. Du, X.-F. Wang, Leaf recognition based on the combination of wavelet transform and Gaussian interpolation. *Adv. Intell. Comput.* (2005)
11. K. Singh, I. Gupta, S. Gupta, SVM-BDT PNN and fourier moment technique for classification of leaf shape. *Int. J. Signal Process. Image Process. Pattern Recogn.* **3**, 67–78 (2010)
12. A.T. Quadri, M. Sirshar, Leaf recognition system using multi-class kernel support vector machine. *Int. J. Comput. Commun. Syst. Eng.* **2**(2), 260–263 (2015)
13. S. Prasad, P. Kumar, R.C. Tripathi, Plant leaf species identification using Curvelet transform, in *2011 2nd International Conference on Computer and Communication Technology (ICCC-2011)*
14. F. Ahmed, A.S.M.H. Bari, A. Shihavuddin, H.A. Al-Mamun, P. Kwan, A study on local binary pattern for automated weed classification using template matching and support vector machine, in *2011 IEEE 12th International Symposium on Computational Intelligence and Informatics (CINTI)* (2011)
15. M. Islam, M. Billah, M. Yousuf, Automatic plant detection using HOG and LBP features with SVM. *Int. J. Comput. (IJC)* **33**, 26–38 (2019)
16. C.A. Priya, T. Balasarayanan, A.S. Thanamani, An efficient leaf recognition algorithm for plant classification using support vector machine, in *International Conference on Pattern Recognition, Informatics and Medical Engineering (PRIME-2012)*, Salem, Tamilnadu, 2012, pp. 428–432
17. S. Prasad, K. Mohan, R. Tripathi, Relative sub-image based features for leaf recognition using support vector machine, in *ACM International Conference Proceeding Series*, 2011, pp. 343–346. <https://doi.org/10.1145/1947940.1948012>
18. T. Munisami, M. Ramsurn, S. Kishnah, S. Pudaruth, Plant leaf recognition using shape features and colour histogram with k-nearest neighbour classifiers. *Proc. Comput. Sci.* **58**, 740–747 (2015). <https://doi.org/10.1016/j.procs.2015.08.095>
19. A. Gupta, B.S. Rai, Recognition of plants by leaf image using nearest neighborhood classification. *Int. J. Technol. Res. Eng.* **1**(9) (2014)
20. M. Shabanzade, M. Zahedi, S. Aghvami, Combination of local descriptors and global features for leaf recognition. *Signal Image Process. Int. J.* **2** (2011). <https://doi.org/10.5121/sipij.2011.2303>
21. Sahay, M. Chen, Leaf analysis for plant recognition, in *2016 7th IEEE International Conference on Software Engineering and Service Science (ICSESS)*, Beijing, 2016, pp. 914–917
22. S. Abbasi, F. Mokhtarian, J. Kittler, Reliable classification of chrysanthemum leaves through curvature scale space (1997)
23. J. Chaki, R. Parekh, S. Bhattacharya, Plant leaf recognition using texture and shape features with neural classifiers. *Pattern Recogn. Lett.* **2015**(58), 61–68 (2015). <https://doi.org/10.1016/j.patrec.2015.02.010>

24. C. Xia, J.-M. Lee, Y. Li, Y.-H. Song, B.-K. Chung, T.-S. Chon, Plant leaf detection using modified active shape models. *Biosys. Eng.* **116**, 23–35 (2013). <https://doi.org/10.1016/j.biosystemseng.2013.06.003>
25. A. Ehsani Rad, Plant classification based on leaf recognition. *Int. J. Comput. Sci. Inf. Secur.* **8** (2010)
26. D. Shitole, F. Tamboli, K. Raj, Ayurvedic herb detection using image processing. *Int. J. Trend Sci. Res. Dev.* **3**, 491–494 (2019). <https://doi.org/10.31142/ijtsrd23605>
27. A. Backes, D. Casanova, O. Bruno, Plant leaf identification based on volumetric fractal dimension. *IPRAI* **23**, 1145–1160 (2009). <https://doi.org/10.1142/S0218001409007508>
28. C. Arrasco, S. Khlebnikov, A. Oncevay, C.B. Castanon, Leaf venation enhancing for texture feature extraction in a plant classification task, in *2018 IEEE Latin American Conference on Computational Intelligence (LA-CCI)*, Guadalajara, Mexico, 2018, pp. 1–4
29. B. Bama, S.M. Valli, S. Raju, V.A. Kumar, Content based leaf image retrieval (CBLIR) using shape, color and texture features. *Indian J. Comput. Sci. Eng.* **2**, 202–211 (2011)
30. P.K. Mishra, S.K. Maurya, R.K. Singh, A.K. Misra, A semi automatic plant identification based on digital leaf and flower images, in *IEEE-International Conference On Advances In Engineering, Science And Management (ICAESM -2012)*, Nagapattinam, Tamil Nadu, 2012, pp. 68–73
31. J. Hu, Z. Chen, M. Yang, R. Zhang, Y. Cui, A multiscale fusion convolutional neural network for plant leaf recognition. *IEEE Signal Process.* (2018)
32. R.X. Hu, W. Jia, H.B. Ling et al., Multiscale distance matrix for fast plant leaf recognition. *IEEE Trans. Image Process.* **21**(11), 4667–4672 (2012)
33. C. Zhang, P. Zhou, C. Li, L. Liu, A convolutional neural network for leaves recognition using data augmentation, in *2015 IEEE International Conference on Computer and Information Technology; Ubiquitous Computing and Communications; Dependable, Autonomic and Secure Computing; Pervasive Intelligence and Computing*, Liverpool, 2015, pp. 2143–2150
34. G.L. Grinblat, L.C. Uzal, M.G. Larese, P.M. Granitto, Deep learning for plant identification using vein morphological patterns. *Comput. Electron. Agric.* **127**, 418–424 (2016)
35. E. Kang, I. Oh, Weak constraint leaf image recognition based on convolutional neural network, in *2018 International Conference on Electronics, Information, and Communication (ICEIC)*, Honolulu, HI, 2018, pp. 1–4
36. S. Prasad, P.P. Singh, Medicinal plant leaf information extraction using deep features, in *TENCON 2017–2017 IEEE Region 10 Conference* (2017)
37. S.H. Lee, C.S. Chan, P. Wilkin, P. Remagnino, Deep-plant: plant identification with convolutional neural networks, pp. 452–456 (2015)
38. A. Devaraj, K. Rathan, S. Jaahnavi, K. Indira, Identification of plant disease using image processing technique, in *2019 International Conference on Communication and Signal Processing (ICCSP)*, Chennai, India, 2019, pp. 0749–0753
39. M. Mehdipour Ghazi, B. Yanikoglu, E. Aptoula, Plant identification using deep neural networks via optimization of transfer learning parameters. *Neurocomputing* **235** (2017)
40. Y.R. Azeez, C. Rajapakse, An application of transfer learning techniques in identifying herbal plants in Sri Lanka, in *2019 International Research Conference on Smart Computing and Systems Engineering (SCSE)*, Colombo, Sri Lanka, 2019, pp. 172–178
41. J. Tan, C. Siow-Wee, S. Abdul Kareem, H.J. Yap, Y.-K. Thai, Deep learning for plant species classification using leaf vein morphometric, in *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 2018, pp. 1
42. C. Amuthalingeswaran, M. Sivakumar, P. Renuga, S. Alexpandi, J. Elamathi, S.S. Hari, Identification of medicinal plant's and their usage by using deep learning, in *2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI)*, Tirunelveli, India, 2019, pp. 886–890
43. A. Aakif, M. Khan, Automatic classification of plants based on their leaves. *Biosyst. Eng.* **139**, 66–75 (2015)
44. B. Şekeroğlu, Y. Inan, Leaves recognition system using a neural network. *Proc. Comput. Sci.* **102**, 578–582 (2016)

45. Y.F. Li, Q.S. Zhu, Y.K. Cao, C.L. Wang (n.d.), A leaf vein extraction method based on Snakes technique, in *2005 International Conference on Neural Networks and Brain*
46. Z.Q. Zhao, B.-J. Xie, Y.-M. Cheung, X. Wu, Plant leaf identification via a growing convolution neural network with progressive sample learning, in *ACCV 2014*
47. S. Albawi, T. Abed Mohammed, S. Alzawi, *Understanding of a Convolutional Neural Network*. <https://doi.org/10.1109/ICEngTechnol.2017.8308186> (2017)
48. D. Kingma, J. Ba, Adam: a method for stochastic optimization, in *International Conference on Learning Representations*
49. S. Kannoja, G. Jaiswal, Effects of varying resolution on performance of CNN based image classification an experimental study. *Int. J. Comput. Sci. Eng.* **6**, 451–456 (2018). <https://doi.org/10.26438/ijcse/v6i9.4514556>
50. A. Mikołajczyk, M. Grochowski, Data augmentation for improving deep learning in image classification problem, in *2018 International Interdisciplinary PhD Workshop (IIPHDW)*, Swinoujście, 2018, pp. 117–122. <https://doi.org/10.1109/IIPHDW.2018.8388338>
51. D. Scherer, A. Müller, S. Behnke, *Evaluation of Pooling Operations in Convolutional Architectures for Object Recognition*, pp. 92–101. https://doi.org/10.1007/978-3-642-15825-4_10
52. <https://alexlenail.me/NN-SVG/AlexNet.html>. Accessed on: 27/09/2020

Design and Implementation of Multi-class Logistic Regression for Effective Classification of Low, Medium and High Risk Lung Cancer Problem



Shivaprasad, P. Mahabaleshwara Bhat, and C. Naveena

1 Introduction

Although taking into account other types of cancers in the USA, lung cancer is probably the deadliest cancer. Every year, it takes more lives than combined colon, prostate, ovarian and breast cancer [1]. In the lungs, two organs in the chest that take in oxygen when inhaled and release carbon dioxide upon exhalation, this form of cancer starts. People who are addicted to smoking are more likely to have lung cancer. People who do not smoke may also develop lung cancer, but the risk is very low. Any of the signs and symptoms of lung cancer include excessive coughing, bleeding while coughing, shortness of breath, chest pain, weight loss without dieting. By destroying the cells that line the lungs, doctors say smoking causes lung cancer. When you inhale cigarette smoke that is full of cancer-causing substances (carcinogens), changes in lung tissue begin almost instantly. Our body would be able to repair this damage initially. But normal cells that line your lungs are progressively weakened with each repeated exposure. Over time, the damage causes cells to behave abnormally and cancer may develop eventually.

1.1 Types of Lung Cancer

Based on the presence of lung cancer cells under the microscope, doctors classify lung cancer into two main forms. Depending on which major form of lung cancer you have, your doctor makes treatment decisions.

Small cell lung cancer: Lung cancer of small cells occurs almost exclusively in heavy smokers and is less common than lung cancer of non-small cells.

Shivaprasad (✉) · P. Mahabaleshwara Bhat · C. Naveena
Electronics and Communication, Vivekananda College of Engineering and Technology, Puttur,
India

Non-small cell lung cancer: For many types of lung cancers that act in a similar way, non-small cell lung cancer is an umbrella word. Squamous cell carcinoma, adenocarcinoma and large-cell carcinoma are non-small cell lung cancers [2].

1.2 Lung Cancer Dataset

The lung cancer dataset is collected form Data World platform (<https://data.world/cancerdatahp/lung-cancer-data>).

The database consists of 25 attributes, 1000 instances of lung cancer, 365 instances of patient with High Risk level, 303 instances of patient with Low Risk level and 332 instances of patient with Medium Risk levels.

1.3 Attribute Information

S. No.	Attribute	Value
1	Patient_id	String
2	Age	Integer
3	Gender	Integer
4	Air_pollution	Integer
5	Alcohol_use	Integer
6	Dust_allergy	Integer
7	Occupational_hazards	Integer
8	Genetic_risk	Integer
9	Chronic_lung_disease	Integer
10	Balanced_diet	Integer
11	Obesity	Integer
12	Smoking	Integer
13	Passive_smoker	Integer
14	Chest_pain	Integer
15	Coughing_of_blood	Integer
16	Fatigue	Integer
17	Weight_loss	Integer
18	Shortness_of_breath	Integer
19	Wheezing	Integer
20	Swallowing_difficulty	Integer
21	Clubbing_of_finger_nails	Integer

(continued)

(continued)

S. No.	Attribute	Value
22	Frequent_cold	Integer
23	Dry_cough	Integer
24	snoring	Integer
25	Level	String

The first five patients' information from the complete lung cancer dataset is addressed in Fig. 1.

2 Literature Survey

Parikh et al. in [3], described the paper, Lung Cancer in India Current status and promising strategies. The paper includes cancer data collected from different states in year 2011. It was observed that the number patients with lung cancer was more in number in the Maharashtra state. Also the author predicted the raise in the number of patients with lung cancer positive by 2020 by artificial intelligence system.

Mohamed et al. in [4] presented a paper, Early detection and prediction of lung cancer using machine learning algorithms. The paper addressed many ML algorithms including Naive Bayes, logistic regression, neural network, KNN classifier, decision tree, SVM, and random forest. It was observed that the accuracy obtained using SVM and logistic regression was best (around 94%) compared to other classification algorithm.

Timor and Gleeson in [5], explored Lung cancer prediction model using machine learning and advance imaging techniques. Paper highlighted the nodule effective classification for lung cancer detection using convolutional neural network (CNN) and support vector machine (SVM) classifier. Dataset A included size-matched data; Dataset B included unmatched data. While investigating the role of nodule size within a machine learning model of nodule malignancy using SVM, an accuracy of 70% on Dataset A and 91% of accuracy on Dataset B was observed.

Nair et al. in [6] proposed a paper, Comparative Study of Lung Cancer Detection using Machine Learning Algorithms. The paper included attributes for classification such as Age, Gender, Air Pollution, Alcohol use, Dust Allergy, Occupational Hazards, Genetic Risk, Chronic Lung Disease, Balanced Diet, Obesity, Smoking, Passive Smoker, Chest Pain, Coughing of Blood, Fatigue, Weight Loss, Shortness of Breath, Wheezing, Swallowing Difficulty, Clubbing of Finger Nails, Frequent Cold,

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
1	Patient ID	Age	Gender	Air Pollution	Alcohol use	Dust Allergy	Occupational Hazards	Genetic Risk	Chronic Lung Disease	Balanced Diet	Obesity	Smoking	Passive Smoker	Chest Pain	Coughing of Blood	Fatigue	Weight Loss	Shortness of Breath	Wheezing	Swallowing Difficulty	Clubbing of Finger Nails	Frequent Cold	Dry Cough	Snoring	Level
2	P1	33	M	2	4	5	4	3	2	2	4	3	2	2	4	3	4	2	2	3	1	2	3	4	Low
3	P2	27	F	3	3	6	3	4	2	2	2	2	3	1	3	7	8	6	2	1	7	2	Medium	2	Medium
4	P3	35	M	4	5	6	5	5	4	6	7	2	3	4	8	8	7	9	2	1	4	6	7	2	High
5	P4	32	F	2	2	7	7	6	7	7	7	7	7	8	4	2	3	1	4	5	6	7	5	High	5
6	P5	48	M	3	4	8	7	7	7	6	7	7	8	7	7	9	3	2	4	3	4	2	4	2	3

Fig. 1 Lung cancer dataset, first five patients' information

Dry Cough, and Snoring. Different classification approaches such as decision tree, Naïve Bayes, logistic regression and SVM were applied upon the above attributes. The accuracy was best while using SVM and was less for logistic regression.

3 Proposed Hypotheses for Multi-class Logistic Regression

The lung cancer dataset includes 23 attributes; hence, the hypothesis for lung cancer detection is of the form,

$$h_{\theta}(x) = \frac{1}{(1 + e^{-\theta^T x})} \quad (1)$$

where

$$\theta^T x = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \theta_3 x_3 + \dots + \theta_{23} x_{23} \quad (2)$$

Here $x_1, x_2, x_3, \dots, x_{23}$ are 23 attributes assisting lung cancer.

The cost function for logistic regression is given by [7]

$$J(\theta) = -\frac{1}{m} \sum_{i=1}^m y^i \log(h_{\theta}(x^i)) + (1 - y^i)(1 - \log(h_{\theta}(x^i))) \quad (3)$$

where m is total number of input instances.

The objective here is to minimize the cost function parameterized by θ , using gradient descent rule [8, 9].

ie. $\min_{\theta} J(\theta)$, where θ_j is computed as follows

$$\theta_j := \theta_j - \alpha \frac{\delta J}{\delta \theta_j}; \quad 0 \leq j \leq 23 \quad (4)$$

Here partial derivatives of cost function parameterized by θ are estimated as follows

$$\frac{\delta J(\theta)}{\delta \theta_0} = \frac{1}{m} \sum_{i=1}^m (h_{\theta}(x^i) - y^i) \quad (5)$$

$$\frac{\delta J(\theta)}{\delta \theta_j} = \frac{1}{m} \sum_{i=1}^m (h_{\theta}(x^i) - y^i) x_j^i; \quad 1 \leq j \leq 23 \quad (6)$$

Algorithm:

Step 1 Initialize all weights θ 's to zero.

- Step 2 Use Eq. (2) to estimate $\theta^T x$.
- Step 3 Obtain the hypothesis $h_\theta(x)$ from Eq. (1).
- Step 4 Compute the cost function $J1(\theta)$ by setting the High Risk Level to 0's and Low, Medium Risk level's to 1's.
- Step 5 Compute the cost function $J2(\theta)$ by setting the Low Risk Level to 0's and High and Medium Risk level's to 1's.
- Step 6 Compute the cost function $J3(\theta)$ by setting the Medium Risk Level to 0's and Low, High Risk level's to 1's.
- Step 7 Compute the gradients from Eqs. (5) and (6) and update the weights using Eq. (4).
- Step 8 Use the updated weights to evaluate cost function $J1(\theta)$, $J2(\theta)$ and $J3(\theta)$.
- Step 9 Go to step 7 until the cost function converges [10].

4 Results and Discussions

Logistic regression and KNN classifier were implemented from scratch without the use of sklearn machine learning package in Python3.6 environment. NumPy Matplotlib essentials for enhancing the speed of the model, data visualization and interpretation were imported while designing the system.

Case (i) Complete lung cancer dataset is considered for classification, without decomposing the data into training set and testing set. The size of the dataset was 1000 samples.

From Table 1, it can be observed that as the number of K neighbors increases the accuracy on different levels of risk pertaining to lung cancer decreases. However, more neighbors need to be considered for effective realization of KNN algorithm on the dataset. Figure 2 shows the relationship between the number of K neighbors and accuracy on different risk levels of lung cancer dataset.

Table 2 provides the accuracy obtained using logistic regression with learning rate set to 0.09. It can be observed that as the number iterations increases, accuracy increases. However, increasing the iterations, more often leads performance damage.

Table 1 KNN classifier applied on lung cancer dataset

Classifier KNN(K)	Accuracy high risk class (365)	Accuracy low risk class (303)	Accuracy medium risk class (332)	Overall accuracy
$K = 2$	100	99.67	100	99.89
$K = 5$	100	99	100	99.67
$K = 10$	100	99	100	99.67
$K = 15$	100	99	100	99.67
$K = 20$	94.8	85.8	82.2	87.6
$K = 25$	100	86.1	82.5	89.5
$K = 30$	100	85.8	82.5	89.44

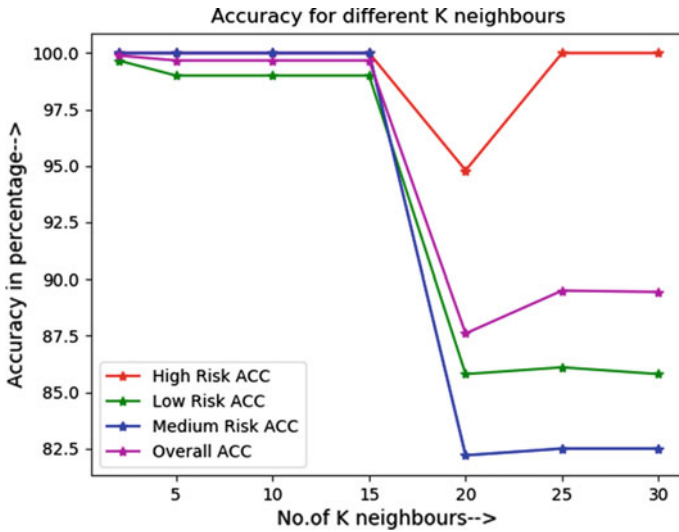


Fig. 2 No. of neighbors versus accuracy for KNN classifier

Table 2 Logistic regression applied on lung cancer dataset

Number of iterations	Min cost function in high, low, medium classes	Accuracy high risk class (365)	Accuracy low risk class (303)	Accuracy medium risk class (332)	Overall Acc
500	0.0384, 0.0635, 0.2321,	100	96.7	90.97	95.88
1000	0.0226, 0.0392, 0.2122	100	99.7	90.97	96.87
1500	0.01614, 0.02842, 0.2019	100	100	93.67	97.89
2000	0.0125, 0.0222, 0.1947	100	100	93.67	97.89
10,000	0.0028, 0.0051, 0.146	100	100	100	100

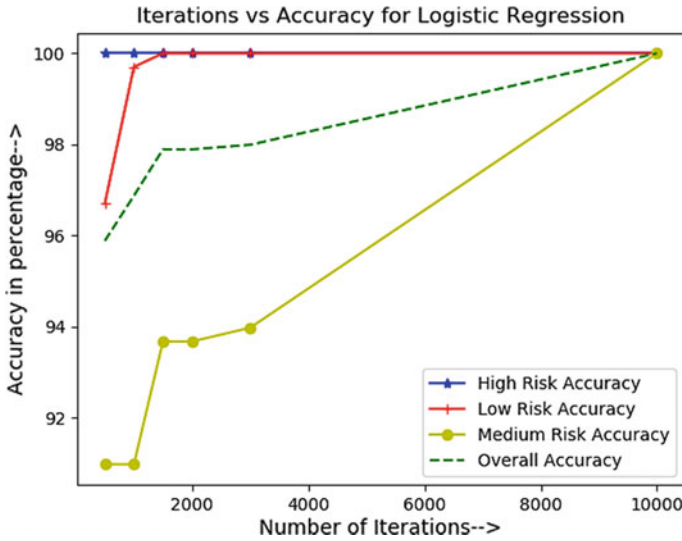


Fig. 3 Iterations versus accuracy for multi-class logistic regression

By repeating the algorithm to perform 10,000 times, 100% accuracy was achievable. Figure 3 shows the relationship between the number of iterations and accuracy using logistic regression on different risk levels of lung cancer data set.

Case (ii) Lung cancer dataset was decomposed into training dataset and testing dataset. A total of 800 samples and 200 samples were considered for training and testing the dataset, respectively. This approach is employed for logistic regression only. Figures 4 and 5 show the accuracy versus iterations plotted for the training and testing lung cancer dataset using multi-class logistic regression with learning rate set to 0.09.

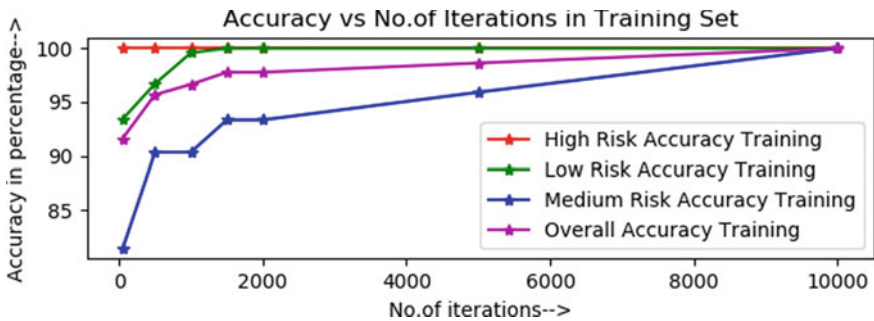


Fig. 4 Iterations versus accuracy for multi-class logistic regression on training dataset

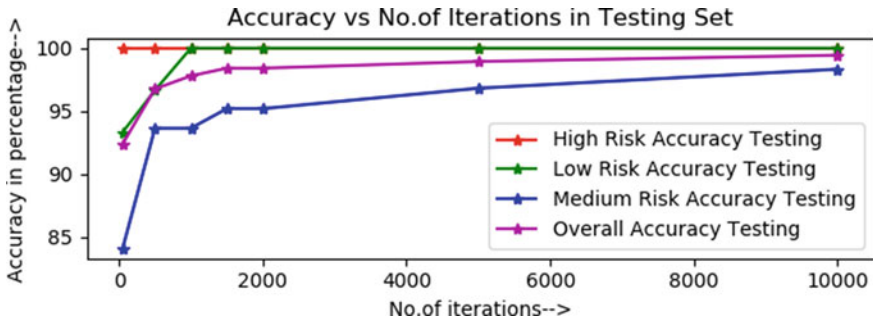


Fig. 5 Iterations versus accuracy for multi-class logistic regression on testing dataset

5 Conclusion and Future Work

The paper provides two learning algorithms for effective classification of lung cancer disease. A total of 23 attributes which facilitates lung cancer disease for 1000 patients were used as input for the classification algorithms. Upon applying the two algorithms on the lung cancer dataset, it was observed logistic regression performed better than KNN for larger iterations. To further increase the accuracy neural network, deep learning architecture could also be used. When it comes to deep learning, the dataset available does not have massive data for effective classification; total data available in this lung cancer dataset is $23 \times 1000 = 23,000$. Neural network may give the best accuracy, but it may require more hidden layer, which definitely degrades the performance.

Acknowledgements We appreciate the Data World platform for sharing and availing the survey lung cancer dataset <https://data.world/cancerdatahp>. Specifically, we appreciate *Prithivraj* for contributing the dataset to Data World.

References

1. World Cancer Day 2020: 1 in 10 Indians Likely to Develop Lung cancer in their Lifetime, Says Study, Feb. 4, 2020. Accessed on Nov 21, 2020 [Online]. Available <https://www.news18.com/news/india/world-cancer-day-2020-1-in-10-indians-likely-to-develop-cancer-in-their-lifetime-says-study-2485857.html>
2. Mayo Clinic Family Health Book, Fifth edn. Rochester, US, 2018
3. P.M. Parikh, A.A. Ranade, B. Govind, N. Ghadyalpatil, R. Singh, R. Bharath, Lung cancer in India: Current status and promising strategies. *South Asian J. Cancer*, 93–95 (2016)
4. A. Mohamed, B. Thar, D. Al-Jumeily, A. Hussain, A. Aljaaf, M. Jamila, Early detection and prediction of lung cancer using machine-learning algorithms applied on a secure healthcare data-system architecture. *Mach. Learn. Comput. Cyber Secur.* 233–257 (2019)
5. K. Timor, F. Gleeson, Lung cancer prediction using machine learning and advanced imaging techniques. *Transl. Lung Cancer Res.* 7, 304–312 (2018)

6. P.R. Radhika, R.A.S. Nair, G. Veena, A comparative study of lung cancer detection using machine learning algorithms, in *2019 IEEE International Conference on Electrical, Computer and Communication Technologies (ICECCT)*, Coimbatore, India, 2019, pp. 1–4
7. V. Alex Smola, *Introduction to Machine Learning* (Cambridge University Press, UK, 2008)
8. T.M. Mitchell, *Machine Learning* (McGraw-Hill, Maidenhead U.K., 1997)
9. C.M. Bishop, *Pattern Recognition and Machine Learning* (Springer, New York, 2006)
10. R.M.C.R. Souza, C. Francisco, Q. Diego, F. Roberta, A multi-class logistic regression model for interval data, in *2018 IEEE International Conference on Systems, Man and Cybernetics*, pp. 1253–1258 (2008)

Voice Controlled IoT Based Grass Cutter Powered by Solar Energy



Mahadevaswamy, Kiran Kumar Humse, K. Chethan, and K. V. Sudheesh

1 Introduction

It is a well-known fact that long ago; the grass was being cut by using blades which highly required time and manpower. Later grass cutters which used fuels came into the market. These used gasoline for their operation which pollutes the environment. These caused severe health effects on human beings. These lawn mowers were massive in size and required manpower for their operation. In today's world, time is the most precious thing. No one has the time to spend on holding the lawn movers and cut the grass. Spending time on lawn movers can be avoided, and for this to happen, this work is planned to make the lawn mower more convenient with more efficiency and can be easily operated and eco-friendly.

These devices use solar energy as their main supply and contribute to reducing the pollution that may be either sound pollution or noise pollution. The proposed device helps in overcoming all the problems and drawbacks of the conventional grass cutters to some extent. During cloudy weather, these devices are operated through batteries, which are charged using the solar panel when there is lots of sunlight.

Mahadevaswamy (✉) · K. K. Humse · K. Chethan · K. V. Sudheesh
Electronics and Communication Engineering, Vidyavardhaka College of Engineering, Mysuru,
Karnataka, India
e-mail: mahadevaswamy@vnce.ac.in

K. K. Humse
e-mail: kirankumarhumse@vnce.ac.in

K. Chethan
e-mail: chethank@vnce.ac.in

K. V. Sudheesh
e-mail: sudheesh.kv@vnce.ac.in

The organization of the paper is as follows; Sect. 2 gives a brief literature survey about traditional speech separation methods. Section 3 describes the methodology followed in this paper. Results are presented in Sect. 4 and finally concluded in Sect. 5.

2 Literature Survey

The grass cutting robot is designed with Atmega328 as a core controller along with solar panel, DC motor for the cutter, batteries, LCD screen, 2D laser scanner, infrared sensor, Blade, and robot wheel with DC motors and gears. The laser scanner is adapted in front of the robot to detect the obstacle. It directs the microcontroller to rotate the vehicles motor in the direction wherever no obstacle is detected. The evaluation of area uncut by the robot is measured for a different location and considered successful after the maximum duration. In this work, the only semi-structured environment is taken into consideration for solar grass cutting machine [1].

The auto-energy-saving robotic lawn mower is developed using ARM9 core family based S3C2450 which is a RISC processor with low power consumption and high-performance application. The robotic lawn mower is tested at real-time to check the obstacle avoidance error rate. The several experimental conditions are taken into account if position error in x -axis and y -axis is too large then heading error is also found maximum and is minimum for lower error rate concerning x - and y -axis. The proposed method provides the correct specified point of angle for which to avoid the obstacles to prevent the occurrence of collision [2].

The design of lawn mower is made up of sensors and transducers, programmable logic controller, analyzer, and actuators and drive system. The prototype model is developed using the above design components. The final prototype is installed and tested, and the result was found that the lawn mower is successfully changing the direction against obstacles after several trials were performed. Further improvement can be seen in the device by replacing the geared motors which are slow, balancing sensors in unflated surface, the solar panel can be incorporated, and multifunctional aspects can also be done, for example, transporting objects, alarming at a condition, etc. [3].

The lawn mower setup consists of an induction motor, a battery, an alternator, and a link mechanism. Casing, handle frames, lift roads, collapsed blades, tires, and collection bags form the external construction of the lawn mower. The design of the lawn mower is discussed briefly in a mathematical model, and the performance test given its cutting efficiency almost near to 90% with least minimal human effort. The result is obtained by plotting the graph force required in Newton versus the effect of mower handle angle on maneuverability in degrees. The work found that beyond 40°, lawn mower is not easy to handle. Above 40°, the lawn mower can operate smoothly in a suitable condition [4].

This paperwork deals with path tracking, localization, and obstacle detection. The subsystem consists of hardware design, software algorithms, user interface, and system architecture. Hardware design consists of a two-axis laser scanner used for obstacle detection. Obstacle detection algorithms made of three functions such as classification, fusing, and filtering. System architecture is incorporated with a laser range finder, a GPS, gyro, and encoders. Software architecture is incorporated with CPU and communication modules along with shared memory. The obstacle detection performance tested for two situations, namely flat terrain and curved terrain. Localization system performances are also tested during GPS outages and evaluated [5].

This paper discusses the autonomous lawn mower, which is presented for ION lawn mower competition at Miami University. The competition consists of design and testing lawn mower devices of known shape and enabled global positioning system with no local remote control. The design of red blade 3 lawn mower consists of sensing, controlling, monitoring, and testing systems. Sensing part is incorporated with custom DGPS, obstacle detection, wheel encoders, and digital compass. Control parts are incorporated with microprocessor and control software, DC motors, and control drive system. Monitoring and testing include remote base station. Control algorithms composed of five divisions include initialization process which provides user inputs, path control which provides the capability of changing the path for obstacles detection: orientation change, obstacle detection, and position change. Further improvement can be made for this lawn mower by replacing the onboard computer with high processing compatible microprocessor [6].

The lawn mower consists of diodes, resistors, relays, LM7805, PIC16F876A, resistors, SIM 800 GSM module, transistors, IFR3205, relays, 4 MHz crystal oscillator, 9 W DC motor, two units of 18 W DC motor, plastic casing, mild steel blade, 1/4" bolts and nuts, 16 × 2 LCD display, pushbuttons, transistors, MOSFET, two units of 18Ah batteries, solar panel, ultrasonic sensor, and wheel encoder. In this work, we design and construct an automated lawn cutter that can cut lawns without the need of building an obstacle or boundary wire around the field to be mowed [7].

The solar grass cutter consists of solar panel, batteries, brushless DC motor, solar charger, circuit breaker, and blades. Photovoltaic principles can be carried out in silicon as a good semiconductor which forms solar panel to absorb radiation from sunlight. Batteries are used as a storage device and supplier of power. DC motor is for controlling the speed of the motor. Circuit breaker act as an electrical switch to protect the circuit from damage due to overload or short circuit. The Scotch yoke mechanism is used in which linear motion of the slider is converted into rotational motion or vice versa. Further improvement can be made by replacing the Scotch yoke mechanism because of less efficiency. Heavy material can also replace with lightweight material in order to increase the efficiency of the DC motor [8].

This is totally based on Internet of Things (IoT) and robotics. A special feature of this grass cutter is that it can be controlled from any part of the world because it is connected from the Internet. We can even control it from our cell phone, which is a very common gadget installed in everyone pockets. In this device, we use ARDUINO IDE software to do all the programming of this device even for the slightest movement of the wheel. An IR and ultrasonic sensor are used to detect the location of the obstacle and to inform the arm to respond. This device is installed with many features such as robotic arm to remove obstacle in its path, ultrasonic sensors to sense an obstacle, solar panels to charge battery, and many more. Also, the main feature given here is that we can trim the grass at a different level according to our choice by using a mobile phone [9].

In this paper, the author gives his review of designing lawn mower with the best use of comparing the blades design, power consumption, and engine type to provide efficient and cost-effective grass cutting machine. There are two types of lawn mower which can be designed with respect to the axis of the rotation of the blade. The author found reel lawn mower is better than the rotary lawn mower among their types of an axis, i.e., horizontal and vertical, respectively. Solar-powered lawn mower design consists of a microcontroller, sensors, battery powered by sunlight, Android smartphone, ZigBee module, and a graphical user interface using MATLAB. The author concluded giving importance to low cost which enabled manual mower is good where the operating cost may be expected to reduce in future days [10]. The block diagram of the design consists of Arduino UNO, solar panel, battery, ultrasonic sensors, and motor driver. Atmega328 is a main 8-bit microcontroller based on Arduino UNO and have six channels 10-bit A/D converter. Arduino UNO is an open-source platform for building the projects. Input voltage level for this microcontroller is 7–12 V. Ultrasonic sensors are for detecting the obstacles by making a small ultrasonic burst and listening to its echo. 150 rpm geared motor is used for robot wheels. L293D motor driver is 16-bit IC for controlling the direction of motors. Further improvements can be made by modifying the size, efficiency by providing good capacity battery, and more sensors are accommodated for better results [11].

The system was made up of Arduino UNO as a microcontroller, DC motor, PIR sensors, solar panel, and a lithium-ion battery. The simple algorithm is used here if the PIR sensor detects any input stop moving the grass cutter until input is zero. It is otherwise moving process continuously. This is a simple and much efficient way for cutting the grass [12, 13].

3 Methodology

In this grass cutter, Node MCU microcontroller is like a brain as it controls the entire working of the device. The batteries used here are rechargeable which uses a monocrystalline solar panel of 15 W capacity. This type of solar panel is one of the best types of solar panel which are available in the market. This monocrystalline

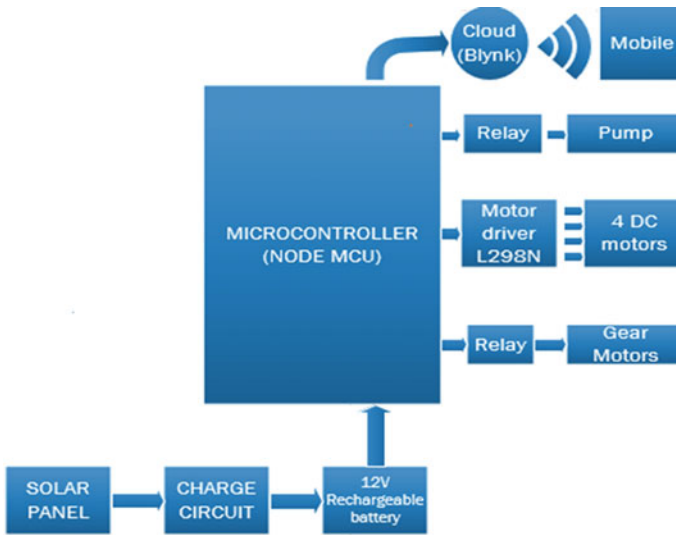


Fig. 1 Architecture of IOT based grass cutter

solar panel has the capacity to withstand harsh weather and climate conditions. So the user need not worry about the device even if it exposed to harsh conditions while the device is in operation. These batteries supply 12 V DC for the functioning of the motor. Brushless DC motor is used which run at a speed of 150–250 rpm.

The design contains microcontroller, DC motor, motor driver, relay, solar panel, and a rechargeable battery, and the microcontroller used here has an ability to connect to a network and also communicates with the prescribed network. The microcontroller is programmed using specific software's and is controlled by the Android/IOS with the help of specific apps [14]. DC motors are attached with wheels and connected to a motor driver whereupon the specific command or pulse cutter is carried on, i.e., forward, backward, right, and left movements. A DC motor is connected to a relay for the rotation movement of the blades which are attached to the DC motor. A solar panel is used to absorb solar energy. A battery which is rechargeable is charged with the help of the charging circuit by the solar energy. The power supply to the whole circuit is given through the 12 V rechargeable battery (Fig. 1).

3.1 Components and Features

Solar panel: Monocrystalline or single jewel sun situated PV sheets are one of the most settled, for the most part reliable, and most beneficial ways to deal with make power from sun arranged essentialness. Here, each PV module is produced from a single silicon valuable stone. The silicon is refined, mollified, and a short time later set into ingots, which are then cut into humble wafers to make solitary cells. The regular

shade of monocrystalline PV module is dull or splendid blue. Monocrystalline sun-based sheets are usually utilized for tremendous extension sun-based applications, for instance, business and private daylight based foundations.

Blynk application for NodeMCU: Blynk application is important and easy to use software with IOS and applications based on the Android OS. Blynk app is utilized to command and control Raspberry Pi, Arduino, and Node MCU using the Web. It is the main dashboard where one can accumulate sensible interfacing. There are three gigantic parts in the Blynk app stage which are:

Blynk App: This helps us in creating user-friendly applications for many projects using different various available widgets.

Blynk Server: This is mainly responsible for the communication which is made between the smartphone and hardware. This can be accomplished by using the Blynk cloud. It is open-source which is capable of easy handling of many IoT devices.

Blynk Libraries: for all the famous equipment stages empower correspondence with the server and procedure all the approaching and outcoming commands.

Battery: A battery device used to build the automated grass cutter is charged using solar energy through solar panel and charging cables. Once the battery is charged to its full capacity, the stored charge can be utilized to run the automated grass cutter. The robot can run till some minimum energy in the battery, it can halt the grass cutting task, the battery can be recharged, and the grass cutting process can be repeated.

The use of multiple batteries can encourage the grass cutting operation to be continuous.

Camera: United vision offers a wide scope of CMOS cameras from the minimal effort to elite cameras with cutting edge highlights for the most requesting applications from applied autonomy to quality control, modern investigation, observation, traffic, coordination, and considerably more. Partnered vision cameras include the most recent top-notch CMOS sensors from prestigious producers which are to guarantee magnificent picture quality. Highlights incorporate VGA to 5 Megapixel goals, Up to 550 casings for every second, Sony, CMOSIS, e2V, and On Semi Python CMOS sensors, GigE Vision or USB3 vision attachment and play interface, 3-year guarantee.

Internet of Things (IoT): The IoT is a strategy of interrelated computational systems and devices, digital and mechanical machines labelled with separate identifiers and capacity exchange information over networks with almost less or no involvement of humans such as humans to humans or humans to machines. These IDs are abbreviated as UIDs. The IoT definition has resulted due to the amalgamation of several technologies such as data analytics, machine learning, sensors and networks, and embedded systems. IoT technology normally refers to smart applications such as house automation, smart mobile phones, smart watch, smart health bands, smart pen, smart lock, smart boards, smart lighting systems, smart fan, smart robots, and smart audio speakers. There are cons present in IoT technology. The most popular limitation is the privacy and security issue over IoT technology. IoT technology finds its usage in various fields such as commercial, infrastructure, industrial, and consumer applications. IOT technology enables devices and also considers their applications in the healthcare industry, particularly in hospitals. It also finds applications in the

field of agriculture, energy management systems, military equipment, exploration of oceans, and remote control. The robot is interfaced with the Google voice assistant software using Blynk and IFTTT software. The robot is thus capable of working according to the English voice commands such as move front, moving back, left turn, right turn, start, and stop.

Node MCU Microcontroller: NodeMCU v3 is an advancement board which runs on the ESP8266 with the Espressif non-operating system SDK and equipment are dependent on the ESP-12 module. The gadget highlights 4 MB of glimmer memory, 80 MHz of framework clock, around 50k of usable SRAM, and an on-chip WiFi handset. Capacity to the Node MCU v3 is provided through the on-board USB smaller scale B connector or straightforwardly by means of the “VIN” pin. The force source is chosen consequently. The gadget can work on an outside gracefully of 6 to 20 V. In the event that it is utilizing more than 12 V, the voltage controller may overheat and harm the gadget. The prescribed range is 7–12 V. The NodeMCU v3 uncovered the sequential port of the ESP8266 module through CH340 USB to UART chip which is likewise associated with the boot pins of the module, taking into account consistent virtualization of the gadget.

Engine driver: L293D is a customary engine driver or engine driver IC which licenses DC engine to drive on either course. L293D is a 16-pin IC which can control a lot of two DC engines all the while toward anyway. It construes that you can control two DC engine with a solitary L293D IC. It pursues the chance of H-interface. H-partner is a circuit which permits the voltage to be flown in either bearing. As you in all likelihood realize voltage needs to adjust its course for having the choice to turn the engine clockwise or anticlockwise way, subsequently H-partner IC is perfect for driving a DC engine.

Relay: It goes about as an exchanging circuit. It is utilized for killing on and a specific circuit when different circuits must be constrained by a solitary sign. Here, we use it as a switch for the engine drivers.

4 Experimental Results

In Fig. 2, we can see the robot that has been completely build for its purpose. In Figs. 3 and 4, we can see that the robot has been tested, and we have recorded positive results.

Here, in Figs. 5, 6 and 7, we can see some of the images and screenshots of the Blynk application which is used to control the device. This application is made to install in our smartphones.

Here we can see Blynk dash board which is displayed. After this, we are supposed to create a new project using this application.

Figures 8, 9, 10, 11 and 12 show how the inputs are given to the Google assistant and how it responds to the voice commands (Figs. 13 and 14).

Table 1 describes the specifications of the voice controlled IOT based grass cutter.

Fig. 2 Grass cutting robot



Fig. 3 Rotating blades



Fig. 4 Grass has been cut



5 Conclusion

An automated grass cutter is developed using IOT technology and solar energy. The robot is able to cut the grass reliably and make the forward movement, backward

Fig. 5 Blynk dash board

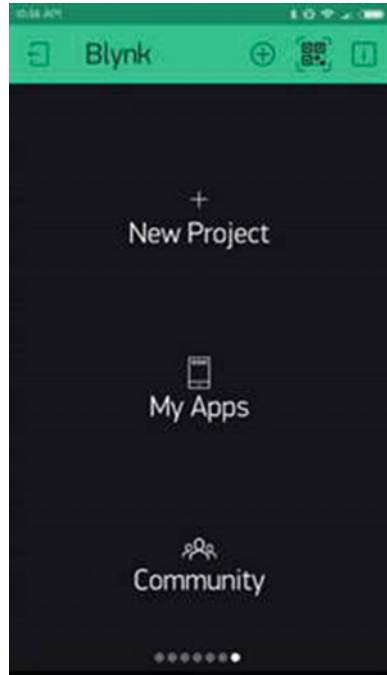


Fig. 6 Creating a new project



Fig. 7 Selection of hardware

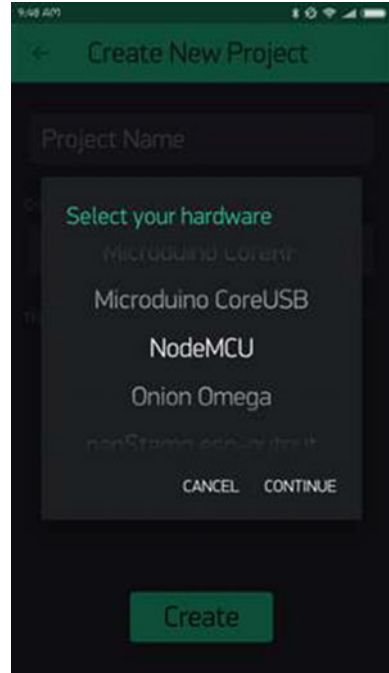
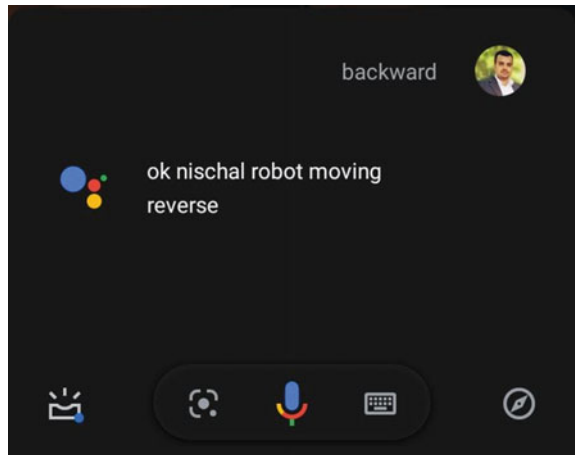


Fig. 8 Robot reverse motion



movement, left turn, right turn, and stop whenever desired. The robot is controlled using a Blynk app, and usage of solar energy as a source of energy makes the robot economically cheap. Also, the grass can be cut from anywhere in the world, and the device can be controlled from distant places using IoT. This grass cutter is a portable, compact and lightweight grass cutter compared to conventional grass

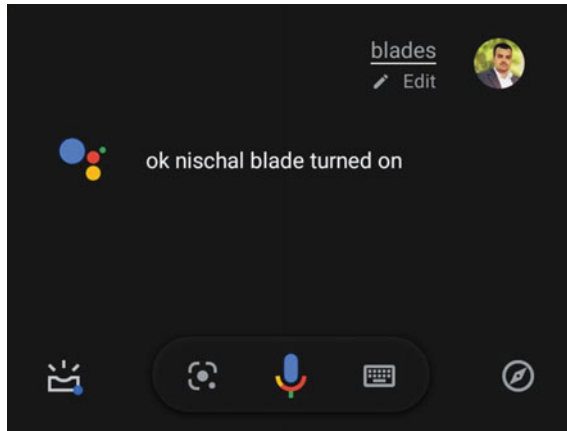


Fig. 9 Blades movement

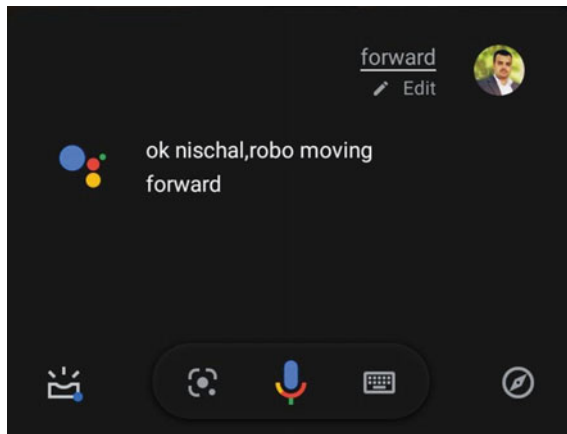


Fig. 10 Robot forward motion

cutters. The Google voice assistant is also interfaced to the grass cutter through Blynk app and IFTTT software. The grass cutter is also controlled using English commands through Google assistant. The obstacle detection feature can be integrated with the proposed system for reliable operations, and the performance of the system can also be evaluated for different shaped grass cutting blades. The offline speech recognition models can be integrated with the robot to develop a grass cutter that can work even in remote locations.

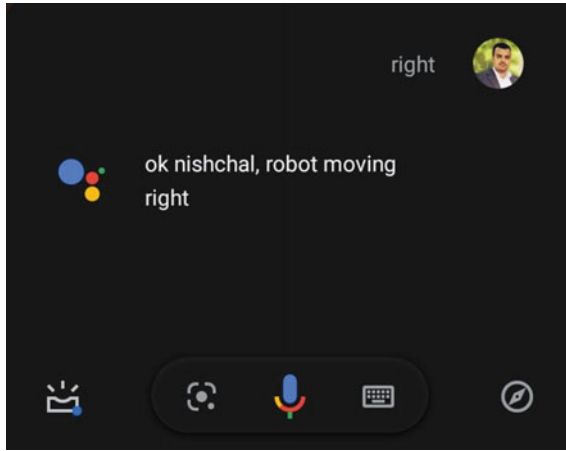


Fig. 11 Robot right movement

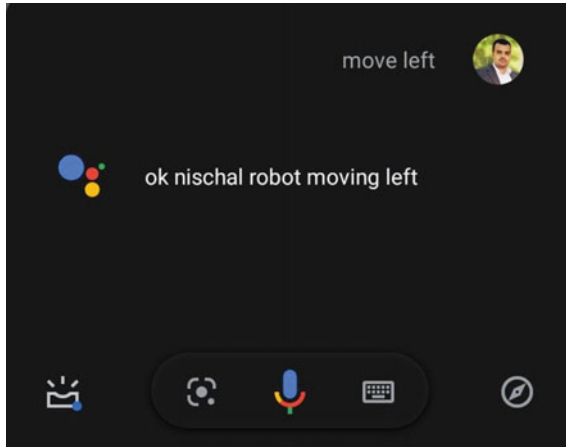


Fig. 12 Robot left movement

Fig. 13 IFTTT Dashboard
page 1

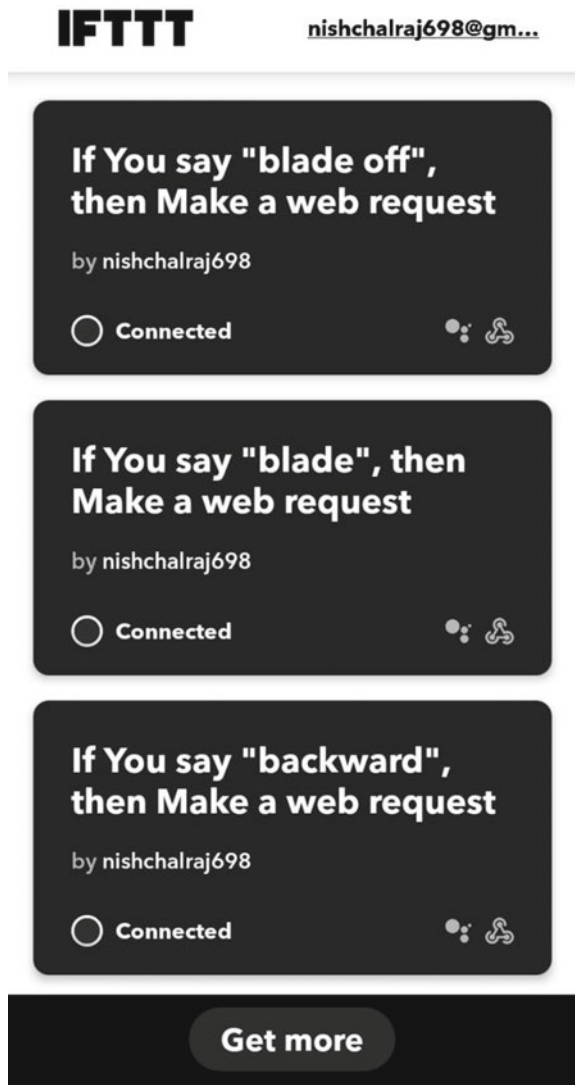


Fig. 14 IFTTT Dashboard
page 2

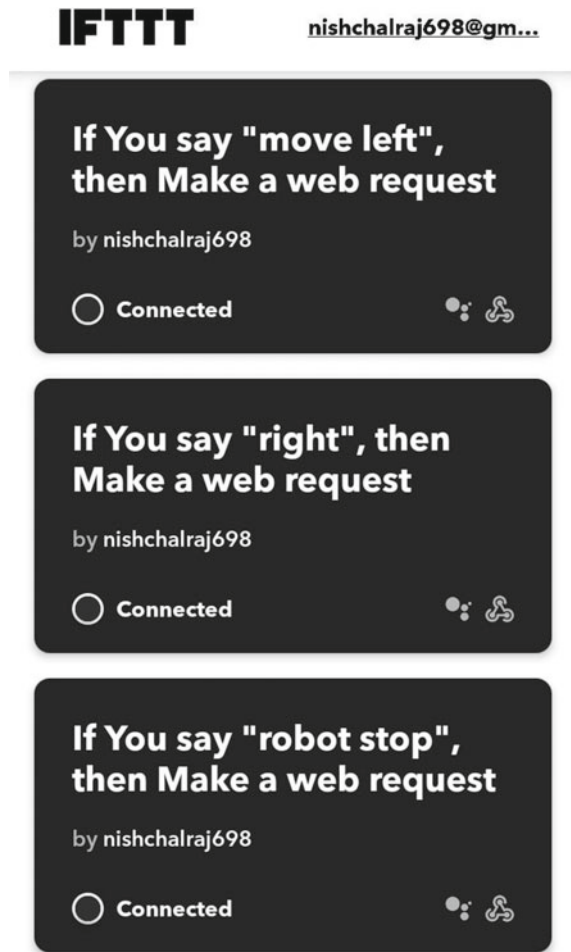


Table 1 Specifications of the IOT based grass cutter

S. No.	Components/software used	Specifications
1	NodeMCU	Tensilica 16-bit RISC CPU, 80 MHz, 12 V
2	Motor driver	L293D
3	Solar panel	12 V, 310 A
4	Battery	12 V, 1.2 mAH
5	Motor	4 DC, 130 rpm
6	Software platform used	Blynk app
6	ASR system used	Google speech recognition [15–18, 20–22]
7	Integration software for ASR and Blynk app	IFTTT software [19]

Acknowledgements Authors would like to extend their sincere thanks to HOD, ECE Department, Vidyavardhaka College of Engineering, Mysuru, for supporting the work.

References

1. B.P. Dilip, N.B. Pagar, V.S. Ugale, S. Wani, M. Sharmila, Design and implementation of automatic solar grass cutter. *Int. J. Adv. Res. Electr. (IJARE)* **6** (2017)
2. B.R. Patil, S.S. Patil, Solar based grass cutting. *Int. J. Electr. Electron. Eng. (IJEET)* (2017)
3. S. Jain, A. Khalore, S. Patil, Self-efficient and sustainable solar powered robotic lawn mower. *Int. J. Trend Res. Dev. (IJTRD)* **2**(6) (2015)
4. M. D'Souza, V.B. Naik, R.V. Bicholkar, Automatic solar grass cutter. *IJSTE Int. J. Sci.*
5. B.P. Dilip, N.B. Pagar, V.S. Ugale, S. Wani, M. Sharmila, Design and implementation of automatic solar grass cutter. *Int. J. Adv. Res. Electr. Electron. Instrument. Eng.* **6**(4) (2017)
6. *Int. J. Sci. Eng. Res.* **10**(4), 1089 (2019). ISSN 2229-5518
7. M. Alam, V.V. Singh, V.Y. Chandan, IOT based grass cutter with solar panel. *IJSER* (2019). <https://www.ijser.org>
8. G.D. Dattatraya, M.V. Mhatarde, L.M. Shrihari, S.G. Joshi, Robotic agriculture machine. *Int. J. Innov. Res. Sci. Eng. Technol.* **3**(4) (2014)
9. S. Yaghoubi, N.A. Akbarzadeh, S.S. Bazargani, S.S. Bazargani, M. Bamizan, A.S. Maryan Irani, Autonomous robots for agricultural tasks and farm assignment and future trends in agro robots. *IJMMEIJENS* **13**(3) (2013)
10. *Int. J. Res. Appl. Sci. Eng. Technol. (IJRASET)*
11. M. Mudda, V. Teja, S. Kumar, P. Kumar, Automatic solar grass cutter **6**(IV) (2018). IC Value: 45.98; SJ Impact Factor: 6.887. ISSN: 2321-9653. Available at www.ijraset.com
12. A.P. Pralhad, B.S. Bhagwan, K.S. Subhash, P.N. Vijay, R.U. Kale, Solar powered automatic grass cutter and pesticide spreading robot. *Int. Res. J. Eng. Technol. (IRJET)* **4**(5) (2017). e-ISSN: 2395-0056, p-ISSN: 2395-0072. www.irjet.net
13. A. Rithvik Reddy, N.V. Chaitanya, P. Abhishek, A. Suvarnamma, ECE Department SRMIST, Autonomous solar based lawn mower. *Int. J. Pure Appl. Math.* **119**(12), 13129–13134 (2018). ISSN: 1314-3395 (on-line version). <https://www.ijpam.eu>
14. Mahadevaswamy et al., Providing natural language interface to database using artificial intelligence. *Int. J. Sci. Technol. Res.* **8**(10), 1074–1078
15. Mahadevaswamy, D.J. Ravi, Performance of isolated and continuous digit recognition system using Kaldi Toolkit. *Int. J. Recent Technol. Eng.* **8**(2S2), 264–271 (2019)
16. Mahadevaswamy, D.J. Ravi, Robust perceptual wavelet packet features for continuous kannada speech recognition using Kaldi toolkit. *Solid State Technol.* **63**(3) (2020)
17. R. C. Naik, R. Bhat, N. M. Nagaraj, K. S. Praveena, U. G. Student. Voice based self assisting robot using firebird V (2019)
18. Speech to text conversion. <https://patents.google.com/patent/US6173259B1/en>, <https://cloud.google.com/speech-to-text/pricing>
19. IFTTT Software. <https://ifttt.com/home>
20. Mahadevaswamy, D. J. Ravi, Performance analysis of adaptive wavelet denosing by speech discrimination and thresholding, 2016 International Conference on Electrical, Electronics, Communication, Computer and Optimization Techniques (ICEECCOT), Mysuru, 2016, pp. 173–178. <https://doi.org/10.1109/ICEECCOT.2016.7955209>
21. Mahadevaswamy, D. J. Ravi, Performance analysis of speech enhancement using adaptive wavelet thresholding, 2016 IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT), Bangalore, India, 2016, pp. 392–396. <https://doi.org/10.1109/RTEICT.2016.7807848>

22. Mahadevaswamy, D. J. Ravi, Performance analysis of LP residual and correlation coefficients based speech separation front end, 2017 International Conference on Current Trends in Computer, Electrical, Electronics and Communication (CTCEEC), Mysore, 2017, pp. 328–332. <https://doi.org/10.1109/CTCEEC.2017.8455039>

Human Body Measurement Extraction from 2D Images



Sachin S. Bhat, Alaka Ananth, Preema Dsouza, K. Sharanyalaxmi, Shreeraksha, and Tejasvini

1 Introduction

Fashion industry is evolving and everyone wants to be fashionable. More than self-expression, fashion boosts confidence and makes to feel fresh among our peers. Not only shopping online but also customers are concerned about how the particular item looks on them before purchasing. But the fit and form of outfit depend on the accuracy of measurement and visualization. Hence, the need for non-contact anthropometric measurement extraction is of very much an interesting area. Extracting the measurements using 3D scanners can be said to be of much accurate one than compared to the 2D-image-based measurements. However, it is not that easy for all small-scale garment industries to acquire a 3D measurement tool besides the reason that is costly. Hence, proposing a system that performs measurement extraction task using only 2D images with improved accuracy is much of a need.

This paper proposes a non-contact body measurement extraction with the help of computer vision. It provides easiest way for implementing the system for non-contact human body measurement and can be easily implemented in the garment industry whether it is small- or large-scale one.

S. S. Bhat (✉) · P. Dsouza · K. Sharanyalaxmi · Shreeraksha · Tejasvini
Department of Electronics and Communication Engineering, Shri Madhwa Vadiraja Institute of Technology and Management, 574115 Bantakal, India

A. Ananth
Department of Information Science and Engineering, NMAM Institute of Technology, 574110 Nitte, India

2 Literature Review

The shape variation of different objects can be measured using principal component analysis (PCA). The feature is extracted using part-based sampling technique. This provides accurate results and is less expensive but allows measuring only five parameters [1]. Body features are extracted from front and side images using background subtraction and contour detection. Canny edge detector is used for contour detection. Selected points from silhouette curve are considered as feature points. MATLAB software is used for evaluation. Even though it seems to be robust and reliable method, the images have to be captured over a white background only [2]. In order to determine threshold automatically, moving object segmentation technique is accommodated which uses Euler computation method since it requires less computational time. This method incorrectly identifies foot and head position and skin color feature is not used. It is not applicable to the dark and shadow region [3]. Body volume can be calculated using curve fitting method. Body volume is represented in pixels by adding up the areas of the slices. This provides precise measurement and the body volume calculated by the proposed method is accurate [4]. Using CNN learning, pose can be estimated and it can capture all the features of input image. However the algorithm seems to be complex [5, 6].

Human body model can be synthesized with regression analysis method. Paper includes human body contour extraction as the initial step and curve fitting method is used to calculate the circumference. Since the regression analysis accuracy is unstable, this method provides inaccurate results [7]. With the help of Muller's method, human body measurements are obtained. The binary image is used to calculate the body measurement with image processing software developed in MATLAB. It processes the input image with low cost and high speed [8]. One of the outstanding methods of measuring the human body parameter was using Alex Net. This initially created 3D model of 2D image (silhouettes were used in order to preserve the identity) and then produced the measurement [9].

3 Methodology

Proposed methodology consists of three parts: Acquiring the image along with its preprocessing, segmenting the foreground part of appearance and extraction of body measurements as shown in Fig. 1.

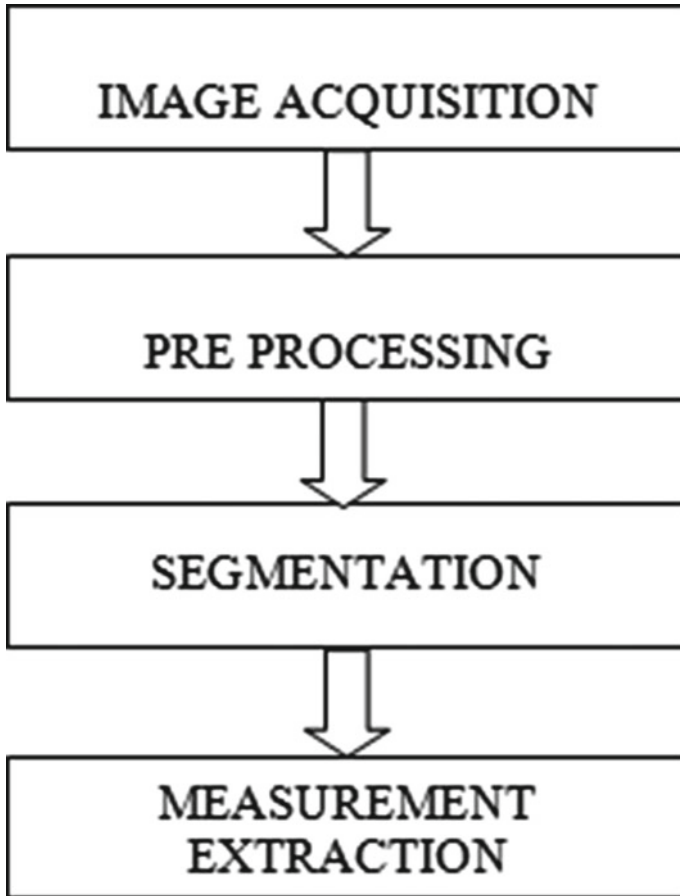


Fig. 1 Flowchart of methodology

3.1 *Image Acquisition and Preprocessing*

Image Acquisition

The measurement extraction system with simple algorithm takes three different input images. Namely an image with checker board, arm spread image and a side image. Where the checker board image is necessary for camera calibration and affine correction processes and arm spread image used to measure the sleeve length and side image is used to measure the waist. Along with these three images, it is also necessary to input the one edge length of the checker board used in the image acquisition process which is very much necessary for camera calibration, applying transformation and also for obtaining pixel to meter converted measurements.

3.2 *Preprocessing*

As for any image processing-based algorithm, preprocessing is one of the necessary steps. The acquired image might not be always up to the required quality of the algorithm to which it is fed into. The preprocessing used in our work involves camera calibration and affine mode correction.

a. *Camera Calibration*

Whenever an image processing algorithm deals with extraction of some information from the image, it is necessary that it obtains the three-dimensional co-ordinates for the particular image. This can be done by camera calibration. However, the calibration process requires any one of the original image parameters to be known so that it can retrieve other image-based parameters that are of concern.

This work uses camera calibration technique that involves checker board. The calibration technique used here is mainly concerned in retrieving the extrinsic parameters of the camera. Extrinsic parameter of a camera is related to real world co-ordinates (in 3D) and camera co-ordinates (in 3D) corresponding to any one point in the image. We use specific methods defined in the OpenCV library for camera calibration process. The acquired image is then fed into the algorithm as arguments.

b. *Affine Mode Corrections*

Affine mode correction is of much necessity to preserve the information about points, straight lines and planes. When a 2D image is captured from a real-world scenario, it often loses its three-dimensional properties. In particular, we can say that if a real world scenario that contains some parallel lines is being captured, then the obtained 2D image will not have those parallel lines at the same place where they are in real-world scenario. This is due to the perception that a 2D capturing device makes some delusions regarding some 3D co-ordinates. This needs to be corrected before one head to any sort of further operation with the image.

This process of correcting the perception occurred in capturing the 3D scenario by a 2D device is done by using perceptive transformation.

c. *Perspective Transformation*

As the name suggests in perspective transformation, perceptive of the image can be changed. Perspective transformation reserves parallelism of an image. Here, the checker board image that is fed as input is used to apply the perceptive transformation to the rest of the image.

The perceptive transform requires four different corner points in the 2D image which is also called as quadrangle vertices. The methods provided in OpenCV for finding the checkerboard edges are used in the algorithm for obtaining these points to give them as inputs for perspective transform function.

3.3 Segmentation

Among image segmentation algorithms, there are two major groups: (a) methods assuming known appearance models and (b) methods estimating appearance models jointly with segmentation. Typically, the first group optimizes appearance log-likelihoods in combination with some spatial regularization. This problem is relatively simple and many methods guarantee globally optimal results. The second group treats model parameters as additional variables transforming simple segmentation energies into high-order NP-hard functionals [10, 11].

The grabcut segmentation method is based on graph cuts. Starting with a user-specified bounding box around the object to be segmented, the algorithm estimates the color distribution of the target object and that of the background using a Gaussian mixture model. This segments the input image iteratively. It uses color data modeling by iterative energy minimization. Iterations are extremely slow. Hence, this segmentation method is discarded.

The Chroma keying method is used to segment the human body from background. In this technique a specific color is removed from the image and is replaced with another color element in specific by black. It is also referred as green screen segmentation. The term green refers to the color of the background used. In Chroma keying method, any background can be used as background but the bright green is usually preferred as it is the color that is farthest away from the human skin tones.

In this method, following transformation is applied on the each pixel of the input image:

$$F(r, g, b) \rightarrow \alpha \tag{1}$$

If $\alpha = 0$, means the pixel is fully in the green screen. If $\alpha > 0$, i.e., 1 means the pixel is in the foreground object.

3.4 Measurement Extraction

Further step is to obtain measurement for waist, shoulder length and sleeve length. Measurement is done by manually obtaining the points on three different images that are fed onto the algorithm each time by considering two images for a particular measurement in order to increase the accuracy. The user is supposed to mark two points in each image that appears. The pixel corresponding to the point that is marked in the image is used for measurement.

Distance between any two points x and y on the image is given by the formula,

$$\text{dist} = x - y \tag{2}$$

The distance obtained after each step is then converted to corresponding meter units.

a. *Waist measurement*

After making the waist in first 2 images distances images $d1$ and $d2$ between waist ends in both are obtained using the Eq. (2). Waist measurement is done by approximating the measurement obtained for 2 images as in ellipse.

The measurement that is obtained between two points marked of two different images is treated as semi-major axes and are used to find the total perimeter of the waist similar the way of finding the perimeter of the ellipse when two semi-major axes lengths are known. The formula is as shown below:

$$\text{Waist_perimeter} = 2\pi \sqrt{\frac{(\text{dist1})^2 + (\text{dist2})^2}{2}} \quad (3)$$

where $d1$ and $d2$ are semi-major axes lengths and $\text{dist1} = d1/2$, $\text{dist2} = d2/2$.

b. *Shoulder and sleeve length measurement*

The shoulder measurement can either be done by using image marked manually or by heuristic selection. The heuristic point selection is done by first finding the head tip in the silhouette image. Traversing horizontally through the edge and finding the point where there is a drastic change in the slope that referring the shoulder edge. In the same way, wrist edge is also found whereas option for manual marking is also given. We prefer marking manually as it is the method in which it is possible to get the exact edges compared to heuristic marking which results in larger error due to incorrect marking. In manual marking, distance between two pixels is obtained in the same way as mentioned in the case of waist measurement. In first image, fall near the both sides of the neck is marked. However, in the next image, shoulder edge is marked. The same is repeated for next two images.

In case of shoulder measurement, total shoulder measurement is done in terms of meter pixel using the formula:

$$\text{Shoulder} = \text{dist1} + \text{dist2} + \text{dist3} \quad (4)$$

dist1 distance from the left shoulder to left fall
 dist2 distance from the right shoulder to right fall
 dist3 distance from left shoulder to right shoulder

Two different shoulder measurements in terms of meter pixel are obtained using four different images that are marked and the final shoulder length is obtained by taking the mean between the two different shoulder measurements obtained as below:

$$\text{Shoulder_measurement} = \text{shoulder1} + \text{shoulder2} \quad (5)$$

Shoulder1 and shoulder2—two different shoulder measurement extracted using the Eq. (4).

To obtain sleeve length measurement, distance between left shoulder and left wrist and right shoulder and right wrist is obtained from two images which are then averaged to get the final result.

$$\text{Sleeve_measurement} = (\text{left_sleeve} + \text{right_sleeve})/2 \tag{6}$$

All the distance measurements done are in meter pixel unit which are then converted to the proper distance unit, i.e., to centimeters using formula:

$$\text{PixelToDistance} = \sqrt{(\text{point}[0] \times \text{mx})^2 + (\text{point}[1] \times \text{my})^2} \tag{7}$$

3.5 Results and Comparison

Initially, three different images of the person whose measurements are to be extracted are given as input. The different images are arm spread image, waist/side image and image with checker board.

Figure 2 shows the effect of applying perspective transformation onto the input checker board. As shown, the result of the transformation is an image with the edges having some positions as in the real-world scenario.

Figure 3 shows the result of segmentation process. Chroma keying segmentation method converts RGB image into a binary image such that the required foreground in represented in binary 1 or white and the green background converted to binary 0 or black.

After applying the algorithm to the input images, extracted measurements are given in Table 1. The table also illustrates the comparison between actual and observed measurements. Using these comparisons, error in different measurements is also calculated.

- OWL Observed waist length (in cm)
- AWL Actual waist length (in cm)
- OSHL Observed shoulder length (in cm)

Fig. 2 Affine more correction: **a** input image, **b** output image

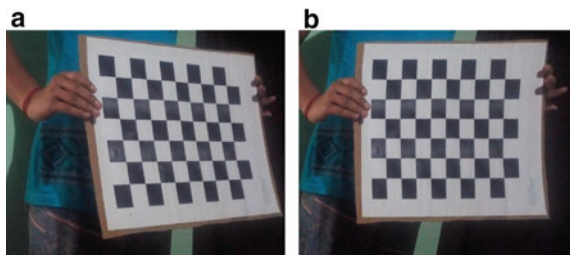


Fig. 3 Segmentation results:
a image with checker board,
b arm spread image

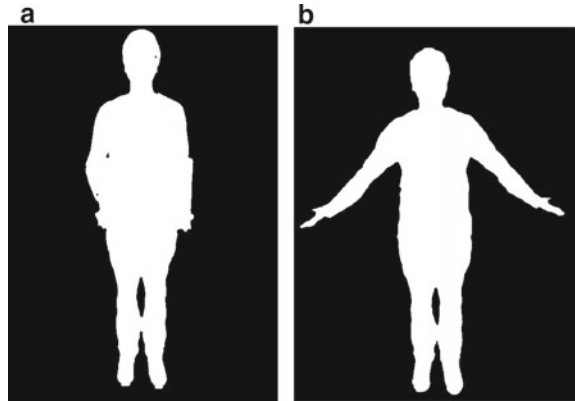


Table 1 Extracted measurements and actual measurements and the mean difference between them

	OWL	AWL	<i>E</i>	OSHL	ASHL	<i>E</i>	OSLL	ASLL	<i>E</i>
Img1	107	92	15	41	38	3	48	50	2
Img2	90	82	8	41	35	6	43	45	2
Mean error			11.5			4.5			2

- ASHL Actual shoulder length (in cm)
- OSL Observed sleeve length (in cm)
- ASL Actual sleeve length (in cm)
- E* Error (in cm)

As shown in the table, the measurement of waist involves more error compared to the other measurements whereas the mean error for the waist measurement is around 11.5 cm, which is calculated by taking the average of the two different waist measurement errors. The reason for maximum error in case of waist measurement is due to improper marking of the points.

The algorithm gives at the max 91% of accuracy for waist measurements, 92% of accuracy for the shoulder measurement and 96% of accuracy for sleeve measurement. In case of shoulder measurement, kinect sensor showed an accuracy around 88% which is less when compared to the implemented algorithm.

4 Conclusion

This work is concerned in extracting the human body measurements with non-contact method. It uses computer vision technique in obtaining the measurements. This method requires background of the input image to be green only. Further segmentation can be generalized so that measurements can be taken with the help of any

kind of background. Also, while selecting the alternate to the green segmentation method, it is necessary to consider that it needs the required level of accuracy in segmentation.

In the proposed algorithm, key point estimation is done both heuristically and manually. Heuristic key point extraction is of less favorable due to the reason that its incorrect recognition of point might lead to large amount of error in the result. Whereas, on the other hand, manual marking of the key point is also considered of less accuracy. Hence, robust key point estimation method can be implemented rather than the one which is used.

As for this algorithm, camera calibration is done using checker board which conditions the image acquisition in particular manner. More implicit method can be implemented for metric calibration so as to remove the dependency from checker board. As the algorithm only extracts the measurements from the acquired image, it can be further generalized by developing a user friendly website that extracts the image and automatically sends the collected data to the garment-maker who fits a garment to the user using that data.

References

1. W.Y. Chang, Y.C.F. Wang, Seeing through the appearance: body shape estimation using multi-view clothing images, in *2015 IEEE International Conference on Multimedia and Expo (ICME)* (IEEE 2015, June), pp. 1–6
2. L. Jiang, J. Yao, B. Li, F. Fang, Q. Zhang, M.Q.H. Meng, Automatic body feature extraction from front and side images. *J. Softw. Eng. Appl.* **5**(12), 94 (2013)
3. C.F. Juang, C.M. Chang, J.R. Wu, D. Lee, Computer vision-based human body segmentation and posture estimation. *IEEE Trans. Syst., Man, Cybern. Part A Syst. Humans* **39**(1), 119–133 (2008)
4. L. Pradhan, G.Song, C. Zhang, B. Gower, S.B. Heymsfield, D.B. Allison, O. Affuso, Feature extraction from 2D images for body composition analysis, in *2015 IEEE International Symposium on Multimedia (ISM)* (IEEE 2015, Dec), pp. 45–52
5. B. Ai, Y. Zhou, Y. Yu, S. Du, Human pose estimation using deep structure guided learning, in *2017 IEEE Winter Conference on Applications of Computer Vision (WACV)* (IEEE 2017, Mar), pp. 1224–1231
6. Y. Liu, Y. Xu, S.B. Li, 2-D human pose estimation from images based on deep learning: a review, in *2018 2nd IEEE Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC)* (IEEE 2018, May), pp. 462–465
7. S. Luo, Application research on 2D image-based non-contact anthropometric technology in clothing E-commerce, in *2017 7th International Conference on Applied Science, Engineering and Technology (ICASET 2017)* (Atlantis Press, 2017, May)
8. A.D. Roknabadi, M. Latifi, S. Saharkhiz, H. Aboltakhty, Human body measurement system in clothing using image processing. *World Appl. Sci. J.* **19**(1), 112–119 (2012)
9. E. Dibra, H. Jain, C. Öztireli, R. Ziegler, M. Gross, Hs-nets: Estimating human body shape from silhouettes with convolutional neural networks, in *2016 Fourth International Conference on 3D Vision (3DV)* (IEEE 2016, Oct), pp. 108–117

10. X. Liang, K. Gong, X. Shen, L. Lin, Look into person: Joint body parsing & pose estimation network and a new benchmark. *IEEE Trans. Pattern Anal. Mach. Intell.* **41**(4), 871–885 (2018)
11. B. Wang, H. Zheng, X. Liang, Y. Chen, L. Lin, M. Yang, Toward characteristic-preserving image-based virtual try-on network, in *Proceedings of the European Conference on Computer Vision (ECCV)*, pp. 589–604 (2018)

Generation of ECG Arrhythmias Using Fourier Analysis



Paresh Praveen, B. Samartha, Vaibhav R. Pai, M. Namith Rao, K. S. Shivaprakasha, and Rekha Bhandarkar

1 Introduction

Electrocardiography (ECG) is the method of analyzing the electrical signals produced from the cardiac muscles of the heart, helpful in determining the heart problems and its conditions. ECG analysis is done in situations where the person has chest pain, shortness of breath, and fatigue. ECG is a boon to mankind in the form of early predictor of deadly heart ailments like coronary artery disease (CAD) and arrhythmia.

In an ECG test, the patient's arms, legs, and chest are linked with sensors called electrodes. The electric potentials generated by heart muscles are monitored to check the heart rate and its state. The recording machines show the graphical representation of the heart rhythm which is printed on paper for better diagnostic analysis [1–3].

A normal ECG wave is as shown in Fig. 1. It composes of a P wave, a QRS complex, and a T wave. It is highly recommended to have the understanding about these waves as prerequisite for ECG analysis. P wave is the small upward wave, which

P. Praveen (✉) · B. Samartha · V. R. Pai · M. Namith Rao · K. S. Shivaprakasha · R. Bhandarkar
Department of Electronics and Communication Engineering, NMAM Institute of Technology,
Nitte (Visvesvaraya Technological University, Belagavi), Udupi, Karnataka, India
e-mail: pareshpraveen05@gmail.com

B. Samartha
e-mail: samartha.bannur@gmail.com

V. R. Pai
e-mail: vaibhavpai99@gmail.com

M. Namith Rao
e-mail: nmith5610@gmail.com

K. S. Shivaprakasha
e-mail: shivaprakasha.ks@nitte.edu.in

R. Bhandarkar
e-mail: rekhabhandarkar@nitte.edu.in

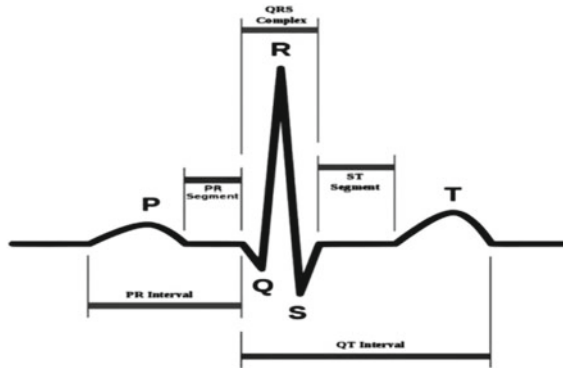


Fig. 1 A normal ECG signal

occurs before atrial systole (contraction of atrial heart chambers). As a result of this, the blood is pushed into the left and right ventricles which lasts for 0.1 s. QRS complex with average duration of 0.3 s is produced before ventricular systole (contraction of ventricles), resulting in blood flow to the arteries. Ventricular repolarization results in generation of T wave. It occurs before the end of ventricular systole which marks the end of ventricular contraction [4, 5]. The duration and length of the wave intervals is a dominant factor in ECG studies.

Various intervals and segments in an ECG wave can be detailed as follows [6].

1. PR interval—It is the interval which starts from P wave and ends at the beginning of the QRS complex.
2. QT interval—The interval starting from QRS complex till the end of T wave.
3. QRS complex—This depicts entire QRS complex.
4. ST segment—The tail of the QRS complex to the head of the T wave marks the ST segment.
5. PR segment—The end of the P wave to the onset of the T wave depicts PR segment.

Arrhythmias are variations in the normal sinus rhythms caused due to improper functioning of the heart which creates fluctuations in the electrical signals. The electrical signals are generated due to depolarization and repolarization of the cardiac muscles. As a result of this, arrhythmias are generated. The arrhythmias are classified on the basis of interval range between two consecutive waves, amplitude variation, absence of P wave, and many other properties [7].

Understanding the features of different arrhythmias is very important for students in medical courses. This paper presents a partial work of the consultancy project for KSHEMA, Mangaluru. The project intends to develop a device that generates normal ECG and various arrhythmias so as to enable students of medical colleges to understand the key features of the same. In this paper, we present the process of generating ECG arrhythmias using Fourier series representation of signals.

The remaining part of the paper is arranged as follows. Section 2 gives a perception about similar works carried out in the near past, and Sect. 3 gives an overview of arrhythmias. Section 4 details the process of generating ECG arrhythmias using Fourier analysis approach. The results and the analysis are discussed under Sect. 5 followed by the conclusions in Sect. 6.

2 Previous Works

There are many methods that have been developed so far for the generation of ECG signal and its arrhythmia. In [8], authors present the generation of an ECG signal using Simulink tool. It modelled PR interval wave, QRS interval wave, ST interval wave, and ECG wave for one cycle using MATLAB Simulink [8]. The ECG signal can also be generated using dynamical functions in MATLAB [9]. A model for generating simple synthetic signals with each particular wave modelled using trigonometric functions or a Gaussian mono pulse was presented in [10].

In [11], the authors have proposed the generation of ECG signals from reaction diffusion model. Analysis of ECG signal using MATLAB and LabVIEW tools was presented in [12]. The work includes the generation and simulation of ECG signal and the detection of abnormalities in ECG.

ECG signals can also be generated using Fourier series. Individual waves of an ECG signal were approximated by mathematical model and were implemented on MATLAB [13]. A single cycle of an ECG signal has triangular and sinusoidal waveforms. In [14], authors have showed that each significant feature of ECG signal can be represented by shifted and scaled versions.

3 ECG Arrhythmias

An ECG signal consists of consecutive heartbeats. Each beat can be observed as series of deflection away from the baseline. It can be represented by five waves. The P wave, T wave, and QRS complex give glimpse about the atrial and ventricular depolarization and repolarization. In this section, the important features used for the classification of arrhythmias are presented.

A. *P wave, PR Segment, and PR interval*

The computation of P wave begins the ECG elucidation. In addition to atrial contraction, the P wave highlights the depolarization of left and right atrium. The starting and end points of P wave and QRS complex decide the segment and interval of PR region. PR segment commences from tail of the P wave and terminates at the start of the QRS complexes [15].

Table 1 Amplitude and timing information

Wave	Amplitude
P wave	0.25 mV
R wave	1.6 mV
Q wave	25% of R wave
T wave	0.35 mV
Interval	Duration
PR wave	0.12–0.2 s
QT wave	0.35–0.44 s
ST wave	0.05–0.15 s
P wave	0.11 s

B. QRS complex

The second wave is the QRS complex. These complex represents the depolarization of the ventricles. It incorporates Q, R, and S waves. First positive and negative deflections in the QRS complex are called R wave and Q wave, respectively. S wave being the next negative deflection comes just after the R wave. Some QRS complexes do not have all three deflections. QRS duration represents the time interval of the entire complex which lasts between 0.06 and 0.1 s. ST segment follows the QRS complex, starting from the end of QRS till the onset of T wave.

C. T wave

Ventricular repolarization reflects the current of T wave in rapid phase. T wave is asymmetric, with steeper downward slope. The polarity of this wave normally follows that of the main QRS deflection. The change to the T wave which follows the ST segment is expected to be even in a normal ECG signal [16].

The amplitude and timing information of ECG signal is given in Table 1 [17–19].

4 Generation of ECG Arrhythmias

In this section, we present the mathematical model used for the generation of the ECG signals. Consider a, b, and L as amplitude, duration, and interval of ECG wave, respectively, we have used Fourier analysis for the development of ECG signals.

The Fourier series equations used to produce the required five waves of an ECG arrhythmia are presented as follows:

P and T waves can be generated using the following Eq. (1):

$$af(x) = a \left(\frac{2}{b\pi} + \sum_{n=1}^{\infty} \frac{4b}{(b - 2n)(b + 2n)} \cos \frac{n\pi}{b} \cos \left(\frac{n\pi x}{L} \right) \right) \tag{1}$$

Equation (2) represents the generations of Q and S waves (negative gradient).

$$f(x) = -\left(\frac{a}{2b} + \sum_{n=1}^{\infty} \frac{2ab}{n^2\pi^2} \left[1 - \cos\left(\frac{n\pi}{b}\right) \cos\left(\frac{n\pi x}{L}\right)\right]\right) \quad (2)$$

R wave can be generated using Eq. (3)

$$f(x) = +\left(\frac{a}{2b} + \sum_{n=1}^{\infty} \frac{2ab}{n^2\pi^2} \left[1 - \cos\left(\frac{n\pi}{b}\right) \cos\left(\frac{n\pi x}{L}\right)\right]\right) \quad (3)$$

Equation (4) is used for the generation of flutter waves (f waves) [20].

$$f(x) = \left(\frac{a}{2b} + \sum_{n=1}^{\infty} \frac{2ab}{n^2\pi^2} \cdot \left(\frac{1}{\pi^2}\right) \sin\left(\frac{n\pi x}{L}\right)\right) \quad (4)$$

Arrhythmia can be uniquely classified based on the key features like rate, rhythm, P wave, lengths of PR wave, and QRS complex [21]. Table 2 presents the signatures of different arrhythmias that are used for generation of the same.

5 Simulation Results

Using the mathematical model developed in Sect. 4, we have generated normal ECG and 17 arrhythmias. MATLAB is used for simulation.

Normal Sinus Rhythm

P waves come before QRS complex which has a range between 60 and 100 bpm with regular rhythm. The interval length of PR and QRS complexes are normal. We have considered a rate of 72 bpm. The waveform is shown in Fig. 2.

Sinus Bradycardia

The rate is less than 60 bpm. QRS complex comes after P wave, and the rhythm is regular. The interval length of PR and QRS complexes are normal. We have considered a rate of 35 bpm. The waveform is shown in Fig. 3.

Sinus Tachycardia

The rate is in between 100 and 160 bpm. QRS complex is after P wave in the wave and is regular in nature. The interval length of PR and QRS complexes are normal. We have considered a rate of 105 bpm. The waveform is shown in Fig. 4.

Sinus Arrhythmia

The rate is in between 60 and 100 bpm. The QRS complex comes after P wave with irregular rhythm. The interval length of PR and QRS complexes are normal. We have considered a rate of 72 bpm. The waveform is shown in Fig. 5.

Sinus pause

Table 2 Features of ECG arrhythmias

S. No.	Name	Rate	Rhythm	P wave	Length of PR interval (in s)	Length of QRS complexes (in s)
1	Normal sinus rhythm	60–100 bpm	Atrium: Regular Ventricular: Regular	P waves are present before QRS complex	0.12–0.2	0.06–0.12
2	Sinus bradycardia	<60 bpm	Atrium: Regular Ventricular: Regular	P waves are present before QRS complex	0.12–0.2	0.06–0.12
3	Sinus tachycardia	101–160 bpm	Atrium: Regular Ventricular: Regular	P waves are present before QRS complex	0.12–0.2	0.06–0.12
4	Sinus arrhythmia	60–100 bpm	Irregular	P waves are present before QRS complex	0.12–0.2	0.06–0.12
5	Sinus pause/arrest	Variable, depends on the frequency	Irregular	P waves are present if QRS complex is present	0.12–0.2	0.06–0.12
6	Supraventricular tachycardia	Atrial: 150–250 bpm Ventricular: 150–250 bpm	Regular	Usually not discernable, especially at the high rate range	Usually not discernable	0.06–0.12
7	Premature atrial complexes (PAC's)	Usually regular but depends on the underlying rhythm	Irregular as a result of PAC	Usually upright but premature and abnormal shape	0.12–0.2	0.06–0.12
8	Idioventricular rhythm	20–40 bpm	Regular	Absent	Not measurable	Wide and unusual (>0.12)
9	Accelerated idioventricular rhythm	41–100 bpm	Regular	Absent	Not measurable	Wide and unusual (>0.12)
10	Agonal rhythm	<20 bpm	Regular	Absent	Not measurable	Wide and unusual (>0.12)
11	Premature ventricular contractions (PVC)	Usually normal	Irregular during PVC's	Absent during PVC's	Not measurable during PVC's	wide and unusual (>0.12)

(continued)

Table 2 (continued)

S. No.	Name	Rate	Rhythm	P wave	Length of PR interval (in s)	Length of QRS complexes (in s)
12	Junctional escape rhythm	40–60 bpm	Regular	Usually inverted or absent, may be before, during or after QRS complex	Will be shortened, if occurs before QRS complex, otherwise not measurable	0.06–0.12
13	Accelerated junctional rhythm	60–100 bpm	Regular	Usually inverted or absent, may be before, during or after QRS complex	Will be shortened, if occurs before QRS complex, otherwise not measurable	0.06–0.12
14	Junctional tachycardia	101–180 bpm	Regular	Usually inverted or absent, may be before, during or after QRS complex	Will be shortened, if occurs before QRS complex, otherwise not measurable	0.06–0.12
15	Premature junctional complex	Depends on underlying rhythm	Irregular due to early beat	Usually inverted or absent, may be before, during or after QRS complex	Will be shortened, if occurs before QRS complex, otherwise not measurable	0.06–0.12
16	Atrial flutter	Atrial: 250–400 bpm, Ventricular: variable	Atrial: Regular, Ventricular: may be irregular	Normal P waves are absent; Flutter waves (f waves or saw tooth pattern)	Not measurable	0.06–0.12
17	Atrial fibrillation	Atrial: 350–400 bpm, Ventricular: variable	Irregularly irregular	Normal P waves are absent; replaced by F waves	Not discernible	0.06–0.12
18	Ventricular tachycardia	101–250 bpm	Atrial rhythm not distinguishable, Ventricular rhythm usually regular	No	Not measurable	Wide and unusual (>0.12)

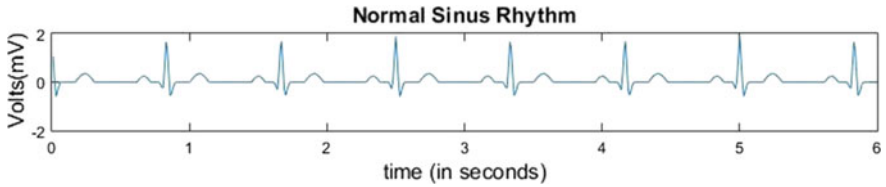


Fig. 2 Normal sinus rhythm

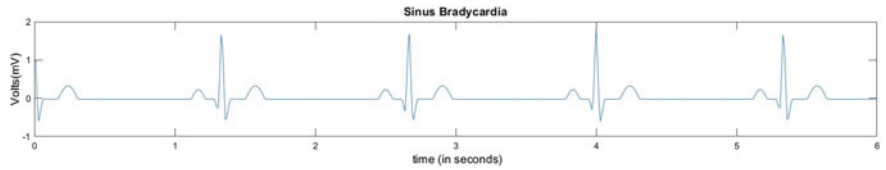


Fig. 3 Sinus bradycardia

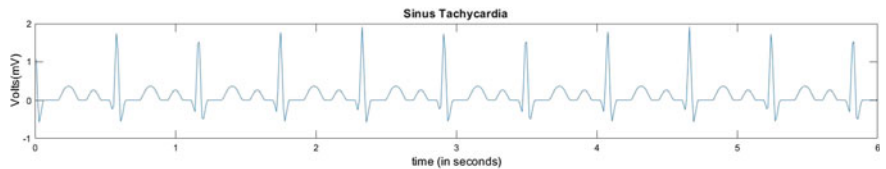


Fig. 4 Sinus tachycardia

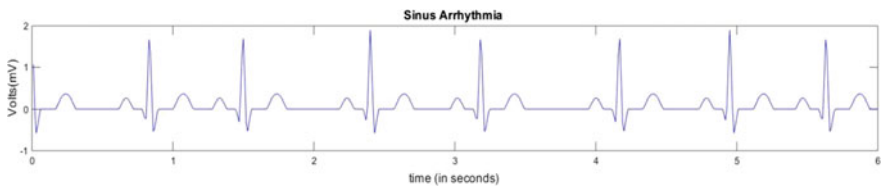


Fig. 5 Sinus arrhythmia

The rate is variable, depends on frequency. The Rhythm is irregular and P waves are present if QRS complexes are present. The interval length of PR and QRS complexes are normal. The waveform is shown in Fig. 6.

Supraventricular Tachycardia

The rate is in between 150 and 250 bpm. Even though the rhythm is regular, the P waves are not noticeable especially at the high rate range. The interval length of PR complexes are usually not discernable and that of QRS complexes are normal. A rate of 225 bps is considered for simulation. The generated waveform is as shown in Fig. 7.

Premature Atrial Complexes (PAC's)

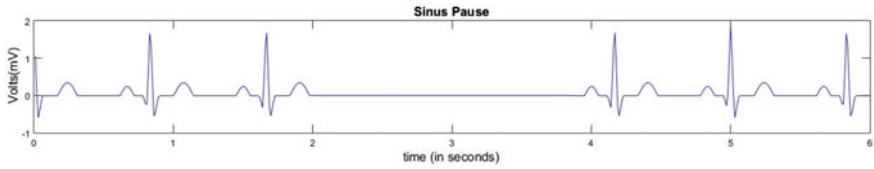


Fig. 6 Sinus pause

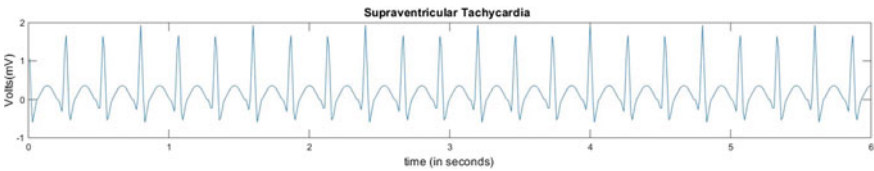


Fig. 7 Supraventricular tachycardia

The rate is usually regular but at times dependent on the rhythms. The interval length of PR and QRS complexes are normal. We have considered a rate of 72 bpm for simulation. The waveform is as shown in Fig. 8.

Idioventricular Rhythm

The rate is in between 20 and 40 bpm. The rhythm is regular, and P waves are absent. The interval length of PR complexes are not measurable and that of QRS complexes are wide and bizarre. A rate of 30 bpm is considered in our simulation. It is as shown in Fig. 9.

Accelerated Idioventricular Rhythm

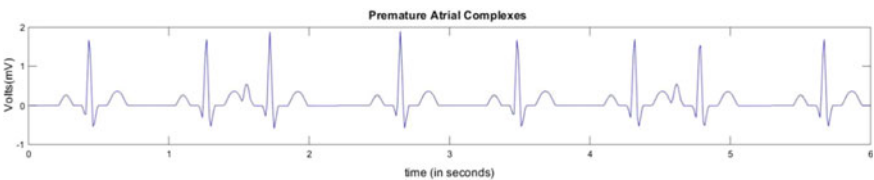


Fig. 8 Premature atrial complexes (PAC's)

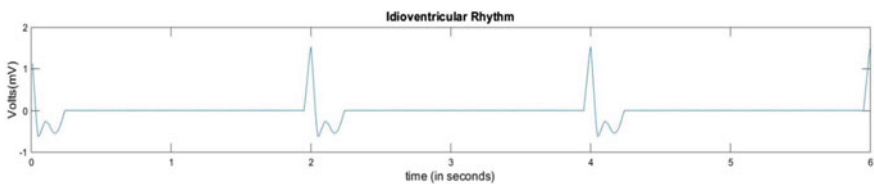


Fig. 9 Idioventricular rhythm

The rate is in between 41 and 100 bpm. The rhythm is regular, and P waves are absent. The interval length of PR complexes are not measurable and that of QRS complexes are wide and bizarre. A rate of 90 bpm is considered. The waveform is shown in Fig. 10.

Agonal Rhythm

The rate is less than 20 bpm. The rhythm is regular, and P waves are absent. The interval length of PR complexes are not measurable and that of QRS complexes are wide and bizarre. The simulated waveform is as shown in Fig. 11.

Premature Ventricular Contractions

The rate is normal. The rhythm is irregular during PVC's, and P waves are absent during PVC's. The interval length of PR complexes are not measurable and that of QRS complexes are wide and bizarre. We have considered 72 rpm for simulation. The waveform is as shown in Fig. 12.

Junctional Escape Rhythm

40 and 60 bpm is the ideal rate of junctional escape rhythm. The rhythm is regular, and P waves are usually upended which may be present or absent. The length of PR complexes will be shortened, if occurs before QRS complex, otherwise not measurable and that of QRS complexes are normal. We have considered a rate of 50 bpm. The waveform is as shown in Fig. 13.

Accelerated Junctional Rhythm

60 and 100 bpm is the expected rate for accelerated junctional rhythm. The rhythm is regular, and P waves are usually upended which may be present or absent. The length of PR complexes will be shortened, if occurs before QRS complex, otherwise not measurable and that of QRS complexes are normal. 72 rpm is considered for our simulation. Figure 14 depicts the generated waveform.

Junctional Tachycardia

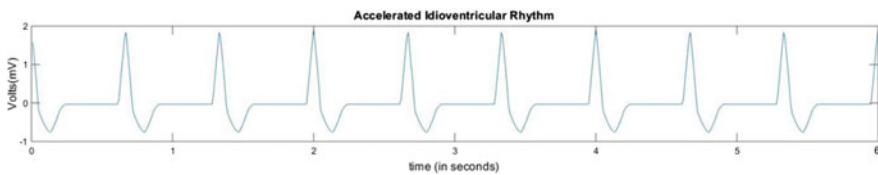


Fig. 10 Accelerated idioventricular rhythm

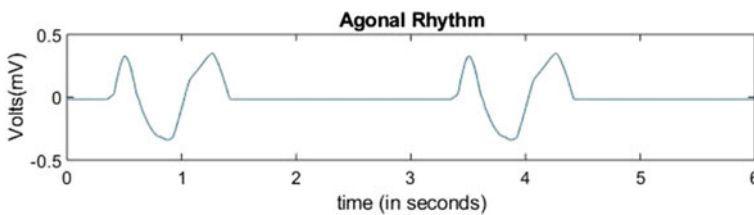


Fig. 11 Agonal rhythm

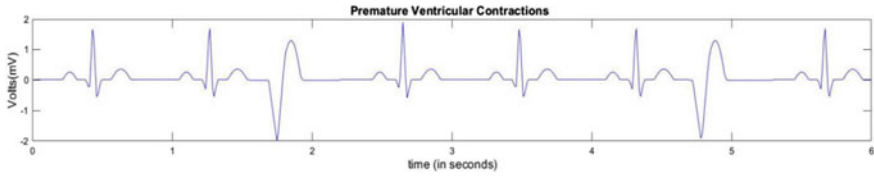


Fig. 12 Premature ventricular contractions

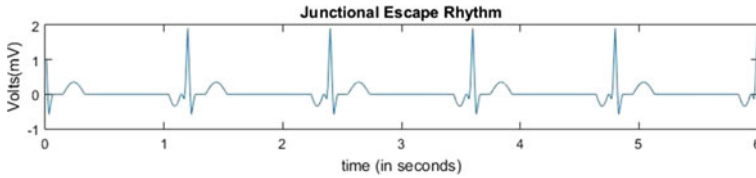


Fig. 13 Junctional escape rhythm

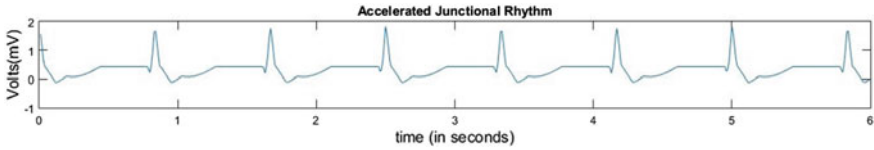


Fig. 14 Accelerated junctional rhythm

The rate is in between 101-180 bpm. The rhythm is regular, and P waves are usually upended which may be present or absent. The interval length of PR complexes will be shortened, if occurs before QRS complex, otherwise not measurable and that of QRS complexes are normal. We have considered a rate of 105 bpm. The simulated signal is shown in Fig. 15.

Premature Junctional Complex

The rate depends on underlying rhythm. The rhythm is irregular because of early beat, and P waves are usually upside down which may be present or absent. The interval length of PR complexes will be shortened, if occurs before QRS complex, otherwise not measurable and that of QRS complexes are normal. We have considered a rate of 72 rpm. The waveform is shown in Fig. 16.

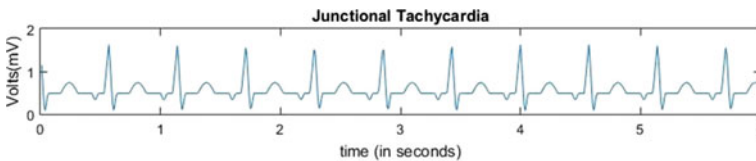


Fig. 15 Junctional tachycardia

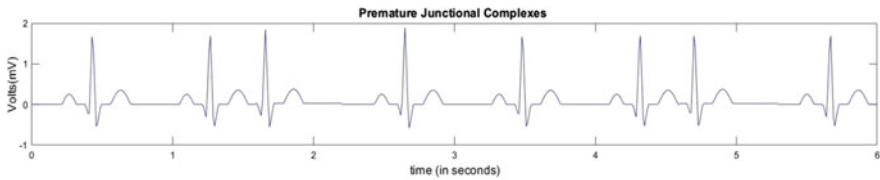


Fig. 16 Premature junctional complex

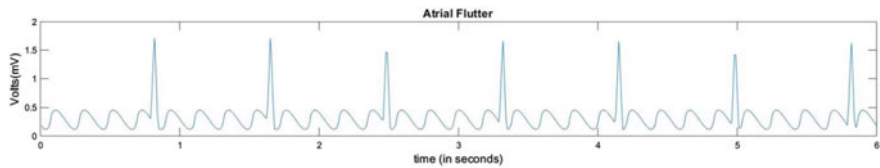


Fig. 17 Atrial flutter

Atrial Flutter

The atrial rate is between 250 and 400 bpm, and the ventricular rate is variable. The atrial rhythm will always be regular, and the ventricular rhythm may be irregular sometimes. P waves are absent. They are replaced by flutter waves (saw tooth waves). The length of the QRS complexes is normal. An atrial rate of 290 bpm and ventricular rate of 72 bpm were considered for simulation. The waveform is shown in Fig. 17.

Atrial Fibrillation

The atrial rate is 350 and 400 bpm, and the ventricular rate is variable. The rhythm is irregular. P waves are absent and are replaced by F waves (fibrillatory waves). The width of the QRS complexes is normal. A ventricular rate of 70 bpm is considered [22]. The waveform is as shown in Fig. 18.

Ventricular Tachycardia

The rate is between 101 and 250 bpm. The atrial rhythm is not distinguishable, and the ventricular rhythm is usually regular. P waves are absent, and the QRS complex is wide and bizarre. We have considered a rate of 145 bpm. The simulated waveform is as shown in Fig. 19.

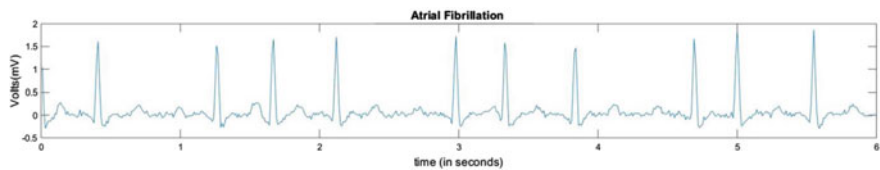


Fig. 18 Atrial fibrillation

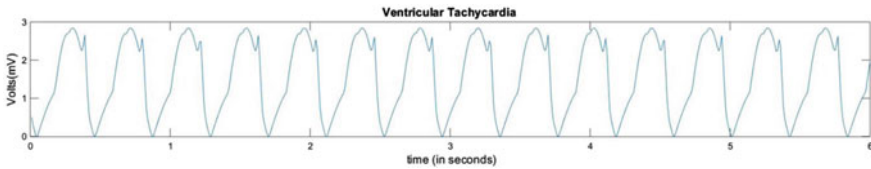


Fig. 19 Ventricular tachycardia

6 Conclusion

This paper presents the generation of ECG arrhythmia on the basis of Fourier series. 18 ECG arrhythmias were generated and presented in the paper. These arrhythmias differ from one another on the basis of amplitude, period, and length of the wave intervals. As the arrhythmias cause the heart to beat in an abnormal pace, pumping action of the heart may be ceased or might lead to complicated heart issues. This paper being a part of the consultancy work provides a good scope for the medical students to understand the signature features of different arrhythmias. The results obtained were validated with the real-time ECG signals [23, 24].

Acknowledgements We would like to thank KSHEMA, Nitte Deemed to be University, Deralakatte, Mangaluru, for the opportunity to develop the device.

References

1. A.K. Joshi, A. Tomar, M. Tomar, A review paper on analysis of electrocardiograph (ECG) signal for the detection of arrhythmia abnormalities. *Int. J. Adv. Res. Electr. Electron. Instrument. Eng.* **3**(10), 12466–12475 (2014)
2. M.K. Sarkaleh, A. Shahbahrami, Classification of ECG arrhythmias using discrete wavelet transform and neural networks. *Int. J. Comput. Sci. Eng. Appl.* **2**(1), 1 (2012)
3. Y. Özbay, R. Ceylan, B. Karlik, A fuzzy clustering neural network architecture for classification of ECG arrhythmias. *Comput. Biol. Med.* **36**(4), 376–388 (2006)
4. T.S. Diehl (ed.), *ECG Interpretation made Incredibly Easy!* (W. Kluwer/L. Williams & Wilkins, 2011)
5. R.S. Mane, A.N. Cheeran, V.D. Awandekar, P. Rani, Cardiac arrhythmia detection by ecg feature extraction. *Int. J. Eng. Res. Appl.* **3**(2), 327–332 (2013)
6. S. Ramesh, T. Sumanam, *Classification of ECG Signals*. PhD diss. (California State University, Sacramento, 2016)
7. F.G. Yanowitz, *Introduction to ECG Interpretation* (LDS Hospital and Intermountain Medical Center, 2012)
8. S. Bhowmick, P.K. Kundu, G. Sarkar, Synthesis of ECG waveform using simulink model, in *2016 International Conference on Intelligent Control Power and Instrumentation (ICICPI)*, pp. 61–64 (IEEE, 2016)
9. J. Ackora-Prah, A.Y. Aidoo, K.B. Gyamfi, *An Artificial ECG Signal Generating Function in MATLABM* (2013)
10. P. Dolinsky, I. Andraš, L. Michaeli, D. Grimaldi, Model for generating simple synthetic ECG signals. *Acta Electrotechnica et Informatica* **18**(3), 3–8 (2018)

11. M.A. Quiroz-Juárez, O. Jiménez-Ramírez, R. Vázquez-Medina, V. Breña-Medina, J.L. Aragón, R.A. Barrio, Generation of ECG signals from a reaction-diffusion model spatially discretized. *Sci. Rep.* **9**(1), 1 (2019)
12. M.K. Islam, G. Tangim, T. Ahammad, M.R. Khondokar, Study and analysis of ecg signal using matlab & labview as effective tools. *Int. J. Comput. Electr. Eng.* **4**(3), 404 (2012)
13. J. Kubicek, M. Penhaker, R. Kahankova, Design of a synthetic ECG signal based on the Fourier series, in *2014 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, pp. 1881–1885 (IEEE, 2014)
14. R. Karthik, *ECG simulation using MATLAB*. www.mathworks.com/matlabcentral/fileexchange/10858 (2003)
15. P. Mayapur, Classification of arrhythmia from ECG signals using MATLAB. *Int. J. Eng. Manage. Res. (IJEMR)* **8**(6), 115–129 (2018)
16. E.J. Luz, W.R. Schwartz, G. Cámara-Chávez, D. Menotti, ECG-based heartbeat classification for arrhythmia detection: A survey. *Comput. Methods Progr. Biomed.* **1**(127), 144–164 (2016)
17. S. Syama, G. Sai Sweta, P.I.K. Kavyasree, K. Jagan Mohan Reddy, Classification of ECG signal using machine learning techniques, in *2019 2nd International Conference on Power and Embedded Drive Control (ICPEDC)*, pp. 122–128 (IEEE, 2019)
18. D.K. Ojha, M. Subashini, Analysis of electrocardiograph (ecg) signal for the detection of abnormalities using matlab. *Int. J. Biomed. Biol. Eng.* **8**(2), 120–123 (2014)
19. C. Saritha, V. Sukanya, Y. Narasimha Murthy, ECG signal analysis using wavelet transforms. *Bulg. J. Phys.* **35**(1), 68–77 (2008)
20. Fourier Series—Sawtooth Wave. <https://mathworld.wolfram.com/FourierSeriesSawtoothWave.html>. Retrieved on 25 Oct 2020
21. S.T. Sanamdikar, S.T. Hamde, V.G. Asutkar, A literature review on arrhythmia analysis of ECG signal. *Int. Res. J. Eng. Technol.* **2**(3), 307–312 (2015)
22. J. Healey, G.D. Clifford, L. Kontothanassis, P.E. McSharry, An open-source method for simulating atrial fibrillation using ECGSYN, in *Computers in Cardiology*, pp. 425–427 (IEEE, 2004)
23. Six Second ECG Rhythm Strips. <https://ecg-educator.blogspot.com/2016/11/ecg-rhythm-strips.html>. Retrieved on 29 Oct 2020
24. Records in the MIT-BIH Arrhythmia Database. <https://www.physionet.org/physiobank/database/html/mitdbdir/records.htm>. Retrieved on 2 Oct 2020

A Survey on Methodologies and Database Used for Facial Emotion Recognition



Shashank M. Gowda and H. N. Suresh

1 Introduction

One of the non-verbal communication by which one understands the mood/mental state of a person is the expression of the face. Face emotion recognition has gained more attention due to its huge application area. Facial emotion recognition is being used in the areas like, which are not limited to, security, biometrics, advanced driver assistance systems [1], human–computer interaction, identification of the mental state [2], face expression synthesis, deception detection, music for mood [3, 4], operator fatigue detection, adult detection, etc. [5–7]. Not only in the technical field, face emotion recognition has been extensively studied and practiced in psychology [8]. As the field of facial emotion recognition has a wide range of application, it is challenging for the researchers to build a system which is robust and gives the desired output under various circumstances.

The problem domain of facial emotion recognition (FER) has been an important area of research for decades. At the very first, Darwin in his book, *The Expression of the Emotions in Man and Animals* (1872/1998), was the one who pointed out that humans across all cultures show the same expression for a particular emotion. Darwin’s theory was verified and supported by various researches in subsequent decades [9–11]. According to the study of Ekman, any individual would show six basic types of facial expressions which are anger, disgust, happiness, surprise, sadness, and fear. The major contributor to these facial expressions are eyes, nostrils, lips, eyebrows, muscle bulging area, and wrinkles on various parts of the face.

S. M. Gowda (✉)
Department of ECE, Yenepoya Institute of Technology, Moodbidri, India
e-mail: shashank.m.gowda.91@gmail.com

H. N. Suresh
Department of EIE, Bangalore Institute of Technology, Bengaluru, India
e-mail: hn.suresh@rediffmail.com

FER is a part of computer vision which uses image processing, machine learning, and deep learning technology [5]. Still, the researchers are being carried out to give a more efficient, robust, and automatic facial emotion recognition system [12]. A very important part of any facial emotion recognition system is to find the face in the given input image after it is acquired from the image acquisition system. This is a very easy task for a human being, but when it comes to computers, it will lead to a very challenging task, and this problem has been into research in the past few decades. The difficulty in detection of the face is imposed by factors such as variations in illumination, pose, orientation, and occlusions [13].

2 Overview of Facial Emotion Recognition System

The facial Emotion recognition system is a computer system that is capable of predicting the emotions of human beings from the input image or video frame. The facial emotion recognition system consists of three major steps which are image pre-processing, image feature extraction, and image classification. The block diagram of the facial emotion recognition system is as shown in Fig. 1. In the figure, it is clear that there are two ways in which the action of facial Emotion recognition can be carried out which are traditional machine learning and modern machine learning.

The traditional machine learning techniques involve image or data collection and pre-processing, Feature extraction from the image, classification based on the extracted features, and finally making the decisions based on the classification [12, 14, 15] are taken about what is the emotion in the image.

In the modern machine learning techniques, the images are fed into deep neural networks like convolutional neural networks (CNN) [16–18] where the input image is classified into any one of the expression classes through an end-to-end method. The paper puts light on both traditional as well as modern machine learning techniques which will serve as a strong base for selecting the technique to be used in facial emotion recognition systems.

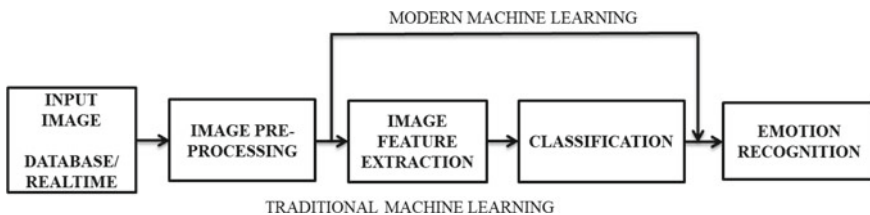


Fig. 1 Block diagram of facial emotion recognition system

3 Methodologies Used in Facial Emotion Recognition

For any computer vision system, the main component is the input image. The input to the system can be from an image acquisition system that takes images in real-time or from a standard database. The first step in image processing system is image acquisition because without image no processing can be done. The image can be taken in two different ways which are posed and real-world. The posed images [19] are taken in a controlled environment, and these images are easy to do pre-processing. The real-world [20] images or in the wild images are taken without any controlled environment. These images will be very challenging to do pre-processing as these scenes may be affected by wrong lighting conditions, angle of variation in head tilt, head rotation, occlusion, etc. The output image which is captured or taken from the database is then passed through various image pre-processing algorithms discussed in the next section.

A. *Image Pre-processing*

In the pre-processing stage, the image from the image acquisition system is processed for noise removal or enhancement, and the output of this stage is either an enhanced image or an image where the only face is cropped. The pre-processing may include the following steps: reducing the noise in the image, enhancing the image, face detection in the image, face segmentation, region of interest (ROI) detection, orientation correction, and face normalization [12, 21, 22]. Viola and Jones [23, 24] proposed an approach that can detect faces with the highest accuracy of about 95% and very little computational complexity, and it is 15 times faster than the existing techniques. Here in this, the Haar features are used to search for the desired object. To get the information about the neutral face, expression removal techniques are employed and morphological image processing methods were used for contrast equalization [6]. The pre-processing steps are important in improving the robustness and accuracy of the face emotion recognition system.

B. *Image Feature Extraction*

The key factor in recognizing facial expressions is the descriptor. The variance must be high for the descriptors of different expressions while the variance must be less for the same expressions. The feature can be extracted from the face by using any of these two techniques, and they are holistic template-based methods and geometric feature-based methods [5, 25]. In the former method, the whole face is considered for the generation of the feature vector, whereas in the latter method, the knowledge of the location of the facial components is used for generating the vector. Principle component analysis (PCA), linear discriminant analysis (LDA) [26], independent component analysis (ICA), and Gabor wavelet analysis [27] algorithms have been used on the face or some specific regions of the face to extract the facial features. The magnitude feature in the Gabor filter gives details about the organization of the face image [19].

LBP was originally proposed for texture analysis [28]. In the LBP method, the thresholding is done between the center pixel and the local pixels, and a binary code is

produced for each LBP feature. To measure the spatial structure around a pixel, LBP operator [15, 29] is the best operator. LBP has higher recognition accuracy even if it has less computational complexity. The LBP operator is an excellent texture descriptor, but it is not a good region descriptor. CS-LBP [30] is a better region descriptor which creates lower binary patterns than LBP. CS-LBP provides better results than LBP. In LBP from three orthogonal planes (LBP-TOP) [31], the spatial and temporal features are taken from LBP, and the dynamic texture features are extracted from the image sequence from three orthogonal planes. Another approach is to calculate the LOP- and LBP-TOP histograms and combine them into a single histogram [32].

A volume LBP (VLBP) developed by [29] combines both motion and appearance together to get the features. While the LBP compares the central value with the neighboring pixels, the LGC gets the histograms from the information about the relationship between neighboring pixels [33, 34]. Some others used the principles of horizontal diagonal and vertical diagonal have been proposed and used for efficient feature extraction [35]. To get better performance when the images have variations in illumination and noise, the local directional pattern approach is used [26]. Due to the advancement in the consumer computer market where more computational capable CPUs and GPUs are included, it gave rise to the development of complex algorithms for detecting the features. Above-discussed methods are used as pre-programmed and extracted features from the object elements.

Histogram of gradient orientations (HOG) [33, 36] feature extraction technique is illumination invariant. The featured can be found from the magnitude or pixel orientation. Using the gradient filter, the gradients of the image are calculated followed by angle and magnitude orientations of the corresponding gradients which are then divided into parts called bins. The variations in HOG can be found as co-occurrence HOG and coherence vector of oriented gradients [37, 38].

Another alternative is to use convolution neural networks algorithm for feature extraction and then feed the extracted features to the image classification step [39].

C. *Image Classification*

In the image classification step of the facial emotion recognition system, the emotions are classified to the respective facial expressions based on the feature information got from the image feature extraction stage. Expression classification is done using various parts of the face such as eyes, mouth, and cheeks. The expressions are classified as six basic expressions which are sadness, happiness, disgust, fear, anger, and surprise, and in some research, neutral face expression detection is also included. Some of the most popular image classification methods used in traditional machine learning are support vector machine (SVM) [40–42], nearest neighbor (NN) [30, 40, 43], hidden Markov model (HMM) [44], and adaboost algorithm [45].

In the emerging modern machine learning techniques, the classification of facial emotion is done in an end-to-end way depending on the automatically determined features using convolutional neural networks (CNN) [46–48]. The softmax loss function is most commonly used in CNN to minimize the cross-entropy between the ground truth and the estimated class. Even if deep learning has powerful feature

learning capability, there exists a problem in doing facial emotion recognition. These modern machine learning techniques need a sufficiently large amount of data in terms of image databases for training to avoid over-fitting. VGGNet, Inception, and AlexNet are most commonly used CNN architectures for image classification [16]. Almost all models using the baseline as VGGNet, Inception, or AlexNet architectures are known to achieve high accuracy on a particular dataset.

4 Facial Expression Databases

Database plays a very important role in any facial emotion recognition system. Having a huge training dataset will affect the accuracy of the machine learning system which is designed for facial emotion recognition. In this section, publically available databases in which the images are collected in a controlled environment or from the images available on the Web are discussed. The datasets that are widely used for validating the algorithms of facial emotion recognition are discussed. Table 1 provides an overview of the datasets in terms of no. of samples, no. of subjects, a condition in which the samples are acquired, and to what expressions the images in the database are annotated. Note that six basic expressions include anger, disgust, surprise, happiness, sadness, and fear.

CK+ [49]: The extended Cohn-Kanade (CK+) database is a laboratory-controlled database. It is the most commonly used database in the facial emotion recognition system. The database contains 593 image sequences of 123 subjects. It contains images of two types' frontal and 30°. The duration of the video sequence varies between 10 and 60 frames. The size of the images is 640×490 pixels. Out of 593 image sequences of 123 subjects, 327 sequences of 118 subjects are labeled. These are labeled based on the facial action coding system (FACS) with seven basic expression labels anger, sadness, disgust, fear, happiness, contempt, and surprise.

JAFFE [19]: The Japanese Female Facial Expression (JAFFE) database was collected at Kyushu University under controlled conditions in laboratory. It contains 213 images of 10 Japanese females where all the images taken are posed. The size of the image is 256×256 pixels. In this database, there are six basic expressions and one neutral expression of each subject. The six basic expressions considered here are anger, happiness, sadness, disgust, fear, and surprise.

MMI [50, 51]: The MMI Facial Expression Database is a laboratory-controlled database having 326 sequences of 32 subjects of both gender (44% female). The age range of the subjects is from 19 to 62. Out of 326 sequences, 213 are labeled with six basic expressions which are fear, happiness, sadness, anger, disgust, and surprise. The size of the images are 720×576 pixels. Out of 326 sequences, 205 sequences are frontal face images. The MMI sequence are onset-apex-offset labeled. This means that the images start with expressions which are neutral and in the middle reaches the peak and reaches to neutral.

FER2013 [20]: The FER2013 is an open-source dataset where the images are taken from the automatic Google image search API. It was introduced in ICML

Table 1 An overview of facial expression datasets

Database	No. of samples	No. of subjects	Condition	Expression annotation
CK + [49]	593 image sequence	123	Laboratory posed	6 basic expressions plus contempt and neutral
JAFFE [19]	213 images	10	Laboratory posed	6 basic expressions plus neutral
MMI [50, 51]	740 images and 2900 videos	25	Laboratory posed	6 basic expressions plus neutral
FER2013 [20]	35,887 images	Image search API	Web posed and Spontaneous	6 basic expressions plus neutral
TFD [39]	112,234 images	Image search API	Amalgamated posed	6 basic expressions plus neutral
CMU Multi-PIE [52, 53]	755,370 images	337	Laboratory posed	Smile, neutral, squint, disgust, scream, and surprised
EmotioNet [54]	1,000,000 images	Image search API	Web posed & Spontaneous	23 basic expressions or compound expressions
AffectNet [56]	450,000 images	Image search API	Web posed & Spontaneous	6 basic expressions plus neutral
ExpW [57]	91,793 images	Image search API	Web posed & Spontaneous	6 basic expressions plus neutral
AFEW [58]	54 Videos	330	Web posed	6 basic expressions plus neutral

2013 challenges. All the images in the dataset are grouped into labels of seven emotions with the file names as 0 to 6. The seven emotions are anger, disgust, sadness, surprise, neutral, fear, and happiness. The size of the images in the dataset is 48×48 pixels. The resulting dataset contains a total of 35,887 images; most of the images are in-the-wild images. Out of 35,887 images in the dataset, 28,709 images are for training, 3589 images are for validation, and 3589 images are test images. The images in each folder are as follows 0:–4593 images–Angry, 1:–547 images–Disgust, 2:–5121 images–Fear, 3:–8989 images–Happy, 4:–6077 images–Sad, 5:–4002 images–Surprise, 6:–6198 images–Neutral.

TFD [39]: The Toronto Face Database (TFD) contains a combination of various facial expression datasets. The total number of images available with this database is 112,234 images. Out of 11,234 images, only 4178 images are labeled with any one of six basic expressions anger, disgust, sadness, surprise, fear, happiness, and one neutral expression. All the faces in the dataset are of size 48×48 pixels.

CMU Multi-PIE [52, 53]: CMU Multi-PIE face database has 755,370 images of 337 subjects. It is a laboratory-controlled database where the images are taken in 15 multiple view angles and 19 different illumination settings. The subjects were instructed to pose for different facial expressions which are, happy, neutral, surprise, angry, disgust, squint, and scream. All the images in the database are labeled for the above six expressions.

EmotioNet [54]: This database is a large-scale database which contains more than one million facial expression images which are collected from the Internet. The images were downloaded using various search engines by automatically inserting the emotion keywords in WordNet [55]. Some example of queries used for the image search are happiness, fear (query by emotion), AU4, AU6 (query by action units number), anxiety, and disapproval (query by keyword). Out of one million images, 950,000 images were auto labeled using the automatic action unit (AU) model [54]. EmotioNet is the best database of FACS model of images in the wild with a great amount of subject variation.

AffectNet [56]: This database has a collection of more than one million facial emotion images from the Internet with faces extracted from landmark points. Out of that, 450,000 images are manually annotated by twelve human experts, and the images are tagged for occlusions in the face. It is the best largest Web-based database.

ExpW [57]: The expression in-the-wild database (ExpW) contains 91,793 images of faces downloaded using Google image search. The motivation for creating a new dataset and the methodology of creating it was proposed in [54]. Each image is manually labeled as one of the six basic expression and one neutral expression categories.

AFEW [58]: Acted facial expressions in the wild database contains the video clips from 54 movies having 330 subjects from the age 1–77 years. The annotations of the clips were done according to the six basic emotions and one neutral expression. This database has spontaneous facial expressions, different resolutions, varied focus, and occlusions of the face which is almost similar to the in-the-wild photos.

5 Conclusion

In the past decade, facial emotion recognition has gained considerable importance from researchers in computer vision. FER is widely used in real-time applications such as medical robotics, driver attention monitoring, detecting deceptions, human–computer interaction, and forensic. This survey paper is useful for new researchers in the field of facial emotion recognition in recognizing the opportunities available and in gaining some insight on what methodologies to use to resolve the challenges. This

survey provides the researcher with the knowledge of various algorithms used for image pre-processing, image feature extraction, and image classification. All these methodologies are mainly based on traditional and modern machine learning, and the accuracy of the recognition depends highly on the availability of a large number of labeled data for training in the form of databases.

For traditional machine learning methodologies, the database with a lesser number of labeled images can be used but for modern machine learning methodologies, and the databases having more number of labeled images for testing will give a better performance. Facial emotion recognition system in general will have three major steps which are image pre-processing, image feature extraction, and image classification. The paper has given extensive survey on various methodologies that can be used in each of the three steps as discussed earlier. The performance of the facial expression system will totally depend on the methodology chosen in the three stages and also will vary for different databases. There is no hard and fast rule for selecting the methodologies, which is where the aspect of research comes into picture. The researcher has to try out the combination of the methodologies and databases relevant to the traditional machine learning techniques to get more accuracy in recognizing the facial expressions. Based on the previous works in the field of facial emotion recognition, Viola–Jones algorithm and local binary pattern were majorly used by researchers in the image pre-processing step. PCA and LBP were majorly used for feature extraction, and it can be noted from the survey that deep learning-based feature extraction can also be used if a large training database is available. SVM and HMM classifiers were the most preferred algorithms for image classification in traditional machine learning methods.

In recent years, the modern machine learning uses deep learning for solving the problems of identification and recognition. Several deep learning techniques are available, of which convolutional neural networks (CNN) is best suited to solve the recognition problems. The research must be done in terms of the architecture of the CNN and also on tweaking the hyper-parameters of the activation function to achieve a good accuracy.

Various databases available are compared against parameters which will be helpful for a new researcher to dive into the field of facial emotion recognition. JAFFE and CK+ databases are the most widely used databases for validating facial emotion recognition problems in traditional machine learning. In the modern machine learning, the databases used for in traditional machine learning cannot be used as it will lead to over-fitting problem. The modern machine learning techniques require huge number of images from the database for proper recognition of the expressions which consumes a lot of computational power and complexity. If the researcher has sufficient computational resources, then some of the mentioned databases which have millions of images can be used for training and validation and achieve good accuracy.

References

1. Z. Zhang, J. Zhang, A new real-time eye tracking for driver fatigue detection, in *6th International Conference ITS Telecommunication. Proceedings*, no. 60572027, pp. 8–11 (2006)
2. M.K. Mandal, R. Pandey, A.B. Prasad, Facial expressions of emotions and schizophrenia: a review. *Schizophr. Bull.* **24**(3), 399–412 (1998)
3. F.H. Hariz, K.N. Upadhyaya, T.F.M. Danish, B. B, S.M. Gowda, Facial emotion recognition for smart applications. *Int. J. Eng. Res. Technol.* **8**(13), 35–37 (2020)
4. A. Dureha, An accurate algorithm for generating a music playlist based on facial expressions. *Int. J. Comput. Appl.* **100**(9), 33–39 (2014)
5. J.S Nayak, P.G, M. Vatsa, M. Reddy Kadiri, S. S, Facial expression recognition: a literature survey. *Int. J. Comput. Trends Technol.* **48**(1), 1–4 (2017). <https://doi.org/10.14445/22312803/ijctt-v48p101>
6. V. Radhamani, G. Dalin, A supporting survey to step into a novel approach for providing automated emotion recognition service in mobile phones, in *Proceedings of 2nd International Conference on Inventive Systems and Control ICISC 2018*, no. Icisc, pp. 35–39 (2018). <https://doi.org/10.1109/icisc.2018.8399108>
7. M.G. Frank, Facial expressions. *Int. Encycl. Soc. Behav. Sci.*, pp. 5230–5234 (2001). <https://doi.org/10.1016/b0-08-043076-7/01713-7>
8. R. Adolphs, How we recognize facial expressions. *Psychol. Sci.* **15**(5), 222–226 (2006)
9. P. Ekman, W.V. Friesen, Constants across cultures in the face and emotion. *J. Personality Soc. Psychol.* **17**(2), 124–129 (1971)
10. P. Ekman, W.V. Friesen, Constants across cultures in the face and emotion. *J. Person. Soc. Psychol.* **17**(2), 124–129 (1971)
11. P. Ekman, P. Ekman, *Psychol. Sci.* (1992). <https://doi.org/10.1111/j.1467-9280.1992.tb00253.x>
12. B. Fasel, J. Luettin, Automatic facial expression analysis: a survey. *Pattern Recognit.* **36**(1), 259–275 (2003). [https://doi.org/10.1016/S0031-3203\(02\)00052-3](https://doi.org/10.1016/S0031-3203(02)00052-3)
13. A. Kumar, A. Kaur, M. Kumar, Face detection techniques: a review. *Artif. Intell. Rev.* **52**(2), 927–948 (2019). <https://doi.org/10.1007/s10462-018-9650-2>
14. C. Chang, Y. Huang, C. Yang, Personalized facial expression recognition in color image, in *2009 Fourth International Conference on Innovative Computing, Information and Control*, pp. 1164–1167 (2009). <https://doi.org/10.1109/icicic.2009.294>
15. R. Verma, M.Y. Dabbagh, Fast facial expression recognition based on local binary patterns, in *Canadian Conference of Electrical and Computer Engineering*, pp. 1–4 (2013). <https://doi.org/10.1109/ccece.2013.6567728>
16. A. Agrawal, N. Mittal, Using CNN for facial expression recognition : a study of the effects of kernel size and number of filters on accuracy. *Vis. Comput.* **2** (2019). <https://doi.org/10.1007/s00371-019-01630-9>
17. J. Zou, X. Cao, S. Zhang, B. Ge, A facial expression recognition based on improved convolutional neural network, in *Proceedings of 2019 IEEE International Conference on Intelligent Applied System Engineering ICIASE 2019*, pp. 301–304 (2019). <https://doi.org/10.1109/iciase.45644.2019.9074074>
18. N. Perveen, D. Singh, C.K. Mohan, Spontaneous facial expression recognition : a part based approach, in *15th IEEE International Conference on Machine Learning Application*, no. March 2018, pp. 819–824 (2016). <https://doi.org/10.1109/icmla.2016.162>
19. M. Lyons, S. Akamatsu, M. Kamachi, J. Gyoba, Coding facial expressions with Gabor wavelets, in *Proceedings of 3rd IEEE International Conference on Automation Face Gesture Recognition, FG 1998*, pp. 200–205 (1998). <https://doi.org/10.1109/afgr.1998.670949>
20. I.J. Goodfellow et al., Challenges in representation learning: a report on three machine learning contests, in *International Conference on Neural Information Processing*, pp. 117–124
21. Y. Ji, K. Idrissi, Automatic facial expression recognition based on spatiotemporal descriptors. *Pattern Recognit. Lett. J.* **33**, 1373–1380 (2012). <https://doi.org/10.1016/j.patrec.2012.03.006>

22. A. Poursaberi, H.A. Noubari, M. Gavrilova, S.N. Yanushkevich, Gauss-Laguerre wavelet textural feature fusion with geometrical information for facial expression identification. *EURASIP J. Image Video Process.* **2012**(1), 1–3 (2012)
23. P. Viola, M. Jones, Rapid object detection using a boosted cascade of simple features, in *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 511–518 (2001)
24. A.D. Sergeeva, A.V. Savin, V.A. Sablina, O.V. Melnik, Emotion recognition from micro-expressions: search for the face and eyes, in *2019 8th Mediterranean Conference on Embedded Computing (MECO)*, MECO 2019—Proceedings, no. June, pp. 1–4 (2019). <https://doi.org/10.1109/meco.2019.8760029>
25. R. Brunelli, T. Poggio, Face recognition: features versus templates. *IEEE Trans. Pattern Anal. Mach. Intell.* **15**(10), 1042–1052 (1993). <https://doi.org/10.1109/34.254061>
26. T. Jabid, H. Kabir, O. Chae, Facial expression recognition using local directional pattern (LDP), in *Proceedings of 2010 IEEE 17th International Conference Image Processing*, pp. 1605–1608 (2010)
27. Y. Tian, T. Kanade, J.F. Cohn, Facial expression analysis, in *Handbook of Face Recognition* (Springer, 2003)
28. T. Ojala, M. Pietikäinen, D. Harwood, A comparative study of texture measures with classification based on featured distributions. *Pattern Recogn.* **29**(1), 51–59 (1996)
29. G. Zhao, M. Pietik, Dynamic texture recognition using local binary patterns with an application to facial expressions. *IEEE Trans. Pattern Anal. Mach. Intell.*, pp. 1–14 (2007)
30. N. Neeru, L. Kaur, Face recognition based on LBP and CS-LBP technique under different emotions, in *2015 IEEE International Conference Computational Intelligence and Computing Research ICCIC 2015* (2016). <https://doi.org/10.1109/iccic.2015.7435803>
31. Y. Guo, Y. Tian, X. Gao, X. Zhang, Micro-expression recognition based on local binary patterns from three orthogonal planes and nearest neighbor method, in *International Joint Conference Neural Networks*, no. 20071152, pp. 3473–3479 (2014)
32. Y. Gizatdinova, V. Surakka, G. Zhao, E. Mäkinen, R. Raisamo, Facial expression classification based on local spatiotemporal edge and texture descriptors, in *Proceedings of the 7th International Conference on Methods and Techniques in Behavioral Research*, pp. 14–17 (2010)
33. J. Kumari, R. Rajesh, K.M. Pooja, Facial expression recognition: a survey. *Procedia Comput. Sci.* **58**, 486–491 (2015). <https://doi.org/10.1016/j.procs.2015.08.011>
34. Y. Tong, R. Chen, Y. Cheng, Facial expression recognition algorithm using LGC based on horizontal and diagonal prior principle. *Opt. Int. J. Light Electron Opt.* (2014). <https://doi.org/10.1016/j.ijleo.2014.04.062>
35. S.S. Dlay, W.L. Woo, J.A. Chambers, Facial expression recognition using local gabor gradient code-horizontal diagonal descriptor, in *2nd IET International Conference Intelligent Signal Processing*, pp. 1–6 (2015)
36. P. Sombatpiboonporn, T. Charoenpong, Human edge segmentation from 2D images by histogram of oriented gradients and edge matching algorithm, in *2019 First International Symposium on Instrumentation, Control, Artificial Intelligence, and Robotics*, pp. 29–32 (2019)
37. R. Rajesh et al., On Experimenting with pedestrian classification using neural network. *IJCNN*, pp. 107–111 (2011)
38. R. Rajesh, K. Rajeev, K. Suchithra, V.P. Lekhesh, V. Gopakumar, N.K. Ragesh, Coherence vector of oriented gradients for traffic sign recognition using neural networks, in *Proceedings of 3rd International Conference on Electronics Computing Technology*, pp. 907–910 (2011)
39. S. Li and W. Deng, “Deep Facial Expression Recognition : A Survey,” *Int. Conf. Neural Inf. Process.*, 2013
40. C. Shan, S. Gong, P.W. McOwan, Facial expression recognition based on Local Binary Patterns: A comprehensive study. *Image Vis. Comput.* **27**(6), 803–816 (2009). <https://doi.org/10.1016/j.imavis.2008.08.005>
41. C.C. Lai, C.H. Ko, Facial expression recognition based on two-stage features extraction. *Optik (Stuttg)* **125**(22), 6678–6680 (2014). <https://doi.org/10.1016/j.ijleo.2014.08.052>

42. A. Ankit, D. Narayan, A. Kumar, Transformation of facial expression into corresponding emoticons. *Int. J. Eng. Adv. Technol.* **8**(5), 256–258 (2019)
43. P. Tarnowski et al., Emotion recognition using facial expressions Emotion recognition using facial expressions. *Procedia Comput. Sci.* **108**, 1175–1184 (2017). <https://doi.org/10.1016/j.procs.2017.05.025>
44. Z. Uddin, T. Kim, B.C. Song, An optical flow feature-based robust facial expression recognition with hmm from video. *Int. J. Innov. Comput. Inf. Control* **9**(4), 1409–1421 (2013)
45. P. Yang, Q. Liu, D.N. Metaxas, *Boosting Coded Dynamic Features for Facial Action Units and Facial Expression* (IEEE, 2007)
46. N.S. Goud, K.R. Reddy, G.A., G.S. Sucheta, Facial emoji recognition. *Int. J. Trend Sci. Res. Dev.* **3**(3), 1330–1333 (2019). <https://doi.org/10.31142/ijtsrd23166>
47. S. Bazrafkan, T. Nedelcu, P. Corcoran, P. Filipczuk, Deep learning for facial expression recognition : a step closer to a SmartPhone that knows your moods, in *IEEE International Conference Consumer Electronics* (2017). <https://doi.org/10.1109/icce.2017.7889290>
48. B. Taha, D. Hatzinakos, Emotion recognition from 2D facial expressions, in *2019 IEEE Canadian Conference on Electrical Computing Engineering*, pp. 1–4 (2019)
49. P. Lucey et al., The Extended Cohn-Kanade Dataset (CK+): A complete dataset for action unit and emotion-specified expression Patrick, in *2010 IEEE Computer Society Conference on Computing Vision Pattern Recognition—Work. CVPRW 2010*, pp. 94–101 (2010)
50. M.F. Valstar, M. Pantic, Induced disgust, happiness and surprise : an addition to the MMI facial expression database, in *Proceedings of 3rd International Workshops Emotions* (Satellite Lr. Corpora Research Emotion Affect, 2010)
51. M. Pantic, M. Valstar, R. Rademaker, L. Maat, Web-based database for facial expression analysis Delft University of Technology, in *IEEE International Conference on Multimedia Expo, 2005. ICME 2005*, pp. 0–4 (IEEE, 2005)
52. A. Mollahosseini, D. Chan, M.H. Mahoor, Going deeper in facial expression recognition using deep neural networks, in *IEEE Winter Conference on Applications Computing Vision* (2016)
53. R. Gross, I. Matthews, J. Cohn, T. Kanade, S. Baker, Multi-PIE. *Image Vis. Comput.* **28**(5), 807–813 (2010). <https://doi.org/10.1016/j.imavis.2009.08.002>
54. C.F. Benitez-quiros, R. Srinivasan, A.M. Martinez, EmotioNet : an accurate, real-time algorithm for the automatic annotation of a million facial expressions in the wild, in *Proceedings of IEEE International Conference on Computing Vision Pattern Recognition (CVPR)*, Las Vegas, NV, USA (2016). <https://doi.org/10.1109/cvpr.2016.600>
55. G.A. Miller, WordNet: a lexical database for english. *Commun. ACM* **38**(11), 39–41 (1995)
56. A. Mollahosseini, S. Member, B. Hasani, S. Member, AffectNet : a database for facial expression, valence, and arousal computing in the wild. *IEEE Trans. Affect. Comput.* **3045**(c), 1–17 (2017). <https://doi.org/10.1109/taffc.2017.2740923>
57. Z. Zhang, P. Luo, C.C. Loy, X. Tang, From facial expression recognition to interpersonal relation prediction. *Int. J. Comput. Vis. Manusc.* (2017)
58. S. Zafeiriou, A. Papaioannou, I. Kotsia, M. Nicolaou, G. Zhao, Facial affect ‘in-the-wild’: a survey and a new database, in *IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pp. 1487–1498 (2016). <https://doi.org/10.1109/cvprw.2016.186>

Intensity-Based Feature Extraction of Real-Time Transformer Oil Images



C. M. Maheshan and H. Prasanna Kumar

1 Introduction

A static electric machine used to transfer power between the two circuits at constant frequency is called transformer. The uninterrupted reliable operation of power transformer ensures the effectiveness of power transfer as well as economic stability. In general, mineral oil extracted from crude petroleum is used for cooling along with insulating medium. Hence, transformer oil is also called as insulating oil or cooling oil or dielectric oil, etc. The life of transformer depends on the performance of cooling oil. During the operation, transformer oil experiences different types of stresses such as thermal, electromechanical, and chemical which degrades the transformer oil. This deformation of the oil alters the properties of oil results in underperformance of the in service transformer [1–4]. Even though there are different classical methods to measure the performance of transformer oil, an image processing techniques emerged as a new non-destructive diagnosis method to determine the performance condition of the transformer [5, 6].

Digital image processing is a method of accumulate the information of an image by performing different operations on it. This technique focuses on improvement or modification or manipulation of picture information for human interpretation as well as to process the image data for storage, transmission, and representation according to computer perception. A smallest point of information of an image is called as pixels or picture elements. Pixels provides the intensity information about the image

C. M. Maheshan (✉)

Department of Electrical Engineering, VTU-RRC-Mysuru (Affiliated to VTU,Belagavi) / UVCE, Regional Resource Centre-Mysuru / K.R.Circle, Mysuru / Bengaluru, India
e-mail: maheshan.cm@gmail.com

H. Prasanna Kumar

Department of Electrical Engineering, University Visvesvaraya College of Engineering(UVCE) , Bangalore University, K.R. Circle, Bengaluru, Karnataka 5600001, India
e-mail: uvcehp@gmail.com

[7]. Intensity-based feature extraction simplifies the complete information about the image. This paper uses Weiner filter, non-local means (NLM) filter, original shock filter (OSF), complex shock filter (CSF), and improved complex shock filters (ICSF) to eliminate the noise in transformer oil images [8–12].

The objective of this paper is to determine the intensity-based GLCM texture features of transformer oil images of different ages using image processing technique. Transformer oil images of different ages were collected from Ref. [13]. GLCM features such as homogeneity, contrast, energy, correlation, and entropy are extracted from the original test image as well as from the enhanced transformer oil images [14, 15]. These are achieved through MATLAB programming tool. Texture features of transformer oil image represent the operating condition of the transformer oil.

This paper organized as follows: Sect. 2 discusses entropy and GLCM features, Sect. 3 provides the comprehensive results and discussions, and in Sect. 4 the conclusion of the paper is presented.

2 Glcm Features

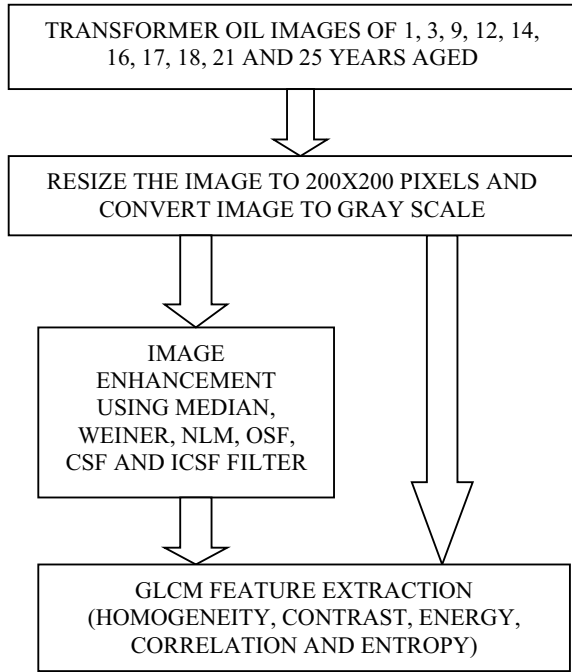
Feature extraction is a technique of extracting visual content of pictures for indexing and retrieval. Texture feature analysis aims in discovering a way of representing the characteristics of an image. GLCM is used to compute special dependence of gray levels in an image.

GLCM is a second-order statistical tool to investigate the texture of an image by the calculations of frequency of occurrence of pixel neighborhood pairs having a specific spatial relationship. GLCM is formed in four directions 0°, 45°, 90°, and 135° with symmetric as well as asymmetric representation. A number of statistical features are extracted from GLCM. In this paper, the features namely homogeneity, contrast, correlation, energy, and entropy evaluated through MATLAB programming. The formulas for calculation of GLCM texture features of an image are given in Table 1. Figure 1 illustrates the GLCM texture feature extraction methodology for transformer oil images of different ages. All the images are resized to 200 × 200

Table 1 GLCM texture feature

S. no.	Texture feature	Formula
1	Homogeneity	$\sum_{i,j=0}^{N-1} \frac{P_{i,j}}{1+(i-j)^2}$
2	Contract	$\sum_{i,j=0}^{N-1} P_{i,j} (i - j)^2$
3	Energy	$\sum_{i,j=0}^{N-1} P_{i,j}^2$
4	Correlation	$\sum_{i,j=0}^{N-1} P_{i,j} \left[\frac{(i-\mu_i)(j-\mu_j)}{\sqrt{(\sigma_i^2)(\sigma_j^2)}} \right]$
5	Entropy	$\sum_{i,j=0}^{N-1} P_{i,j} (-\ln P_{i,j})$

Fig. 1 GLCM texture feature extraction methodology



pixels and preprocessed through different image enhancement filters. GLCM features are extracted for raw as well as enhanced images.

3 Results and Discussions

In this section, the visualization as well as numerical results of GLCM texture features for transformer oil gray converted as well as enhanced gray images of different ages are presented. The acquired images are converted into gray images, and these images are enhanced through median, Wiener, non-local means, original shock filter, complex shock filter, and improved complex shock filters. In this paper, transformer oil images at ages of one, three, nine, twelve, fourteen, sixteen, seventeen, eighteen, twenty-one, and twenty-five years are considered. Figures 2, 3, 4, 5, 6, 7, 8, 9, 10 and 11 demonstrates the gray images along with histograms of transformer oil test images working under operation at different ages ranging from one-year to the twenty-five-year-old. The visualization results assessed with image histogram through the occurrence of pixel intensity values for all the test images. The analysis of intensity-based features of transformer oil images through GLCM texture features is relayed on the occurrence of symmetric or asymmetric intensities around the neighborhood pixels. MATLAB programming approach is adopted to assess each of the test images.

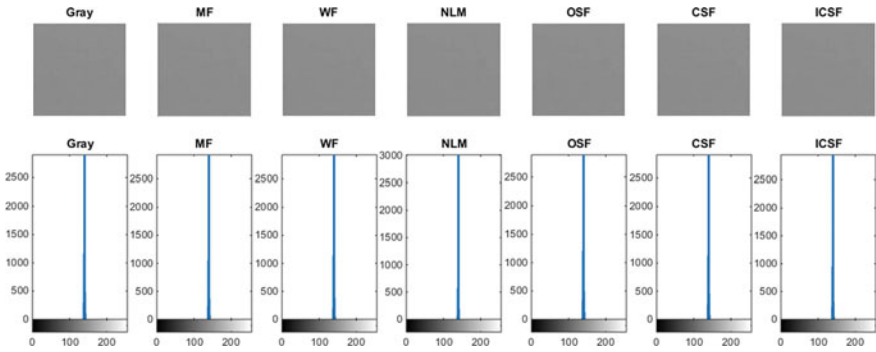


Fig. 2 Transformer oil one-year-old gray and filtered gray images along with its histogram

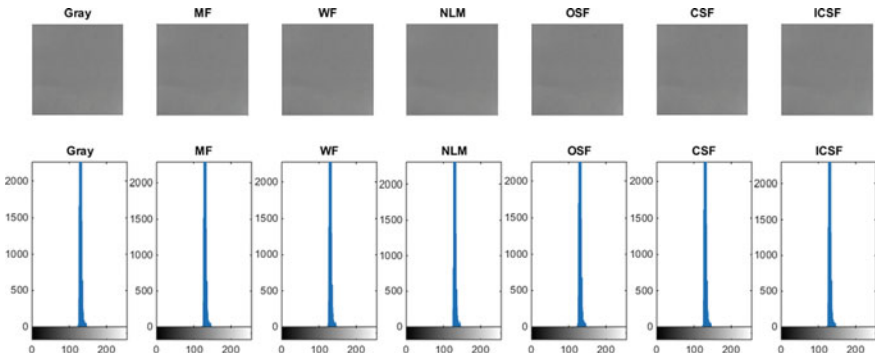


Fig. 3 Transformer oil three-year-old gray and filtered gray images along with its histogram

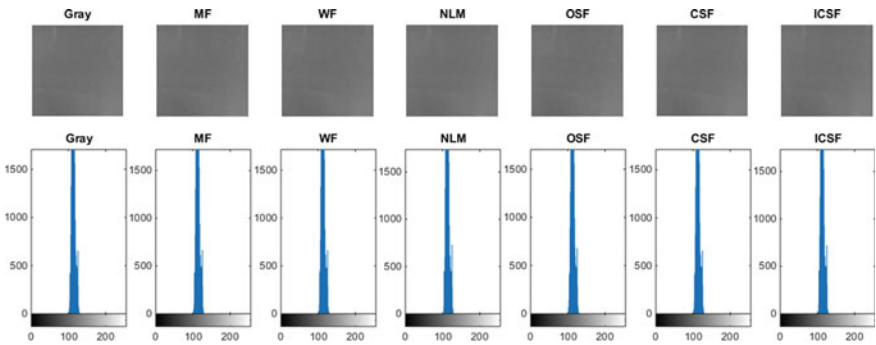


Fig. 4 Transformer oil nine-year-old gray and filtered gray images along with its histogram

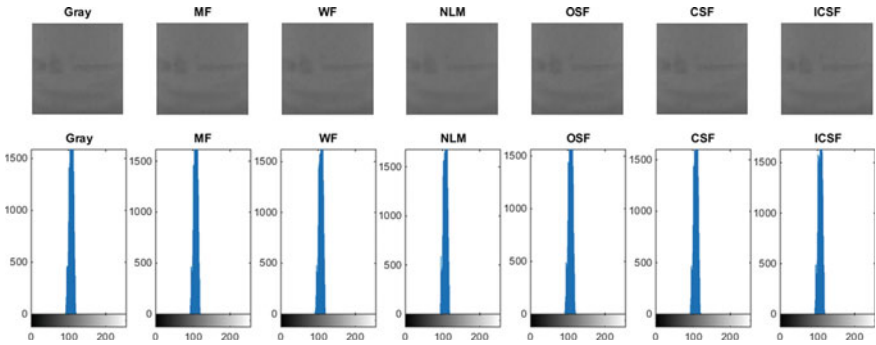


Fig. 5 Transformer oil twelve-year-old gray and filtered gray images along with its histogram

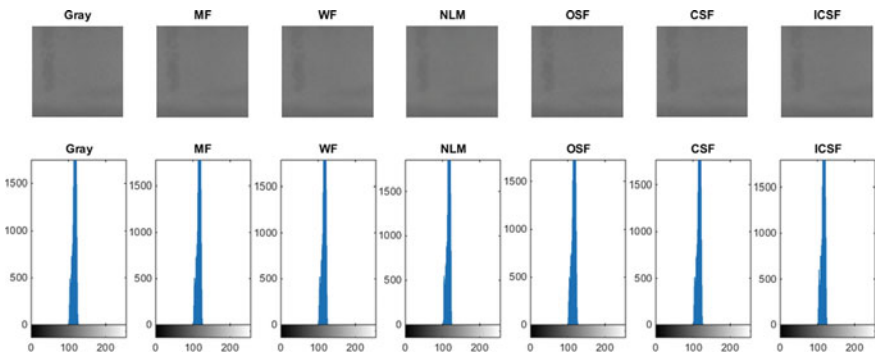


Fig. 6 Transformer oil fourteen-year-old gray and filtered gray images along with its histogram

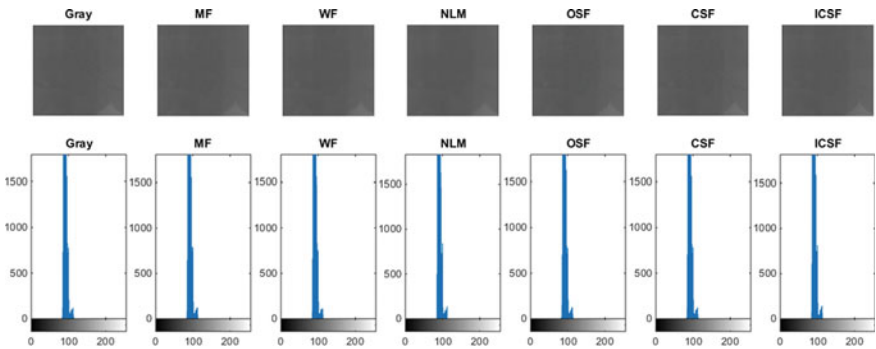


Fig. 7 Transformer oil sixteen-year-old gray and filtered gray images along with its histogram

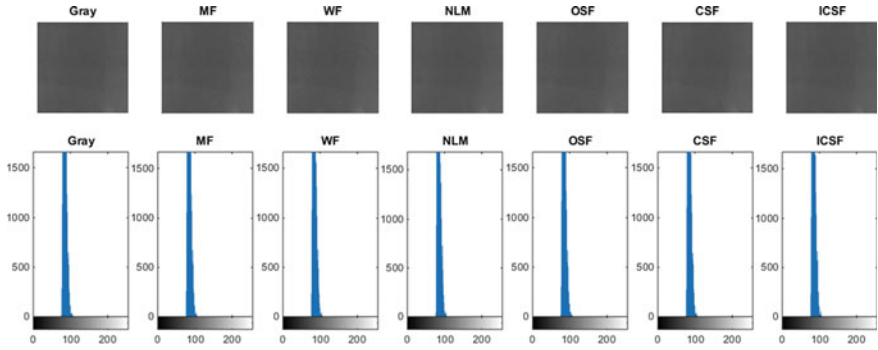


Fig. 8 Transformer oil seventeen-year-old gray and filtered gray images along with its histogram

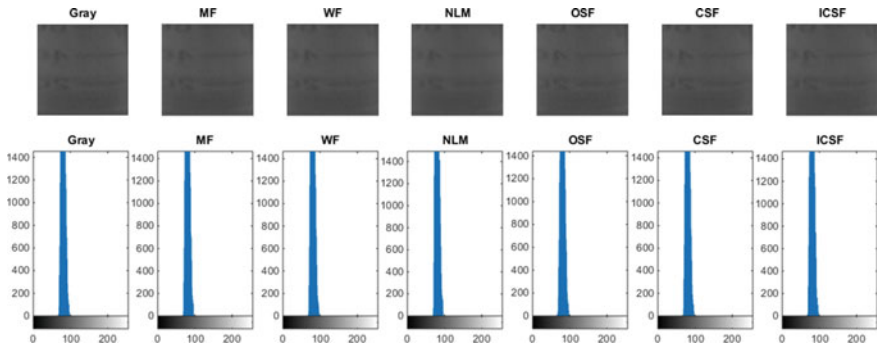


Fig. 9 Transformer oil eighteen-year-old gray and filtered gray images along with its histogram

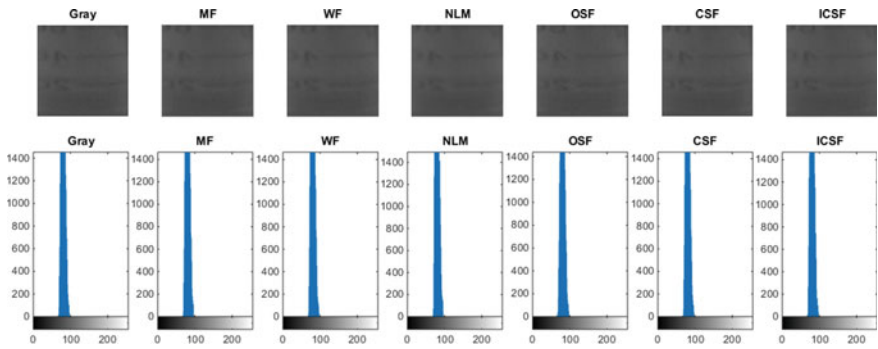


Fig. 10 Transformer oil twenty-one-year-old gray and filtered gray images along with its histogram

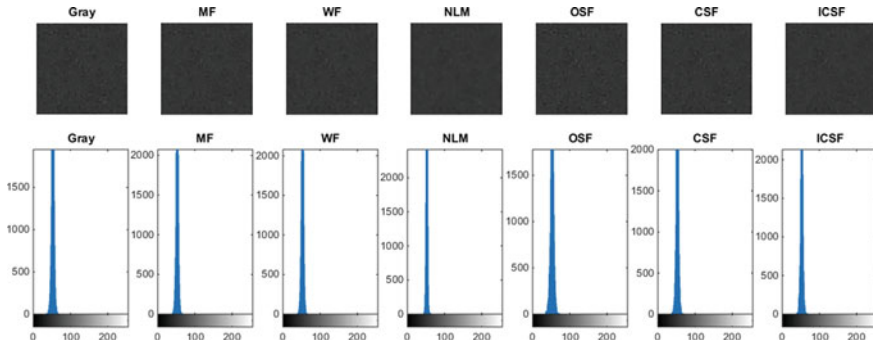


Fig. 11 Transformer oil twenty-five-year-old gray and filtered gray images along with its histogram

The results of GLCM texture features such as homogeneity, contrast, correlation, energy, and entropy of all images are presented in Tables 2, 3, 4, 5, 6, 7, 8, 9, 10 and 11. It can be observed that the texture feature values represent the approximate textures of transformer oil images operating under different ages. Hence, it can be possibly classified the age of the oil with respect to these GLCM texture features. From this, it can be predicted to replace or filter the transformer oil.

4 Conclusions

The intensity-based features extraction was evaluated and presented. This paper demonstrates visualization results of transformer oil gray as well as enhanced images. The image histogram quantifies the significant occurrence of pixel values. In this paper, GLCM texture features estimated in order to notice and discriminate the textures of the transformer oil images under ages. GLCM texture features such as homogeneity, contrast, correlation, energy, and entropy are evaluated for all the test images. The results indicate that the features evaluated were significantly identifying texture features. This paper concludes that the performance of GLCM texture features quantify the essential features of test images with and without noise eliminating filters. The extension of this study can be utilized to monitor the condition of transformer oil of in service transformers.

Table 2 GLCM texture feature of one year aged transformer oil

Sl. no.	GLCM texture features	One year aged oil	Median filter	Weiner filter	NLM filter	OSF	CSF	ICSF
1	Homogeneity	0.04357	0.04353	0.04358	0.04357	0.04357	0.04357	0.04357
2	Contrast	0.004202	0.004205	0.004206	0.004209	0.004203	0.004203	0.004203
3	Correlation	293.3197	298.1569-	297.2241	295.4628	294.9237	294.1903	293.9123
4	Energy	2.4268e-05	2.4271e-05	2.4269e-05	2.4268e-05	2.4269e-05	2.4269e-05	2.4269e-05
5	Entropy	0.033731	0.034534	0.035221	0.032340	0.032250	0.032739	0.032893

Table 3 GLCM texture feature of three year aged transformer oil

Sl. no.	GLCM texture features	Three year aged oil	Median filter	Weiner filter	NLM filter	OSF	CSF	ICSF
1	Homogeneity	0.044358	0.044312	0.044362	0.044357	0.044358	0.044359	0.044359
2	Contrast	0.009779	0.009889	0.009898	0.009876	0.009886	0.009897	0.009940
3	Correlation	6641.141	6639.534	6640.087	6641.272	6641.155	6641.112	6640.872
4	Energy	2.5011e-05	2.5014e-05	2.5012e-05	2.5011e-05	2.5012e-05	2.5012e-05	2.5012e-05
5	Entropy	0.030377	0.03303	0.03275	0.03279	0.03221	0.03310	0.03293

Table 4 GLCM texture feature of nine year aged transformer oil

Sl. no.	GLCM texture features	Nine year aged oil	Median filter	Weiner filter	NLM filter	OSF	CSF	ICSF
1	Homogeneity	0.044681	0.044629	0.044682	0.044677	0.044681	0.044678	0.044672
2	Contrast	0.021356	0.021506	0.021536	0.021527	0.021567	0.02153	0.021487
3	Correlation	6577.03	6575.391	6575.948	6577.072	6577.005	6577.014	6577.439
4	Energy	2.5040e-05	2.5042e-05	2.5040e-05	2.5039e-05	2.5041e-05	2.5040e-05	2.5040e-05
5	Entropy	0.040396	0.040532	0.040554	0.040171	0.040813	0.040638	0.040363

Table 5 GLCM texture feature of twelve year aged transformer oil

Sl. no.	GLCM texture features	Twelve year aged oil	Median filter	Weiner filter	NLM filter	OSF	CSF	ICSF
1	Homogeneity	0.04406	0.044018	0.044064	0.044067	0.044059	0.044061	0.044058
2	Contrast	0.0017644	0.0017355	0.0017482	0.0017374	0.0017337	0.0017376	0.0017680
3	Correlation	6685.081	6683.145	6683.824	6684.95	6685.089	6685.142	6686.108
4	Energy	2.5054e-05	2.5055e-05	2.5053e-05	2.5050e-05	2.5056e-05	2.5054e-05	2.5053e-05
5	Entropy	0.0454238	0.04236364	0.04232632	0.04163234	0.04298412	0.04252854	0.04217208

Table 6 GLCM texture feature of fourteen year aged transformer oil

Sl. no.	GLCM texture features	Fourteen year aged oil	Median filter	Weiner filter	NLM filter	OSF	CSF	ICSF
1	Homogeneity	0.044019	0.043977	0.04402	0.044011	0.04402	0.04402	0.044017
2	Contrast	0.042918	0.046128	0.046108	0.045342	0.044197	0.043511	0.043123
3	Correlation	6650.85	6648.855	6649.61447	6651.146	6650.932	6650.79	6650.875
4	Energy	2.5038e-05	2.5040e-05	2.5038e-05	2.5035e-05	2.5040e-05	2.5038e-05	2.5037e-05
5	Entropy	0.040963	0.0403396	0.0399231	0.03906750	0.040713	0.0401358	0.0408131

Table 7 GLCM texture feature of sixteen year aged transformer oil

Sl. no.	GLCM texture features	Sixteen year aged oil	Median filter	Weiner filter	NLM filter	OSF	CSF	ICSF
1	Homogeneity	0.044335	0.044284	0.044339	0.044341	0.044336	0.0443349	0.044328
2	Contrast	0.003416	0.003876	0.0039112	0.003523	0.003527	0.003562	0.003499
3	Correlation	6717.505	6715.60919	6716.114	6716.422	6717.292	6717.465	6717.694
4	Energy	2.5064e-05	2.5065e-05	2.5063e-05	2.5061e-05	2.5064e-05	2.5063e-05	2.5063e-05
5	Entropy	0.042173	0.038943	0.038950	0.038534	0.039153	0.039080	0.038789

Table 8 GLCM texture feature of seventeen year aged transformer oil

Sl. no.	GLCM texture features	Seventeen year aged oil	Median filter	Weiner filter	NLM filter	OSF	CSF	ICSF
1	Homogeneity	0.044254	0.044211	0.044259	0.044257	0.044254	0.044255	0.044256
2	Contrast	0.0029451	0.0028845	0.0028951	0.0029158	0.0029132	0.0029045	0.0028647
3	Correlation	6726.811	6724.716	6725.372	6726.442	6726.555	6726.699	6726.601
4	Energy	2.5072e-05	2.5074e-05	2.5071e-05	2.5070e-05	2.5073e-05	2.5072e-05	2.5071e-05
5	Entropy	0.0572501	0.058856	0.056973	0.059775	0.057646	0.057640	0.057964

Table 9 GLCM texture feature of eighteen year aged transformer oil

Sl. no.	GLCM texture features	Eighteen year aged oil	Median filter	Weiner filter	NLM filter	OSF	CSF	ICSF
1	Homogeneity	0.044254	0.044205	0.044255	0.044244	0.044248	0.044251	0.0442439
2	Contrast	0.00451373	0.00440081	0.0043950	0.00435942	0.00444906	0.00444783	0.00456736
3	Correlation	6695.149	6692.47	6693.741	6694.723	6694.956	6694.821	6693.9475
4	Energy	2.5106e-05	2.5107e-05	2.5104e-05	2.5098e-05	2.5110e-05	2.5105e-05	2.5104e-05
5	Entropy	0.031476	0.037351	0.037180	0.031537	0.034273	0.033866	0.036911

Table 10 GLCM texture feature of twenty-one year aged transformer oil

Sl. no.	GLCM texture features	Twenty-one year aged oil	Median filter	Weiner filter	NLM filter	OSF	CSF	ICSF
1	Homogeneity	0.043937	0.043898	0.043944	0.04394	0.043976	0.043945	0.0439292
2	Contrast	0.0034101	0.00352813	0.00351735	0.00352435	0.00399516	0.00365297	0.00347477
3	Correlation	6690.185	6688.722	6689.888	6691.8396	6691.928	6691.816	6694.1191
4	Energy	2.5099e-05	2.5078e-05	2.5075e-05	2.5081e-05	2.5091e-05	2.5093e-05	2.5095e-05
5	Entropy	0.0460504	0.045281	0.045173	0.045243	0.045995	0.045297	0.045747

Table 11 GLCM texture feature of twenty-five year aged transformer oil

Sl. no.	GLCM texture features	Twenty-five year aged oil	Median filter	Weiner filter	NLM filter	OSF	CSF	ICSF
1	Homogeneity	0.044101	0.04406	0.044098	0.0440887	0.044126	0.044093	0.0440867
2	Contrast	0.001323	0.0014051	0.001451	0.001395	0.001381	0.0013593	0.0013472
3	Correlation	6671.649	6666.849	6670.508	6673.2135	6675.025	6674.199	6678.1423
4	Energy	2.5417e-05	2.5340e-05	2.5337e-05	2.5203e-05	2.5583e-05	2.5395e-05	2.5332e-05
5	Entropy	0.0396753	0.039252	0.039242	0.038642	0.039920	0.039326	0.039013

References

1. J. J. Winders, Jr., *Power Transformers: Principles and Applications* (Marcel Dekker, Inc. New York, 2002)
2. Waldemar Ziomek, Transformer electrical insulation. *IEEE Trans. Dielectrics Electr. Insul.* **19**(6), 1841–1842 (2012)
3. Bureau of Reclamation Engineers, *Transformer Basics, Maintenance and Diagnostics* (Hydroelectric Research and Technical Services Group Denver, Colorado, 2005)
4. T.O. Rose, Mineral insulating oils in transformers. *IEEE Electr. Insul. Mag.* **14**, 6–16 (1998)
5. C.M. Maheshan, H. Prasanna Kumar, A review of image processing applications in fault detection of transformer oil. *Int. J. Res. Advent Technol. ICTTSTM*, pp. 433–437 (2018)
6. R.C. Gonzalez, R.E. Woods, *Digital Image Processing* (Addison-Wesley Publishing Company, 2002)
7. M. Sonka, V. Hlavac, R. Boyle, *Image processing, analysis and machine vision*, 3rd edn. (Cengage Learning, Boston, 2013)
8. J. Chen et al., New insights into the noise reduction Wiener filter. *IEEE Trans. Audio Speech Lang. Process.* **14**, 1218–1234 (2006)
9. A. Buades, B. Coll, J.M. Morel, A non-local algorithm for image denoising, in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 2, pp. 60–65 (2005)
10. S. Osher, L.I. Rudin, Feature-oriented image enhancement using shock filters. *SIAM J. Numer. Anal.* **27**, 919–940 (1990)
11. C.M. Maheshan, H.P. Kumar, Image enhancement of transformer oil images using improved complex shock filter. *IETA Instrumentation Measure Metrologie.* **18**(4), 407–412 (2019)
12. C.M. Maheshan, H.P. Kumar, Performance of image pre-processing filters for noise removal in transformer oil images at different temperatures. *SN Appl. Sci.* **2**, 67 (2020)
13. M.H. Sidram, S. Hegde, A novel approach of transformer oil quality analysis using image processing, in *International Conference on Advances in Engineering & Technology (ICAET)*, pp 52–57 (2014)
14. R.M. Haralick, Statistical and structural approaches to texture. *Proc. IEEE* **67**(5), 786–804 (1979)
15. P.S. Shijin Kumar, V.S. Dhruvan, Extraction of texture features using GLCM and shape features using connected regions. *Int. J. Eng. Technol.* **8**(6), 2926–2930 (2016)

A Comparative Analysis of Statistical Model and Spectral Subtractive Speech Enhancement Algorithms



K. Rohith and Rekha Bhandarkar

1 Introduction

According to the World Health Organization, an estimated 360 million people across the globe suffer from varying types of hearing disabilities. In India, around 63 million are approximated to suffer from hearing disabilities [1]. Hearing impairments can be classified into partial and severe hearing disabilities. Hearing Aids and Cochlear Implants help these individuals cope up with their disabilities. The Hearing Aids help the people with partial hearing impairments, whereas the Cochlear Implants are particularly useful for those suffering from severe hearing impairment or almost complete deafness [2].

The Hearing Aids predominantly focus on amplification of the sound signals before being delivered to the inner ear. The Cochlear Implants help in restoring the sense of hearing to the individuals with profound deafness by bypassing the inner ear and directly stimulating the auditory nerves connected to the brain. However, many of the Cochlear Implant recipients complain that the speech sounds heard are unnatural and pitch perception is generally poor particularly in outdoor noisy environments. The presence of the noises in the environment namely stationary and non-stationary noises poses a massive challenge for the individuals with both partial and profound hearing disabilities in understanding the speech. Thus, the speech enhancement algorithms must be capable of improving the perceptual aspect of the speech degraded by the noises. The key challenge in the design of an efficient speech signal enhancement algorithm is in nullifying the background noise content in the

K. Rohith (✉) · R. Bhandarkar
Department of ECE, NMAM Institute of Technology, Nitte (Visvesvaraya Technological University, Belagavi), Udipi, India
e-mail: rohith.k@vnce.ac.in

R. Bhandarkar
e-mail: rekhabhandarkar@nitte.edu.in

noisy speech without inducing any perceptible distortion in the processed signal [3, 4].

Understanding of the different kind of noises, differences between their temporal domain and spectral domain features and range of noise magnitudes help us in designing efficient speech enhancement algorithms. The two categories of noises are stationary noise and non-stationary noise. The stationary noises do not significantly alter their characteristics with time such as the fan noise in a living area. The non-stationary noises such as restaurant noise with many people talking simultaneously at a time in the background along with many other additive noises constantly change its characteristics with respect to time. Thus, suppressing non-stationary noise that is continuously varying is furthermore challenging than the task of nullifying the stationary noise [5]. Figures given below decompose the structure and distribution of the noise signal. A unique feature of the different type of noises is in their shape of spectrum, and it corresponds to the noise energy spread in the spectral domain. For instance, the major energy content of the wind noise is condensed in the lower band of frequencies, normally fewer than 500 Hz. Another non-stationary restaurant noise holds a broader range of frequencies in its frequency domain spectrum. Figure 1 shows the time domain spectrums of the car, train and restaurant noise taken from the NOIZEUS corpus.

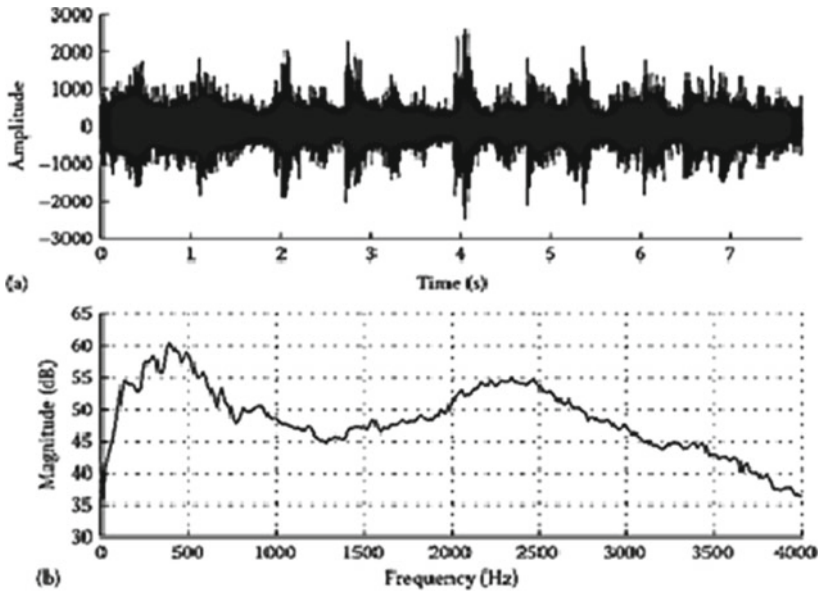
2 Speech Enhancement Algorithms

Speech signal enhancement algorithms can be classified into three major classes namely statistical-model-based algorithms, subspace algorithms and spectral subtractive algorithms [6, 7].

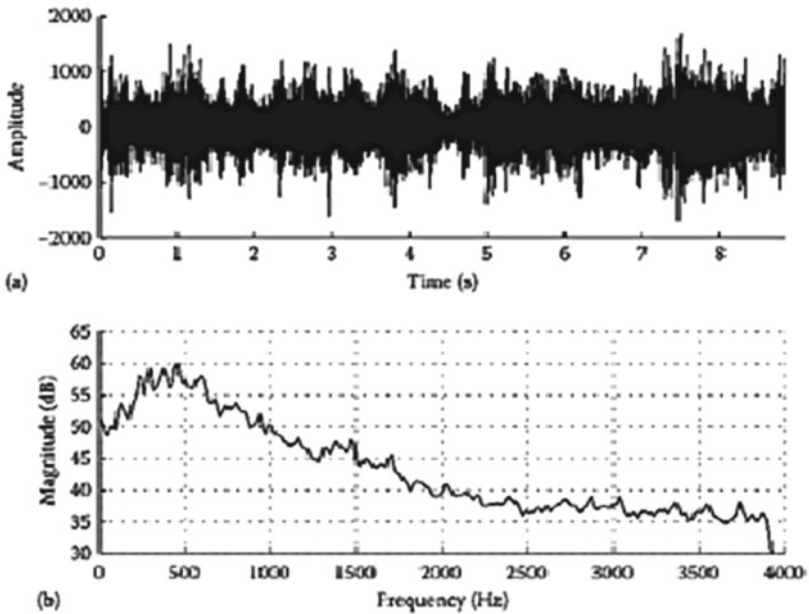
The spectral subtraction algorithms are the simplest methods to implement and are built on the estimate of noise signal spectrum and the subsequent subtraction of noise content signal from the noisy speech. Statistical model algorithms consider the statistical estimation framework with the given set of datasets, corresponding spectral domain coefficients of the noisy speech signal to compute the transformation coefficients of the noise free enhanced signal, whereas the subspace algorithms are constructed on the linear algebraic principles that the enhanced speech signal may be narrowed down to a noisy sub-space of the Euclidean space [8].

2.1 *Spectrum Subtraction Algorithm*

It is based on the principle that considering the extra noise content, we are able to predict the clean speech signal spectrum by subtracting an approximate of noise signal spectrum from the noisy speech spectrum. Noise signal spectrum is approximated and updated during periods, when the speech signal is absent and during intervals where speech is paused. The presumption made is that noise is initially stationary in



(a) Example noise from a train; (b) its long-term average spectrum



(a) Example noise from a restaurant; (b) its long-term average spectrum

Fig. 1 Noise spectrums of car, train and restaurant noises [5]

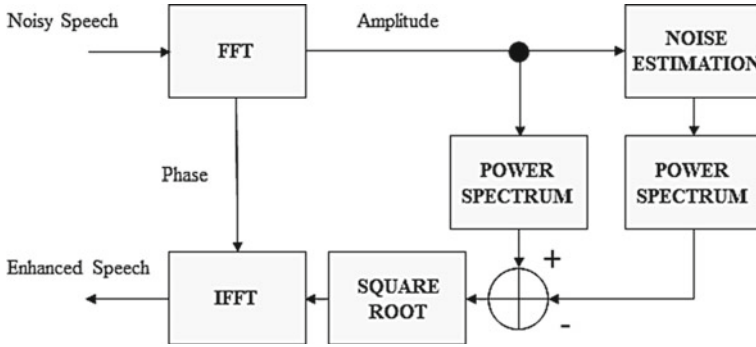


Fig. 2 Spectral subtraction speech enhancement method

nature and that the noise spectrum do not change drastically between the updating periods [9]. The enhanced speech signal is obtained by computing the ID-FT of the approximated enhanced signal spectrum using the phase component of the noisy signal. The spectral subtraction technique is a straightforward one as it involves only a Fourier transformation and an inverse Fourier transformation (Fig. 2).

The noisy input signal $z(n)$ is considered to be composed of additive noise signal $v(n)$ and the original clean speech signal $x(n)$.

$$z(n) = x(n) + v(n) \tag{1}$$

Taking DT-FT on both sides and expressing it in the polar-form we get:

$$Z(\omega) = X(\omega) + V(\omega) \tag{2}$$

$$Z(\omega) = |Z(\omega)| e^{j\phi_y(\omega)} \tag{3}$$

where $|Z(\omega)|$ is the magnitude component of $Z(\omega)$, $\phi_y(\omega)$ is phase component of the noise corrupted speech signal $Z(\omega)$.

The noise signal spectrum $V(\omega)$ is also given in terms of its magnitude and phase spectra as:

$$V(\omega) = |V(\omega)| e^{j\phi_d(\omega)} \tag{4}$$

Estimated clean speech signal spectrum is obtained using the equation:

$$\hat{X}(\omega) = \left[|Z(\omega)| - |\hat{V}(\omega)| \right] e^{j\phi_y(\omega)} \tag{5}$$

The spectral subtraction process must be carried out thoughtfully to avoid any original speech signal distortion [10]. If the algorithm subtracts too much, then some

actual speech content might be removed, whereas if the algorithm subtracts too little, then enhanced signal may contain much of the interfering noise.

2.2 Wiener Filter Algorithm

An LTI system gives an output signal $y(n)$ by accepting the input signal $x(n)$. The designed system output signal $\hat{z}(n)$, is as close approximation to the desired signal $z(n)$. This is accomplished by estimating the error $e(n)$ and reducing it to as small a value as possible. Wiener filter is an optimal filter that reduces the estimation error [11]

In speech enhancement applications [5], the $y(n)$ is the input signal to the Wiener filter as shown in Fig. 3 which is combination of original speech signal and noise content.

$$y(n) = x(n) + v(n) \tag{6}$$

where $v(n)$ is the noise component and $x(n)$ is the original clean speech. The desired signal $z(n)$ in Fig. 3 is the clean and enhanced version of $x(n)$, i.e., $z(n) \approx x(n)$. Wiener filter focuses on generating the clean signal $x(n)$ estimation. We get the Wiener filter transfer function as:

$$H(\omega k) = P_{xx}(\omega k) / [P_{xx}(\omega k) + P_{vv}(\omega k)] \tag{7}$$

The Wiener filter focuses on the portions of the noisy speech signal spectrum where the signal-to-noise ratio is high and allows it and attenuates the portions of the noisy speech signal spectrum where the signal-to-noise ratio is low, i.e., attenuation amount applied by the Wiener filter is proportional to the SNR of each frequency component.

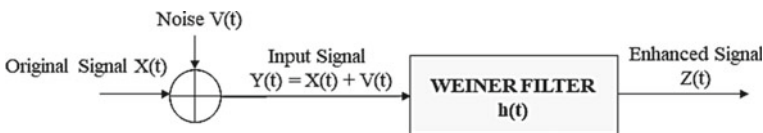


Fig. 3 Wiener filter method

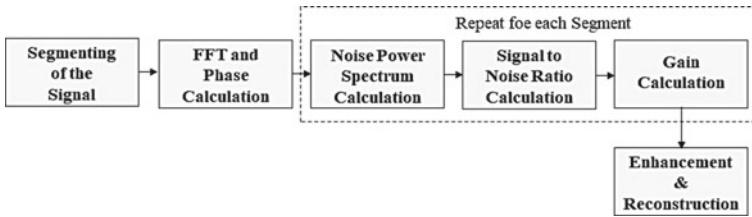


Fig. 4 Short-time spectral amplitude MMSE method

2.3 Short-Time Spectral Amplitude MMSE Algorithm

The STSA-MMSE method focuses on finding an average observatory probability density function of the enhanced speech signal spectrum amplitude from the given the probability density function of the original speech and noise signals [12] (Fig. 4).

Using noise power spectrum, the priori signal-to-noise ratio (ξ_k) and the posteriori signal-to-noise ratio (γ_k) are calculated, the gain (G) of the signal is then updated using the signal-to-noise ratios and the signal is then cleaned by combining the FFT of each of the frames, the clean signal is reconstructed using the overlap and add method utilizing the phase of the FFT. The performance of the STSA-MMSE technique for noise reduction depends on the estimation of the priori signal-to-noise ratio [13].

Furthermore, the priori signal-to-noise ratio approximation relies on the noise estimation and speech signal variance. Conventionally, noise component variance is estimated during the paused portion of the speech presuming that the noise considered is stationary in nature such as fan noise in a living room, etc. Better estimation of the noises helps in further improving the performance of STSA-MMSE method.

NOIZEUS is a noisy dataset recorded at Dallas University of Texas at (UTD) used extensively for the analysis of speech signal enhancement algorithms performance among the research groups. The NOIZEUS corpus has 30 audio files generated by three female speakers and three male speakers and is distorted by various physical world noises at 15, 10, 5 and 0 dB SNRs. The noise content in the NOIZEUS was taken from the AURORA corpus which contains airport, train-station noise, suburban train noise, multi-speaker babble noise, car, restaurant, exhibition hall and street noises, etc Tables 1 and 2 give the list of sentences in the audio files recorded in the NOIZEUS corpus [14] (Fig. 5).

3 Quality Measurements

Speech quality evaluation can be done using signal-to-noise ratio for the implemented speech enhancement algorithms. A well-known standard SNR definition is given by the [7] Global SNR (GSNR) as:

Table 1 List of sentences used in NOIZEUS

Filename	Speaker	Gender	Sentence text
sp01.wav	CH	M	The birch canoe slid on the smooth planks
sp02.wav	CH	M	He knew the skill of the great young actress
sp03.wav	CH	M	Her purse was full of useless trash
sp04.wav	CH	M	Read verse out Loud for pleasure
sp05.wav	CH	M	Wipe the grease off his dirty face
sp06.wav	DE	M	Men strive but seldom get rich
sp07.wav	DE	M	We find joy in the simplest things
sp08.wav	DE	M	Hedge apples may stain your hands green
sp09.wav	DE	M	Hurdle the pit with the aid of a long pole
sp10.wav	DE	M	The sky that morning was clear and bright blue
sp11.wav	JE	F	He wrote down a long list of items
sp12.wav	JE	F	The drip of the rain made a pleasant sound
sp13.wav	JE	F	Smoke poured out of every crack
sp14.wav	JE	F	Hats are worn to tea and not to dinner
sp15.wav	JE	F	The clothes dried on a thin wooden rack

Table 2 List of sentences used in NOIZEUS

Filename	Speaker	Gender	Sentence text
sp16.wav	KI	F	The stray cat gave birth to kittens
sp17.wav	KI	F	The lazy cow Lay in the cool grass
sp18.wav	KI	F	The friendly gang left the drug store
sp19.wav	KI	F	We talked of the sideshow in the circus
sp20.wav	KI	F	The set of china hit the floor with a crash
sp21.wav	SI	M	Clams are small, round, soft and tasty
sp22.wav	SI	M	The line where the edges join was clean
sp23.wav	SI	M	Stop whistling and watch the boys march
sp24.wav	SI	M	A cruise in warm waters in a sleek yacht is fun
sp25.wav	SI	M	A good book informs of what we ought to know
sp26.wav	TI	F	She has a smart way of wearing clothes
sp27.wav	TI	F	Bring your best compass to the third class
sp28.wav	TI	F	The club rented the rink for the fifth night
sp29.wav	TI	F	The flint sputtered and lit a pine torch
sp30.wav	TI	F	Let's all join as we sing the last chorus

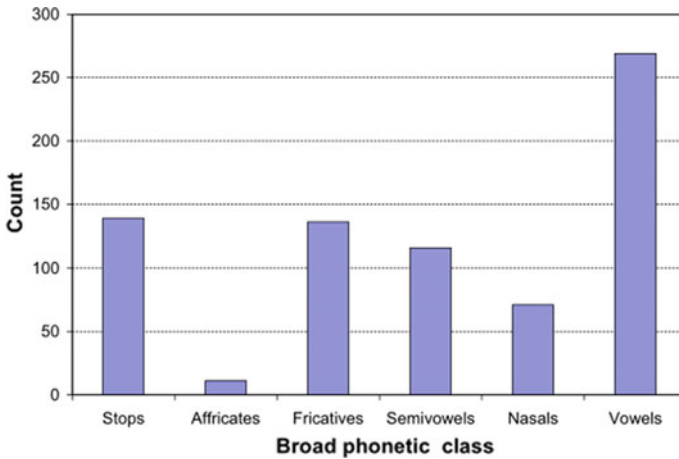


Fig. 5 Phonetic classes distribution in the NOIZEUS corpus

$$G\text{SNR} = [10 \log_{10}(\sigma_s^2)/(\sigma_n^2)] \tag{8}$$

where σ_s^2 is the related to the speech signal power and σ_n^2 corresponds to the noise power. It focuses on the evaluation of signal-to-noise ratio through the speech activity portion of the enhanced analyzed signal [15–17] (Figs. 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18 and Table 3).

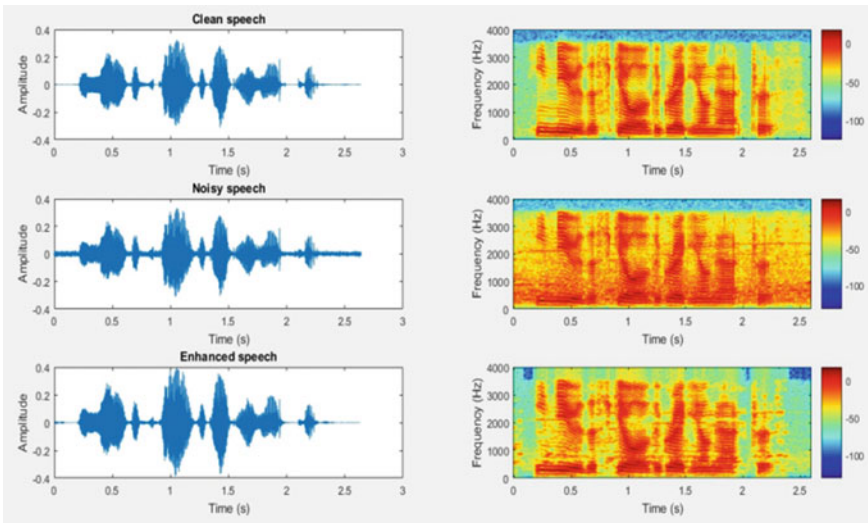


Fig. 6 Speech enhancement algorithm—time and frequency spectrums in MATLAB—using spectral subtraction [15 dB airport noise]

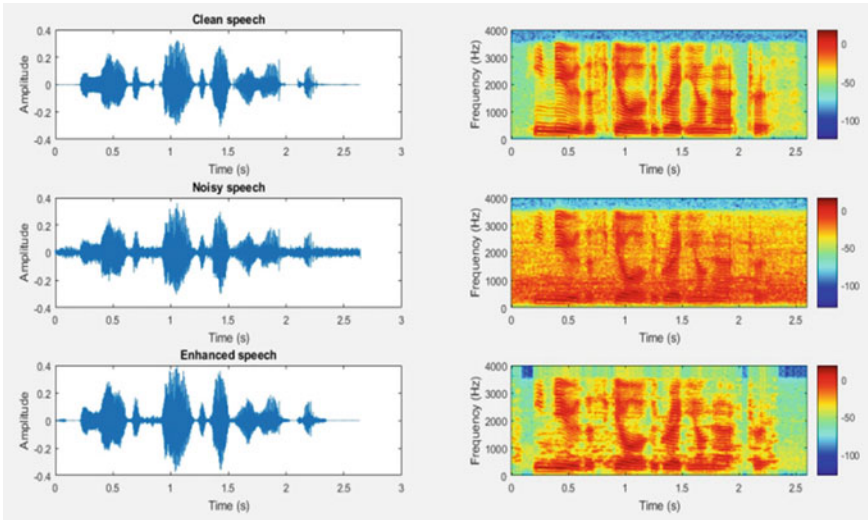


Fig. 7 Speech enhancement algorithm—time and frequency spectrums in MATLAB—using spectral subtraction [10 dB airport noise]

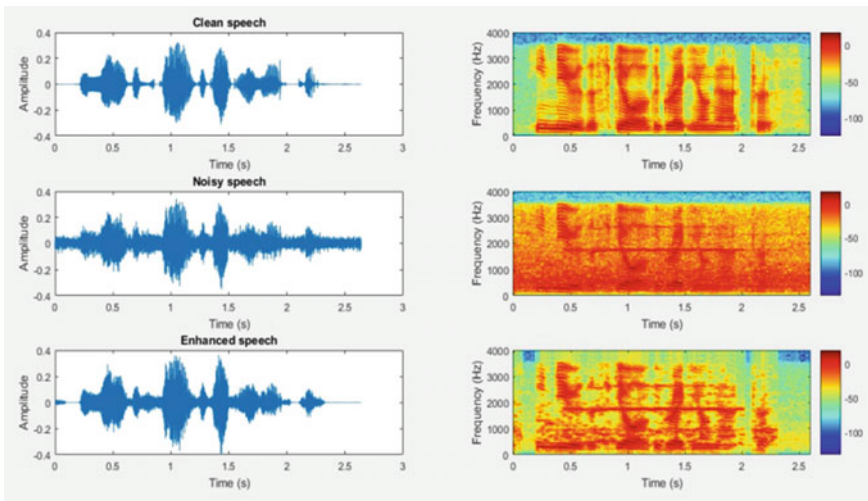


Fig. 8 Speech enhancement algorithm—time and frequency spectrums in MATLAB—using spectral subtraction [5 dB Airport noise]

4 Results and Discussion

The spectral subtraction method showed better performance as compared to Wiener filter technique and short-time spectral amplitude MMSE technique under the noisy

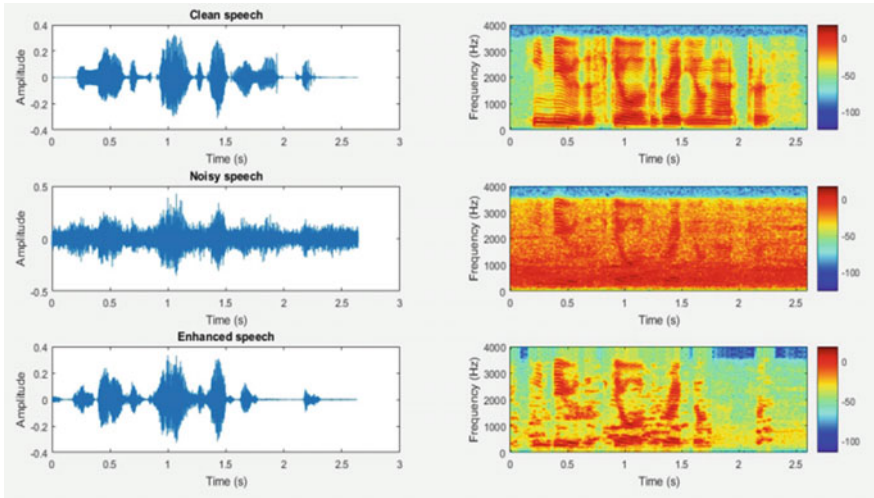


Fig. 9 Speech enhancement algorithm—time and frequency spectrums in MATLAB—using spectral subtraction [0 dB Airport noise]

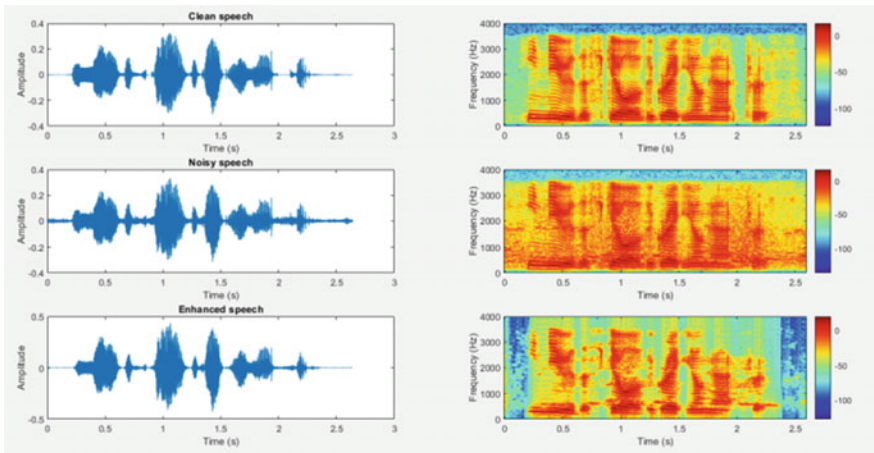


Fig. 10 Time and frequency spectrums in MATLAB—speech enhancement using Wiener filter [For restaurant noise of 15 dB SNR]

environments considered. The short-time spectral MMSE method showed second best performance followed by Wiener filter method.

The SNR used for the evaluation of these algorithms using NOIZEUS speech database for performance testing of speech enhancement algorithms. Furthermore, Speech Intelligibility Evaluation and Perceptual Quality Analysis can be carried out for the evaluating of these algorithms.

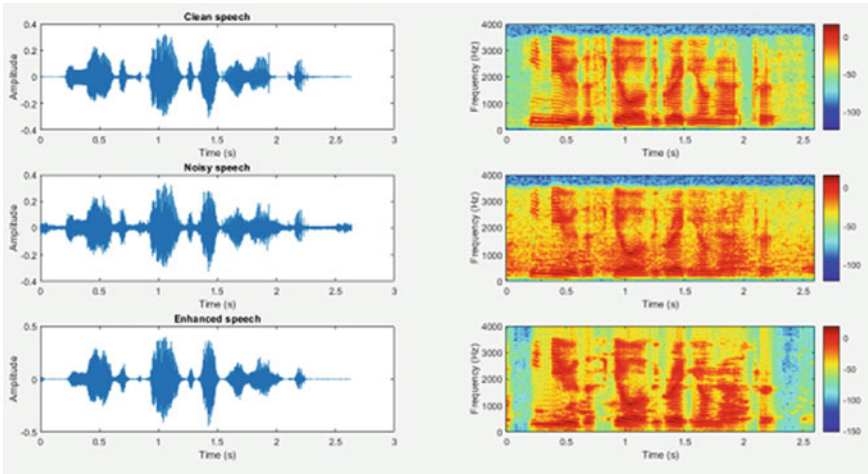


Fig. 11 Time and frequency spectrums in MATLAB—speech enhancement using Wiener filter [For restaurant noise of 10 dB SNR]

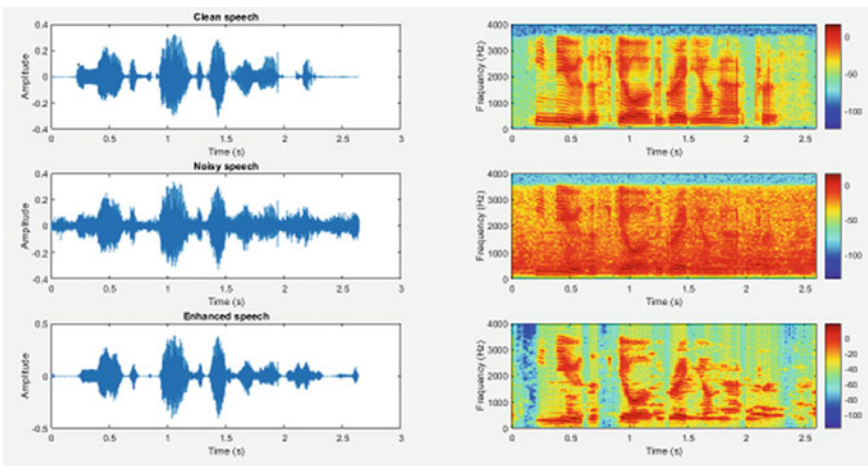


Fig. 12 Time and frequency spectrums in MATLAB—speech enhancement using Wiener filter [For restaurant noise of 05 dB SNR]

PRAAT is a useful speech signal analysis tool [18, 19]. It can be used for spectrographic analysis, articulatory synthesis, manipulating and neural networks. Figure 19 shows the temporal spectrum of the clean speech sample (sp02.wav) and its corresponding spectrogram. Spectrogram gives the optical representation of the sound signals in a space with time in the horizontal dimension and frequency in the vertical dimension. The formant frequencies of the speech signal can be seen in the Spectrogram highlighted by red coloured line with formant frequency f1 around 310 Hz

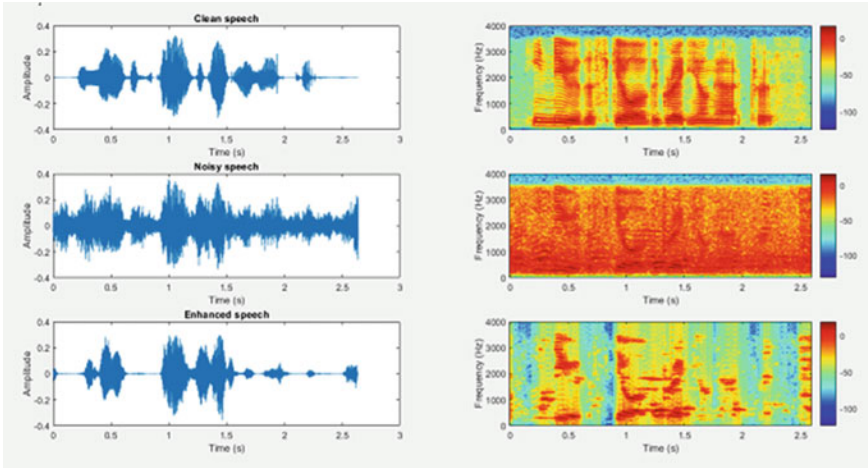


Fig. 13 Time and frequency spectrums in MATLAB—speech enhancement using Wiener filter [For restaurant noise of 0 dB SNR]

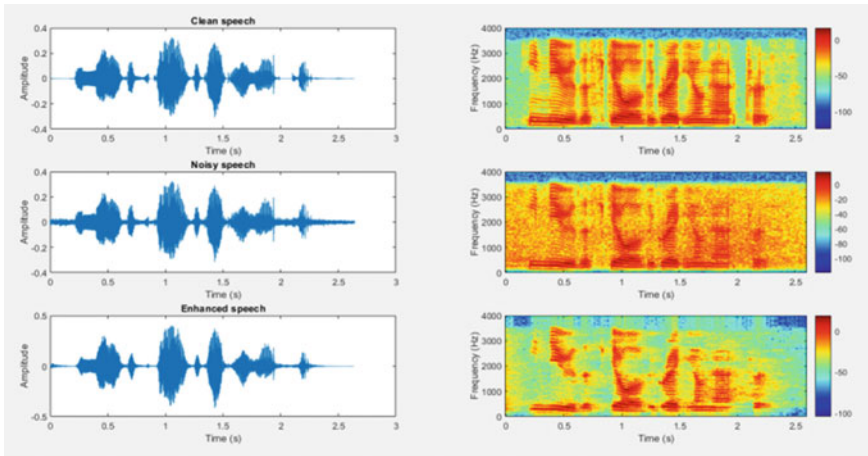


Fig. 14 Time and frequency spectrums in MATLAB—speech enhancement using STSA MMSE [For street noise of 15 dB SNR]

and formant frequency f_5 around 3.3 kHz. Yellow coloured lines indicate the signal intensity variation of the speech signal considered (Fig. 20).

Cochleogram relates to the properties of the cochlea and is a time frequency representation mimicking the components of the cochlea in the human hearing system [20, 21].

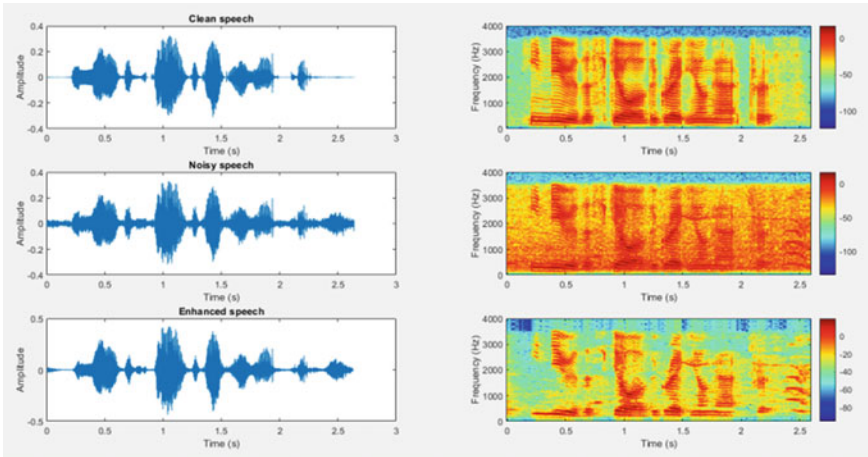


Fig. 15 Time and frequency spectrums in MATLAB—speech enhancement using STSA MMSE [For street noise of 10 dB SNR]

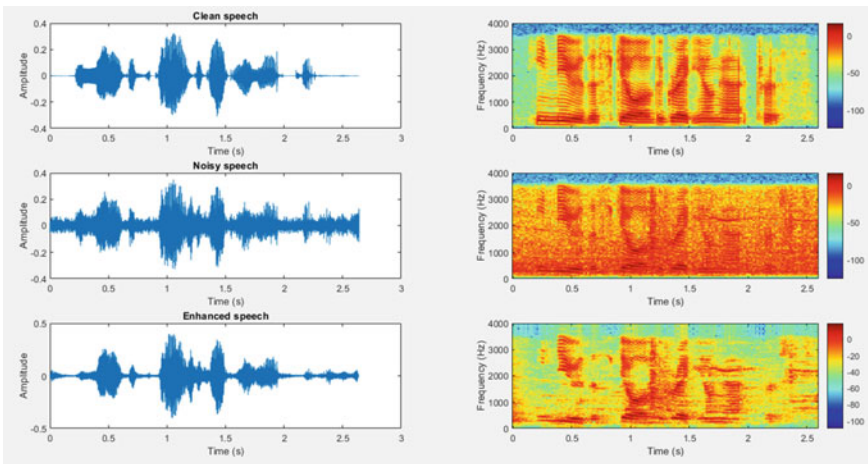


Fig. 16 Time and frequency spectrums in MATLAB—speech enhancement using STSA MMSE [For street noise of 05 dB SNR]

Figure 21 shows the spectrogram of airport noise with 15 dB SNR. As it is seen from the diagram, the noise occupies a larger bandwidth and its energy spectrum is spread throughout the entire length of the signal.

Figure 22 shows the spectrogram of the noisy signal with 15dB airport noise. As it is seen from the diagram, the noise has combined with the speech signal and it occupies a larger bandwidth and its energy spectrum is spread throughout the entire

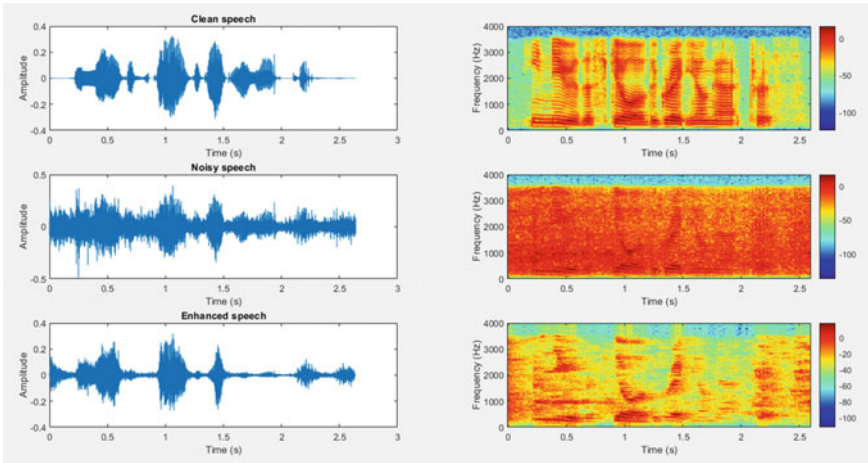


Fig. 17 Time and frequency spectrums in MATLAB—speech enhancement using STSA MMSE [For street noise of 0 dB SNR]

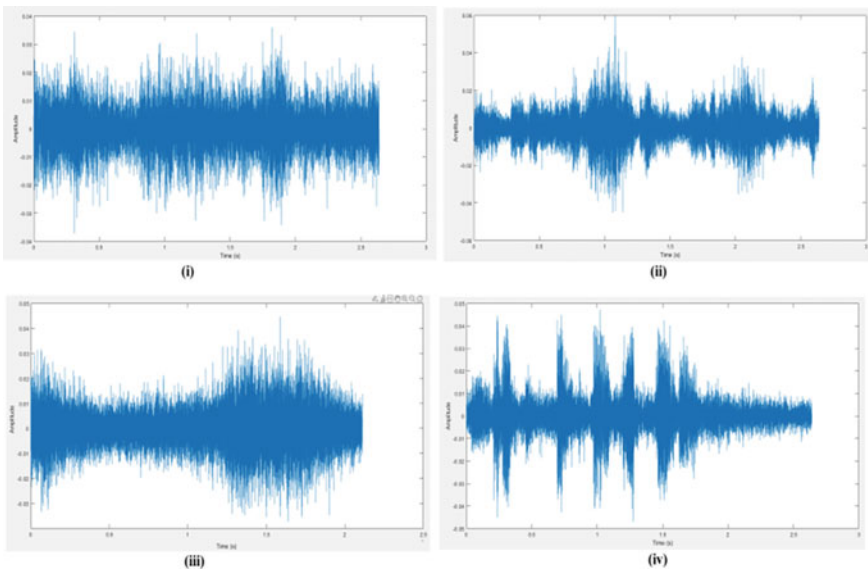


Fig. 18 Time domain spectrums of various noises—(i) airport noise_15 dB; (ii) restaurant noise_15 dB; (iii) street noise_15 dB; (iv) train station noise_15 dB

length of the signal. Figure 23 shows the spectrogram of speech enhancement using spectral subtraction for airport noise of 15dB SNR. The noise components are greatly suppressed as seen from the diagram.

Table 3 Performance evaluation of algorithms using SNR

For input SNR of 15 dB			
Noise	Output SNR		
	WF	SS	STSA
Airport noise	15.71	17.33	16.23
Restaurant noise	15.53	16.25	15.66
Street noise	15.15	15.82	15.53
Train Station	15.36	17.10	15.68
For input SNR of 10 dB			
Noise	Output SNR		
	WF	SS	STSA
Airport noise	11.16	12.19	11.00
Restaurant noise	11.85	12.42	11.74
Street noise	10.83	11.28	11.01
Train Station	8.77	11.53	10.07
For input SNR of 5 dB			
Noise	Output SNR		
	WF	SS	STSA
Airport noise	5.38	7.66	6.30
Restaurant noise	6.26	6.88	6.23
Street noise	5.72	6.35	6.25
Train Station	6.78	8.37	8.06
For input SNR of 0 dB			
Noise	Output SNR		
	WF	SS	STSA
Airport noise	3.46	4.12	3.81
Restaurant noise	2.28	2.78	2.10
Street noise	2.89	3.29	3.54
Train Station	2.91	3.35	2.91

5 Conclusion and Future Work

Three speech signal enhancement algorithms such as short-time spectral amplitude MMSE algorithm, Weiner filter algorithm and spectral subtraction algorithm are implemented and tested using NOIZEUS corpus for SNRs at 15, 10, 5 and 0 dB for various types of noises which are predominantly non-stationary noises. The frequency domain, time domain spectrums, spectrogram of noisy speech and noise reduced speech were extracted to visualize and evaluate the algorithms.

The spectral subtraction method showed better performance as compared to Wiener filter technique and short-time spectral amplitude MMSE technique under

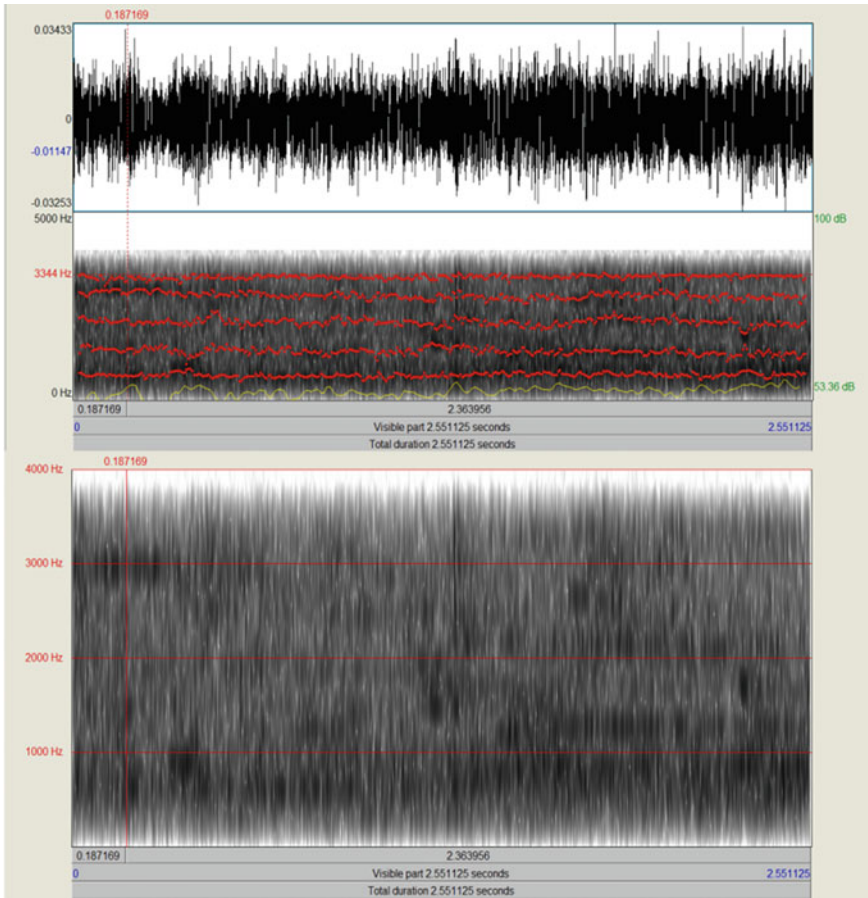


Fig. 21 Spectrogram—noise signal—airport noise of 15 dB SNR

the noisy environments considered. The short-time spectral MMSE method showed second best performance followed by Wiener filter method. To further analyze the performance of these speech enhancement algorithms, speech intelligibility and perceptual evaluation can be done as the future perspective.

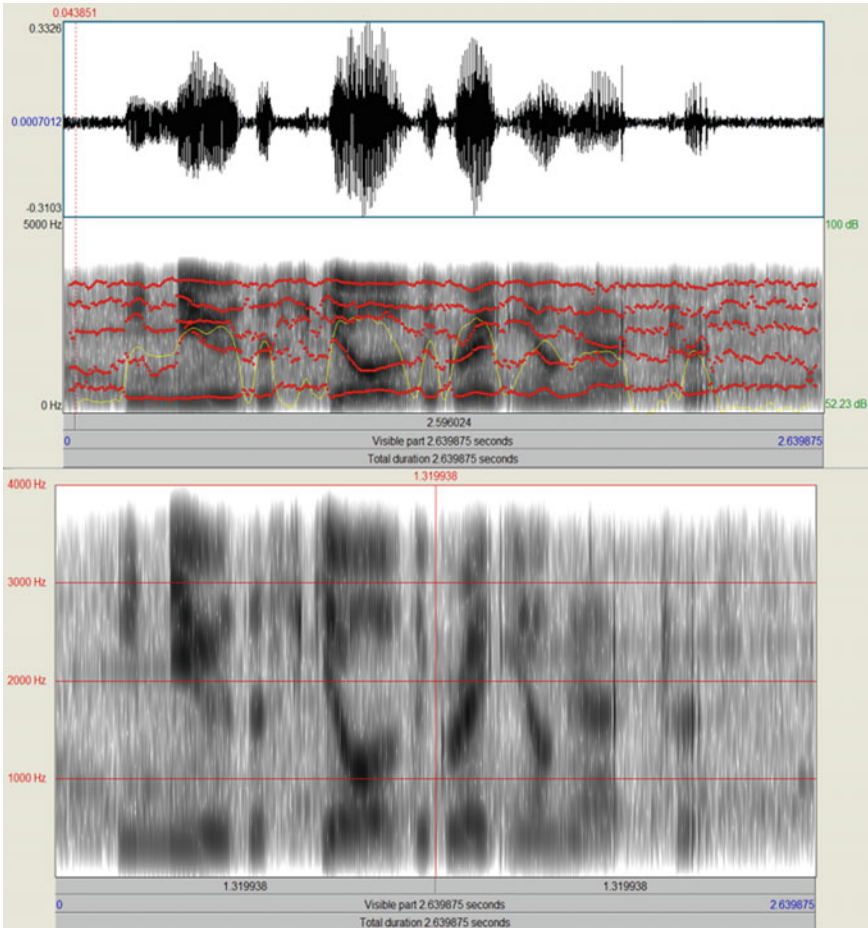


Fig. 22 Spectrogram—noisy speech [sp02 with airport noise of 15 dB SNR]

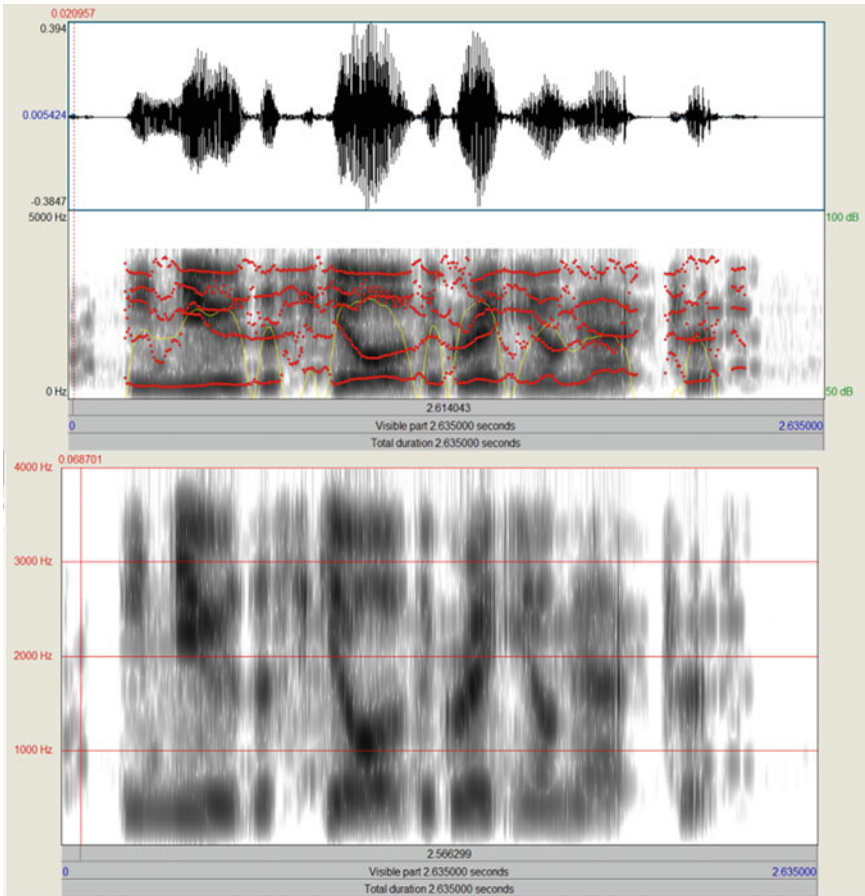


Fig. 23 Spectrogram—speech Enhancement using spectral subtraction [For airport noise of 15 dB SNR]

References

1. S. Varshney, Deafness in India. *Indian J. Otol.* **22**(2) (2016). Published by Wolters Kluwer–Medknow
2. P.C. Loizon, Mimicking the human ear. *IEEE Sig. Process. Mag.* 101-130 (1988). Department of Electrical Engineering, University of Texas at Dallas, USA
3. D. Wang, J.H.L. Hansen, Speech enhancement for cochlear implant recipients. *J. Acoust. Soc. Am.* **143**, 2244 (2018)
4. J.T. Rubinstein, How cochlear implants encode speech. *Curr. Opinion. Otolaryngol. Head Neck Surg.* **12**(5), 444–448 (2004). <https://doi.org/10.1097/01.moo.0000134452.24819.c0>
5. P.C. Loizou, *Speech Enhancement Theory and Practice*, 2nd ed (CRC Press Taylor & Francis Group, 2013)
6. N. Upadhyay, A. Karmakar, The spectral subtractive-type algorithms for enhancing speech in noisy environments, in *Presented at the 2012 1st International Conference on Recent Advances in Information Technology (RAIT)*, March 2012. <https://doi.org/10.1109/rait.2012.6194534>

7. S. Rangachari, P.C. Loizou, A noise-estimation algorithm for highly non-stationary environments. *Speech Commun.* **48**(2), 220–231 (2006). <https://doi.org/10.1016/j.specom.2005.08.005>
8. P.C. Loizou, A. Lobo, Y. Hu, Subspace algorithms for noise reduction in cochlear implants. *J. Acoust. Soc. Am.* (2005)
9. M. Krawczyk-Becker, D. Fischer, T. Gerkmann, Utilizing spectro-temporal correlations for an improved speech presence probability based noise power estimation, in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, April 2015. <https://doi.org/10.1109/icassp.2015.7177992>
10. Y. Hu, P.C. Loizou, Subjective comparison and evaluation of speech enhancement algorithms. *Speech Commun.* (2007)
11. Y. Hu, P. Loizou, Subjective comparison of speech enhancement algorithms, in *Proceedings of IEEE International Conference on Acoustic, Speech, Signal Processing*, vol. I, Toulouse, France (2006), pp. 153–156
12. J. Hovorka, Methods for evaluation of speech enhancement algorithms. *Adv. Military Technol.* **4**(2) (2009)
13. R. Yao, Z. Zeng, P. Zhu, A priori SNR estimation and noise estimation for speech enhancement. *EURASIP J. Adv. Sig. Process.* (2016). <https://doi.org/10.1186/s13634-016-0398-z>
14. IEEE Subcommittee, IEEE recommended practice for speech quality measurements. *IEEE Trans. Audio Electroacoust.* **AU-17**(3), 225–246 (1969)
15. V. Martin, P. Petr, Methods for speech SNR estimation: evaluation tool and analysis of VAD dependency. *Radio Eng.* **14**(1) (2005)
16. A. Pandey, V.J. Mathews, Low-delay signal processing for digital hearing aids. *IEEE Trans. Audio Speech Lang. Process.* **19**(4) (2011)
17. P. Papadopoulos, A. Tsiartas, J. Gibson, S. Narayanan, A supervised signal-to-noise ratio estimation of speech signals, in *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Florence (2014), pp. 8237–8241. <https://doi.org/10.1109/icassp.2014.6855207>
18. Ambalika, Er.S. Saini, Speech analysis in Praat tool using hybrid filter. *Int. J. Adv. Res. Electron. Commun. Eng. (IJARECE)* **5**(10) (2016). ISSN 2278-909X
19. B. Shishir et al., GUI based performance analysis of speech enhancement techniques. *Int. J. Sci. Res. Publ.* **3**(9) (2013)
20. H. Chaurasiya, Time-frequency representations: spectrogram, cochleogram and correlogram. *Procedia Comput. Sci.* **167**, 1901–1910 (2020). ISSN 1877-0509. <https://doi.org/10.1016/j.procs.2020.03.209>
21. A. Tjandra, S. Sakti, G. Neubig, T. Toda, M. Adriani, S. Nakamura, Combination of two-dimensional cochleogram and spectrogram features for deep learning-based ASR, in *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Brisbane, QLD (2015), pp. 4525–4529. <https://doi.org/10.1109/icassp.2015.7178827>

Design of Electronic Instrumentation for Isotope Processing



V. Arunprasath

1 Background, Motivation and Objective

A new isotope processing facility (IPF) is coming up in BARC. It is an upgraded and revamped facility with a purpose to augment the existing capacity to process iodine isotope samples to 100 Curie. Iodine-131 has become one of the most important isotopes applied in medicine and industry [1, 2]. The facility is planned to cater to all such needs of the society. The main objective of the design of this facility is to prevent against loss of partial containment during I-131 processing and protection against internal and external exposures during processing of I-131. I-131 processing facility is an enclosed facility fabricated out of lead filled panels and stainless steel sheets. The isotope facility is being designed as per the applicable safety codes and guides. Well-designed ventilation system exists for this facility for prevention of airborne contamination due to I-131 and for the comfort of the personnel in the working area. The exhaust air after filtration through installed charcoal filters in the facility is discharged to the environment through 120 m height stack with advancements in isotope handling, monitoring of parameters and operations.

The motivation for the work is that the instrumentation and control systems of the existing old facility are of traditional type and do not have provisions for automatic logic, safety interlocks and control. The objective is to have the new facility designed with modern and efficient sensors, signal processing circuits and recording systems. Controls and interlocks are designed with taking into considerations ease of operations, handling and importantly safety of the personnel and the environment. The facility is provided with extensive instrumentation for monitoring pressure, temperature, flow and level in different systems including cooling water system, compressed air system, Argon system and more importantly ventilation systems of the operating chamber where isotopes are processed.

V. Arunprasath (✉)

Research Reactor Design and Projects Division, Bhabha Atomic Research Centre, Mumbai, India
e-mail: arunv@barc.gov.in

2 Contributions/Methods

The contributions involve design and providing instrumentation and control system catering to the needs of this new revamped isotope processing facility.

This process involves activities of selection of appropriate sensors and instruments, data acquisition and recording system, defining and configuring the system architecture and safety analysis for given postulating failure initiating conditions and defining the consequences and proposed implementation strategies. All of these are described in the following.

2.1 Isotope Facility System Description

The new isotope facility is for radiochemical processing of neutron irradiated TeO_2 powder and production of I-131 isotope samples, which are used for various applications in medicine and industry. The facility operates by heating the irradiated Tellurium dioxide (TeO_2) powder in an induction furnace and passing it through separators, and a series of traps to extract I-131. An ionization chamber calibrated to I-131 is used to measure the iodine collected. It is provided with various systems like iodine extracting system, cooling water system, compressed air system, Argon system and more importantly ventilation systems. Structurally, the isotope processing facility consists of two enclosures. The primary enclosure consists of three chambers and surrounding lead shielding. Two chambers are for processing of I-131 and the third one for loading and unloading of samples and is called loading unloading chamber. The secondary enclosure houses the primary enclosure (including surrounding lead shielding) and it provides access to personnel entry and material movement through two nos. of doors. It acts as a working area. The chambers of primary enclosure are kept at 25–50 mm of water column (WC) vacuum under normal operating conditions so that direction of airflow is always from secondary enclosure to inside the primary enclosure, thus ensuring radioactivity release outside the chambers does not happen enclosure in case of any breach of boundary. The exhaust fans provided for primary enclosure keep the chambers under desired vacuum. The ventilation fans of primary chambers are capable of maintaining desired vacuum in worst case. The secondary enclosure ventilation consists of exhaust fans and supply grill. Enclosure is maintained under induced suction by operation of exhaust fans and thus maintained under slightly negative pressure with respect to outside environment. The schematic diagram of the ventilation system and the enclosures is shown in Fig. 1.

The cooling water system is to provide constant cooling for the induction furnace coils while the system is under operation. It also consists of an overhead tank which provides cooling in case of the pump trip. Argon gas system is used for integrity and vacuum test of the complete radiochemical processing system prior to starting the isotope extraction process. Compressed air is used at a particular pressure with ejector to create vacuum necessary in the system for the iodine extraction process.

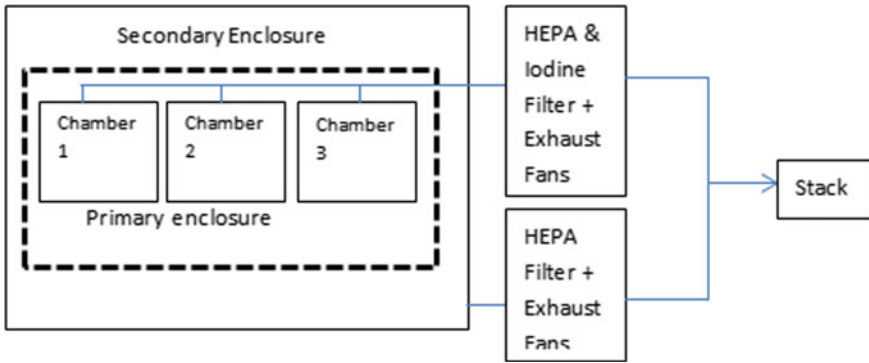


Fig. 1 Schematic diagram of the ventilation system and the enclosures

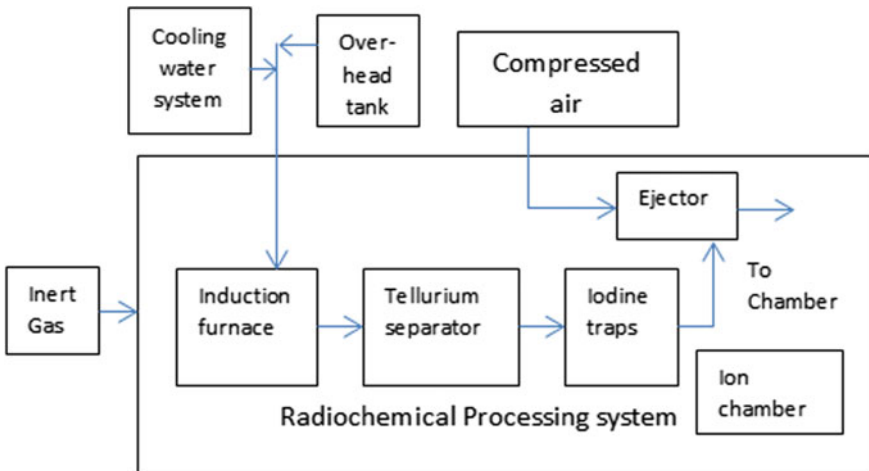


Fig. 2 Schematic diagram of the process systems and the radiochemical processing system

The schematic of the process systems and the radiochemical processing system is shown in Fig. 2.

2.2 Instrumentation and Control System Description

Instrumentation and control system in the new isotope processing facility is provided to monitor various process parameters and provide indications and controls for safe and efficient operation. The system consists of various sensing devices/transmitters, signal conditioning modules, interlock logics and readout/indication devices [3, 4].

In IPF, standard 4–20 mA loop philosophy is followed. Transmitters are used for all parameters which are required to be monitored remotely. A typical instrumentation loop in IPF consists of sensor at the field end. Transmitters are used to convert the sensor signal to standard 4–20 mA signal which is used for various functions like indication, recording, control, generation of alarms/interlocks, etc [5, 6]. For generation of alarms and interlocks, alarm trip units (ATUs) are used. An ATU is a current switch which generates contact output(s) when the input current signal reaches the pre-set threshold (set point). Contact output(s) of ATUs are fed to alarm annunciator and/or interlock logics. All transmitted parameters are fed to the recording system installed in the facility and displayed in the panel. Various process parameters measured are flow, temperature, pressure, differential pressure and level. Design of instrumentation is based on relevant standards. Design parameters are selected as per the values of parameters being measured under various operational conditions [7–9]. The configuration and architecture of the entire instrumentation and control system is provided in Fig. 3.

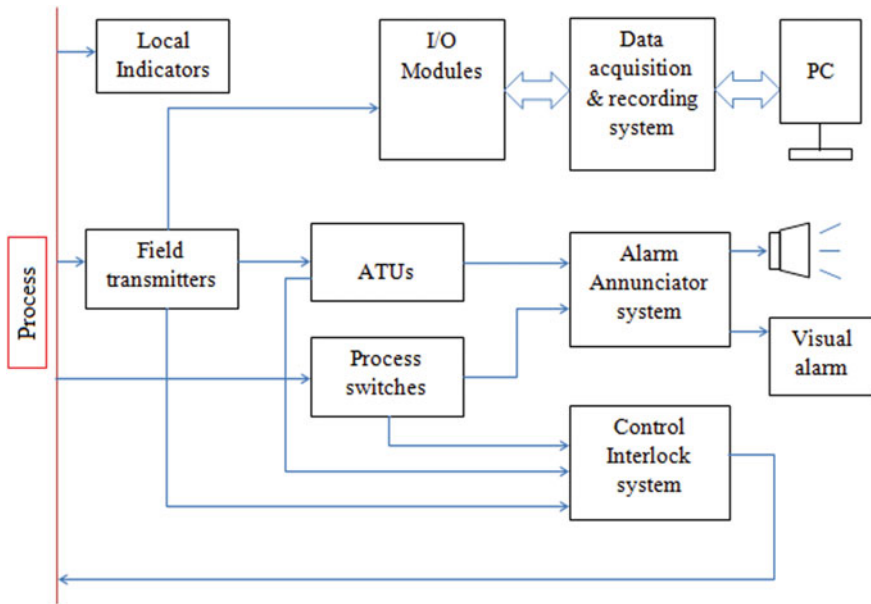


Fig. 3 Configuration and architecture of the instrumentation and control system

2.3 Instrument Selection and Description

2.3.1 Process Sensors

(a) Flow Sensors

The secondary enclosure has three exhaust grills and two exhaust fans to provide air exhaust flow of 3000 m³/h (CMH). There are two supply grills/louver for air supply inside the secondary enclosure. The secondary enclosure is maintained at about -2 to -3 mm of WC with respect to outside environment pressure, whereas the vacuum inside the primary chambers is -25 to -50 mm of WC. This ensures air flow is always from secondary enclosure to glove boxes. The exhaust fan takes suction from the secondary enclosure through suction grills and a header. This flow measurement is done using averaging pitot tube which is a differential pressure-based flow measurement device. It is designed to produce a differential pressure of 14 mm WC for a maximum flow of 3000 CMH [10]. Glass tube rotameter is used for monitoring gas flow locally. Monitoring of cooling water flow is essential to ensure heat removal in the furnace. Metal tube rotameter with transmitter is used for this purpose. An alarm is generated for operator action when this transmitter reading indicates low flow. Furnace trip is generated when the flow reaches very low set point.

(b) Pressure Sensors

Pressure measurement in the new isotope facility is mainly done using bourdon type pressure gauges and pressure transmitters. Pressure transmitters are used for measuring individual primary chamber pressures, individual primary fan suction and discharge pressures. For differential pressure measurements, DP transmitters are used. Differential pressure across primary high efficiency particulate air (HEPA) filter and secondary HEPA filter are monitored using DPTs. Argon inlet pressure is monitored using a U tube manometer. Monitoring of cooling water pressure is also essential along with the flow in ensuring adequate heat removal in the furnace. On low pressure, alarm is annunciated for necessary corrective operation, and on very low pressure, an interlock is executed to trip the furnace for safety purpose. This is accomplished using pressure switch having two single pole double throw (SPDT) hermetically sealed field adjustable microswitches generating potential free contacts.

(c) Temperature Sensors

Temperature measurement in the new isotope facility is mainly done using three wire resistant temperature detectors [RTDs]. Individual primary chamber ambient temperatures are continuously monitored. The three wire RTDs provide good accuracy and repeatability at reasonable cost and have been adopted by all major temperature transmitter manufacturers as the standard type of sensor for most applications. RTDs are provided with thermowells for strength and ease of maintenance purposes.

Thermowells are designed using ASME PTC 19.3 TW-2010 [11]. RTD signals are directly connected to the recording system for monitoring and display.

(d) *Level Sensor*

Status monitoring of cooling water overhead tank level is essential since the tank performs the furnace cooling function in the absence of pumps. In order to ensure the tank level is full, a low alarm is provided at 90% level. This alarm contact generation is achieved by using a float type level switch. The sensor is compact and rugged consisting of a free moving pivoted float assembly and a switch enclosure in non-magnetic material to achieve undisturbed magnetic flux. It employs dual magnets, one carried by the float arm and other by the contact carriers housed in the switch enclosure. A change in liquid level brings the like poles of dual magnets opposite to each other and resulting repulsion force ensures a changeover of contacts with snap action. The magnetic coupling is glandless to prevent any leakage. The sensor is mounted on the side of the storage tank with a bulkhead facilitating ease of maintenance and not hindering access of the tank from top. The sensor is having 1 SPDT microswitch generating potential free contact. The isotope processing can be started only after tank 'low level' alarm is cleared.

(e) *Ion Chamber*

An ionization chamber calibrated for I-131 is used to measure the activity trapped inside the iodine trap [12]. A 20 cm³ ionization chamber is fitted inside the production cell very close to the iodine trap and is connected to a precision electrometer fitted on the control panel using a low noise Triax cable.

2.3.2 Instrument Panel

The instrument panel is of size 1600 mm (height) × 600 mm (width) × 500 mm (depth). Panel size has been chosen to accommodate instruments like relays, ATUs, recording system and its associated input/output (IO) modules, alarm annunciator, 24 V DC power supply and all other accessories. Proper illumination, ventilation and locking arrangements are provided for the panel. The panel is of enclosed type. It is made of sheet steel and is coated with RAL 7035. All parts including side panels, top plate, bottom plate, front and rear doors are made of sheet steel of adequate thickness. Base structure of 100 mm has been provided for the floor standing panel.

2.3.3 Signal Conditioners

(a) *Alarm Trip Units*

In the new isotope facility, alarm trip units (ATUs) are used to generate contacts for various alarms and interlock logics. ATUs are current switches that operate on the

input 4–20 mA signal and generate contact outputs. ATUs having two independent set points and two independent contact outputs are used. ATUs used are externally powered and are DIN rail mountable. Signal isolation is ensured between the current input side, power supply side and the two outputs.

(b) *Relays*

Control logics and interlocks are implemented using hardwired relay logic circuits. The relays used are epoxy sealed relays with four change-over contacts. From a given changeover contact, either normally open (NO) or normally closed (NC) contact is selected as per the requirement. Relays used are DIN rail mountable.

2.3.4 Power Supply

All the transmitters and signal conditioning modules are powered by 24 V DC, 60 W supply. This is a compact and DIN rail mounted power supply. Sensors like level switch and pressure switch & relays are interrogated using 24 V DC. The recording system is powered by single phase 230 V AC. Alarm annunciator, panel illumination light and ventilation fans operate on 230 V AC.

2.3.5 Alarm Annunciator System

Alarm annunciator system is designed to alert the operator of any abnormal behavior of various process system parameters. It comprises of an annunciator system and associated alarm windows. It receives the status of parameters from various systems. Whenever this system registers a change in the status of any parameter (from normal to off-normal or vice-versa), it provides a corresponding visual indication on the relevant display window. It also provides two distinct audio signals (for normal state to alarm state change and for alarm state to normal state change) to attract the operator's attention. It uses ring back sequence [13]. Total number of windows is 30. Alarm accept, reset and test features are provided. Apart from the process parameter alarms, the system also receives input from the dedicated fire detection system of the plant for annunciating fire alarm. Following are the important alarms provided in the facility.

1. Primary Chamber 1 vacuum low (15 mm WC).
2. Primary Chamber 2 vacuum low (15 mm WC).
3. Loading/unloading chamber vacuum low (15 mm WC).
4. Cooling water flow low (30 lpm).
5. Cooling water pressure low (3 kg/cm² (g)).

2.3.6 Data Acquisition and Recording System

A continuous universal monitoring and recording system is provided to acquire, display and record all the parameters in the facility. The system is provided with IO modules capable of measuring universal inputs. The IO modules communicate to the system via Modbus functionality. It is mounted in the instrument panel and all the signals measured in the plant are wired to it. RTDs are directly connected using three wire input channels. The system is provided with a memory card of 1 GB storage capacity for data portability. It also has Ethernet and USB communications. The system also has dedicated software which can be installed in a PC for online monitoring, generating spreadsheets, charting and analysis.

2.3.7 Control Interlocks

Control interlocks ensure safe and efficient operation of the facility and safety of the personnel involved. Some of these are achieved using electrical equipment status and some are implemented using process parameters measured. These are given below.

Following interlocks are provided for ensuring the safe operation:

- (a) In case primary exhaust fan trips, the operating exhaust fan of secondary enclosure will also trip. The converse is not true. This ensures that primary chamber is always under more negative pressure than secondary.
- (b) In case primary exhaust fan trips due to any reason except power failure, the standby fan will start automatically.
- (c) In case of vacuum in any of the three primary chambers drops and reaches 10 mm of WC, the solenoid valve on that chamber's emergency exhaust line will open to improve vacuum resulting due to higher exhaust flow.
- (d) In case of primary exhaust fan trip and the standby fan is not started within one minute, the furnace will trip.
- (f) The induction furnace cannot be started if the cooling water pressure is below 2.5 kg/cm² and flow is less than 25 lpm.
- (g) The isotope processing can be started only after overhead storage tank 'low level' alarm is cleared.

2.4 Safety Analysis of Postulated Conditions

The instrumentation and control system design is based on the principles of fail-safe philosophy, testability and maintainability to ensure safety and reliability.

The processing and logic circuits are kept energized during normal operation and made to de-energize to achieve fail-safe philosophy. Solenoid valves used in the chambers' primary ventilation lines are normally open type.

In case of unavailability of power supply to any transmitter/relay, fail-safe outputs will be provided. Field instruments are selected with necessary environmental protection. Hence, loss of ventilation will not have any significant effect.

In the panel, ventilation is provided by fan operating on 230 V AC supply. The panel mountable instruments are selected so as to withstand the maximum temperature of 45 °C under cooling failure.

Filed sensors are selected taking in to account the process environment including radiation. Hence, radiation exposure within the design basis will not have any significant effect on the field sensors.

For pressure/DP/flow measurements, impulse lines are drawn outside the shielded chambers till the transmitter mounting pipes, which have relatively lesser exposure to radiation.

Equipment and components related to interlock logics are located in accessible areas and are hence not expected to face any radiation exposure.

Failure of sensors/logic circuits of safety-related systems will lead to fail-safe outputs. Hence, radiological consequences on system failure are unlikely.

2.5 Material Selection

Material selection for various instruments is done considering the process environment and the ambient environment. Stainless steel is used for flow sensors, impulse tubing, fittings and thermowells considering like corrosion resistance and strength. Panel enclosure and supports are made of steel and are coated for good weather resistance and finish. Fire Retardant Low Smoke (FRLS) material cables are used for fire resistance. Necessary environment protection is ensured by selecting instruments with suitable Ingress protection following IEC IP code 60529.

2.6 Implementation and Installation Aspects

The instrumentation and control system is also designed keeping in mind aspects related to implementation and installation which are briefed below.

For the averaging pitot tube, straight lengths are provided for the necessary uncertainty level as recommended by ISO 5167. For other field sensors like RTDs, mounting requirements like insertion length, insertion angle and mounting location (elbow/in pipe) are considered [4]. Mounting of transmitters in pipes/wall is done such that they facilitate easy maintenance. Local display instruments like pressure gauges and rotameters are housed in the field area. All other equipment like ATUs, relays, power supply and the recording system and its modules are located in the panel. Cables and tubing are laid with trays having adequate space and are routed along the plant mindful of space requirements and other equipment. Mechanical strength and vibration aspects of thermowells for RTDs are taken care of by

following ASME PTC 19.3-TW 2010. The panel is designed as per IEC codes and its structural supports for standing are provided based on its floor loading. Transmitters are either mounted on 2 inch pipe fixed to the wall or directly mounted on the wall with rigid supports. Impulse tubing is routed and interconnected with double compression fittings. Cables and tubing are laid with cable trays based on estimated space requirements and support strength.

Surveillance activities like monitoring, instrument checks, functional tests, and calibration and response time checks as applicable will be carried out at scheduled intervals to establish the healthiness of instruments [7].

For achieving the desired functional requirements of the facility, proper installation and commissioning of the Instrumentation and control systems is necessary. Hence, a separate commissioning procedure will be prepared at site which includes the following requirements [14].

- (a) Components mounted on the panel, instruments mounted on the pipes and equipment, etc. are checked as per applicable general arrangement drawings/bill of material and instrument listing.
- (b) Tagging details are verified with applicable documents.
- (c) Panel/rack grouting is checked for their rigid support, proper continuity to ground and cable glanding.
- (d) Cable terminations are checked as per terminal lists.
- (e) The wiring for panel and instruments is checked as per wiring diagrams.
- (f) Insulation resistance of panels/equipment is checked to meet system design requirements.
- (g) Operation of power supply units and voltage regulation is checked as per applicable specification/data sheets.
- (h) Integrated functional test is carried out to verify all the operational logics, indications, annunciators, displays and recording to meet system design intent.
- (i) Alarm functions are simulated and verified.
- (j) Calibration for all the sensors is verified to meet system design requirements.
- (k) The equipment is commissioned as per the recommended/approved commissioning procedures and Instruction manuals.

3 Conclusions

Brief description of the isotope processing systems, all the salient features of the instrumentation and control system including sensor selection and description, signal processors, alarm annunciator system, data acquisition and recording system, control interlock logics and their implementation for the new isotope facility were described in the paper.

Following are the ultimate functional requirements of instrumentation and control system in the facility.

- (a) To provide necessary information about various process parameter values and statuses.

- (b) To incorporate automated control and operator information-assistant system.
- (c) To implement the essential interlock logics.
- (d) To provide data acquisition, display and recording for all the plant parameters those are monitored.
- (e) To ensure safe and efficient operation of the facility and safety of the personnel by implementing above.

The conceptualization, design, detailing, configuration of architecture, selection of instruments, selection of materials, engineering of safety and reliability into the system, safety analysis using postulated conditions and defining consequences, design of housing and support structures, proposed installation and commissioning features as described in the paper, all facilitate achieving a performance at a level above par the functional requirements of the new isotope facility.

References

1. K. Hora, *Iodine Production and Industrial Applications*. (IDD Newsletter, 2016)
2. M. Gourani, B. Nabardi, H. Farahani, *Technology of iodine-131 production and its application* (Inst YadernojFizikiNYaTs RK, Kazakhstan, 2005)
3. D.M. Considine, *Process/Industrial Instruments and Controls Handbook* (1999)
4. B. Liptak, *Instrument Engineers' Handbook: Process Measurement and Analysis*, vol. 1, 4th ed (2003)
5. ANSI/ISA-S51.1, *Process Instrumentation Technology* (1993)
6. W. Ibrahim, H. Sallam, Hany, Instrumentation and controls architectures in new NPPs. *Int. J. Nucl. Knowl. Manag.* **6**, 283–302 (2014). <https://doi.org/10.1504/ijnkm.2014.062830>
7. H.M. Hashemian, *Maintenance of Process Instrumentation in Nuclear Powerplants* (Springer, Berlin, 2006), pp. 89–14
8. *Modern Instrumentation and Control for Nuclear Power Plants: A Guidebook*. Technical Report Series 387. (IAEA, 1999)
9. IAEA-TECDOC-973, Research reactor instrumentation and control technology (1995)
10. R. Miller, *Flow Measurement Engineering Handbook* (1996)
11. ASME PTC 19.3 TW-2010, Thermowells, Performance Test Codes (ASME, 2010)
12. P. Singh, R.K.B. Yadav, S. Anilkumar, R.K. Gopalakrishnan, S. Chakraborty, *Improved Monitoring Procedure for Iodine-131 in Radiochemical Process Laboratory* (Indian Association for Radiation Protection, India, 2016)
13. ISA-18.1-1979 (R1992), Annunciator sequences and specifications (1992)
14. K. Lawry, D.J. Pons, *Integrative Approach to the Plant Commissioning Process*. *J. Ind. Eng.* **2013**, 12 (2013). Article ID 572072

Visual Cryptography Using Hill Cipher and Advanced Hill Cipher Techniques



Jagadeesh Basavaiah, Audre Arlene Anthony,
and Chandrashekar Mohan Patil

1 Introduction

Cryptography refers to technique of protection of information and communications by using codes such that the information can be read and processed by the person who is intended for it. In the field of computer science, cryptography deals with secured data and communication methods which are derived from concepts of mathematics and a set of calculations based on rules known as algorithms, which is used in converting messages into a form that is very difficult to decipher. The deterministic algorithms are made use of, for the generation of cryptographic keys, data privacy verification and protection, digital signing, browsing over the internet, and confidential transactions.

Cryptography is in the area of cryptanalysis and cryptology. It consists of methods like words being merged with images, microdots and few other techniques to secure information during transmitting or storing. In the era of computer-centric world, scrambling of plaintext into ciphertext by a process known as encryption and decryption is the reverse process associated with cryptography. Individuals who are involved in working and practicing in this field are called cryptographers. The general cryptography system block diagram is shown in Fig. 1.

Four objectives of modern cryptography are as follows:

1. **Confidentiality:** Information should not be understood by any other person except the person who is intended for it.

J. Basavaiah (✉) · A. A. Anthony · C. M. Patil

Department of Electronics and Communication Engineering, Vidyavardhaka College of Engineering, Mysore, India

e-mail: jagadeesh.b@vvce.ac.in

A. A. Anthony

e-mail: audre.arlene@vvce.ac.in

C. M. Patil

e-mail: patilm@vvce.ac.in

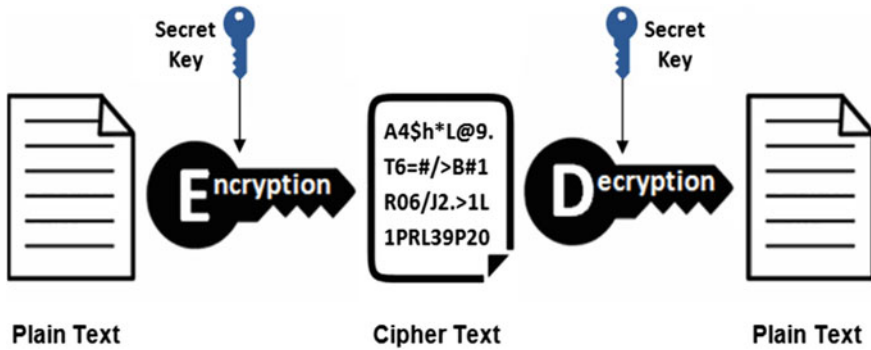


Fig. 1 Block diagram of cryptography system

2. **Integrity:** Information is not supposed to be altered during the process of storing or transmitting between sender as well as the intended receiver.
3. **Non-repudiation:** Sender or creator of information must not deny the purpose in transmitting or creating the information in future.
4. **Authentication:** Both sender and receiver should confirm their identity, source and destination of the information.

Cryptosystems use a set of protocols called cryptographic algorithms in order to encrypt and decrypt the information for secured transmission or communication among the computer systems, the devices and the applications. The cipher suite utilizes the first algorithm to perform the encryption process and the second algorithm to authenticate the information, and finally, the last one is for exchanging the key. The process has been embedded into protocols written in the software which runs on the operating systems and the networked computer systems, which will involve the private and public key generation process for the process of data encryption/decryption, digital signing and verification for the process of message authentication and the key exchange.

The stability of the process of encryption method lies in the algorithm, key secrecy, key length, initialization vectors, and how all these work together. Stability in encryption refers to how difficult it is to find out the key or algorithm, which is not made public. The stability correlates to the time and processing power required to find or break the key. Breaking a key is usually accomplished by a Brute-force attack, which is nothing but trying each and every possible key values for decrypting till the resulting plaintext is significant and meaningful. Subjected to the length of the key and algorithm, the task of decryption can be simple or close to impossible. One more stability factor is the work factor, which is defined as an estimate of the effort taken by an attacker to pierce an encryption method. The stability of the security mechanism should be correlated to the sensitivity of the information that is being encrypted. Every encryption mechanism has its purpose and place. Even though if the algorithm is thorough and complex, there exist other problems within encryption which can deteriorate the strength and stability of encryption method. Since the

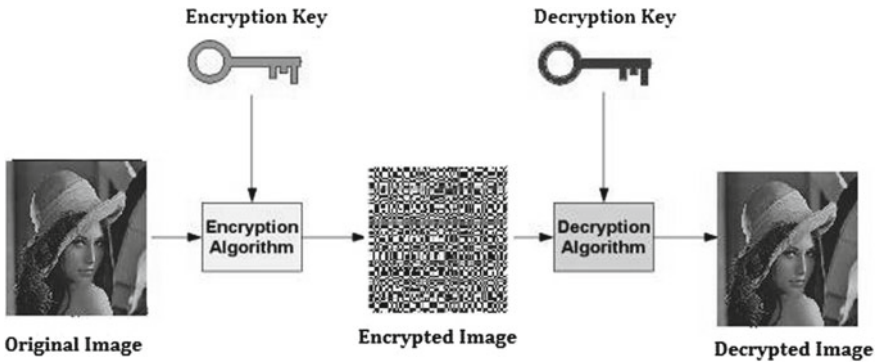


Fig. 2 Visual cryptography

key is the secret value required to encrypt and decrypt the information, inappropriate security of the key will weaken strength and stability of encryption. An exceptionally strong and stable algorithm can be developed using a large and random key value and a large key space which are the necessities for a strong and stable encryption.

Visual cryptography [8] is a type of cryptographic method that lets visual information such as text or images to be encrypted such that decrypted information will be an image. In the present era of the digital world, securing the digital images has become very significant due to swift growth of the Internet. Securing of the digital images has engrossed much requirement, and the different encryption principles and methods are being proposed to secure the digital images. Image encryption techniques involve in converting an image into another which would be difficult in understanding. In contrast, the process of image decryption will recover the original image from the image that is encrypted. Many encryption standards for encrypting and decrypting of image data are present and there is no single encryption algorithm which would satisfy for all the image types. Figure 2 shows the process of visual cryptography.

Many encryption algorithms that are designed for the digital images were suggested during the mid-1990s. The two main classifications of the image encryption algorithms are: (a) chaos-based selective or non-selective methods and (b) non-chaos selective methods. Many of the algorithms are being specifically designed for compressed or uncompressed image formats, and few algorithms are for the format complaints. Some methods will offer light encryption where as other methods will offer strong encryption. Some algorithms can be scaled and can have various modes that range from weaker to stronger encryption.

2 Literature Survey

Hammoudi and Melkemi [1] proposed a visual cryptography technique which was pictogram-based, which is used in generating shares that is being textured with the

aesthetic and customizable rendering. The robustness feature present in the proposed technique toward the automatic decoding of the secret message was shown. Investigational results reveal the concrete personalized shares [7] and the applicative potentials for creative as well as the security domains. In future, it was suggested to examine the personalization of visual cryptography shares by the exploitation of a mix of varied pictograms.

Zhang and Wang [2] proposed an algorithm of multiple-images encryption (block cipher) which is having the chaotic system and the 3D permutation model. Firstly, chaotic systems are used in generating chaotic sequences and images. Later, using 3D permutation model the sender will permute the plain image. Finally, the sender will perform the XOR logic function on the permuted images. Thus, the comparison of the simulation process and the algorithm proposed possesses high security, efficiency and desirable encryption images.

Maisa'a Abid et al. [3] suggested a novel method where the combination between the concealment of information and the encryption would make it difficult in identifying that data which has to be transmitted via the network system. The study uses two-force secure (2FS) for encrypting the images, which means, on the image the SF is twice frequent, which helps in obtaining dominant encryption and then the process of secret message being concealed is done in the cryptography of that particular image which would be implemented by means of a secret key, and the stego-encryption image is being transformed for the sake of storage of Internet of things in the database of IoT, when the user is in requirement of some information then it could be easily accessed in the Internet of things. The result obtained is to be evaluated using different measures, which includes mean square error, peak signal noise ratio, entropy correlation, coefficient and histogram. The developed algorithm is fast, efficient, highly secured, robust and transparent in nature.

Wadhwa et al. [4] suggested a method for the design of the one-way cryptographic hash function and the block ciphering scheme which is based on the hash codes. In this work, the message is being divided into different blocks with each of the block being separately processed by the chaotic systems. Few transitional hashes have been created by the use of the advanced control and input parameters. These two hash codes are being used in order to generate the final hash. Results of the experiments justify the outstanding performance of the chaotic hash method that was proposed. Besides, the hash code that has been generated is being applied for realization of an image block ciphering technique. The process of encryption is usually plain image dependent thus it shows reasonable encryption effect which is apt for real-world applications.

Singla et al. [5] proposed a new algorithm toward the construction of an efficient cryptographic hash function by the usage of a four-layer chaotic neural network. This proposed hash function will satisfy the security necessities of diffusion and confusion, and the mechanism will allow the flexibility in the length of hash value and that will make it resistant to the birthday attack for the hash lengths which would be longer than the 128 bits. Furthermore, the reduction of the running time of the neural network is done by using the parallel processing. The statistical analysis of

the algorithm that is being proposed would prove that it is one of the good possibility for designing of cryptographic hash function.

3 Proposed Methodology

A symmetric key model has been used for encryption and decryption of the image. The model for image encryption and decryption is as shown in Fig. 3.

A. Hill Cipher Technique

The Hill Cipher has been a well-known symmetric block cipher cryptosystem which is having a number of advantages in encryption of data. It is simple as it employs matrix multiplication and matrix inversion operation for the process of enciphering and deciphering. The operating speed and the throughput of the algorithm is high. Nevertheless, the Hill Cipher algorithm is not capable of encrypting images that consists of huge areas of single color code. Hence, all the features of the images will not be hid which reveals the patterns present in the plaintext. Furthermore, it could be broken without much difficulty with a plaintext attack that is known showing weak security.

Matrix manipulation is the core of Hill Cipher algorithm. In the process of encryption, algorithm will take m successive plaintext letters and as a replacement for that, it will substitute m cipher letters. For each character, numerical values are allocated

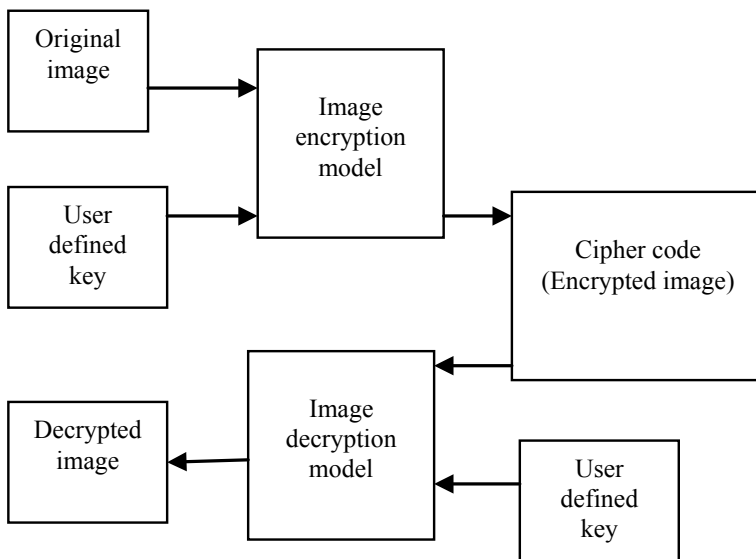


Fig. 3 Image encryption and decryption model

such as $a = 0, b = 1, \dots, Z = 25$ [5, 9], the substitution of ciphertext letters as a replacement of plaintext will lead to m linear equations. The system of m linear equations for $m = 3$ is defined as follows:

$$\begin{aligned} C_1 &= (K_{11}P_1 + K_{12}P_2 + K_{13}P_3) \text{ mod } 26 \\ C_2 &= (K_{21}P_1 + K_{22}P_2 + K_{23}P_3) \text{ mod } 26 \\ C_3 &= (K_{31}P_1 + K_{32}P_2 + K_{33}P_3) \text{ mod } 26 \end{aligned} \quad (1)$$

Equation (1) can be articulated as $C = KP$, where C and P are column vectors of length 3, which denotes the plaintext and ciphertext correspondingly, and K is encryption key which is a 3×3 matrix. All the operations are done with mod 26. Decryption uses the inverse of the encryption key matrix K . The matrix K has an inverse K^{-1} defined by

$$KK^{-1} = K^{-1}K = I \quad (2)$$

where ' I ' indicates the identity matrix. Usually, the inverse of the matrix does not always exist, and if it exists, then it satisfies Eq. (2). The inverse matrix K^{-1} is applied to the ciphertext, and then the plaintext is recuperated. Generally, for encryption and decryption equations can be written as follows:

For encryption:

$$C = E_k(P) = K_p \quad (3)$$

For decryption:

$$P = D_k(C) = K^{-1}C = K^{-1}K_p = P \quad (4)$$

If m is the length of the block, then 26^m different m letters blocks are possible, and each of them is considered as a letter in a 26^m letter alphabet [6]. Hill Cipher method accounts to a mono-alphabetic substitution on the alphabet.

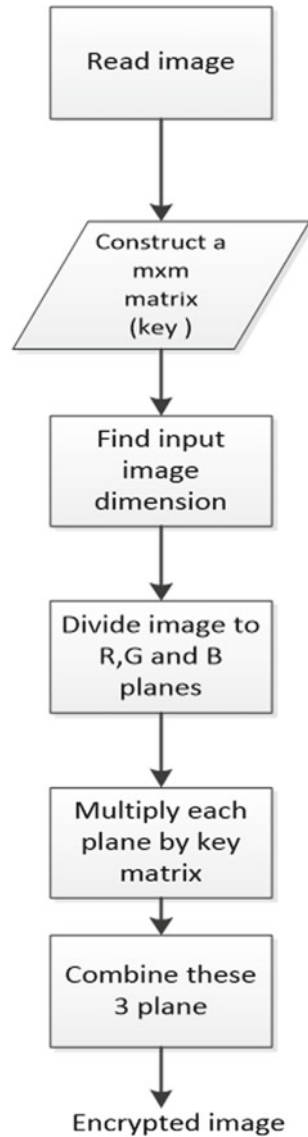
The flowchart for image encryption using Hill Cipher method is as shown in Fig. 4.

The flowchart for image decryption using Hill Cipher algorithm is as shown in Fig. 5.

B. Advanced Hill Cipher Algorithm

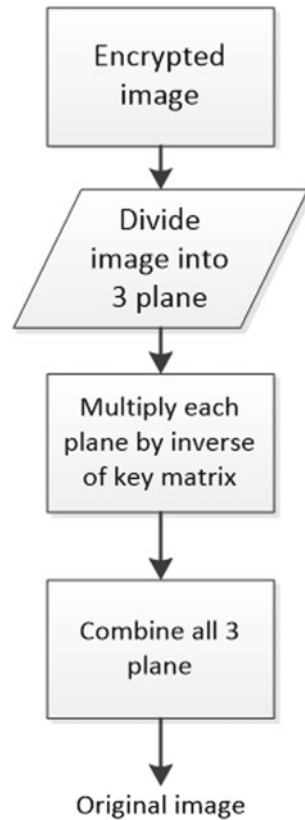
The Hill Cipher algorithm will be a symmetric key algorithm which will have many advantages compared to others. The inverse matrix will be used in decrypting of the plaintext. Therefore, it is very clear that the decryption of the encrypted text cannot be done if the matrix that is being used is not an invertible matrix. In case of the generation of the self-invertible matrix, the matrix that is being used for the process of encryption should be self-invertible by itself. Hence, it is not required to determine the inverse of that particular matrix during the process of the decryption. Furthermore, this method will remove the complexity present in the computation that

Fig. 4 Hill Cipher encryption algorithm



is involved in determining the inverse of the matrix while performing the decryption technique. The main aim is to perform the process of image encryption using a technique that is not the same as the conventional Hill Cipher technique. Advanced Hill Cipher encryption technique for image cryptography is proposed that makes use of an involutory key matrix. Modular arithmetic is utilized to create the involutory matrix.

Fig. 5 Hill Cipher decryption algorithm



Hill Cipher can be used for encryption and decryption of the grayscale images along with color images. Similarly, advanced Hill Cipher algorithm can be employed to perform the encryption and decryption process of the grayscale along with color images. In case of grayscale images, the modulus will be 256. In case of the color images, the color image will be divided into R-G-B components and each component is encrypted individually. Finally, encrypted components will be concatenated to get the color image which is encrypted. Advanced Hill Cipher is fast encryption method that overcomes problems associated with encryption of an image with homogeneous background. This paper presents a comparative analysis of Hill Cipher and advanced Hill Cipher encryption techniques. The encrypted images at the output show that the proposed advanced Hill Cipher technique is relatively robust and reliable. Figure 6 shows the advanced Hill Cipher algorithm.

Algorithm for Advanced Hill Cipher Encryption:

Step 1: Construct an involutory key matrix with the dimension of $m \times m$.

Step 2: The plain image will be divided into symmetric blocks of dimension $m \times m$.

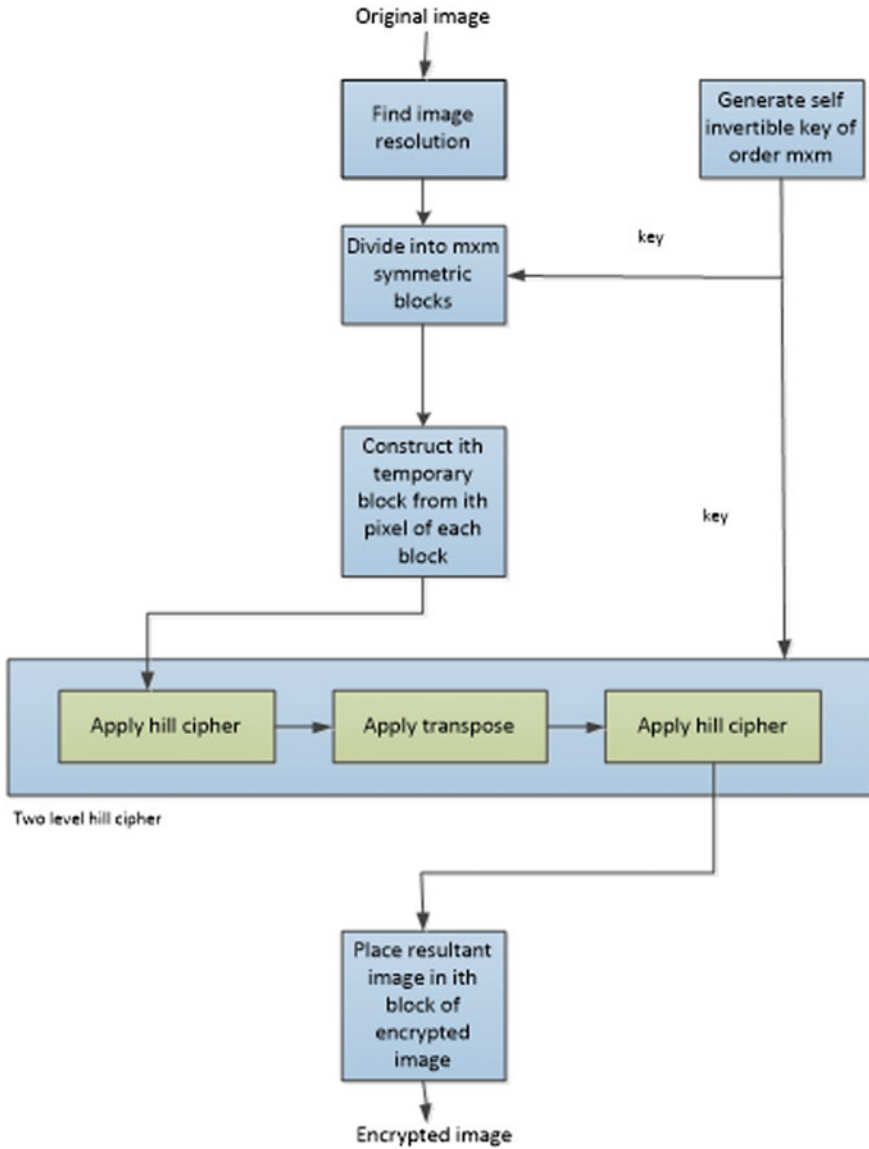


Fig. 6 Advanced Hill Cipher encryption algorithm

Step 3: The i th pixels present in each block will be used together in forming a temporary block.

- a. Hill Cipher technique is applied onto the temporary block.
- b. The resultant matrix is transposed and Hill Cipher is again applied to the matrix.

Step 4: The resultant matrix which is obtained is positioned in the i th block of the encrypted image.

For decryption the reverse process of encryption is to be employed.

4 Results and Discussions

The results of the process of encryption by making use of Hill Cipher and advanced Hill Cipher algorithms are shown in Table 1.

The inputs to the decryption algorithm will be the encrypted images and the outputs will be the original image. The results of decryption using Hill Cipher and advanced Hill Cipher algorithms are as shown in Table 2.

The good encryption technique must be robust in nature against all kinds of Brute-force, cryptanalytic and statistical attacks. In this paper, the performance analysis of both the algorithms is carried out.

A. Histogram analysis

The investigation reveals that the histograms of the new input and the encrypted images in the R-G-B planes will be obtained which is as shown in Figs. 8 and 9. It is pretty clear that from the histograms that is being obtained the encrypted images will have an histogram which will nearly distributed uniformly and noticeably different from that of the original image input. Therefore, the encrypted image will not be able to provide any of the traces which employ the statistical attack on the proposed encryption of the image. It is very difficult to obtain the original image information

Table 1 Results of Hill Cipher and advanced Hill Cipher algorithm


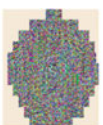
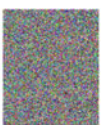


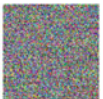

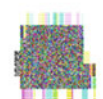
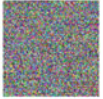
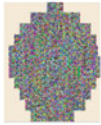
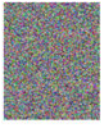


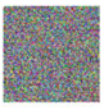

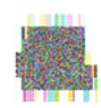


Image	Input Image	Encryption using Hill Cipher Algorithm	Encryption using Advanced Hill Cipher Algorithm
Logo			
Nike			
Dog			

Table 2 Decryption using Hill Cipher and advanced Hill Cipher algorithms

Input to Decryption using Hill Cipher Algorithm	Input to Decryption using Advanced Hill Cipher Algorithm	Output Image
		
		
		

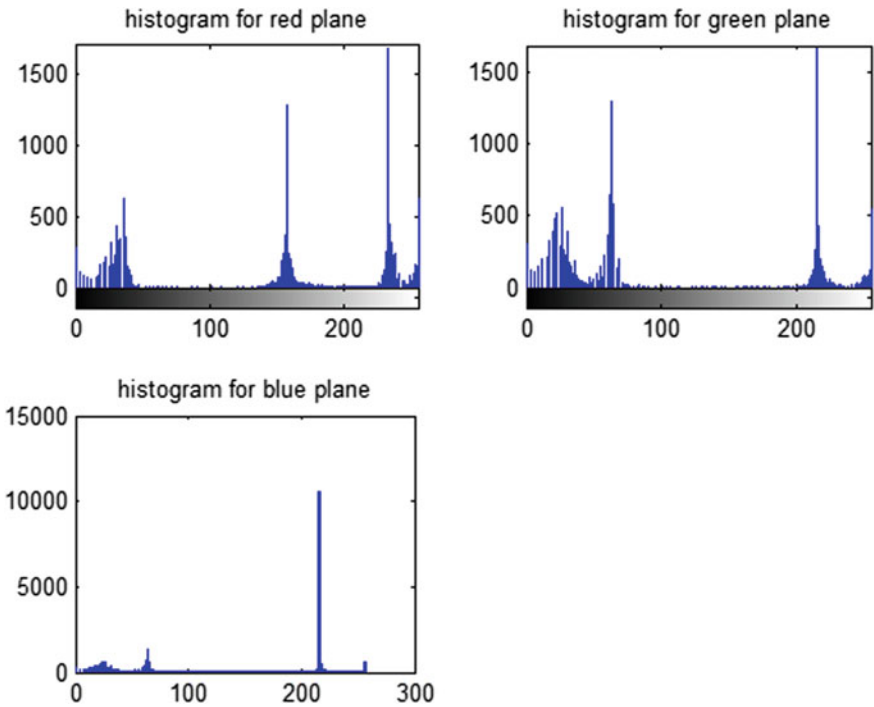


Fig. 8 Histogram of the input logo image

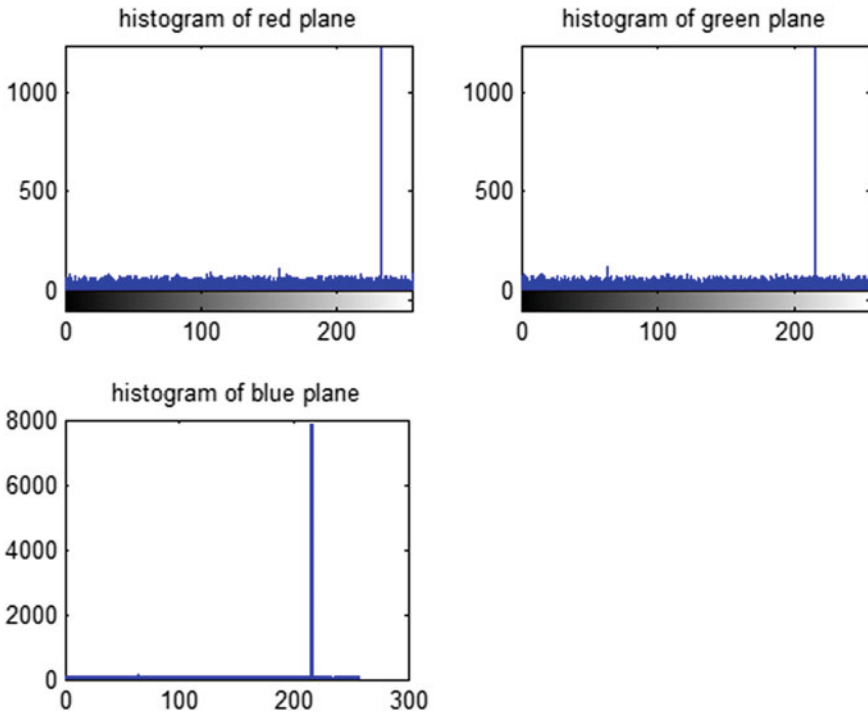


Fig. 9 Histogram of the encrypted logo image

from the histogram analysis. The properties suggest that the proposed encryption method will have a high security against the statistical attacks.

B. Entropy

Entropy is defined as a measure of the uncertainty which is being associated with random variables. In the Hill Cipher technique, the mutual data decreases amongst encrypted image variables besides the entropy value will be increased. The secured cryptosystem ought to satisfy the condition entropy of ciphered image which is not going to provide any information regarding the plain images. The following equation is used to calculate the entropy:

$$H_e = \sum_{k=0}^{G-1} P(k) \cdot \log_2(P(k)) \tag{5}$$

where

H_e = Entropy of image

G = Gray value of an input image (0–255).

$P(k)$ = Probability of the occurrence of symbol k .

Table 3 Entropy calculation table for Hill Cipher and advanced Hill Cipher algorithms

Image	Entropy		
	Original image	Encrypted image using Hill Cipher method	Encrypted image using advanced Hill Cipher method
Logo	7.4299	7.9969	7.9974
Nike	5.7001	7.7104	7.9975
Dog	7.6007	7.9983	7.9988

If in case each symbol would have an equal probability, then the entropy $H_e= 8$ which would correspond to some random source.

Table 3 shows the entropy calculation table for Hill Cipher and advanced Hill Cipher algorithms.

Entropy values are calculated using both Hill Cipher and advanced Hill Cipher method. Entropy values of encrypted image should be greater than the original image, and then the encrypted image will not impart information of original image. Here, we see that encrypted image has more entropy value than that of original image. Also, we see in advanced Hill Cipher entropy value of encrypted image is more than that of Hill Cipher. Hence, advanced Hill Cipher is secure.

C. Correlation

Correlation is defined as the statistical measurement of the relationship between two different variables. Generally, correlation ranges between +1 and -1. The zero correlation will indicate that there will not be any relationship between the variables present. A correlation of -1 will indicate a perfect negative correlation, which means that when one variable increases, the other variable decreases. A correlation of +1 will indicate a perfect positive correlation, which means that together both variables will move in the same direction. A correlation coefficient of value 1 means that those two numbers will be perfect correlation: If one goes up so will the other and if there is a change in one then there will be a multiple of the change in the other variable. A correlation coefficient of value -1 will indicate that the numbers are in perfect inversed correlation. Growth in one variable will be a negative multiple of the growth in the other variable. The correlation coefficient of value zero indicates that there is no relation between two numbers. The non-zero correlation coefficient indicates that there is relation between numbers, but unless the coefficient is either 1 or -1 there might be other influences and the relationship between the two numbers will not be fixed. Therefore, if one number is known the other number can be estimated, but then not with certainty. If correlation coefficient is closer to zero, then uncertainty will be greater, and lower correlation coefficients indicate that the relationship is not useful.

Table 4 shows the correlation calculation table for Hill Cipher and advanced Hill Cipher algorithm.

Compared to Hill Cipher method, advanced Hill Cipher gives less correlation, hence, this technique is more secure than the Hill Cipher method.

Table 4 Correlation calculation table for Hill Cipher and advanced Hill Cipher algorithms

Image	Correlation	
	Hill Cipher method	Advanced Hill Cipher method
Logo	0.5528	0.0087
Nike	-0.0061	-0.0046
Dog	0.0046	0.0035

5 Conclusion

The Hill Cipher algorithm will not be appropriate for the encryption of the images which will consist of huge areas of a single color code. Therefore, it is not going to hide all the features of the particular image that discloses the patterns present in the plaintext. Additionally, with the known plaintext attack, the algorithm is easily broken down which reveals the weak security of the data. On the other hand, advanced Hill Cipher algorithm will definitely be more secured than the Hill Cipher algorithm which would Brute-force attacks. In the advanced Hill Cipher algorithm, a Brute-force attacks would need $2^{7+8*(\frac{n}{2})^2}$ number of keys being generated, where n indicates the order of the key matrix. The advanced Hill Cipher algorithm is a faster encryption technique which would provide reasonable results which is against the Hill Cipher technique. The scheme is likewise resilient against the known plaintext attacks. This algorithm will work for the gray scale as well as for the color images. The result analysis shows that the advanced Hill Cipher technique is much more secured than Hill Cipher technique.

As a future directive, it is suggested to use advanced encryption algorithms for secured image cryptography.

References

1. K. Hammoudi, M. Melkemi, Personalized shares in visual cryptography. *Journal of Imaging* **4**(11), 126 (2018). <https://doi.org/10.3390/jimaging4110126>
2. X. Zhang, X. Wang, Multiple-image encryption algorithm based on the 3D permutation model and chaotic system. *Symmetry* **10**(11), 660 (2018). <https://doi.org/10.3390/sym10110660>
3. K.A. Maisa'a Abid, A. Alabaichi, A.S. Abbas, Dual method cryptography image by two force secure and steganography secret message in IoT. *TELKOMNIKA (Telecommun. Comput. Electron. Control)* **18**(6), 2928 (2020). <https://doi.org/10.12928/telkomnika.v18i6.15847>
4. S. Wadhwa, M. Ahmad, H. Vijay, Chaotic Hash function based plain-image dependent block ciphering technique, in *2016 International Conference on Advances in Computing, Communications and Informatics (ICACCI)* (2016). <https://doi.org/10.1109/icacci.2016.7732117>
5. P. Singla, P. Sachdeva, M. Ahmad, Exploring chaotic neural network for cryptographic Hash function, in *Lecture Notes in Electrical Engineering Emerging Trends in Computing and Communication*, pp. 143–48 (2014). https://doi.org/10.1007/978-81-322-1817-3_16
6. S.K. Panigrahy, On the privacy protection of biometric traits: palmprint, face, and signature, in *Communications in Computer and Information Science* (2009)

7. K. Hammoudi, M. Melkemi, Personalized shares in visual cryptography. *J. Imag.* **4**(11), 126 (2018)
8. D. Jena, S.K. Jena, A novel visual cryptography scheme, in *2009 International Conference on Advanced Computer Control* (2009)

Analysis of Power Adaptation Techniques Over Beaulieu-Xie Fading Model



Veenu Kansal and Simranjit Singh

1 Introduction

The channel capacity is a significant factor in wireless communication system to describe the maximum achievable data rate of the system. It is a useful performance metric by which we can design new techniques to enhance the spectral efficiency of today's communication systems such as 5G networks and indoor, future femtocells, and high speed train scenarios [1, 2]. In such scenarios, the received signal contains multiple signals, and can be composed of specular components or diffuse components or both. The accurate representation of the distribution received signal power in these types of fading environment is difficult by using the existing models because the modeling of such a fading channel requires adaptive parameters for both specular power and fading severity. Recently, a new fading channel was proposed in [3], which can characterize the scenario of such environmental fading conditions. It consists of a flexible fading parameter ' m ' (like the one in Nakagami- m fading) which controls the severity of fading and ' K ' (like the one in Rician fading) which represents the power of specular components. This new proposed fading channel called Beaulieu-Xie (BX) fading that is based on the non-central chi distribution and is very useful in studying the received power distribution in case of high speed vehicles, small cells area, femtocells, heterogeneous networks, etc., with more accuracy. Although κ - μ , generalized Ricean fading distributions are also based on non-central chi distribution, the reason why BX fading is different i.e., its fading power is bounded. Looking at the second-order moment it becomes evident that unlike the κ - μ or generalized Ricean fading, and the total power of the NLOS component is invariant [4]. In literature, an

V. Kansal · S. Singh (✉)

Department of ECE, Punjabi University Patiala, Patiala, India

e-mail: simranjit@live.com

V. Kansal

e-mail: veenukansal@outlook.com

increasing number of researchers have turned their attention on the analysis of BX fading model [4–13].

The asymptotically analysis of correlated BX fading channel with equal gain combining and selection diversity is reported in [5, 6]. Further in [7], the analysis on BX fading channel is extended for second-order statistics such as level crossing rate and average fade duration. The effective rate of BX fading channel by considering the delay occurring in the received signal is investigated in [8]. In [9], the author investigated the performance of energy detectors over BX fading model by using the moment generating function (MGF) method. Furthermore, the error rate analysis of various binary and M-ary phase-shift keying (PSK) modulation schemes is evaluated over this fading channel in [10, 11].

Recently, new research is addressed on the decode-and-forward-based mixed BX fading for w dual-hop transmission system [12] and spectral efficiency analysis with maximal ratio combining with different power and rate transmission protocols [13]. The main motive of this work is to present the analysis for the capacity of BX fading with adaptive transmission protocols with low mathematical complexity. In this work, the expressions for capacity with optimal rate adaptation (ORA) by keeping constant transmitter power, capacity with optimal power and rate adaptation (OPRA), channel inversion with fixed rate (CIFR), and truncated channel inversion with fixed rate (TIFR) over the BX fading model are derived. The derived expressions are obtained in form of elementary functions which makes the analysis easy and tractable. We also demonstrate the usefulness of our derived expressions by including the capacity expressions of classical models such as Rice, Rayleigh, and Nakagami-m fading as special cases of BX fading.

In literature, the capacity analysis under adaptive transmission techniques for the different fading channels such as Rayleigh fading, α - η - κ - μ fading, α - η - μ fading, Fisher-Snedecor F fading, fluctuating two ray fading, κ - μ /inverse gamma and η - μ /inverse gamma composite fading models , etc., was reported in [14–21].

2 Beaulieu-Xie Fading Model

The fading model consists of multiple specular components with diffused power. The PDF of the signal-to-noise ratio (SNR) can be evaluated by [8]:

$$p_{\gamma}(\gamma) = \eta^{\frac{m+1}{2}} \exp(-mK - \eta\gamma) \times \left(\frac{\gamma}{mK}\right)^{\frac{m-1}{2}} I_{m-1}\left(2\sqrt{mK\eta\gamma}\right) \quad (1)$$

where γ is instantaneous SNR, $\eta = L_1^{m-1}(-mK)/\bar{\gamma}$, $L_{\alpha}^n(\cdot)$ denotes the associated n th order Laguerre polynomial [13], ' m ' controls the shape of PDF, ' K ' denotes the ratio of specular power to diffuse power, ' $\bar{\gamma}$ ' represents the average received SNR, and $I_{m-1}(\cdot)$ is the modified Bessel function of first kind with order of $(m - 1)$.

The expression of the PDF in (1) can be rewritten by using the property [22, (Eq. 8.445)] as:

$$p_\gamma(\gamma) = \eta^m \gamma^{m-1} \exp(-mK - \eta\gamma) \sum_{g=0}^{\infty} \frac{(mK\eta\gamma)^g}{g! \Gamma(m+g)} \quad (2)$$

where $\Gamma(\cdot)$ is gamma function.

3 Capacity Analysis of BX Fading Model

The expressions for capacity over the BX fading model with several adaptive transmission schemes are derived in this section.

A. *Optimal Simultaneous Power and Rate Adaptation (OPRA):*

In this technique, it is considered that the information of channel must be available to both the transmitter and the receiver. The capacity of a fading model for power and rate adaptation, C_{opra} (bits/s) can be given by [14]:

$$C_{\text{opra}} = B \int_{\gamma_0}^{\infty} \log_2 \left(\frac{\gamma}{\gamma_0} \right) p_\gamma(\gamma) d\gamma \quad (3)$$

where B defines the channel bandwidth (in Hz) and γ_0 indicates threshold value of SNR. The γ_0 must satisfy the condition given as below [15]:

$$\int_{\gamma_0}^{\infty} \left(\frac{1}{\gamma_0} - \frac{1}{\gamma} \right) p_\gamma(\gamma) d\gamma = 1 \quad (4)$$

By inserting the value of (2) into (4), then by using the identity [22, Eq. (3.351.2)], we can solve the integral for γ_0 as:

$$\sum_{g=0}^{\infty} \frac{(mK)^g \exp(-mK)}{\Gamma(\varphi)g!} \left(\frac{\Gamma(\varphi, \eta\gamma_0)}{\gamma_0} - \frac{\Gamma(\varphi-1, \eta\gamma_0)}{1/\eta} \right) = 1 \quad (5)$$

where $\varphi = m + g$.

It is noticed that there is a unique value of γ_0 which can satisfy (5). An asymptotic evaluation of (5) demonstrates that as the average SNR, $\bar{\gamma}$ tends to ∞ , γ_0 tends to 1. We cannot solve the γ_0 analytically, but can be found by using software such as MATLAB or MATHEMATICA.

Now, to solve (3), substitute the value of $p_\gamma(\gamma)$ from (2) into (3), we get:

$$C_{\text{opra}} = \frac{B}{\ln 2} \sum_{g=0}^{\infty} \frac{(mK\eta)^g \eta^m}{\Gamma(\varphi)g!} \int_{\gamma_0}^{\infty} \ln\left(\frac{\gamma}{\gamma_0}\right) \gamma^{\varphi-1} \exp(-mK - \eta\gamma) d\gamma \quad (6)$$

By taking $\gamma/\gamma_0 = t > 0$, as $\gamma \rightarrow \gamma_0$, then $t \rightarrow 1$, and as $\gamma \rightarrow \infty$, $t \rightarrow \infty$. Now, Eq. (6) can be written as:

$$C_{\text{opra}} = \frac{B}{\ln 2} \eta^m \gamma_0^m \exp(-mK) \sum_{g=0}^{\infty} \frac{(mK\eta\gamma_0)^g}{\Gamma(\varphi)g!} \times \int_1^{\infty} t^{\varphi-1} \ln(t) \exp(-\eta\gamma_0 t) dt \quad (7)$$

By employing the identity as provided in [14, Eq. (64)], the capacity OPRA per bandwidth is obtained as:

$$\frac{C_{\text{opra}}}{B} = \frac{\exp(-mK)}{\ln 2} \sum_{g=0}^{\infty} \frac{(mK)^g}{g!} \sum_{h=0}^{\varphi-1} \frac{\Gamma(h, \eta\gamma_0)}{h!} \quad (8)$$

where $\Gamma(\cdot, \cdot)$ is the upper incomplete gamma function.

The resulting expression in (8) is a generalized expression, and therefore some of the existing fading channels can be deduced from this expression as special cases of the BX fading by substituting specific value of fading parameters K, m such as Rayleigh fading (constant $K = 0, m = 1$), Rician fading ($K = \text{any value}, m = 1$), Nakagami- m fading ($K = 0, m > 0.5$).

B. Constant Transmit Power

In the ORA technique with constant transmit power, we consider that the statistics of channel is only available at the receiver. With this ORA scheme, the rate of transmitter must be adapted to maintain a fixed power level. The capacity of the channel with constant transmits power scheme, C_{ora} (bits/s), is evaluated by [16]:

$$C_{\text{ora}} = B \int_0^{\infty} \log_2(1 + \gamma) p_{\gamma}(\gamma) d\gamma \quad (9)$$

By substituting the value of $p_{\gamma}(\gamma)$ from (2) into (9), we get the expression as:

$$C_{\text{ora}} = \frac{B}{\ln 2} \exp(-mK) \sum_{g=0}^{\infty} \frac{(mK)^g \eta^{\varphi}}{\Gamma(\varphi)g!}$$

$$\times \int_0^\infty \ln(1 + \gamma) \exp(-\eta\gamma) (\gamma)^{\varphi-1} d\gamma \tag{10}$$

With the aid of formula given as [14, Eq. (78)]:

$$\int_0^\infty t^{u-1} \ln(1 + t) \exp(-\mu t) dt = (u - 1)! \exp(\mu) \sum_{h=1}^u \frac{\Gamma(-u + h, \mu)}{\mu^h} \tag{11}$$

By using (11), we can solve (10) as:

$$\frac{C_{\text{ora}}}{B} = \frac{\exp(-mK + \eta)}{\ln 2} \sum_{g=0}^\infty \frac{(mK)^g}{g!} \sum_{h=1}^\varphi (\eta)^{\varphi-h} \Gamma(h - \varphi, \eta) \tag{12}$$

The resulting expression of the capacity with ORA technique in (12) is also a generalized expression and that can be used to study the existing fading channels as special cases of the BX model.

C. Channel Inversion with Fixed Rate (CIFR)

In the CIFR system, it is considered that the transmitter side has complete knowledge about the channel statistics. To keep a fixed SNR at receiver, the power of transmitter must be adjusted accordingly. So, it is needed to invert the fading effects of the model which can be adopted by using CIFR policy. The channel capacity C_{cifr} (bits/s) with this policy can be evaluated by [15]:

$$C_{\text{cifr}} = B \log_2 \left(1 + \frac{1}{E[1/\gamma]} \right) \tag{13}$$

where $E[1/\gamma] = \int_0^\infty \frac{P_\gamma(\gamma)}{\gamma} d\gamma$.

By using the PDF given in (1) and the formula [22, (Eq. 6.6432)], [22, Eq. (9.202.2)], we can evaluate the $E[1/\gamma]$ as:

$$E[1/\gamma] = \frac{\eta}{m - 1} M(1, m, -mK) \tag{14}$$

where $M(.,.;.)$ is confluent hypergeometric function.

Now, by using (14) in (13), the final exact expression for CIFR per bandwidth can be evaluated as:

$$\frac{C_{\text{cifr}}}{B} = \log_2 \left(1 + \frac{m - 1}{\eta M(1, m, -mK)} \right) \tag{15}$$

D. Truncated Channel Inversion with Fixed Rate (TIFR)

In truncated channel inversion policy, the inversion takes place in the channel fading only when the output SNR is more than threshold SNR, γ_0 . The capacity of channel with TIFR, C_{tifr} (bits/s) can be evaluated by [17]:

$$C_{\text{tifr}} = B \log_2 \left(1 + \frac{1}{\int_{\gamma_0}^{\infty} \frac{p_{\gamma}(\gamma)}{\gamma} d\gamma} \right) (1 - P_{\text{out}}) \tag{16}$$

where P_{out} denotes the outage probability. It can be described as the probability which exists when the instantaneous SNR, γ falls below a certain predefined threshold, γ_0 [13]. The outage probability is given by:

$$P_{\text{out}} = 1 - \int_{\gamma_0}^{\infty} p_{\gamma}(\gamma) d\gamma \tag{17}$$

Using (2) in (17), we obtain:

$$P_{\text{out}} = 1 - \left(\exp(-mK) \sum_{g=0}^{\infty} \frac{\eta^{\varphi} (mK)^g}{g! \Gamma(\varphi)} \int_{\gamma_0}^{\infty} \gamma^{\varphi-1} \exp(-\eta\gamma) d\gamma \right) \tag{18}$$

By solving the integral in (18) with the help of [22, Eq. (3.351.2)], the resulting expression for the outage probability is obtained as:

$$P_{\text{out}} = 1 - \exp(-mK) \sum_{g=0}^{\infty} \frac{(mK)^g \Gamma(\varphi, \eta\gamma_0)}{g! \Gamma(\varphi)} \tag{19}$$

Now, by substituting (19) into (16) and on solving the integral term given in (16), the final expression for TIFR per bandwidth is obtained as:

$$\begin{aligned} \frac{C_{\text{tifr}}}{B} = \log_2 & \left(1 + \frac{g! \Gamma(\varphi) \exp(mK)}{\eta \sum_{g=0}^{\infty} (mK)^g \Gamma(\varphi - 1, \eta\gamma_0)} \right) \\ & \times \sum_{g=0}^{\infty} \frac{(mK)^g \Gamma(\varphi, \eta\gamma_0)}{g! \exp(mK) \Gamma(\varphi)} \end{aligned} \tag{20}$$

This is the final exact closed form expression for TIFR.

4 Numerical Results and Discussion

The various graphs are plotted to investigate the capacity of given fading channel. The final expression obtained for OPRA, ORA and TIFR is expressed in terms of infinite series. We discussed the convergence of derived expressions [as provided by Eqs. (8), (12), and (20)] by considering three different combinations of fading parameter ‘ K ’ and ‘ m ’ is shown Table 1. It shows that the derived expressions only need 26 terms in the summation series to converge to an accuracy of 10^{-4} for all the three considered cases.

Figures 1 and 2 demonstrate the results for outage probability (P_{out}) versus average SNR by using expression (19) for combinations of (K, γ_0) and (m, γ_0) respectively with the constant total average fading power (AFP) of 7.78 dB. Figure 1 illustrates the plot for outage probability for several values of K and γ_0 with constant m , i.e., $m = 3$. From Fig. 1, we observe that the increase in K and/or a decrease in γ_0 , leads to a decrease in the outage probability.

Figure 2 illustrates the results for outage probability versus average SNR for several values of m and γ_0 with a constant K factor = 2 dB. It is observed that there is a noteworthy decrease in the outage probability as γ_0 decreases. Also, when the fading severity parameter, ‘ m ’ increases, the outage probability decreases. From the

Table 1 Number of terms required for convergence of the capacity for OPRA, ORA, and TIFR for different cases to achieve accuracy of 10^{-4} for all considered cases

Parameters	OPRA	ORA	TIFR
$K = 1, m = 1$	9	9	8
$K = 2, m = 2$	16	16	16
$K = 3, m = 3$	26	25	25

Fig. 1 Outage probability against average SNR for K, m

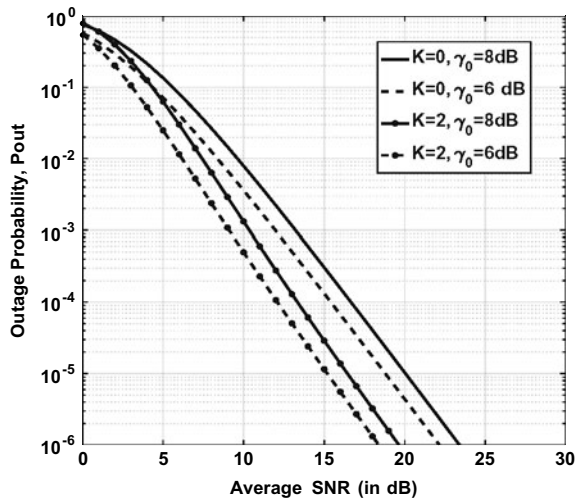


Fig. 2 Outage probability versus average SNR for m and γ_0

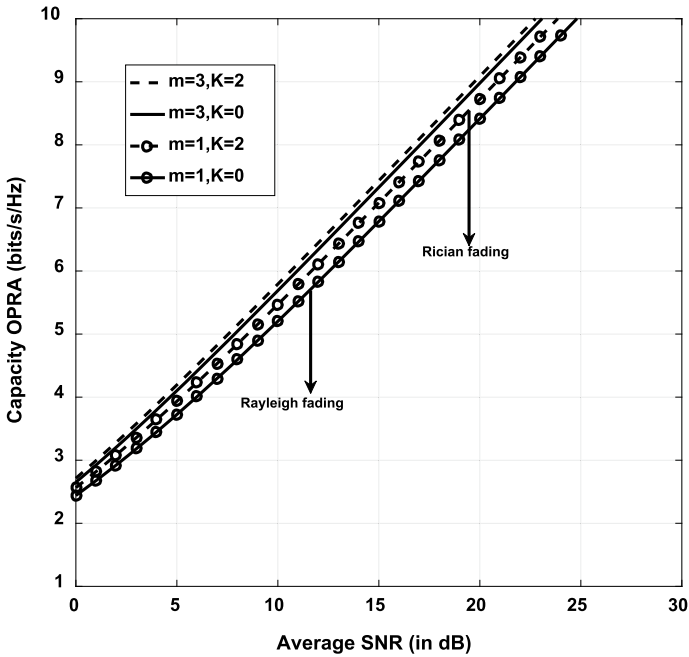
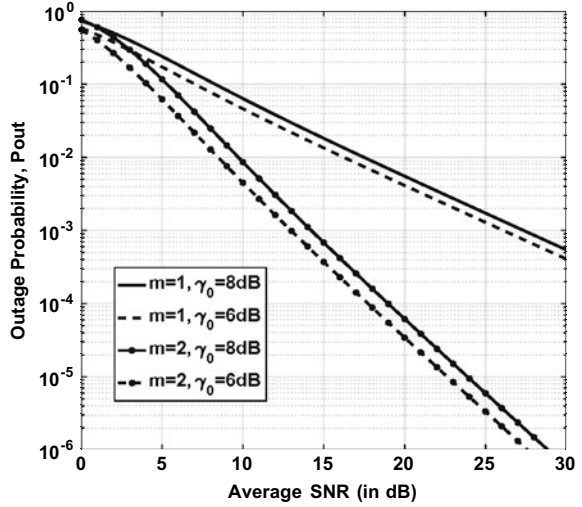


Fig. 3 Capacity OPRA versus average SNR

above observations, it is interpreted that the outage probability curves shift to right as the threshold increases and all the curves share the same behavior as anticipated.

Figures 3 and 4 show the behavior of the channel for the capacity OPRA and ORA against average SNR respectively with constant total AFP of 7.78 dB. To make the figures more legible, we have shown the results only for four several combinations of K and m . It is found that the capacity improves by increasing fading parameter, K and/or m . The curve with $K = 0, m = 1$ shows the result for Rayleigh fading which is special case of the BX fading. The curve with $K = 2, m = 1$ shows the results for Rician fading as special case of the BX fading. From the above analysis, it can be concluded that the derived results are closely matched with the existing fading models.

By using Eqs. (15) and (20), the results plotted for both CIFR and TIFR versus average SNR, respectively as shown in Fig. 5. In case of CIFR, a huge loss of capacity is shown for Rayleigh fading. Figure also illustrated that the spectral efficiency achieved by using CIFR technique and can be improved by using TIFR technique in which a cut-off SNR is used to increase the capacity at the cost of an increased outage probability.

By varying average SNR, the results for the capacity TIFR against the threshold SNR are plotted in Fig. 6. It is observed that the cut-off SNR that maximize the capacity increases as average SNR increases. The curves for comparison of the capacity for different adaptive transmission techniques are plotted in Fig. 7 for $K, m = 2, 7$ with AFP of 3.01 dB. By observing the plot, it is clear that the performance of channel capacity with OPRA and ORA techniques is better than CIFR and TIFR.

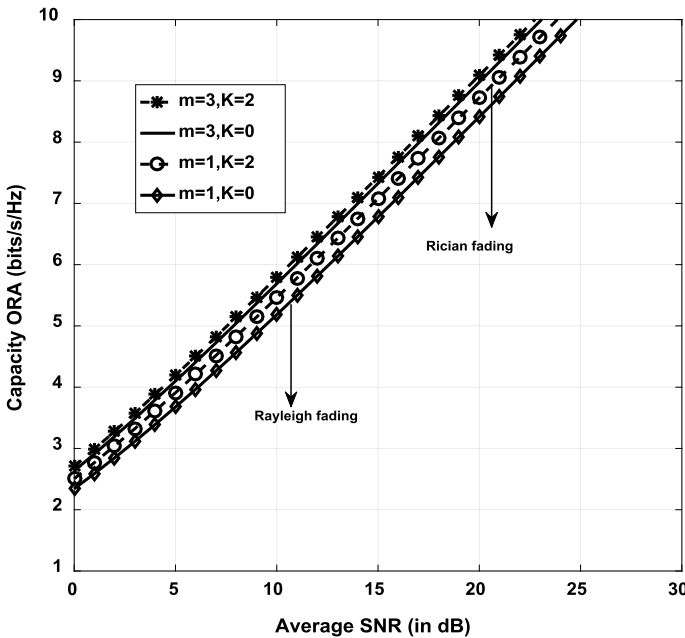


Fig. 4 Capacity ORA against average SNR over BX fading channel

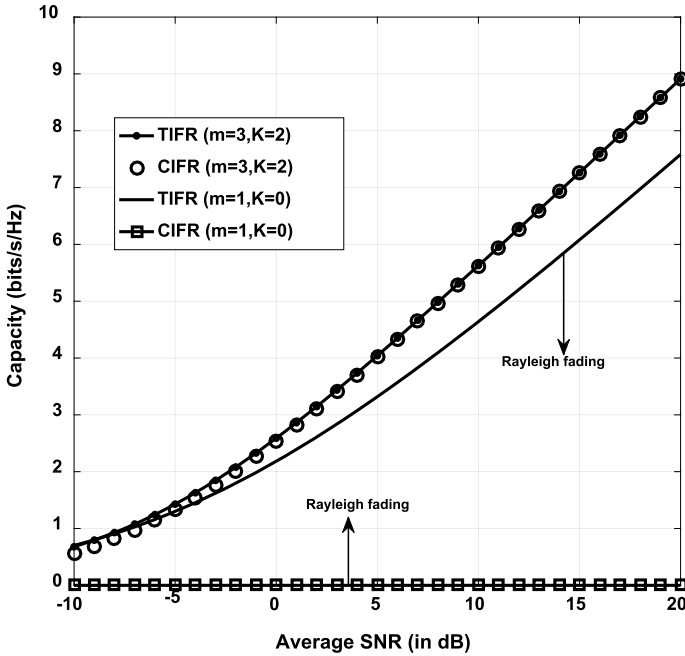


Fig. 5 Capacity for both CIFR and TIFR versus average SNR over BX fading channel

Table 2 shows the variation in capacity with various adaptation transmission techniques for different value of average SNR by keeping constant AFP of 6.02 dB and fading parameters ($K = 2, m = 3$). From the table, it can be easily deduced that the capacity of the channel increases by using OPRA technique as compared to other given adaptive techniques.

5 Conclusion

This work presented the analysis of the BX fading with adaptive transmission protocols via formulated the closed-form expressions for OPRA, ORA, CIFR, and TIFR. The effect of both the independent fading parameters, specular power (K) and fading severity (m) has been studied by plotting the capacity. From the results, it is observed that the capacity increases with reduction in the fading severity and/or increase in the specular power. The obtained numerical results for special cases of BX fading are compared with the fading channels already exist in literature such as Rayleigh and Rician fading channel. The final expressions are in the form of simple mathematical functions and conducive to further analysis, especially in the cases of 5G networks where the cells are highly dense.

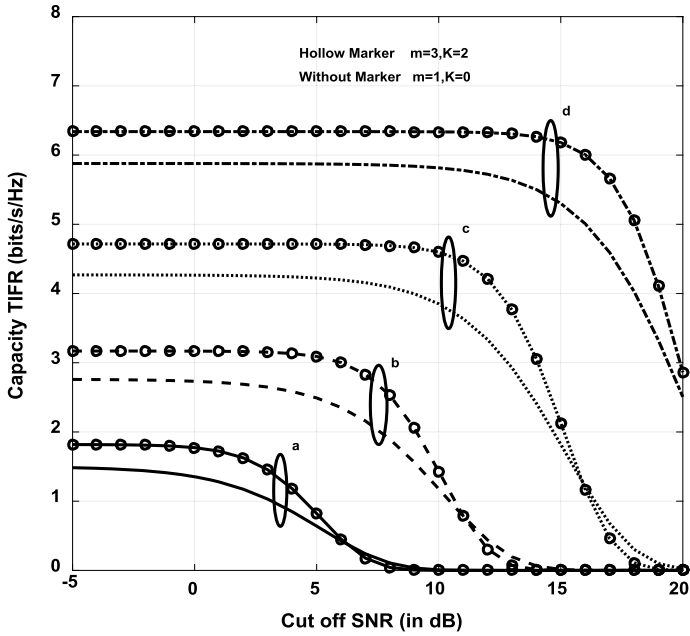


Fig. 6 Capacity TIFR versus Cutoff SNR for constant AFP = 3.01 dB **a** $\bar{\gamma} = 5$ dB, **b** $\bar{\gamma} = 10$ dB, **c** $\bar{\gamma} = 15$ dB, **d** $\bar{\gamma} = 20$ dB

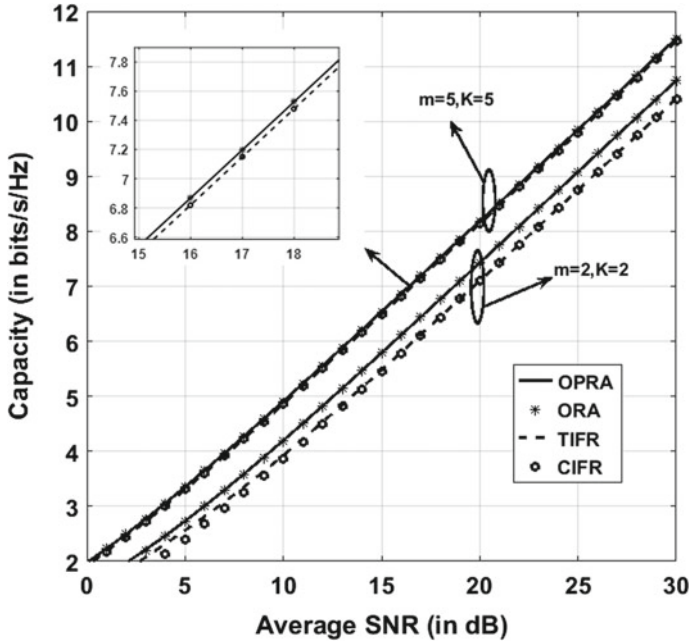


Fig. 7 Comparison plot for capacity with various adaptive techniques over the BX fading channel

Table 2 Variation in capacity with adaptive transmission techniques for $K = 2, m = 3$ and constant AFP of 6.02 dB at different average SNRs

Parameters	OPRA	ORA	CIFR	TIFR
Average SNR = 0 dB	2.247	2.235	2.069	2.136
Average SNR = 5 dB	3.652	3.650	3.474	3.488
Average SNR = 10 dB	5.222	5.211	5.043	5.046
Average SNR = 15 dB	6.852	6.848	6.674	6.675

Acknowledgements This work is funded by Visvesvaraya Ph.D. Scheme for Electronics & IT, MeitY, Govt. of India under grant no. VISPHD-MEITY-1674.

References

1. Q.Y. Liu, M. Wang, Z.D. Zhong, Statistics of capacity analysis in high speed railway communication system. *Tamkang J. Sci. Eng.* **14**(3), 209–215 (2011)
2. Y. Dong, P. Fan, K.B. Letaief, High-speed railway wireless communications: efficiency versus fairness. *IEEE Veh. Technol.* **63**(2), 925–930 (2014)
3. N.C. Beaulieu, X. Jiandong, A novel fading model for channels with multiple dominant specular components. *IEEE Wirel. Commun. Lett.* **4**(1), 54–57 (2014)
4. A. Olutayo, J. Cheng, J.F. Holzman, Performance bounds for diversity receptions over a new fading model with arbitrary branch correlation. *EURASIP J. Wirel. Commun. Network.* **2020**(1), 1–26 (2020)
5. A. Olutayo, J. Cheng, J.F. Holzman, Asymptotically tight performance bounds for equal-gain-combining over a new correlated fading channel, in *15th Canadian Workshop on Information Theory (CWIT)*, pp. 1–5 (2017)
6. A. Olutayo, J. Cheng, J.F. Holzman, Asymptotically tight performance bounds for selection diversity over Beaulieu-Xie fading channels with arbitrary correlation, in *IEEE International Conference on Communications (ICC)*, pp. 1–6 (2017)
7. A. Olutayo, H. Ma, J. Cheng, J.F. Holzman, “Level crossing rate and average rate duration for the Beaulieu-Xie fading model. *IEEE Wirel. Commun. Lett.* **6**(3), 326–329 (2017)
8. V. Kansal, S. Singh, Analysis of effective capacity over Beaulieu-Xie fading model, in *Proceedings of IEEE 4th International WIECO—ECE*, pp. 207–210, Dehradun, India
9. D. Singh, H.D. Joshi, Generalized MGF based analysis of line-of-sight plus scatter fading model and its applications to MIMO OFDM systems. *AEU-Int. J. Electron. Commun.* **91**, 110–117 (2018)
10. V. Kansal, S. Singh, Analysis of binary PSK modulations over the line-of-sight plus scatter fading model, in *Data and Communication Networks, Advances in Intelligent system and Computing*, pp. 13–20 (2019)
11. V. Kansal, S. Singh, Analysis of Average Symbol Error Probability of MDPSK MFSK and MPSK in the Beaulieu-Xie Fading, in *6th Edition of International Conference on Wireless Networks & Embedded Systems (WECON)*, pp. 11–14 (2019)
12. J. Hu, Z. Zhang, J. Dang, L. Wu, G. Zhu, Performance of decode-and-forward relaying in mixed Beaulieu-Xie and w dual-hop transmission systems with digital coherent detection. *IEEE Access* **7**, 138757–138770 (2019)
13. V. Kansal, S. Singh, Capacity analysis of maximal ratio combining over Beaulieu-Xie fading. *Ann. Telecommun.*, pp. 1–8 (2020)

14. M.S. Alouini, A.J. Goldsmith, Capacity of Rayleigh fading channels under different adaptive transmission and diversity-combining techniques. *IEEE Trans. Veh. Technol.* **48**(4), 1165–1181 (1999)
15. X. Li, X. Chen, J. Zhang, Y. Liang, Y. Liu, Capacity Analysis of $\alpha - \eta - \kappa - \mu$ fading channels. *IEEE Commun. Lett.* **21**(6), 1449–1452 (2017)
16. T. Aldalgamouni, A.M. Magableh, S. Mater et al., Capacity analysis of $\alpha - \eta - \mu$ channels over different adaptive transmission protocols. *IET Commun.* **11**(7), 1114–1122 (2017)
17. H. Al-Hmood, H.S. Al-Raweshidy, Unified analysis of channel capacity under different adaptive transmission policies. *Electronics Letters.* **56**(2), 87–89 (2020)
18. S.K. Yoo, P.C. Sofotasios, S.L. Cotton, S. Muhaidat, F.J. Lopez-Martinez, J.M. Romero-Jerez, G.K. Karagiannidis, A comprehensive analysis of achievable channel capacity in F composite fading channels. *IEEE Access* **7**, 34078–34094 (2019)
19. H. Zhao, L. Yang, A.S. Salem, M.S. Alouini, Ergodic capacity under power adaptation over Fisher Snedecor F fading channels. *IEEE Commun. Lett.* **23**(3), 546–549 (2019)
20. H. Zhao, Z. Lui, M.S. Alouini, Different power adaptation methods over fluctuating two-ray fading channels. *IEEE Commun. Lett.* **8**(2), 592–595 (2019)
21. D. Pant, P.S. Chauhan, S.K. Soni, S. Naithani, Channel capacity analysis of wireless system under ORA scheme over $\kappa - \mu$ /inverse gamma and $\eta - \mu$ /inverse gamma composite fading models. in *International Conference on Electrical and Electronics Engineering (ICE3)*, pp. 425–430 (2020)
22. I.S. Gradshteyn, I.M. Ryzhik, *Table of Integrals, Series and Products*, 5th edn. (Academic, San Diego, CA, 1994)