Bibhu Prasad Swain
Uday Shanker Dixit   *Editors*

Recent Advances in Electrical and Electronic Engineering

Select Proceedings of ICSTE 2023

Springer

Lecture Notes in Electrical Engineering

Volume 1071

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. **

Bibhu Prasad Swain · Uday Shanker Dixit
Editors

Recent Advances in Electrical and Electronic Engineering

Select Proceedings of ICSTE 2023

Springer

Editors
Bibhu Prasad Swain
Department of Physics
National Institute of Technology Manipur
Imphal, India

Uday Shanker Dixit
Department of Mechanical Engineering
IIT Guwahati
Guwahati, India

Preface

Recent Advances in Electrical and Electronic Engineering: Select Proceedings of ICSTE 2023 comprises a collection of research and review articles presented at the 1st International Conference on Sci/Tech and Engineering (ICSTE 23). The present volume comprises select proceedings from electrical electronics, electronics and communication engineering and computer science and engineering presentations. The major topics in computer science and engineering are digital recognition of hand gestures, artificial intelligence, image spam classification, textual content analysis for emotion identification, artificial neurons, smart healthcare, and machine learning. The major topics from electrical engineering include gallium nitride (GaN) metal-oxide-silicon field-effect transistors (MOSFET)-based Class E converter, high-power wireless power transfer (WPT) systems, real-time simulation of solar photovoltaic systems, and floating inductance simulator. Promising topics such as microstrip and co-planar waveguide, wireless sensor network, deep learning machine intelligence models, algorithm analysis in non-orthogonal multiple access (NOMA), paper-based capacitive sensors, UV photo-detection and electronically tunable oscillator are presented from electronics and communication engineering.

Eminent speakers like Prof. Manoj Gupta, National University of Singapore, Singapore; Prof. Bharat Dahiya Thammasat University, Thailand; Prof. Surender Kumar, Ex. Prof. (BIT Mesra), India; Prof. Sawan Suman Sinha, IIT Delhi, India; Prof. Abhishek Dey, Indian Association for the Cultivation of Science, Kolkata, India; Prof. Bimlesh Kumar, IIT Guwahati, India; and Dr. Sanjeet Kumar Dwivedi, Green Hydrogen-based Danish MNC, Denmark, shared their knowledge and experience. The conference was attended and enriched by participants from numerous institutes such as IITs, NITs, NEHU, BIT, VIT, MIT Manipur, Manipur University, IIEST Kolkata, and abroad; several speakers deliberated on their research works. In addition, the paper presentations were accompanied by six keynotes address from leading academic and industrial researchers around the globe. Out of 268 papers, 138 papers were selected for publication in the ICSTE 23 Proceedings. The paper presentations took place on eight different tracks with parallel main sessions. On the ICSTE 23 platform, we got the opportunity to promote the national campaign "G20" in 2023, which is being observed under the chairmanship of India.

The review committee has done an excellent job in reviewing articles and approving high-quality research articles to be published in the conference proceedings. The editors are thankful to all the faculty members and students of various committees for their dedication in making it a very successful conference. Thanks to the editing and printing support staff of Springer for making the publication of this book possible. We sincerely hope that it will inspire researchers/scientists to explore the potential in the diverse upcoming research fields.

Imphal, India   Dr. Bibhu Prasad Swain
                                                              bpswain@nitmanipur.ac.in

Guwahati, India   Dr. Uday Shanker Dixit
                                                                    uday@iitg.ac.in

Contents

Computer Science and Engineering

About the Editors

Dr. Bibhu Prasad Swain is currently an associate professor at the Department of Physics, National Institute of Technology Manipur. He obtained his B.Sc. (Physics) from Utkal University, Bhubaneswar; M.Sc. (Physics) from the National Institute of Technology Rourkela; M.Tech. (Materials Science) from Barkatullah University; and Ph.D. from the Indian Institute of Technology Bombay. His major areas of research interests include large bandgap semiconductors, mechanical hard materials, nanostructured materials, graphene-based supercapacitor, photovoltaic materials, biocompatible coating, and modeling of advanced MOSFET. He has published 150 papers in respected international journals and four edited books. He received the Brain Korea 21 Fellow and Japanese Society of promotion of Science Postdoctoral Fellow from Materials Research, Seoul National University, and Government of Japan in the years 2007 and 2008. Currently, he is an editorial board member of the *Nanoscience and Nanotechnology—Asia* and the guest editor in *Materials* journals in MDPI publications.

Dr. Uday Shanker Dixit is currently a senior professor (HAG) at the Department of Mechanical Engineering, Indian Institute of Technology Guwahati. He obtained his B.Tech. (Mechanical Engineering) from the Indian Institute of Technology Roorkee, M.Tech. (Mechanical Engineering) from the Indian Institute of Technology Kanpur, and Ph.D. from the Mechanical Engineering, Indian Institute of Technology Kanpur. His major areas of research interest include plasticity, metal forming, laser-based manufacturing, and finite element modeling and optimization. He has published 146 papers in reputed journals and authored/edited 19 books. He is an honorary fellow member of the Indian Welding Society and felicitation by the National Coordinator of NPTEL for getting larger number of viewers of his course on Engineering Mechanics. He was the director of CIT, Kokrajhar (2014–2015), a BOG member of many institutes, and a guest editor for many special issues of reputed journals.

Electrical Engineering and Electronics & Communication Engineering

Cost-Effective Solution of Grid Dispatch Problems Using a Novel Rao-4 Algorithm

Shuvam Sahay, Ramanaiah Upputuri, Pooja Kumari, and Niranjan Kumar

1 Introduction

Four different issues covered under grid dispatch (GD) are the economic operation of power plants, the minimum rate of emission, minimum load voltage deviation, and improved voltage stability. The economic operation of power plants means the production of electrical energy at minimum fuel cost subjected to linear and nonlinear constraints. It is considered an optimization problem that is formed to minimize the fuel cost by regulating the active and reactive power flow within a prescribed limit and also fulfill the load demand balance. This process is carried out by either including the valve-point effect or excluding it. Nowadays, we are completely dependent on the thermal power plant due to continuously increasing power demand. Thermal power plant requires the opening and closing of steam valve for regulating the steam input to the turbine for controlling the speed of the generator. Therefore, it creates an impact on the generation of electrical energy indirectly [1]. Coal is one of the important fuels used in the thermal power plant for generating steam. It emits harmful gases like sulfur dioxide, nitrogen dioxide, and carbon dioxide into the atmosphere. It leads to ozone layer depletion and rise in the level of sea due to melting of ice. Therefore, it is important to reduce the emission level under permissible limit. It is treated as an optimization problem in which the emission level is expressed in terms of the active power output [2]. Transmission lines and distribution lines require voltage management at various stages to maintain allowed voltage levels at the final consumer's location. If the voltage of the system deviates from its nominal value, the performance and lifespan of the device deteriorate. Motors operating below normal voltage draw significant currents and may overheat, even with the rated horsepower

S. Sahay (✉) · R. Upputuri · P. Kumari · N. Kumar
Department of Electrical Engineering, National Institute of Technology, Jamshedpur, India
e-mail: 2019rsee010@nitjsr.ac.in

load. Actual line losses depend on both reactive and real line power flow. Line-end voltages affect reactive line flow. It is mandatory to reduce the load voltage deviation within the prescribed limit [3]. Because of the rise in power consumption, the power system is now operating much closer to the threshold at which it can remain stable. As a result of improvements made to the power system through a variety of methods, voltage instability and line overloading have emerged as significant challenges. The unbalanced reactive power that results when a power system is under stress is one of the major contributors to voltage instabilities. Voltage instability may cause load shedding in a region, or tripping of transmission lines and other equipment by their protective systems, resulting in cascade outages. The voltage stability index is the parameter that indicates the ability of a particular system to withstand voltage instability [4].

The GD problem has been solved using a variety of standard optimization techniques such as linear programming (1979), Newton methods (1992), interior point methods (1998), and dynamic programming (2001). They do not provide any assurance that the global optimal solution will be found, and the computations required to implement them are complex and time-consuming. When it comes to discrete variables, they just are not equipped to handle them [5].

Numerous powerful metaheuristics have been created in recent years. Some of these have been applied with remarkable success to the GD problems. Recent illustrations of metaheuristics being used for GD issues are the squirrel search algorithm (SSA) [6], ant colony optimization (ACO) [7], bat algorithm (BA) [8], teaching learning-based algorithm (TLBO) [9], crow search algorithm (CSA) [10], harmony search algorithm (HSA) [11], grasshopper optimization (GO) [12], non-sorting genetic algorithm (NSGA) [13], mine blast algorithm (MBA) [14], group search algorithm (GSA) [15], whale optimization algorithm (WOA) [16], sine–cosine algorithm (SCA) [17], backtracking search optimization (BSA) [3], and rot tree optimization (RTO) [4]. But because there are so many different goals that need to be met to solve GD problems, there is no single algorithm that is best suited to doing so. A new algorithm that can effectively address most of the GD problems is thus always needed. These algorithms require some specific control parameters. Researchers feel difficulty in tuning these control parameters to get an accurate solution. Power system researchers feel difficulty in implementing these techniques for solving these complex power system issues. They may converge to local optimum value and solutions given by these techniques possess a lesser degree of precision [18].

Therefore, this paper proposes a technique that is not inspired by any animals, insects, birds, musical instruments, chemical, and physical reactions, etc. It requires only common control parameters like population size and termination criteria. It is very easy for researchers to implement these techniques for complex systems. Additionally, in addition to the basic OPF problem being examined in this study, several sophisticated formulations with non-convex cost functions and other goal functions are also taken into account, along with the environmental concern requiring emission reduction.

The remaining sections of this research work are structured as follows. In Sect. 2, the GD formulation is briefly presented. Following that, Sect. 3 presents the description of the proposed method known as Rao-4. Following that, Sect. 4 discusses the outcomes of applying Rao-4 method to solve various GD problem scenarios. Finally, conclusions are drawn.

2 Problem Formulation

The aim of grid dispatch problem (GD) is to determine the finest settings of control variables that reduce the cost of active power generation, transmission losses, and load voltage deviation and maintain higher voltage stability within system's operational boundaries. It can be represented mathematically [8].

$$\text{Minimize } m\,(p,\,q) \tag{1}$$

$$\text{Subject to } e\,(p,\,q)\,=\,0; \tag{2}$$

$$\text{And } i\,(p,\,q)\,\leq\,0; \tag{3}$$

where p and q are the set of dependent and independent variables e and i are set the equality and inequality constraints.

2.1 Control Variables

These are the load flow equations, and these are the sets of variables that may be changed to make them satisfied. The OPF issue formulation includes a set of control variables that are as follows [1].

Hence, p can be expressed as

$$p = \left[P_{G1} \ldots P_{Gnpv},\, V_{G1} \ldots V_{Gnpv},\, Q_{C1} \ldots Q_{Cnqc},\, T_1 \ldots T_{nt} \right], \tag{4}$$

where *npv*: Number of generator bus; *nqc*: Number of shunt capacitor; *nt*: Number of transformer tapping. P_G: Active power output; V_G: Generator voltage; Q_C: Shunt compensation; T: Tap value.

2.2 Constraints

There are mainly two types of constraints defined for GDPs which are given below.

2.2.1   Equality Constraints

Given below is a set of constraints that, when satisfied, will encapsulate the standard nonlinear power flow equations that are used to regulate the power system [10].

$$\sum P_i = P_d + P_l, \tag{5}$$

$$\sum Q_i = Q_d + Q_l, \tag{6}$$

where P_i and Q_i: Input power; P_d and Q_d: Power demand; P_l and Q_l: Power loss

2.2.2   Inequality Constraints

A set of limitations that represent the system's operational and physical boundaries are described in the following way [12]:

$$V_{Gi}^{\min} \leq V_{Gi} \leq V_{Gi}^{\max}, \quad i = 1, \ldots, N_G, \tag{7}$$

$$Q_{Gi}^{\min} \leq Q_{Gi} \leq Q_{Gi}^{\max}, \quad i = 1, \ldots, N_{G,} \tag{8}$$

$$T_i^{\min} \leq T_i \leq T_i^{\max}, \quad i = 1, \ldots, N_T, \tag{9}$$

$$P_{Gi}^{\min} \leq P_{Gi} \leq P_{Gi}^{\max}, \quad i = 1, \ldots, N_{G,} \tag{10}$$

$$Q_{shi}^{\min} \leq Q_{shi} \leq Q_{shi}^{\max}, \quad i = 1, \ldots, N_{C,} \tag{11}$$

$$V_{\text{load}i}^{\min} \leq V_{\text{load}i} \leq V_{\text{load}i}^{\max}, \quad i = 1, \ldots, N_L, \tag{12}$$

$$P_{LFi} \leq P_{LFi}^{\max}, i = 1, \ldots, N_l. \tag{13}$$

The violation in the above limits is incorporated into the objective function by introducing a quadratic penalty term as shown in Eq. (18). These terms are evaluated according to Eqs. (15), (16), and (17), respectively [13]. The selection of the value of the penalty factor is taken from [3].

$$\Delta V_{\text{Load}} = V_{\text{Load}}^{\min} - V_{\text{Load}}\big) \text{if } V_{\text{Load}} < V_{\text{Load}}^{\min}$$
$$\& \ \Delta V_{\text{Load}} = V_{\text{Load}}^{\max} - V_{\text{Load}}\big) \text{if } V_{\text{Load}} > V_{\text{Load}}^{\max} \tag{14}$$

$$\Delta Q_G = Q_G^{\min} - Q_G \big) if\ Q_G < Q_G^{\min}\ \&\ \Delta Q_G = Q_G^{\max} - Q_G \big) if\ Q_G > Q_G^{\max}$$
(15)

$$\Delta S_f = S_f^{\max} - S_f \big) if\ S_f^{\max} < S_f.$$
(16)

2.3 Objective Function

This article takes into consideration a number of different objective functions, including cost reduction (with valve-point effect), improvement of voltage profile, enhancement of voltage stability, and reduction of emissions. In addition to that, a number of different combinations of these objective functions are also taken into consideration [14].

$$\text{penalty} = \mu \sum_{i=1}^{vnpq} \Delta V_{\text{Load}i}^2 + \psi \sum_{i=1}^{vnpv} \Delta Q_{Gi}^2 + \phi \sum_{i=1}^{vnl} \Delta S_{Li}^2$$
$$L = L_{\text{obj}} + \text{penalty}$$
(17)

3 Rao-4 Algorithm

Rao-4 algorithm is a metaphor-less algorithm and independent of any definite parameters. It demands parameters like population size and iteration count only. The finest outcome given by this method depends on the interaction of the present solution with both the best and worst solution simultaneously but in previous three versions present value interacts with the best value or worst value including randomly picked solution [19].

The following is an outline of the primary phases involved in carrying out an optimization of an objective function.

P-1: Clarify the population size, number of control variables and their upper and lower limits and the criteria for termination.

P-2: Evaluate the objective function from randomly selected initial population.

P-3: Find out the best and worst solution from the present population on the basis of their function value. For minimization problem, the best solution corresponds to the lowest function value, whereas the worst solution corresponds to the highest function value.

P-4: Find out the present solution during the ith iteration for complete population. Suppose $H(i, j)$ is the value of the jth variable during the ith iteration, $b(j)$ is the best value of the jth variable, $w(j)$ is the best value of the jth variable, and H_{new}

(i, j) is the new value of the jth variable during the ith iteration. Hence, the new value can be calculated using the equation given below

$$H_{\text{new}}(i, j) = H(i, j) + r1 * (b(j) - w(j)) + 0.5 * (r2 * (w(j) - H(i, j))) \\ + r3 * (b(j) - H(i, j))) - r4 * (w(j) - H(i, j)), \qquad (18)$$

where $r1$, $r2$, $r3$, and $r4$ are the random lies in the range [0, 1].

P-5: Evaluate the objective function for updated value of control variables and check whether it is higher or smaller than the old solution. If it is higher, reject the present solution. If it is smaller, accept the present solution. If the criteria for stopping the process are met, stop the process and report the finest solution. If not, repeat P-3.

4   Simulation and Results

The Rao-4 algorithm has been utilized on an IEEE 30-bus system in this paper. In addition to this, seven additional case studies are studied. These cases begin with the most fundamental cases and then progress through increasing levels of complexity while taking into account a variety of other objectives.

IEEE 30-bus system

It has a total load demand of 283.40 MW. It has a total of 30 buses, 41 branches, 6 alternators, 4 transformers tapping, and 2 shunts. The buses 1, 2, 5, 8, 11, and 13 all have alternators linked to them. After finding the location of weak buses in the aforementioned testing system, it is determined that the shunt would be connected to buses 10 and 24. The 11, 12, 15, and 36 bus lines all have connections to transformers [8]. Cost and emission coefficient employed in this paper is taken from [1].

C-1: Optimal power flow (OPF)-based cost reduction

This study begins its investigation with the base case, which entails finding the lowest possible value for a quadratic function that represents the amount of money spent on fuel for power generation without consideration of valve-point effect. The mathematical expression for this scenario is given below.

$$C_t = \sum_{i=1}^{Ng} \alpha_i + \beta_i P_{gi} + \gamma_i P_{gi}^2 + \text{penalty}, \qquad (19)$$

where α_i; $\beta_i \gamma_i$ are the fuel cost coefficient of the ith generator [8]. The proposed Rao-4 method is applied to this scenario, and the finest control variables that are obtained through this technique are reported in Table 1. The optimal generation fuel cost that is obtained is found to be 798.043 dollars per hour. In addition, a rough sketch of reduction process of this function is presented in Fig. 1.

Table 1 Optimal value of control variables and its corresponding objective function value

	C-1	C-2	C-3	C-4	C-5	C-6	C-7
P-1	175.560	169.400	163.919	193.676	184.7932	190.2804	115.6753
P-2	49.7317	46.220	43.236	45.5109	23.4484	49.5328	55.7087
P-3	21.1198	23.020	26.727	19.8451	27.917	17.2312	28.2089
P-4	18.7782	17.000	35.000	13.6477	26.8038	14.9298	34.9852
P-5	15.5861	21.080	10.298	10.009	15.9447	10.1993	26.4183
P-6	12.0588	16.010	12.000	12.0465	14.2516	12.4894	28.3035
V-1	1.1000	1.039	1.0998	1.0581	1.033	1.0685	1.0944
V-2	1.0082	0.985	1.0872	1.0374	0.9512	1.0111	1.0818
V-3	1.0532	1.006	1.0475	1.0049	1.0146	1.0312	1.0502
V-4	1.0450	1.007	1.0999	1.1	1.0112	1.094	1.0963
V-5	1.0574	1.022	1.0674	1.0609	1.0827	1.046	0.95
V-6	0.9505	1.030	1.0194	0.9537	0.9965	0.977	0.9521
T-1	0.9851	0.933	0.9165	1.0996	0.9778	0.989	1.1
T-2	1.0860	1.012	0.9187	1.0571	1.0114	0.9017	0.932
T-3	0.9292	0.996	0.9001	0.9089	0.9486	0.9314	1.0692
T-4	1.0340	0.959	0.9005	0.985	0.9435	0.9008	1.064
Qc-1	4.0327	4.420	4.2233	2.0993	1.9233	2.5809	5
Qc-2	0.0000	3.504	0	4.7406	3.4104	2.7881	4.4361
Qc-3	3.6155	1.136	3.0061	0.039	2.1534	3.1165	1.3786
Qc-4	4.2111	3.429	4.9952	1.2743	1.5045	5	2.015
Qc-5	0.0009	3.281	0	1.6849	4.2657	0	0.072
Qc-6	0.0005	3.510	3.5356	4.9546	1.3937	4.7247	4.8415
Qc-7	0.2049	4.670	0	0.0168	3.1572	0	2.0388
Qc-8	4.9890	1.141	4.9865	4.4176	4.8428	1.6038	4.7711
Qc-9	0.4744	2.210	1.3821	0.2996	1.5295	3.1192	0.3893
Cost	**798.043**	**807.000**	**805**	**830.054**	**858.120**	**840.500**	**836.804**
VD	0.5066	**0.164**	2.2486	0.4688	**0.1497**	0.741	1.5072
L_{max}	0.1415	0.135	**0.1246**	0.1364	0.134	**0.1275**	0.1386
P_{loss}	9.4345	9.371	7.7793	11.3354	9.5785	11.2628	5.8999
Emission	0.3612	0.342	0.3323	0.4185	0.3862	0.4077	**0.2444**
Function Value	798.043	873.100	1428.363	830.054	918.957	1478.995	1056.764

C-2: OPF-based cost reduction and voltage profile improvement

One of the most crucial and essential indicators of the security and reliability of power systems is bus voltage. In case C-1, only the cost function is considered without regulating the load voltage from the standard value. Thus, by taking into
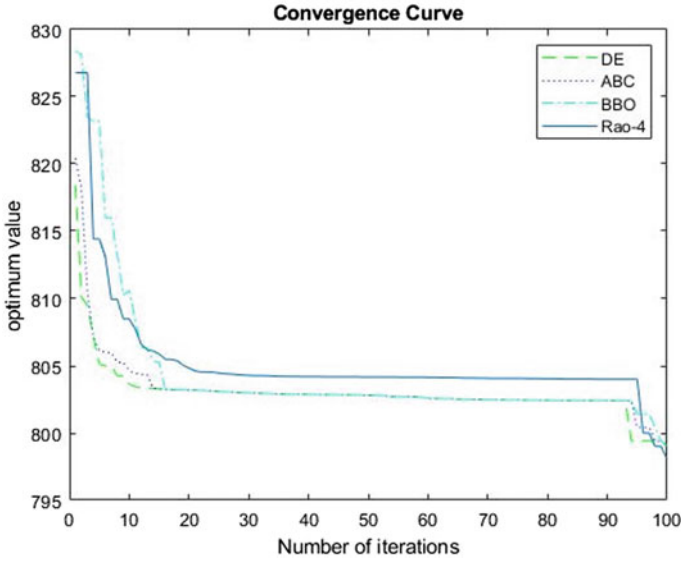
Fig. 1 Convergence curve for C-1

account a dual objective function, this example seeks to minimize fuel expense while maintaining a flatter voltage profile. The voltage profile is enhanced by decreasing the load voltage deviation from ideal value which is shown below.

$$F_{vd} = \sum_{l=1}^{N_{\text{load}}} \text{abs}(V_l - 1)/N_{\text{load}}. \tag{20}$$

The mathematical expression for this scenario is given below. [3]:

$$C_t = \left(\sum_{i=1}^{Ng} \alpha_i + \beta_i P_{gi} + \gamma_i P_{gi}^2 \right) + \delta_{\text{VD}} F_{\text{VD}} + \text{penalty}. \tag{21}$$

δ_{VD} is a scaling factor that is used to maintain the balance between both objective functions. The value of the scaling factor is selected on experimental trial basis. Here, we have considered the value of δ_{VD} is 400. The proposed Rao-4 method is applied to this scenario, and the optimal control variables that are acquired through optimization are reported in Table 1. The voltage deviation reduces from 0.5046 to 0.1640 as compared with C-1. In addition, a rough sketch of reduction process of this function is presented in Fig. 2. The voltage profile is improved further as compared with C-1 which is presented in Fig. 3.
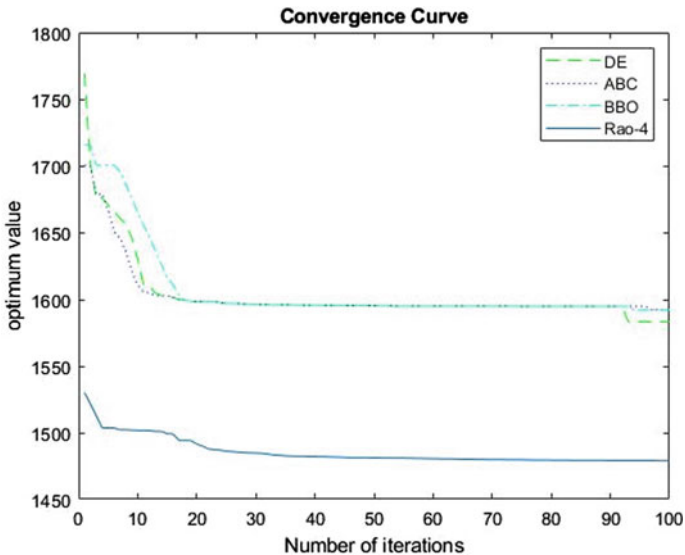
Fig. 2 Convergence curve for C-2

C-3: OPF-based cost reduction and voltage stability enhancement

It is important to examine the ability of a system to withstand any voltage instability. A voltage stability index L created by Kessel and Glavitch and characterized using local indicators L_k is provided by

$$L_k = [1 - \sum_{i \in N_g} \frac{(F_{ji} V_i \{\alpha_{ij} + (\delta_i - \delta_j)\})}{V_j}, \ L = \min[\max L_k], \qquad (22)$$

where F is evaluated by the partial inversion of impedance matrix which is given below.

$$F_{ji} = -[Y_{ij}]^{-1}[Y_{ji}]. \qquad (23)$$

The L indicator ranges from 0 to 1. It states that lower the value of L, system will be more stable.

The mathematical expression for this scenario is given below [3].

$$C_t = \left(\sum_{i=1}^{Ng} \alpha_i + \beta_i P_{gi} + \gamma_i P_{gi}^2 \right) + \delta_{VSI} L + \text{penalty}. \qquad (24)$$

δ_{VSI} is a scaling factor that is used to maintain the balance between both objective functions. The value of the scaling factor is selected on an experimental trial basis.
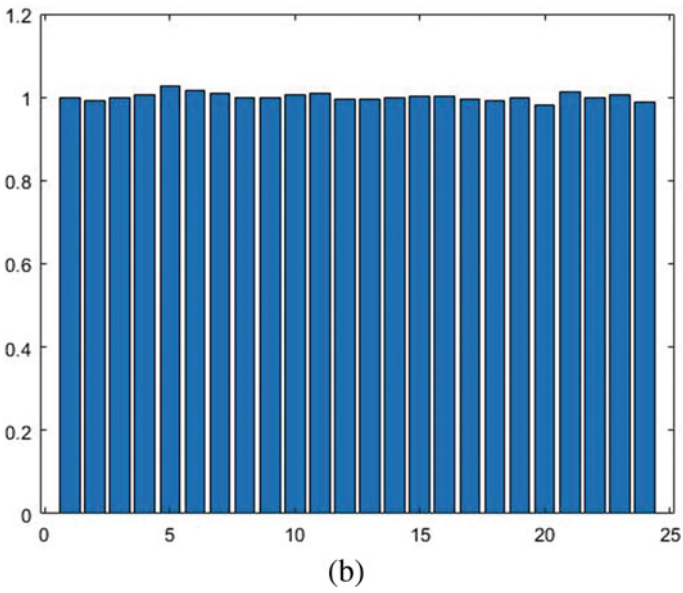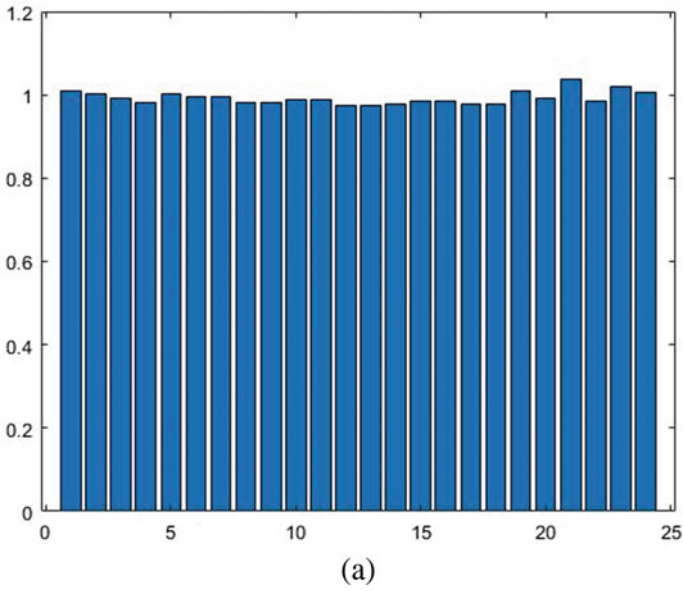
Fig. 3 **a** Voltage profile for C-1, **b** voltage profile for C-2
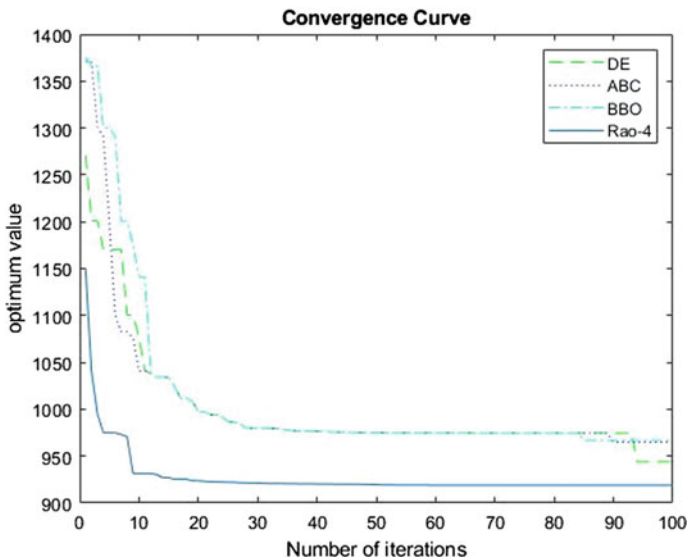
Fig. 4 Convergence curve for C-3

Here, we have considered the value of δ_{VSI} 5000. The proposed Rao-4 method is applied to this scenario, and the optimal control variables that are acquired through optimization are presented in Table 1. It has been observed that L reduces from 0.1415 to 0.1246 as compared with C-1. In addition, a rough sketch of reduction process of this function is presented in Fig. 4.

C-4: OPF-based practical cost reduction

The valve-point effect must be taken into consideration for a more accurate and effective modeling of generating cost function by adding repeating rectifying sinusoidal term to the cost function. The mathematical expression for this scenario is given below [3].

$$C_t = \sum_{i=1}^{Ng} \alpha_i + \beta_i P_{gi} + \gamma_i P_{gi}^2 + \text{abs}(\zeta_i \sin(\lambda_i (P_{gi}^{\min} - P_{gi}))) + \text{penalty}, \qquad (25)$$

where ζ_i and γ_i: Fuel coefficients; and $P_{gi}^{\min}$: Minimum value of active power of ith generating unit. The proposed Rao-4 method is applied to this scenario, and the finest control variables that are obtained through this technique are reported in Table 1. The optimal generation fuel cost that is obtained is found to be 830.054 dollars per hour. In addition, a rough sketch of reduction process of this function is presented in Fig. 5.
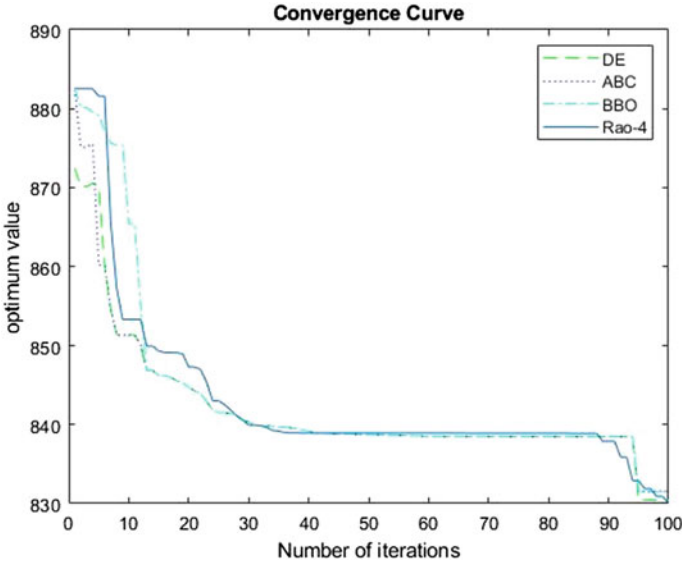
Fig. 5 Convergence curve for C-4

C-5: OPF-based practical cost reduction and voltage profile improvement

In this case, practical fuel cost is minimized along with the improvement in the load voltage profile. Table 1 provides the best outcomes for this situation utilizing Rao-4. In addition, a rough sketch of reduction process of this function is presented in Fig. 6.

C-6: OPF-based cost reduction and voltage stability enhancement

Similar to C-3, the goal in this scenario is to minimize fuel cost while taking the valve-point effect and the voltage stability index into account. The best outcomes in this situation, as determined by Rao-4, are listed in Table 1. In addition, a rough sketch of reduction process of this function is presented in Fig. 7.

C-7: OPF-based Cost reduction including emission

It is mandatory to reduce emissions in light of rising political concerns about the environment. The total ton/h emission of air pollutants from fossil fuel-powered thermal units, compounds of sulfur, nitrogen and carbon with oxygen, can be stated as

$$E_d = \sum_{j=1}^{Ng} 10^{-2}(\alpha_j + \beta_j P_{gj} + \gamma_j P_{gj}^2) + \eta_j e^{(\delta_j P_{gj})}, \tag{26}$$

where E_d is rate of emission (tons/hr.), α_j, β_j, Υ_j, η_j and δ_j are emission coefficients, and P_{gj} is active power output of jth generating unit.

The mathematical expression for this scenario is given below [3].
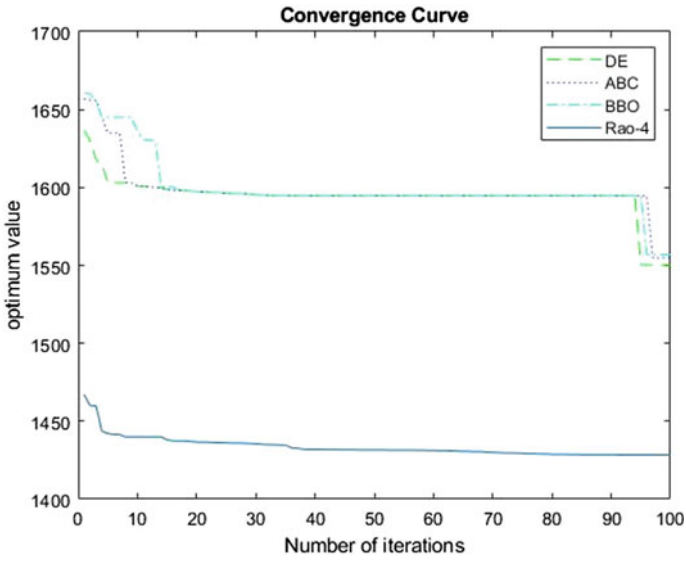
Fig. 6 Convergence curve for C-5
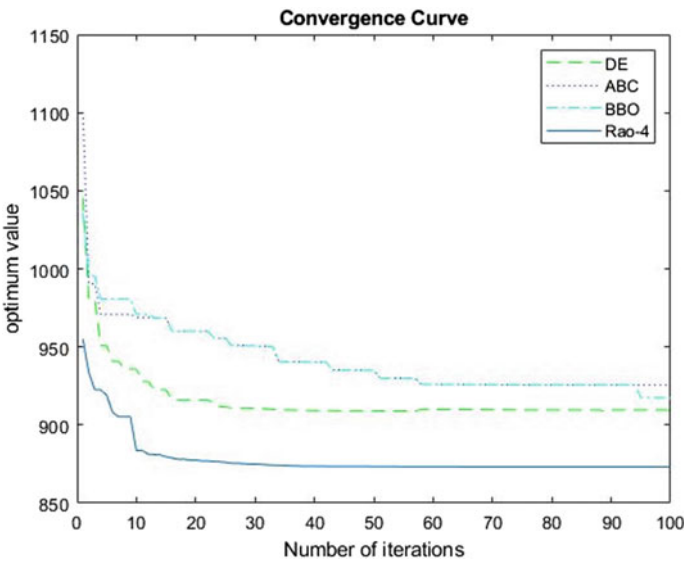
Fig. 7 Convergence curve for C-6
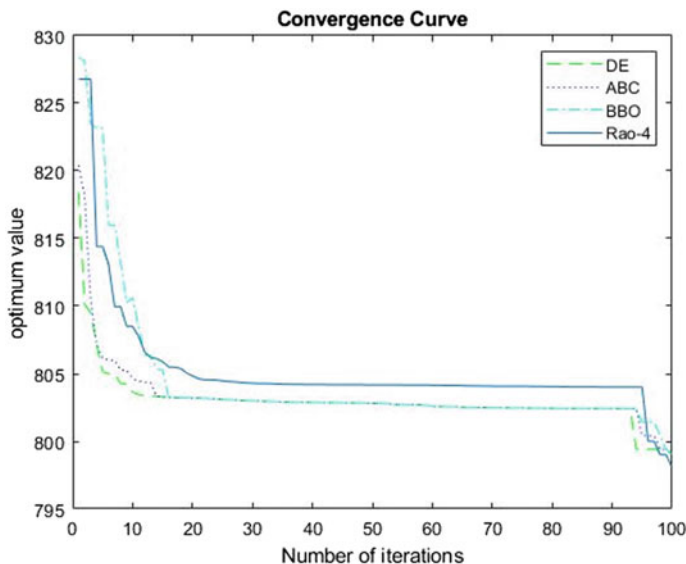
Fig. 8 Convergence curve for C-7

$$C_t = \left(\sum_{i=1}^{Ng} \alpha_i + \beta_i P_{gi} + \gamma_i P_{gi}^2 \right) + \delta_{em} E_d + \text{penalty}, \qquad (27)$$

where δ_{em} is a scaling factor which is used to maintain balance between both objective function. The value of the scaling factor is selected on experimental trial basis. Here, we have considered the value of δ_{em} is 900. The proposed Rao-4 method is applied to this scenario, and the optimal control variables that are acquired through optimization are reported in Table 1. In addition, a rough sketch of reduction process of this function is presented in Fig. 8.

5   Performance Evaluation Study

The Rao-4 has been contrasted to a number of optimization methods, including DE, ABC, and BBO, in order to assess its performance. Table 2 presents the findings of a statistical study. We enter five values—the minimum (or best), the mean (or average), the median, the maximum (or worst), and the standard deviation—into each cell that represents the performance of the proposed method in a particular scenario. This table shows how effective the Rao-4 is in solving a number of GD issues of various degrees of complexity. For all circumstances, the average values are quite close to the best values.

Table 2 Comparison of proposed techniques with other well-known optimization techniques

Case	Notation	Rao-4	DE	ABC	BBO
C-1	Minimum	798.0430	799.0376	799.0541	799.1267
	Mean	798.3180	799.3047	799.6945	801.1927
	Median	798.2250	799.0458	799.4613	801.1287
	Maximum	799.0460	801.5552	802.6327	803.1429
	Standard Deviation	0.2934	0.6624	0.8145	1.0251
C-2	Minimum	873.1000	909.08550	925.3049	917.2192
	Mean	879.6300	935.20760	944.2215	960.8904
	Median	879.5160	917.51730	943.3027	954.5105
	Maximum	887.8870	1068.5257	1006.1873	1080.675
	Standard Deviation	5.4003	42.4846	14.7217	37.754
C-3	Minimum	1428.363	1550.3449	1554.8035	1557.027
	Mean	1436.180	1555.5616	1565.6099	1566.820
	Median	1437.200	1553.4312	1565.8584	1566.800
	Maximum	1439.666	1578.0121	1574.1667	1593.783
	Standard Deviation	3.640	5.8042	5.9094	6.2569
C-4	Minimum	830.0540	830.4425	831.5783	831.4581
	Mean	832.3160	831.4997	834.4691	835.8153
	Median	832.2270	830.5125	834.1686	835.3208
	Maximum	834.4120	842.7195	839.0831	842.5715
	Standard Deviation	1.4871	3.0912	1.9432	2.6118
C-5	Minimum	918.9570	943.9238	964.6089	966.8975
	Mean	927.4350	975.6403	986.1737	1000.332
	Median	926.2025	958.2037	987.3366	989.3155
	Maximum	939.5480	1096.034	1012.5744	1110.738
	Standard Deviation	6.9130	40.579	11.5230	35.4894
C-6	Minimum	1478.995	1583.3209	1592.8876	1592.2125
	Mean	1484.150	1590.4131	1604.9151	1607.7481
	Median	1484.460	1589.8157	1605.4779	1607.0377
	Maximum	1490.270	1623.2834	1618.1911	1624.5010
	Standard Deviation	3.357	8.0913	7.3870	8.5617
C-7	Minimum	1056.764	1077.3561	1077.4222	1077.7629
	Mean	1058.070	1079.1245	1077.8749	1079.0579

<div align="right">(continued)</div>

Table 2 (continued)

Case	Notation	Rao-4	DE	ABC	BBO
	Median	1057.800	1077.4365	1077.7441	1079.1963
	Maximum	1059.784	1096.8296	1079.4747	1079.1963
	Standard Deviation	1.301	4.7910	0.4630	0.7013

6   Conclusion

In this study, a recently developed method known as the Rao-4 is used to solve a number of Grid dispatch issues. The performance of the suggested technique has been examined using the IEEE 30-bus system and seven cases. The findings are also contrasted with those produced by employing other recently developed optimization methods such as DE, ABC, and BBO. This research work states that Rao-4 algorithm is a highly efficient and robust algorithm for resolving grid dispatch problems. It does not converge to local optima. It does not converge prematurely. It is very simple and easy to implement this technique for complex problems. It is suggested that in future research, a multi-objective Rao-4 algorithm will be hybridized with other techniques to increase the search ability and used to address grid dispatch problems.

References

1. Gherbi YA, Lakdja F, Bouzeboudja H, Gherbi FZ (2019) Hybridization of two metaheuristics for solving the combined economic and emission dispatch problem. Neural Comput Appl 31(12):8547–8559. https://doi.org/10.1007/s00521-019-04151-7
2. Mahdi FP, Vasant P, Abdullah-Al-Wadud M, Kallimani V, Watada J (2019) Quantum-behaved bat algorithm for many-objective combined economic emission dispatch problem using cubic criterion function. Neural Comput Appl 31(10):5857–5869. https://doi.org/10.1007/s00521-018-3399-z
3. Chaib AE, Bouchekara HREH, Mehasni R, Abido MA (2016) Optimal power flow with emission and non-smooth cost functions using backtracking search optimization algorithm. Int J Electr Power Energy Syst 81:64–77. https://doi.org/10.1016/j.ijepes.2016.02.004
4. Labbi Y, Attous DB, Gabbar HA, Mahdad B, Zidan A (2016) A new rooted tree optimization algorithm for economic dispatch with valve-point effect. Int J Electr Power Energy Syst 79:298–311. https://doi.org/10.1016/J.IJEPES.2016.01.028
5. Fathy A, Abdelaziz A (2018) Single-objective optimal power flow for electric power systems based on crow search algorithm. Arch Electr Eng 67(1):128–138. https://doi.org/10.24425/118996
6. Sakthivel VP, Suman M, Sathya PD (2021) Combined economic and emission power dispatch problems through multi-objective squirrel search algorithm. Appl Soft Comput 100:106950. https://doi.org/10.1016/j.asoc.2020.106950
7. Abdulrasool AQ, Al-Bahrani LT (2021) Multi-objective constrained optimal power flow based on enhanced ant colony system algorithm. In: 12th International symposium on advanced topics in electrical engineering, ATEE 2021. https://doi.org/10.1109/ATEE52255.2021.9425082

8. Rugema FX, Yan G, Mugemanyi S, Jia Q, Zhang S, Bananeza C (2020) A Cauchy-Gaussian quantum-behaved bat algorithm applied to solve the economic load dispatch problem. IEEE Access 9:3207–3228. https://doi.org/10.1109/access.2020.3034730
9. Maity D, Banerjee S, Chanda CK (2019) Bare bones teaching learning-based optimization technique for economic emission load dispatch problem considering transmission losses. Iran J Sci Technol Trans Electr Eng 43:77–90. https://doi.org/10.1007/s40998-018-0158-1
10. Melguizo FJ, Jurado F, Ebeed M, Naresh S, Damodar Reddy PM, Reddy K (2019) Optimal power flow solution using crow search algorithm. IOSR 09:20–26
11. Rezaie H, Kazemi-Rahbar MH, Vahidi B, Rastegar H (2019) Solution of combined economic and emission dispatch problem using a novel chaotic improved harmony search algorithm. J Comput Des Eng 6(3):447–467. https://doi.org/10.1016/j.jcde.2018.08.001
12. Taher MA, Kamel S, Jurado F, Ebeed M (2019) Modified grasshopper optimization framework for optimal power flow solution. Electr Eng. https://doi.org/10.1007/s00202-019-00762-4
13. Alawode KO, Adegboyega GA, Abimbola Muhideen J (2018) NSGA-II/EDA Hybrid evolutionary algorithm for solving multi-objective economic/emission dispatch problem. Electric Power Compon Syst 46(10):1158–1170. https://doi.org/10.1080/15325008.2018.1488302
14. Ali ES, Abd Elazim SM (2018) Mine blast algorithm for environmental economic load dispatch with valve loading effect. Neural Comput Appl 30(1):261–270. https://doi.org/10.1007/s00521-016-2650-8
15. Daryani N, Zare K (2018) Multiobjective power and emission dispatch using modified group search optimization method. Ain Shams Eng J 9(3):319–328. https://doi.org/10.1016/j.asej.2016.03.001
16. Faseela CK, Vennila H (2018) Economic and emission dispatch using whale optimization algorithm (WOA). Int J Electr Comput Eng (IJECE) 8(3):1297–1304. https://doi.org/10.11591/ijece.v8i3.pp1297-1304
17. Patel N, Bhattacharjee K (2018) A comparative study of economic load dispatch using sine cosine algorithm. Sci Iranica 27(3):1467–1480. https://doi.org/10.24200/SCI.2018.50635.1796
18. Pulluri H, Naresh R, Sharma V (2018) A solution network based on stud krill herd algorithm for optimal power flow problems. Soft Comput 22(1):159–176. https://doi.org/10.1007/S00500-016-2319-3
19. Rao RV, Keesari HS (2021) A self-adaptive population Rao algorithm for optimization of selected bio-energy systems. J Comput Des Eng 8(1):69–96. https://doi.org/10.1093/jcde/qwaa063

Comparative Analysis of Gate Driver Circuits for GaN MOSFET-Based Class E Resonant Inverter

Vikram Kumar Saxena and Kundan Kumar

1 Introduction

Modern industries depend upon electronic devices which work on a high frequency. The advantages of higher operating frequency of devices are reduction in the size of capacitors, inductors, and transformers. Class E resonant inverter is highly efficient on high frequency, which makes it suitable for the electronic florescent lamp (low voltage), radio transmitters, induction welding, high-frequency heaters, and fiber optics applications. Class E resonant inverters have the advantage of reliable power transfer, improved bandwidth, smaller switching losses, and increased efficiency [1].

In recent past decades, GaN high electron mobility transistors (HEMTs) have shown an encouraging performance for high-frequency applications. Its input capacitor and ON time resistance are very smaller than the SiC and Si MOSFETs under the same operating conditions [2, 3]. Therefore, it is predictable that GaN HEMTs can be adapted for high switching frequencies around tens of megahertz (MHz). GaN devices can operate very efficiently at high power with high frequencies [4, 5]. So it is obvious that in near future, it may obsolete the Si and SiC devices from the market.

The input capacitance C_{iss} of MOSFET is represented by the combination of the gate-to-source capacitance C_{gs} and gate-to-drain capacitance C_{gd}, i.e., $C_{iss} = C_{gs} + C_{gd}$. To raise the potential voltage of the gate from zero to a given voltage, the capacitors of MOSFET should be charged up very fast enough to operate the converters [6, 7]. Hence, to feed the energy to the MOSFET's gate, the driver circuit provides gate charge Q_g to the C_{iss}. Figure 1. shows the MOSFET's gate charge characteristics, which clearly explain that C_{iss} has to charge up until the threshold voltage is reached.

V. K. Saxena (✉) · K. Kumar
Department of Electrical Engineering, National Institute of Technology Manipur, Imphal, Manipur 795004, India
e-mail: vikramkumarsaksena@gmail.com

Fig. 1 Gate charge
characteristics of MOSFET
[6]
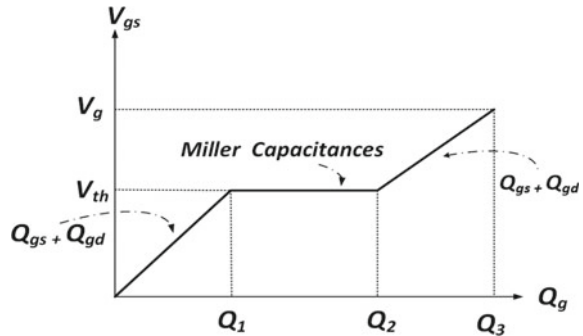
The gate driver circuit of MOSFETs provides the charging energy to the input capacitance (C_{iss}) of MOSFET for the turn ON and turn OFF transition in the form of Q_g. Normally, during discharging of C_{iss}, the energy is dissipated in the gate resistor [8]. Further, dissipated power loss will again increase as frequency will be raised. The motivation behind this paper is to reduce the gate power loss by recycling it, and the efficiency of gate drivers can be improved for GaN devices [6, 9]. Hence, transient resonant gate driver circuit has been compared with the conventional driver circuit and simulated in LTspice. In this work, a loss analysis of a resonant transient gate driver is also compared with a conventional totem pole resonant driver circuit.

In detail, the paper is summarized as follows. Section 2 describes the class E inverter and its modeling. Section 3 demonstrates different types of driver circuits. The simulation results and comparison of various gate driver circuits for class E resonant inverter are presented in Sect. 4. Finally, Sect. 5 concludes the paper.

2   Class E Resonant Inverter

A converter has more switching losses at higher frequency (~MHz ranges). Therefore, to overcome this switching loss, the concept of zero voltage switching (ZVS) and zero derivative voltage switching (ZDVS) technique can be employed. The time error can also be reduced during turn ON operation by incorporating ZVS and ZDVS techniques. The ZVS technique is fully utilized in the class E resonant inverter in which the commutation takes place at zero switching voltage which makes it more efficient [1]. In Fig. 2, a class E resonant inverter is shown in which a parallel resonant capacitor (C_{sh}) can be seen which absorbs the output capacitance of the switch. The class E resonant inverter's passive components are shown in Fig. 2, such as choke inductor L_{ch}, resonance inductor L_r, resonance capacitor C_r, and load resistance R_L.

Fig. 2 Circuit diagram of class E resonant inverter [1]
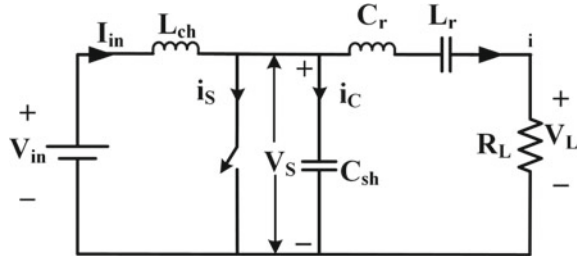
3 Different Types of Gate Driver Circuit

3.1 Basic Principle of Resonant Gate Driver Circuit

In the conventional totem pole gate driver circuit, there is a series combination of MOSFETs S_1 and S_2 demonstrated in Fig. 3. In this driver circuit, the C_{iss} of the main MOSFET Q is charged and discharged by using the gate signal of the driver circuit. In Fig. 4a, b, a comparison is shown for voltage source and current source driver circuit waveforms, respectively [5]. In totem pole conventional gate driver circuit [3, 9] gate loss and conduction loss can be calculated by using (1) and (2), respectively, as follows:

$$P_{\text{gate}} = V_g Q_g f_s = V_g^2 C_{\text{iss}} f_s. \tag{1}$$

$$P_q = I^2 R_{\text{on}}, \tag{2}$$

Fig. 3 Conventional totem pole gate driver circuit [7]
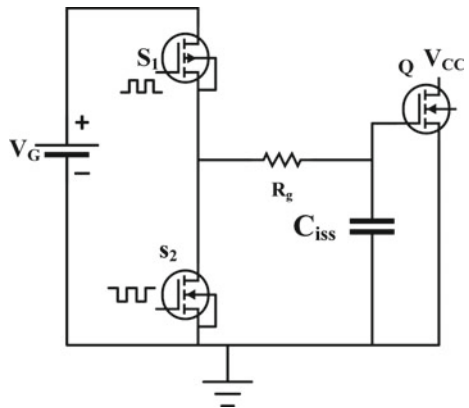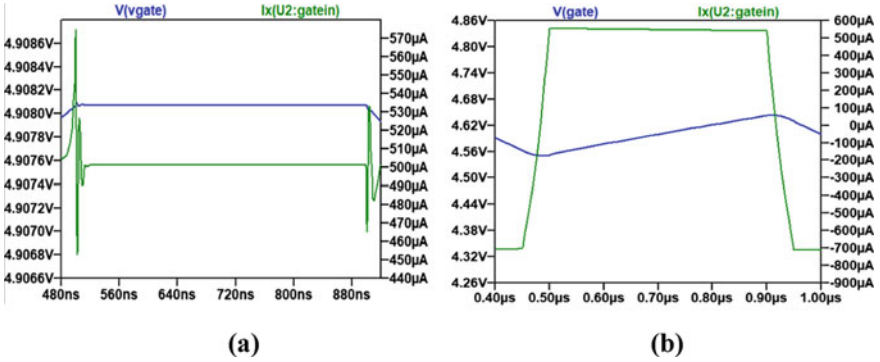
Fig. 4 For conventional gate driver circuit: **a** gate voltage and current by using voltage, source **b** gate voltage and current by using current source driver

where V_g is gate voltage, i.e., $V_g = V_{gON} - V_{gOFF}$. The gate power loss is directly proportional to the frequency; therefore in case of higher frequency, gate power loss will be more [10–12].

3.2 Resonant Transient Gate Driver Scheme

In order to overcome the above problem, a slight modification may be done (as shown in Fig. 5) known as a resonant transient driver (RTD) circuit. For Si MOSFETs, this RTD circuit design process was first described in [6]. The resonant transient circuit consists of the series combination of L_r and C_r which charges and discharges the input capacitance (C_{iss}) of GaN MOSFET by constant current I. When the MOSFET's gate is fully charged to V_{gs}, then S_1 is turned ON and a lower impedance path is achieved by the voltage source of gate of MOSFET, and the inductor current increases linearly simultaneously. The transition phenomenon from ON to OFF is achieved by turning OFF S_1, and S_2 is turned ON when the gate is fully discharged, and the gate and source of MOSFET are shorted. The function of C_r is to eliminate the dc component of the voltage of the inductor (L_r) [6].

Fig. 5 Modified conventional gate driver resonance transient circuit [6]
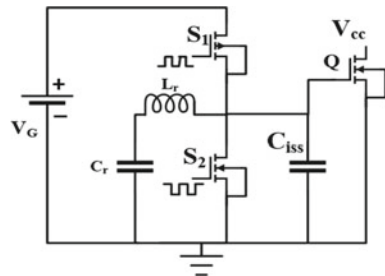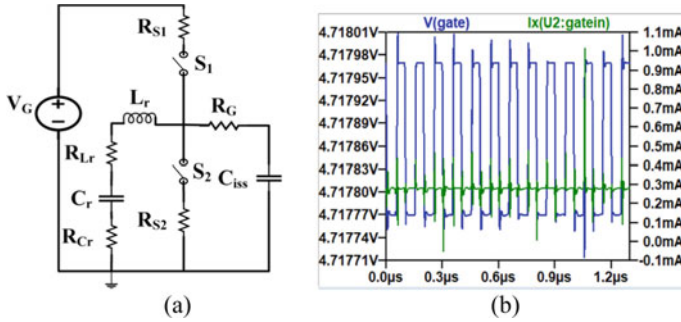
Fig. 6 Resonant transient driver: **a** equivalent circuit, and **b** gate voltage and gate current [6]

In Fig. 6a, equivalent resonant transient driver circuit is represented, while its gate voltage and current waveform are shown in Fig. 6b. Comparing Fig. 6b with Fig. 4a, b, the improvement of gate voltage and current can be easily visualized. Loss calculation for resonant transient driver circuit is as follows:

The length of transition interval τ is considered as $0.01\ \mu$s, so the relative length of transition interval $p = \tau * f = 0.1$. The value of resonating inductor L_r is selected by using (3), and the value of resonating capacitance C_r is chosen with the help of (4).

$$L_r = \frac{D(1-D)p}{2 * C_{\text{iss}} * f^2} \tag{3}$$

$$L_r C_r \gg \frac{1}{f} \tag{4}$$

$$R_x = R_{\text{lr}} + R_{\text{cr}} + 2R_{\text{on\_ds}}(1 - 6p) + 6R_g p. \tag{5}$$

In (5), equivalent resistance R_x is computed by adding the ESR of L_r and C_r, i.e., R_{Lr} and R_{Cr}, respectively, drain to source on resistance $R_{\text{on\_ds}}$, and gate resistance R_g of MOSFETs. The gate power loss of resonant transient driver circuit can be calculated by using (6).

$$P_{g\_\text{total}} = \frac{1}{3}R_x \left(\frac{C_{\text{iss}}V_{\text{gs}}f}{p}\right)^2. \tag{6}$$

4 Simulation Results and Comparison of Various Gate Driver Circuits for Class E Resonant Inverter

A totem pole resonant driver and resonant transient driver circuits are used to operate GaN-based class E resonant inverter. The simulation work is carried out using LT spice software considering the parameters of class E resonant inverter as listed in Table 1. From the simulation results, it is observed that the resonant transient driver circuit has improved output voltage and output current profile as compared to conventional totem pole gate driver circuit as shown in Figs. 7 and 8.

Switch gate driving loss and conduction loss can be calculated using (7) and (8).

$$P_{\text{GS}} = V_{\text{GS}} Q_g f, \tag{7}$$

Table 1 Class E resonant inverter parameters using conventional totem pole and resonant transient driver circuits

Parameters (RMS value)	Symbol	Conventional driver circuit	RT driver circuit
Rated power	P_o	200 W	200 W
Output voltage	V_l	52.996 V	53.562 V
Output current	(I_l)	3.7031 A	3.6811 A
Drain to source voltage	(V_{ds})	78.347 V	78.389 V
Drain to source current	(I_{ds})	7.2051 A	7.0165 A
Gate voltage	(V_g)	4.3267 V	4.2249 V
Gate current (I_g)	(I_g)	580.07 mA	580.4 mA
Calculated output power	$P_{\text{o,calc}}$	196.25 W	197.17 W
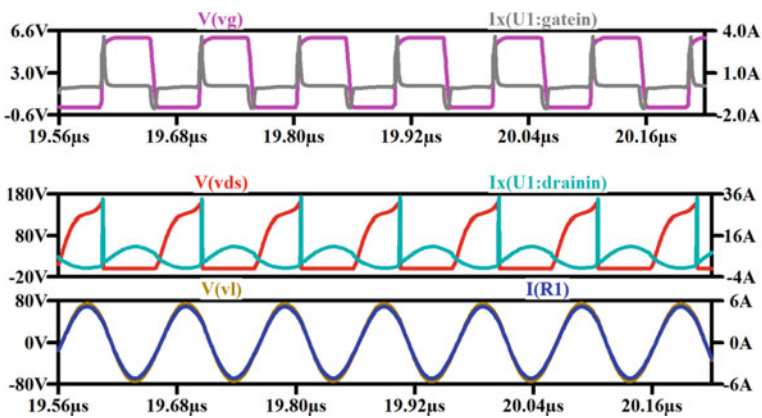
Fig. 7 Waveforms of class E resonant inverter using resonant transient driver circuit
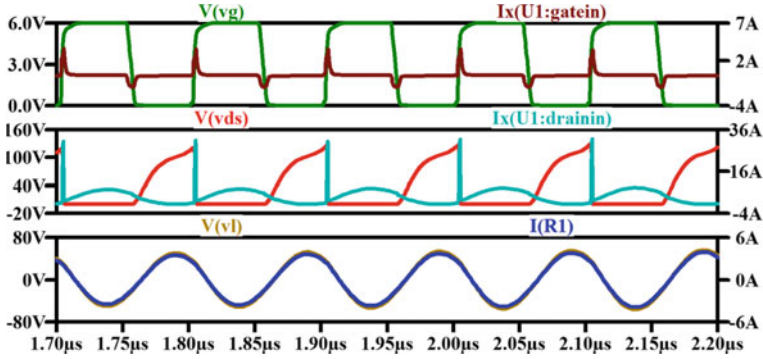
Fig. 8 Waveforms of class E resonant inverter using conventional totem pole driver circuit

$$P_{r\text{DS}} = r_{\text{DS}}\, I_{\text{SRMS}}^2. \tag{8}$$

The power losses in ESR of L_{ch}, C_{sh}, C_r, and L_r are given by (9), (10), (11), and (12).

$$P_{rL_{\text{ch}}} = r_{L_{\text{ch}}} I_{L_{\text{ch}}\text{RMS}}^2, \tag{9}$$

$$P_{rC_{\text{sh}}} = r_{C_{\text{sh}}} I_{\text{CRMS}}^2, \tag{10}$$

$$P_{rL_r} = \left(r_{L_r} I_m^2\right)/2, \tag{11}$$

$$P_{rC_r} = \left(r_{C_r} I_m^2\right)/2, \tag{12}$$

$$P_{t_{\text{off}}} = \frac{\omega t_f}{12} P_o. \tag{13}$$

In (13), turn OFF switching loss is expressed as the turn ON losses are negligible due to ZVS operation. Hence, the total power loss may be summed up as in (14).

$$P_{\text{LS}} = P_{rL_{\text{ch}}} + P_{r\text{DS}} + P_{rC_{\text{sh}}} + P_{rL_r} + P_{rC_r} + P_{\text{GS}} + P_{t_{\text{off}}}. \tag{14}$$

Using the (7)–(14) and simulated data, the total power loss is determined and shown in Table 2.

The class E inverter's efficiency can be estimated using (15).

$$\eta = \frac{1}{1 + \frac{P_{\text{LS}}}{P_R}}. \tag{15}$$

Table 2 Different power losses for conventional totem pole and resonant transient driver circuits

Parameters (RMS value)	Symbol	Conventional driver circuit (W)	RT driver circuit (W)
Gate loss	P_{GS}	0.1591	0.1517
Conduction loss	P_{rDS}	0.2595	0.2461
Turn OFF switching losses	$P_{t_{off}}$	1.0112	0.2315
ESR loss in L_r	P_{rL_r}	0.6753	0.7140
ESR loss in C_r	P_{rC_r}	0.6753	0.7140
ESR loss in L_{ch}	$P_{rL_{ch}}$	0.7125	0.6213
ESR loss in C_{sh}	$P_{rC_{sh}}$	0.2530	0.1530
Total losses	P_{LS}	3.7459	2.8316
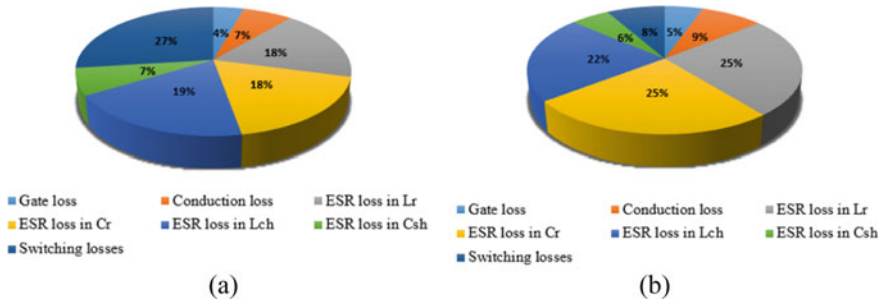
Fig. 9 Pie chart for power losses in: **a** resonant transient driver circuit, **b** conventional driver circuit

The pie charts representation of various loss components of both the considered drivers as given in Table 2 is presented in Fig. 9a, b, respectively.

5 Conclusion

In this work, a totem pole resonant gate and transient resonant gate drivers are implemented to compare the performance of the class E resonant inverter circuit. The power loss modeling for both the driver circuits is demonstrated. Further, the performance of the class E resonant inverter is investigated using both the driving circuits. The simulation work is carried out with the help of LTspice simulation. The overall efficiency of the class E resonant inverter is obtained using conventional totem pole driver circuit and transient resonant gate driver circuit that are 98.12% and 98.63%, respectively. The improved efficiency in the case of the transient resonant gate driver can also be maintained in a wide range of switching frequencies. Hence, it can be concluded that the performance of the class E resonant inverter using transient resonant driver circuit is better as compared to conventional driver circuit.

References

1. Kumar K, Saxena VK (2022) LTspice simulation analysis for loss estimation of GaN MOSFET based class E resonant inverter. In: 2022 IEEE global conference on computing, power and communication technologies (GlobConPT), New Delhi, India, pp 1–6
2. Keshmiri N, Wang D, Agrawal B, Hou R, Emadi A (2020) Current status and future trends of GaN HEMTs in electrified transportation. IEEE Access 8:70553–70571
3. Kumar K, Santra SB (2018) Performance analysis of a three-phase propulsion inverter for electric vehicles using GaN semiconductor devices. IEEE Trans Ind Appl 54(6):6247–6257
4. Kumar K, Badugu V, Devanshu A, Kumar A, Choudhury TR (2022) Power loss estimation of isolated coupled inductor based dual active bridge converter using GaN devices. In: IEEE international conference on industrial electronics: developments & applications (ICIDeA-2022), Bhubaneswar, 15–16 Oct 2022, pp 42–247
5. Kumar K (2021) Performance analysis of onboard GaN-based synchronous rectifier used in dynamic wireless charging system for electric vehicles. In: Proceedings of IEEE international conference on smart technology for power, energy and control (STPEC 2021), Bilaspur, India, 19–22 Dec 2021, pp 1–6
6. Maksimovic D (1991) A MOS gate drive with resonant transitions. In: PESC '91 record 22nd annual IEEE power electronics specialists conference. IEEE, pp 527–532
7. Sun B, Zhang Z, Andersen MA (2019) A comparison review of the resonant gate driver in the silicon MOSFET and the GaN transistor application. IEEE Trans Ind Appl 55(6):777–7786
8. Chennu JVPS, Maheshwari R, Li H (2017) New resonant gate driver circuit for high-frequency application of silicon carbide MOSFETs. IEEE Trans Industr Electron 64(10):8277–8287
9. Seidel A, Wicht B (2018) Integrated gate drivers based on high-voltage energy storing for GaN transistors. IEEE J Solid-State Circ 53(12):3446–3454
10. Zhang Z, Liu C, Wang M, Si Y, Liu Y, Lei Q (2021) A Novel current-source-based gate driver with active voltage balancing control for series-connected GaN HEMTs. IEEE Open J Power Electron 2:346–367
11. Zhu J et al (2021) Study and implementation of 600-V high-voltage gate driver IC with the common-mode dual-interlock technique for GaN devices. IEEE Trans Industr Electron 68(2):1506–1514
12. Chen J, Luo Q, Huang J, He Q, Sun P, Du X (2020) Analysis and design of an RC snubber circuit to suppress false triggering oscillation for GaN devices in half-bridge circuits. IEEE Trans Power Electron 35(3):2690–2704

High-Power Wireless Power Transfer (WPT) System Using Series-Series (SS) Topology

Prabhat Chandra Ghosh

1 Introduction

A high-frequency AC voltage source is necessary in a wireless power transfer (WPT) system in order to transmit power from the source to the load. Since high-frequency AC voltage cannot be produced by a commercial power source, the WPT transformer needs to undergo the proper conversion. Therefore, in order to produce a high-frequency alternating magnetic field, the current flowing in the primary coil of a WPT transformer must be a high-frequency alternating current. An AC/DC rectifier and a high-frequency DC/AC inverter are needed in order to convert low-frequency (50 Hz) AC voltage into high-frequency AC current. A WPT transformer also needs a lot of power that is delivered at a high frequency and efficiency. This demand can only be satisfied by specially constructed high-frequency voltage source inverters. In general, WPT transformers work at or close to their resonance frequency. As a result, the primary coil's current must be sinusoidal. Passive components made up of capacitors and inductors are required to build a high-frequency inverter with a sinusoidal output current. The resonance frequency of these passive networks of the WPT transformer is chosen to be the same as the switching frequency of the switching components of a high-frequency inverter. The square wave voltage is the inverter circuit's output voltage. This square wave voltage is then applied to the WPT transformer's resonant network, which filters out all but the frequencies that coincide with the passive networks' resonant frequencies. Resonant inverters are inverters that make utilization of passive resonant networks. These resonant inverters may achieve a variety of zero-voltage and zero-current switching states by carefully selecting the component values of the passive networks, the switching frequency, and

P. C. Ghosh (✉)

Department of Electrical Engineering, College of Military Engineering, Pune,
Maharashtra 411031, India
e-mail: pcghoshcme@gmail.com

its duty cycle because they are switched-mode circuits as well. Under ideal conditions, voltage source resonant inverters can attain 100% efficiency. However, due to the high-frequency switching losses and resistive losses of the passive parts, this is nearly impossible to achieve. Voltage source resonant inverters' typical efficiency ranges from 85 to 95% [1, 2].

2 Full-Bridge Resonant Inverter for the WPT System Using SS Topology

Figure 1 illustrates a series-series (SS) topology-based WPT system. A constant current source converter is one that uses SS topology. This indicates that the secondary current $i_2(t)$ is constant regardless of the load when the inverter's switching frequency is equal to the resonance frequency [3, 4]. In the SS topology, the secondary current $i_2(t)$ can be given as follows:

$$i_2(t) = \frac{v_1(t)_{(1)}}{j\omega_0 M},$$
(1)

where $v_1(t)_{(1)}$ is the fundamental component of primary voltage $v_1(t)$, ω_0 is the resonance frequency, and M is the mutual inductance of the WPT transformer as given in Fig. 1.

The output current $i_2(t)$, is supplied to the input of the diode rectifier. The output capacitor C_0 functions as a low-pass filter for pulsating DC output current of the diode rectifier. In order to minimize ripples in the DC load current and voltage across the resistive load, a very large filter capacitor is provided. Consequently, the secondary voltage $v_2(t)$ of the transformer is a square wave. From (1) and Fig. 4, it can be shown that secondary current $i_2(t)$ leads the primary voltage's fundamental component $v_1(t)_1$ by $\frac{\pi}{2}$ radians when the inverter's switching frequency is equal to
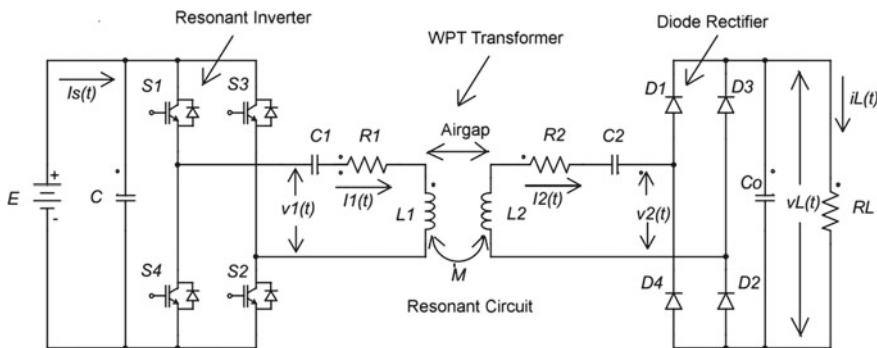
Fig. 1 Circuit schematic of a WPT system using series-series (SS) compensation

the resonant frequency. Hence, $i_2(t)$ can be given as:

$$i_2(t) = I_2 \sin\left(\omega t + \frac{\pi}{2}\right), \tag{2}$$

where I_2 is the maximum value of secondary current $i_2(t)$.

The WPT transformer's secondary voltage $v_2(t)$ can now be given as:

$$v_2(t) = \sum_{n=1,3,5....} \frac{4V_L}{n\pi} \sin\left(n\omega t + \frac{\pi}{2}\right) = \frac{4V_L}{n\pi} \sum_{n=1,3,5....} \sin\left(n\omega t + \frac{\pi}{2}\right), \tag{3}$$

where V_L is the average value of DC load voltage $V_L(t)$.

From (3), fundamental component of $v_2(t)$ can be given as:

$$v_2(t)_{(1)} = \frac{4V_L}{\pi} \sin\left(\omega t + \frac{\pi}{2}\right). \tag{4}$$

The filter capacitor C_0 filters the secondary current $i_2(t)$ after the diode rectifier has rectified it. The average rectified current value $i_2(t)$ is equal to the average value of the load current I_L and can be given as:

$$I_L = 0.637 I_2. \tag{5}$$

3 Finite Element Analysis (FEA) of the Proposed WPT System

For the proposed WPT system, the mutual inductance, coupling coefficient, and self-inductances of the main and secondary coils of the transformer have been calculated using the Maxwell finite element tool and magneto-static solution method. The FEM model of the WPT transformer is depicted in Fig. 2. In post processing, change in ampere turns of transmitter and receiver coils have been done in the ratio of 1:1.5 and therefore WPT transformer turns ratio is changed to 11:25 which is not shown in Fig. 2 (before processing) [5].

Table 1 lists the parameters and results of finite element analysis (FEA). Since the solution of this finite element approach is derived through numerical manipulation of various electric and magnetic values at various finite model elements, it is a precise and straightforward method to determine the self and mutual inductances. The values of the magnetic flux density at various model elements are shown in Fig. 3.

Fig. 2 Ansys Maxwell 3D
model of the proposed WPT
transformer
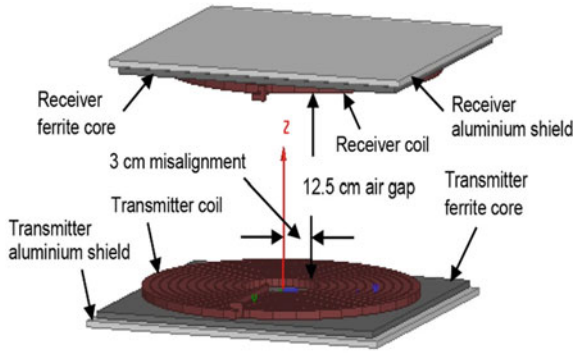
Table 1 Input parameters for Ansys Maxwell finite element analysis (FEA) and output results

FEA input parameters	Values	FEA output parameters	Values
Transformer turns ratio (Post processing ampere turns adjustment)	11:25	Self-Inductance of the transmitting coil	48.47 µH
The inner radius of the transmitting coil	5 cm	Self-Inductance of the receiving coil	109.8 µH
The outer radius of the transmitting coil	18 cm	Mutual inductance	22.55 µH
Dimension of ferrite core (cm)	38 × 38 × 0.5	Coupling coefficient	0.31
Dimension of aluminum shield (cm)	40 × 40 × 0.5		
Air gap	12.5 cm		
Lateral misalignment	3 cm		
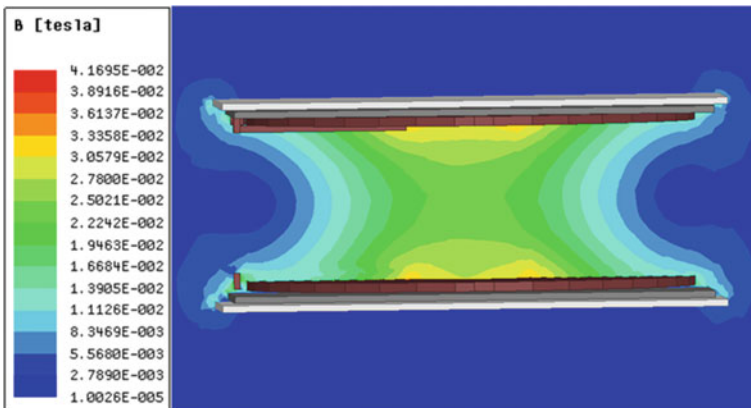
Fig. 3 Distribution of flux density in the air gap
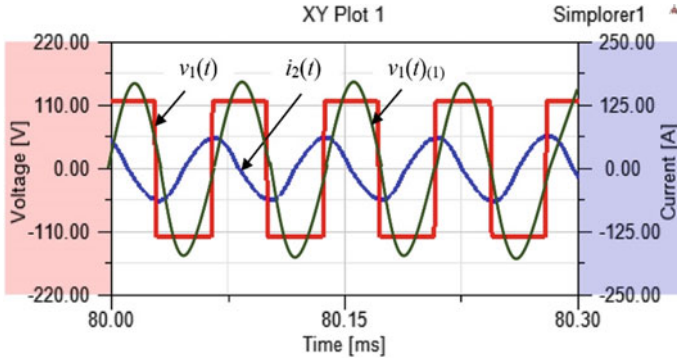
Fig. 4 Primary voltage $v_1(t)$, fundamental component of primary voltage $v_1(t)_{(1)}$, and load current $i_2(t)$

4 Circuit Simulation for the Proposed System

By integrating the FEM model with the electrical circuit model, circuit simulation is performed for the proposed system using Ansys Simplorer software to evaluate the electrical transfer characteristics [5, 6]. In this investigation, a full-bridge inverter with SS compensation scheme was used to simulate a WPT system.

The corresponding circuit diagram of the suggested system is shown in Fig. 1. The switching frequency and the system's resonant frequency in this investigation are both 14 kHz [7]. As a result, the transmitter and receiver coils' compensating capacitor values were selected in accordance with the resonance frequency [8, 9]. Table 2 shows the inputs for circuit simulation. In both transmitting and receiving coils, a series-series (SS) compensation scheme has been implemented. A thorough analysis of the proposed WPT system is conducted, taking into account an air gap of 12.5 cm and a lateral misalignment of 3 cm. High-power applications were the focus of this simulation. The suggested system has an input power of 7.6 kW and an output power of 5.5 kW for a load resistance of 3.35 Ω. Consequently, as stated in Table 3, 72.36% energy efficiency is attained for the mutual inductance of 22.55 μH, and the lateral misalignment of 3 cm. 140 V is the recommended DC output voltage. Primary and secondary voltage, current, and power are shown in Figs. 5 and 6, respectively. Table 3 shows the results of the circuit simulation.

5 Conclusions

A high-frequency resonant inverter topology suitable for high-power WPT applications was presented in this study. The high-frequency inverter is crucial to a WPT system to successfully transmit power from the source to the load by supplying high-frequency voltage to the WPT transformer. First, a thorough analysis and study of

Table 2 Parameters and their values for circuit simulation using Ansys Simplorer

Circuit parameters	Values
Input voltage	120 V DC
Load resistance, R_L	3.35 Ω
Self-inductance of the transmitter coil, L_1	48.47 µH
Self-inductance of the receiver coil, L_2	109.76 µH
Mutual inductance, M	22.55 µH
Compensating capacitance (transmitter), C_1	2.7 µF
Compensating capacitance (receiver), C_2	1.17 µF
Coupling coefficient	0.31
R_1 and R_2 (resonant circuit)	4 mΩ and 3 mΩ
Switching frequency	14 kHz
Resonance frequency	14 kHz

Table 3 Ansys Simplorer circuit simulation results

Load resistance (Ω)	DC output (V)	Input power (kW)	Output power (kW)	Efficiency (%)
3.35	140	7.6	5.5	72.36
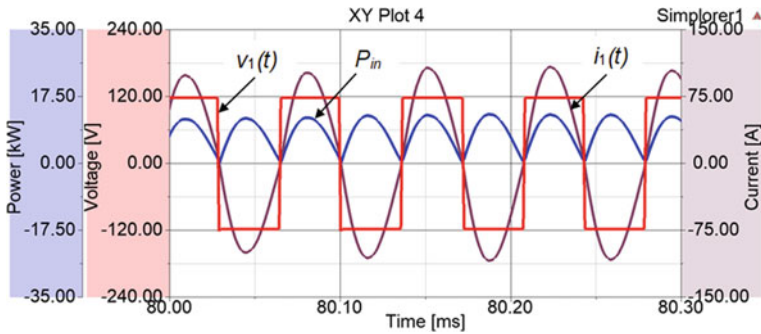
Fig. 5 Voltage, current, and power in primary side
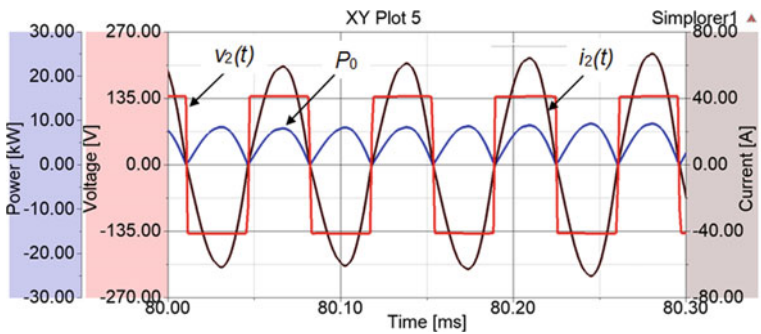
Fig. 6 Voltage, current, and power in secondary side

the full-bridge inverter are conducted. A full-bridge inverter with SS compensation topology has been simulated to validate the inverter topology and design process for high-power WPT applications. The prototype's simulated waveforms and the theoretical analysis agree very well. Primary voltage and primary current are in phase with each other and sinusoidal. Therefore, the supply of any reactive power from the source is not required. Finally, it can be said that the suggested high-frequency inverter with SS topology is well suited for high-power WPT applications, including electric vehicle charging systems.

References

1. Choi J, Tsukiyama D, Tsuruda Y, Davila JMR (2018) High-frequency, high-power resonant inverter with eGAN FET for wireless power transfer. IEEE Trans Power Electron 33:1890–1896
2. Jou HL, Wu JC, Wu KD, Kuo CY (2021) Bidirectional DC–DC wireless power transfer based on LCC-C resonant compensation. IEEE Trans Power Electron 36:2310–2319
3. Kunwar A, Sheldon SW (2019) Design guidelines to avoid bifurcation in a series-series compensated inductive power transfer system. IEEE Trans Industr Electron 66:3973–3982
4. Zhang W, Mi CC (2016) Compensation topologies of high-power wireless power transfer systems. IEEE Trans Veh Technol 65:4768–4778
5. Wang J, Ho SL, Fu W, Kit CT, Sun M (2011) Finite-element analysis and corresponding experiments of resonant energy transfer for wireless transmission devices. IEEE Trans Magn 47:1074–1077
6. Song D, Shi F, Dai S, Liu L (2019) Design and analysis of a wireless power transmission system with magnetic coupling resonance in the weak-coupling region. Chin J Electr Eng 5:51–60
7. Jiang Y, Wang L, Wang Y, Liu J, Wu M, Ning G (2019) Analysis, design, and implementation of WPT system for EV's battery charging based on optimal operation frequency range. IEEE Trans Power Electron 34:6890–6905
8. Onai K, Ojo JO (2022) Performance analysis and design of frequency controlled series-series compensated inductive power transfer system for electric vehicle battery charging. IEEE Trans Ind Appl 58:962–973
9. Zeng H, Wang X, Peng FZ (2020) An efficient inductive power transfer topology for electric vehicle battery charging. IEEE Trans Ind Appl 56:6925–6936

A Multi-criteria Decision-Making (MCDM) Approaches for Systematic Analysis and Ranking of Solar Power Plant Site Using ANN

Hormi Kashung and Benjamin A. Shimray

1 Introduction

Electrical energy is essential and has become basic need for mankind. But dependence on fossil fuels makes it very hard to utilize clean energy causing global warming, climate changes, etc. [1]. The 1997 adoption of the Kyoto Protocol to the United Nations Framework Convention on Climate Change landmarked a major turning point to promote Renewable Energy [2, 3]. One of the few nations to have made the essential preparations for the production, transmission, and distribution of renewable energy is India. The availability of electrical energy also affects a nation's social and economic conditions [4]. Renewable energy like solar energy plays an important role to increase energy independence. India is a country that has enormous potential for solar energy and because of its location, access to sun energy all year long and the biggest land area [5]. India stands fourth in generating solar energy in the world. India has also invested a huge sum of money to generate 100 GW by 2022 in the national solar mission. Therefore, it is essential to recognize the best location for the nation's solar plant installation. However, research is currently being conducted in several Indian states [6]. Choosing solar farm location can be problematic because it has to be checked whether it can be extended in the future [7]. Thus, choosing a location for a solar farm is one of the most essential concerns to generating or maximizing the overall effectiveness. India is a nation that receives 5000 trillion kWh of solar insolation annually. The barren lands in the country can be accessed by up to 750 GW of solar energy. According to the Government of India's Ministry of New and Renewable Energy (MNRE-GOI), the Thar Desert can generate up to 142 GW of solar energy [8]. Four states of India have been taken for case study, i.e., Chandigarh, Gujarat, Manipur, and Kerala in this paper. The research methodology is explained

H. Kashung · B. A. Shimray (✉)
Department of Electrical Engineering, NIT Manipur, Imphal, Manipur 95004, India
e-mail: benjaminshimray@gmail.com

in detail for the presented work. Further, multi-layer perceptron back-propagation and genetic algorithm optimization technique is used for comparative analysis.

2 Research Methodology

This section examines the pertinent research on choosing a solar farm and outlines the knowledge gaps that must be filled for the project to be completed. Social, technological, economic, environmental, and political (STEEP) factor identification is the primary and the most crucial stage in laying the groundwork for a new solar farm. The architecture of STEEP is shown in Fig. 1.

2.1 Detecting Solar Hotspots in India

Since 2005 India's power generation has grown at a compound annual growth rate (CAGR) of 5.2%. Still, more than 400 million people lack the opportunity to use electricity. The IEPR 2006 (Integrated Energy Policy) has predicted greater than 800,000 megawatts (MW) by 2032 [9]. An average of 66 MW of solar applications are installed in the country [10]. The goals outlined by "Solar India" launched in 2010 are to meet the ambitious target of 22,000 MW grid-connected and 2000 MW off-grid solar production by 2025. It is crucial to locate the country's solar hotspots to access the potential and variability of the solar resources in India to realize the vision of a "**Solar India**" [11].
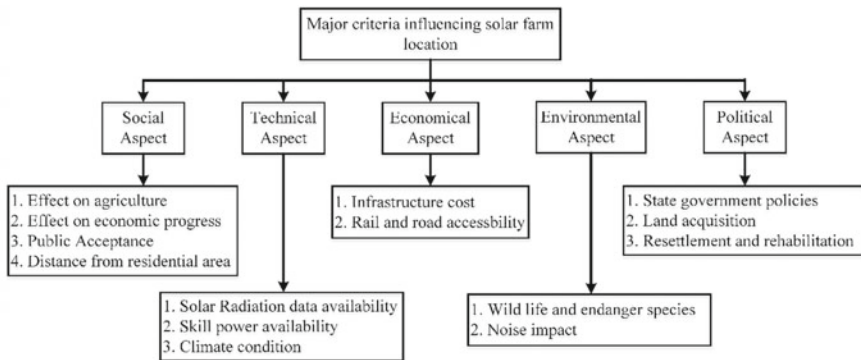
Fig.1 Major criteria influencing the location of the solar farm

3 The Research Framework of the Present Study

Using a thorough literature review methodology and recommendations from experts, it begins the choice of STEEP criterion influencing the choice of solar farms in the context of India. The creation of a decision hierarchy utilizing MLPBP and MLP-GA technique with STEEP criteria and sub-criteria are adopted in the present study. Experts are engaged once again to address any discrepancies that may appear in MLPBP results. The right place is then suggested after certain potential locations are taken into account based on MLPBP results and MLP-GA. Figure 2 shows the proposed framework for the proposed investigation [12].

3.1 MLP-BP

A supervised network that arranges different layers of neurons is used by MLP. In Fig. 3, MLP consists of an input layer, a hidden layer which is the network's brain, and an output layer, i.e., three layers. It is a network administrator which is topologically organized into numerous sections of neurons. Each layer's neurons are linked to the layer of every other layer's neurons in the $(i + 1)$th order. Back-propagation, a training procedure for MLP, must be used to calculate this connection weight. The back-propagation (BP) technique is based on the error-correction principle, which is used to train the network for learning. The weights of the hidden layer are updated using Eq. (1). The following steps are used to carry out the detailed training procedure.

$$v_{ij}(n + 1) = v_{ij}(n) + \eta * \partial_j(n) * x_i(n), \tag{1}$$
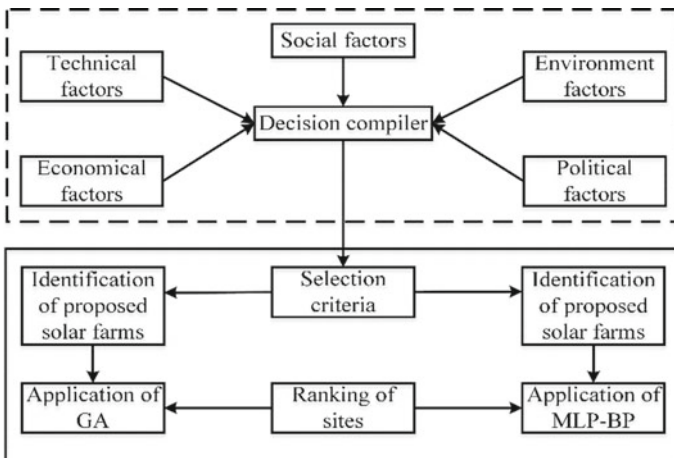
Fig. 2 Framework for the proposed investigation
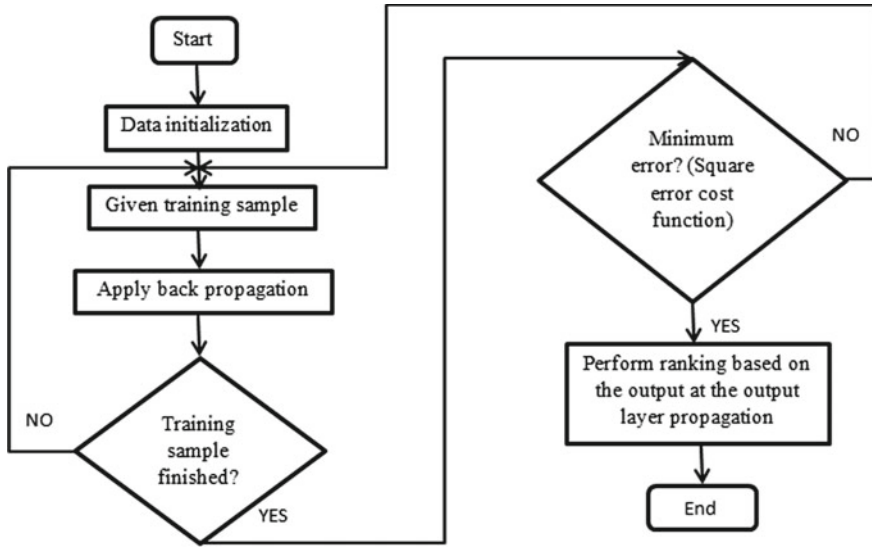
Fig. 3 Working flowchart for MLP-BP

$$w_{jk}(n+1) = v_{jk}(n) + \eta * \partial_j(n) * y_i(n). \tag{2}$$

$$\partial_k(n+1) = (\partial_k - o_k) * (1 - o_k) * O_k, \tag{3}$$

$$\partial_{j'} = \sum_{k=0}^{k} \partial_k * w_{jk}, \tag{4}$$

$$\partial_j = (1 - y_i) * y_j * \partial_{j'}. \tag{5}$$

Using the sigmoid function, learn the output of every layer of neurons for every input. Accordingly, Eq. (6) gives the input layer's output, and Eq. (7) gives the output layer's output (Fig. 4).

$$s_1 = \text{sigmoid}\left(\sum_{i,j} v_{ij} * x_i\right), \tag{6}$$

$$s_2 = \text{sigmoid}\left(\sum_{j,k} w_{jk} * y_j\right). \tag{7}$$
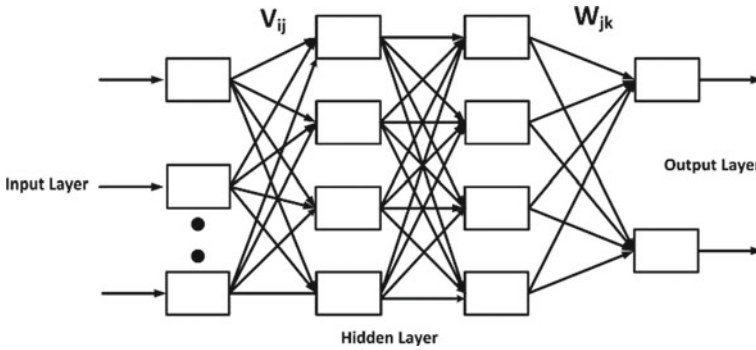
Fig. 4 MLP neural network

Where the weight of ith input to the jth hidden neuron layers is denoted by v_{ij}, ∂_j is the error signal produced by the jth hidden neuron, y_i is the output of the jth hidden neuron, η is the learning rate, o_k is the output of the kth output neuron, and ∂_k is the desired output kth output neuron, x_i is the ith Input [13].

Equations (1) and (2) are used to change the weights of every neuron's hidden and output layers. The erroneous signals from the hidden neurons are fed back to the hidden layer by the output layer in the BP method. The procedure is repeated until the error is lower than a predetermined threshold for the complete input–output pattern. For minimization, using the squares error cost function in (8).

$$E = \frac{1}{2} \sum_{k=0}^{k} (\partial_k - o_k)^2. \tag{8}$$

In Table 1, the different parameters are tabulated for MLP-BP algorithm.

Table 1 Climatic parameters of specific sites

Location	GHI	DNI	TI	K	S	T	RH	AQI	DH	WS
Manipur (Phangrei)	1.045	1.317	3.08	0.52	9.06	23	54	64	76	5.5
Kerala (CIAL)	5.571	3.934	5.45	0.56	13.02	32	78	87	16.5	12.3
Gujarat (Charanka)	5.542	4.867	6.42	0.55	13.4	34	55	15	19	13.5
Chandigarh (NChoe)	4.153	3.446	6	0.62	12.14	29	63	170	17	10

3.2 MLP-GA Algorithm

Figure 5 gives the flow chart of MLP-GA implementation. By designating the geno-
type of the GA as the weight list, an MLP is evolved. A binary number can be used
to represent each weight. To indicate the weights of the connections between the
layers of the ANN, each solution will be a bit string. Each training input size used
in this work is 11 in size. There are four hidden neurons and two output neurons
in total. Eleven total weights (TW) are assigned based on the size of each training
input. There are four and two hidden neurons and output neurons, respectively. Total
weights are provided by Eq. (9).

$$TW = (I * AB + AB * QN), \tag{9}$$

where the size of input pattern is denoted by I, AB is the of hidden neurons, and
QN is the number of output neurons. The current work's overall weight is 52. The
equation of length of gene (LG) is given by:
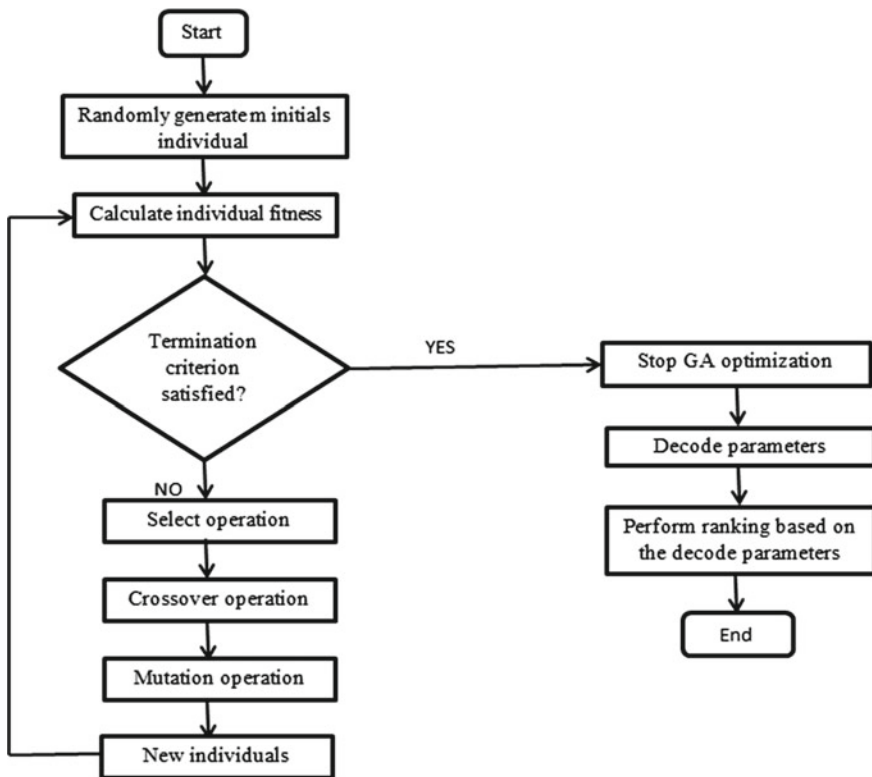
Fig. 5 Working flowchart for MLP-GA

$$\text{LG} = [P * (I * \text{AB} + \text{AB} * \text{QN})]. \tag{10}$$

where P is the number of bits per weight. Each weight is represented using a 16-bit binary number, i.e., $P = 16$, and hence, the length of the gene is $(\text{LG}) = 832$. Equation (11) shows the reconstruction of the phenotype from the genotype:

$$y_m = \sum_{k=1}^{P} b_{mk2^{-2}}, \tag{11}$$

where $b_{mk2^{-2}}$ is the m weight's k bit.

$$w_S = y_m * D + F, \tag{12}$$

where w_S is the amount of weight in the string, D is the scaling factor, and F is the shifting factor. The output layer and hidden layer's output are given as follows.

$$s_1 = \left(\sum_{m,j} v_{jm} * x_{pm} \right), \tag{13}$$

where $y_j = $ sigmoid function (S1), x_{pm} is the input:

$$s_2 = \left(\sum_{j,k} w_{kj} * y_j \right). \tag{14}$$

The mistake is eventually updated using Eq. (15):

$$E = \frac{1}{2} \sum_{k=0}^{k} (\partial_k - o_k)^2, \tag{15}$$

where $o_k = $ sigmoid 1 (function of unipolar activation is sigmoid), ∂_k is the previously determined desired output equation fitness, and N represents the number of training samples to determine the string fitness $= \frac{1-E}{N}$.

4   Result and Analysis

In Table 1, MLP-GA and MLP-BP parameter are shown. It consists of 11 inputs, 4 hidden neurons, and 2 output neurons. The target result is set as $(1, 0)$ for good, $(0, 1)$ for fair, and $(0, 0)$ for bad. Table 1 shows the data for the solar power plant site selected in a recent study. After the application of the proposed MLP-GA and conventional

MLP-BP, the results obtained are as in Table 2. Effective implementation of the algorithms for 40,000 iterations is illustrated as under:

For MLP-BP:

Weight "*W*" (for 40,000 iterations) in the output layers are: −11.250484, −6.032527, −5.925765, −5.783226, −5.201148 and 5.376030, −4.364461, −4.718610, −4.377911, 5.598781.

Weights for connections in the hidden layer are: −0.310785, −0.372383, −0.397257, 0.095255, −0.406114, −0.531399, −0.327321, 0.506233, 0.277953, 0.512685, −0.456857 and 0.143809, −0.552357, 0.374949, 0.110471, 0.015784, −0.061495, −0.159856, −0.614426, −0.127249, 0.211235, 0.327446 and −0.163731, −0.132746, −0.136705, 0.392167, −0.311036, −0.442995, 0.156651, −0.154446, −0.030061, −0.233577, −0.083521 and 0.371313, −0.296220, 0.133337, −0.136436, −0.304520, −0.255625, −0.298221, −0.038948, −0.331417, −0.081242, −0.137536.

For MLP-GA:

Weight "*W*" (For 40,000 iteration) in the output layers are *W*: 5.003967, −3.188782, −0.614929, 4.794006, 3.460999, and −7.992859, 9.524231, −8.113098, 3.690796, 4.251099.

Weights for connections in the hidden layer are:

2.979126, −0.974121, 0.061646, −4.470215, −6.751404, −1.930542, −4.538269, −5.680542, −5.046082, 5.933533, 9.702148, −1.342163, −1.699219,

Table 2 Ranking and comparison of MLP-BP and MLP-GA

MLP Algorithm (N0.of iterations)		Manipur (Phangrei)	Kerala (CIAL)	Gujarat (CHARANKA)	Chandigarh (N-CHOE)
MLP - BP	For 20,000 iterations	Out o/p 1:0.003346 o/p 2:0.444520	o/p 1:0.942892 o/p 2:0.013951	o/p 1:0.803346 o/p 2:0.044520	o/p 1:0.003346 o/p 2:0.444520
	For 40,000 iterations	o/p 1:0.002354 o/p 2:0.444541	o/p 1:0.076785 o/p 2:0.325028	o/p 1:0.003854 o/p 2:0.444541	o/p 1:0.002354 o/p 2:0.44541
	Ranking	3	1	2	3
MLP - GA	For 20,000 iterations	o/p 1:0.000000 o/p 2:0.000000	o/p 1:0.994136 o/p 2:0.000000	o/p 1:0.568620 o/p 2:0.000098	o/p 1:0.003500 o/p 2:0.000008
	For 40,000 iterations	o/p 1:0.000000 o/p 2:1.000000	o/p 1:0.999240 o/p 2:0.000000	o/p 1:0.483500 o/p 2:0.009798	o/p 1:0.000000 o/p 2:0.999798
	Ranking	4	1	2	3

9.031982, 7.217712, 9.011230, −5.486450, 6.130371, −3.273926, 6.184692 and 2.531433, −4.640808, 6.295166, −2.701721, −5.285339, 9.144287, −9.393616, 2.718811, 5.208740, 4.013367, −9.048157, 5.948486, −4.125366, −5.840454, −7.821655, 9.327087, 4.200745, −4.916992, 4.541321, 4.348145 and −1.596985, 8.518372, 7.958069, −3.184814, 5.506897, 1.583862, −8.864136, 7.969360, −5.873108, 7.953491, −9.496460, 8.467102, 1.188660, 1.690979, −4.476318, 2.637634, 0.491028, −0.689697, 6.616211, 9.660339 and 2.225342, 9.621887, 1.385193, −9.605713, 1.617737, 7.885437, 0.466919, −0.914307, 1.846008, −9.912720, −2.230835, 6.101990, −5.264893, 6.053162, 4.118652, 9.774475, 6.222534, 4.756470, −0.056458, 1.686096.

5 Conclusion

In this paper, both the quantitative and qualitative features required for the site selection of the plant are discussed. The proposed MLP-GA demonstrates its ability to correctly rank appropriate sites for the development of solar energy projects. Our findings are objective, and various crucial factors, including both quantitative and qualitative details regarding the proposed solar power facility, were taken into account. As the needs of the business sector change, considerations like internal return rate (IRR) and systemic advantages may be considered in future. However, the focus of our current study is on environmental effects, which are crucial to long-term sustainability. The adoption of these methods will result in an expansion of the work discussed in this study by using MLP-BP and fuzzy MLP-GA in combination to determine the best places to build solar power plants.

References

1. http://mnre.gov.in/file-manager/UserFiles/state-power-policies/Haryana-Solar-Power-Policy.pdf.
2. Sindhu S, Nehra V (2017) Investigation of feasibility study of solar farms deployment using hybrid AHP-TOPSIS analysis: case study of India. Renew Sustain Energy Rev 73:496–511. Elsevier
3. United Nations framework convention on climate change. Kyoto protocol (1997)
4. Choudhary D, Shankar R (2012) A steep-fuzzy AHP-TOPSIS framework for evaluation and selection of thermal power plant location: a case study from India. Energy, pp 510–521, Elsevier
5. Verma M, Ghritlahre HK, Bajpai S (2021) A case study of optimization of a solar power plant sizing and placement in Madhya Pradesh, India using multi-objective genetic algorithm. Ann Data Sci, pp 1–34
6. Thapar S, Sharma S (2016) Economic and environmental effectiveness of renewable energy policy instruments: best practices from India. Renew Sustain Energy Rev 66:487–498. Elsevier
7. Sindhu S, Nehra V, Luthra S (2016) Identification and analysis of barriers in implementation of solar energy in Indian rural sector using integrated ISM and fuzzy MICMAC approach. Renew Sustain Energy Rev 62:70–88

8. Tahri PM (2015) The evaluation of solar farm locations applying geographic information system and multi-criteria decision-making methods: case study in southern Morocco. Renew Sustain Energy Rev 51:1354–1362. Elsevier
9. Sahu BK (2016) Solar energy developments, policies and future prospectus in the state of Odisha, India. Renew Sustain Energy Rev 61:526–536
10. Benjamin AS, Manglem K, Khelchandra K, Mehta RK (2017) Ranking of sites for installation of hydropower plant using MLP neural network trained with GA: a MADM approach. Comput Intell Neurosci Article ID 4152140:8
11. Kar SK, Sharma A, Roy B (2016) Solar energy market developments in India. Renew Sustain Energy Rev 62:121–133
12. Singh A, Vats G, Khanduja D (2016) Exploring tapping potential of solar energy: prioritization of Indian states. Renew Sustain Energy Rev 58:397–440
13. Shimray BA, Singh KM, Khelchandra T, Mehta RK (2017) Ranking of sites for installation of hydropower plant using MLP neural network trained with GA: a MADM approach. Comput Intell Neurosci

Real-Time Simulation of an Islanded Mode Solar PV System with New Elman NN-Based MPPT

Bappa Roy, Shuma Adhikari, Aribam Deleena Devi, and Kharibam Jilenkumari Devi

1 Introduction

Solar power is one of the most important and easily available renewable energies in most parts of the world. By converting the sunlight and ultraviolet radiation of solar energy, a solar cell generates electrical power [1]. The generation of power from the PV cell depends completely on solar irradiance and the temperature of the PV panel, and it gives a steady output. To achieve maximum power (MP) in this power extraction process, the maximum power point tracking (MPPT) system is necessary. The MPPT helps to increase the productivity of the PV module [2].

Various techniques are applied around the world to generate MP from solar irradiation by researchers. P&O and INC methods are mostly used to track the peak power because of their simplicity and low maintenance [3–6]. Artificial neural network algorithm-based MPPT system has more advantages compared to conventional MPPT techniques [7]. Conventional methods have some disadvantages likes as they are slow in tracking the peak point of power when there is a sudden change in irradiation [8]. Different fuzzy logic techniques [9–13] and different optimization techniques such as particle swarm optimization and optimization of the colony were

B. Roy (✉) · S. Adhikari
Department of Electrical Engineering, National Institute of Technology Manipur, Imphal, Manipur 795001, India
e-mail: bapparoy11101994@gmail.com

A. D. Devi
Department of Electrical Engineering, National Institute of Technology Silchar, Assam 788010, India

K. J. Devi
Department of Electronics and Communication Engineering, National Institute of Technology Manipur, Imphal, Manipur 795001, India

applied to get a fast response in tracking. Machine learning [14] and deep learning-based [15] MPPT techniques have also been applied for tracking the peak power by predicting the reference voltage.

This paper proposes an MPPT system based on the new Elman neural network to extract MP from irradiation and improve the operation of the solar-battery system. The studied model consists of a PV panel along with an energy-storing system linked to a three-phase load through a DC/AC inverter. Both the PV and the battery are linked to two different boost converters to manage power flow. For the real-time verification of the studied model, real-time design and simulation are performed in RT-LAB with a real-time simulator OP4510.

2   System Configuration and Modeling

An islanded mode solar PV model is given in Fig. 1. A solar PV panel and battery energy storage are included in the studied model to feed the load.

The components of the studied model are described below.

2.1   Solar PV and Boost Converter

The solar system's output is linked to a converter (boost) as shown in Fig. 1 to rise the voltage level for controlling the output. A solar cell equivalent circuit is given in Fig. 2. It consists of a source of current, diodes, resistance in series, and resistance in parallel.

The mathematical formulation for the current output of a cell of the PV panel is expressed in Eq. (1).
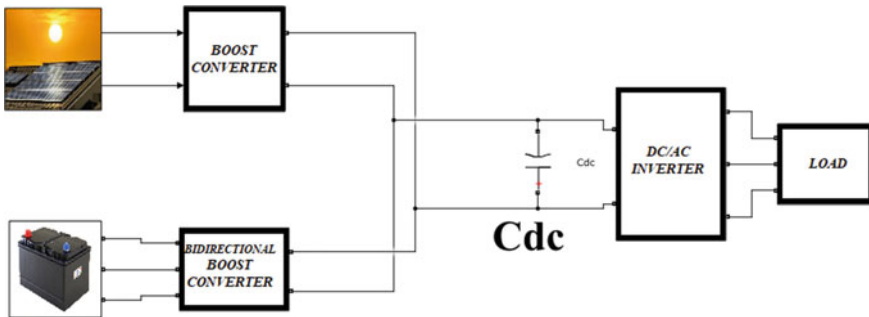
Fig. 1 Typical islanded solar PV battery system

Fig. 2 Solar cell equivalent circuit
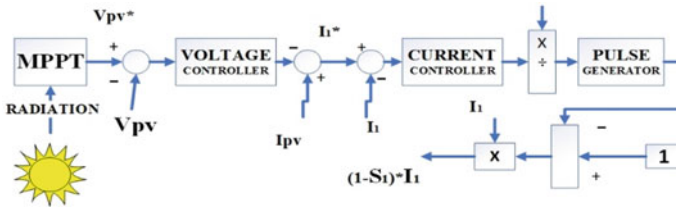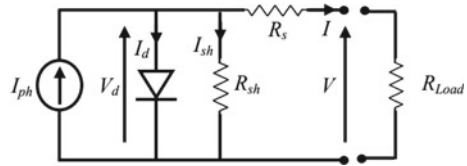
Fig. 3 DC-DC boost converter of solar

$$I = I_{\text{ph}} - I_d \cdot \left[\exp\left(\frac{q.(V + R_s.I)}{nKT} \right) - 1 \right] - \frac{V + R_s.I}{R_{\text{sh}}}, \tag{1}$$

where the current of a solar array is denoted by I, the output voltage is denoted by V, the light-generated current is represented by I_{ph}, I_d denotes a diode reverse saturation current, the electronic charge is denoted by q, n is the deviation factor of an ideal diode, K is the constant (Boltzmann's), the cell temperature is denoted by T, R_s is the resistance connected in series, and R_{sh} is the resistance connected in parallel. Figure 3 shows the working strategy for the regulation of the PV boost converter.

The input, solar irradiance led through the maximum power point tracker to calculate the PV reference voltage ($V_{\text{pv}}{}^*$) along with its maximum power. The actual voltage signal is compared with the calculated voltage (reference) and goes through a voltage PI controller for finding the current reference. Then the error in current goes through a current controller to produce the controlled voltage for monitoring the unidirectional converter of the solar unit as shown in Fig. 3.

2.2 Solar PV MPPT and New Elman Neural Network

MPPT allows the PV panel to track the MP point by computing the reference voltage. The maximum voltage at which PV panels can extract optimum power is called the peak power point. In this study, a new Elman neural network-based algorithm is used to find the MP point. The components for building the network are given in Table 1, and a training performance plot is given in Fig. 4.

For the training of the network, three layers of activation function along with one output function are applied. The first layer is the Tansig (activation) function with 100 hidden neurons, the second layer is the Logsig (activation) function with

Table 1 New Elman neural network components

Number of layers	Training algorithm	Max epochs	Training rate
4	Scaled conjugate gradient	5000	0.05
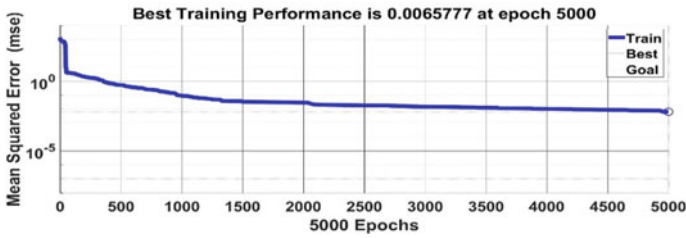
Fig. 4 Network training performance

50 hidden layers, the third layer is the Tansig (activation) function with 25 hidden layers, and the fourth layer or the output layer is the Purelin linear activation function with one hidden layer. A total of 10,000 historical data is been considered for the training of the network, where 15% of data is considered for testing and 15% of data is considered for validation.

2.3 Battery Energy Storage System and Bidirectional Boost Converter

Depending upon the weather condition, the solar unit gives its output, and because of that, the supply of power can become unstable. To overcome this problem of solar PV, an energy storage system is installed with the solar unit in the studied model. A bidirectional converter is connected to the battery to regulate the power flow depending on the surplus energy and load demand. The bidirectional boost converter control scheme is shown in Fig. 5.

The energy storage system stores energy when the generation of power is more than the demand, and it starts discharging when the solar power generation is lower
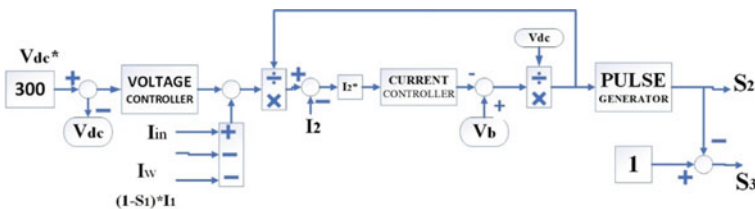
Fig. 5 Bidirectional DC/DC boost converter control

than the load. In control of the DC/DC boost converter, the actual DC voltage is equated with the reference DC voltage to get the mismatch value and it goes over the DC-link controller (voltage) to generate the current reference value of the DC-link. By comparing the actual current and reference current, an error in battery current is obtained, and then for generating the controllable battery voltage, the current error goes through the current PI controller.

2.4 Real-Time Modeling of the Studied Simulink Model

For the real-time justification of the studied Simulink model, MATLAB/Simulink environment along with OPALRT-LAB is used to design the studied model. The simulator OP4510 is applied for the simulation. A diagram of the real-time setup for real operation is given in Fig. 6. Where a real-time target simulator, OP4510, is connected to the host system via an Ethernet cable, which includes MATLAB and RT-LAB.

For the real-time design of any MATLAB/Simulink model, models need to be divided into subsystems as shown in Fig. 7. There are a minimum of two subsystems: one for the computational parts of the model and another for the input-output components of the model. Depending upon the difficulty of the Simulink model and depending on the number of cores of the real-time simulator, the number of subsystems in the model can be more than two but, in a model, only one input–output subsystem can be made. Every input signal coming to any subsystem of the model should go through an 'OpComm' block, and here naming of the subsystems has an important role. Without giving the proper and particular names of the subsystems, the OpComm block cannot be added inside any subsystems. The name of the input–output subsystem must be the 'SC_subsystem,' the nearest subsystem (computational subsystem) name of the input–output subsystem must be 'SM-subsystem,' and the rest of the subsystems can be named as 'SS-subsystem.'
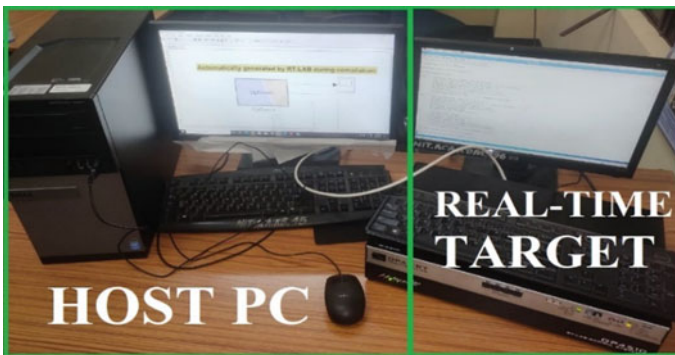
Fig. 6 Real-time simulation setup
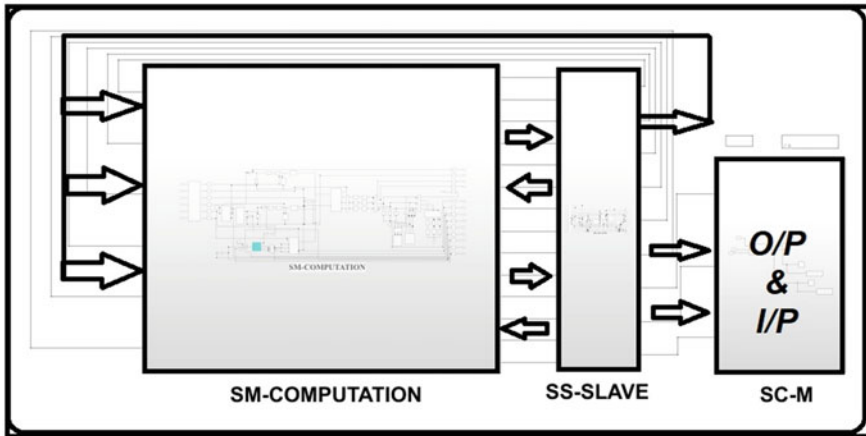
Fig. 7 Real-time modeling

3 Results and Discussion

The studied model is tested in dual simulation platforms, online mode, and offline mode. The online environment for the simulation included MATLAB/Simulink, and in the offline mode simulation, MATLAB/RT-LAB environment is considered. In the online mode, real-time simulation is carried out with the help of a real-time simulator, OP4510 as shown in Fig. 6.

3.1 Offline Mode

Offline mode responses are given in Fig. 8.

Simulation Time, $t = 0$ to 0.2 s: At starting 0.2 s of the simulation time, load is 1 kw and the solar is supplying nearly 2.2 kw of power. So, some surplus energy is available because the amount of generated solar power is more than the load. With that remaining energy, the battery starts charging itself. Figure 8a shows the PV, battery, and load power curves with simulation time.

Simulation Time, $t = 0.2$–0.4 s: At 0.2 s of simulation time, the load starts increasing and it reaches up to 2.2 kw. From 0.2 s to 0.4 s of simulation time, the load is around 2.2 kw, and at the same time, power from the solar PV is also 2.2 kw as shown in Fig. 8a. So, because of the lack of surplus power supply from the generation, the battery begins to discharge. Battery voltage and current response are shown in Fig. 8c, d.

Simulation Time, $t = 0.4$–0.6 s: At this time duration, the demand is at its peak, nearly 3 kw, and the solar PV output is only at around 2.2 kw. The energy storage system stops charging at this time duration and continuously discharges

(a)

(b)

(c)

(f)

(d)

(g)

(e)

(h)
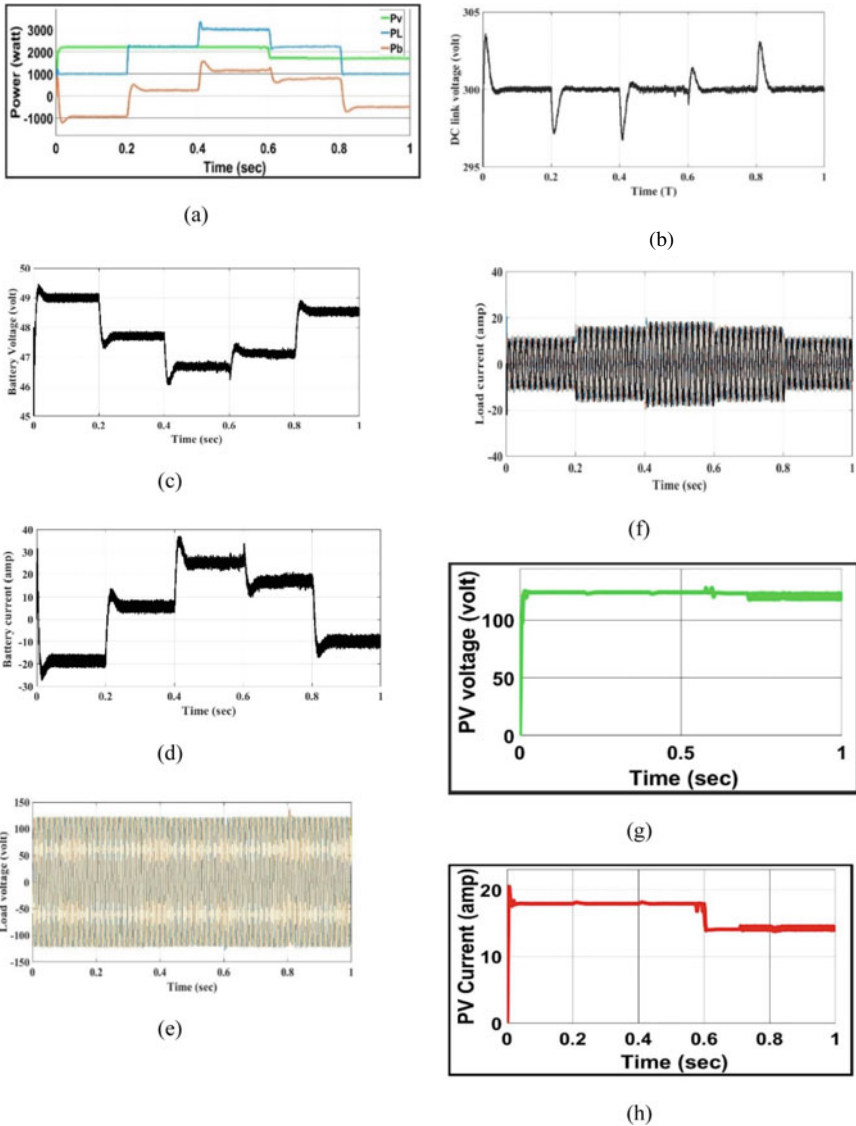
Fig. 8 **a** Power, **b** voltage (DC-link), **c** voltage (battery), **d** current (battery), **e** load voltage, **f** load current, **g** PV voltage, and **h** PV current

power because the generated power from the PV is less compared to the load. The DC-link voltage is maintained at 300 V as shown in Fig. 8b.

Simulation Time, $t = 0.6$–0.8 s: The load as well the PV generation of power drops at this time duration. The load power is 2.2 kw, and the solar power generation is nearly 1.7 kw. The generation of power is less than the load, and the required
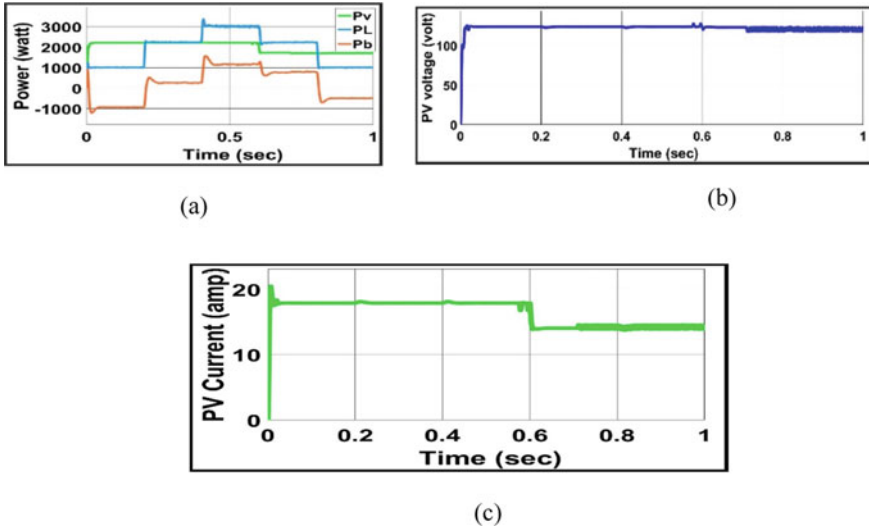
Fig. 9 **a** Power, **b** PV voltage, and **c** PV current

amount of power is provided by the battery storage system. The solar PV voltage and current curve with time are shown in Fig. 8g, h.

Simulation Time, $t = 0.8$–1 s: The load falls to 1 kw, but the generation of power from the PV stays at 1.7 kw. So, the energy storage starts charging by the surplus energy. Figure 8e, f show the load voltage and load current curve, respectively.

3.2 Online Mode

In this section, real-time simulation outcomes are discussed and given in Fig. 9 for the output power of the studied model, PV voltage, and PV current.

3.3 Comparison Between the Offline Mode and Online Mode Results

In this section, a comparison between the real-time (online mode) and MATLAB (offline mode) simulation results for the studied model is discussed in Table 2.

Table 2 Comparison of online and offline simulation

Sl. No.	Parameter	Offline mode	Online mode
1	Change in output response	Response of the studied model is taken after MATLAB/Simulink simulation and shown in Fig. 8. From Fig. 8a, it can be realized that the power flow is managed suitably depending on the load demand. So, the system worked as it was designed	For the online mode or real-time simulation of the studied model, OPALRT simulator along with MATLAB is used and shown in Fig. 9. It is visible by comparing Fig. 9a with Fig. 8a that both the power output curves are almost the same

4   Conclusion

A solar PV system including a battery energy storage system is built in MATLAB/Simulink environment with its coordination control to manage power flow. A vector control technique is implemented to achieve the control of the load-side converter. A new Elman neural network-based MPPT system is considered for the PV to track maximum power. Simulation results show that the studied system is working accurately, and the flow of power is managed accurately to fulfill the load demand. The studied model is also implemented in OPALRT-LAB with the help of a real-time simulator, OP4510, to validate the simulation results. It is observed from the comparison between the offline and online mode results that the PV source and battery system can feed the load according to its control technique in real time.

References

1. Kumar R, Choudhary A, Koundal G, Sing A (2017) Modelling/simulation of MPPT techniques for photovoltaic system using Matlab. IJARCSSE 7(4):178–187
2. Podder AK, Roy NK, Pota HR (2019) MPPT methods for solar PV systems: a critical review based on tracking nature. IET Renew Power Gener 13(10):1615–1632
3. Kamarzaman NA, Tan CW (2014) A comprehensive review of maximum power point tracking algorithms for photovoltaic systems. Renew Sustain Energy Rev 37:585–598
4. Reisi AR, Moradi MH, Jamasb S (2013) Classification and comparison of maximum power point tracking techniques for photovoltaic system: a review. Renew Sustain Energy Rev 19:433–443
5. Bendib B, Belmili H, Krim F (2015) A survey of the most used MPPT methods: conventional and advanced algorithms applied for photovoltaic systems. Renew Sustain Energy Rev 45:637–648
6. Srivastava A, Nagvanshi A, Chandra A, Singh A, Roy AK (2021) Grid integrated solar PV system with comparison between Fuzzy logic controlled MPPT and P&O MPPT. In: 2021 IEEE 2nd international conference on electrical power and energy systems (ICEPES), pp 1–6
7. Bendib B, Krim F, Belmili H, Almi MF, Bolouma S (2014) An intelligent MPPT approach based on neural-network voltage estimator and fuzzy controller, applied to a stand-alone PV system. In: 23rd International symposium on industrial electronics (ISIE), pp 404–409

8. Jyothy LP, Sindhu MR (2018) An artificial neural network based MPPT algorithm for solar PV system. In: 2018 4th International conference on electrical energy systems (ICEES), pp 375–380
9. Chen Y, Jhang Y, Liang R (2016) A fuzzy-logic based auto-scaling variable step-size MT method for PV systems. Sol Energy 126:53–63
10. Alajmi BN, Ahmed KH, Finney SJ, Williams BW (2011) Fuzzy-logic-control approach of a modified hill-climbing method for maximum power point in microgrid standalone photovoltaic system. IEEE Trans Power Electron 26:1022–1030
11. Robles Algarín C, Taborda Giraldo J, Rodriguez Alvarez O (2017) Fuzzy logic based MPPT controller for a PV system. Energies 10(12):2036
12. Rai RK, Rahi OP (2022) Fuzzy logic based control technique using MPPTfor solar PV system. In: 2022 First international conference on electrical, electronics, information and communication technologies (ICEEICT), pp 1–5
13. Hassan SZ, Li H, Kamal T, Arifoglu U, Mumtaz S, Khan L (2017) Neuro-fuzzy wavelet based adaptive MT algorithm for photovoltaic systems. Energies 10:394
14. Takruri M, Farhat M, Barambones O, Ramos-Hernanz JA, Turkieh MJ, Badawi M, AlZoubi H, Abdus Sakur M (2020) Maximum power point tracking of PV system based on machine learning. Energies 13(3):692
15. Phan BC, Lai YC, Lin CE (2020) A deep reinforcement learning-based MPPT control for PV systems under partial shading condition. Sensors 20(11):3039

Different Strategies for Estimation of State of Charge for Battery Packs of Electric Vehicle

Pooja Kumari, Durgesh Choudhary, Shuvam Sahay, and Niranjan Kumar

1 Introduction

The significant carbon gas emissions from internal combustion engine (ICE)-based conventional vehicles and industries are causing environmental pollution [1]. Internal combustion engines (ICEs) in conventional vehicles have a significant impact on this problem, causing the greenhouse effect and having a noteworthy emission rate. Because EVs using electrical energy to drive their motors efficiently, environmental researchers are paying close attention to the development of electric vehicles (EV) [2]. EVs are a promising technology for creating a sustainable transportation system in the long run because of their negligible to zero carbon emissions, minimal noise, tremendous effectiveness, and malleability in grid administration and connections [3]. In EV, battery is playing a wide role. One of the obstacles facing the development of electric vehicles is the state of battery technology. Research on battery management is vital, both theoretically and practically, in order to accurately predict the level of charge of a battery [4]. Thus, the concept of electric vehicles (EVs) led to the development of several energy storage technologies, including batteries, super- or ultra-capacitors, and fuel cells [5]. Hybrid electric vehicles (HEVs), a modification of these EVs, are a mix of the aforementioned sources with greater efficiency and appropriate assistance in maintaining continuous power supply in vehicles. Very cleverly, correct HEV installation can lower the rate of CO_2 and CO emissions. According to the power source, electric vehicles can generally be divided into the following categories: fuel cell vehicles (FCVs), photovoltaic electric vehicles (PEVs), plugin hybrid electric vehicles (PHEVs), battery-powered electric vehicles (BEVs), and solely hybrid electric vehicles (HEVs) [6].

P. Kumari (✉) · D. Choudhary · S. Sahay · N. Kumar
National Institute of Technology, Jamshedpur, Jharkhand, India
e-mail: pooja.student.kumari.1@gmail.com

PEV adoption will have a substantial impact on the electric load profiles of power systems, resulting in unpredictable and dynamic electricity demands that will put additional strain on the grid. Therefore, it is getting more difficult to coordinate EV charging systems by designing high-efficiency energy management controllers (EMC). Battery packs for EVs typically consist of dozens of individual cells connected in series and parallel [3]. Literature has suggested many algorithms for management of energy, including rule-based and optimization methods. The construction of rule-based systems involves use of pre-established policies to determine how much energy is coming from each power source [7]. As EV charging management applications involve highly dynamic systems, it is impossible to determine the best behavior for these systems using optimization approaches [8]. The literature is overflowing with in-depth analyses of the effects of PEV charging on the distribution system. Researchers have emphasized the advantage of deploying vehicle-to-grid (V2G) technology, which involves releasing some of the power stored in PEV batteries into the grid, to lower the peak demand of microgrid. It provides a summary of the effects of EV charging as well as the benefits and drawbacks of charging in V2G mode. Recent research has analyzed how V2G service impacts the rate at which the PEV battery degrades. A review of the effects of V2G on electrical power networks may be found in [9]. Recently, there has been a lot of interest in lithium-ion battery (LIB) packs, especially for use in EV applications. LIBs are placed in series and parallel to make the battery packs for EVs. The battery management system (BMS) controls and monitors SOC, SOH, SOT, etc., for each and every cell of the battery pack; hence, BMS becomes more complicated [10–12].

The ratio of the current stored to the cell's entire charge capacity is known as state of charge (SOC). In order to ensure that the driver receives accurate SOC data, a trustworthy SOC estimator must be integrated with the BMS because the SOC cannot be measured by using sensors directly [13]. Figure 1 represents the basic elements and primary duties of the BMS.
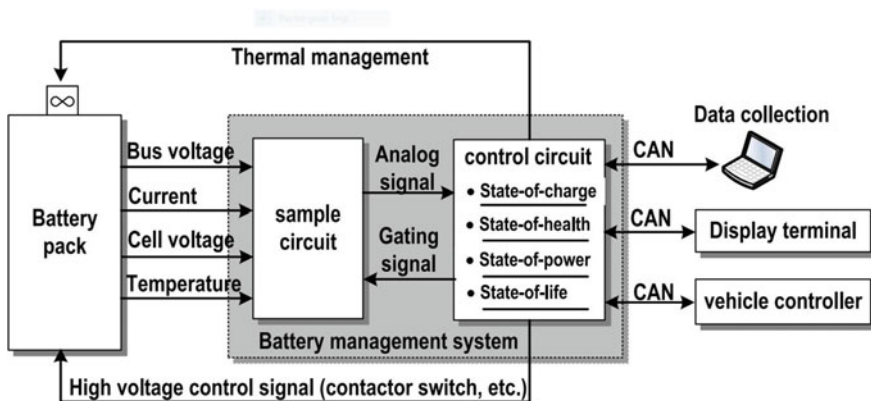
Fig. 1 BMS basic elements and primary duties

SOC of battery can typically be estimated using variety of methods. SOC can be calculated by utilizing adaptive filters based on battery model and observers and physically measured properties. These techniques include Kalman filters, the least squares, particle traditional filter, adaptive Luenberger observer, and black-box models based on machine learning in which measured data including voltage, current, temperature, and charge status are used to train machine learning models [7–14].

Due to the unavoidable variations in cell properties, the SOC of each cell must be taken into account when evaluating SOC for a battery pack. Procedures of SOC estimation are used by pack estimation algorithms, which may also include other characteristics to reduce computing cost and for boosting accuracy. The simplest estimation algorithms ignore variances in SOC of individual cells and aggregate all the cells to estimate SOC into one. It is also possible to utilize a unique SOC estimator for every cell, but the BMS may not be able to handle the computational complexity. To fix this, the pack may choose a reference cell and employ a higher bandwidth, more precise SOC estimating approach for this cell. The remaining cells may therefore employ lower bandwidth, less precise SOC algorithms, lowering the computational needs. Methods for estimating the SOC of a pack may also use difference models, which calculate the SOC of each cell [15].

2 Estimation of SOC for Battery Packs

Battery packs are often made up of number of cells to satisfy the needs of EVs [16]. The battery pack contains significant complicated properties. SOC of a battery pack is not fundamental in the same way as they are for a single cell. Precise capacity and SOC measurements of the battery pack are extremely difficult because the battery packs itself having unpredictable qualities in terms of battery capacity, resistance, voltage, and other aspects. A SOC estimation flow chart has been represented in Fig. 2.

Although several initiatives have been done to get precise SOC estimates for battery packs, they may be divided into several categories.

3 Cell-Based Calculation Methods

It typically has three different realizations:

(a) The "big cell" approach treats the battery pack as a single giant cell and calculates the battery's state of charge (SOC) using the voltage and current. However, the idiosyncrasies that affect a cell's efficiency have been overlooked. While it may be easier to calculate, it is evident that it cannot ensure the safe use of a battery pack.
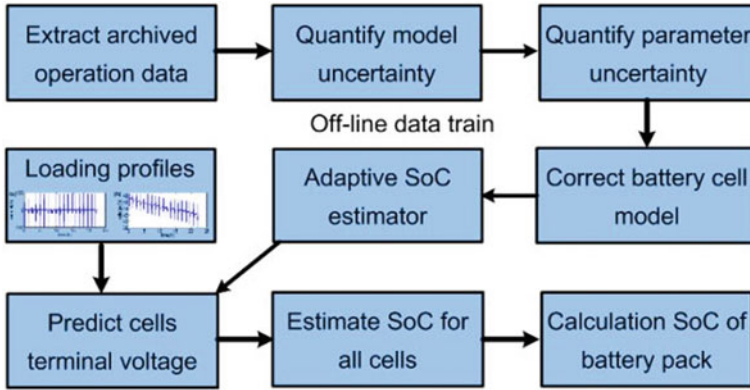
Fig. 2 SOC estimation flowchart

(b) "Short-board effect" method: This determines the battery pack's SOC using the extreme cell. The SOC of the battery pack is specifically shown by the cell with the lowest voltage during the discharge process and by cells with greatest voltage during the charging process. Obviously, it may increase the battery pack's safety, but for widely used battery packs with operating ranges of 30% to 80% SOC, this approach will lower the battery pack's energy use.

(c) One-by-one calculation method: Firstly, it determines the SOC of the battery pack before estimating the SOC for each cell in the battery pack. This sort of approach may provide the needed estimate accuracy, as expected. However, because of the high computational cost, BMS in EVs cannot use it.

4   Methods Based on Process of Screening

Given that each battery cell in the pack has the same capacity, resistance, etc., the SOC of a single cell is used to represent the SOC of the whole pack. Researchers have discovered that a further screening mechanism is employed to choose the battery cells used in the final battery pack. All of the SOC estimate errors were found to be less than 2% over the whole battery pack, as seen by the results. As time passes, this method's performance will drop below that of the "big cell" technique. As the disparities between the cells grow, so does the uncertainty in the estimate.

5   Methods for Bias Correction

The SOC estimation is completed using the corrected model after first creating a nominal model and after that using the bias-correction approach to detect the discrepancy between the nominal model and the battery cell as shown in Fig.3.
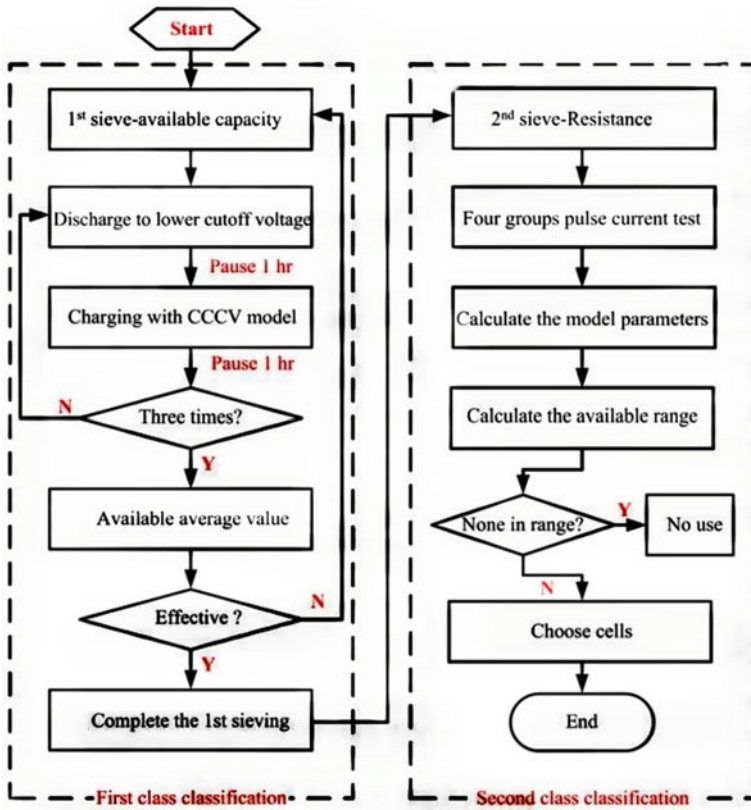
Fig. 3 Cells filtering approach

The SOC may be determined using the SOC of each cell [17].

$$U_t^t = U_{oc} - U_{D1} - \ldots - U_{Dn} - i_L R_i + \delta\left(C_{\text{rated}}^j, Z^j, \Delta Q^j\right),$$

where δ is the function of cell charge–discharge rate. It should be noticed that the battery pack's cell number is indicated by the superscript j. To determine model and parameter uncertainty is a known issue. In this research, a response surface approximation approach for the calculation of the bias function was developed using RBF neural network. The results reveal that this strategy performs well. It should be emphasized that either the average pack or the special cell model may be used to construct the nominal model. This approach may enhance the battery model's real-time performance while lowering the computational cost.

It is an approach that shows promise for resolving the strong time-varying, nonlinear, and non-uniform properties of battery packs. The computational cost must

Table 1 Tabular representation of different types of methods for SOC estimates of battery packs

Methods		Remarks
Cell-based calculation methods	Big cell method	• Less calculations needed • It cannot guarantee the battery pack's safety application
	Short-board effect method	• It may increase the battery pack's safety • This approach will lower the battery pack's energy use
	One-by-one calculation method	• High computational cost and accuracy
Methods based on process of screening		• SOC estimation errors of less than 2% • The estimated inaccuracy will increase as the cell's differences become larger
Methods for bias correction		• Lowering the computational cost • Promise for resolving the strong time-varying, nonlinear, and non-uniform properties of battery packs

be significantly decreased if an electric vehicle has a large enough number of battery cells (Table 1).

6 Issues and Challenges

Li-ion batteries have a lot of advantages, including high potential, density, small size, absence of effect of memory, quick charging, handling of high load, wider range of operating temperature, absence of need for air conditioning, high life cycle, extended service, and replacement intervals. Hence, it is necessary to identify the issues and find the appropriate remedies in order to make the battery to be an acceptable replacement for other batteries.

The excitability of Li-ion batteries, custom and cost issues, the memory effect for partially multiple sequential charging cycles, environmental impact, and recycling are some of these issues [18]. Figure 4 depicts how temperature affects the charge capacity of several rechargeable batteries during current feed and release. The following are brief sketches of the major difficulties.

(1) **Temperature**

All batteries often experience temperature generation brought on by chemical reactions. Unusual temperature destroys the battery and damages its chemical components. More importantly, the secondary lithium battery must have the temperature
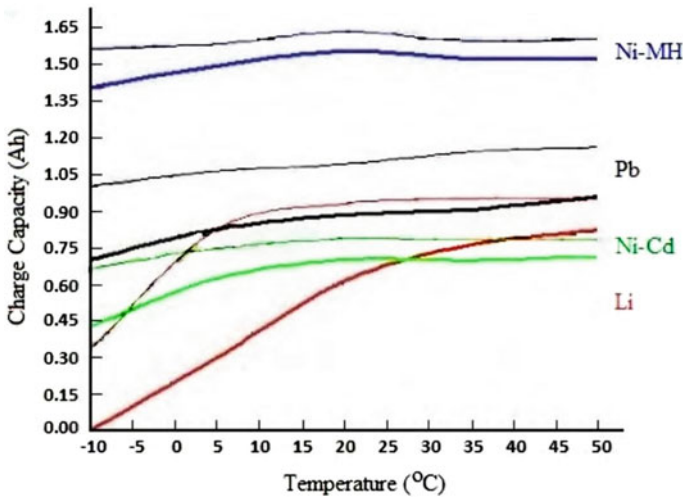
Fig. 4 Variations of capacity for different batteries with temperature

control system. The battery must often be used in both high and low temperatures. With regard to temperature, the battery's charging and discharging current as well as power handling capabilities are decreased for the low-temperature impact due to the low rate of reactions [19]. On the other side, a greater battery temperature creates certain challenging conditions that result in anomalous chemical behavior and cause the battery to explode. Ignition processes may leads to decrease in power, and results in an increase in temperature and the occurrence of a thermal runway. By adopting the proper safeguards, the thermal runaway process must be prevented. Unlike standard batteries, the Li-ion battery's capacity increases with temperature; however, this behavior shortens the battery's lifespan [20].

(2) **Safety**

Li-ion battery safety is a major worry for engineers and scientists. Due to its explosive properties, unlike NiCd batteries, LIBs are normally used with high power demands and considerable safety precautions. In addition, a single Li-ion cell is never used. The Li-ion battery needs extra components and circuits to monitor the battery's state and functioning while safeguarding against faults. Because of this, creating LIBs is more challenging than creating and organizing batteries. The risk factors for a cell that is overcharged or undercharged as compared to the SOC standard threshold are shown in Fig.5. BMS may be used to safely operate rechargeable battery cells while charging–discharging within the optimal operating range of SOC which is 20–90% [21].

(3) **Life Cycle**

Each cell's loss function due to voltage and heat effects has been examined; however, these losses still shorten the battery's useful life. It is important to note that a battery
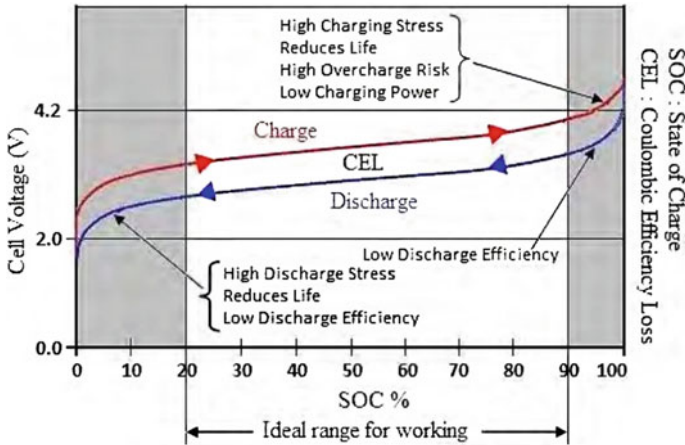
Fig. 5 SOC estimation for protection of battery cell

cell has an irreversible capacity loss if it is used outside of its normal working range. The cumulative effect of this anomaly is to shorten battery life, which might eventually lead to the battery's full failure. How much time cells spend at extreme temperatures and the chemistry they exhibit [22] determine the actual life. The lifetime drops gradually below 10 degrees Celsius owing to anode plating, while it decreases dramatically beyond 60 °C due to chemical breakdown. Due to reduced lifetime and wasteful thermal management, the recommended operating temperature should be in range.

(4) **Memory Effect**

Memory impact causes the battery to carry less memory. The impact is brought on by irregular charging and discharging patterns or by several cycles of charging and discharging during normal operation [23]. In electric vehicles (EVs), this happens when the engine is in generator mode and charges the battery during braking activities, but subsequently discharges the battery during acceleration. The memory effect is triggered in EV batteries by the repetitive charging and discharging that occurs intermittently. The voltage profile of a LIBs may be used to notice this since the battery seems to be quickly losing charge due to voltage depression or variation [24].

(5) **Cost**

Ten years ago, the cost of lithium was quite expensive. However, owing to increased use, demand, and manufacturing at the moment, its price is now more negotiable. A Li-ion battery pack costs between 25 and 30% of what an electric automobile costs [25]. LIB's market prices are falling over the last few years; the price has decreased by over 15% in each of the previous three years. According to a McKinsey & Company report, it might decrease by more than 25% by 2020. According to Bloomberg New Energy Finance's forecast, the cost of a LIB will drop to one-fourth of its current price

by 2030, and its production will increase as a result of the decline in its manufacturing cost. The market demand for consumer utilities, technological breakthroughs, their usage, the expansion of recycling, geopolitical concerns, budget allocation, environmental effect, etc., will all have a significant influence on the supply of lithium and its price.

(6) **Environmental Impact**

Oil costs are rising, and there is a huge need of energy for sustainable transportation, which has led to a trend toward electrification in cars like EVs, HEVs, and PHEVs Toyota predicts that, by 2020, more than 7% of all worldwide transportation will be by electric vehicle [26]. While reducing the need for oil-based cars is great for the environment, the LIB's production and disposal processes result in carbon dioxide emissions [27]. According to the findings, those who deal with cobalt and nickel metal compounds throughout their production, processing, and use may be at risk for developing respiratory, pulmonary, and neurological diseases [28]. By reducing demand for nickel and cobalt and saving virgin resources, recycling technology developed by a LIB might make this threat obsolete [29].

(7) **Recycling**

LIBs are more in demand right now since they will be used in future transportation and other applications [30]. Hydrometallurgical recovery, pyrometallurgical recovery, and direct recycling are the three fundamental recycling methods. The pyrometallurgical recovery process operates at high temperature, and direct recycling enables a larger proportion of battery materials to be recovered. Batteries that have reached the end of their lifespans might perhaps be reconditioned or given new life with fresh electrolytes in the near future [31].

7 Conclusion

A crucial feature of a BMS used in electric cars is battery SOC estimate. In this study, common SOC estimation techniques are reviewed and compared. Different SOC estimation techniques for battery cells and battery packs have been extensively reviewed and summarized. LIBs are now gaining popularity as alternative energy sources for EVs, electronic gadgets, tools, and utilities. Devices powered by LIBs do not need regular recharges because of their great energy density. These batteries are a popular option for usage in portable devices because of their small weight. Since their release, Li-ion battery market prices have been steadily declining. By 2030, prices are expected to have decreased much further, to be around 75% of what they are now. However, the cost is still high.

This research provided a thorough analysis of Li-ion batteries and other battery types. Additionally, a performance comparison is provided. Li-ion battery safety issues and harmful environmental impacts were discussed in this assessment with

some criticism. In order to enhance the performance of the LIBs in EV applications, a general overview of BMS is provided. For future research and development, issues and problems related to safety, environmental consequences, performance, and applications are underlined. Through this article, the many noteworthy and focused recommendations for future technical advancement and Li-ion battery production have been emphasized.

- By putting each cell through a process of deep charge–discharge cycle, while employing the best charge–discharge management approach, the memory effect might be reduced.
- Utilizing lithium material electrode to its full potential in accordance with proper demand and application is crucial.
- Due to environmental and health concerns, there is a need for advanced and regulated processes during battery manufacturing and disposal.
- Due to safety concerns, it is necessary to take into account the development of BMS technologies with low maintenance and operating costs.
- In most nations, it is crucial to have access to processing equipment for battery recycling in order to reserve raw resources. So, it should have to be solved.
- The use of refurbished used batteries has to be explored to increase second-hand battery use.

These recommendations provide a significant contribution to the development of LIB technology and made the environment pollution free by using electric vehicles.

References

1. Ramoni MO, Zhang H-C (2013) End-of-life (EOL) issues and options for electric vehicle batteries. Clean Technol Environ Policy 15(6):881–891
2. Hu X, Zou C, Zhang C, Li Y (2017) Technological developments in batteries: a survey of principal roles, types, and management needs. IEEE Power Energy Mag 15(5):20–31. https://doi.org/10.1109/MPE.2017.2708812
3. Zhang M, Fan X (2020) Review on the State of charge estimation methods for electric vehicle battery. World Electr Veh J 11:23. https://doi.org/10.3390/wevj11010023
4. Hu L, Ye Y, Bo Y, Huang J, Tian Q, Yi X, Li Q (2022) Performance evaluation strategy for battery pack of electric vehicles: online estimation and offline evaluation. Energy Rep 8(Supplement 4):774–784. ISSN 2352-4847. https://doi.org/10.1016/j.egyr.2022.02.026, https://www.sciencedirect.com/science/article/pii/S2352484722002736
5. Hannan MA, Hoque MM, Peng SE, Uddin MN (2017) Lithium-ion battery charge equalization algorithm for electric vehicle applications. IEEE Trans Ind Appl 53(3):2541–2549. https://doi.org/10.1109/TIA.2017.2672674
6. Manzetti S, Mariasiu F (2015) Electric vehicle battery technologies: from present state to future systems. Renew Sustain Energy Rev 51:1004–1012. ISSN 1364-0321
7. Xing Y, Ma EWM, Tsui KL, Pecht M (2011) Battery management systems in electric and hybrid vehicles. Energies 4(11):1840–1857
8. Zhang C, Wang LY, Li X, Chen W, Yin GG, Jiang J (2015) Robust and adaptive estimation of state of charge for lithium-ion batteries. IEEE Trans Ind Electron 62(8):4948–4957. https://doi.org/10.1109/TIE.2015.2403796

9. Das S, Acharjee P, Bhattacharya A (2021) Charging scheduling of electric vehicle incorporating grid-to-vehicle and vehicle-to-grid technology considering in smart grid. IEEE Trans Ind Appl 57(2):1688–1702. https://doi.org/10.1109/TIA.2020.3041808

10. Zhang W, Wang L, Wang L, Liao C, Zhang Y (2022) Joint state-of-charge and state-of-available-power estimation based on the online parameter identification of lithium-ion battery model. IEEE Trans Ind Electron 69(4):3677–3688. https://doi.org/10.1109/TIE.2021.3073359

11. Bruen T, Marco J (2016) Modelling and experimental evaluation of parallel connected lithium ion cells for an electric vehicle battery system. J Power Sources 310:91–101. ISSN 0378-7753

12. Rothgang S, Baumhöfer T, van Hoek H, Lange T, De Doncker RW, Sauer DU (2015) Modular battery design for reliable, flexible and multi-technology energy storage systems. Appl Energy 137:931–937 ISSN 0306-2619

13. Yang R, Xiong R, He H, Mu H, Wang C (2017) A novel method on estimating the degradation and state of charge of lithium-ion batteries used for electrical vehicles. Appl Energy 207:336–345. ISSN 0306-2619

14. Álvarez Antón JC, García Nieto PJ, Blanco Viejo C, Vilán Vilán JA (Dec 2013) Support vector machines used to estimate the battery state of charge. IEEE Trans Power Electron 28(12):5919–5926. https://doi.org/10.1109/TPEL.2013.2243918

15. Naguib M, Kollmeyer P, Emadi A (2021) Lithium-Ion battery pack robust state of charge estimation, cell inconsistency, and balancing: review. IEEE Access 9:50570–50582. https://doi.org/10.1109/ACCESS.2021.3068776

16. Xiong R, Sun F, Gong X, He H (2013) Adaptive state of charge estimator for lithium-ion cells series battery pack in electric vehicles. J Power Sources 242:699–713. ISSN 0378-7753

17. Sun F, Xiong R (2015) A novel dual-scale cell state-of-charge estimation approach for series-connected battery pack used in electric vehicles. J Power Sources 274:582–594. ISSN 0378-7753

18. Xiong R, Tian J, Mu H, Wang C (2017) A systematic model-based degradation behavior recognition and health monitoring method for lithium-ion batteries. Appl Energy 207:372–383. ISSN 0306-2619

19. Tarascon J-M et al (2010) Hunting for better li-based electrode materials via low temperature inorganic synthesis. Chem Mater 22(3):724–739

20. Lv S, Wang X, Lu W, Zhang J, Ni H (2022) The influence of temperature on the capacity of lithium ion batteries with different anodes. Energies 15(1):60. https://doi.org/10.3390/en15010060

21. Berecibar M, Gandiaga I, Villarreal I, Omar N, Van Mierlo J, Van den Bossche P (2016) Critical review of state of health estimation methods of Li-ion batteries for real applications. Renew Sustain Energy Rev 56:572–587. ISSN 1364-0321

22. Hoque MM, Hannan MA, Mohamed A (2016) Optimal CC-CV charging of lithium-ion battery for charge equalization controller. 2016 international conference on advances in electrical, electronic and systems engineering (ICAEES), pp 610–615. https://doi.org/10.1109/ICAEES.2016.7888119

23. Huggins RA (2006) Mechanism of the memory effect in Nickel electrodes. Solid State Ionics 177:2643–2646

24. Sasaki T, Ukyo Y, Novák P (2013) Memory effect in a lithium-ion battery. Nat Mater 12(6):569–575. https://doi.org/10.1038/nmat3623. Epub 2013 Apr 14 PMID: 23584142

25. Xiaoming Zhu, Xiaoling Liu, Wenwen Deng, Lifen Xiao, Hanxi Yang, Yuliang Cao, Perylenediimide dyes as a cheap and sustainable cathode for lithium ion batteries, Materials Letters, Volume 175,2016, Pages 191–194, ISSN 0167-577X.

26. Fan B, Chen X, Zhou T, Zhang J, Xu B (2016) A sustainable process for the recovery of valuable metals from spent lithium-ion batteries. Waste Manage Res 34(5):474–481. https://doi.org/10.1177/0734242X16634454

27. Dunn JB, Gaines L, Sullivan J, Wang MQ (2012) Impact of recycling on cradle-to-gate energy consumption and greenhouse gas emissions of automotive lithium-ion batteries. Environ Sci Technol 46(22):12704–12710. https://doi.org/10.1021/es302420z. Epub 2012 Oct 30 PMID: 23075406

28. Tolomeo R, De Feo G, Adami R, Sesti Osséo L (2020) Application of life cycle assessment to lithium ion batteries in the automotive sector. Sustainability 12(11):4628
29. Yuan X, Liu X, Zuo J (2015) The development of new energy vehicles for a sustainable future: a review. Renew Sustain Energy Rev 42:298–305. ISSN 1364-0321
30. Gaines L (2014) The future of automotive lithium-ion battery recycling: charting a sustainable course. Sustain Mater Technol 1–2:2–7. ISSN 2214-9937
31. Shahriar SM, Bhuiyan EA, Nahiduzzaman M, Ahsan M, Haider J (2022) State of charge estimation for electric vehicle battery management systems using the hybrid recurrent learning approach with explainable artificial intelligence. Energies 15:8003. https://doi.org/10.3390/en15218003
32. Li J, Wang L, Lyu C, Pecht M (2017) State of charge estimation based on a simplified electrochemical model for a single LiCoO$_2$ battery and battery pack. Energy, Elsevier 133(C):572–583

Novel Current Mode First-Order Filter and Oscillator

Ashok Kumar and Ajay Kumar Kushwaha

1 Introduction

First-order filters play a key role in the construction of higher order filter topologies and noise reduction in analogue signal processing circuits. In some applications, first-order LP and HP filters with different APFs are employed to separate signals based on frequency. While the signal's amplitude is held constant, its phase is altered. The unwelcome phase change that results from signal processing ideas is equalised by the first-order AP-type filter. Other uses include making multiphase and quadrature oscillators as well as premium factor frequency selective filters. For LP and HP filters, first-order, second-order, third-order and so on, are all potential orders. While band-pass filters and band-stop filters can be created by combining two first-order filters, they must be second-order or higher.

Due to the many advantages offered by CM signal processing, first-order filters using various current mode ABBs have recently been developed [1–12]. Key advantages of signal processing in the current domain include reduced complexity, high frequency operation and low voltage operation. A new era in electronically controlled filter design was particularly introduced by DX-MOCCII [11, 12]. A CM filter should, theoretically, have low input and high output impedances to make cascading with other CM circuits easier.

The design of a quadrature oscillator and first-order universal filter is discussed in Sect. 2. We present the simulation results in Sect. 3. The paper is concluded in Sect. 4.

A. Kumar (✉)
National Institute of Technology, Srinagar, India
e-mail: ashok.kumar@nitsri.ac.in

A. K. Kushwaha
Bharati Vidyapeeth (Deemed to Be University) College of Engineering, Pune, India

2   Circuit Description

There are several uses for the DX-MOCCII, a particularly practical and versatile analogue building block, in analogue signal processing [11, 12]. Figures 1 and 2 depict the DX-MOCCII symbolic representation and CMOS structure, respectively. The below matrix can be described the optimal DX-MOCCII equation.

$$
\begin{bmatrix} I_Y \\ V_{XP} \\ V_{XN} \\ I_{ZP1} \\ I_{ZP2} \\ I_{ZN1} \\ I_{ZN2} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} V_Y \\ I_{XP} \\ I_{XN} \end{bmatrix}
\tag{1}
$$

It has non-inverting (X_P) and inverting (X_N) terminals that reflect current to the appropriate Z terminals $(Z_{P1}, Z_{P2}$ and $Z_{N1}, Z_{N2})$. High output impedance at Z_{P1},

Fig. 1 Symbol of DX-MOCCII
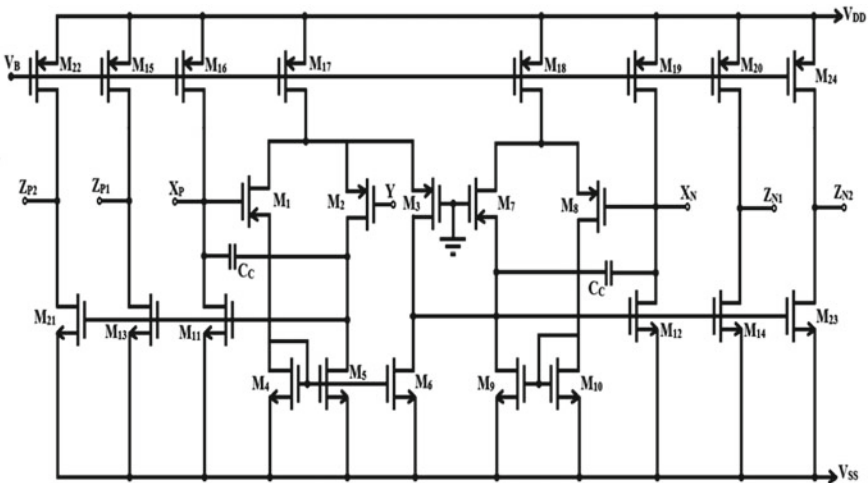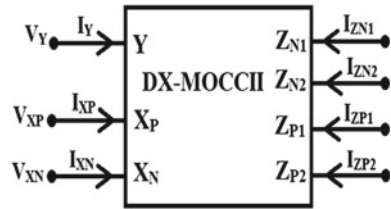
Fig. 2 MOS-based DX-MOCCII

Z_{P2} and Z_{N1}, Z_{N2} terminals. Low impedance at X_P and X_N terminals, and high input impedance at the Y terminal make up the active building block.

2.1 Presented First-Order Universal Filter

Figure 3 shows a new CM first-order multifunction filter using one DX-MOCCII, one capacitor and two grounded resistors.

Routine study of Fig. 3 results in the following transfer function, assuming ideal DX-MOCCII.

$$\frac{I_{LP}}{I_{IN}} = \frac{I_{R1}}{I_{IN}} = \frac{1/C_1 R_1}{S + 1/C_1 R_1} \tag{2}$$

$$\frac{I_{HP}}{I_{IN}} = \frac{I_{C1}}{I_{IN}} = \frac{s}{S + 1/C_1 R_1} \tag{3}$$

$$\frac{I_{AP}}{I_{IN}} = \frac{S - 1/C_1 R_1}{S + 1/C_1 R_1} \tag{4}$$

From Eqs. (2, 3 and 4), which are derived from the proposed circuit, it is observed that CM performs LP, HP and AP (assuming $R_1 = 2R_2$) operation with pole frequency as

$$\omega_\mathrm{o} = 1/C_1 R_1 \tag{5}$$

The value of the capacitor (C_1) individually can be changed to tune the pole frequency of the filter, and the phase of the AP filter is calculated as

$$\angle\phi = -2 \tan^{-1} \frac{\omega}{C_1 R_2} \tag{6}$$

Fig. 3 CM first-order multifunction filter
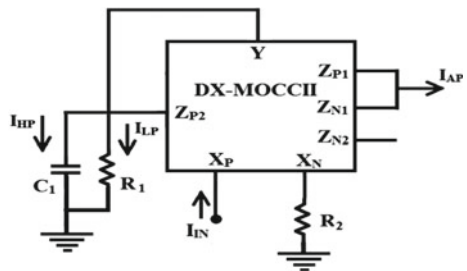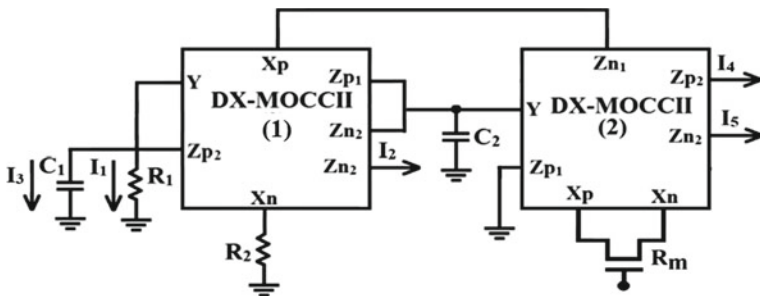
Fig. 4 Proposed oscillator circuit

From Eq. (6), the suggested all pass filter offers a phase shift at the output terminal (I_{AP}) between $0°$ and $-180°$.

The filters in Eq. (5), have the following pole ω_o sensitivities:

$$S_{C1}^{\omega_o} = -1 \quad \text{and} \quad S_{R_1}^{\omega_o} = -1 \tag{7}$$

from Eq. (7), that pole frequency (ω_o) sensitivities are low, or unity in magnitude.

2.2 *Proposed Quadrature Oscillator*

The proposed CM sinusoidal oscillator is presented in Fig. 4. The oscillator design uses two DX-MOCCII blocks and passive components.

The characteristic equation for the oscillator circuit in Fig. 4 is:

$$s^2 + s\left[\frac{1}{R_1 C_1} - \frac{2}{R_m C_2}\right] + \frac{2(R_1 - R_2)}{R_1 R_2 R_m C_1 C_2} = 0 \tag{8}$$

The condition of oscillation can be expressed as.

$$\text{Co:} \quad R_m C_2 \geq 2 R_1 C_1 \tag{9}$$

The corresponding oscillation frequency (FO) is expressed as.

$$\text{Fo:} \quad \omega_o = \left[\frac{2(R_1 - R_2)}{R_1 R_2 R_m C_1 C_2}\right]^{1/2} \tag{10}$$

The proposed MSO circuit can generate five different currents (I_1, I_2, I_3, I_4 and I_5). It should be observed that all current outputs are available at terminals with high impedance, making it simple to cascade them into other circuits. It shows that the phase of the output currents is multiphase in nature.
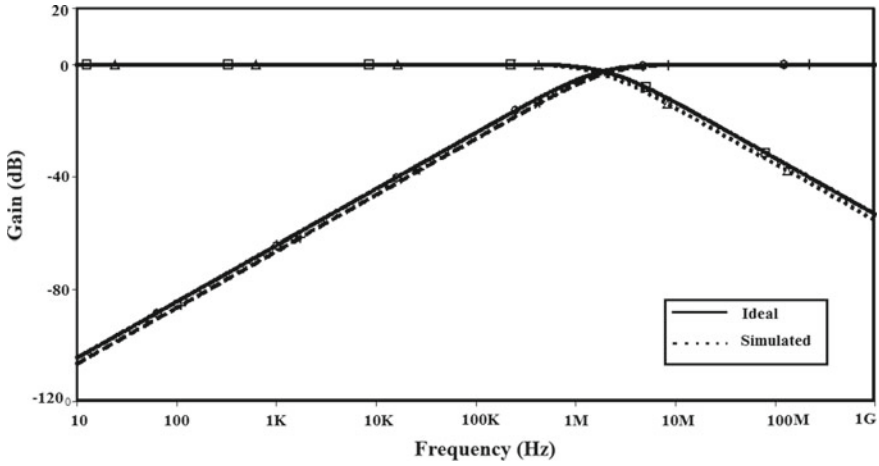
Fig. 5 Response of LP and HP filter

3 Simulation Results

The TSMC 0.25 μm CMOS model parameters SPICE simulation tool is used to evaluate the CM first-order universal filter presented in Fig. 3 to support the theoretical analysis. The MOS transistor aspect ratio is mentioned in Ref. [12]. $V_{DD} = -V_{SS} = 1.25$ V have been the DC supply voltages which been employed. The proposed filter is built with $C = 10$ pF, $R_1 = 7.5$ K and $R_2 = 3.75$ K for a pole frequency of $f_o = 2.12$ MHz. The observed LP and HP responses are shown in Fig. 5. The amplitude and phase responses of an AP filter are shown in Fig. 6. The results of the simulation and the theoretical analysis agree well.

The multiphase sinusoidal oscillator is designed using passive components $C_1 = 20$ pF, $C_2 = 40$ pF, $R_1 = 1$ kΩ $= R_3 = 5$ kΩ and $R_2 = 2.5$ kΩ, with an oscillation frequency (f_o) of 1.59 MHz, for analysis purposes. Figures 7 and 8 show the simulated outputs of a CM multiphase sinusoidal oscillator. The initial output response is shown in Fig. 7, and the steady state output response are shown in Figs. 8 and 9, respectively.
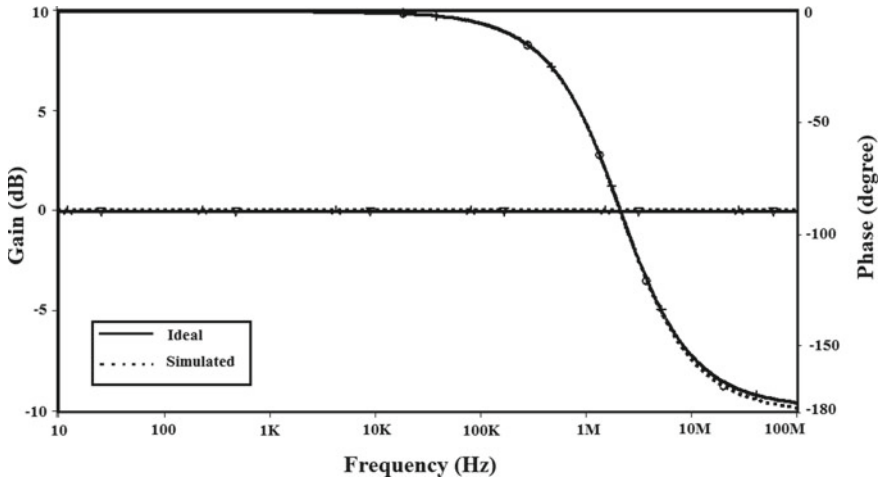
Fig. 6 Response of current mode AP filter
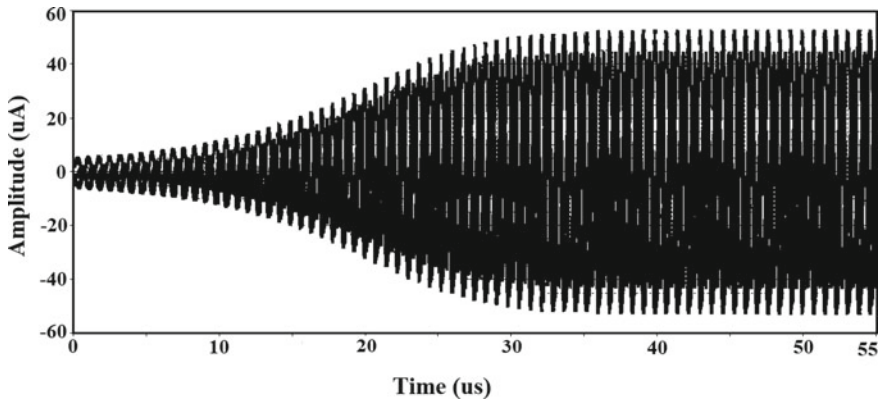
Fig. 7 Initial state of waveform
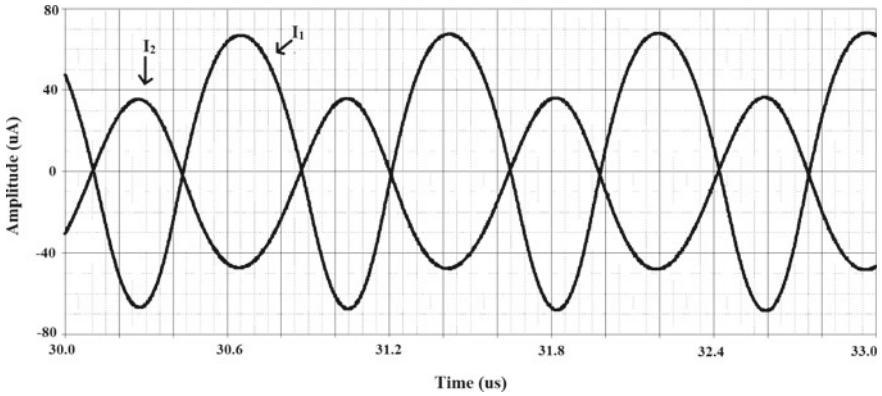
Fig. 8 Steady-state response of oscillator outputs as quadrature
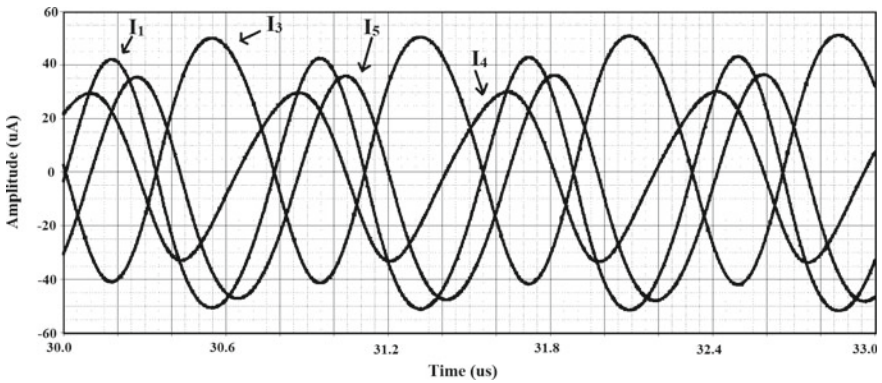
Fig. 9 Steady-state response of oscillator output

4 Conclusion

The DX-MOCCII, grounded capacitors and grounded resistors are combined to provide a novel current mode first-order universal filter. High pass, low pass and all pass responses for CM are provided by the described filters. Filter sensitivity is low. A new sinusoidal oscillator based on DX-MOCCII is also proposed. The oscillator frequency is also externally tunable without disturbing the circuit topology. The block has been designed with the PSPICE tool using 0.25 μm CMOS TSMC model parameters. The performance of the filter and oscillator is examined with the Cadence OrCAD simulator. Simulation results verify the theory.

References

1. Bhaskar DR, Raj A, Senani R, Kumar P (2022) CFOA-based simple mixed-mode first-order universal filter configurations. Int J Circuit Theory Appl 50(7):2631–2641
2. Chaturvedi B, Mohan J, Jitender (2022) First-order current-mode fully cascadable all-pass frequency selective structure, its higher-order extension and tunable transformation possibilities. J Circ, Syst Comput 31(2)
3. Rohilla K, Pushkar KL, Kumar R, Raj A (2022) Resistorless first-order universal filter structures employing OTAs with independent controllability of gain and pole frequency. IETE J Res 1–21
4. Chaturvedi B, Mohan J, Kumar A, Pal K (2022) Current-mode first-order universal filter and its voltage-mode transformation. J Circ, Syst Comput 29(09)
5. Yuce E, Minaei S (2021) A new first-order universal filter consisting of two ICCII+ s and a grounded capacitor. AEU-Int J Electron Commun 137
6. Yucel F (2021) A DVCC-based current-mode first-order universal filter. J Circ, Syst Comput 30(16)
7. Dogan M, Yuce E (2022) A first-order universal filter including a grounded capacitor and two CFOAs. Analog Integr Circ Sig Process 112:379–390
8. Barile G, Safari L, Pantoli L, Stornelli V, Ferri G (2021) Electronically tunable first order AP/LP and LP/HP filter topologies using electronically controllable second-generation voltage conveyor (CVCII). Electron 10(7)
9. Jaikla W, Buakhong U, Siripongdee S, Khateb F, Sotner R, Silapan P, Suwanjan P, Chaichana A (2021) Single commercially available IC-based electronically controllable voltage-mode first-order multifunction filter with complete standard functions and low output impedance. Sensors 21(21)
10. Kumar A, Kushwaha AK, Paul SK (2018) DXCCII-based first order voltage-mode all-pass filter. Lect Notes Electr Eng 436:709–717
11. Zeki A, Toker A (2003) The dual-X current conveyor (DXCCII): a new active device for tunable continuous-time filters. Int J Electron 89(12):913–923
12. Kumar A, Kushwaha AK, Paul SK (2021) Electronically tunable mixed mode quadrature oscillator using DX-MOCCII. J Circ, Syst Comput 30(01)

Eigenfrequencies of Functionally Graded Plates with Rectangular Cutout

Kushal Jana, Aditi Majumdar, and Salil Haldar

1 Introduction

Nowadays, functionally graded material (FGM) has become very popular over traditional composites for its various advantages. FGM (such as SiC-C, TiC-Ni) is used as a thermal barrier due to its ability to withstand high temperatures. Some FGM is able to limit fracture growth. So it is used as a penetration-resistant material in bullet-proof vests. FG porous composite structures are excellent for energy-absorbing systems, sound absorbers, construction materials, etc. The main difference between FGM and composite is gradually introducing one material into another without losing any of their respective properties. Vibration analysis of structural members such as plate is very important because if the natural frequency matches the external dynamic load, that can lead to catastrophic failures. Idea of FGM was introduced in 1984. Since then, there have been many studies on FGM plate using the Isogeometric method [1], Rayleigh–Ritz, Navier solution, FEM, Finite Strip Approach and dynamic stiffness method (DSM). Recently Kumar and Jana [2] used DSM on stepped FGM plates using FSDT. They have considered the displacement field concerning the physical surface and assumed levy-type edge conditions. Ramteke and Panda [3] did a free vibration analysis of multi-directional FGM structures using higher-order shear deformation theory with porosity in consideration. Tabatabaei and Fattahi [4] used Continuum 3D eight-noded reduced integration element for finite element analysis

K. Jana (✉) · S. Haldar
Indian Institute of Engineering Science and Technology, Shibpur, West Bengal, India
e-mail: iamjanakushal@gmail.com

S. Haldar
e-mail: salilhaldar@gmail.com

A. Majumdar
Techno India Institute, Salt Lake, West Bengal, India
e-mail: aditimech@gmail.com

for the FGM plate. But there have been very few papers, which considered gradual cutout size increase, rectangular cutouts, cracked corners, etc. on FGM plates. In this paper finite element formulation based on FSDT using a "nine-noded isoparametric element" has been implemented for the FGM plate.

2   Finite Element Formulation

$$
\begin{aligned}
U &= u_0(x, y) + z\theta_x(x, y) \\
V &= v_0(x, y) + z\theta_y(x, y) \\
W &= w(x, y)
\end{aligned}
\tag{1}
$$

The displacement field equation based on Mindlin's 1st order shear deformation theory as fellows [5, 9]: i.e., displacement is expressed in terms of $[u, v, w, \theta_x, \theta_y]$ where u and v are the axial displacement in two axial directions, w is the transverse displacement, θ_x and θ_y are the rotation of two orthogonal axes. From that, we get a strain displacement relationship

$$
\begin{aligned}
&\epsilon_x = u_{0,x} + z\theta_{x,x},\ \epsilon_y = v_{0,y} + z\theta_{y,y}, \\
&\gamma_{xy} = u_{0,y} + v_{0,x} + z(\theta_{x,y} + \theta_{y,x}), \\
&\gamma_{xz} = \theta_x + w_{,x},\ \gamma_{yz} = \theta_y + w_{,y}, \\
&\epsilon = \left[\epsilon_x \epsilon_y \gamma_{xy} \gamma_{xz} \gamma_{yz}\right]^{\mathrm{T}} = \left[\epsilon_x^0 \epsilon_y^0 \gamma_{xy}^0 \gamma_{xz}^0 \gamma_{yz}^0\right]^{\mathrm{T}} \\
&\quad + z\left[\epsilon_x^1 \epsilon_y^1 \gamma_{xy}^1 \gamma_{xz}^1 \gamma_{yz}^1\right]^{\mathrm{T}}
\end{aligned}
\tag{2}
$$

Stress–strain relationship for FGM is as below and k_s is the shear correction factor

$$
\left[\sigma_x \sigma_y \sigma_{xy} \sigma_{xz} \sigma_{yz}\right]^{\mathrm{T}} = \left[C\right]\left[\epsilon_x^0 \epsilon_y^0 \gamma_{xy}^0 \gamma_{xz}^0 \gamma_{yz}^0\right]^{\mathrm{T}}
\tag{3}
$$

where non-zero elements of $\left[C\right]$ are [2, 6],

$$
\begin{aligned}
C(1,1) &= C(2,2) = E(z)/(1 - v^2); \\
C(1,2) &= C(2,1) = vE(z)/(1 - v^2); \\
C(5,5) &= C(6,6) = k_s E(z)/2(1 + v) = k_s \frac{1 - v}{2} Q11
\end{aligned}
\tag{4}
$$

For P-FGM plate material properties (young's modulus, mass density) varies along thickness (z) [2, 4, 5]

$$
P(z) = P_m + (P_c - P_m)\left(\frac{1}{2} + \frac{z}{h}\right)^n \text{ where, } -\frac{h}{2} \leq z \leq \frac{h}{2}
\tag{5}
$$

Substituting the value of $E(z)$ layer-wise from (9), the coefficients of stiffness of the FGM plate can be obtained as below [2, 5, 7]:

$$A11 = \int_{-\frac{h}{2}}^{-\frac{h}{2}} C(1, 1)dz$$

$$= \int_{-\frac{h}{2}}^{\frac{h}{2}} \frac{E(z)}{1 - v^2}dz = \frac{1}{1 - v^2}(\int_{-\frac{h}{2}}^{\frac{h}{2}} \left((E_c - E_m)\left(\frac{1}{2} + \frac{z}{h}\right)^n + E_m\right)dz$$

$$B11 = \int_{-\frac{h}{2}}^{\frac{h}{2}} zQ11dz = \int_{-\frac{h}{2}}^{\frac{h}{2}} \frac{zE(z)}{1 - v^2}dz, D11 = \int_{-\frac{h}{2}}^{\frac{h}{2}} z^2Q11dz = \int_{-\frac{h}{2}}^{\frac{h}{2}} \frac{z^2E(z)}{1 - v^2}dz$$

$A_{12}, A_{66}, B_{12}, B_{11}, D_{11}, D_{12}, D_{66}$, etc. can be obtained similarly.
Now strain energy [5, 7],

$$U = \frac{1}{2}\int_{\Omega}\{\epsilon^{0T}A\epsilon^0 + \epsilon^{1T}B\epsilon^0 + \epsilon^{0T}B\epsilon^1 + \epsilon^{1T}D\epsilon^1 + \gamma^{0T}A_s\gamma^0\}d\Omega \qquad (6)$$

So, the strain energy of the element

$$U^e = \frac{1}{2}d^{eT}\int_{\Omega^e}\{B_m^{eT}AB_m^{(e)} + B_b^{eT}BB_m^e + B_m^{eT}BB_b^e + B_b^{eT}DB_b^e + B_s^{eT}A_sB_s^e\}d\Omega^e d^e \quad (7)$$

As, $\epsilon^0 = D_m u = D_m[N]d^e = B_m^e d^e$, $\epsilon^1 = B_b^e d^e$, $\gamma^0 = B_s^e d^e$.
Where $B_m^e(3 \times 45)$, $B_b^e(3 \times 45)$, $B_s^e(2 \times 45)$ are membrane, bending and shear components of the strain displacement matrix, respectively.

$$K^e = K_{mm}^e + K_{bm}^e + K_{mb}^e + K_{bb}^e + K_{ss}^e$$

$$= \int_{\Omega^e}\left(B_m^{eT}AB_m^e + B_b^{eT}BB_m^e + B_m^{eT}BB_b^e + B_b^{eT}DB_b^e + B_s^{eT}A_sB_s^e\right)d\Omega^e$$

$$= \int_{-1}^{1}\int_{-1}^{1}\left(B_m^{eT}AB_m^e + B_b^{eT}BB_m^e + B_m^{eT}BB_b^e + B_b^{eT}DB_b^e + B_s^{eT}A_sB_s^e\right)detJd\xi d\eta \quad (8)$$

The kinetic energy takes the form

$$KE = \frac{1}{2}\int_{V}\rho\left((\dot{u} + z\dot{\theta}_x)^2 + (\dot{v} + z\dot{\theta}_y)^2 + \dot{w}^2\right)dV$$

$$= \frac{1}{2}\int_{\Omega}\left(\mathcal{I}_0(\dot{u}^2 + \dot{v}^2 + \dot{w}^2) + 2\mathcal{I}_1(\dot{u}\dot{\theta}_x + \dot{v}\dot{\theta}_y) + \mathcal{I}_2(\dot{\theta}_x^2 + \dot{\theta}_y^2)\right)d\Omega$$

$$= \frac{1}{2}\int_{\Omega}\dot{u}^T[M]\dot{u}d\Omega \qquad (9)$$

where $\dot{u}^T = \begin{bmatrix} \dot{u}\dot{v}\dot{w}\,\dot{\theta}_x\,\dot{\theta}_y \end{bmatrix}$ and the inertia matrix [M] is a (5 × 5) symmetric matrix with non-zero elements,

$M(1, 1) = M(2, 2) = M(3, 3) = \mathcal{I}_0$, $M(1, 4) = M(2, 5) = \mathcal{I}_1$, $M(4, 4) = M(5, 5) = \mathcal{I}_2$ And the inertia terms can be calculated as [1, 8, 9]

$$\mathcal{I}_i = \int\limits_{-\frac{h}{2}}^{\frac{h}{2}} \rho(z)z^i dz = \int\limits_{-\frac{h}{2}}^{\frac{h}{2}} \left(((\rho_1 - \rho_2)\left(\frac{1}{2} + \frac{z}{h}\right)^n + \rho_2 \right) z^i dz, i = 0, 1, 2 \qquad (10)$$

where ρ_C and ρ_m are the mass densities of constituents of FGM.

$$u_0 = \sum_{i=1}^{9} \check{N}_i(\xi, \eta)u_i, \, v_0 = \sum_{i=1}^{9} \check{N}_i(\xi, \eta)v_i, \, w_0 = \sum_{i=1}^{9} \check{N}_i(\xi, \eta)w_i$$

$$\theta_x = \sum_{i=1}^{9} \check{N}_i(\xi, \eta)\theta_{xi}, \, \theta_y = \sum_{i=1}^{9} \check{N}_i(\xi, \eta)\theta_{yi}$$

This can be written in a matrix format with a shape function matrix as

$$\{u\} = \left\{u_0 v_0 w_0 \theta_x \theta_y\right\}^T = [N]\{d^e\} \qquad (11)$$

where [N] is a matrix (5 × 45) that contains the shape or interpolation functions Ň for each DOFs

$$\begin{bmatrix} \hat{u}\hat{v}\hat{w}\hat{\theta}_x\hat{\theta}_y \end{bmatrix} = \begin{bmatrix} u_1 \ldots u_n v_1 \ldots v_n w_1 \ldots w_n \theta_{x1} \ldots \theta_{xn}\theta_{y1} \ldots \theta_{yn} \end{bmatrix}^T$$

The mass matrix can be carried out from a finite element approximation of kinetic energy as

$$KE^e = \frac{1}{2}d^{eT} \int\limits_{\Omega^e} [N]^T[M][N]d\Omega^e \dot{d}^e \qquad (12)$$

and the element mass matrix (in natural coordinates) becomes

$$\begin{bmatrix} M^e \end{bmatrix} = \int\limits_{\Omega^e} [N]^T[M][N]d\Omega^e = \int\limits_{-1}^{1}\int\limits_{-1}^{1} [N]^T[M][N]\det J d\xi d\eta \qquad (13)$$

Assembling global stiffness and global mass matrix, the equation for free vibration of plate becomes,

$$\begin{bmatrix} K_g \end{bmatrix}\{d\} = \omega^2 \begin{bmatrix} M_g \end{bmatrix}\{d\} \qquad (14)$$

After applying proper boundary conditions, using Matlab's in-built eigen function, eigenfrequencies are obtained.

3 Results and Discussion

For the present work, a Matlab program has been developed using a "nine-noded isoparametric element". To demonstrate the performance of this code, first, the FGM plate has been validated with previously published results for a selection of thickness ratios and power indexes. After that, new examples of cutout FGM plates have been shown, including different cutout sizes, arrangement locations, thickness ratios, aspect ratios, etc. Present FGM is consist of Al and Al_2O_3 whose properties are shown in Table 1.

3.1 Simply Supported Square Functionally Graded Plate with Different Power Index and Thickness Ratio

Kumar et al. [9] used Galerkin Vlasov's method based on higher-order shear deformation theory for FGM rectangular plates and tapered plates. Thai et al. [8] used a unique sinusoidal shear deformation theory for the analysis of FGM plates. It is observed in Table 2 that as the power index "n" increases, the natural frequency decreases for the FGM plate. FGM plate has a natural frequency between metal and ceramic. In a nutshell, we can say the present finite element program based on FSDT is simple and accurately predicts the natural frequency for functionally graded plates with cutouts.

Table 1 FG material properties

ρ_{Al}	2707 kg/m^3	E_{Al}	70 GPa	νAl_2O_3	0.3
ρAl_2O_3	3800 kg/m^3	$E Al_2O_3$	380 GPa	νAl_2O_3	0.3

Table 2 Variation of non-dimensional frequencies $\left(\overline{\omega} = \omega h\sqrt{\frac{\rho_c}{E_c}}\right)$ with power index for simply supported FGM plate ($a/b = 1$)

h/a	Theory	Power index (n)			
		0	1	4	10
0.05	Kumar [9]	0.0148	0.0113	0.0098	0.0094
	Thai [8]	0.0148	0.0113	0.0098	0.0094
	Present	0.0148	0.0113	0.0098	0.0094
0.1	Kumar [9]	0.0577	0.0442	0.0381	0.0364
	Thai [8]	0.0577	0.0442	0.0381	0.0364
	Present	0.0577	0.0442	0.0382	0.0366

Table 3 Variation of non-dimensional frequencies $\left(\overline{\omega} = \omega h \sqrt{\frac{\rho_c}{E_c}}\right)$ with cutout size of clamped FGM plate ($a/b = 1$, $h = 0.1b$)

Cutout size	Power index (n)				
	1	2	3	4	5
0.0a × 0.0a	0.0760	0.0690	0.0665	0.0653	0.0645
0.1a × 0.1a	0.0756	0.0686	0.0661	0.0649	0.0641
0.2a × 0.2a	0.0775	0.0704	0.0678	0.0666	0.0658
0.3a × 0.3a	0.0855	0.0775	0.0747	0.0733	0.0725
0.4a × 0.4a	0.1019	0.0924	0.0890	0.0873	0.0862
0.5a × 0.5a	0.1316	0.1192	0.1147	0.1124	0.1110
0.6a × 0.6a	0.1854	0.1677	0.1610	0.1576	0.1553
0.7a × 0.7a	0.2907	0.2626	0.2512	0.2452	0.2412
0.8a × 0.8a	0.5324	0.4799	0.4567	0.4435	0.4345

3.2 FGM Plate with Gradually Increasing Central Cutout

Square FGM plate of thickness $h = 0.1b$ clamped on all sides. The central cutout is gradually increased from 0.0a × 0.0a up to 0.8a × 0.8a and also, the power-law index changed from 1 to 5. It is presented in Table 3. The non-dimensional eigenfrequency reduces with cutout size up to a particular cutout dimension, but after that frequency increases with an increase in cutout size. As cutout size grows regardless of power index, frequency reduces up to a specific cutout dimension because the stiffness falls faster than the mass. But as the cutoff dimension exceeds 0.1a × 0.1a, mass reduces and stiffness rises, which causes the eigenfrequency to rise.

3.3 Cutout FGM Plate with Inner Edge Support

In this example, free vibration analysis was done on the FGM plate ($h = 0.01b$) with the central cutout of size 0.5a × 0.5a with a varying aspect ratio (a/b). The internal, i.e., cutout edges are always clamped but the plate is placed in two different boundary conditions (SSSS–all edges are simply supported; CCCC–all edges are fixed). From Fig. 1, it is seen that the lowest eigenfrequency obtained at $n = 5$ and the rate of change of eigenfrequency is higher in the lower aspect ratio in both the boundary conditions.
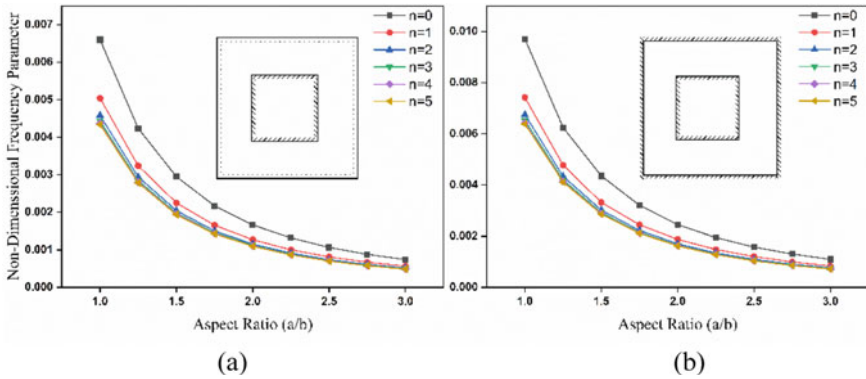
Fig. 1 Variation of non-dimensional frequencies with aspect ratio when central cutout edges of FGM plate are fixed for **a** SSSS-CCCC condition and **b** CCCC-CCCC condition ($a/b = 1$, $h = 0.01b$)

3.4 Square FGM Plate with a Variety of Cutout Orientations

Three FGM plates with different cutout arrangements have been observed in this example for 3 different support conditions (SSSS, CCCC and CSCS-opposite pairs are clamped and simply supported) where in every case the total cutout area remains the same, i.e., $0.06a^2$. For all cases as the power index increases, non-dimensional eigenfrequency decreases (Table 4) and the highest frequency is observed in Case 1 though the difference is only 1.35% compared to Case 1 and Case 3.
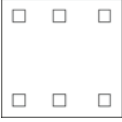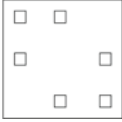
3.5 FGM Plate with a Cutout Along the Right Edge

Here, an FGM plate with an eccentric cutout has been analyzed for free vibration analysis for different aspect ratios. The aspect ratio (a/b) varies from 1.0 to 3.0 keeping the thickness ratio unchanged ($h/b = 0.01$). The vibration analysis was done for 3 different boundary conditions and two different coutout sizes ($0.6a \times 0.6b$, $0.9a \times 0.6b$) as shown in Table 5. As the aspect ratio gradually increases, the non-dimensional frequency decreases irrespective of power index & cutout size. But the rate of change of non-dimensional frequency due to increasing aspect ratio significantly decreases as the power index increases.

3.6 Square FGM Plate with Cracked Corners

A square functionally graded plate ($a \times a$) of thickness $0.01a$ has a central cutout of size ($a/2 \times a/2$). The outer edges of this plate are simply supported and the internal

Table 4 Non-dimensional frequency parameter $\left(\overline{\omega} = \omega h \sqrt{\frac{\rho_c}{E_c}}\right)$ for square FGM (Al/Al$_2$O$_3$) plate with different cutout orientation ($a/b = 1$, $h/b = 0.1$)

$h/b = 0.1$	Support	Power Index (n)				
		1	2	3	4	5
Case 1	CCCC	0.074	0.067	0.064	0.063	0.063
	SSSS	0.041	0.038	0.036	0.036	0.035
	CSCS	0.060	0.055	0.053	0.052	0.051
Case 2	CCCC	0.073	0.066	0.064	0.063	0.062
	SSSS	0.042	0.038	0.037	0.036	0.036
	CSCS	0.059	0.054	0.052	0.051	0.051
Case 3	CCCC	0.073	0.066	0.064	0.063	0.062
	SSSS	0.042	0.038	0.037	0.036	0.036
	CSCS	0.060	0.054	0.052	0.051	0.051

edges are fixed. As the concentration of stress is very high at the corner of the cutouts to relieve the high stress, cracked cutout corners are introduced at each corner. Zhang et al. [10] modeled these cracked corners as small square cutouts with free edges. First results obtained from the present finite element formulation compared with Zhang's Hencky bar-net model for isotropic cracked corner plates for three different cracked corner lengths. It is known that, at $n = 0$, the FGM plate behaves like an isotropic ceramic plate. So taking power index $n = 0$, Table 6 shows excellent accuracy of the present formulation. Next, the FGM plate with cracked corners has been analyzed for different power indexes for the previously mentioned three cases ($r = 0$, a/10 and a/5), which is presented in Table 7. It is observed that, as cracked length r increases, the frequency decreases similar to the isotropic case. Non-dimensional eigenfrequency decreases because of the reduction of strain energy. Also as power index n increases, the frequency decreases as stiffness decreases. The first six non-dimensional frequency and corresponding mode shapes for $r = $ a/5 and power index $n = 5$ have been showcased in Table 8.

Table 5 Non-dimensional frequency $\left(\overline{\omega} = \omega h \sqrt{\frac{\rho_c}{E_c}}\right)$ for different aspect ratios of FGM plate at three different boundary conditions respectively (h = 0.1b)

Boundary condition										
		CCCC			SSSS			CSCS		
Cutout Size ($c \times d$)	a/b	Power index (n)								
		1	3	5	1	3	5	1	3	5
0.6a × 0.6b	1.0	0.086	0.075	0.073	0.046	0.040	0.039	0.067	0.059	0.057
	1.5	0.066	0.058	0.056	0.034	0.030	0.029	0.042	0.037	0.036
	2.0	0.059	0.051	0.050	0.029	0.025	0.025	0.032	0.028	0.028
	2.5	0.055	0.049	0.047	0.026	0.023	0.022	0.032	0.028	0.028
	3.0	0.054	0.047	0.046	0.024	0.021	0.021	0.028	0.024	0.024
0.9a × 0.6b	1.0	0.068	0.059	0.056	0.068	0.059	0.056	0.083	0.072	0.070
	1.5	0.042	0.035	0.032	0.042	0.035	0.032	0.052	0.045	0.044
	2.0	0.028	0.024	0.022	0.028	0.024	0.022	0.038	0.033	0.032
	2.5	0.021	0.017	0.016	0.021	0.017	0.016	0.030	0.025	0.024
	3.0	0.016	0.013	0.012	0.016	0.013	0.012	0.024	0.020	0.019

Table 6 The non-dimensional fundamental frequency parameter $\left(\overline{\omega} = \omega a^2 \sqrt{\frac{\rho h}{D}}\right)$ for the square isotropic plate with the cracked corner in SSSS-CCCC condition
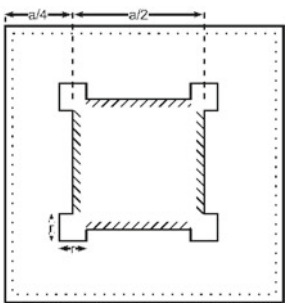
Source	$r = 0$	$r = a/10$	$r = a/5$
Present	218.109	172.6661	136.0642
Zhang et al. [10]	218.35	174.75	138.01

Table 7 Non-dimensional frequency parameter $\left(\overline{\omega} = \omega h \sqrt{\frac{\rho_c}{E_c}}\right)$ for square FGM (Al/Al2O3) plate with the cracked corner in SSSS-CCCC condition

r	Power index (n)					
	0	1	2	3	4	5
0	0.00660	0.00504	0.00458	0.00444	0.00438	0.00434
a/10	0.00523	0.00399	0.00363	0.00351	0.00346	0.00344
a/5	0.00412	0.00315	0.00286	0.00277	0.00273	0.00271

Table 8 First six non-dimensional frequency and corresponding mode shapes for $r = $ a/5 and power index $n = 5$

0.0027

0.0027

0.0027

0.0027

0.0044

0.0045
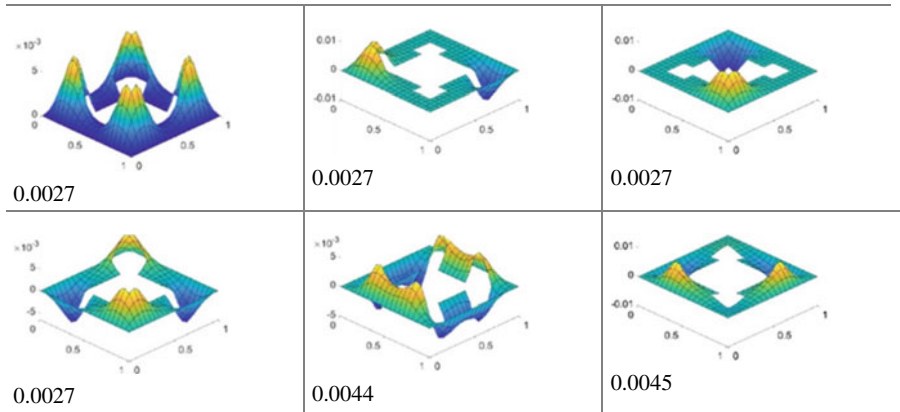
4   Conclusion

The present analysis investigates the dynamic features of FGM plates with finite element method based on first-order shear deformation theory. This inquiry includes FGM plates with rectangular and square cutouts under varied support situations, cutout arrangements, aspect ratio, thickness ratio, power index. This study also includes cracked corner and internal edge support. It is observed that, as cracked length increases frequency decreases similar to isotropic case. Non-dimensional eigenfrequency decrease with increase of power index. When cutout size exceeds a certain dimension, eigenfrequency frequency increases with cutout size. As the aspect ratio gradually increases non-dimensional frequency decreases irrespective of power index. This study adds fresh insights to the literature on the functionally graded material plate.

References

1. Tran LV, Ferreira AJ, Nguyen-Xuan H (2013) Isogeometric analysis of functionally graded plates using higher-order shear deformation theory. Compos B Eng 51:368–383
2. Kumar R, Jana P (2022) Free vibration analysis of uniform thickness and stepped P-FGM plates: a FSDT-based dynamic stiffness approach. Mech Based Des Struct Mach 1–30
3. Ramteke PM, Panda SK (2021) Free vibrational behaviour of multi-directional porous functionally graded structures. AJSE 46:7741–7756
4. Tabatabaei SS, Fattahi AM (2022) A finite element method for modal analysis of FGM plates. Mech Based Des Struct Mach 50:1111–1122
5. Reddy JN (2003) Mechanics of laminated composite plates and shells: theory and analysis. CRC press
6. Majumdar A, Manna MC, Haldar S (2010) Bending of skewed cylindrical shell panels. Int J Comput Appl 1:89–93
7. Ferreira AJ (2009) MATLAB codes for finite element analysis. Springer
8. Thai H-T, Vo TP (2013) A new sinusoidal shear deformation theory for bending, buckling, and vibration of functionally graded plates. Appl Math Model 37:3269–3281
9. Kumar V, Singh SJ, Saran VH, Harsha SP (2020) Exact solution for free vibration analysis of linearly varying thickness FGM plate using Galerkin-Vlasov's method. Proc Inst Mech Eng, Part L: J Mater: Des Appl 235:880–897
10. Zhang YP, Wang CM, Pedroso DM, Zhang H (2018) Extension of Hencky bar-net model for vibration analysis of rectangular plates with rectangular cutouts. J Sound Vib 432:65–87

Floating Inductance Simulator with EXCCTAs

Y. Shantikumar Singh, Ashish Ranjan, Shuma Adhikari,
and Benjamin A. Shimray

1 Introduction

With the development of recent IC technology, an active inductance simulator (AIS) has become an essential element in designing a dynamic network. Mainly, these have been used to create various analog active circuits, such as filters, oscillators, and impedance cancelation electrical systems. Therefore, several actively simulated FIs with different functional elements exist in the open literature [1–16]. The simulators presented in [4] utilized two to three CCCII to get electronically tunable AIS. The available configuration has been simulated floating AIS using three active and passive elements in [5]. Moreover, with two CFOA and five passive components used in [6], it significantly has an excessive number of passive elements. Wherein [7, 8], the floating AIS implementations employing two active and three passive elements are presented but cannot have tunability properties. Although the realizations of [14–18], the floating AIS utilized two or more active and few passive components, they have matching conditions. Moreover, the AIS implementation with one or two active elements, one or two resistors, and one capacitor has been presented in [19, 20]. However, they have obtained only grounded mode AIS. Recently, the EXCCTA has been widely developed as a valuable active element in the analog signal processing application circuit design [19]. The EXCCTA has been proposed with the extension of CCTA through an extra input terminal; namely, an X terminal is added at the input

Y. S. Singh (✉) · A. Ranjan
Department of Electronics and Communication Engineering, National Institute of Technology
Manipur, Lamgol 795004, India
e-mail: yshantikumar99@gmail.com

S. Adhikari · B. A. Shimray
Department of Electrical Engineering, National Institute of Technology Manipur, Lamgol 795004,
India

91

stage. Therefore, by taking advantage of the additional X terminal in EXCCTA many authors have designed various ASPCs with the electronically tunable property.

This paper aims to design floating AIS without matching conditions. The presented floating AIS has utilized two EXCCTAs and three passive elements. The presented inductance is a tunability property through the bias current (I_B) of the EXCCTAs. The floating AIS is tested in the applications of fourth-order LPF and tunable band rejection filter circuits. The workability of the introduced FI is verified with the help of SPICE simulation and experimental responses.

2 Proposed Floating Simulator

The EXCCTA is an electronically tunable active element with nine terminals. It combines EXCCII active element and OTA in a single IC form. The terminal characteristics of EXCCTA are obtained in matrix form:

$$
\begin{bmatrix}
I_Y \\
V_{XP} \\
V_{XN} \\
I_{ZP+} \\
I_{ZP-} \\
I_{ZN+} \\
I_{ZN-} \\
I_O
\end{bmatrix}
=
\begin{bmatrix}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & g_m & 0 & 0 & 0 & 0
\end{bmatrix}
\begin{bmatrix}
V_Y \\
I_{XP} \\
I_{XN} \\
V_{ZP+} \\
V_{ZP-} \\
V_{ZN+} \\
V_{ZN-} \\
V_O
\end{bmatrix}
\tag{1}
$$

The electrical behavior of the EXCCTA active element is given in Fig. 1, and its CMOS structure of EXCCTA is constructed with CMOS of [19]. The transconductance parameter (K) is constructed with transistors (M29-M30). The OTA current output depends on the Z_{p+} terminal voltage. The I_o of the OTA is given by Eq. (2). The realization of the macro-model of EXCCTA using commercially available ICs is constructed the same as [19].

$$
I_o = g_m(V_{zp+}) = \sqrt{2 I_{Bias} K_i}(V_{zp+}),
\tag{2}
$$

where $K_i = \mu C_{0x}$ W/2L, ($i = 29, 30$) W and L denote the channel width and length, C_{ox} is the gate oxide capacitance per unit area, and μ is the carrier mobility. From (2), it is observed that EXCCTA has tenability property through bias current.

The presented electronically tunable floating inductance simulator with two EXCCTA and three components is shown in Fig. 2. By simple circuit analysis (Fig. 2) with $I_{B1} = I_{B2}$, the obtained admittance of the floating AIS is given as:

Fig. 1 Circuit symbol of EXCCTA
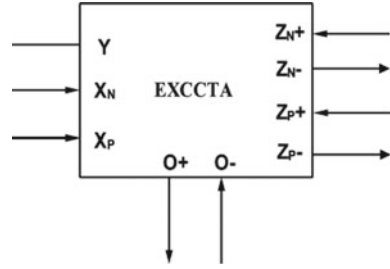
$$Y_{in}(s) = \frac{2g_m}{sCR} \begin{bmatrix} +1 & -1 \\ -1 & +1 \end{bmatrix}. \tag{3}$$

The floating AIS provides the equivalent inductance value as:

$$L_{eq} = \frac{CR}{2g_m}. \tag{4}$$

The active and passive sensitivities of presented floating AIS are obtained as:

$$S_{g_m}^L \big|_{FI} = -1, \ S_C^L \big|_{FI} = S_R^L \big|_{FI} = 1 \tag{5}$$
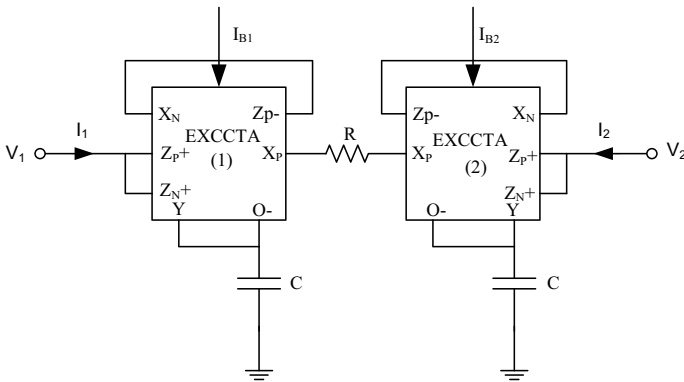
Fig. 2 Proposed FIS with two EXCCTA

3 Influences of Non-Ideal as Well as Parasitic

With the influence of non-idealities such as α, β, and γ in the active elements, the terminal responses have deviated from the ideal one [19]. Therefore, in this EXCCTA active element also present this non-idealities parameter and its terminal relationship along with these non-idealities are given as:

$$V_X = \beta V_Y, \ I_{XN} = \alpha I_{ZN}, \ I_{XP} = \alpha I_{ZP}, \ I_O = \gamma g_m V_{ZP}, \ I_Y = 0. \tag{6}$$

With these influences, the new obtain input admittance function is given as

$$Y_{in}(s) = \frac{2\alpha\beta\gamma g_m}{sCR} \begin{bmatrix} +1 & -1 \\ -1 & +1 \end{bmatrix}. \tag{7}$$

Above these influences, the parasitic element influences present in the circuit, such as parasitic resistor (R_x) and parasitic capacitor (C_x). These parasitic elements are associated with each terminal in parallel form. Detailed structures of the parasitic component present in EXCCTA are well presented in [19]. The proposed floating AIS through parasitic elements are shown in Fig. 3. The parasitic element equivalent for resistance and capacitance at $Z_{N+}(1)$ and $Z_{P+}(1)$, $Y(1)$ and O-(1), $X_{P+}(1)$, and $X_{P+}(2)$, $Y(2)$ and O-(2), and $Z_{N+}(2)$ and $Z_{P+}(2)$, connections are found as $\{R_{eq1} = R_{eq5} = R_{ZP+}||R_{ZN+}$, and $C_{eq1} = C_{eq5} = C_{ZP+} + C_{ZN+}\}$, $\{R_{eq2} = R_{eq4} = R_Y||R_{O-}$, and $C_{eq2} = C_{eq4} = C + C_Y + C_{O-}\}$, and $R_{eq3} = R + 2R_{XP+}$, respectively, where $R_{XP+(1)} = R_{XP+(2)} = R_X + sL_X$. By considering the parasitic effects, the new input current equation is obtained as

$$I_1 = \frac{V_1}{Z_1} + \frac{V_1 - V_2}{Z_2 Z_3}, \quad \text{and} \quad I_2 = \frac{V_2}{Z_5} + \frac{V_2 - V_1}{Z_4 Z_3}.$$
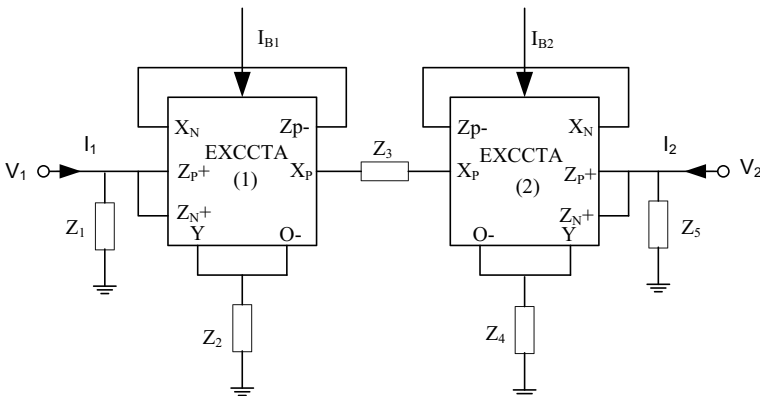
Fig. 3 Proposed FIS with parasitic elements

By using the above equations where Z_1 and Z_5 are very small as well as $Z_2 = Z_5$, then the new obtain input admittance including these effects is given as

$$Y_{in}(s) = \frac{2g_m}{Z_2 Z_3} \begin{bmatrix} +1 & -1 \\ -1 & +1 \end{bmatrix}, \tag{8}$$

where $Z_2 = R_{eq2} \| C_{eq2}$, $Z_3 = R_{eq3}$

4 Performance Verification Using PSPICE and Experimental Results

The verification of the introduced floating AIS is carried out using PSPICE simulation as well as an experimental platform. For the PSPICE simulation platform, the CMOS-based EXCCTA is simulated using 0.18 μm TSMC parameter and transistor dimensions [19]. The V_{DC} and I_B are chosen as ±0.9 V and $I_{B1} = I_{B2} = 24.5$ μA, $I_A = 100$ μA, respectively. The corresponding transconductance value is obtained as 3.674 mS. The introduced floating AIS is tested using bias current and passive component values as $I_B = 40$ μA, ($g_m = 4.69$ mS), $I_A = 100$ μA, and $C = 50$ pF, $R = 4$ kΩ. The simulation and ideal frequency domain response (FDR) of introduced floating AIS in terms of phase and magnitude are illustrated in Fig. 4a–b. Figure 5a–b represents the tunable magnitude and phase responses with different bias currents such as $I_B = \{40$ μA, 50 μA, 60 μA\}. From Fig. 4b, it is noted that the range of frequency of introduced floating AIS is 700 kHz to 57.46 MHz. The power utilization of the introduced floating AIS is found as 2.626 mW. Moreover, the transient response of introduced floating AIS is carried out and has the phase difference of input voltage and current of 88.7° which is illustrated in Fig. 6. Figure 7 displays the Lissajous figure between voltage and current that validate the working principle of the inductor.
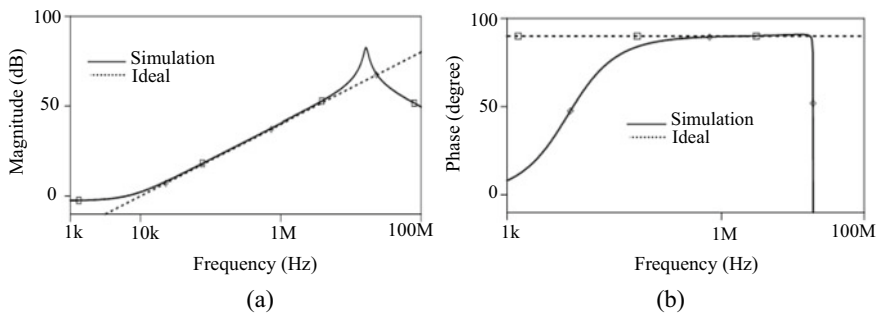
Fig. 4 FDR of introduced floating AIS: **a** Magnitude and **b** Phase
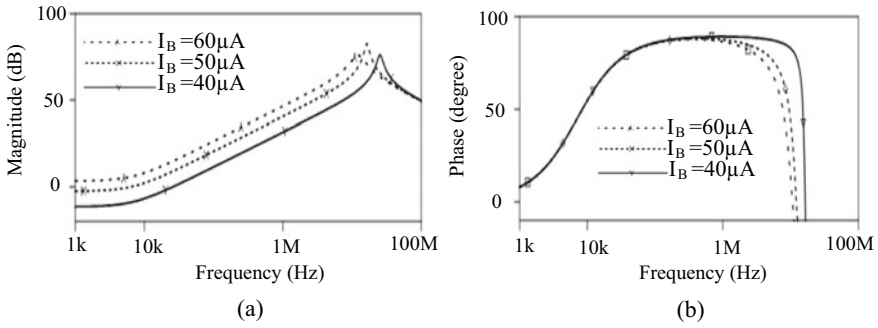
Fig. 5 Tunable FDR of introduced floating AIS: **a** Magnitude and **b** Phase
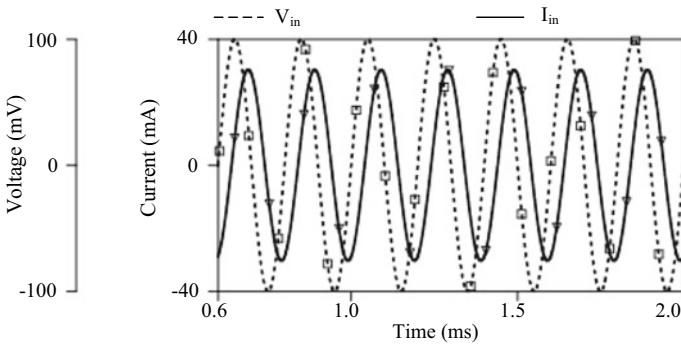
Fig. 6 Transient response of proposed floating AIS
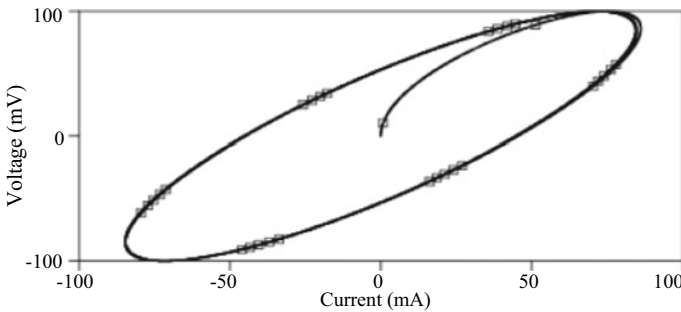
Fig. 7 Lissajous figure between voltage and current

To check the functionality of introduced floating, AIS is utilized in various applications such as fourth-order LPF and second-order BRF, shown in Fig. 8a–b, respectively. A simple routine analysis is performed in Fig. 8a and obtains its $V_{out}(s)/V_{in}(s)$ function as:
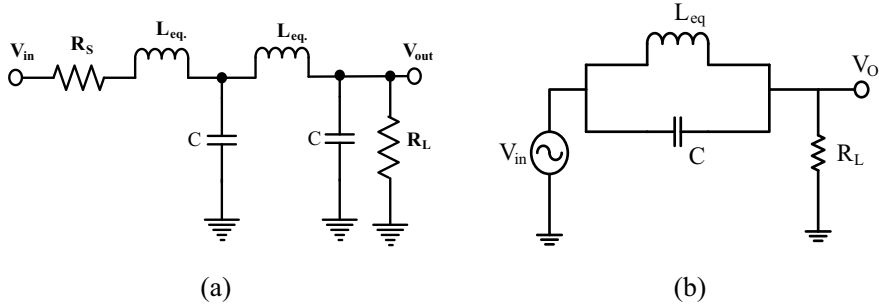
Fig. 8 Application circuits of introduced floating AIS: **a** fourth-order LPF and **b** second-order BRF

$$T(s)|_{\text{4th - order}} = \frac{R_{\text{Load}}}{D_4 s^4 + D_3 s^3 + D_2 s^2 + D_1 s + R_{\text{Load}}}, \tag{9}$$

where D_i denotes the denominator coefficients and obtained as $D_4 = L_{eq}^2 C^2 R_{\text{Load}}$, $D_3 = L_{eq}^2 C + L_{eq} C^2 R_S R_{\text{Load}} + R_S L_{eq} C$, $D_2 = L_{eq} C R_{\text{Load}} + L_{eq} C R_S R_{\text{Load}}$, and $D_1 = L_{eq} + C R_S R_{\text{Load}}$.

Similarly, the same procedure of circuit routine analysis is applied in the second-order band rejection filter circuit of Fig. 8b, and its transfer function is obtained as:

$$T(s)|_{\text{2nd - order}} = \frac{s^2 R_L L_{eq} C + R_L}{s^2 R_L L_{eq} C + s L_{eq} + R_L}. \tag{10}$$

Figure 9a shows the frequency response of fourth-order LPF by choosing the passive components as $R_S = 100\ \Omega$, $C_1 = C_2 = 1\text{nF}$, $R_{\text{Load}} = 1\ \text{k}\Omega$, and $L_{eq} = 0.363\ \text{mH}$ ($I_{B1} = I_{B2} = 50\ \mu\text{A}$, $I_A = 100\ \mu\text{A}$, $C = 50\ \text{pF}$, $R = 4\ \text{k}\Omega$), respectively. Figure 9b illustrates the FDR of the second-order BRF with passive elements as $C = 1\text{nF}$, $R_L = 1\ \text{k}\Omega$, and $L_{eq} = \{0.454\ \text{mH}, 0.363\ \text{mH}, 0.302\ \text{mH}\}$, respectively. From Fig. 9b, it is noted that the introduced floating AIS is electronically tunable through bias current I_B.

In addition to the SPICE response, an experimental setup is obtained with locally available ICs such as AD844 and LM13700 through the macro-model of [19]. The commercial ICs are applied with a ± 10 V DC power supply, and the I_B of OTA (LM13700) is set at 367 μA. The external components are chosen as $R = 4\ \text{k}\Omega$ and $C = 10\text{nF}$. To test the configuration, a sinusoidal input signal is applied with $2V_{\text{p–p}}$ at 50 MHz in the proposed structure. The obtained transient response of the voltage and current relationship is depicted in Fig. 10. From Fig. 10, it is noted that the phase difference is found to be 86.5°. Finally, a brief comparison of the proposed CMOS-based AIS with the current active inductor is observed in Table 1. It includes the number of functional elements, the number of external components count, the nature of the simulator, free from matching, the technology used, supply voltages,
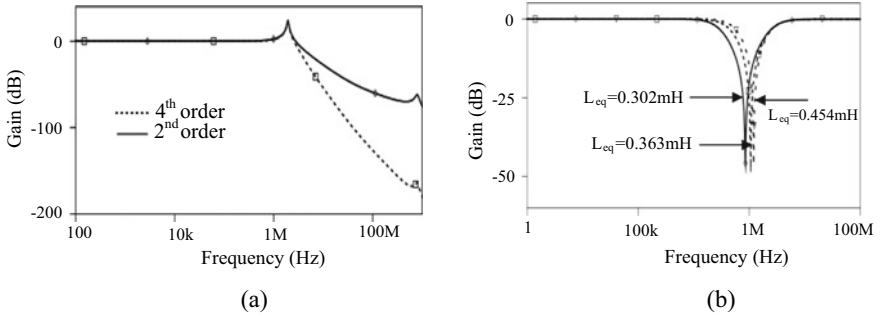
Fig. 9 FDR of application circuit: **a** fourth-order LPF and **b** second-order BRF
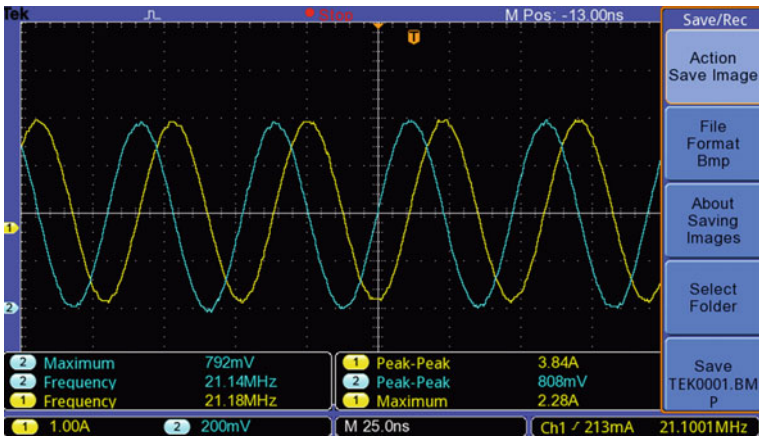
Fig. 10 Experimental response of voltage and current characteristics

and tunability. Based on this attention, the authors acknowledge that there is no floating AIS in the open literature using two EXCCTA and a few passive elements.

Table 1 A brief comparison between proposed floating AIS with existing FS

Ref.	Name of elements: No. of elements	No. of compt. used	Tunability	Tech. used	Power supply	Sim./ Expt	L_{eq} Mode	Free from matching
[4]	CCCII:3	1	Yes	ALA400	±2.5 V	Both	F	No
[5]	CFOA:3	3	No	AD844	±15 V	Sim	F	Yes
[6]	CFOA:2	3/5	No	AD844	±10 V	Sim	F	No
[7]	MCFOA:2	3	No	0.25 µm	±1.5 V	Both	F	Yes
[8]	DVCC:2	3	No	0.35 µm	±1.5 V	Sim	F	Yes
[9]	FTFNTA:2	3	Yes	0.18 µm	±1.65 V	Sim	F	Yes
[10]	VD-DIBA:2	2	Yes	0.18 µm	±0.9 V	Sim	F	Yes
[11]	CCII:4	4	No	AD844	±15 V	Sm	F	Yes
[12]	CFOA:2	3	No	AD844	±15 V	Sim	F	Yes
[13]	CDBA:3/4	5	Yes	0.5 µm	±2.5 V	Sim	F	Yes
[14]	VDBA:2	2	Yes	0.25 µm	±0.75 V	Both	F	Yes
[15]	FTFN:2	5	No	3 µm	±5 V	Sim	F	No
[16]	CC-CFA:3	1	Yes	BiCMOS	±1.5 V	Sim	F	Yes
[17]	VCII+:3	1	No	0.18 µm	±0.9 V	Sim	F	Yes
Prop	EXCCTA:2	3	Yes	0.18 µm	±1.65 V	Both	F	Yes

5   Conclusion

A floating inductance simulator with two EXCCTA active and three passive elements is realized. The proposed floating AIS functionality verification is well carried out through CMOS-based design. To the authors' knowledge, such floating AIS using EXCCTA active elements are not available in the open literature. Therefore, it filled the gap by realizing new EXCCTA-based floating AIS and added it to the current catalog of its analog application circuits. SPICE simulation results have agreed with the theoretical expectation. Lastly, experimental results are well verified the functionality of the proposed design.

References

1. Singh YS, Ranjan A, Adhikari S, Shimray BA (2023) A tunable resistorless floating inductance simulator using MO-DXCCTA. In: Mishra B, Tiwari M (eds) VLSI, microwave and wireless technologies. Lecture notes in electrical engineering, vol 877. Springer, Singapore, pp 73–82. https://doi.org/10.1007/978-981-19-0312-0_9
2. Singh YS, Ranjan A, Adhikari S, Shimray BA (2022) A modern VDCCTA active element and its electronic application. J Circ, Syst, Comput 2350053. https://doi.org/10.1142/S0218126623500536

3. Singh YS, Ranjan A, Adhikari S, Shimray BA (2022) A lossless active inductor design using single ZC-VDCC: grounded and floating mode. IETE J Res 1–15. https://doi.org/10.1080/03772063.2022.2130828

4. Minaei S, Yuce E (2008) Realization of tunable active floating inductance simulators. Int J Electron 95(1):27–37. https://doi.org/10.1080/00207210701809333

5. Abuelma'atti MT, Dhar SK (2016) New CFOA-based floating immittance emulators. Int J Electron 103(12):1–14. https://doi.org/10.1080/00207217.2016.1138544

6. Abuelma'atti MT, Dhar SK, Khalifa ZJ (2017) New two-CFOA-based floating immittance simulators. Analog Integr Circ Sig Process 91:479–489. https://doi.org/10.1007/s10470-017-0956-9

7. Yuce E, Minaei S (2008) A modified CFOA and its applications to simulated inductors, capacitance multipliers, and analog filters. IEEE Trans Circ Syst-I 55(01):266–275. https://doi.org/10.1109/TCSI.2007.913689

8. Yuce E (2010) A novel floating simulation topology composed of only grounded passive components. Int J Electron 97(3):249–262. https://doi.org/10.1080/00207210903061907

9. Singh R, Prasad D (2020) Comment floating simulated inductance circuits using FTFNTAs. Int J Electron 107(9). https://doi.org/10.1080/00207217.2020.1726495

10. Jaikla W, Bunrueangsak S, Khateb F, Kulej T, Suwanjan P, Supavarasuwat P (2021) Inductance simulators and their application to the 4th order elliptic lowpass ladder filter using CMOS VD-DIBAs. Electronics 10(6):684. https://doi.org/10.3390/electronics10060684

11. Yuce E, Cicekoglu O (2006) Novel floating inductance and FDNR simulators employing CCII+s. J Circ, Syst, Comput 15(1):75–81. https://doi.org/10.1142/S0218126606002964

12. Bhaskar DR, Senani R (2013) Simulation of a floating inductance: a new two CFOA-based configuration. In: Fifth international conference on computational intelligence, modelling and simulation, pp 381–383. https://doi.org/10.1109/CIMSim.2013.67

13. Keskin AU, Hancioglu E (2005) CDBA-based synthetic floating inductance circuits with electronic tuning properties. ETRI J 27(2):239–242

14. Tangsrirat W (2019) Actively floating lossy inductance simulators using voltage differencing buffered amplifiers. IETE J Res 65(4):446–459. https://doi.org/10.1080/03772063.2018.1433082

15. Kilic R, Ugur Cam MA, Kuntman H (2002) Improved realization of mixed-mode chaotic circuit. Int J Bifurcation and Chaos 12(06):1429–1435. https://doi.org/10.1142/S0218127402005236

16. Singh A, Jain MK, Wairya S (2019) Novel lossless grounded and floating inductance simulators employing a grounded capacitor based in CC-CFA. J Circuits, Syst, Comput 28(6):1950093. https://doi.org/10.1142/S0218126619500932

17. Safari L, Barile G, Colaiuda D, Stornelli V, Ferri G (2022) Realization of an electronically tunable resistor-less floating inductance simulator using VCII. Electronics 11(3):312

18. Tarunkumar H, Singh YS, Ranjan A (2019) An active inductor employing a new four terminal floating nullor transconductance amplifier (FTFNTA). Int J Electron 107(5):683–702. https://doi.org/10.1080/00207217.2019.1672807

19. Faseehuddin M, Sampe J, Shireen S, Md Ali SH (2018) Lossy and lossless inductance simulators and universal filters employing a new versatile active block. J Microelectron, Electron Compon Mater 48(2):97–113

20. Cicekoglu MO (1998) Active simulation of grounded inductors with CCII+s and grounded passive elements. Int J Electron 85(4):455–462. https://doi.org/10.1080/002072198134003

Comparative Analysis of Microstrip and Co-Planar Waveguide-Fed Printed Monopole Antenna for Ultra-Wideband Application

Samom Jayananda Singh, Rajesh Kumar, Kharibam Jilenkumari, and M. M. Dixit

1 Introduction

Ultra-wideband antennas (UWBs), which offer maximum bandwidth, rise in gain and narrow radiated power, are constantly rising as a result of the tremendous expansion in communication systems, from old-fashioned landlines to modern wireless gadgets. To meet demands for clear resolution and data rates, modern communication technology is always being improved. For maximum data rate wireless communication systems (WCS), antenna scientists must develop tiny antennas on printed circuit boards while maintaining essential broadband features. The advancement of various modern communication systems has increased the progress of multifunctional antennae. Earlier, wireless systems had a single antenna with defined radiation characteristics. The selection of resonant frequency is a theory that has given rise to new technologies for applications in tiny multiband systems. As a result, designing tunable and frequency tunable antennas have gained popularity [1]. In this communication, a small SWB polarization antenna with double-band capabilities has been examined. The suggested antenna has duplex band-rejection characteristics that encompass the WLAN band and X-band satellite communication, and it offers an unusually high impedance BW from 1.2 to 25 GHz. The suggested antenna is a strong option for polarization diversity applications since it has a minimal ECC

S. J. Singh (✉) · R. Kumar · K. Jilenkumari · M. M. Dixit
Department of Electronics and Communication Engineering, NERIST, Nirjuli, Arunachal Pradesh, India
e-mail: jayanandmetal.inc@gmail.com.com

Department of Electronics and Communication Engineering, NIT, Imphal, Manipur, India

Department of Mathematics, NERIST, Nirjuli, Arunachal Pradesh, India

of 0.025 for the SWB frequency range. This antenna may be utilized for the spectrum used in cognitive radio due to its huge bandwidth [2]. A novel microstrip-fed antenna with a planar shape of size $24 \times 28 \times 1.6$ mm^3 is proposed. The V-structure patch, microstrip-fed line and partial ground plane construction make up the antenna structure. A frequency rejection feature that can reject the frequency range from 5.15 to 5.825 GHz is obtained by adding a U-structure slot to the patch [3]. Reference [4] describes the impact of a straightforward ground slot monopole antenna fed with microstrip. Its main purpose is for the UWB application's antenna. An antenna with a T-shape (gap) that is CPW-fed, one-step patched, and has filtering properties are developed for double resonating frequencies of 3.5 GHz and 5 GHz, covering the frequency ranges from 3.2 GHz to 3.5 GHz) and 4.7 GHz to 5.6 GHz, respectively [5]. Diego et al. [6] produced a wide band E structure printed antenna with an approximate impedance bandwidth of 29.8% by creating a zigzag groove in the patch. A U-shaped slot-loaded inverted disc antenna with a maximum bandwidth of 24.2 per cent was created by Kaur et al. [7]. Similar to a slot, fractal, or metamaterial, an ultra-wideband antenna can be made. There have been several published UWB antenna configurations [8] through [9]. Radiator with disc patches and a CPW-fed, concentrically filled antenna for UWB applications. With the installation of a flawed ground plane, it can improve the frequency quality of the antenna [10].

2 Proposed Antenna Configuration

The suggested antenna geometry is shown in Figs. 1 and 2. The FR-4 epoxy-coated substrate has the following measurements.

Optimized sizes for the ground plane are $L_{ge} = 19.5$ mm and $W_{ge} = 11.4$ m to enhance bandwidth for both microstrip fed and CPW fed. This comparative study of both the fed that is microstrip fed as well as CPW fed is going to analyse in terms of bandwidth and return loss.

The following equations, Eqs. (1) and (2), can be used to determine how the proposed monopole antenna with a circular disc-shaped patch should be constructed

Fig. 1 Microstrip-fed printed monopole antenna
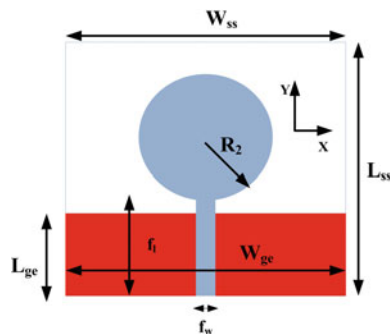
Fig. 2 CPW-fed printed
monopole antenna
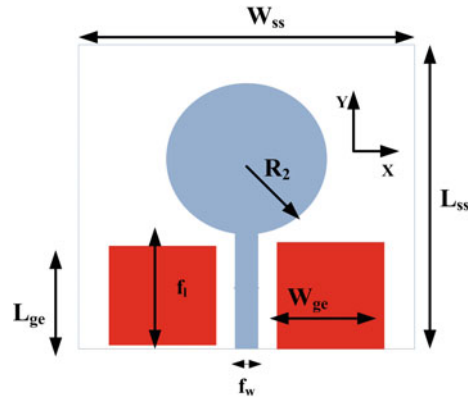
[6].

$$R_2 = \frac{R_{ef}}{\sqrt{1 + \frac{2h}{\pi \varepsilon_r R_{ef}} \left[\ln\left(\frac{1.57 R_{ef}}{h}\right) + 1.78 \right]}} \qquad (1)$$

$$R_{ef} = \frac{8.79 \times 10^9}{f_{sr} \sqrt{\varepsilon_r}}, \qquad (2)$$

where h is the substrate's height in mm, ε_r is the substrate's dielectric constant, and f_{sr} is the resonance frequency.

The low cut-off frequency of the antenna can be calculated using the usual formula provided for predicting the low cut-off frequency of printed monopole antennas. A cylindrical monopole antenna may be utilized with the appropriate modifications [6–9, 11, 12]. These equations are valid for an antenna with a monopole structure and a planar model.

$$f_{ef} = \frac{c}{\lambda_{ll}} = \frac{7.2}{(H + R_{ef} + f_{ll})} \text{GHz} \qquad (3)$$

Compared to planar antennas, which have a single sheet of dielectric on the antenna and have circularly formed monopole characteristics. Here, f_l stands for feed length to match the 50-Ω input impedance. The dielectric substrate increases the antenna's effective size, which lowers the lower band edge frequency.

3 Result

Initially, changing the dimensions of the ground plane makes a major contribution to monopole antennas.

Table 1 shows the various ground structures in terms of return loss and fractional bandwidth for CPW fed. Reducing the ground plane increases the bandwidth, and lowering the return loss after reaching a critical dimension will decrease the bandwidth and no improvement in bandwidth. Maximum fractional bandwidth of 89% is observed in the dimension of $L_{ge} = 15.9$ mm, $W_{ge} = 11.4$ mm with return loss -41 dB. From Table 2, it is noticed microstrip-fed antenna gives a lower performance in comparison to CPW fed. In microstrip fed, it gives fractional bandwidth of 87% which is lower when compared to CPW fed and return loss of -38 dB. In both the cases, optimized dimensions are $L_{ge} = 15.9$ mm and $W_{ge} = 11.4$ mm (Fig. 3).

The space in between the ground plane and patch provides for better improvement of the bandwidth. In Tables 2 and 3, it is showing that the smaller the gap, the more the bandwidth. Figure 4 shows VSWR frequencies result, and VSWR provides transmitted radiated wave and its returning wave. Its value should be the lowest as possible in order to promote fair radiation. CPW fed gives and microstrip-fed VSWR responses are good, but CPW fed gives better which value is 1.018 in contrast to microstrip fed that is 1.65.

Figure 5 shows the performance of gain for both the fed that microstrip fed and CPW fed, and the graph shows that the performance is better in the case of CPW fed reaching a maximum gain of 8.77 dBi, whereas for microstrip fed it only reaches upto 6.45 dBi.

Figures 6 and 7 show the radiation pattern of E-field and H-field for both the microstrip fed and CPW fed. A uniform omnidirectional pattern can be seen in CPW

Table 1 Optimized dimension of the proposed antenna

Parameters	Value(mm)	Parameters	Value(mm)
W_{ss}	35	W_{ge}	11.5
L_{ss}	30	f_l	18
R_2	8.5	f_w	1.6
L_{ge}	19.5	h	1.6

Table 2 Return loss for various ground dimensions for CPW-fed antenna

Ground plane dimension (mm)	Operating centre frequency (GHz)	$|S_{11}|$dB	-10 dB Bandwidth (GHz)	Fractional bandwidth (GHz)
$L_{ge} = 15.9$ $W_{ge} = 13.4$	$f_{cn1} = 7.4$ $f_{cn2} = 15.4$	-25.30 -17.97	(5.1–12.9) (14.2–19)	86.6 28.9
$L_{ge} = 15.9$ $W_{ge} = 12.4$	$f_{cn1} = 6$ $f_{cn2} = 16$	-27.69 -31.9	(5.2–12.8) (14.0–18.0)	84 25
$L_{ge} = 15.9$ $W_{ge} = 11.4$	$f_{cn1} = 6$ $f_{cn2} = 16$	-41.86 -39.91	(5.2–13.6) (14.1–18.2)	89 25
$L_{ge} = 15.9$ $W_{ge} = 10.4$	$f_{cn1} = 6$ $f_{cn2} = 17.4$	-41.58 -26.68	(5.3–13.5) (14.4–19.7)	87 29.5
$L_{gg} = 15.9$ $W_{gg} = 9.4$	$f_{cen1} = 6$ $f_{cen2} = 15.8$	-21.97 -21.97	(5.4–13.4) (14.6–17.0)	85 15
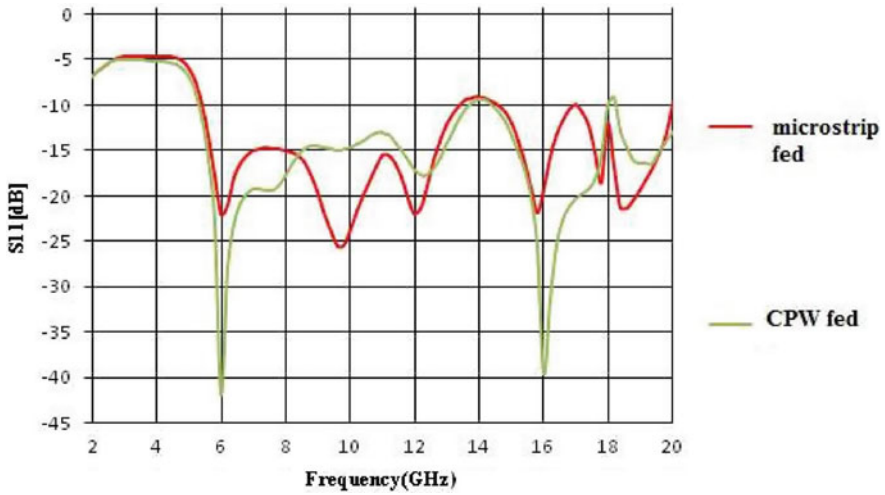
Fig. 3 Return loss versus frequency (GHz)

Table 3 Return loss for various ground dimensions for microstrip-fed antenna

Ground plane dimension	Operating centre frequency (GH)	$|S_{11}|$dB	−10 dB Bandwidth (GHz)	Fractional bandwidth (GHz)
L_{ge} = 15.9 mm W_{ge} = 13.4 mm	f_{cn1} = 7.4 f_{cn2} = 15.4	−23.30 −15.97	(4.9–12.1) (13.2–19)	79.6 28.9
L_{ge} = 15.9 mm W_{ge} = 12.4 mm	f_{cn1} = 6 f_{cn2} = 16	−25.69 −31.9	(5.2–12.8) (12.0–18.0)	70.4 25
L_{ge} = 15.9 mm W_{ge} = 11.4 mm	f_{cn1} = 6 f_{cn2} = 16	−39.5 −37.91	(5.13–13.6) (14.1–18.2)	80 25
L_{ge} = 15.9 mm W_{ge} = 10.4 mm	f_{cn1} = 6 f_{cn2} = 17.4	−39.58 −24.68	(5.3–13.5) (14.4–19.7)	78 29.5
L_{ge} = 15.9 mm W_{ge} = 9.4 mm	f_{cn1} = 6 f_{cn2} = 15.8	−19.97 −19.97	(3.59–13.4) (14.6–15.10)	77 15

fed not in the case of microstrip fed, and this indicates that CPW has a better radiation pattern than microstrip fed.
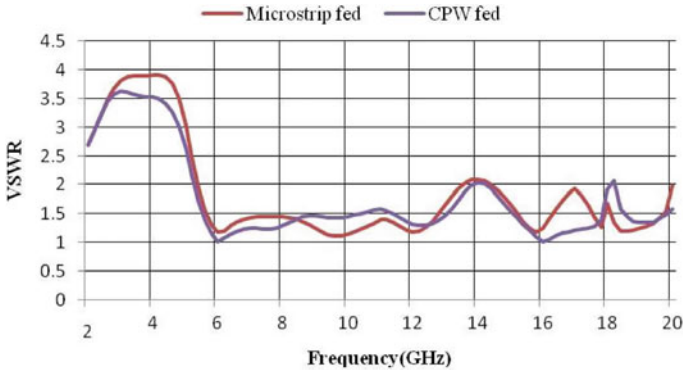
Fig. 4 VSWR versus frequency (GHz)
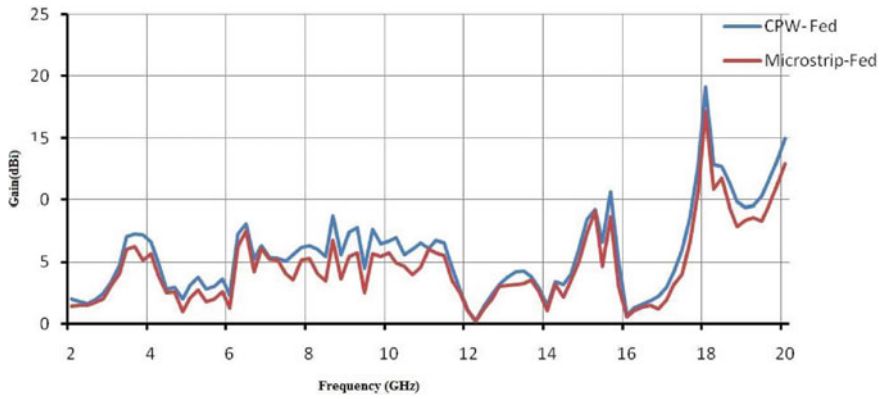
Fig. 5 Gain (dBi) versus frequency (GHz)

Fig. 6 Radiation pattern
microstrip fed
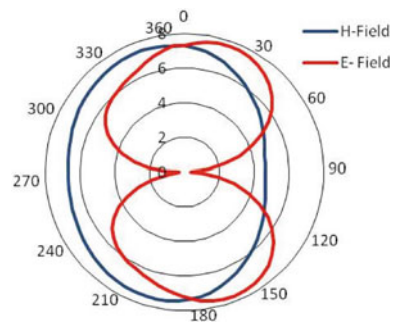
Fig. 7 Radiation pattern
CPW fed
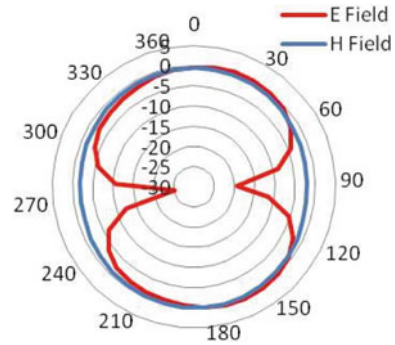
4 Conclusion

The CPW-fed antenna gives a better impedance bandwidth when compared to microstrip fed. Not only bandwidth in terms of return loss absolute value of −41 dB can be achieved in case of CPW fed. Regarding gain CPW can be reached a gain of 8.77 dBi which is a good value of gain and a fair uniform radiation pattern can be achieved. In both cases, reduction in ground plane helps in improving the antenna performance. When ground plane's size is further decreased again after reaching a critical level, the bandwidth naturally initiates to decline.

References

1. FCC report and order on ultra wideband technology, Federal communication commission, Washington, DC, USA (2002)
2. Chinnagurusamy B, Perumalsamy M, Thankamony Sarasam AS (2021) Design and fabrication of compact triangular multiband microstrip patch antenna for C- and X-band applications. Int J Commun Syst 1–8. https://doi.org/10.1002/dac.4939
3. Singh SJ, Kumar R, Dixit MM (2022) Printed monopole antenna design with CPW fed for ultra wideband application. J Phys: Conf Ser 2236(1):012010. https://doi.org/10.1088/1742-6596/2236/1/012010
4. Balanis CA (2016) Antenna theory: analysis and design, 4th edn. Wiley, New Jersey
5. Mahendran K, Gayathri R, Sudarsan H (2020) Design of multi band triangular microstrip patch antenna with triangular split ring resonator for S band, C band and X band applications. Microprocess Microsyst 80:103400. https://doi.org/10.1016/j.micpro.2020.103400
6. Kundu S (2019) Experimental study of a printed ultra-wideband modified circular monopole antenna. Microw Opt Technol Lett 61(5):1388–1393. https://doi.org/10.1002/mop.31736
7. Elfergani IT, Rodriguez J, Otung I, Mshwat W, Abd-Alhameed RA (2018) Slotted printed monopole UWB antennas with tuneable rejection bands for WLAN/WiMAX and X-band coexistence. Radio Eng 27(3):694–702. https://doi.org/10.13164/re.2018.0694
8. Murugan NA, Balasubramanian R, Patnam HR (2018) Printed planar monopole antenna design for ultra-wideband communications. Radioelectron Commun Syst 61(6):267–273
9. Bakariya PS, Dwari S, Sarkar M (2015) Triple band notch UWB printed monopole antenna with enhanced bandwidth. AEU-Int J Electron Commun 69(1):26–30. https://doi.org/10.1016/j.aeue.2014.07.023

10. Singh SJ, Kumar R, Dixit MM (2022) Study analysis of printed monopole antenna for C and X band application. In: Proceeding of CECNet. IOS Press, pp 112–118
11. Carver K, Mink J (1981) Microstrip antenna technology. IEEE Trans Antennas Propag 29(1):2–24
12. Li Z, Zhu X, Yin C (2019) CPW-fed ultra-wideband slot antenna with broadband dual circular polarization. AEU-Int J Electron C 98:191–198
13. Kanagasabai M, Sambandam P, Mohammed GNA et al (2020) On the design of frequency reconfigurable tri- band miniaturized antenna for WBAN applications. AEU-Int J Electron C 127:153450. https://doi.org/10.1016/j.aeue.2020.153450
14. Desai A, Patel R, Upadhyaya T, Kaushal H, Dhasarathan V (2020) Multi-band inverted E and U shaped compact antenna for digital broadcasting, wireless, and sub 6 GHz 5G applications. AEU-Int J Electron C 123:153296. https://doi.org/10.1016/j.aeue.2020.153296
15. Varshney G (2021) Tunable terahertz dielectric resonator antenna. SILICON 13:1907–1915
16. Gangwar AK, Alam MS (2019) A miniaturized quad-band antenna with slotted patch for WiMAX/WLAN/WiMAX/WLAN/GSM applications. AEU-Int JElectron C. https://doi.org/10.1016/j.aeue.2019.152911

Estimation of the Post-burning Area of the Fire Hazard Severity Zone in California from Landsat 8 OLI Images Using Deep Learning Machine Intelligence Model

Mohan Singh, Kapil Dev Tyagi, and Arti Joshi

1 Introduction

Within the terrestrial ecosystem, fire is a fundamental phenomenon that affects sustainable construction and atmospheric composition [1]. In recent decades, land management, population changes, and climate change have altered the function of fire, causing new anomalies in fire behavior that has hitherto unheard-of effects on the ecosystem, society, and the climate [2–4]. In 2015, there were 98 thousand hectares of forest that were impacted by fire, according to a study from the Food and Agriculture Organization (FAO). These forests were mostly found in tropical areas, where at least 4.0% of the total area of forests was consumed by fire [5]. It is still difficult for existing models to accurately describe these intricate physical and biological processes, which determine the dispersion of fire sizes, daily fluctuations in fire behavior, and fire spread rate [6]. New measurements are needed to define fire regimes (such as intensity, frequency, and severity) over different time periods and geographic regions. New effective techniques are needed to assess climate, air quality analysis, and assess the effects of fire on ecosystems, to accurately assess the damage caused by fire. Although it is common to predict the behavior of individual fires and regional ensembles using fire spread models, whether they are based

M. Singh (✉)
Department of ECE, G. L. Bajaj Institute of Technology and Management, Greater Noida 201310, India
e-mail: mohansingh571@gmail.com

M. Singh · K. D. Tyagi
Department of ECE, Jaypee Institute of Information Technology, Noida 201309, India

A. Joshi
Department of Geography, D.S.B. Campus, Nainital, Uttarakhand 263002, India

on classical mechanics algorithms or observational formulations. The assessment of their effectiveness is frequently constrained by a dearth of high-quality evidence [7–12]. In the past, statistics on the fire boundary and size were acquired through ground and aircraft observations. Since the late 1970s, satellite remote sensing instruments, particularly Landsat's instruments, have offered an alternative and reliable data source for mapping fire areas and fire intensity [13]. Early applications of satellite datasets frequently handled fire detection as a collection of discrete, pixel-level occurrences on a global grid, neglecting geographical and temporal relationships between those information. The features of particular fires have been tracked using pixel-level burning data in recent works employing concepts from object-oriented categorization and contextual growth [14–16].

2 Literature Review

Any state's burned region may be examined using satellite imagery. For which the image has to be classified first and the analysis of the burning area depends on the accuracy of the classification of the image. Therefore, a good classifier is needed before studying the burning area. The methodologies and approaches for classifying satellite images are discussed in this section. According to Abburu, S. et al., there are primarily three methods—manual, hybrid, and automatic—for classifying pixel values of an image into meaningful classes. Unsupervised and supervised automatic processes are the two categories within it. A binary decision tree, a neural network, and other techniques are used in supervised classification. K-means, SVM, and other algorithms are used in unsupervised classification and also offer a list of highly helpful datasets [17]. However, all unsupervised learning strategies were inadequate when bigger datasets are available [18]. Because of advancements in deep learning techniques and parallel computing, these distant images may be readily classified by initializing weights in training levels so that class predicting in later deep learning layers can become more precise [19]. Deep learning methods abound in literature and include things like VGG, AlexNet, ResNet, and GoogleNet. This article contains information and resources for remotely sensed research [20]. A thorough review of available deep learning approaches for hyperspectral classifications was published in 2017 [21]. The Landsat 8 OLI satellite picture is divided into three primary categories based on the characteristics it contains: low features, mid characteristics, and high features [22]. B. Saikat et al. employed the SAT6 dataset to do this classification using a deep belief network (DBN) and CNN network, achieving an efficiency of 97.946% [23].

In this section, a tabular comparison of the previously conducted researches is provided in Table 1. Yongguan Xiao et al. proposed a division algorithm and apply to vegetated areas for edge identification using "IKONOS stereo pair" to generate city model [24]. Giorgos Mountrakis et al. proposed a support vector machine (SVM) learning approach and used it for evaluating remote sensing implementation [25]. Manoj Pandya et al. introduced K-means algorithm for satellite image classification

[26]. Harish Kundra and Harsh Sadawarti proposed a particle swarm optimization (PSO) and cuckoo search (CS) for remote sensing uses [27]. Nikita Aggarwal et al. proposed an object-based and pixel-based satellite image classification [28]. Zhang, Ce, et al. proposed MLP-CNN approach based on deep spatial features and used for remote sensing [29].

The structure of this article is as follows: Sect. 3 of this study elaborates on the proposed technique. In Sect. 4, research study province zone is discussed. Section 5 provides an exhaustive examination of the results with comparisons. Finally, the paper's efficacy and its application serve as the paper's final conclusion.

3   Applied Methodology

The following steps of the approach, which is suggested for classifying the multispectral Landsat 8 OLI image collection and estimation of the burned area in California, are displayed in Fig. 1 and discussed.

3.1   Create Training Dataset

First the Landsat 8 OLI multispectral image with 11 bands is downloaded. Seven bands of equal resolution are selected from it and combined with the help of ERDAS imagine. The region of interest is selected and separated from the stacked seven bands image. After preprocessing, states (bareland/road/sand, burned area, urban area, vegetation, and water) are defined in it and then samples are taken for all the states. After defining the states, features are extracted from all samples of all states. We create state id for all states, and all the features of samples are shuffled together. The training dataset is prepared, which is size of 711 MB and having 1,048,575 attributes [37–40] as shown in Fig. 1.

3.2   Create Target Dataset

After creating training dataset, target dataset is created using target image from fired zone of the California. First features are extracted from the target image along with geometry coordinates. All the attributes are concatenated together. Finally, the target dataset is obtained with size of 64.8 MB having 1,011,752 attributes, which is finally applied to the trained model and the classified image is obtained [38, 39].

Table 1 Comparative overview of existing methods

Study year	Name of authors	Used method	Work that is being proposed	Concluding remarks
2015	Abburu and Golla [17]	Classification of satellite images	A comparison of automated satellite image categorization algorithms has been presented	The review helps researchers choose the best satellite image categorization methodology or approach for their needs
2015	Abedi et al. [30]	k-nearest neighbor (k-NN)	To improve categorization, the study used KNN to estimate essential forest parameters such as height, densities, and basal area, tree DBH, and cover types	KNN offers accurate classification and usable maps
2015	Ahmed et al. [31]	Backpropagation neural network (BPNN)	NN has been used for a number of applications including speech recognition and computer vision	BPNN is a self-adaptive dynamic system that works well for satellite image classification
2015	Upadhyay et al. [32]	Feedforward neural network (FFNN)	ANN toolkit designed FFNN and error backpropagation	The classification accuracy of ANN declines with neuron count
2016	LeCun et al. [33]	Deep CNN	CNN architectures have a significant number of meta-parameters	The trained networks encode pictures' spatio-spectral information very well
2016	Aggarwal et al. [28]	Object-based classification (OBC) and Pixel-based classification (PBC)	The study focused on object-oriented and pixel-oriented classification	According to the research, object-based classification is more precise than pixel-based classification
2016	Jog et al. [34]	FCM clustering with both KNN and multi-class SVM	The segmented images were obtained using K-means and Fuzzy C-means clustering methods	FCM with multi-class SVM outperforms KNN

(continued)

Table 1 (continued)

Study year	Name of authors	Used method	Work that is being proposed	Concluding remarks
2016	Sabanci et al. [35]	k-NN, multilayer perceptron (MLP), and Naive Bayes	Training and testing are the two stages of the UCI forest type mapping dataset. For training 38% data and for testing 62% data	MLP and KNN are more accurate than others, with 90.4% and 89.10% accuracy, respectively
2017	Cheng et al. [36]	Deep feature learning-based methods, methods based on unsupervised feature learning, handcrafted feature-based methods	Initially, a comprehensive assessment of existing remote sensing work was presented	The research can assist researchers portray features more influentially and multi-view
2018	Zhang et al. [29]	MLP-CNN	Based on deep spatial feature	Used for rural and urban areas

3.3   DLMI Classifier

The deep learning model proposed in this paper has a total of eight layers and is divided into the following three sections: The first section is for the attributes input, which consists of seven nodes in total, and the input attributes (IA) are applied to each node. All the attributes have been taken according to the seven layers of the given satellite image such as IA1, IA2, IA3, IA4, IA5, IA6, and IA7 as shown in Fig. 2. The second part of the model is the hidden section which consists of six layers, namely HL1, HL2, HL3, HL4, HL5, and HL6. Each layer has 240 nodes. All layers in the hidden section are activated with the ReLU function. The model is optimized for high accuracy with minimal error. The third section is the output section which has five nodes in total, and each node represents a state. The output section of the model is activated by the Softmax function. Finally, all the attributes of the target image are classified into these five states (bareland/road/sand, burned area, urban area, vegetation, and water) as shown in Fig. 7.

3.4   Post-Classification Processing

The designed model is trained from the training dataset. After saving the trained model, the target image is applied to the trained model. Finally, the classified image is obtained which is in shapefile format. Shapefile format is converted into geo-tiff file format with the help of QGIS Desktop, which is called processed image and shown in Fig. 8.

Landsat 8 OLI image for California fire hazard zone with 11 bands

↓

Stack all first 7 bands using Erdas imagine

↓

Select fire burned hazard severity zone using Erdas Imagine

↓

Pre-Processing

↓

Define no. of states and sampled data for each state

↓

Extract attributes from each sample for each state

↓

Create state ID for each state

↓

Join together all the attributes for all states

↓

Training Dataset

↓

Create and Define the model	Model Initialization	Model Validation	Model optimization	Training+Testing+Validation dataset	Target Dataset	Extract attributes for target image	Pre-processing	Target image for fired zone of California
				DLMI Classifier				
				Trained Model				

↓

Prediction of the state

↓

Geometry mapping for each state

↓

Classified Image for California hazard zone

↓

Post classified Image
(shapefile format)

↓

Convert shapefile into geo-tiff format using QGIS desktop

↓

Post classification analysis

↓

Post-burning area of the fire hazard severity zone in California

Fig. 1 Applied methodology

Fig. 2 DLMI classifier

4 Research Study Province Zone

The frequency and intensity of wildfires have recently increased due to years of extreme heat and persistent drought and both of these are related to climate change. Because of this the risk of wildfires is increasing throughout California. In 2016, California saw its hottest summer (June–August) on record, with a nationwide average global temperature that was 3.3 °F higher than normal (75.5 °F). Extreme temps in California are affecting soils and plants to lose moisture early in the spring and remain dry during the fall, which makes the landscape more combustible for a longer period of the year. This is in addition to decreased precipitation. The Blue Cut Fire has consumed thousand acres area as of August 2016. It is the fourteenth fire to rank among California's list of top 20 wildfires after 2000. That's why we chose California for research study.

In August 2016, to analyze the Cajon Pass fire in California, we downloaded a pre-fire image on 17th July 2016, which can be seen in Fig. 3, and another post-fire image downloaded on 18th August 2016, which is shown in Fig. 4.
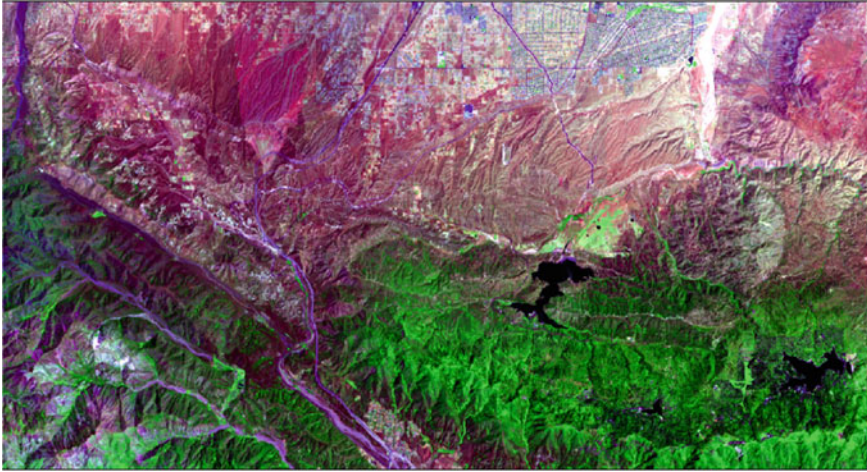
Fig. 3 Pre-fire image on 17–07-2016
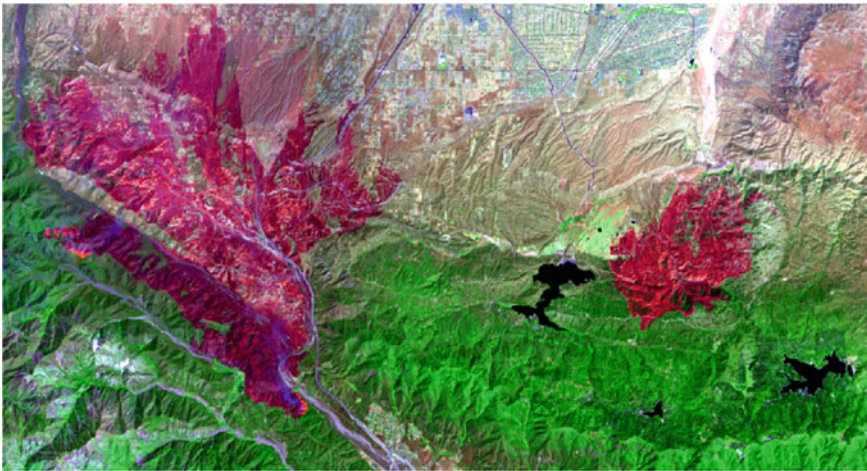
Fig. 4 Post-fire image on 18–08-2016

5 Outcomes Analysis and Discussions

The proposed model is trained with 120 epochs on all 1,048,575 attributes of the training dataset and after which the trained model is tested and evaluated on the test dataset. Finally, the model is validated and optimized for high accuracy and minimum error. The accuracy and model loss of the proposed model are shown in Figs. 5 and 6. The accuracy of the model is obtained at 99.4712% with 0.53% baseline error, which is 8.54% times more accurate as compared to other existing models.
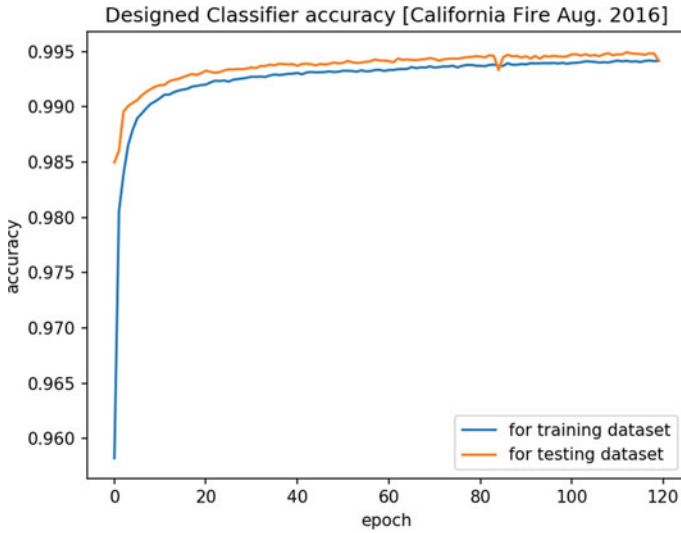
Fig. 5 Proposed model accuracy
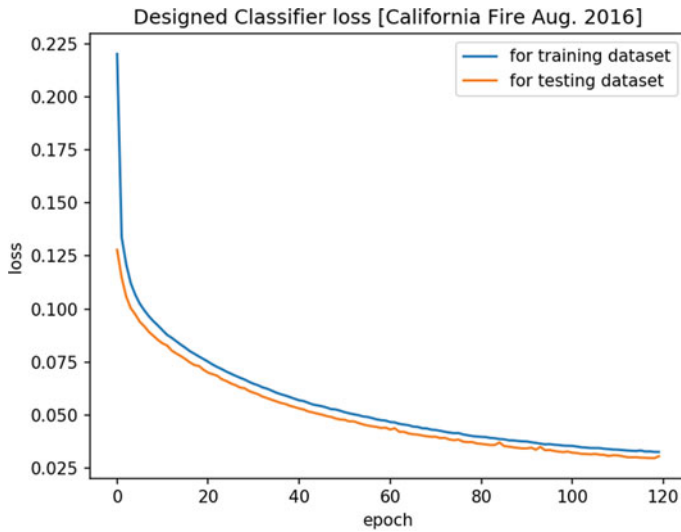
Fig. 6 Proposed model loss

The target dataset is applied to the trained model. The model classifies the target dataset into five states (bareland/road/sand, burned area, urban area, vegetation, and water) as shown in Fig. 7. This classified image is in shapefile format which is further processed for analysis.
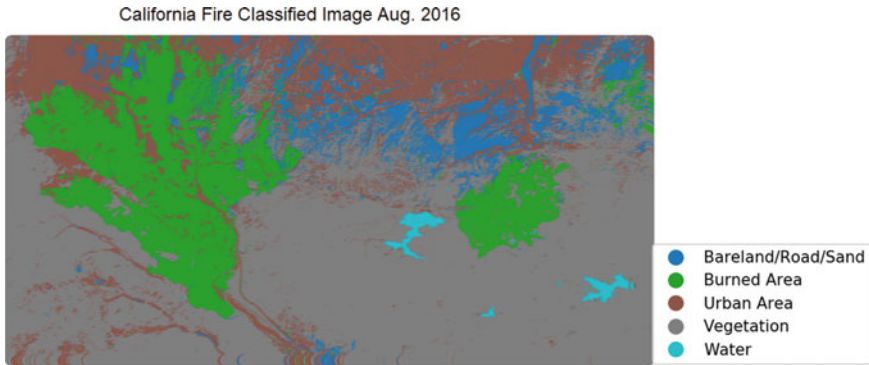
Fig. 7 Post-fire classified image
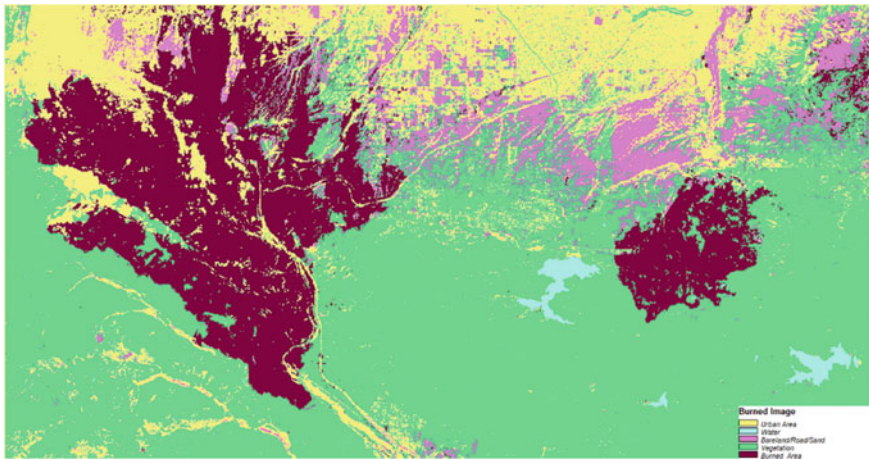
Fig. 8 Post-classified processed image

5.1 Post-classification Analysis

The post-classified image is not suitable for analysis due to being in shapefile format. So it is first converted into geo-tiff format with the help of QGIS Desktop. This geo-tiff format is used for processing and getting post-processed image and shown in Fig. 8.

QGIS Desktop is used to analyze the post-processed image. After analyzing the processed image, the area for each state is obtained, in which 77.229470 km^2 for bareland/road/sand, 171.730934 km^2 for burned area, urban 160.744804 km^2 for area, 461.707482 km^2 for vegetation, and 7.212165 km^2 for water is obtained as shown in Table 2. In this way, the area flowing from the severe fire near Cajon

Pass in California has been estimated, which is estimated to be 171.730934 km^2 (42,435.637956 Acres) as shown in Fig. 9.

A table called a confusion matrix is used to describe how well a classification system performs. The confusion matrix in Table 3 illustrates how a classification algorithm is performed in terms of visualization and summary. The proposed model is designed for five states, and the target dataset is classified into those five states. It can be estimated by looking at the confusion matrix of the model that how high is the accuracy of the model with minimum error. The classification report of the proposed model is shown in Table 4. Table shows precision, recall, f1-score, and support records for the proposed model. For the proposed model, we got f1-scores 0.99, 0.99, 0.99 for the corresponding Accuracy, Macro-Avg, and Weighted Avg. The classification accuracy for the proposed model is 99.4712%.

Comparative study of the proposed model with other earlier approaches is given in Table 5. The classification accuracy of the multilayer perceptron and multilayer perceptron-CNN for the dataset are 81.620% and 90.930%, respectively, whereas

Table 2 Post-burned image data analysis

Class	Name	Count	Cumul	Min	Center	Max	Area (km^2)
1	Bareland/Road/Sand	4,488,437	41,302,731	187.10	191.95	196.80	77.229470
2	Burned area	9,980,691	9,980,691	158	162.85	167.70	171.730934
3	Urban area	9,342,197	51,064,086	245.30	250.65	256	160.744804
4	Vegetation	26,833,603	36,814,294	177.40	182.25	187.10	461.707482
5	Water	419,158	41,721,889	216.20	221.05	225.90	7.212165
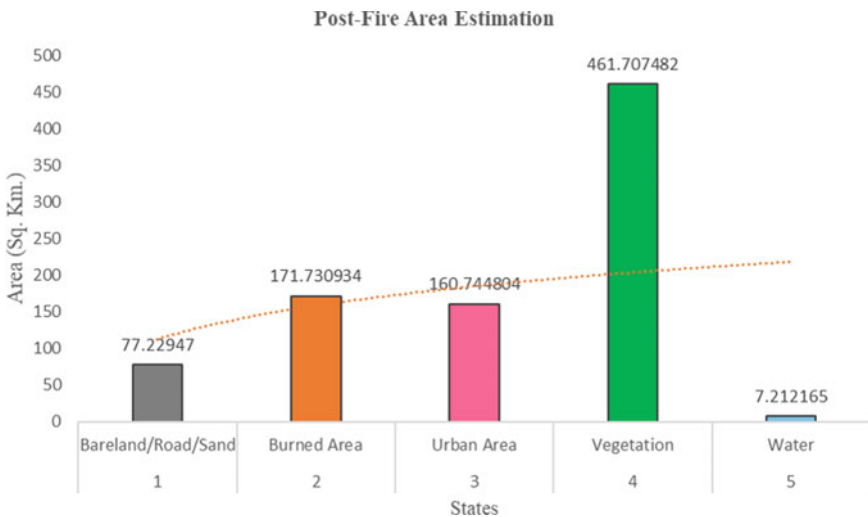
Fig. 9 Post-fire area estimation

Table 3 Confusion matrix

Classes	Bareland/Road/Sand	Burned area	Urban area	Vegetation	Water
Bareland/Road/Sand	332,899	108	549	340	0
Burned area	752	307,840	231	376	0
Urban area	2235	0	199,474	1869	0
Vegetation	247	122	2659	968,014	76
Water	0	22	0	102	14,062

Table 4 Classification report

	Precision	Recall	f1-score	Support
1	0.99	1.00	0.99	333,896
2	1.00	1.00	1.00	309,199
3	0.98	0.98	0.98	203,578
4	1.00	1.00	1.00	971,118
5	0.99	0.99	0.99	14,186
Accuracy			0.99	1,831,977
Macro-Avg	0.99	0.99	0.99	1,831,977
Weighted Avg	0.99	0.99	0.99	1,831,977
Classification Accuracy:	0.994712			

Table 5 Comparative study analysis

Models	Classification accuracy (%)
Multilayer perceptron [41]	81.620
Multilayer perceptron-CNN [29]	90.930
Proposed model	99.471

for the proposed model we got the model accuracy of 99.4712% with baseline error of 0.53%. The accuracy of the proposed model, which is 8.54% more accurate as compared to MLP-CNN and MLP.

6 Conclusion

Active fire is one of the most dangerous risks for wildlife and the environment. The advancement of artificial intelligence technology has greatly aided the improvement of deep learning techniques in image processing and computer vision. In this paper, we proposed a deep learning machine intelligence algorithm to estimate post-fire burning areas using Landsat 8 OLI multispectral dataset. The DLMI model is successfully evaluated in samples with clouds and isolated fire pixels from main fire zone

and estimated 171.730934 km^2 (42,435.637956 acres) post-burning area of the fire hazard severity zone in California. The accuracy of the multilayer perceptron (MLP) and multilayer perceptron-CNN (MLP-CNN) are found as 81.620% and 90.930%, respectively, whereas for the proposed model we got the accuracy of 99.4712% which is 8.54% is more accurate than MLP-CNN and MLP. This work represented a new step in the direction of employing DL approaches for post-burning area estimation using satellite images, which has not been addressed in earlier studies. The results of this study can be applied to successfully estimate the post-burning area of the fires and diminish their damage.

References

1. Bowman DM, Balch JK, Artaxo P, Bond WJ, Carlson JM, Cochrane MA, D'Antonio CM, DeFries RS, Doyle JC, Harrison SP, Johnston FH, Pyne SJ (2009) Fire in the Earth system. Science 324(5926):481–484. https://doi.org/10.1126/science.1163886
2. Walker XJ, Baltzer JL, Cumming SG, Day NJ, Ebert C, Goetz S, Johnstone JF, Potter S, Rogers BM, Schuur EA, Turetsky MR, Mack MC (2019) Increasing wildfires threaten historic carbon sink of boreal forest soils. Nature 572(7770):520–523. https://doi.org/10.1038/s41586-019-1474-y
3. Kirchmeier-Young MC, Gillett NP, Zwiers FW, Cannon AJ, Anslow FS (2019) Attribution of the influence of human-induced climate change on an extreme fire season. Earth's Future 7(1):2–10. https://doi.org/10.1029/2018EF001050
4. Yoon JH, Wang SYS, Gillies RR, Hipps L, Kravitz B, Rasch PJ (2015) Extreme fire season in California: a glimpse into the future? Bull Am Meteorol Soc 96(12):S5–S9. https://www.jstor.org/stable/26233131
5. Nesha MK, Herold M, De Sy V, Duchelle AE, Martius C, Branthomme A, Pekkarinen A (2021) An assessment of data sources, data quality and changes in national forest monitoring capacities in the Global forest resources assessment 2005–2020. Environ Res Lett 16(5):054029. https://doi.org/10.1088/17489326/abd81b
6. Hantson S, Arneth A, Harrison SP, Kelley ID, Prentice IC, Rabin SS, Archibald S, Mouillot F, Arnold SR, Artaxo P, Bachelet D, Yue C (2016) The status and challenge of global fire modelling. Biogeosciences 13(11):3359–3375. https://doi.org/10.5194/bg-13-3359-2016
7. Linn RR (1997) A transport model for prediction of wildfire behavior. New Mexico State University
8. Coen JL, Cameron M, Michalakes J, Patton EG, Riggan PJ, Yedinak KM (2013) WRF-Fire: coupled weather–wildland fire modeling with the weather research and forecasting model. J Appl Meteorol Climatol 52(1):16–38. https://doi.org/10.1175/JAMC-D-12-023.1
9. Finney MA (1998) FARSITE, fire area simulator—model development and evaluation (No. 4). US Department of Agriculture, Forest Service, Rocky Mountain Research Station
10. Finney MA, Grenfell IC, McHugh CW, Seli RC, Trethewey D, Stratton RD, Brittain S (2011) A method for ensemble wildland fire simulation. Environ Model Assess 16(2):153–167. https://doi.org/10.1007/s10666-010-9241-3
11. Andrews PL (2013) Current status and future needs of the BehavePlus fire modeling system. Int J Wildland Fire 23(1):21–33. https://doi.org/10.1071/WF12167
12. Clements CB, Zhong S, Goodrick S, Li J, Potter BE, Bian X, Heilman WE, Charney JJ, Perna R, Jang M, Lee D, Aumann G (2007) Observing the dynamics of wildland grass fires: FireFlux—a field validation experiment. Bull Am Meteorol Soc 88(9):1369–1382. https://doi.org/10.1175/BAMS-88-9-1369

13. Eidenshink J, Schwind B, Brewer K, Zhu ZL, Quayle B, Howard S (2007) A project for monitoring trends in burn severity. Fire Ecol 3(1):3–21. https://doi.org/10.4996/fireecology.0301003
14. Andela N, Morton DC, Giglio L, Paugam R, Chen Y, Hantson S, Van Der Werf GR, Randerson JT (2019) The global fire atlas of individual fire size, duration, speed and direction. Earth Syst Sci Data 11(2):529–552. https://doi.org/10.5194/essd-11-529-2019
15. Artés T, Oom D, De Rigo D, Durrant TH, Maianti P, Libertà G, San-Miguel-Ayanz J (2019) A global wildfire dataset for the analysis of fire regimes and fire behaviour. Sci Data 6(1):1–11. https://doi.org/10.1038/s41597-019-0312-2
16. Lizundia-Loiola J, Otón G, Ramo R, Chuvieco E (2020) A spatio-temporal active-fire clustering approach for global burned area mapping at 250 m from MODIS data. Remote Sens Environ 236:111493. https://doi.org/10.1016/j.rse.2019.111493
17. Abburu S, Golla SB (2015) Satellite image classification methods and techniques: a review. Int J Comput Appl 119(8). https://doi.org/10.5120/21088-3779
18. Elkholy MM, Mostafa MS, Ebeid HM, Tolba M (2022) Unsupervised hyperspectral band selection with deep autoencoder unmixing. Int J Image Data Fusion 13(3):244–261. https://doi.org/10.1080/19479832.2021.1972047
19. Vali A, Comai S, Matteucci M (2020) Deep learning for land use and land cover classification based on hyperspectral and multispectral earth observation data: a review. Remote Sens 12(15):2495. https://doi.org/10.3390/rs12152495
20. Zhu XX, Tuia D, Mou L, Xia GS, Zhang L, Xu F, Fraundorfer F (2017) Deep learning in remote sensing: a comprehensive review and list of resources. IEEE Geosci Remote Sens Mag 5(4):8–36. https://doi.org/10.1109/MGRS.2017.2762307
21. Ghamisi P, Plaza J, Chen Y, Li J, Plaza AJ (2017) Advanced spectral classifiers for hyperspectral images: a review. IEEE Geosci Remote Sens Mag 5(1):8–32. https://doi.org/10.1109/MGRS.2016.2616418
22. Zhang F, Du B, Zhang L (2014) Saliency-guided unsupervised feature learning for scene classification. IEEE Trans Geosci Remote Sens 53(4):2175–2184. https://doi.org/10.1109/TGRS.2014.2357078
23. Basu S, Ganguly S, Mukhopadhyay S, DiBiano R, Karki M, Nemani R (2015) Deepsat: a learning framework for satellite imagery. In: Proceedings of the 23rd SIGSPATIAL international conference on advances in geographic information systems, pp 1–10. https://doi.org/10.1145/2820783.2820816
24. Hemati M, Hasanlou M, Mahdianpari M, Mohammadimanesh F (2021) A systematic review of Landsat data for change detection applications: 50 years of monitoring the earth. Remote Sens 13(15):2869. https://doi.org/10.3390/rs13152869
25. Mountrakis G, Im J, Ogole C (2011) Support vector machines in remote sensing: a review. ISPRS J Photogramm Remote Sens 66(3):247–259. https://doi.org/10.1016/j.isprsjprs.2010.11.001
26. Pandya M, Baxi A, Potdar MB, Kalubarme MH, Agarwal B (2013) Comparison of various classification techniques for satellite data. Int J Sci Eng Res 4(1). https://doi.org/10.21608/ijicis.2021.78853.1098
27. Kundra H, Sadawarti H (2015) Hybrid algorithm of cuckoo search and particle swarm optimization for natural terrain feature extraction. Res J Inf Technol 7(1):58–69. https://scialert.net/abstract/?doi=rjit.2015.58.69
28. Aggarwal N, Srivastava M, Dutta M (2016) Comparative analysis of pixel-based and object-based classification of high resolution remote sensing images—a review. Int J Eng Trends Technol 38(1):5–11. https://doi.org/10.14445/22315381/IJETT-V38P202
29. Zhang C, Pan X, Li H, Gardiner A, Sargent I, Hare J, Atkinson PM (2018) A hybrid MLP-CNN classifier for very fine resolution remotely sensed image classification. ISPRS J Photogramm Remote Sens 140:133–144. https://doi.org/10.1016/j.isprsjprs.2017.07.014
30. Abedi R, Bonyad AE (2015) Estimation and mapping forest attributes using k nearest neighbor, method on IRS-P6 LISS III satellite image data. Ecologia Balkanica 7(1)

31. Ahmed B, Al Noman MA (2015) Land cover classification for satellite images based on normalization technique and artificial neural network. Int Conf Comput Inf Eng (ICCIE) 138–141. https://doi.org/10.1109/CCIE.2015.7399300

32. Upadhyay A, Singh SK, Singh P (2015) Comparative study of artificial neural network based classification of 1RS LISS-III satellite images. Int Conf Green Comput Internet Things (ICGCIoT) 961–965. https://doi.org/10.1109/ICGCIoT.2015.7380601

33. LeCun Y, Bengio Y, Hinton G (2015) Deep learning. Nature 521(7553):436–444. https://doi.org/10.1038/nature14539

34. Jog SM, Dixit M, Rajgopalan A, Ranade SD (2016) KNN and SVM based satellite image classification. Int J Innovative Res Electr, Electron, Instrum Control Eng 4(6):1588–2162. https://doi.org/10.17148/IJIREEICE.2016.4637

35. Sabancı K, Ünlersen MF, Polat K (2016) Classification of different forest types with machine learning algorithms

36. Cheng G, Han J, Lu X (2017) Remote sensing image scene classification: benchmark and state of the art. Proc IEEE 105(10):1865–1883. https://doi.org/10.1109/JPROC.2017.2675998

37. Singh M, Tyagi KD (2021) Pixel based classification for Landsat 8 OLI multispectral satellite images using deep learning neural network. Remote Sens Appl: Soc Environ 24:100645. https://doi.org/10.1016/j.rsase.2021.100645

38. Singh M, Tyagi KD (2021) Multispectral remote sensing image classification using modern machine intelligence approach. In: 2021 7th international conference on signal processing and communication (ICSC). IEEE, pp 216–221. https://doi.org/10.1109/ICSC53193.2021.9673205

39. Singh M, Tyagi KD (2022) Detection of expanded reformed geographical area in bi-temporal multispectral satellite images using machine intelligence neural network. J Indian Soc Remote Sens 50(4):623–633. https://doi.org/10.1007/s12524-021-01476-6

40. Singh M, Tyagi KD, Singh A, Singh KK (2020) Detection of changes in Landsat images using hybrid PSO-FCM. Procedia Comput Sci 167:423–430. https://doi.org/10.1016/j.procs.2020.03.251

41. Del Frate F, Pacifici F, Schiavon G, Solimini C (2007) Use of neural networks for automatic classification from high-resolution images. IEEE Trans Geosci Remote Sens 45(4):800–809. https://doi.org/10.1109/TGRS.2007.892009

Algorithm Analysis in NOMA

Moirangthem Rushdie Devi and Aheibam Dinamani Singh

1 Introduction

With the concern of provision of low latency, high capacity and massive connection in wireless communication, non-orthogonal multiple access (NOMA) is being considered as one of the techniques for 5G wireless networks [1]. Before NOMA came into the picture, wireless communication systems have been utilizing orthogonal multiple access (OMA) techniques. In OMA [2], allocation of resources to multiple UE's takes place orthogonally. OMA techniques include time-division multiple access (TDMA), frequency-division multiple access (FDMA), code-division multiple access (CDMA). In TDMA, multiple users undergo time-division technique by distributing the same frequency channel. In FDMA, multiple UEs follow the frequency-division multiplexing technique where communications are allowed only during their particular frequency slot. In CDMA, multiple users share the entire time and frequency, they are differentiated by the codes. As the OMA technique considers orthogonality, each UE has been allocated one resource block at a time which in turn does not give the desired high capacity and low latency [3].

Contrary to OMA, NOMA provides sharing of resources to multiple UEs which ensures high spectral efficiency, low transmission latency by allocating one frequency channel at the same time. NOMA techniques include power-domain NOMA and code-domain NOMA [4]. In code-domain NOMA, multiplexing is based on different code levels whereas in power-domain NOMA, multiplexing is done based on different power levels. This paper focuses mainly on power-domain NOMA that utilizes superposition coding (SC) and successive interference cancellation (SIC) at the transmitter and receiver side, respectively. NOMA allows multiple UEs to transmit and receive

M. R. Devi (✉) · A. D. Singh
Department of Electronics and Communication Engineering, National Institute of Technology, Imphal, Manipur 795004, India
e-mail: rushdie997@gmail.com

information simultaneously using the same frequency. This feature is only possible due to SC and SIC [5].

1.1 Benefits of NOMA

This subsection presents how NOMA is superior to OMA in several ways [6, 7], such as.

a. NOMA achieves higher spectral efficiency by utilizing the same frequency and time resource for multiple users and interference mitigation through SIC,
b. NOMA supports higher connection density by superimposing the signal of multiple UE's on the same resource block,
c. When compared to OMA, NOMA has lower latency as it does not require separate time slot for transmitting information and
d. As less power is assigned to stronger UE and more power is assigned to weaker UE. Thus, NOMA maintains user fairness.

2 NOMA System Model

In this section, an overview of uplink and downlink NOMA is introduced, as shown in Fig. 1. For simplicity purpose, the system model of NOMA is analysed with a single antenna at the base station (BS) and two users (UE).

2.1 Downlink NOMA Network

In downlink NOMA, the transmitter is the BS while the receiver will be the two UEs. The BS transmits the superposed signal to both the UEs. The multiple UEs sharing the

Fig. 1 Uplink and downlink communication
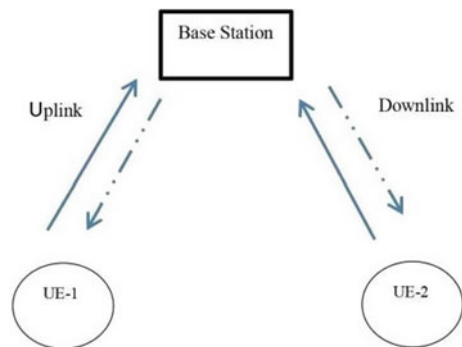
same time and frequency resources are then retrieved at the receiver side. Hence, this process caused increased in spectral efficiency [8]. The superposed signal, which is the combination signals of the two UEs, is allocated with different power coefficients. Power coefficients are allocated according to their channel condition or the interval between the BS and the UE's, in inversely proportional way. The UE which is far from the BS is allocated larger power and lesser power to the other. Also, the channel gain is quasi-static, i.e. constant over the entire transmission time interval [4]. The sum of P_i is equal to P_{total} [9].

Here, assuming that UE-1 is closer to the base station, so it is allocated lower power in comparison with UE-2, which is farther from base station.

The superposed signal is represented as [10, 11].

$$x_s = \sum_{i=1}^{N} P_i x_i, \tag{1}$$

where P_i indicates the allocated power for symbol x_i of the ith UE, and N denotes the number of UE's.

At the receiver side of downlink NOMA, the decoding of UE's message from the superposed signal takes place. This process is single input multiple output (SIMO) [4]. The received signal at the ith UE's [12, 13] is

$$y_i = \sum_{i=1}^{N} h_i \sqrt{a_i x_i} + n_i, \tag{2}$$

where a_i is the power scaling factors express in terms of amplitude, and h_i is the channel gain experienced by ith UE.

2.2 Uplink NOMA Network

In uplink NOMA, the transmitter will be the UE's while the receiver is the BS. Depending upon the channel condition or the distance between the BS and the UEs, signals of the UEs are transmitted with different power levels. The user experiencing lower channel gain transmits higher power, whereas the user that experienced higher channel gain transmits low power [14]. Assuming the same for uplink case, the channel is quasi-static [4]. Now, the superposed signal from both the users is being received by the BS which will be decoded accordingly. This process is multiple input single output (MISO) [4]. In both uplink and downlink, both the user signals are weighted with different powers.

The superposed signal received by the BS [14, 15] is.

NOMA signal $\longrightarrow$ Far user signal decoding $\longrightarrow$ Far user data

Fig. 2 Far-user processing

NOMA signal $\longrightarrow$ Far user signal decoding $\longrightarrow$ Reconstruct far user signal

Near user data $\longleftarrow$ Near user signal decoding $\longleftarrow$ Subtract from NOMA signal
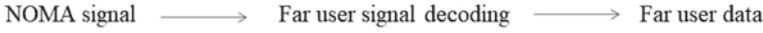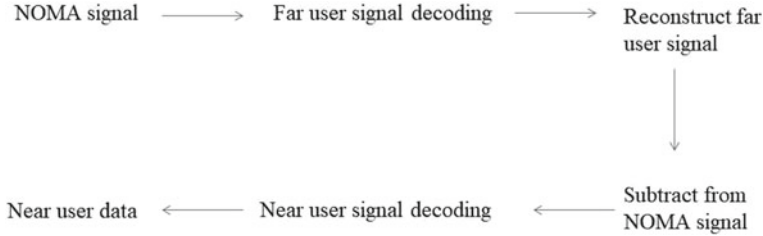
Fig. 3 Near-user processing

$$x_{Bs} = \sum_{i=1}^{N} P_i x_i + n. \tag{3}$$

The superposed signal comprises of UE-1's signal x_1 and UE-2's signal x_2 and also the noise, n. Following the power-domain NOMA principle [15], at the receiver side, i.e. at the BS, it has to perform SIC of the superposed signal which is transmitted from the UE's according to their respective power levels. A simple figure showing the processing of far-user signal in Fig. 2, and near-user processing in Fig. 3, are shown, respectively.

3 Algorithm Analysis in Power-Domain NOMA

In this section, some ideas about SC and SIC are discussed. These two main techniques play a major role in appreciating power-domain NOMA.

3.1 Superposition Coding (SC)

Superposition coding is a process of simultaneously communicating multiple users' information at the same time by a single source [1]. Simply, it is power domain multiplexing. SC process is always implemented at the transmitter side, whether it may be uplink or downlink communication.

The process of superposition coding is as follows:

a. Consider two users x_1 for UE-1 and x_2 for UE-2 which are going to communicate simultaneously,
b. The user's data x_1 and x_2 undergo digital modulation before transmission,

c. The user's data x_1 and x_2 are multiplied with the required power scaling factors, which are expressed in terms of amplitude. The power scaling factors for UE—i must follow the condition that $\sum_{i=1}^{N} a_i = 1$ [1] and

d. The user's data along with the power scaling factors are then added to form the SC signal.

3.2   Successive Interference Cancellation (SIC)

Successive interference cancellation is an algorithm where information is successively decoded according to their power levels [1], while the rest are treated as interference [16]. It is used for detecting the desired signals. SIC process is implemented at the receiver side always.

When SIC is applied, the UE signal which has the largest power is decoded first, while treating the rest as interference [9]. The required signal is then subtracted from the combined signal in order to decode the next UE's signal and so on. Before applying SIC, UEs are sequenced in accordance with their respective signal strength so that the stronger signal is decoded first by the receiver [7]. In brief [17], the process of decoding the superposed signal is expressed as follows:

a. The superposed signal x_{Bs} is received and is first decoded by demodulation technique. From this step, the user's signal which has been allocated higher power is detected and treated the rests as interference,

b. The decoded signal is then multiplied with its respective weight and then subtracted from x_{Bs} and

c. By applying the demodulation technique to the result from step (b), gives the resulting user signal which has been allocated lower power.

4   Result

The simulation analysis was carried out in MATLAB R2019a. For the simulation of the two UE signals, the value of a_1 and a_2 has been allocated 0.75 and 0.25, respectively. It is assumed that the transmission bandwidth and power for the overall system is one Hertz and one Watt, respectively.

In Fig. 4, user x_1 and x_2 are allocated 16 bits which undergoes digital modulation, multiplied with its corresponding weight and added to give the superposed signal. Figure 5 shows the graphical representation of x_1 and x_2 signals. Figure 6 shows the simulated result of the superposed signal, x_{Bs}. Figures 7 and 8 show the decoded signal of user x_1 and x_2.

$x_1 = 1\ 0\ 1\ 0\ 0\ 0\ 0\ 0\ 1\ 1\ 1\ 1\ 0\ 0\ 0\ 0$ $x_2 = 0\ 1\ 1\ 0\ 1\ 1\ 1\ 1\ 0\ 0\ 0\ 0\ 1\ 1\ 1\ 1$

Digital modulation

$x_1 = +1\ -1\ +1\ -1\ -1\ -1\ -1\ -1\ +1\ +1\ +1\ +1\ -1\ -1\ -1\ -1$ $x_2 = -1\ +1\ +1\ -1\ +1\ +1\ +1\ +1\ -1\ -1\ -1\ -1\ +1\ +1\ +1\ +1$

Multiplied with its
weight and added.

$x_{BS} = 0.366\ -0.366\ 1.366\ -1.366\ -0.366\ -0.366\ -0.366\ -0.366\ 0.366\ 0.366\ 0.366\ 0.366\ -0.366\ -0.366\ -0.366$
-0.366
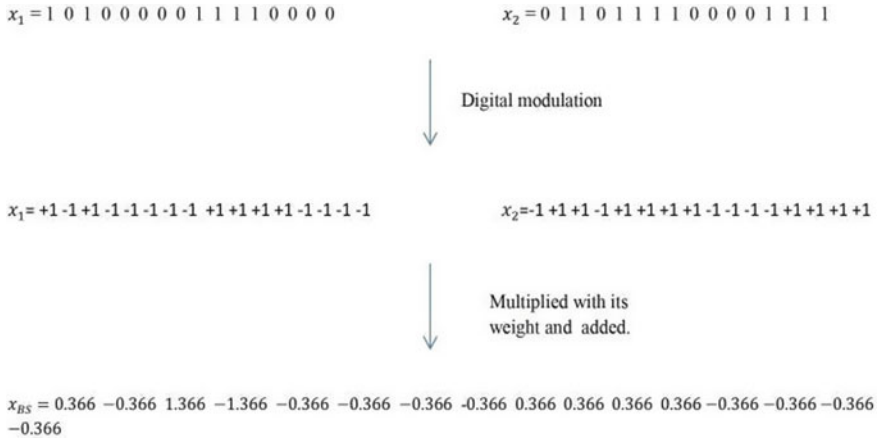
Fig. 4 User data x_1 and x_2 and the SC signal xBS
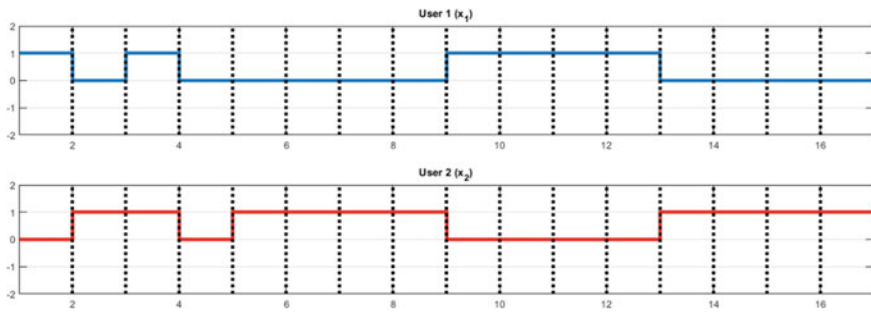
Fig. 5 Graphical representation of x_1 and x_2 signal
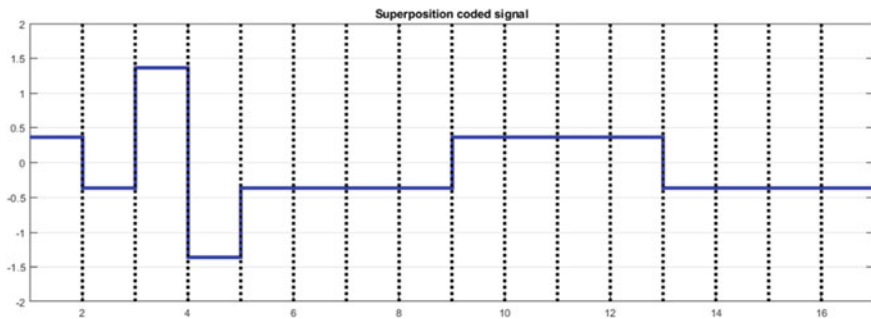
Fig. 6 Simulated result of the SC signal
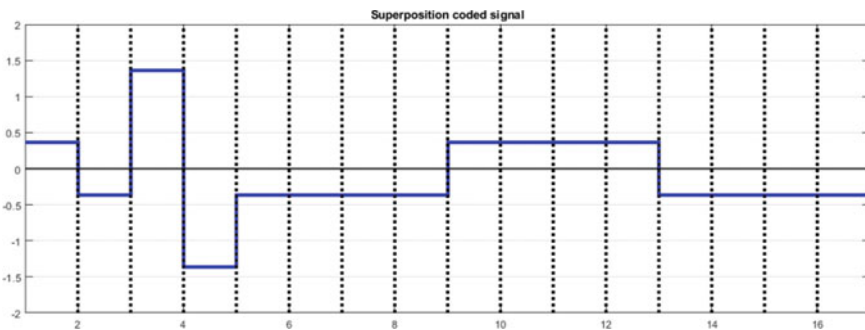
Fig. 7 Result of the far-user ($x_1\prime$) after applying SIC
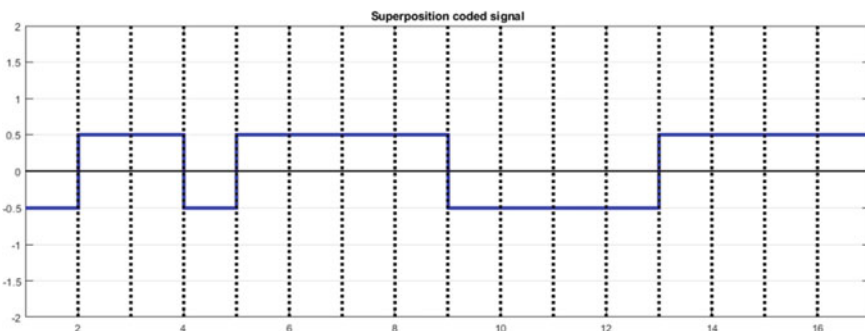
Fig. 8 Result of the near-user ($x_2\prime$) after applying SIC

5 Conclusion

In this paper, two UEs are considered for the analysis. The UEs are allocated 16 bits of data which are multiplied with its corresponding weight to give the perfect SC at the transmitter. SIC is then applied to the resultant SC signal to give error-free user signals at the receiver. Thus, it illustrates how NOMA requires SC at the transmitter side and SIC at the receiver side in order to give every individual user separate messages.

References

1. Islam SMR, Avazov N, Dobre OA, Kwak K (2017) Power-domain non-orthogonal multiple access (NOMA) in 5G systems: potentials and challenges. IEEE Comm Surveys Tut 19(2):721–742
2. Tabassum H, Ali MS, Hossain E, Hossain MJ, Kim DI (2017) Uplink vs. downlink NOMA in cellular networks: challenges and research directions. In: 85th Vehicular Technology Conference 2017, VTCSpring, pp 1–7. IEEE, Sydney

3. Dai L, Wang B, Ding Z, Wang Z, Chen S, Hanzo L (2018) A survey of non-orthogonal multiple access for 5G. IEEE Comm Surveys Tut 20(3):2294–2323

4. Srivastava S, Dash PP, Kumar S (2020) International symposium on 5G & beyond for rural upliftment, pp 28–38. River Publishers, Dhanbad

5. Islam SMR, Zeng M, Dobre OA (2017) NOMA in 5G systems: exciting possibilities for enhancing spectral efficiency. IEEE 5G Tech Focus 1(2)

6. Wang P, Xiao J, Ping L (2006) Comparison of orthogonal and non-orthogonal approaches to future wireless cellular systems. IEEE Veh Technol Mag 1(3):4–11

7. Dai L, Wang B, Yuan Y, Han S, Chih-lin I, Wang Z (2015) Non-orthogonal multiple access for 5G: solutions, challenges, opportunities, and future research trends. IEEE Comm Magazine 53(9):74–81

8. Higuchi K, Benjebbour A (2015) Non-orthogonal multiple access (NOMA) with successive interference cancellation for future radio access. IEICE Trans Comm 98(3):403–414

9. Bhatia V, Swami P, Sharma S, Mitra R (2020) Non-orthogonal multiple access as an enabler for massive connectivity for 5G and beyond networks. J Indian Inst Sci 100(8):337–348

10. Ali NA, Mourad HM, ElSayed HM, Soudani ME, Amer HH, Daoud RM (2016) General expressions for downlink signal to interference and noise ratio in homogeneous and in heterogeneous LTE-Advanced networks. J Adv Res 7(6):923–929

11. Saito Y, Kishiyama Y, Benjebbour A, Nakamura T, Li A, Higuchi K (2013) Non-orthogonal multiple access (NOMA) for cellular future radio access. In: 77th Vehicular Technology Conference (VTC Spring), pp 1–5. IEEE, Germany

12. Zhang X, Xie J, Yue X, Kang S (2021) Effective capacity analysis of NOMA networks with short packets. Appl Sci 11(23):11438

13. Sadia H, Zeeshan M, Sheikh SA (2018) Performance analysis of downlink power domain NOMA under fading channels. In: 12th International Conference ELEKTRO, pp 1–6. IEEE, Czech Republic

14. Mahmoud A, Mesut T, Selahattin G, Kurt K, Kucur O (2018) A tutorial on nonorthogonal multiple access for 5G and beyond. Wireless Comm Mobile Comp (5)

15. Al-Imari M, Xiao P, Imran MA, Tafazolli R (2014) Uplink non-orthogonal multiple access for 5G wireless networks. In: 11th International Symposium on Wireless Communications Systems (ISWCS), pp 781–785. IEEE, Spain

16. Srikamu C, Jayabharathy R (2022) Comparative analysis of ergodic sum capacity of cooperative NOMA aided with spatial modulation. Wireless Pers Commun 123(4):3771–3786

17. Benjebbour A, Saito Y, Kishiyama Y, Li A, Harada A, Nakamura T (2013) Concept and practical considerations of non-orthogonal multiple access (NOMA) for future radio access. In: 2013 International Symposium on Intelligent Signal Processing and Communication Systems, pp 770–774. IEEE, Japan

A New High-Order Electronically Tunable Oscillator (HO-ETO)

Manoj Joshi, Rakesh Kumar, Mohan Singh, and Shilpa Choudhary

1   Introduction

The level of integration of VLSI technology is now capable of reaching 100 million single-chip transistors. Due to the efficient use of BJTs, MOSFETs, Bi-CMOS, and high-speed GaAs devices, integrated circuits (ICs) design has become technology specifically, regardless of advancements in VLSI technology. To increase the effectiveness of analogue designers in ICs-+, nowadays only active components are employing in VLSI circuit design. One of the most often used active components such as operational amplifiers (Op-Amps) is essential for the creation and processing of analogue signal circuits. Op-Amps laid the groundwork for the creation of several circuits and systems. However, the high-frequency performance of these circuits is constrained by their fixed gain bandwidth multiplication, power requirements and low slew rate Op-Amps [1].

The approach suggested resolving this problem has advanced signal processing in its current mode (CM). Therefore, current is selected as the active parameter rather than voltage. Current-mode (CM) processing, which uses current signals for signal processing, has recently gained popularity as an alternative design technique. CMOS techniques can greatly increase device efficiency in terms of bandwidth, signal linearity, slew rate and power consumption [2]. Sedra and Smith introduced the first current-mode circuits in 1968, which sparked the creation of the current conveyor

M. Joshi (✉) · R. Kumar

Department of Electronics and Communication Engineering, JSS Academy of Technical Education, Noida, India

e-mail: manojjoshi1506@gmail.com

R. Kumar

Department of ECE, Delhi Technological University (DTU), New Delhi, India

M. Singh · S. Choudhary

Department of ECE, G.L. Bajaj Institute of Technology and Management, Noida, India

(CC). The three generations of current-mode active block as first generation as CCI, second generation as CCII and third generation as CCIII are available in the literature [1, 3]. They have quite varied port characteristics and were originally made reported in 1968, 1970 and 1995, respectively. After that, there have been whispers of the existence of new analogue building components including the CCCII [4, 5], DO-CCII and OTA [6, 7], CCII [8], current mode [9, 10], CDTA [11], and DDCC [12], differential voltage CC [13], dual-X current conveyor TA [14], fully differential CCII, inverting second-generation CCII, fully balanced second-generation CC, second-generation dual-X current conveyor (DXCCII), CCC-transconductance amplifier (CCCCTA) [15], current differencing TA [16], VDBIA [17], two multiple outputs CCIIs [18], ZC-CG-CDBA[19], MCCFTA [20], DXCCII [21], etc., are used to design the higher-order oscillator. However, some other blocks such as CCDDCC [22], CFOA and OTA [23], MOCC, FTFN, current follower transconductance amplifier (TA), controlled fully balanced CCII, differential voltage CCII, inverting CCIII, current through TA, DX-current-controlled differencing TA [24], OTRA [25] and DX-current-controlled TA [26] are also available in the scientific literature for designing the different analogue circuits.

The oscillator is a key component of signal-processing circuits in our contemporary analogue design environment. The main objective is to produce a high-order electrically tunable oscillator by developing an active component known as the extra-X current conveyor (EX-CCII). For instance, the quadrature oscillator is used in telecommunications because the circuit produces two sinusoids with a 90-degree phase difference. It is also used for measuring reasons in vector generators and selective voltmeters. Therefore, quadrature oscillators are an essential part of many communication and measurement systems. Designing a higher-order oscillator is one of the key goals of this project because it offers good characteristics with less distortion.

2   Oscillator Basic

The oscillator is a key component of signal-processing circuits in our contemporary analogue design environment [1]. The main objective is to produce a high-order electrically tunable oscillator by developing an active component known as the extra-X CCII. For instance, the HO-ETO is used in telecommunications because the circuit produces two sinusoids with a phase difference. Therefore, HO-ETO is an essential part of many communication and measurement systems. Designing a higher-order oscillator is one of the key goals of this research because it offers good characteristics with less distortion. An essential component that is widely used in electrical and communication engineering applications is the oscillator. A consistent continuous signal is produced by an electrical circuit. It has active and passive circuit components and produces repeated waveforms at the output even without a direct external input signal. Fundamentally, oscillators change a DC unidirectional current flow into continuous periodic with the desired frequency, as determined by the elements in its

Fig. 1 Typical oscillator topology
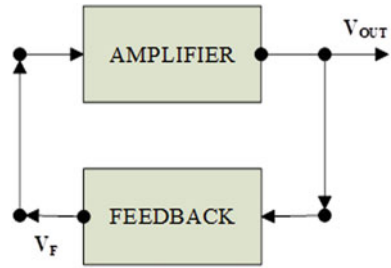
circuit. Oscillators are simply amplifier circuits that produce continuous waveform by sending some of the output with the help of positive feedback back into the input. By making up for circuit losses, an amplifier is used to amplify with the help of an active element, it retains the same-phase feedback signal which is responsible for maintaining the oscillations. The power supply's electrical noise causes oscillations to begin as soon as it is turned ON in the system (Fig. 1).

Within the loop, this noise signal moves around, is amplified, and quickly collapses into a sine wave with a single frequency. The closed-loop gain of the typical oscillator topology is obtained as:

$$G = \frac{A}{1 + A\beta},\tag{1}$$

where A and β are the forward voltage and feedback gain of the amplifier. The oscillations are amplified in this case if $A\beta > 1$; however, they are suppressed if $A\beta < 1$ and, $A\beta = 1$ results in oscillations with constant amplitude as shown in Fig. 2, respectively. On another way, the oscillation fades away, distorted and the sufficient oscillatory circuit can only be created if the feedback gain is low, high and unity, respectively.
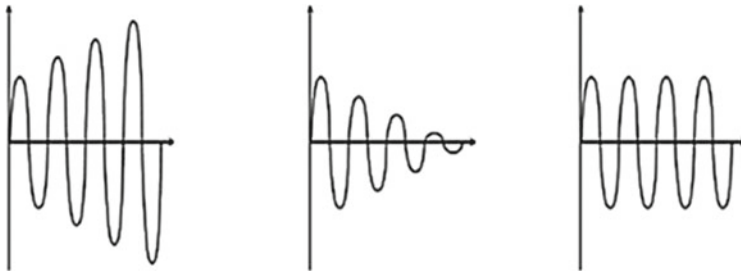
Fig. 2 Response of typical oscillator

2.1 Types of Oscillators

Although there are several varieties of oscillators, they can be broadly divided into two groups: harmonic oscillators (sometimes called linear oscillators) and relaxation oscillators. In addition, energy is frequently transferred from one energy-storing component to another in a harmonic oscillator, and the feedback loop controls the oscillation frequency, which is determined by the time constants involved in the charging and discharging of the phase. The different types of oscillators such as RC phase shift, LC (Hartley, Colpitts, Clapp) oscillators, Wien bridge, crystal oscillators, Gunn oscillators and ring oscillators are reported in scientific literature.

2.2 Application of Oscillators

Oscillators are a simple and inexpensive way to create a signal at a specific frequency. As an example, an RC oscillator generates low-frequency signals, an LC oscillator generates high-frequency signals, and an oscillator based on an Op-Amp generates steady frequencies. With potentiometer setups, the oscillation frequency can be tuned by adjusting the component value. Oscillators are frequently employed in clock pulses, alarms, buzzers, metal detectors, inverters and colourful lighting, among other audio and video systems (radios, televisions and other communication devices).

3 Circuit Description of Proposed Oscillator

The new EX-CCII is implemented by using CCII as shown in Fig. 3a with symbol. The CMOS-based CCII and aspect ratio used in the simulation are taken from [2–28]. The third-order oscillator satisfying the novel circuit presented in this article is depicted in Fig. 3b. The proposed HO-ETO is designed with only three grounded capacitors, and the tuning property is achieved by using the register connected in X Terminal of CCII. In the proposed HO-ETO, both current and voltage modes are supported by the suggested design. The simulation is performed by using Cadence software with active block CCII in order to verify the findings of the theoretical analysis. The third-order oscillator equation can be obtained by manipulating the terms mathematically and rearranging them as follows:

$$s^3(C1C2C3R3R2R1) + s^2\begin{pmatrix} C1C2R1R2 + C1C3R1R3 - R1R3C2C3 \\ -C3C2R2 - C2C3R2R3 - R3C3^2 \end{pmatrix}$$
$$+ s(C1R1 - C2R1 - 2C2R2 - C3R3) + 1 = 0. \quad (2)$$

From Eq. (2), the frequency of oscillation is given by:
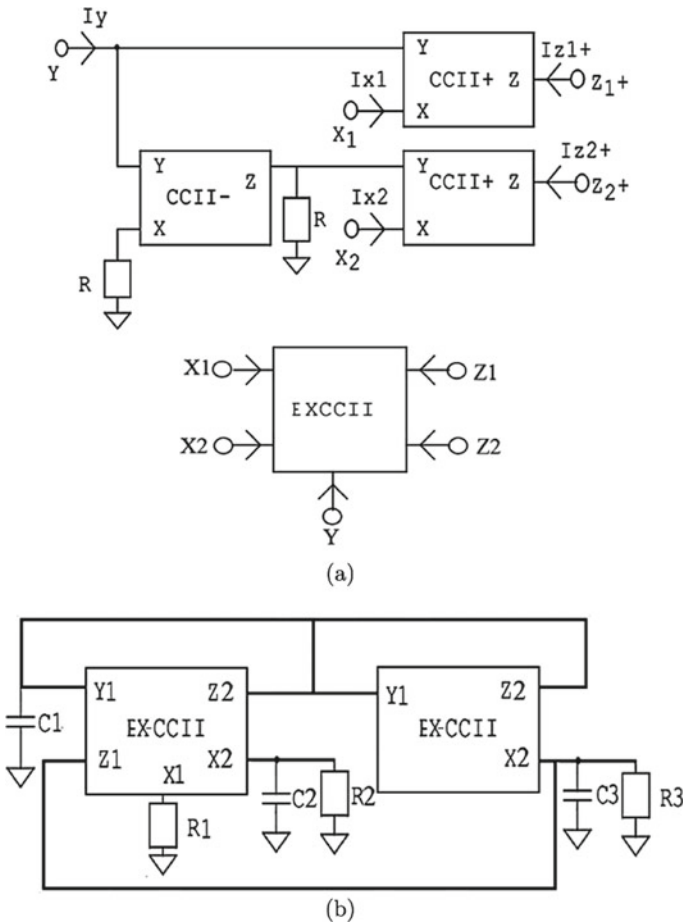
Fig. 3 Proposed **a** extra-X CCII using CCII and its symbol and **b** proposed ETHO

$$f = \frac{1}{\sqrt{C1C2R1R2 + C1C3R1R3 - R1R3C2C3 - C3C2R2 - C2C3R2R3 - R3C3^2}}$$
$$(3)$$

or

$$f = \sqrt{\frac{C1R1 - C2R1 - 2C2R2 - C3R3}{C1C2R3R2R1}}$$
$$(4)$$

And we choose the condition of oscillation as $C_3 = 2C_1$.

Fig. 4 Current output
waveforms at C_1, C_2 and C_3
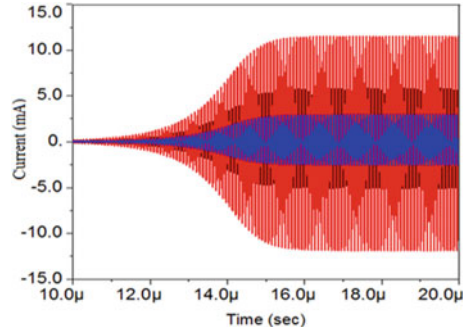
Fig. 5 Voltage output
waveforms at C_3
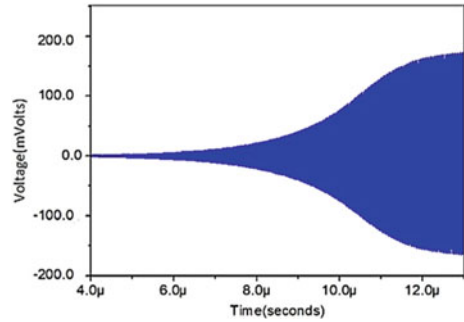
4   Simulation Results

The theoretical analysis of the proposed HO-ETO is well-supported by simulation
verification. The first step in the verification procedure is the PSPICE simulation,
which employs the TSMC 0.18 μm CMOS technology parameter with biasing V_{DD}
$= -V_{SS} = 1.25$ V. By choosing the values of the passive components ($R_1 = R_2 =$
40 kΩ, $R_3 = 90$ kΩ, $C_1 = C_2 = 2$ pF, $C_3 = 4$ pF), suggested simulation results are
based on the circuit shown in Fig. 3. The voltage (V_{C3}) and current waveforms (at
C_1, C_2 and C_3 terminal) produced by the proposed ETHO are depicted in Figs. 4 and
5, respectively. For the operating frequency in MHz, the gain-frequency plot (FFT)
is shown in Fig. 6.

Finally, the comparison of high-order electronically tunable oscillator (HO-ETO)
is performed with different oscillators which shows the following advantages of
proposed circuit as compared to [8, 11, 12, 17–19, 22, 24, 27] such as suitable for
high-frequency operation, electronically tunable, only three grounded capacitor and
resister are used as shown in Table 1.

Fig. 6 FFT: gain vs frequency plot
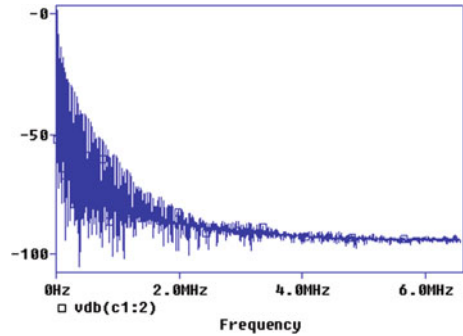
Table 1 Comparison with different high-order oscillator

Ref	Active block	Bias voltage	Passive elements	Electronic tuning	PD (mW)	Freq (Hz)
[8]	CCII-3	12 V	8	N	–	K
[27]	DVCC-3	2.5 V	6	N	–	M
[12]	CCCCTA-2	5 V	0	Y	–	K
[11]	CDTA-2	–	3	N	–	K
[17]	VDIBA-2	0.9 V	4	Y	–	K
[18]	CCII-2	1.25 V	6	Y	–	K
[19]	ZC-CG-CDBA-2	–	5	Y	–	M
[22]	CCDDCC-1	1.25 V	5	Y	1.32	M
[24]	DXCCTA-1	1.25 V	5	Y	1.44	M
Prop	EX-CCII-2	1.25 V	6	Y	1.01	M

5 Conclusion

An electronically tuned third-order oscillator is developed by using CCII-based EX-CCII, and its mathematical analysis is presented in this scientific work. The simulation results are supported by theoretical analyses. The simulation results are also demonstrated by using the CCII-based implementation of EX-CCII. In addition, PSPICE software is used to verify each suggested design. The theoretical analysis and simulation findings are in agreement. It is demonstrated that the suggested design can function roughly between KHz and MHz by tuning the passive components. The outcomes of the simulation demonstrate the suggested design's ability to function in both current and voltage modes.

References

1. Sedra A, Smith K (1970) A second generation current conveyor and its applications. IEEE Trans Circuit Theory. 17:132–134. https://doi.org/10.1109/TCT.1970.10830672
2. Joshi M, Ranjan R (2022) Low power chaotic oscillator employing CMOS. Integrat 85:57–62. https://doi.org/10.1016/j.vlsi.2022.02.011
3. Senani R et al (2016) Sinusoidal oscillators and waveform generators using modern electronic circuit building blocks. Springer, Cham, Switzerland. https://doi.org/10.1007/978-3-319-237 12-1
4. Joshi M, Ranjan A (2020) Current-controlled chaotic Chuas circuit using CCCII. Adv Comm Comp Tech 668:535–545. https://doi.org/10.1007/978-981-15-5341-741
5. Joshi M, Bhatt V, Ranjan A, Benjwal P (2017) Realization of Colpitts oscillator using second generation current controlled current conveyor. In: Proceedings of the Second International Conference on Research in Intelligent and Computing in Engineering (RICE-2017). PTI 49–52. https://doi.org/10.15439/2017R02
6. Siripruchyanun M, Jaikla W (2009) Cascadable current-mode biquad filter and quadrature oscillator using DO-CCCIIs and OTA. Circuit Syst Signal Process 28:99–110. https://doi.org/10.1007/s00034-008-9072-5
7. Khan IA, Khawaja S (2010) An integrable gm-C quadrature oscillator. Int J Elect 87:1353–1357. https://doi.org/10.1080/002072100750000150
8. Horng JW, Hou CL, Chang CM et al (2005) Quadrature oscillators using CCIIs. Int J Electron 92:21–31. https://doi.org/10.1080/00207210412331332899
9. Maheshwari S, Verma R (2012) Electronically tunable sinusoidal oscillator circuit. Act Passive Electron Compon 99(2012):1–6. https://doi.org/10.1155/2012/719376
10. Maheshwari S (2009) Current-mode third-order quadrature oscillator. IET Circuits Devices Syst 4:188–195. https://doi.org/10.1049/iet-cds.2009.0259
11. Horng JW (2009) Current-mode third-order quadrature oscillator using CDTAs. Act Passive Electron Compon 2009:1–5. https://doi.org/10.1155/2009/789171
12. Chiu W, Liu SI, Tsao HW, Chen JJ (1996) CMOS differential difference current conveyors and their applications, IEEE Proc Circuits Devices Syst 143:91–96. https://doi.org/10.1049/ip-cds:19960223
13. Elwan HO, Soliman AM (1997) Novel CMOS differential voltage current conveyor and its applications. IEEE Proc Circuits Devices Syst 144:195–200. https://doi.org/10.1049/ip-cds:19971081
14. Kumar A, Chaturvedi B (2018) Novel CMOS dual-X current conveyor transconductance amplifier realization with current-mode multifunction filter and quadrature oscillator. Circuit Syst Signal Process 37:2250–2277. https://doi.org/10.1007/s00034-017-0680-9
15. Lawanwisut S, Siripruchyanun M (2009) High output-impedance current-mode third-order quadrature oscillator based on CCCCTAs, TENCON. In: 2009—IEEE Region 10 Conference, pp 1–4. IEEE Press, Singapore. https://doi.org/10.1109/TENCON.2009.5395961
16. Lahiri A (2009) Novel voltage/current-mode quadrature oscillator using current differencing transconductance amplifier. Analog Integr Circ Signal Process 61:199–203. https://doi.org/10.1007/s10470-009-9291-0
17. Pushkar KL, Bhaskar DR (2019) Voltage-mode third-order quadrature sinusoidal oscillator using VDIBAs. Analog Integr Circ Sig Process 98:201–207. https://doi.org/10.4236/cs.2017.812021
18. Horng JW (2011) Current/voltage-mode third order quadrature oscillator employing two multiple outputs CCIIs and grounded capacitors. Indian J Pure Appl Phys 49:494–498
19. Biolek D, Lahiri A, Jaikla W et al (2011) Realization of electronically tunable voltage-mode/current-mode quadrature sinusoidal oscillator using ZC-CG-CDBA. Microelectron J 42:1116–1123. https://doi.org/10.1016/j.mejo.2011.07.004
20. Khaw-Ngam K, Kumngern M, Khateb F (2017) Mixed-mode third-order quadrature oscillator based on single MCCFTA. Radioengineering. 26:522–535. https://doi.org/10.13164/re.2017.0522

21. Zeki A, Toker A (2005) DXCCII-based tunable gyrator. Int J Elect Comm 59:59–62. https://doi.org/10.1016/j.aeue.2004.11.004
22. Yadav SK, Joshi M, Ranjan A (2020) A current tunable third-order oscillator using CCDDCC. In: Sharma DK, Balas VE, Son LH, Sharma R, Cengiz K (eds) Micro-electronics and telecommunication engineering. Lecture Notes in Networks and Systems. Springer, Singapore, p 106. https://doi.org/10.1007/978-981-15-2329-8-338
23. Joshi MK, Thakur P (2021) Design and implementation of third order oscillator using CFOA and OTA. In: 2021 International Conference on Electrical, Computer and Energy Technologies (ICECET), pp 1–5. IEEE Press. https://doi.org/10.1109/ICECET52533.2021.9698785
24. Laishram R, Joshi M, Ranjan A, A new high frequency MOS-C TOETQO employing DXCCTA. Int J Elect Lett. https://doi.org/10.1080/21681724.2021.2025435
25. Joshi M, Ranjan A (2000) An autonomous chaotic and hyperchaotic oscillator using OTRA. Analog Integr Circ Sig Process 101:401–413. https://doi.org/10.1007/s10470-019-01395-0
26. Singh KB, Joshi M, Ranjan A (2020) An MOS-C multifunction filter employing DXCCTA for high-frequency operation. In: Mallick PK, Meher P, Majumder A, Das SK (eds) Electronic systems and intelligent computing. Lecture Notes in Electrical Engineering, Springer, Singapore, p 686. https://doi.org/10.1007/978-981-15-7031-5 13
27. Chen HP, Hwang YS, Ku YT (2017) Voltage-mode and current-mode resistorless third-order quadrature oscillator. Appl Sci 7:179. https://doi.org/10.3390/app7020179
28. Maheshwari S (2013) Current conveyor all-pass sections: brief review and novel solution. Sci World J 2013:1–7. https://doi.org/10.1155/2013/429391

Analysis and Simulation of Misalignment Issues in Dynamic Wireless Charging for Electric Vehicles

Kundan Kumar and Ngangoiba Maisnam

1 Introduction

The emerging technologies in the energy and transportation sectors, such as electric vehicles, provide several advantages in terms of the economy and environment-friendly. The charging technology of an EV during the motion can be defined as dynamic wireless charging (DWC). With the help of DWC technology, various challenges may be resolved such as the volume of the battery and the inability to drive EVs over long distances [1]. The designing of transmitting coils and their configurations are very much important for the smooth charging of various types of electric vehicles maintaining higher efficiency [2].

The researchers have used various topologies to study the misalignment of the DWC system. A technique for constructing a transmitting coil utilizing a compensation network has been presented by S. Li et al. in consideration of transfer performance, electrical stress, and zero voltage switching (ZVS) [2]. The coupling coefficient and self-inductance calculations are presented by Sampath et al. using the finite element method (FEM). Furthermore, an optimization methodology against misalignment for DWC is discussed [3]. A three-dimensional analytical rectangular coil model has been proposed by Kushwaha et al. for computing mutual inductance for various misalignments [4]. This paper introduces a 3-D model of a circular helix coil for both transmitting (T_x) and receiving (R_x) coils. Both the T_x and R_x coils are placed in an air medium.

Further, analysis to get comparable findings in accordance with mutual inductance, magnetic flux, and coupling coefficient, k, is presented with various measurements along the y-axis (longitudinal distance) and z-axis (vertical distance), i.e., air-gap distance.

K. Kumar (✉) · N. Maisnam
National Institute of Technology Manipur, Imphal 795004, India
e-mail: kundankumar.nitm@gmail.com

In this paper, misalignment issues of DWC for an EV are analyzed and simulated. The overview of the DWC system and its working are described in Sect. 2. The significance of coupling pads and misalignment are discussed in Sect. 3. The simulation analysis and various results for modeled coil structures are presented in Sect. 4. The summary of the paper is concluded in Sect. 5.

2 Overview of DWC and Its Working

The DWC system is the present demand for EV technology which strengthens the charging infrastructures. In the DWC system, the vehicle battery can be charged even though it is in motion. It is important to give attention to a multi-objective design strategy for the DWC system that takes into account both the coupling coils and the compensation networks. When designing a DWC system, the LCC network is employed as a compensation network to achieve constant output current and zero voltage switching (ZVS) conditions [2].

Figure 1 exhibits the DWC system's schematic representation. N transmitting coils comprise the system, and each coil has dimensions of length (L_x) and width (B_x) (whereas $x = 1, 2,\ldots N$). The W is the distance between two consecutive transmitting coils. L is the overall length of N coils including distances between consecutive transmitting coils. There is a single receiving coil placed on the vehicle denoted by k which is shown in Fig. 1 by the red periphery. The LCC compensation network is used to account for the reactive components of a single transmitting coil. To adjust the power-supplying coils, N number of switches ($S_1, S_2,\ldots S_N$) are connected with the LCC networks and the high-frequency inverters. The detecting devices, i.e., sensors, are installed beside each transmitting coil to locate the exact position of the receiving coil. The activation and deactivation of each transmitter coil can be managed with the help of artificial intelligence (AI) techniques. However, the detection and sensing control are not included in this work.
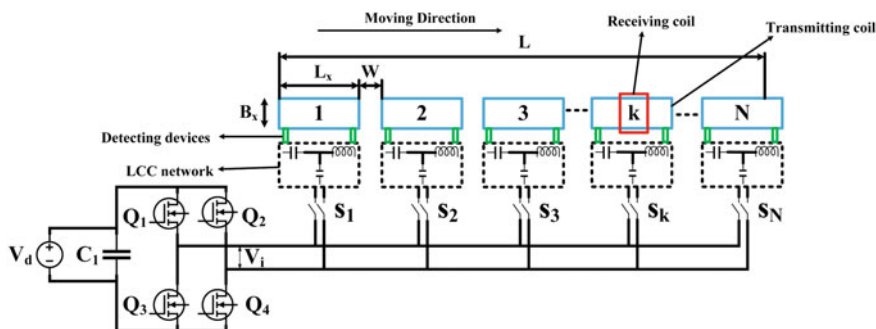
Fig. 1 Dynamic wireless charging (DWC) system

For simplicity, two transmitting coils (T_x) and one receiving coil (R_x) system are considered in this study and the circuit diagram is given in Fig. 2. The LCC network is employed in Fig. 2, where L_{mT_1} and R_1 are the transmitting coil 1's inductance and resistance, respectively, whereas for the transmitting coil 2 inductance and resistance can be termed as L_{mT_2}, and R_2, respectively. On the other hand, the inductance and resistance of the receiving coil are represented by L_R and R_R, respectively. The input impedances of transmitting coils 1 and 2 are denoted by Z_1 and Z_2, respectively, while for the receiving coil Z_R. The load impedance of receiving side is denoted by Z_L. The different values of the LCC network can be estimated by the following formula.

$$L_{m_{T\_n}} = \frac{X_T}{\omega} \tag{1}$$

$$C_{TT\_n} = \frac{1}{\frac{\omega}{X_T}} \tag{2}$$

$$C_{TR\_n} = \frac{1}{\frac{\omega}{(\omega L_{T\_n} - X_T)}}, \tag{3}$$

Where ω is the angular frequency, X_T is the equivalent impedance of the transmitting coil, $L_{m_{T\_n}}$ is the self-inductance of the transmitting coil (where $n = 1$ and 2), C_{TT_n} and C_{TR_n} is the capacitance of the transmitting coil (where $n = 1$ and 2). The high-frequency inverter (HFI) output voltage is denoted by V_i which is fed to transmitting coils and the transmitting current by I_n (where $n = 1$ and 2), then X_T is given by

$$X_T = \frac{V_i}{I_n} \tag{4}$$
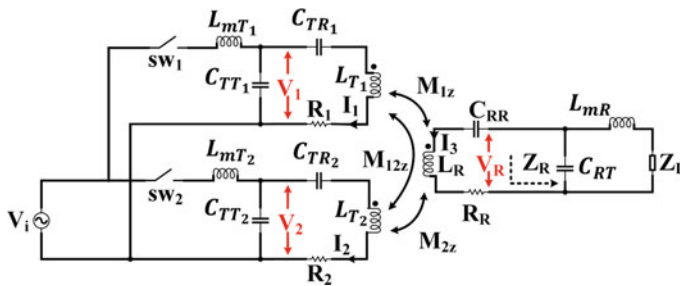
Fig. 2 Circuit diagram of the DWC system

2.1 Operation of the DWC System

The HFI is connected to both the transmitting coils, and activation of both the trans-
mitting coils is maintained by turning on and off the switches sw_1 and sw_2 [2, 3].
The operation of the segmented DWC system is categorized into three stages which
are given as follows.

STAGE 1: $sw_1 = ON$, $sw_2 = OFF$.

In this stage, the system's KVL equation is expressed in matrix form as given by,

$$\begin{bmatrix} Z_1 & j\omega M_{12z} & j\omega M_{1z} \\ j\omega M_{12z} & Z_2 - jX_T & j\omega M_{2z} \\ j\omega M_{1z} & j\omega M_{2z} & Z_R \end{bmatrix} \begin{bmatrix} I_1 \\ I_2 \\ I_3 \end{bmatrix} = \begin{bmatrix} V_1 \\ 0 \\ 0 \end{bmatrix}, \tag{5}$$

where the mutual inductances M_{1z} and M_{2z} are those between T_x coil 1 and R_x coil
and between T_x coil 2 and R_x coil, respectively. The mutual inductance of both the
T_x coils is M_{12z}. The currents flowing through the T_x and R_x coils are

$$I_1 = \frac{V_s}{jX_T} \tag{6}$$

$$I_2 = -I_1 . \frac{\omega^2 M_{12z} M_{2z} + j\omega M_{12z} Z_R}{\omega^2 M_{2z}^2 + (Z_2 - jX_T) Z_R} \tag{7}$$

$$I_3 = \frac{j\omega(M_{1z} I_1 + M_{2z} I_2)}{Z_R}. \tag{8}$$

The voltage induced in the receiving coil is given by

$$V_R = \omega M_{1z} I_1 + \omega M_{2z} I_2 \tag{9}$$

STAGE 2: $sw_1 = ON$, $sw_2 = ON$.

In this stage, the system's KVL equation is expressed in matrix form as given by,

$$\begin{bmatrix} Z_1 & j\omega M_{12z} & j\omega M_{1z} \\ j\omega M_{12z} & Z_2 & j\omega M_{2z} \\ j\omega M_{1z} & j\omega M_{2z} & Z_R \end{bmatrix} \begin{bmatrix} I_1 \\ I_2 \\ I_3 \end{bmatrix} = \begin{bmatrix} V_1 \\ V_2 \\ 0 \end{bmatrix}. \tag{10}$$

The currents of transmitting and receiving coils are given by

$$I_1 = \frac{V_s}{jX_T} \tag{11}$$

$$I_2 = \frac{V_s}{jX_T} \tag{12}$$

$$I_3 = \frac{j\omega(M_{1z}I_1 + M_{2z}I_2)}{Z_R} \tag{13}$$

The receiving coil's induced voltage can be represented by (9).

STAGE 3: $sw_1 = $ OFF, $sw_2 = $ ON.

Similarly, for the third stage, the above analysis can be carried out by exchanging the subscripts "1" and "2."

3 Coupling Pad and Misalignment

The most crucial component of the DWC system is the coupling pads. For coupling pads, it is important to have high-quality factor Q, high misalignment tolerance, and high coupling coefficient k [5–10].

The above-said parameters depend on the inductor's shape, core material, and spacing between the coils (i.e., T_x and R_x coils) [6–8]. The efficiency of linked inductors in a DWC system depends upon the inductor's quality factor and magnetic coupling coefficient. The geometric average of the transmitter and receiver quality factor denoted as Q is given as:

$$Q = \sqrt{Q_T Q_R} \tag{14}$$

$$Q_T = \frac{\omega L_T}{R_T} \tag{15}$$

$$Q_R = \frac{\omega L_R}{R_R}, \tag{16}$$

where Q_T and Q_R represent the transmitting and receiving coil quality factors, respectively. The equivalent inductance of the T_x and R_x coils is L_T and L_R, respectively (where $T = 1$ and 2). The equivalent resistance of the T_x and R_x coils is termed R_T and R_R, respectively (where $T = 1$ and 2).

It is necessary to design the inductor with a high level of self-inductance, low series resistance, and a relatively high frequency to get a high Q [9]. So, lowering the coil's resistance is necessary to raise Q. The value of coil self-inductance can be computed by (17) [7].

$$L = \frac{\mu N^2 A}{l}, \tag{17}$$

where l is the mean length, N is the number of turns, A is the conductor's cross-sectional area, and μ represents the flux path's magnetic permeability.

3.1 Circular Helix Coil

The parameters considered for designing the DWC system are shown in Table 1. Figure 3 shows the model representation of the circular helix coil, where the coil's dimensions are shown in Fig. 3a, and a 3-D representation of the coil's circular helix is presented in Fig. 3b. The transmitter (T_x) coil is presented in blue color, and the receiver (R_x) coil is in red color that is physically spaced by an air gap. The transmitter coil is energized by the HFI current, whereas the receiver coil is not connected to any active sources. Power is transferred in the DWC system via a magnetic field that the transmitting coil generates and couples to the receiving coil via an air medium. A magnetic connection between coils is necessary for the DWC system to transfer the power maintaining optimum efficiency [11–13]. The orientation and the flux linkage of the coils affect the performance of the DWC system [14].

	Sl. no.	Parameters	Values	Unit
Table 1 Design specification of circular helix coil	1	Wire radius	3	mm
	2	Start helix coil radius	40	mm
	3	No. of turns	3	
	4	Current	5	A
	5	Coil pitch	8.01	
	6	Terminal coil diameter	2	mm
	7	Terminal coil length	500	mm
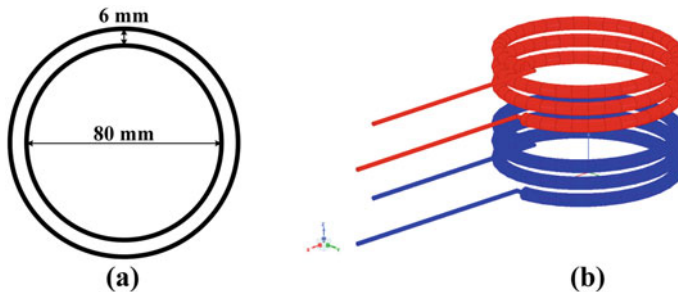
Fig. 3 Model representation **a** dimension of circular helix coil, and **b** 3-D geometry of circular helix coil
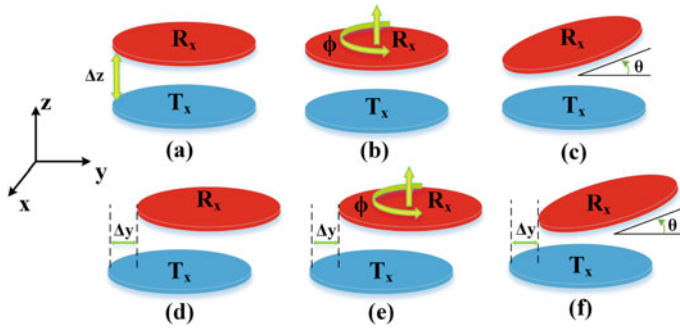
Fig. 4 Various types of misalignments **a** vertical, **b** planar, **c** angular, **d** longitudinal, **e** planar and longitudinal, and **f** angular and longitudinal [14]

3.2 Misalignment

The various types of misalignment are shown in Fig. 4. It can be observed that due to different diverse positions of the receiving (R_x) coil, different misalignments can categorize such as vertical, planar, angular, longitudinal, planar and longitudinal, and angular and longitudinal [4]. If the receiving coil moves in the vertical direction, then the deviation is represented by Δz and is called vertical misalignment as shown in Fig. 4a. In case of planer misalignment, the receiving coil is rotated with an angle ϕ (i.e. $0 < \phi \le 2\pi$) parallel to the transmitting coil at a fixed center with constant height as shown in Fig. 4b. In Fig. 4c, the receiving coil deviates at an angle θ (where $0 < \theta \le \pi$) by maintaining the same center point by both the T_x and R_x coils without any vertical misalignment which is named as angular misalignment [15–17]. When the receiving coil travels along the longitudinal direction, several sorts of variations are conceivable. In this kind of misalignment, the T_x and R_x coil's centers are displaced by Δy in the y-direction, as shown in Figs. 4d–f. The receiving coil is misaligned longitudinally in Fig. 4d when it is shifted in the y-axis with a constant height. The receiving coil in Fig. 4e is shifted in the y-axis and rotated at a fixed height, which is referred to as planar and longitudinal misalignment. In the case of angular and longitudinal misalignment, the receiving coil is deviated by an angle θ and displaced in the y-axis as shown in Fig. 4f [16].

4 Simulation Analysis and Result Discussion

The simulation studies are carried out by considering one transmitter coil as well as one receiver coil to observe the misalignment of the DWC system. ANSYS MAXWELL 3-D is used to validate the specified analytical model. Both the coils (i.e., T_x and R_x) have been simulated with the consideration of an air-gap distance of
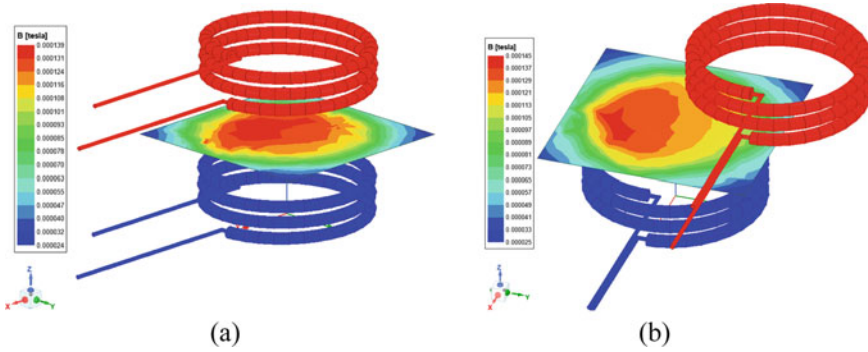
Fig. 5 Magnetic field density between T_x and R_x with the variation of **a** 0 mm y-axis and 65 mm z-axis. **b** 50 mm y-axis and 65 mm z-axis

65 mm on the z-axis (vertical) while in the y-axis (longitudinal) distance is considered as 0 and 50 mm which are displayed in Fig. 5a, b, respectively. Figure 5 also depicts the magnetic field density between the transmitter and receiver coils. The highest magnetic density is observed in the heated region, which is represented by the red zone in Fig. 5a, with a value of 0.000139 T (Tesla) at an air-gap (vertical) distance of 65 mm in the z-axis and no fluctuation in the y-axis, while the lowest is observed in the blue zone, with a value of 0.000024 T (Tesla). Likewise, the hot region in Fig. 5b exhibits the highest magnetic density at an air-gap (vertical) distance of 65 mm on the z-axis and a misalignment of 50 mm on the y-axis, with a value of 0.000145 T (Tesla) and the value of the lowest zone, blue, is 0.000025 T (Tesla). Examining Fig. 5a, b shows a difference of 0.000006 T.

Further, in the event of a vertical variation, the receiving coil shifts along the z-axis, while they move along the y-axis when there is a longitudinal misalignment. By changing the z-axis and y-axis distances, several values are simulated in accordance with the mutual inductance, flux linkage, magnetic field density, and coupling coefficient.

The parameters of the considered analytical model are used to observe the coupling coefficient, mutual inductance, and magnetic flux density of different variations in the y-axis (longitudinal distance) and z-axis (vertical distance) [18, 19]. The coupling coefficient k of transmitting coil input (T_{x_in}) and receiving coil input (R_{x_in}) is shown in Fig. 6. As the distance between the coils rises, the value of k decreases. Multiple results were computed for different variations in the y-axis for a significant distance in the z-axis. The coupling coefficient is shown by the red curve at an air-gap (vertical) distance of 35 mm for the variations in the y-axis, the green curve at an air-gap (vertical) distance of 45 mm, the blue curve at an air-gap (vertical) distance of 55 mm, and the orange curve at an air-gap (vertical) distance of 65 mm.

It is observed that the self-induced flux on the transmitting coil is constant while the flux linkage is varying along with the movement of receiving coil as shown in Fig. 7.
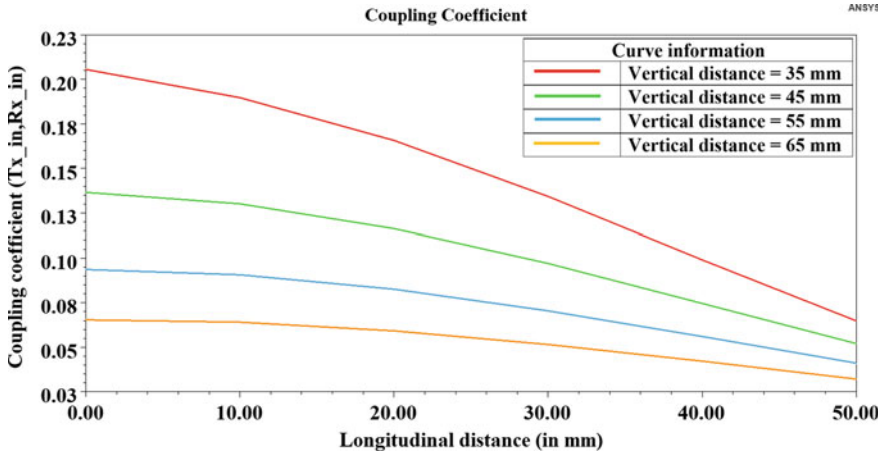
Fig. 6 Coupling coefficient of T_{x_in} and R_{x_in}
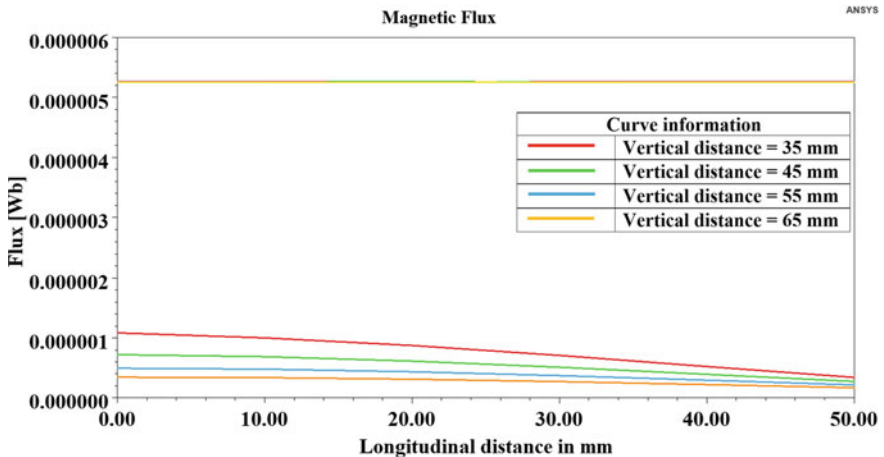
Fig. 7 Flux linkage in T_x and R_x coils

The red curve indicates the flux in the receiving coil at a 35 mm air-gap (vertical) distance and varies along the longitudinal axis. The green curve, which varies along the longitudinal axis, depicts the flux in the receiving coil at a 45 mm air-gap (vertical) distance. The blue curve, which varies along the longitudinal axis, depicts the flux in the receiving coil at a 55 mm vertical (air-gap) distance. The orange curve, with changes in the longitudinal axis, depicts the flux in the receiving coil at a 65 mm vertical (air-gap) distance. In Table 2, different values of mutual inductance between the T_x and R_x coils are surmised for divergence in the receiving coil. The mutual inductance value is decreasing as the R_x coils travel away from the T_x coil [20–23]. The visual representation of the mutual inductance between the T_x and R_x coils is

Table 2 Mutual inductance between T_x and R_x coil

Sl. no.	y-axis [mm]	M [nH] (for $z =$ 35 mm)	M [nH] (for $z =$ 45 mm)	M [nH] (for $z =$ 55 mm)	M [nH] (for $z =$ 65 mm)
1	0.00	216.84	144.20	98.89	69.02
2	10.00	200.16	193.55	95.87	67.63
3	20.00	174.81	122.75	87.25	62.51
4	30.00	141.73	102.34	74.39	54.41
5	40.00	104.60	78.66	59.19	44.52
6	50.00	68.33	54.88	43.36	33.93

y-axis: longitudinal distance, z-axis: air-gap distance (i.e. vertical distance), and M: mutual inductance between T_x and R_x
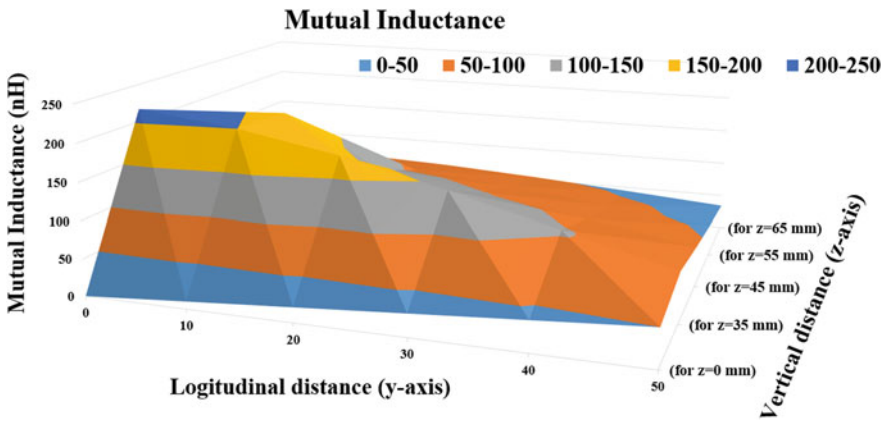
Fig. 8 Graphical view of mutual inductance

shown in Fig. 8. The graph clearly demonstrates that the mutual inductance values reflect the distance between the coils.

5   Conclusion

In this paper, the misalignment issues that arise in dynamic wireless charging (DWC) are investigated in detail. The impact of the DWC system and its working with the consideration of coupling coefficients, flux linkage, as well as mutual inductance are discussed. A circular helix coil is considered for analyzing the DWC system, and hence, an ANSYS MAXWELL 3-D model is created. The copper material is chosen for both the coils (T_x and R_x) and placed in an air medium for the simulation modeling. In the simulation studies, different misalignments are considered by varying the distances along the longitudinal and vertical, i.e., y-axis and z-axis, respectively,

and the variant in the mutual inductance, coupling coefficient, flux linkage, and distribution of magnetic flux density is observed. It is observed that the increase in deviation from the proper alignment leads to a decrease in mutual inductance, coupling coefficient as well flux linkage.

References

1. Patil D, McDonough MK, Miller JM, Fahimi B, Balsara PT (2018) Wireless power transfer for vehicular applications: overview and challenges. IEEE Trans Transp Elect 4(1):3–37
2. Li S, Wang L, Tao C, Li F, Wang L (2019) Designing of the transmitting coils and compensation network of a segmented DWPT system. In: 2019 IEEE 15th Brazilian Power Electronics Conference and 5th IEEE Southern Power Electronics Conference (COBEP/SPEC), pp 1–5
3. Sampath JPK, Alphones A, Vilathgamuwa DM (2016) Coil optimization against misalignment for wireless power transfer. In: 2016 IEEE 2nd Annual Southern Power Electronics Conference (SPEC), pp 1–5
4. Kushwaha BK, Rituraj G, Kumar P (2017) 3-D analytical model for computation of mutual inductance for different misalignments with shielding in wireless power transfer system. IEEE Trans Transp Elect 3(2):332–342
5. Zaheer A, Covic GA, Kacprzak D (2014) A bipolar pad in a 10-kHz 300-W distributed IPT system for AGV applications. IEEE Trans Ind Electron 61(7):3288–3301
6. Covic GA, Kissin MLG, Kacprzak D, Clausen N, Hao H (2011) A bipolar primary pad topology for EV stationary charging and highway power by inductive coupling. In: Proc IEEE Energy Convers Congr Expo, Phoenix, AZ, USA, pp 1832–1838
7. Kim S, Covic GA, Boys JT (2017) Tripolar pad for inductive power transfer systems for EV charging. IEEE Trans Power Electron 32(7):5045–5057
8. Chowdary KV, Kumar K, Behera RK, Banerjee S (2020) Overview and analysis of various coil structures for dynamic wireless charging of electric vehicles. In: 2020 IEEE International Conference on Power Electronics, Smart Grid and Renewable Energy (PESGRE2020), pp 1–6
9. Kiani M, Jow U-M, Ghovanloo M (2011) Design and optimization of a 3-coil inductive link for efficient wireless power transmission. IEEE Trans Biomed Circuits Syst 5(6):579–591
10. Moon S, Kim B-C, Cho S-Y, Ahn C-H, Moon G-W (2014) Analysis and design of a wireless power transfer system with an intermediate coil for high efficiency. IEEE Trans Ind Electron 61(11):5861–5870
11. Lin FY, Carretero C, Covic C, Boys J (2016) Reduced order modelling of the coupling factor for varying sized pads used in wireless power transfer. IEEE Trans Transport Elect 99:1
12. Fotopoulou K, Flynn BW (2011) Wireless power transfer in loosely coupled links: Coil misalignment model. IEEE Trans Magn 47(2):416–430
13. Diep NT, Trung NK, Minh TT (2019) Power control in the dynamic wireless charging of electric vehicles. In: 2019 10th International Conference on Power Electronics and ECCE Asia (ICPE 2019 - ECCE Asia), pp 1–6
14. Abou Houran M, Yang X, Chen W (2018) Magnetically coupled resonance WPT: review of compensation topologies, resonator structures with misalignment, and EMI diagnostics. Electronics 7:296
15. Ramezani A, Narimani M (2019) High misalignment tolerant wireless charger designs for EV applications. IEEE Transp Elect Conf Expo (ITEC) 2019:1–5
16. Shin Y, Hwang K, Park J, Kim D, Ahn S (2019) Precise vehicle location detection method using a wireless power transfer (WPT) system. IEEE Trans Veh Technol 68(2):1167–1177
17. Ong A, Jayathuathnage PKS, Cheong JH, Goh WL (2017) Transmitter pulsation control for dynamic wireless power transfer systems. IEEE Trans Trans Elect 3(2):418–426
18. Razu MRR et al (2021) Wireless charging of electric vehicle while driving. IEEE Access 9:157973–157983

19. Triviño A, Sánchez J, Delgado A (2022) Efficient methodology of the coil design for a dynamic wireless charger. IEEE Access 10:83368–83378
20. Chowdary KVVSR, Kumar K (2022) Assessment of dynamic wireless charging system with the variation in mutual inductance. In: Proc IEEE Indian Council International conference (INDICON-2022), Kochi, India, 24–26 November, pp 1–6
21. Kumar K, Chowdary KVVSR, Nayak BK, Mali V (2022) Performance evaluation of dynamic wireless charging system with the speed of electric vehicles. In: Proc IEEE Indian Council International conference (INDICON-2022), Kochi, India, 24–26 November, pp 1–6
22. Zhang X, Meng H, Wei B, Wang S, Yang Q (2019) Mutual inductance calculation for coils with misalignment in wireless power transfer. J Eng 2019:1041–1044
23. Kumar K, Chowdary KVVSR, Mali V, Kumar RR (2021) Analysis of output power variation in dynamic wireless charging system for electric vehicles. In: Proc of IEEE International Conference on Smart Technologies for Power, Energy and Control (STPEC 2021), Bilaspur, India, 19–22 December, pp 1–6

An Asymmetric CPW-Fed Compact Dual-Band Circularly Polarized Monopole Antenna

Mehtaz Marin, Reshmi Dhara, and Sanoj Mahato

1 Introduction

Today's wireless communication systems use dual-band dual-polarized (DBDP) antennas to meet the diverse requirements of various devices. Due to the enormous data capacity and high-resolution requirements for current wireless communication, circular polarization (CP) is a very desirable property of antennas. To generate the circular polarization, it is needed to maintain two orthogonal electric fields that should be equal in magnitude and 90° phase difference among them. Circular polarized antenna is more popular nowadays in wireless communication systems because it is not necessary to arrange the orientation of the transmitting and receiving antenna. Additionally, circular polarized antenna overcomes multipath fading, enhances system performance and can also provide better mobility with weather penetration compared to linearly polarized antenna. As a result, compared to dual-band antennas that simultaneously operate at two different frequency bands in dual circular polarization, dual-band antennas with two linear polarizations are much less prevalent. It is not necessary to use as many antennas when using a multiband antenna that also performs dual-polarization performance. Various researchers [1–5] have investigated various DBDP antennas that generate LHCP and RHCP polarization.

An asymmetric CPW-fed compact dual-band circularly polarized monopole antenna is proposed in this paper influenced by the above-mentioned antenna characteristics. The proposed antenna has DBDP with LHCP over both lower- and higher-frequency regions. The simulated antenna can accommodate 6.22–8.29 GHz and

M. Marin · R. Dhara (✉) · S. Mahato

Department of Electronics and Communication Engineering, National Institute of Technology Sikkim, Ravangla, India

e-mail: reshmidhara@nitsikkim.ac.in

15.99–18.58 GHz. The design has been adequately enhanced in such a manner so as to create both circular polarizations.

2 Design Procedure of Antenna

Figure 1 depicts the suggested simulated dual-band CPW antenna's design that shows the top and bottom views of the antenna. The proposed antenna's dimensions are 32.62 × 22.69 mm^2. Rogers RO4003C substrate is used in implementing the antenna having a thickness of $h = 0.813$, $\varepsilon_r = 3.38$ and loss tangent $\tan\delta = 0.0027$. The final geometrical parameters after optimization appear in Table 1. The different design progression of the designed antenna is shown in Fig. 2. Figure 3a, b shows the comparative graphs of the return loss and the ARBW of the various antenna stages. All three antennas are operating at a frequency of 3.5 GHz.

In the first step indicated here as Antenna 1, a straight microstrip line monopole antenna is designed with the length of approximately half wavelength with a full ground as depicted in Fig. 2. The straight microstrip radiating patch at the centre is excited using 50 Ω microstrip feed-line on the top of the substrate to produce a CP at this resonant frequency. However, Fig. 3 shows that the impedance matching and ARBW are quite weak at this stage. Therefore, it requires more current paths and
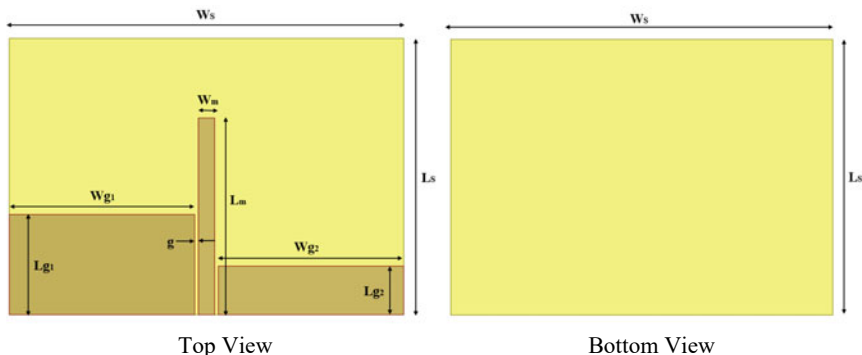
Fig. 1 Dimensions of the antenna

Table 1 Dimensions of designed simulated antenna

Parameters	Value (mm)	Parameters	Value (mm)
h	0.813	Lg$_1$	8
g	0.3	Wg$_2$	15.32
Ws	32.62	Lg$_2$	5.93
Ls	22.69	Wm	1.35
Wg$_1$	15.32	Lm	16.17
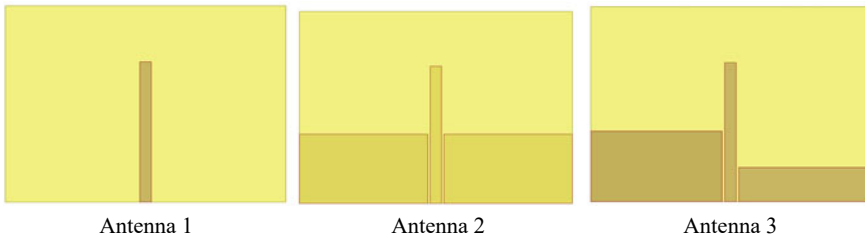
Fig. 2   Different stages of the proposed antenna
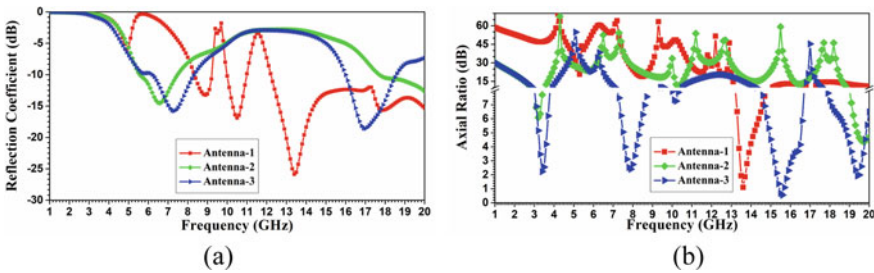
Fig. 3   Comparisons of **a** IB and **b** ARBW

coupling to meet multi-band features [6]. So, to attain better coupling two symmetric rectangular CPW-fed grounds are designed [7] at both the sides of the microstrip feed and the gap between the ground and the patch is *g = 0.3* that gives better IBW and ARBW compared to the earlier stage. Since the generated ARBW doesn't satisfy the 3 dB criteria, therefore to improve better performance of the ARBW, the third stage is designed. In the third stage, the two asymmetric ground plane [8] is used which gives two IBW and two CP bandwidth within 6.22–8.29 GHZ and 15.99–18.58 GHz IBW. Since it achieved our goals, it was clear to choose, finalize and analyse this design's functionality.

3   Results and Discussions

Ansys high-frequency structure simulator (HFSS) was used to design the antenna in order to generate an electromagnetic simulation of the suggested antenna and to compare the findings. The proposed antenna generates dual impedance bandwidth of 6.22–8.29 GHZ and 15.99–18.58 GHz as shown in Fig. 4a which are resonating at 7.25 and 17.28 GHz at a percentage of 28.55% and 14.98%. The dual circularly polarized ARBW ranges from 7.68 to 8.08 GHz and 15.11 to 16.19 GHz as shown in Fig. 4b which are resonating at 7.84 and 15.65 GHz at a percentage of 6.12% and 6.90%.
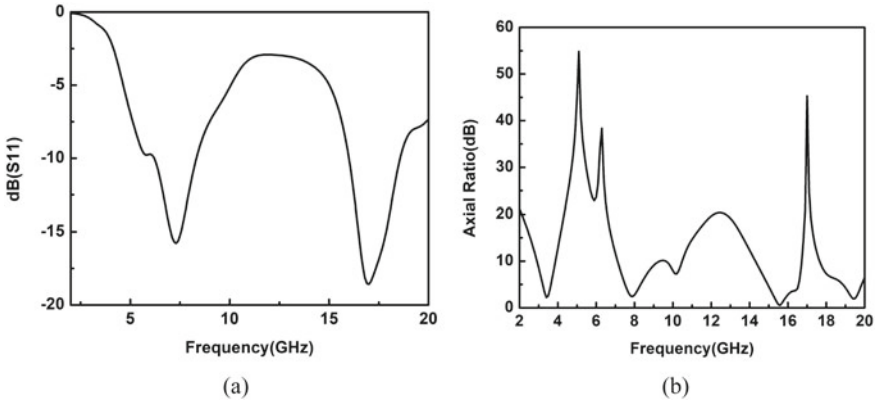
Fig. 4 **a** S_{11} and **b** ARBW of the simulated antenna
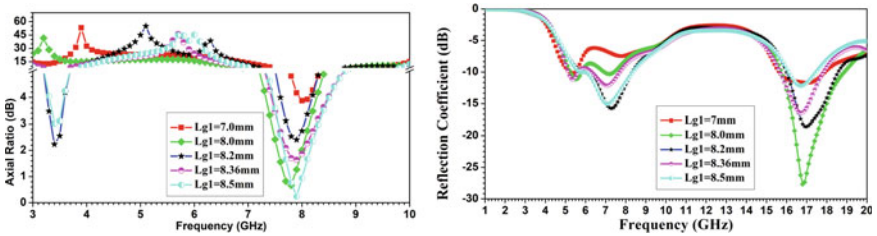
Fig. 5 Variations of Lg1

3.1 Parametric Optimization Results

3.1.1 Variation of Lg$_1$

The impact of length of the leftmost CPW-fed ground (Lg$_1$) is observed to have small effect on S_{11} and AR. It is observed from Fig. 5 that Lg$_1$ is responsible for creating the lower CP band.

3.1.2 Variation of Lg$_2$

The impact of length of the rightmost CPW-fed ground (Lg$_2$) is observed to have small effect on S_{11} and AR. It is observed from Fig. 6 that Lg$_2$ is responsible for creating the higher CP band.

Figure 7 illustrates a quantitative observation that helps explain how the two CP modes at 7.8 GHz and 15.6 GHz are generated. According to the normalized currents distribution shown in the above figures, dual CP modes are achieved at fcp$_1$
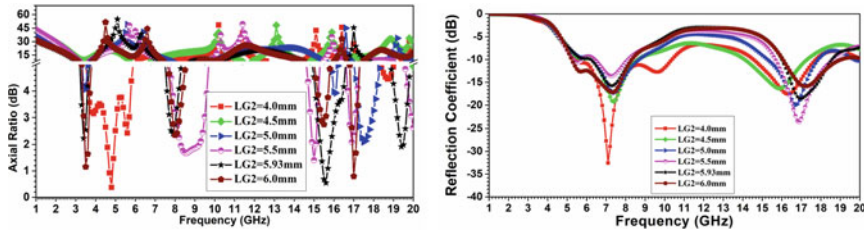
Fig. 6 Variations of Lg2

= 7.8 GHz and fcp$_2$ = 15.6 GHz, which are both LHCP, at four separate points in time at $t = 0$, $t = T/4$, $t = 3\,T/4$ and $t = T$.

Figure 8a, b is radiating at fc$_1$ = 7.8 GHz, and Fig. 8c, d is radiating at fc$_2$ = 15.6 GHz. Therefore, it is evident from the below-mentioned figures that both of the CP radiation patterns are LHCP. Both of the radiations that were detected at two CP resonating frequencies are LHCP, with polarization disparities of 17 dBi and 30 dBi.

Figure 9a, b, which examine the radiation properties of the antenna, shows the three-dimensional directivity pattern distributions for two CP resonant frequencies labelled as fc$_1$ = 7.8 GHz and fc$_2$ = 15.6 GHz, respectively. Figure 8a, b shows that the x- and y-directed currents are nearly equal in amplitude and 90° phase at fc$_1$ = 7.8 GHz and fc$_2$ = 15.6 GHz, resulting in circular polarization. Thus, it can be said that the antenna emits circular polarization radiation at fcp$_1$ and fcp$_2$.

When real, the simulated input impedance corresponds to a 50 Ω microstrip feed-line for resistance and reactance when it is imaginary The IB match of the impedance

(a) t=0, t=T/4, t=3T/4 and t=T

(b) t=0, t=T/4, t=3T/4 and t=T
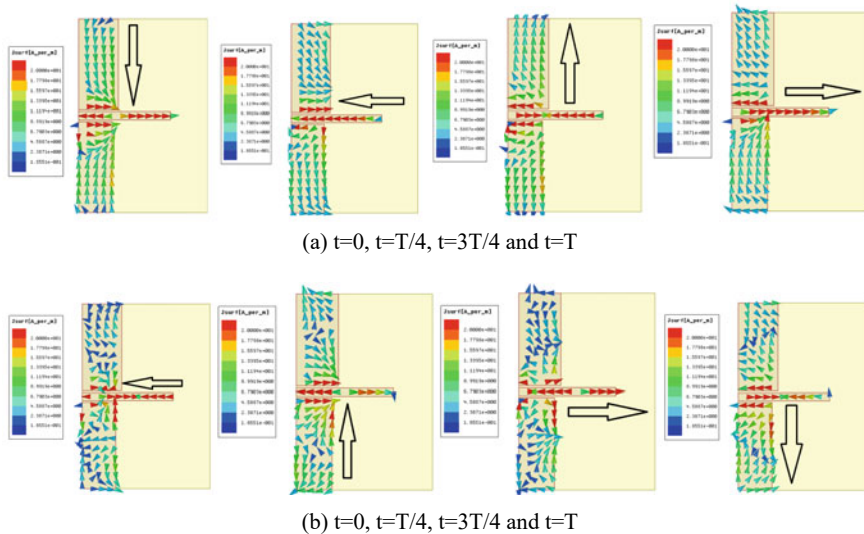
Fig. 7 Current distribution simulation at **a** fc$_1$ = 7.8 GHz, **b** fc$_2$ = 15.6 GHz

Frequency = 7.8 GHz   Frequency = 15.6 GHz
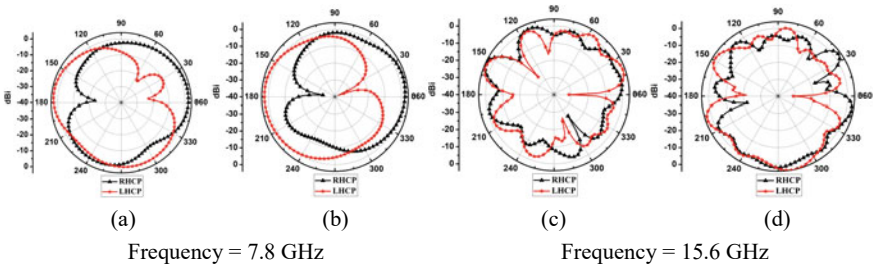
Fig. 8 LHCP and RHCP radiation patterns for **a, c** φ = 0° and **b, d** φ = 90° planes
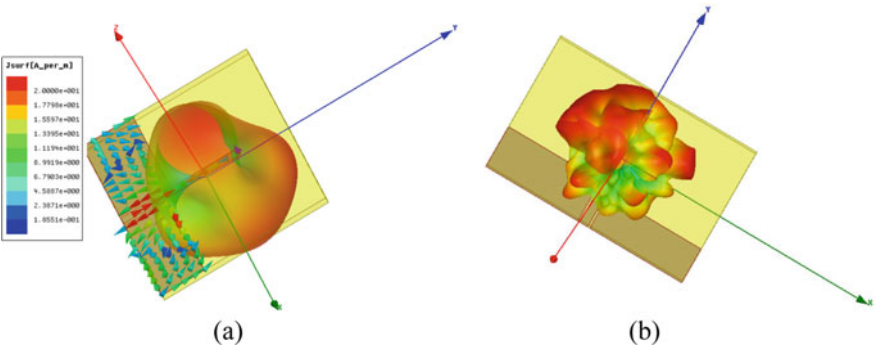
Fig. 9 Total 3D directivity patterns of the simulated antenna **a** $fc_1 = 7.8$ GHz and **b** $fc_2 = 15.6$ GHz

is excellent as shown in Fig. 10 because the resistance component of the impedance is closer to 50Ω and the reactance part is closer to 0 Ω. The normalized impedance (Z11) values for resonance frequencies of 7.3 GHz and 17 GHz are near to 1 as shown in Fig. 11, and the values for the complex reflection coefficient magnitude are likewise quite small. As a result, the above resonance frequencies show best matching as the real portion of Z11 becomes closer to 50 Ω, and the imaginary component approaches to 0 Ω in this case.

The peak gain of the simulated CP antenna is shown in Fig. 12. The above figure shows that the simulated antenna's gain is at its highest at 15.6 GHz or 10.7 dBi.

Figure 13 shows that inside the two circular polarized bands, the Ex/Ey magnitude is roughly identical to 1 dB or 0 dB, and that the phase difference between the two bands is also nearly 90°. Thus, it can be seen that these dual bands satisfy the requirements of circular polarization.

Table 2 shows the comparison related to different DBDP antennas with the simulated antenna. It can be observed from the above table that the IBW and ARBW of the simulated antenna are very good compared to the other referred antennas. Additionally, the simulated antenna is substantially smaller than the other referred antennas in terms of size.
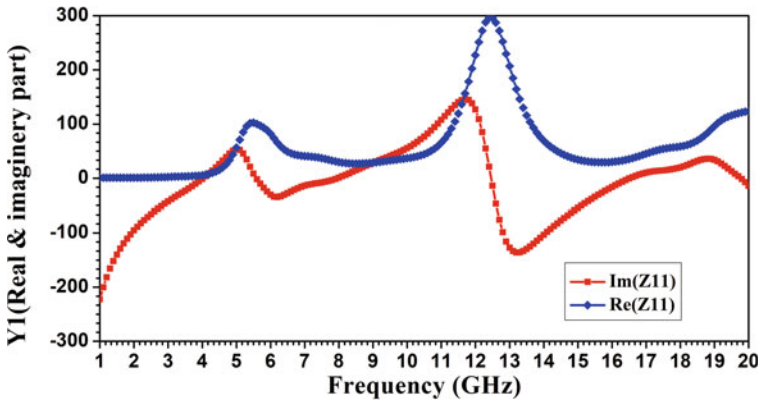
Fig. 10 Real and imaginary part of for the antenna impedance
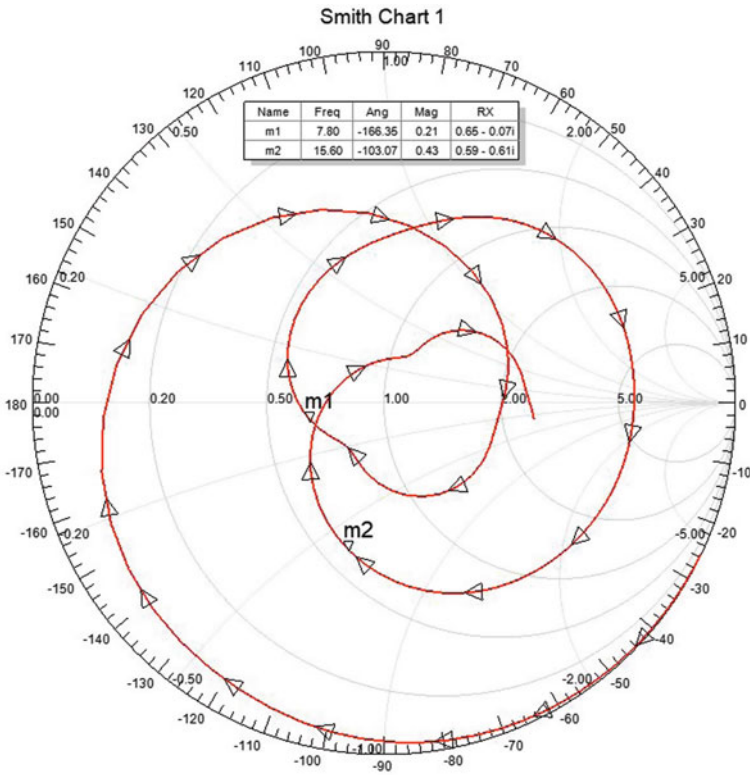
Fig. 11 Smith chart for the implemented antenna
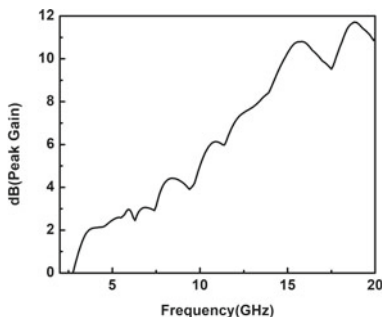
Fig. 12 Peak gain of the simulated antenna
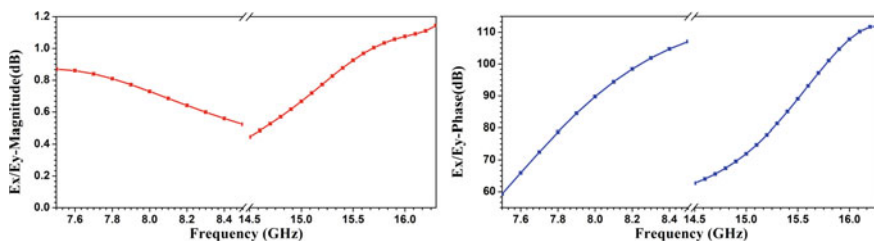
Fig. 13 E_X/E_Y ratio of magnitude and phase plots of the simulated antenna

Table 2 Comparison related to DBDP antenna with the simulated antenna

Reference	Year	IBW (MHz)	ARBW (MHz)	Size (mm^2)
[1]	2014	339.1932	250.360	40 × 45
[2]	2018	240.2730	260.2480	43 × 49
[3]	2019	48.150	50.170	50 × 50
[4]	2020	90.270	62.93	77 × 103
[5]	2021	224.420	132.770	70 × 70
Proposed work	2022	2070.2590	480.1080	32.62 × 22.69

4   Conclusion

An antenna is designed here that shows dual-band polarization with dual circular polarization characteristics with LHCP at both higher- and lower-frequency regions. The proposed antenna is of the size 32.62 × 22.69 mm^2 (0.674 × 0.468 $\lambda_0{}^2$) Dual IBWs are produced, with one at a lower-frequency band of 28.55% and another at a higher-frequency band of 14.98%. The lower resonating frequency band might be utilized for broadcasting and fixed mobile application, and higher resonating band can be used for broadcasting and fixed radio-location application.

References

1. Lu JH, Liou CW (2014) Planar dual-band circular polarization monopole antenna for wireless local area networks. IEEE Antennas Wirel Propag Lett 14:478–481
2. Yang D, Zhai H, Xue K, Xi L, Wu X, Li Y (2019) A dual-band circularly polarized planar monopole antenna for wireless local area network/worldwide interoperability for microwave access applications. Microw Opt Technol Lett 61(2):399–404
3. Li JF, Wu DL, Zhang G, Wu YJ, Mao CX (2019) A left/right-handed dual circularly-polarized antenna with duplexing and filtering performance. IEEE Access 7:35431–35437
4. Xu R, Gao SS, Li J, Wei K, Luo Q (2020) A reconfigurable dual-band dual-circularly polarized antenna for vehicle global navigation satellite system application. IEEE Trans Veh Technol 69(10):11857–11867
5. Dhara R (2021) Dual Band Dual Polarized Planar Monopole Antenna for L, and C Band Applications. Prog Electromag Res Lett 96:65–72. https://doi.org/10.2528/PIERL20121903
6. Ullah U, Koziel S (2019) A geometrically simple compact wideband circularly polarized antenna. IEEE Antennas Wirel Propag Lett 18(6):1179–1183
7. Zheng JX, Lai JY, Wang CJ (2018) A compact CPW-fed monopole antenna for dual-band circular polarization. In: 2018 IEEE International Symposium on Electromagnetic Compatibility and 2018 IEEE Asia-Pacific Symposium on Electromagnetic Compatibility (EMC/APEMC), pp 977–980. IEEE
8. Balanis CA (2005) Antenna theory, analysis, and design, 3rd edn. Wiley, New York

Massive MIMO Systems Precoder Design Under Various Environments

R. Srividhya and M. Anto Bennet

1 Introduction

In the current wireless scenario where the number of users and achievable rate has considerably increased and it incorporates multi-user MIMO system [1]. The MIMO technology enables multiple transmission of signals to large number of user terminals by employing multiple antenna elements at the transmitter. The organization of such multiple signals from multiple user terminals is coined as beamforming. The generalization of beamforming is also known as precoding. Fading is an important parameter in any wireless system that should be reduced significantly. MIMO technology plays a vital role in minimizing the fading effects in wireless channels by allowing multiple transmissions of the same signal by employing large number of antenna arrays at the base station. At the mobile receiver terminal (MRT), the losses incurred due to fading is significantly reduced by multiple transmission of the same signal and the rate at which the signals are combined given by maximum ratio combining (MRC). If the channels between the transmitter and receiver are assumed to be virtual in a MIMO system, then the throughput of the system can be maximized. Figure 1 illustrates the single-user MIMO (SU-MIMO) where only a single user can be provided with the access at a particular time and termed as conventional MIMO scheme. Since the single-user MIMO had several drawbacks and did not address many of practical applications so the researchers moved to a change in MIMO system.

Figure 2 illustrates the multi-user MIMO along with the space division multiple access (SDMA) [2]. A multi-user MIMO system can serve more than two user terminals but the major drawback is computational complexity. In addition, in such wireless systems that employ multiple user MIMO the drawback is the interface

R. Srividhya (✉) · M. A. Bennet
Department of Electronics and Communication Engineering, Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Chennai, Tamil Nadu 600054, India
e-mail: vidhsdec9@gmail.com

Fig. 1 Single-user MIMO
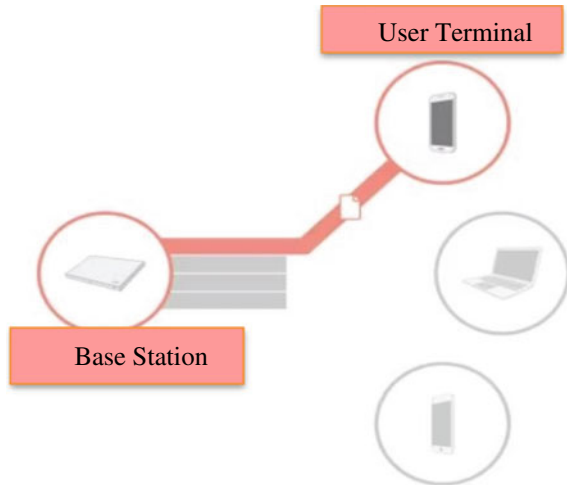
between the user terminals. As the number of user terminals increases, the interface maintenance between the terminals is difficult since they are controlled by a single base station. This limitation is resolved by employing signal processing techniques at the transmitting side of a MIMO system. So, if the channel state information (CSIT) is known at the transmitter, then the gain and throughput of the system can be significantly improved compared to conventional schemes.

Wireless communication technology in recent days has gained significant attention as it could incorporate hundreds to thousands of antennas at the base station as shown in Fig. 3. The concept of massive MIMO system is typically used for 5G and 6G applications by providing coverage to smaller L number of users per cell by means of a base station along with acceptable N number of antennas, $N \ll L$ [3]. The coverage provided by a massive MIMO system is significantly very high when compared with multi-user MIMO. Various signal processing techniques are performed at the transmitter in order to improve the number of users and gain of the system to enhance the throughput.

Fig. 2 Multi-user MIMO
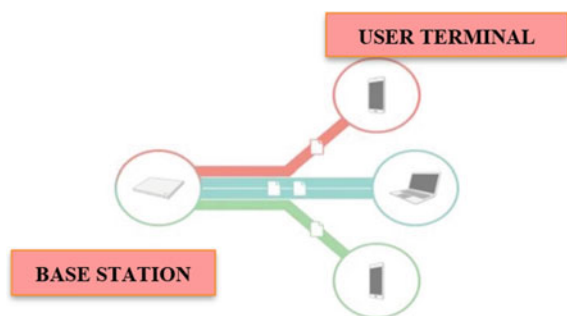
Fig. 3 Massive MIMO
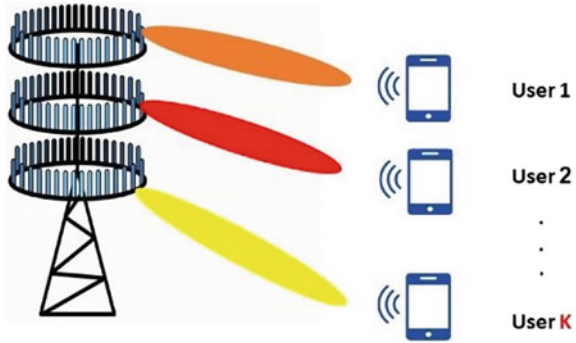
The preprocessing stage of the RF system employs a transmit signal processing technique such as precoding. The spectral efficiency of a MIMO system can be improved significantly by knowing the channel state information (CSI) at the transmitter and by performing precoding techniques at the base station. In conventional techniques, the complexity of the system is addressed at the receiver side which is transformed to the base station by performing signal processing techniques at the transmitter side such as precoding in a massive MIMO system. The next promising issue in most real-time scenarios is to maintain a perfect CSI at the transmitter for different wireless propagation channels. Many researchers have found that MIMO precoding techniques significantly improve the system performance and throughput by varying the beam angles and leading them to a specific user terminal. The precoding techniques of a massive MIMO system involve infinite antennas that enhance the throughput and capacity of the system by reducing the properties of fading and interference. The precoding is broadly divided into several categories such as linear and nonlinear techniques. Precoding along with beamforming is used together in Wi-Fi, 4G, and 5G systems. Multi-user interference (MUI) can be mitigated by using simple precoding algorithm called as zero forcing (ZF) [4]. This paper proposes different precoding techniques for wireless communication channels under various environments for a massive MIMO-based OFDM system. The simulation results are then compared with other precoding techniques in terms of bit error rate (BER) and signal-to-noise ratio (SNR). Then the best precoding algorithm is identified and employed for various applications.

2 Linear Precoding Algorithms

Linear precoding is also known as beamforming, and it is a particular subclass of schemes transmission that serves multiple users at the time. The given weights of user are selected in precoding to maximize the ratio between the signal gain at this user and the interference generated at other users (with some weights of users) plus noise. Linear precoding is more attractive for large-scale MIMO. In multi-user

Fig. 4 Massive MIMO with linear precoding
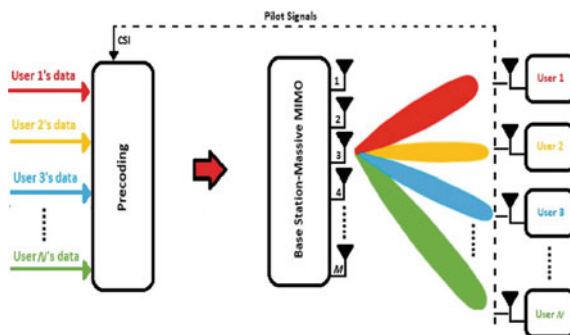
MIMO system, it provides high capacity with space division multiple access [19]. The channel state information if available at base station (BS) or access point (AP), then multi-user information can be suppressed. Precoding allows to perform many complex processing at base station or AP. It helps simplification at users end, the linear precoding techniques. But nonlinear techniques yield better performance (Fig. 4).

For obtaining accurate gain and phase the multi-user MIMO and massive MIMO incorporates precoding schemes (hybrid beamforming). The Ns data streams are passed into the base band precoder and the weights are computed to be passed into number of RF chains and then to the analog precoder as shown in Fig. 5 to compute the channel matrix [H].

In massive MIMO systems, the technique of precoding enables multiple versions of data from different user terminals by assigning weights to every signal that originates from every antenna element separately. The main principle of the precoder is to determine the precoding weight vector in such a way that it improves the link throughput of the system at the receiver terminal. The authors of [5] Keke Zu have presented many linear precoding techniques in a detailed manner. He concluded by the fact that the computational complexity of the MIMO system is directly proportional to the number of antenna array elements at the base station. Thus, there is always a trade-off between the complexity and efficiency of the massive MIMO system.

Fig. 5 Digital precoding
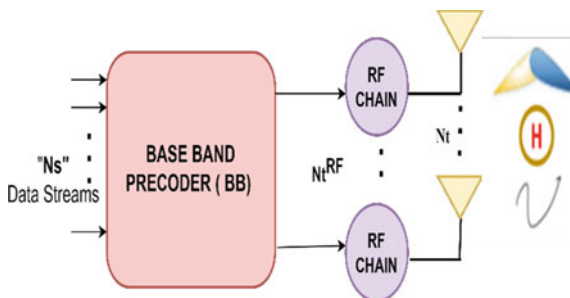
2.1 Minimum Mean Square Error (MMSE) Algorithm

The MMSE algorithm is based on the computation of mean square error and finding the minimum among them for a MIMO system. Bahrami et al. [6] illustrated that the linear precoding techniques such as maximum ratio transmission (MRT), zero forcing (ZF), and MMSE are low complexity algorithms and simple but with the drawback is poor performance. In conventional systems, the reversal property of the channel is used but MMSE-based precoding technique employs regularized version of channel inversion as function of balancing the magnitude of the signal along with its alignment and noise compression termed as tuning in MMSE expression. Specifically tuning a function in MMSE is to vary the parameters in a regularized manner. For practical applications, MMSE-based precoding mainly concentrates on the antenna array structure and the number of antenna elements used. For a MMSE-based MIMO system, the precoding matrix is,

$$P_{\text{MMSE}} = \frac{1}{\beta} H^H \left(\text{HH}^H + \frac{\text{Nu}}{P_{\text{tr}}} I_{\text{Nu}} \right)^{-1} \tag{1}$$

Then, the vector at the receiver terminal of a MMSE-based precoder is,

$$y = \frac{1}{\beta} H \left[H^H \left(\text{HH}^H + \frac{\text{Nu}}{P_{\text{tr}}} I_{\text{Nu}} \right)^{-1} \right] x + n. \tag{2}$$

where

$$\beta = \sqrt{\frac{tr \left(H^H \left(\text{HH}^H + \frac{\text{Nu}}{P_{\text{tr}}} I_{\text{Nu}} \right)^{-1} \right)}{P_{\text{tr}}}}. \tag{3}$$

From the above expressions, it is evident that MMSE-based precoders usually offer a trade-off between the maximum ratio transmission (MRT) and zero forcing (ZF) algorithm [12] and it also proves that the performance of the system has significantly improved even if the wireless channel is noisy when compared with zero forcing (ZF) technique [13]. The major advantage of MMSE is that it eliminates the inter-user interference (IUI) and offers an optimal achievable rate even in the presence of noise [14] for all types of propagation channels.

2.2 Maximum Ratio Transmission Algorithm

Maximum ratio transmission algorithm performs beamforming at the transmitter side of a MIMO system by employing several antennas at the transmitter and receiver

side for successful transmission of data streams. It allows the system to make use of the same receiver by improving the diversity order [19]. The MRT algorithm is considered for special types of scenario such as two receive antennas and infinite number of transmit antennas for which the probability of error is computed and found to be less compared to other conventional algorithms.

For a MRT-based MIMO system, the precoding matrix is,

$$P_{\text{MRT}} = \frac{1}{\beta} H^H \tag{4}$$

$$\beta = \sqrt{\frac{\text{tr}(BB^H)}{P_{\text{tr}}}} \tag{5}$$

Also,

$$B = H^H \tag{6}$$

Then, the vector at the receiver terminal of a MRT-based precoder is,

$$y = \frac{1}{\beta} HH^H x + n \tag{7}$$

$$\beta = \sqrt{\frac{\text{tr}(HH^H)}{P_{\text{tr}}}} \tag{8}$$

A MIMO system consists of several transmitting schemes that were discussed in many precoding literatures. Wittenben proposed a transmitting algorithm that employs a delayed version of the signal [7, 8]. Also, Seshadri and Winters [9, 10] proposed a similar algorithm by slightly varying the delayed signal by transmitting the repeated versions of the signal continuously through multiple antennas at the base station at different time intervals. Additionally, Alamouti [11] presented a transmit diversity scheme that uses a couple of symbols that are to be transmitted by employing two antennas at the base station initially and again transmits the modified version of the pair of symbols that arrive at the diversity based on maximum ratio transmission (MRT). The proposed algorithm is very simple and provides high efficiency compared to conventional algorithms. Though the preliminary objective of such precoding schemes is to maximize the SNR, it also provides sub-optimal values of achievable sum rate.

Authors of [15] proposed signal processing techniques at the base station for every antenna terminal. The MRT is more adaptable, and its efficiency is near to the optimal performance for any system that has limited noise component [16]. But for such systems to operate efficiently the columns of the propagation matrix need to be orthogonal elements.

2.3 Zero Forcing Algorithm

The multi-user interference (MUI) can be nullified in a multi-user MIMO communication system by employing signal processing techniques to the antenna elements at the base station. The zero forcing (ZF) precoders require pseudo-inversion of the channel matrix [H] known at the transmitter. This algorithm is typically employed for nullifying the inter-user interference (IUI) by re-transmitting the same version of every signal to the receiver. For a ZF-based MIMO system, the precoding matrix is,

$$P_{ZF} = \frac{1}{\beta} H^H \left(HH^H\right)^{-1}, \tag{9}$$

where

$$\beta = \sqrt{\frac{\text{tr}(BB^H)}{P_{tr}}} \tag{10}$$

Also,

$$B = H^H (HH^H)^{-1}. \tag{11}$$

Then, the vector at the receiver terminal of a ZF-based precoder is,

$$y = \frac{1}{\beta} H \left[H^H \left(HH^H\right)^{-1} \right] x + n, \tag{12}$$

where

$$\beta = \sqrt{\frac{\text{tr}(HH^H \left(HH^H\right)^{-2}}{P_{tr}}} \tag{13}$$

The data rate of zero forcing (ZF) algorithm is very high when compared with MRT algorithm [17]. This algorithm completely minimizes the IUI, and there is a trade-off between complexity and efficiency of the MIMO system. For interference-controlled system, the ZF algorithm provides optimal results when compared with MRT and MMSE [18].

3 Results and Analysis

The simulation results were obtained under different environments such as urban, sub-urban, and rural areas. The number of transmitting antennas at the base station was kept at $T = 50$, and the number of user terminals was $N = 4$ for all the scenarios.

The simulation was computed for a massive MIMO system with OFDM type of modulation that consists of fifty-two sub-carriers and 64-bit quadrature amplitude modulation (QAM) modulated symbol under various propagation scenarios. All the simulation results were obtained using MATLAB 2020b.

The analysis has been conducted using the linear precoding algorithms such as zero forcing, minimum mean square error, maximum ratio transmission under the urban scenario with number of transmitting base station antennas $T = 50$, number of users being $N = 4$. The MMSE algorithm is observed to be performing well under this scenario with BER $= 10^{-5}$ and SNR $= 16$ dB as shown in Fig. 6. The comparative analysis of various linear algorithms such as ZF, MRT, and MMSE is tabulated in Fig. 7. For urban area scenario, the MMSE algorithm is observed to be performing well under this scenario with BER $= 10^{-5}$ and SNR $= 21$ dB as shown in Fig. 8.

Further the analysis has been conducted using the linear precoding algorithms such as zero forcing, minimum mean square error, maximum ratio transmission under the sub-urban scenario with number of transmitting base station antennas $T=50$, number of users being $N=4$.

The comparative analysis of various linear algorithms such as ZF, MRT, and MMSE is tabulated in Fig. 9. for sub-urban area scenario.

The analysis has been conducted using the linear precoding algorithms such as zero forcing, minimum mean square error, maximum ratio transmission under the rural scenario with number of transmitting base station antennas $T = 50$, number of users being $N = 4$. The MMSE algorithm is observed to be performing best under this scenario with BER $= -2511.88$ and SNR $= 14$dB as shown in Fig. 10.
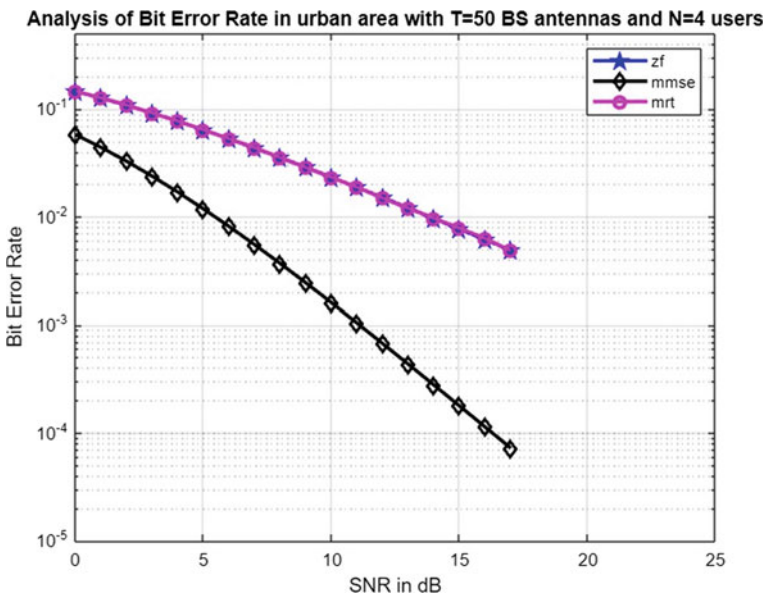
Fig. 6 Urban area scenario

ALGORITHM	NO.OF BASE STATION ANTENNAS	NO.OF USER TERMINALS	BIT ERROR RATE (BER)	SNR (dB)
Zero Forcing (ZF)	50	4	-10^-2	17
Maximum Ratio Transmission (MRT)	50	4	-10^-2	17
Minimum Mean Square Error (MMSE)	50	4	-10^-4	16

Fig. 7 Urban area scenario—comparison of ZF, MRT, and MMSE
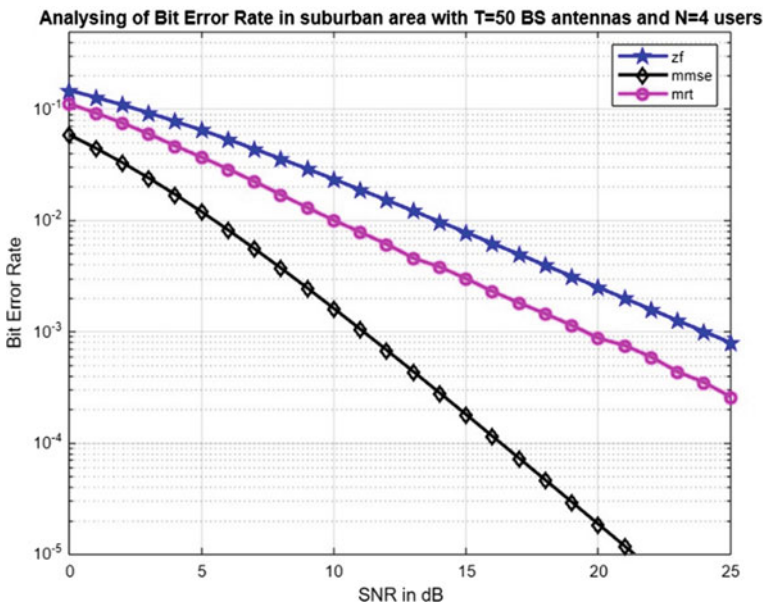
Fig. 8 Sub-urban area scenario

The comparative analysis of various linear algorithms such as ZF, MRT, and MMSE is tabulated in Fig. 11 for rural area scenario.

ALGORITHM	NO.OF BASE STATION ANTENNAS	NO.OF USER TERMINALS	BIT ERROR RATE (BER)	SNR (dB)
Zero Forcing (ZF)	50	4	-10^{-3}	25
Maximum Ratio Transmission (MRT)	50	4	-10^{-4}	25
Minimum Mean Square Error (MMSE)	50	4	-10^{-5}	21

Fig. 9 Sub-urban area scenario—comparison of ZF, MRT, and MMSE
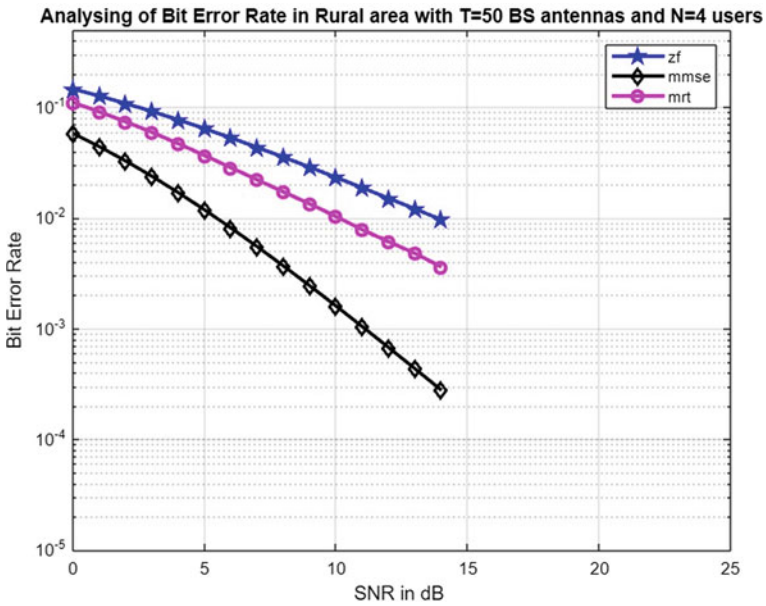
Fig.10 Rural area scenario

ALGORITHM	NO.OF BASE STATION ANTENNAS	NO.OF USER TERMINALS	BIT ERROR RATE (BER)	SNR (dB)
Zero Forcing (ZF)	50	4	-100	15
Maximum Ratio Transmission (MRT)	50	4	-325.5	14
Minimum Mean Square Error (MMSE)	50	4	-2511.88	14

Fig. 11 Rural area scenario—comparison of ZF, MRT, and MMSE

4 Conclusion

This work compares the massive MIMO system incorporated with different linear precoding algorithms such as MRT, MMSE, and ZF under different environments with QAM type of modulation. The bit error rate of all these algorithms under different scenario was tabulated. The matrix inversion is employed to reduce computational complexity of the MMSE-based precoder. The simulation results of MRT under urban scenario proved to be good. To conclude, by comparing different scenarios, it is observed that MMSE has less SNR and less BER among MRT and ZF precoding algorithms. Hence for rural environment, the precoder can be designed using MMSE or RZF algorithm is employed.

5 Future Scope

High-speed Internet and fast access to the network have drastically increased in recent years. The rapid evolution of the mobile technology tremendously increases the number of users and fast services in all the fields. These requirements lead to the Massive MIMO system, and for efficient functioning of such large system, signal processing techniques were incorporated at the base station. Such signal processing techniques were precoding, scheduling, and channel estimation. Linear precoding algorithms, with which we can simulate for a greater number of subscribers. The future work can be extended by employing deep learning algorithms to overcome the computational complexity of these conventional algorithms.

References

1. Qiao X, Zhang Y, Yang L (2018) Conjugate gradient method based linear precoding with low-complexity for massive MIMO systems. In: Proc IEEE 4th International Conference Computer Communication (ICCC), pp 420–424
2. Albreem MA, Juntti M, Shahabuddin S, Massive MIMO detection techniques: a survey. IEEE Commun Surveys Tuts 21(4):3109–3132
3. Lee B (2017) 'Simplified antenna group determination of RS overhead reduced massive MIMO for wireless sensor networks.' Sensors 18(2):84
4. Bahrami H-R, Le-Ngoc T (2008) Maximum ratio combining precoding for multi-antenna relay systems. In: 2008 IEEE International Conference on Communications, pp 820–824. IEEE
5. Emil EGL, Marzetta TL (2016) Massive MIMO: ten myths and one critical question. IEEE Comm Magazine 54(2):114–123
6. Bai Y, Liang Z, Zhai C, Xin Y, Li W (2019) Joint precoding using successive over-relaxation matrix inversion and Newton iteration for massive MIMO systems. In: Proc 11th International Conference Wireless Communication Signal Process (WCSP), pp 1–5
7. Seshadri N, Winters J (1993) Two signaling schemes for improving the error performance of FDD transmission systems using transmit antenna diversity. In: Proceedings IEEE Vehicular Technology Conference (VTC), May, pp 508–511
8. Fatema N, Hua G, Xiang Y, Peng D, Natgunanathan I (2018) 'Massive MIMO linear precoding: a survey.' IEEE Syst J 12(4):3920–3931
9. Albreem MA, Juntti M, Shahabuddin S (2019) Massive MIMO detection techniques: a survey. IEEE Commun. Surveys Tuts. 21(4):3109–3132
10. Boccardi F, Huang H (2006) Optimum power allocation for the MIMO-BC zero-forcing precoder with per-antenna power constraints. In: 2006 40th Annual Conference on Information Sciences and Systems, pp 504–504. IEEE
11. Björnson E, Hoydis J, Sanguinetti L (2017) Massive mimo networks: spectral, energy, and hardware efficiency. Found Trends Sig Process 11(3–4):154–655
12. Zheng K, Zhao L, Mei J, Shao B, Xiang W, Hanzo L (2015) Survey of large-scale MIMO systems. IEEE Comm Surveys Tut 17(3):1738–1760
13. Wittneben A (1991) Base station modulation diversity for digital simulcast. In: Proceedings IEEE Vehicular Technology Conference (VTC), May, pp 848–853
14. Fatema N, Hua G, Xiang Y, Peng D, Natgunanathan I (2018) Massive MIMO linear precoding: a survey. IEEE Syst J 12(4):3920–3931
15. Zu K (2013) Novel efficient precoding techniques for multiuser MIMO systems. University of York, PhD diss.
16. Winters J (1994) The diversity gain of transmit diversity in wireless systems in Rayleigh fading. In: Proc ICC/SUPERCOMM, New Orleans, LA, vol 2, pp 1121–1125
17. Chandrasekaran K, Kandasamy P, Ramanathan S (2020) Deep learning and reinforcement learning approach on microgrid. Int Trans Electr Energ Syst, e12531. https://doi.org/10.1002/2050-7038.12531
18. Alamouti S (1998) A simple transmit diversity technique for wireless communications. IEEE J Select Areas Commun 16:1451–1458
19. Subitha D (2020) Analysis of linear precoding techniques for massive MIMO-OFDM systems under various scenarios. In: IOP Conference Series: Materials Science and Engineering, vol 1084. ICCSSS

A Miniaturized Modified Koch Fractal Antenna with Partial Ground Plane for Wideband Application

Tejaswi Kumar, Reshmi Dhara, and Sanoj Mahato

1 Introduction

Wireless communication technology is quickly evolving these days. After G. Marconi successfully sent and received radio waves in 1901, communication transitioned from telegraphy to wire. The integrated antenna, intended for a tiny form factor, low profile, and low manufacturing cost with broadband properties, is a vital component of numerous communication systems. A fractal antenna is one that has a Koch geometry of fractals and can cause a significant return loss. The antenna is a fundamental component of a communication system and is crucial in wireless communication systems; wireless communication is inconceivable without appropriate antenna design and configuration [1]. An antenna, according to IEEE standards, is a device for transmitting or receiving unguided waves [2]. However, Koch fractal antennas are becoming more popular, and they offer numerous benefits, including small size with multiple resonant frequencies, gain optimization, reduced quality factor (Q), and low-cost improvement of bandwidth of Koch fractal antennas based on FR4-epoxy substrate [3, 4]. The recommended antenna is developed through a series of iterations ranging from 0 to 3. In the third iteration, a defective ground plane is added to achieve wideband characteristics and to enhance bandwidth, the area of the ground plane is changed. Several researchers have researched various fractal antennas [5–8] to generate broad impedance and lower resonant frequency. In this study, a small, basic Koch monopole antenna with partial ground plane has been built as a result of the preceding research effort. More fractal ground planes on the left side of the design have been included to increase impedance bandwidth.

T. Kumar · R. Dhara (✉) · S. Mahato

Department of Electronics and Communication Engineering, National Institute of Technology Sikkim, Imphal, India

e-mail: reshmidhara@nitsikkim.ac.in

1.1 Design Configuration

This research presents an optimum ground structure fractal antenna design. The hexagonal patch was built as the initial phase in the design of the proposed antenna using HFSS. The proposed antenna is built on a low-cost FR4 glass epoxy substrate with a 1.6 mm thickness, a loss tangent of 0.02, and a resonance frequency of 3.2 GHz. According to simulations, the patch side length and radius are both 12.6 mm. Figure 1 shows the 0th iteration of the research work of hexagonal-shaped radiator antenna. Figure 1c depicts on the back-to-back iteration of the proposed work. Therefore by adding the meander-line structure on the fractal antenna of all six faces of hexagonal radiator (Figs. 2 and 3). The proposed fractal antenna fundamental design with the whole ground plane is altered. The defective ground structure with length (G_L) and width (G_W) improves return loss and bandwidth over the frequency range of 2.62–4.48 GHz, as shown in Fig. 4. Table 1 depicts the geometry of the suggested fractal antenna, the revised ground plane, and the size of the partial ground plane are all discussed. The projected antenna's simulated impedance bandwidth runs from 2.62 GHz to 4.48 GHz, or 52% of the total bandwidth, and its resonance frequency is 3.55 GHz.

To obtain a fractal antenna dimension, first compute the dimensions at a frequency of 2.62 GHz, a dielectric constant of 4.4, and a loss tangent of 0.02. The antenna design parameters for theoretical computing are depicted in the given equation.

Patch width formulae (W)

$$W = \frac{1}{2fr\sqrt{\mu0\varepsilon0}}\sqrt{\frac{2}{\varepsilon r + 1}} \tag{1}$$

Effective Dielectric Constant $(\varepsilon_{\mathrm{reff}})$
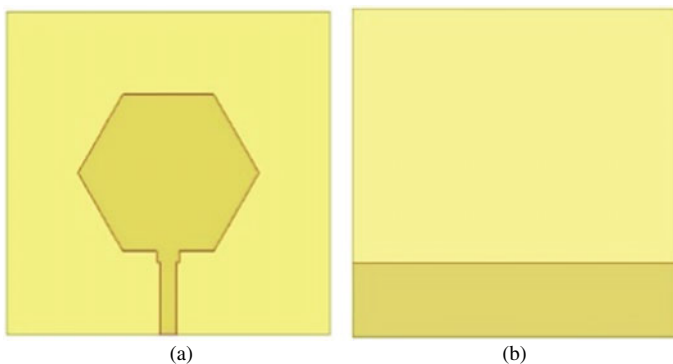
Fig. 1 Ant #1 zero iteration **a** front view **b** back view
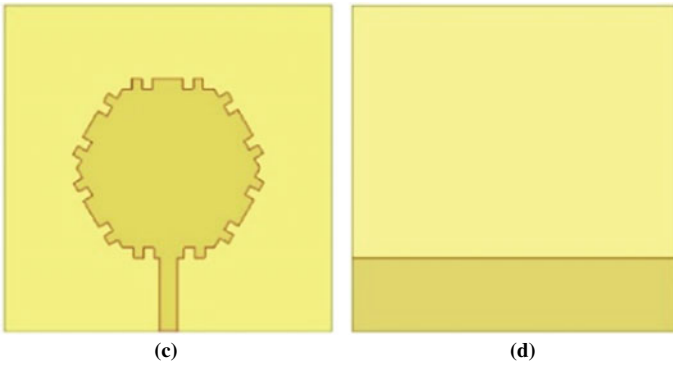
Fig. 2 Ant #2 1st iteration **c** front view **d** back view
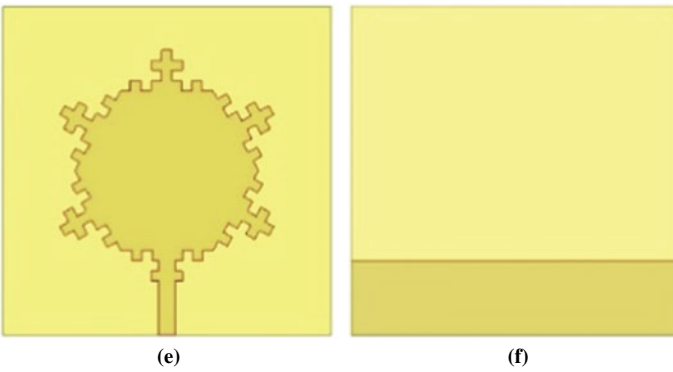
Fig. 3 Ant #3 2nd iteration **e** front view **f** back view
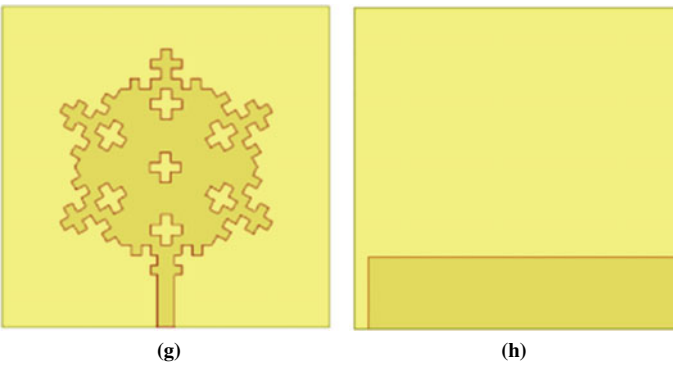
Fig. 4 Ant #4 3rd iteration **g** front view **h** back view

Table 1 Proposed antenna dimension optimization

Parameters for antenna structure	Explanation	Value (mm)	Parameters for antenna structure	Explanation	Value (mm)
$P_G = P_Y$	Radius	12.6	$Q_1 = Q_3$	Mnd. element	4.2
Q_2	Mnd, line	1.4	$Q_{S1} = Q_{S2}$	Mnd. slots	1.4
S_W	Width sub	45	S_L	Length sub	44.9
F_W	Feedline width	2.4	F_L	Feedline length	10.15
h	Height sub	1.6	G_L	Defective plane length	10.12
G_W	Defective plane width	43	g_1	Horizontal width	2.56

Mnd: Meander

$$\varepsilon_{\text{reff}} = \frac{\varepsilon r + 1}{2} \frac{\varepsilon r - 1}{2} \left(1 + 12 \frac{h}{w}\right)^{-\frac{1}{2}} \tag{2}$$

Effective Length formulae (L_{eff})

$$L_{\text{eff}} = \frac{\lambda g}{2} \frac{\Delta L}{h} = 0.412 \frac{(\varepsilon_{\text{reff}} + 0.3)\left(\frac{w}{h} + 0.264\right)}{(\varepsilon_{\text{reff}} - 0.258)\left(\frac{w}{h} + 0.8\right)} \tag{3}$$

Patch length formulae (L)

$$L = L_{\text{eff}} - 2\Delta L \tag{4}$$

Substrate length formulae

$$L_{\text{sub}} = L + 12 \times h \tag{5}$$

Substrate width formulae

$$W_{\text{sub}} = W + 12 \times h \tag{6}$$

When the lower frequency f_L is 2.62 GHz, we get: The length of the radiator is 26.92 mm, the length of the substrate is 46.2 mm, the width of the radiator is 34.84 mm, and the width of the substrate is 54.2 mm. As a result of this usual design method, the antenna footprint is 2484 mm² (46 × 54 mm²). In this research, Ansys HFSS optometric tools are employed to improve the parameter. The antenna's size has been reduced by 18.66% after the design has been modified.

Figure 2c depicts the second iteration of the proposed fractal antenna, in which the slots are introduced by meander components of comparable scale. Additionally, this geometry structure is improved by adding + shaped slots to the radiator, in
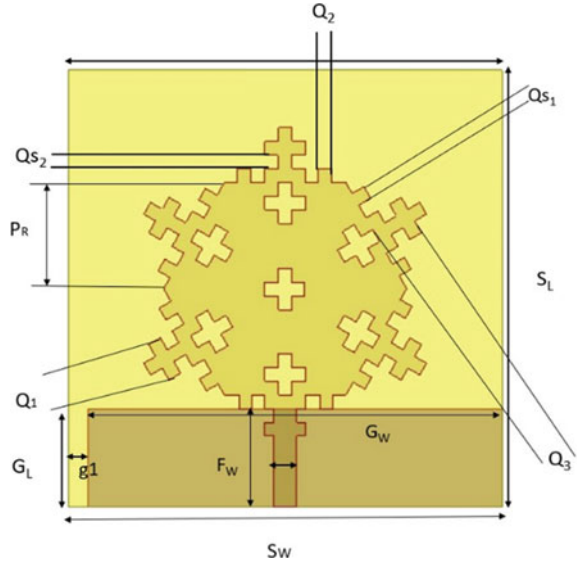
which the radiating patch of the third iteration and ultimate shape, as illustrated in Fig. 5. The configuration of fractal antenna is designed for many iterations process, including 0th, 1st, 2nd, and 3rd, where with the help of iteration method it was able to improve impedance bandwidth. A notch is introduced on the leftmost side of the partial ground plane, which provides the broadest impedance bandwidth compared to all other antennas.

1.2 Result and Discussion

The low and high frequencies of 1 and 7 GHz are utilized to examine numerous essential performance characteristics for antenna design. The suggested antenna has improved from the first to the last three generations. In the final version, a defective plane is also integrated to rise of a bandwidth frequency, as shown in a given Fig. 6, with an impedance bandwidth frequency range of 3.55 GHz.

As the result, the comparison of the proposed work of fractal antenna from different iteration methods is added with the respective form of different types of defective plane and they are by adding the notch structure on ground portion of fractal antenna as shown in Fig. 7, respectively (Table 2).

As a result, the recommended fractal antenna ground plane length, or '$g1$,' was taken into account for the parametric analysis and is an important quantity to examine the antenna's properties. The recommended fractal antenna's bandwidth and return loss calculations for various ground lengths are given in Fig. 8.

Fig. 6 Designing of antenna characteristics at a lower resonant frequency of 2.6 GHz
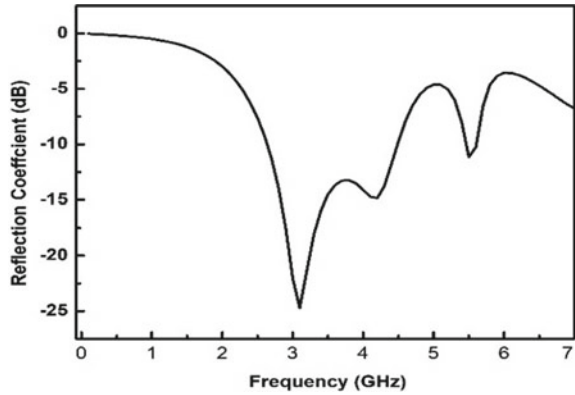
Fig.7 Comparison shows the implemented antenna Ant#1 to #5 with a simulated return loss up to 7 GHz
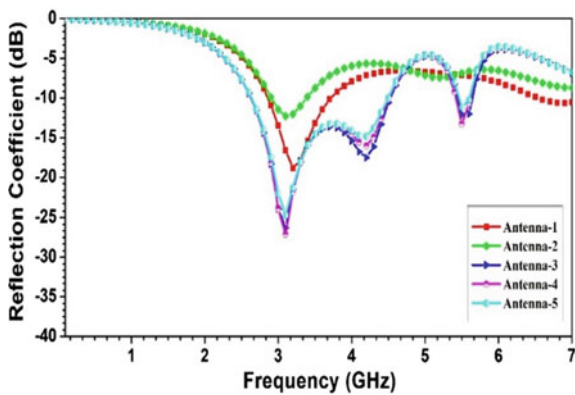
Table 2 Achievement of the suggested dractal antenna using various ground parts

Fractal Antenna structure	Operating frequencies (GHz)	Bandwidth (%)	Bandwidth (GHz)
Ant #1	2.85–3.71	26.21	3.2
Ant #2	2.89–3.38	15.62	3.1
Ant #3	2.64–4.53	52.71	3.0
Ant# 4	2.62–4.52	52.80	3.0
Ant# 5	2.62–4.48	53.0	3.2

The width of the notch 'g1' is taken as the suggested dimension to see the parametric form of study to recommended as fractal antenna. The fundamental Fig. 8 to evaluate wide impedance bandwidth for purposed work and to notice that when g1 = 2 mm delivers the enlarge bandwidth compared to other iteration methods used on hexagonal radiator. The applied work of the radiation pattern is indicated by well-defined. This gap creates a series capacitance which create the coupling between the radiator and ground.
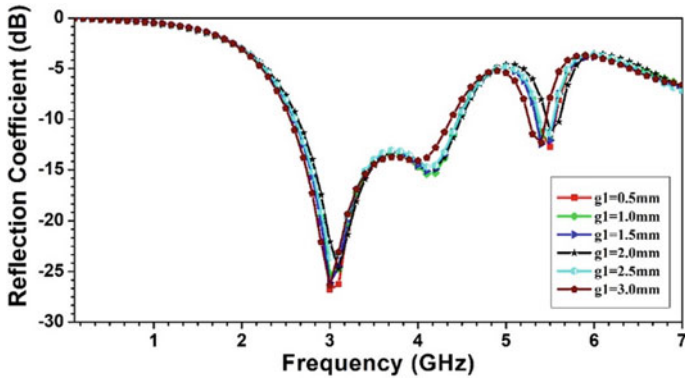
Fig. 8 Proposed fractal antenna with 'g1' variation vs frequency
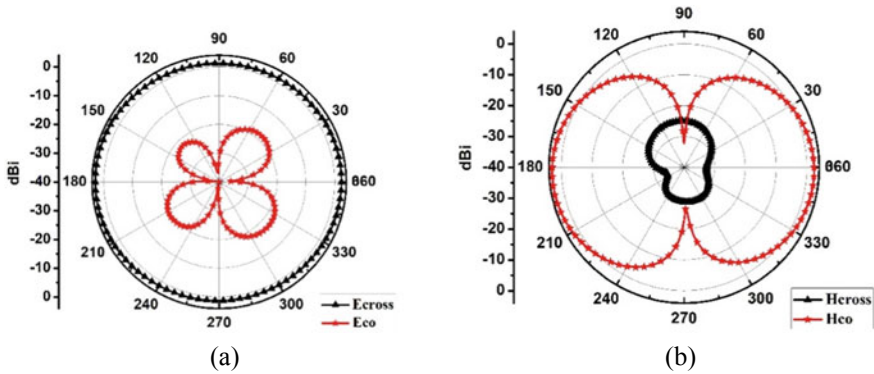
Fig. 9 Radiation pattern simulated for **a** the XZ (0°) and **b** the YZ (90°) plane

Simulated radiation pattern at 3.2 GHz is shown in Fig. 9a, b. The ECO/ECROSS and HCO/HCROSS plane radiation is 35 dBi greater in the widest direction, as shown in Fig. 9a, b. Figure 10 displays the frequency-dependent peak gain simulation where the maximum peak gain inside the impedance band was recorded at 3.2 GHz as 5.0 dBi with a radiation efficiency of 96%. This result is ideal for wireless technology.

There are many applications that could benefit from a fractal antenna. There are several ideas where fractal antenna could make a real impact to a device integrating global perspective satellites. Table 3 depicts the simulated antenna performance with comparison between some recently reported fractal antenna, compare to all structure proposed work gives reasonable impedance bandwidth at lower resonant frequency as well as reduction in size also.

(a)   (b)
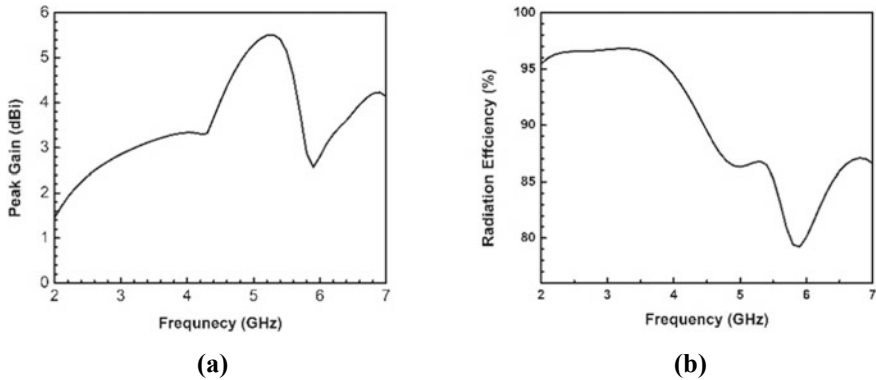
Fig. 10 **a** Peak gain simulation **b** radiation efficiency with respect to frequency

Table 3 Evaluation of achievement against recommended antenna

Reference nos.	Size (mm^2)	Operating frequency(GHz)	Structure
[5]	26 × 30	3.1–10.6	Hexagonal micro-strip fractal antenna for UWB
[6]	40 × 80	1.45–4.52	Koch snowflake fractal antenna
[7]	20 × 40	2.5–11	Hybrid Sierpinski fractal UWB MIMO antenna
[8]	36 × 34	1.2–10.9	Minkowski fractal adding TCM analysis
Proposed work	**45 × 44.9**	**2.62–4.51**	**Miniaturized modified Koch fractal antenna**

2 Conclusion

Many variations of fractal geometry antenna have been incorporated into the design of antennas. The dimension shows the value of 44.9×45 mm^2 wideband application-optimal fractal antenna design with an altered form of ground structure as proposed. The number of iterations on a fractal antenna design affects the return loss for higher frequencies in a multiband antenna. The length of the reverberation perimeter also affects the return loss and the resulting frequency. Several iterations can be studied to understand the trends governing the antenna to better understand the physics of the problem. After computing the provided equation, the suggested work fractal antenna with size reduction is 18.66%, although this work may be used for the antenna is constructed at 2.5 GHz in order to cover long-term evolution (LTE) communication system or that can be used for WiMAX (2.5–2.7) GHz (3.3–3.8 GHz). It is also employed in several C-band applications.

References

1. Balanis CA (2015) Antenna theory: analysis and design. Wiley
2. Singh A, Singh S (2015) A novel CPW-fed wideband printed monopole antenna with DGS. AEU-Inter J Elect Comm 69(1):299–306
3. Dhara R (2021) A compact dual band dual polarized monopole antenna with enhanced bandwidth for C, X, and Ku band applications. Prog Electromag Res Lett 96:65–72
4. Dhara R, Mitra M (2020) A triple-band circularly polarized annular ring antenna with asymmetric ground plane for wireless applications. Eng Rep 2(4):e12150
5. Sawant KK, Kumar CS (2015) CPW fed hexagonal micro strip fractal antenna for UWB wireless communications. AEU-Int J Elect Comm 69(1):31–38
6. Choukiker YK, Behera SK (2017) Wideband frequency reconfigurable Koch snowflake fractal antenna. IET Microwaves Antennas Propag 11(2):203–208
7. Sampath R, Selvan KT (2020) Compact hybrid Sierpinski Koch fractal UWB MIMO antenna with pattern diversity. Int J RF Microwave Comput Aided Eng 30(1):e22017
8. Dhara R, Govil MC, Kundu T (2021) Application of theory of characteristics modes for bandwidth enhancement of a miniaturized Minkowski fractal antenna. IETE J Res, 1–16

Comparative Study on YOLOv2 Object Detection Based on Various Pretrained Networks

Richard Ningthoujam, Keisham Pritamdas, and Loitongbam Surajkumar Singh

1 Introduction

In recent years, there has been a huge increase in the data space of images and videos. Such data may be from storage devices or real-time cameras of advanced embedded systems. Thereby object detection and classification in this ever-growing data have become one of the greatest challenges in computer vision, which requires accuracy and speed. Redmond et al. [1] introduced YOLO in 2015 as one of the fastest in object detection by achieving up to 155 fps in its smaller fast version. The base YOLOv1 also achieved a speed of 45 fps which is very high and an mAP of 63% which is a good level in PASCAL VOC 2007 when the state of the art of that time fast R-CNN has 71% mAP at a lower speed of 5 fps. Later, Redmon and Farhadi [2] proposed another version of YOLOv2 in 2016 as an improved version of YOLOv1 with an accuracy of 76.8 at 67 fps on the VOC 2007 dataset. Nowadays, we also have various types of pretrained networks available around us. Each layer of such a neural network model is capable of extracting different types of features from input images during training. Such pretrained networks can be used with YOLO as transfer learning where the pretrained network portion will be used for feature extraction and YOLO layers for the detection as depicted in Fig. 1a.

Considering the benefits of YOLO, various research works have been carried out recently. Loey et al. [3] proposed a detection model for face masks based on YOLOv2 by using ResNet50 as a deep transfer learning model for feature extraction. On their novel dataset of people wearing a facemask, the detector model achieved 81% mAP. Itakura and Hosoi [4] used the YOLOv2 in the detection of a 3D image reconstructed from a 360-degree camera and achieved an F-score of 0.94, which is considered good and could use in place of the costly LiDAR for wide range applications. Huang et al.

R. Ningthoujam (✉) · K. Pritamdas · L. S. Singh
Department of ECE, NIT Manipur, Manipur 795004, India
e-mail: richard.ningthoujam@gmail.com

| yolov2Conv1 |
| yolov2Batch1 |
| yolov2Relu1 |
| yolov2Conv2 |
| yolov2Batch2 |
| yolov2Relu2 |
| yolov2ClassConv |
| yolov2Transform |
| yolov2OutputLayer |

(a)　　　　　　　　　　(b)　　　　　　　　　　(c)
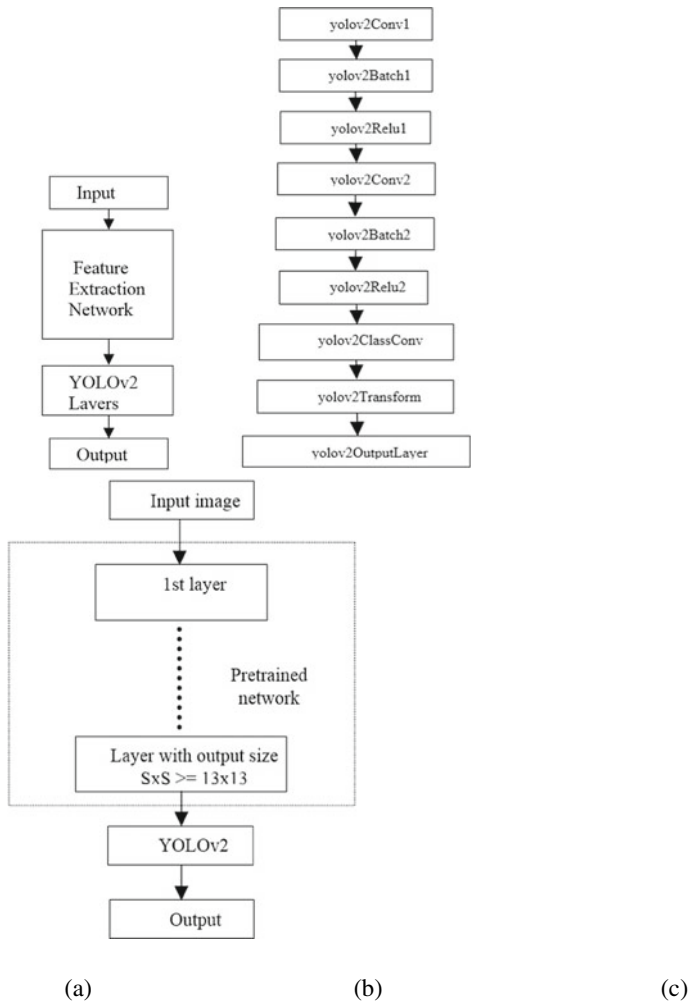
Fig. 1 a YOLO detection model. **b** YOLOv2 layers. **c** YOLO detection model with pretrained network as feature detector

[5] proposed a detection model called M-YOLO which is based on YOLOv3 having MobileNetv3 as a feature extractor. The model achieved 94.96% average precision with 10 fps which is better than the results of YOLOv3 on the same dataset of night-time vehicle images. In another work proposed by Ren et al. [6], the detection model uses the fire module of SqueezeNet in the feature extraction layers, thereby reducing the number of parameters dramatically from the previous YOLO-PC model. It can count people at 41 fps with an average precision of 72%. With an aim of a compressed detection model, Wei Fang et al. [7] came up with Tinier-YOLO which is 4 × reduced model size compared to the Tiny-YOLO which was already small

and fit for embedded devices. To reduce the model size, it uses fire modules of SqueezeNet by strictly monitoring the number of modules and their positions in the network. It achieved 25 fps real-time performance on Jetson TX1 with 65.7% mAP on PASCAL VOC and 34% on COCO datasets. Zhao et al. [8] experimented on a dataset of 2419 wild animal images by using their MobileNet-YOLO, the mAP got increased by 7.7% and speed by 3 fps compared to basic YOLO. In the recent work of Shoai and Sayed [9], Inceptionv3 is found to be used as a feature extractor in brain tumor detection applications. The YOLO-Inceptionv3 model has 174 layers in total with 165 layers of inceptionv3 and 09 layers of Tiny YOLOv2 [10]. The Inceptionv3 layers are composed of 01 input, 50 convolution, 50 batch normalization, 50 ReLU, 03 max pool, 06 depth concatenation, and 05 average pool.

Over recent years, the autonomous driving system has become a trending topic which is creating a space for researchers. One of the major challenges is the detection of vehicles in real-time which requires both accuracy and speed. YOLO is one of the favorite techniques for such conditions. So for our work, a single-class vehicle dataset [11] is used as a common dataset for the comparison of the proposed YOLO models. In this work, eight YOLOv2 detection models are implemented based on eight different types of pretrained neural networks. We use MATLAB R2021b, and the pretrained networks are taken from the deep learning toolbox. This paper is organized as follows: Sect. 2 describes the architectural development of the models, YOLO layer structures, and their connection to pretrained networks are discussed. In Sect. 3, the experimental results are given and discussed in terms of mean average precision (mAP). The conclusion of the research is presented in Sect. 4.

2 Proposed Detection Models

2.1 YOLOv2

The YOLOv2 has a total of 9 layers comprising a few convolution layers, a transform layer, and an output layer as shown in Fig. 1b. Instead of going pixel-by-pixel for detection, YOLO divides the input image into S × S grids. The grid which has the center position of the object is considered responsible for detection. Each grid predicts B number of bounding boxes. And each bounding box has a box confidence score, four location information (x, y, w, h), and the probability of each of the classes in the dataset. For example, if there are 3 bounding boxes and 2 classes in the dataset. Then, each grid will have $3 \times (5 + 2)$ tensors. For S × S grids, the total will be S × S × 3 × (5 + 2) tensors. With the concept of anchor boxes in YOLOv2, the top-k bounding boxes are predicted that have the best match of the desired object where k is the number of anchor boxes. The prediction of the anchor boxes is done by running a K-means clustering algorithm, considering the intersection-over-union (IOU) of the predicted box with the ground truth. In YOLOv2, S × S is taken as 13 × 13 [2] which can also be assumed as the feature extraction network has an output of size

13×13 to be connected as an input to the further layers of detection. In our work, the output layer of the pretrained network is considered at an output size of 13×13 or closely bigger as shown in Fig. 1c.

2.2 Pretrained Networks for the YOLOv2 Detection Model

Residual Network (ResNet50). Residual network (ResNet) is a type of deep learning network based on residual connection [12]. Residual connection is created from a specific layer to a point in further layers by skipping some layers in between as shown in Fig. 2. The residual connection suppresses the problem of vanishing/ exploding gradient. The ResNet50 architecture has a total of 177 layers with 16 residual bottlenecks. The 141st layer of the ResNet50 having an output of $14 \times 14 \times 1024$ is then connected as an input to the first layer of YOLOv2 layers creating a network model as shown in Fig. 1 © by considering the $S \times S$ as 14×14.

AlexNet. AlexNet is one of the most popular convolutional neural networks in computer vision which also won the LSVRC in 2012. It achieved higher accuracy on visual recognition tasks performed on images of a huge ImageNet dataset, thereby showing good feature extraction. It has a total of 25 layers including the output layer, having 8 layers with weights, i.e., 5 convolutional and 3 fully connected layers. To use the AlexNet with YOLO for feature extraction, the 15th layer of the AlexNet being the last layer with size 13×13 is connected with the first layer of YOLOv2.

Fig. 2 Residual connection
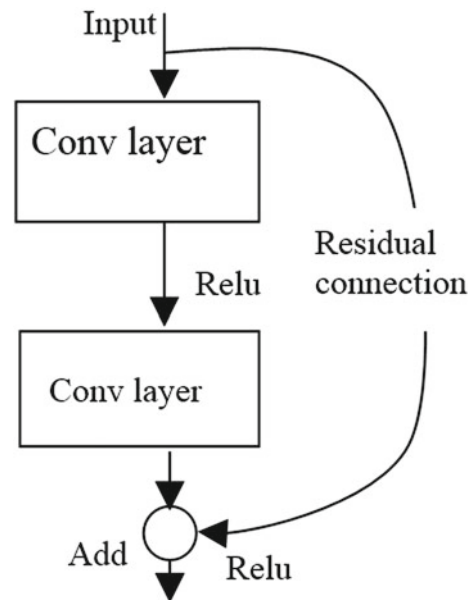
Fig. 3 Inception module
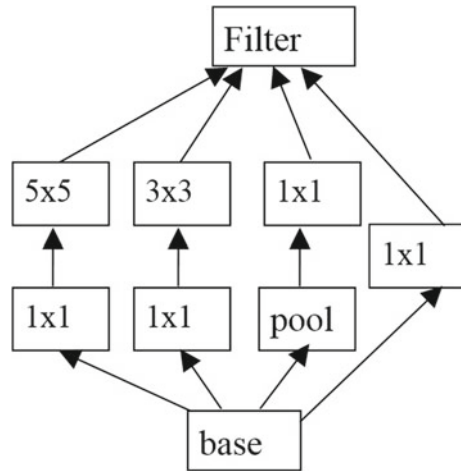
GoogleNet. GoogleNet showed a significant decrease in error rate compared to AlexNet and also won the ILSRVC in 2014. It has a special component called the inception module where convolution is done using kernels of different scales (5 × 5, 3 × 3, 1 × 1) in parallel and later concatenated [13]. As the network is computation intensive, it also uses 1 × 1 filters to reduce the number of parameters from the previous layer as depicted in Fig. 3. It has a total of 144 layers with 9 inception blocks. As the output of 110th is having a size of 14 × 14 which is closest to 13 × 13, it is connected as an input to the YOLOv2 detection layers.

MobileNet v2. MobileNet was introduced as an optimized neural network architecture for mobile vision applications [14]. MobileNet v2 inherited the depth-wise separable convolution technique from its previous version MobileNet v1 where a single convolution kernel is applied per channel instead of a single 3D convolution. It dramatically reduced the number of network parameters compared to normal convolution, which resulted in reduced computation complexity and resources. It has a special block called an inverted residual block (bottleneck block) [5] where a residual connection is made between two convolutional blocks of narrow channels. Inside the residual block, the narrow channel from an input is expanded in depth by pointwise convolution, and depth-wise separable convolution is applied. Both the convolutions are followed by ReLU6 (ReLU having positive values limited to 6). Then, a pointwise convolution is applied to get a channel depth that matches the dimension of the residual connection. The last convolution does not have a non-linear layer as shown in Fig. 4. MobileNet v2 has a total of 154 layers, but layers up to 116 are used in the proposed YOLO model.

ShuffleNet. ShuffleNet is another neural network that focuses on mobile device platforms like smartphones, drones, and robots. It outperformed MobileNet v1 by achieving higher accuracy. It also outperformed AlexNet in terms of speed up to 13 × at comparable accuracy [15]. The architecture has a special unit called the

Fig. 4 Inverted residual block
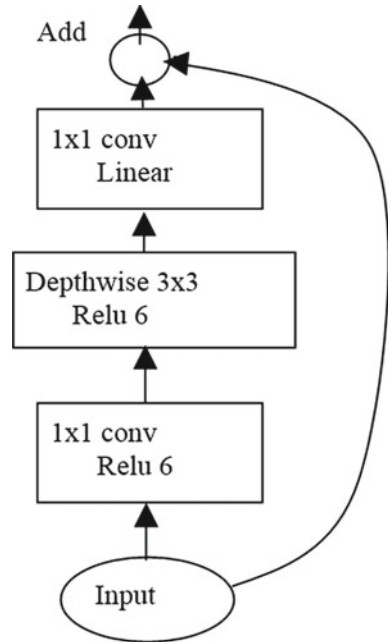
ShuffleNet unit which is a slight modification in the bottleneck block of MobileNet. ShuffleNet replaces the 1×1 convolutions with group convolutions and added a channel shuffle layer after the first 1×1 group convolution layer as shown in Fig. 5. Normally, output, and input channels are related only within their respective channel groups and block inter-channel information. But with channel shuffle operation, the convolution will have data from all the channel groups from the input resulting in fully related output channels. It has 172 layers in total, but layers up to layer 127 are considered for feature extraction in the YOLO model.

SqueezeNet. SqueezeNet aims at the compression of a neural network so that it can be deployed in embedded systems. In an embedded platform, smaller networks have some advantages: (1) During distributed training, it requires less communication across the servers. (2) It takes less bandwidth in exporting a new model from the cloud to the device. (3) It is more viable in deploying to limited memory devices like FPGA. SqueezeNet [16] achieved $50 \times$ fewer parameters and model size less than 0.5 MB at the accuracy of AlexNet on ImageNet. It has a special unit called a fire module comprising a squeeze and expand layer as shown in Fig. 6. The squeeze layer is made of 1×1 filters smaller number than expand layer, which mainly helps in limiting the number of input channels to the next layer. The expand layer comprises of 1×1 and 3×3 filters working in parallel. SqueezeNet has 68 layers in total where 65th layer is connected as an input to the detection layers.

DenseNet201. DenseNet was proposed to improve accuracy by solving the vanishing gradient in high-level neural networks. As the path from the input layer to the output

Fig. 5 ShuffleNet unit
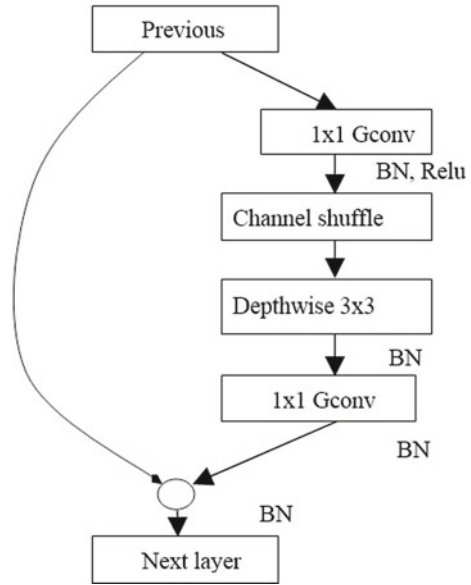
Fig. 6 Fire module of SqueezeNet
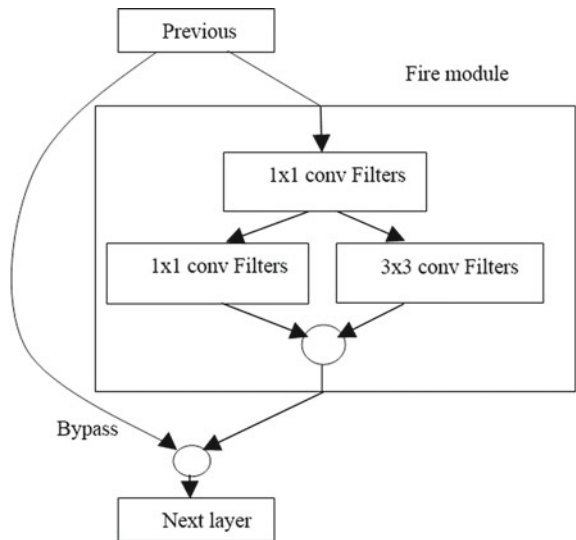
layer gets longer, the information vanishes before reaching the final layers. To overcome this problem, each layer receives additional inputs from all the preceding layers and performs depth concatenation. Unlike ResNet which has a residual connection from a specific layer and is added elementwise, DenseNet has a composition layer at the residual connection. The composition layer [17] has three layers: BN, ReLU, and 3 × 3 convolutions, which is repeated for every layer concatenation as shown

in Fig. 7. As the inputs are taken from preceding layers, the actual network looks compact and thin with only a few kernels. Such kernels will be the additional kernels in the subsequent layers maintaining a network growth rate of k. If L is the number of layers, then the DenseNet will have $L(L + 1)/2$ direct connections. If the network gets bigger and bigger, there is a change in feature map explosion. In order to solve this, the dense blocks are created such that each dense block have a specific number of layers, and the output of such dense block will be given to a 1×1 convolution layer as a transition layer which may limit the number of filters. Then, it is followed by max pooling for the reduction of feature maps size. It has a total of 708 layers, but layers up to 478 will be connected to YOLO for the proposed model.

Inceptionv3. Inceptionv3 is an optimized version of the predecessor inception version that was already very efficient. The modification of the network includes (i) factorization of large convolution kernels into smaller kernels, where 5×5 kernels are replaced with 3×3 s and 1×1 which helps in reducing the computation parameters. (ii) Kernels are factorized spatially into asymmetric kernels where the 3×3 kernel is replaced with two kernels 1×3 and 3×1. (iii) Auxiliary classifiers are added between layers during training so that the loss will be added to the loss of main network loss. (iv) Efficient grid size reduction where convolution and max pooling blocks are in parallel and later concatenated [18] as shown in Fig. 8. If there are D

Fig. 7 Dense block of 4 layers with a growth rate of 3
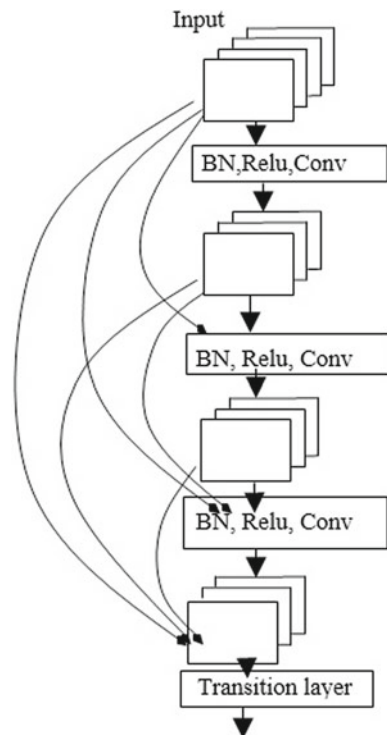
Fig. 8 Inceptionv3 module
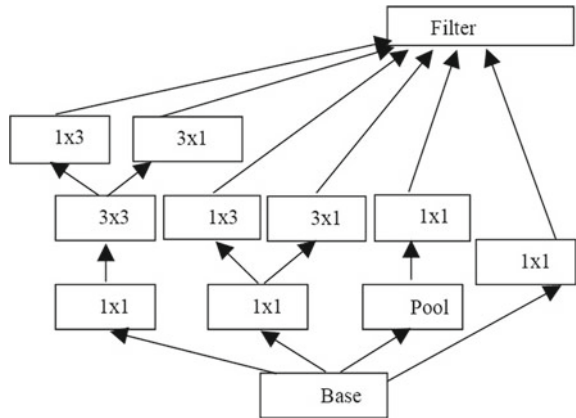
$\times$ *D* grids with k filters, then after reduction, it gives $D/2 \times D/2$ with 2 k filters. Though Inceptionv3 has 42 layers which are higher than its previous versions, it has lower computation cost and more accuracy. For the YOLO model, 229th layer from inceptionv3 network is connected as an input to the YOLO layers.

3   Experimental Results

The output of the YOLOv2 detection model based on ResNet50 is shown in Fig. 9. Objects in the images are detected and classified. The box indicates the location of the detected object with a confidence score and the color of the box indicates the class of the object. E.g., the left side image has a detected object which is 96.63% likely to be a vehicle. Detection of multiple objects in a single image is shown in the right side figure with different confidence scores.

An experiment is also conducted by taking the output of the feature extractor network, one at a size close to 13×13 and another at the lower layer close to 7×7 of the original network. It is found that the model whose feature output layer is taken at a size around 13×13 performs better than those taken at around 7×7 which is much deeper in the network as given in Table 1. As a reason, it can be concluded that features from layers much further than 13×13 do not give a good result; hence, it is better to take an output size close to 13×13.

The average precisions of the eight YOLOv2 models are presented in Table 2. MobileNetv2 is found to be having highest precision among them by achieving 81.64% with a model size of 23.6 MB. ShuffleNet being the compressed version of MobileNetv2 has the lowest size of 6.16 MB with an average precision of 75.30%. As observed in the bar graphs shown in Fig. 10, the ShuffleNet has a good trade-off between model size and precision.

Fig. 9 Vehicle detection with YOLOv2 model

Table 1 Comparison based on feature extraction output layer taken at different sizes

Output layer	ResNet 50	MobileNet v2	DenseNet 201
Size around 13×13	80.81	81.64	80.67
Size around 7×7	64.80	58.38	53.25

Table 2 Average precisions and model sizes of the eight YOLO detection models based on eight different types of networks

Sl. no.	YOLO network	Average precision (in %)	Model size (in MB)
1	ResNet50	80.81	97.8
2	AlexNet	74.45	12.5
3	GoogleNet	79.11	56.9
4	MobileNetv2	81.64	23.6
5	ShuffleNet	75.30	6.16
6	SqueezeNet	68.76	19.5
7	DenseNet201	80.67	92.2
8	Inceptionv3	75.54	69.9
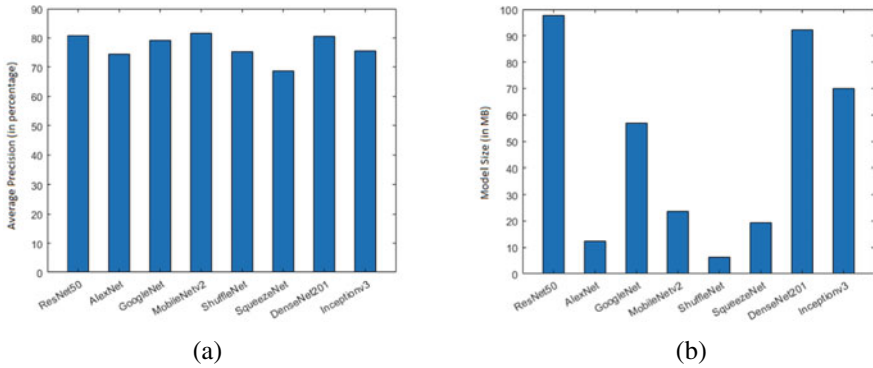
Fig. 10 Results in bar graph representation. **a** Average precisions for the proposed detection models. **b** Trained model sizes of the proposed detection models

4 Conclusion

In this paper, we have implemented various object detection models based on YOLOv2 by using various types of pretrained networks. After training the detection models on a common dataset of vehicle images, they are able to detect the vehicles in present in the test images. The average precisions and model sizes are compared as a performance parameters. Some models are found to have higher average precisions by having bigger model size and vice versa. So, there's a trade-off between model size and average precision. It is found that ShuffleNet gives the lowest-sized trained model with considerably good average precision, such a model will be most suited when deployed in embedded systems and mobile devices platforms where the resources are limited.

References

1. Redmon J, Divvala S, Girshick R, Farhadi A (2015) You only look once: unified, real-time object detection
2. Redmon J, Farhadi A (2016) YOLO9000: better, faster, stronger
3. Loey M, Manogaran G, Taha MHN, Khalifa NEM (2021) Fighting against COVID-19: a novel deep learning model based on YOLO-v2 with ResNet-50 for medical face mask detection. Sustain Cities Soc 65. https://doi.org/10.1016/j.scs.2020.102600
4. Itakura K, Hosoi F (2020) Automatic tree detection from three-dimensional images reconstructed from 360 spherical camera using YOLO v2. Remote Sens (Basel) 12
5. Huang S, He Y, Chen XA (2021) M-YOLO: a nighttime vehicle detection method combining mobilenet v2 and YOLO v3. In: Journal of physics: conference series. IOP Publishing Ltd. https://doi.org/10.1088/1742-6596/1883/1/012094
6. Ren P, Wang L, Fang W, Song S, Djahel S (2020) A novel squeeze YOLO-based real-time people counting approach. Int J Bio-Inspired Comput 16:94–101. https://doi.org/10.1504/ijbic.2020.109674

7. Fang W, Wang L, Ren P (2020) Tinier-YOLO: a real-time object detection method for constrained environments. IEEE Access 8:1935–1944

8. Zhao T, Yi X, Zeng Z, Feng T (2021) MobileNet-Yolo based wildlife detection model: a case study in Yunnan Tongbiguan Nature Reserve, China. J Intell Fuzzy Syst 41:2171–2181. https://doi.org/10.3233/JIFS-210859

9. Shoaib M, Sayed N (2022) YOLO object detector and inception-V3 convolutional neural network for improved brain tumor segmentation. Traitement du Signal 39:371–380. https://doi.org/10.18280/ts.390139

10. Li G, Song Z, Fu Q (2018) A new method of image detection for small datasets under the framework of YOLO network. In: Proceedings of 2018 IEEE 3rd advanced information technology, electronic and automation control conference, IEAEC 2018. Institute of Electrical and Electronics Engineers Inc., pp 1031–1035

11. https://www.mathworks.com/supportfiles/vision/data/yolov2ResNet50VehicleExample_19b.mat

12. He K, Zhang X, Ren S, Sun J (2015) Deep residual learning for image recognition

13. Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A (2014) Going deeper with convolutions

14. Sandler M, Howard A, Zhu M, Zhmoginov A, Chen LC (2018) MobileNetV2: inverted residuals and linear bottlenecks

15. Zhang X, Zhou X, Lin M, Sun J (2017) ShuffleNet: an extremely efficient convolutional neural network for mobile devices

16. Iandola FN, Han S, Moskewicz MW, Ashraf K, Dally WJ, Keutzer K (2016) SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5MB model size

17. Huang G, Liu Z, van der Maaten L, Weinberger KQ (2016) Densely connected convolutional networks

18. Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna Z (2015) Rethinking the inception architecture for computer vision

Data Transmission and Optimization of Energy in Smart Campus Using LoRaWAN Industrial IoT Technology

Ramasamy Mariappan, Ch. S. V. N. S. L. Amulya, M. Yogisri Vasanthi, Pothapu Aditya, and Ch. Sai Manohar

1 Introduction

It was estimated that nearly 45.12% people all over the world with 3.5 billion mobiles are eager to remotely control their appliances using their mobiles [1]. They are interested in controlling and monitoring their appliances at a low power and low cost, in a hazardless way from a long distance. The objective is to show a practical solution of design and implementation, which enables the users to easily build a private LoRa network [2] for several applications. This can be achieved by bringing forward a better software architecture of LoRa network server. LoRaWAN technology is used widely for Low-Power Wide-Area Networks (LPWANs). It offers coverage for low-power gadgets, making it a perfect communication system for IoT applications.

LoRa is better than already existing technologies [3] like Zigbee or other Bluetooth technologies which control smart appliances within a short range such as smart fans, smart lights, etc., which are readily available in the market. LoRa is better as it can be implemented by linking the existing electrical appliances with the LoRa system to get the same efficient result by increasing the controlling range to several kilometers. Hence, it is not necessarily needed to dispose the existing appliances and purchase more expensive smart appliances to have a smart campus. LoRa technology is currently being used for many applications worldwide. It is ideally used in applications where small chunks of data need to be transmitted with low bit rates. There are few license-free bands, but 868 MHz is the popular band in Europe region. Europe band is being used for our application which is said to be 863–870 MHz.

R. Mariappan (✉)
School of Computer Science and Engineering, Vellore Institute of Technology, Vellore 632014, Tamil Nadu, India
e-mail: prof.mariappan.r@gmail.com

Ch. S. V. N. S. L. Amulya · M. Y. Vasanthi · P. Aditya · Ch. S. Manohar
 Dept of ECE, VR Siddhartha Engg College, Vijayawada, India
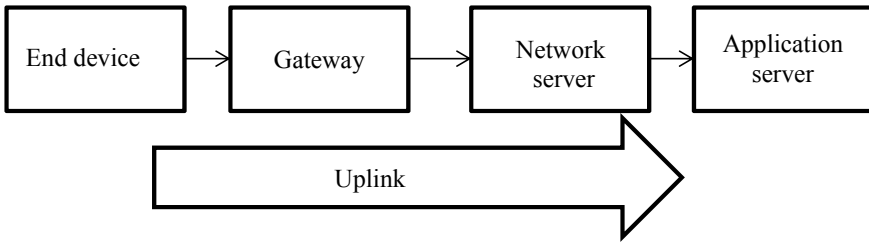
Fig. 1 Uplink transmission
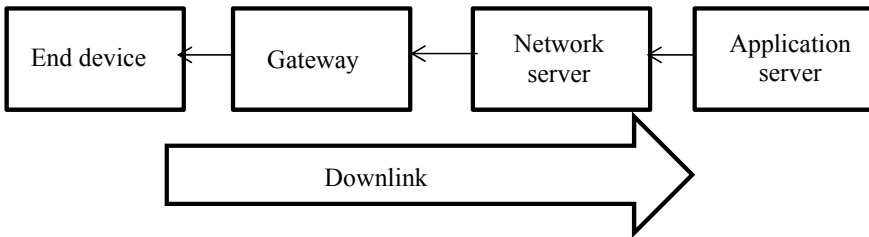
Fig. 2 Downlink transmission

The major components in any LoRa-based network include end device, gateway, network server, and application server. All the components communicate with each other. End device sending data to gateway is called uplink transmission and gateway sending data to end device is called downlink transmission as shown in Figs. 1 and 2.

1.1 End Device

Dragino LoRa Sensor Node (LSN50) is the end device being used. Usually there are 3 types of end devices: class A, class B, and class C [4]. The class of the device will decide when it will receive downlinks. LSN50 is class A by default. It consumes the lowest energy but take long intervals of time between uplinks.

1.2 Gateway

LPS8 is the LoRaWAN open-source gateway being used. A gateway is used as a pass over between a LoRa wireless network and an Internet Protocol (IP) network. The LoRa gateway is used to send the sensor data from electrical device to the cloud/server, i.e., application server. Configuration and registration of gateway must be

done in order to access the gateway. For this, the IP address of gateway must be known. Here, 'Fallback IP Address' method is used. The gateway is registered in The Things Network (TTN) using the unique gateway EUI ID. The data can be seen in the TTN only when status of gateway is 'Connected'.

1.3 Network and Application Server

The main role of a server is to receive, store, and send data within a network. The network server connects end devices, gateways by ensuring reliable and secure data transfer all over the LoRaWAN network. The data from end devices is processed by application server and it is also responsible for uplink data decryption, decoding, downlink queuing, downlink data encoding, and encryption.

2 Literature Survey

1. An idea is presented regarding smart automation using LoRa technology. Components used are mobile phone, ESP32 at the transmitter end. LoRa module and ESP32 at the receiver end. They have suggested a system with adaptable design evident to be very effectual in monitoring and controlling of home appliances. Its main limitation was the loss of data during transmission, especially to a farther distance, which could be dealt with by using several gateways.
2. They highlight the design and implementation of a realizable solution, providing its users to build a private and secured LoRa network for diverse IoT applications. The LoRa network server is categorized into four separate modules below the architecture. The data-streaming-based messaging system is used for inter-module interaction to ensure scalability and flexibility.
3. They have provided an outline of research published between 2015 and September 2018 and available through the IEEE Explore and Google scholar databases. A comprehensive tutorial on the introduction to the LoRa and LoRaWAN functions is given. In addition to this literature review of the LoRaWAN, other suggested solutions that significantly expand the LoRaWAN standard are being explored.
4. The literature survey explores the scalability challenges faced when deploying LoRaWAN in ultra-dense IoT networks where there is a need for massive LoRa connectivity. This survey also provided compelling research insights to help researchers design efficient and scalable LoRaWAN systems.
5. They provided an analysis of five LPWAN driver technologies, including NB-IoT, Telensa, Ingenu (RPMA), SigFox, and LoRa/LoRaWAN. They categorized recent attempts to enhance LoRa/LoRaWAN performance in the matter of power usage, pure data extraction rates, network scalability, network coverage, security

and service quality. Providing possible methods by analyzing recent approach solutions.

This paper aims at monitoring and controlling of appliances in campus. To enable long range, CSS modulation [5] is used. CSS modulation enables transmission and reception of data over long distances. It is robust against radio frequency (RF) noise. The gateway LPS8 falls under class A mode. Class A mode consumes the minimum power possible. So, our work provides a feasible solution for optimization of energy along with monitoring and controlling of appliances in campus remotely over long distances using LoRaWAN IoT technology (Table 1).

3   Proposed System

The system is proposed so that a user should be able to monitor and control the electronic appliances of a room from a remote place with the help of a web application. The monitoring and controlling of appliances are done remotely with the support of a gateway. The entire proposed system is split into two modules. They are transmitter module and receiver module.

Transmitter Module: The transmitter module comprises of LoRa shield with antenna, Arduino Uno Rev3 and sensors (DHT11, ultrasonic, PIR, ALS). The sensors are connected to the LoRa module. Sensors read the data like temperature, distance, light intensity, motion, etc., and transmit the data in the form of packets. These sensors are connected in the room which is intended to control by the user.

Receiver Module: The receiver module comprises of relay module which acts as an electrical switch to control appliances.

Our work is 'Smart Campus' which is an application of industrial LoRaWAN IoT technology. Here the sensors are connected in the classrooms, laboratories, office rooms, etc. These sensors read the data in the room. The relay modules are connected in the respective rooms which are used for ON/OFF controlling of appliances like lights, fans, and air-conditioners. A web application is developed for the monitoring and controlling of the appliances. So, the user would be able to see the sensor readings in the web application from a remote place (within the range of gateway). This system helps to reduce the power consumption and helps for optimization of energy in campus module.

3.1   Circuit Explanation

The circuit is divided into three parts: transmitter module, receiver module, and the dragino gateway. The circuit diagram is shown in Fig. 3.

The transmitter module contains a LoRa shield with antenna, Arduino Uno board and sensors (DHT11, ultrasonic, PIR, ALS). The sensors are connected to the LoRa

Table 1 Literature survey

Title	Methodology	Components	Advantages	Disadvantages
[1] Smart Monitoring and Controlling of Appliances Using LoRa-Based IoT System (9 March 2021)	LoRa-based automation system For long-range communication, it is a smart IoT system with LoRa wireless communication module and ESP32 module	• ESP32 • LoRaWAN • RYLR896 transceiver • Sensors (DHT11, Flame sensor, MQ2, Laser sensor)	LoRa has the capability to cover an extensive area using low-power economical wireless connectivity	Non-uniform division of the data transmissions caused by nodes, Seasonal effects
[2] Design and Implementation of Open LoRa for IoT (July 22, 2019)	Connector, Join Server, Central Server, and Network Controller are the four modularized and low-coupling modules into which the LoRa Network Server is divided	• LoRaWAN • SX1255 transceivers • SX1301 digital baseband processor	High performance and flexibility	-
[3] A Survey of LoRaWAN for IoT: From Technology to Application (November 16, 2018)	Frequency Shift Chirp Modulation (FSCM) Adaptive Data Rate (ADR) mechanism LoRa PHY error model	• LoRaWAN	For the application cases which have communication asymmetry, LoRaWAN is useful	LoRaWAN is vulnerable to mobility, it worsens the performance of end nodes when the reception conditions are bad

(continued)

Table 1 (continued)

Title	Methodology	Components	Advantages	Disadvantages
[4] A Survey on Scalable LoRaWAN for Massive IoT: Recent Advances, Potentials, and Challenges (February 22, 2022)	The LoRaWAN network consists of 3 main components: Network servers (NS), Eds, and GWs; which feature in a star-of-star topology	• LoRaWAN • NB-IoT • LPWAN	Focuses on LoRaWAN scalability issues facing its deployment for massive IoT	Energy utilization is the principal issue facing the effective deployment of CSMA methods in LoRaWAN networks
[5] A Survey on LoRaWAN Technology: Recent Trends, Opportunities, Simulation Tools, and Future Directions (January 5, 2022)	NB-IoT, Telensa, SigFox, Ingenu (RPMA), and LoRaWAN are said to be the five selected driving LPWAN technologies	• IoT and M2M devices • Gateways (GW) • End devices (ED) • Network Server (NS)	LoRaWAN proposes many benefits in terms of low bit rate, energy consumption, wide coverage, simplicity, and ease of management	Network administration, high-density LoRaWAN installations and device interoperability, scalability in a large-scale
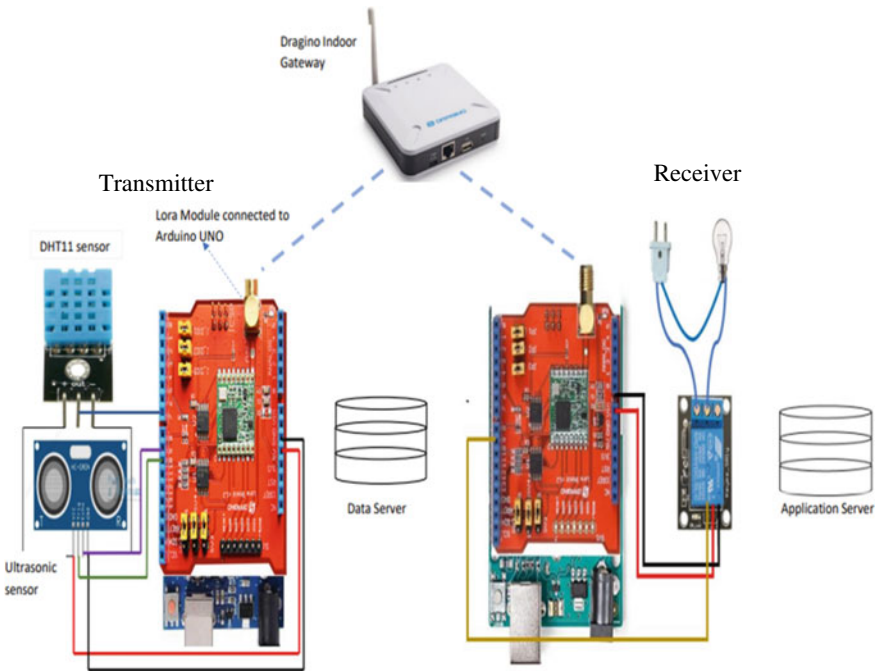
Fig. 3 Circuit diagram

module. This setup is placed in the room which is planned to control. The LoRa shield is used to transmit data over long distances at lower data rates. The signal is subjected to Chirp Spread Spectrum (CSS) modulation which makes the signal to transmit over a long distance by making the signal robust to noise. The data transmitted is stored in the data server.

When a user performs an action through the web application then the relay in the receiver module executes the switching operation of the respective appliance. LoRa gateway is used to transmit sensor data to the data server or cloud. This data is used by the user to monitor the appliances. The data server, TTN is used as application server.

3.2 Methodology

Chirp spread spectrum (CSS)

LoRa exclusively uses a modulation scheme called Chirp Spread Spectrum (CSS). Spread spectrum means the signal is spread in the frequency domain. The main use of spread spectrum modulation is to increase the bandwidth of input signal. CSS modulation scheme uses chirp pulses to encode the information as shown in Fig. 4. Chirp pulse is a sinusoidal signal. Chirp's frequency increases or decreases with time and it is also known as sweep signal. There are two types of chirp pulses called up-chirp and down-chirp.

The advantages of CSS modulation are:

1. Resistant to doppler effect
2. Robust against RF noise
3. Encryption is easier
4. Transmission and reception over long distances.
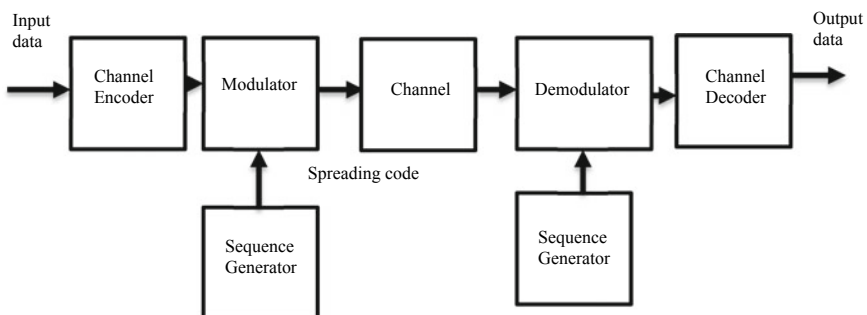
Fig. 4 Block diagram for CSS modulation
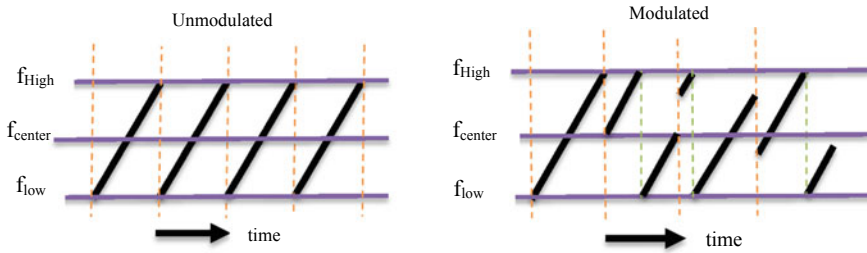
Fig. 5 Unmodulated signal and modulated signal

Block diagram

The relation between symbol rate, bandwidth, chirp rate, and spreading factor is given by

$$\text{Symbol rate} = \frac{\text{BW}}{2^{\text{SF}}} \tag{1}$$

$$\text{Chirp rate} = \text{Bandwidth} \times \text{Symbol Rate} = \text{BW} \times \frac{\text{BW}}{2^{\text{SF}}} \tag{2}$$

Encoding of information onto the carrier signal is done by the end node which has a radio module. Transmission and reception of this signal is done by a gateway. The radio module present at the gateway decodes the modulated signal and extracts information. Chirp signals are used as carrier signals for encoding.

Modulation means how bits are modified to RF energy. The signal which comes after subsequent up-chirps followed by two down-chirps is identified as the message signal. Figure 5 shows the unmodulated and modulated chirp signals.

The LoRa transmission efficiency is greatly affected by the modulation process. There are four parameters that effect the modulation process. They are maximum frequency, minimum frequency, input signal, and spreading factor [6]. Bandwidth depends on both maximum and minimum frequencies.

The relation between maximum rates of information transmitted when noise is present over a channel of defined bandwidth is explained by the Shannon-Hartley theorem.

The maximum rate of transmitted information is called the channel capacity in bits per second denoted by C and the relation is given in (3).

$$C = B \times \log_2\left(1 + \frac{S_p}{N_p}\right) \tag{3}$$

where B is the bandwidth of channel; S_p, N_p are average signal power and average noise power, respectively. The ratio of S_p, N_p is called Signal to Noise Ratio (SNR).

$$\frac{C}{B} = 1.433\left(\frac{S}{N}\right) \tag{4}$$

$$\frac{C}{B} \sim \frac{S}{N} \tag{5}$$

From Eq. (5), it can be said that on increasing the channel bandwidth, channel capacity increases. Also, in order to obtain higher channel capacity, SNR must be increased. SNR and throughput possess a linear relation which means on increasing SNR, greater throughput can be achieved.

Latency is another factor that effects the performance of LoRa transmission. It is the time taken to transmit the data and receiving the response. It should be maintained minimum.

To compute data rate or bit rate R_b:

$$R_b = \text{SF} \times \frac{\text{BW}}{2^{\text{SF}}} \times \frac{4}{(4 + \text{CR})} (\text{bits/sec}) \tag{6}$$

Band Width (BW) in Hertz.
Spreading Factor (SF): 7 to 12.
Code Rate (CR): 1 to 4.

If bandwidth increases, bit rate or data rate increases. If spreading factor increases, bit rate or data rate decreases. If SF increases by 1 unit, SNR changes by -2.5 db.

The message 'b' consists 'L' identically distributed bits. Channel coding sets up correlation between bits. Reducing this correlation performs an XOR of 'b^c' with a whitening sequence 'w'. The interleaver's output is a deterministic shamble of the bits in 'b^w' which is represented as 'b^i'. The groups of spreading factor bits are made, yielding a sequence of symbols 'c'. 'c' is translated into signal s(t) using CSS modulator, which is viewed as a corrupted signal r(t) as in Fig. 6.
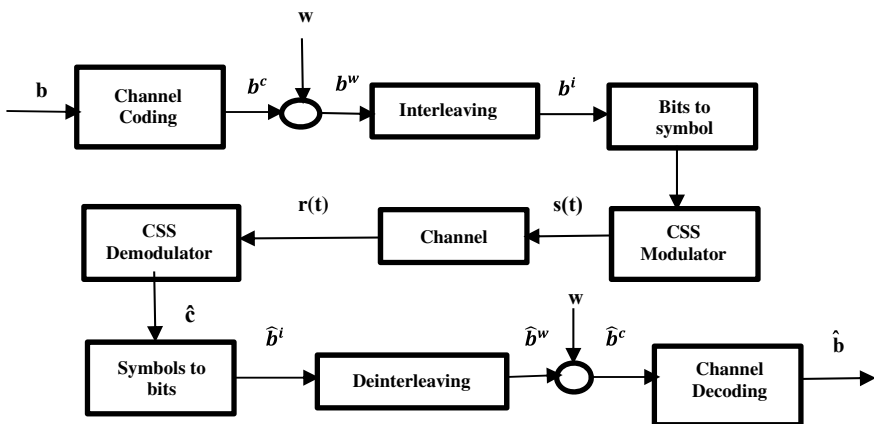
Fig. 6 Architecture of LoRa transceiver

Fig. 7 Transmitter end
circuit connection
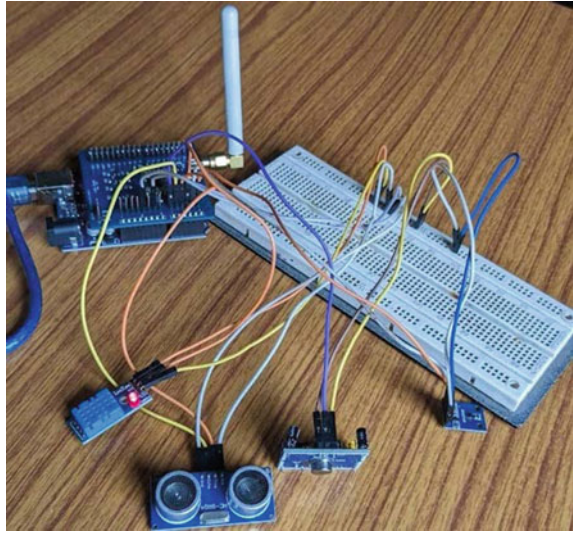
3.3 Data Transmission

The transmitter end connection diagram is shown in Fig.7.

The process of sensor data transmission through LoRa shield is given in Fig. 8. in the form of flowchart.

LoRa.h library is used to initialize LoRa and transmit data using radio frequencies. Serial Peripheral Interface (SPI.h) library is used to communicate with multiple peripheral devices quickly over short distances. CSS modulation uses orthogonal spreading factors which enables the network preserve battery life. Here the input signal is multiplied with a pre-defined code called spreading code which is at higher rate which enables to transmit over longer ranges. Each bit in the spreading code is called a chirp. LoRaWAN makes use of unlicensed radio spectrum. The frequency used is 866 MHz. In process to setup LoRa, if the frequency is not equal to 866 MHz, then the starting of LoRa will be failed. After reading sensor data, the data packets will be transferred via LoRa with the help of LoRa beginpacket() and the transfer of sensor data packets is stopped with the help of function LoRa endpacket().

4 Results and Discussion

Lower the SF, shorter the distance communications, it leads to higher data rates, lesser latency (time on air), and lesser energy consumption. Decreasing the SF by 1 will half the energy consumption, half the air on time and double the data rates and vice versa (Figs. 9, 10, 11 and 12).

Fig. 8 LoRa transmitter flow chart
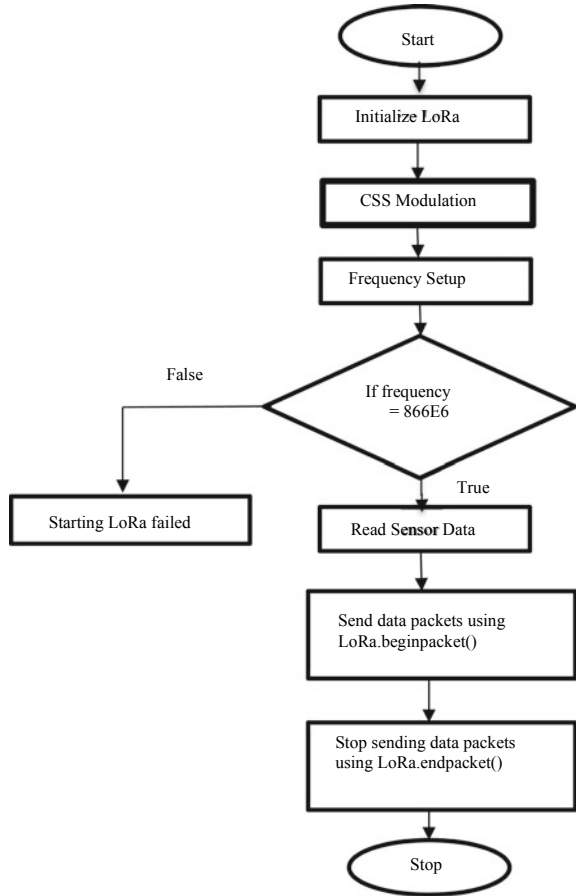
Fig. 9 DHT11 transmitted data


```
COM8

LoRa Transmitter
Distance is: 2 inches
Distance is: 2 inches
Distance is: 2 inches
Distance is: 1 inches
Distance is: 5 inches
Distance is: 7 inches
Distance is: 9 inches
Distance is: 81 inches
Distance is: 81 inches
Distance is: 81 inches
Distance is: 81 inches
Distance is: 81 inches
Distance is: 81 inches
Distance is: 81 inches
Distance is: 82 inches
Autoscroll  Show timestamp
```

Fig. 10 Ultrasonic transmitted data


```
COM8

LoRa Sender
Sending packet:
Humidity: 68% Temperature:29.70C

Sending packet:
Humidity: 68% Temperature:29.70C

Sending packet:
Humidity: 68% Temperature:29.80C

Sending packet:
Humidity: 68% Temperature:29.80C

Autoscroll  Show timestamp
```

Fig. 11 PIR transmitted data

```
COM8

Motion Ended!
Motion Detected
Motion Ended!
Motion Detected
Motion Ended!
Motion Detected
Motion Ended!
Motion Detected
Motion Ended!
Motion Detected
Motion Ended!
Motion Detected
Motion Ended!

☑ Autoscroll ☐ Show timestamp
```
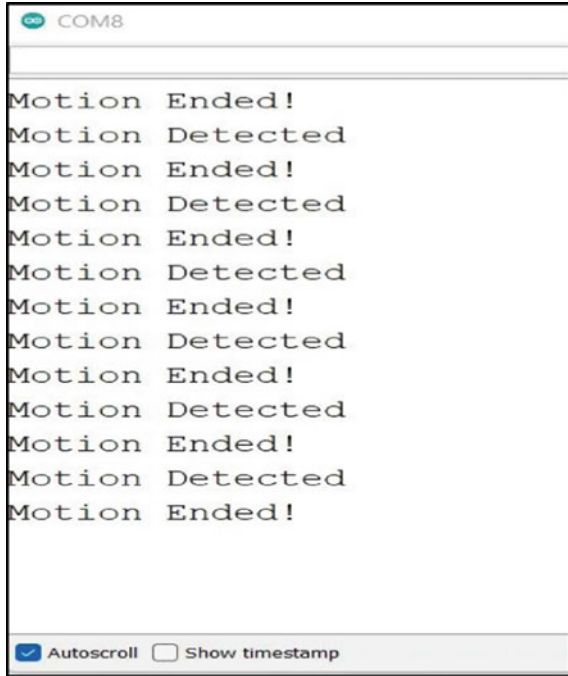
Fig. 12 ALS transmitted data

```
COM8

LoRa Sender
Light = 69.17 lx

Light = 70.83 lx

Light = 72.50 lx

Light = 75.00 lx

Light = 88.33 lx

Light = 500.00 lx

Light = 650.00 lx

Light = 555.00 lx
```
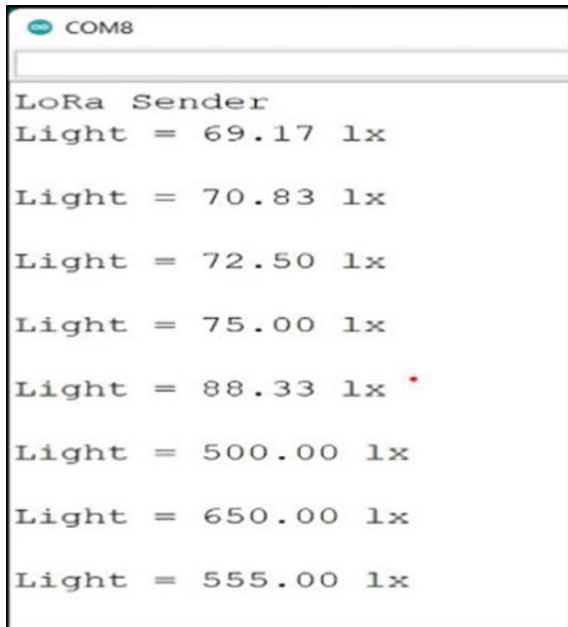
5   Conclusion

The LoRaWAN has many advantages which includes low-power consumption, long-range transmission, secure connectivity, possibility of private deployment geolocation without GPS and it is open accessible. In this paper, an intelligent automation system is discussed for remote control of appliances in home, institute or industry using LoRaWAN IoT technology. The system can monitor and control from 3 to 12 km using LoRa-based wireless communication. By continuously monitoring the sensor data one can get to know the status of appliances connected and these appliances are controlled from remote distances.

References

1. Nur-A-Alam, Ahsan M, Based MA, Haider J, Rodrigues EMG (2022) Smart monitoring and controlling of appliances using LoRa based IoT system. J MDPI-Des
2. Zhou Q, Zheng K, Hou L, Xing J, Xu R (2019) Design and implementation of open LoRa for IoT. In: Conference paper, IEEE
3. Haxhibeqiri J, De Poorter E, Moerman I, Hoebeke J (2018) A survey of Lora WAN for IoT: from technology to application. J MDPI-Sens
4. Jouhari M, Amhoud EM, Saeed N, Alouini MS (2022) A survey on scalable Lora WAN for massive IoT: recent advances, potentials and challenges. In: Conference paper, arXiv
5. Almuhaya MA, Jabbar WA, Sulaiman N, Abdulmalek S (2022) A survey on LoRaWAN technology: recent trends, opportunities, simulation tools and future directions. J MDPI-Electron
6. Reynders B, Pollin S (2016) Chirp spread spectrum as a modulation technique for long range communication. In: IEEE symposium on communications and vehicular technology

An Implementation of Differential Difference Voltage Difference Transconductance Amplifier (DD-VDTA) and Its Application as a Dual Output Integrator

Prerna Rana and Ashish Ranjan

1 Introduction

In recent time, a wide variety of active elements with different characteristics is enriched in the scientific literature based on fundamental active block [1]. These designs are based on the modification of the basic elements like Voltage Feedback Amplifier (VFA) [2], Current Feedback Amplifier (CFA) [3], Operational Transconductance Amplifier (OTA) [4], and Current-Conveyors (CC) [5] which fulfils the modern circuit design speciality like low power consumption, high slew rate, few more [6, 7]. A very popular active element "VDTA" [8] exhibits a simple design approach for the generation of integrators, bi-quad filters, oscillator, inductance, and frequency dependent negative resistor (FDNR) simulator using a single active block with a few passive elements. This design can be compared to the active elements like second generation Current Conveyor (CC-II) [9], third generation Current Conveyor (CC-III) [10], Differential Difference Current Conveyor (DDCC) [11], Differential Voltage Current Conveyor (DVCC) [12], Current Differencing Buffered Amplifier (CDBA) [13], second generation Dual X Current Conveyor (DXCC-II) [14], and offers a better results in comparison to the [9–14]. In addition to this, this VDTA can also be used for transconductance mode applications as its input ports is voltage and output is current. Moreover, integrator design [15–18] and Wave Active Filter (WAF) [19–21] are also available. By keeping in line of the features of VDTA, the authors have made an attempt to design an advance version of VDTA with addition differential difference property and collectively called as "DD-VDTA". This research paper contributes a design for DD-VDTA which enables an extra features of differential

P. Rana (✉) · A. Ranjan
Department of Electronics and Communication, National Institute of Technology, Manipur 795004, India
e-mail: ranaprerna135@gmail.com

difference property and owes all the good characteristics of VDTA block with its circuit description in Sect. 2. The proposed DD-VDTA design is translated for the active integrator for lossy, lossy subtraction and lossless in Sect. 3. The workability test of DD-VDTA is examined in both DC and AC conditions in Sect. 4. In addition, simulation results of integrator are also verified and ends with a comparative table. Moreover, this proposed block can be used to synthesize integrators having two outputs, i.e. inverting and non-inverting output through the same design and to design WAF with lower number of active and passive elements.

2   Circuit Description of DD-VDTA

The proposed new active element DD-VDTA is a combination of DDA [22] and VDTA [23] which offers the features of DDCC and an OTA in a single circuit design. The CMOS structure of DD-VDTA is shown in Fig. 1.

The characteristics equation of DD-VDTA can be characterized as:

$$I_z = g_{m1}(V_p - V_n),\tag{1}$$

$$I_{xn} = -g_{m2}V_z,\tag{2}$$
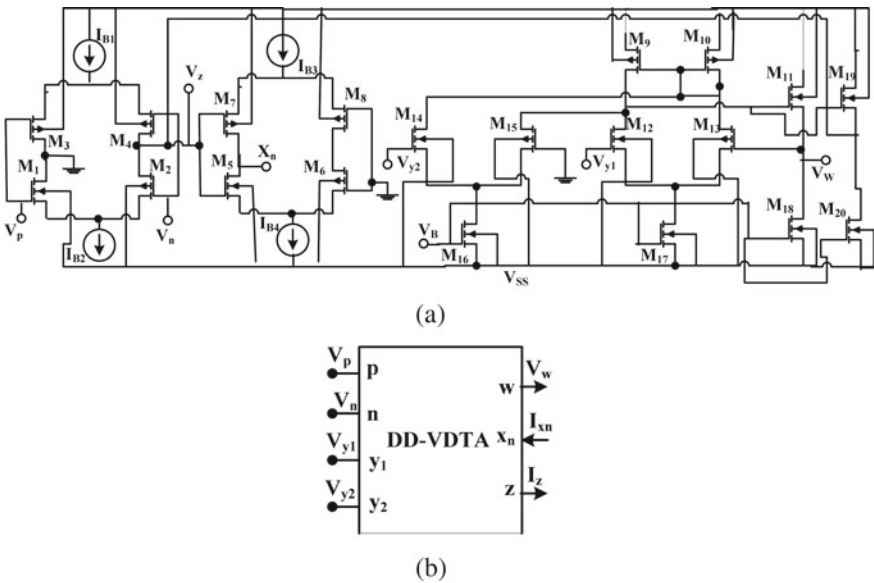
Fig. 1 DD-VDTA a CMOS structure b block diagram

$$I_w = I_z, \tag{3}$$

$$V_w = V_{y1} - V_{y2}, \tag{4}$$

First stage consists of VDTA having transistors M_1 to M_8, where the output and input stages are made up of floating current sources. It has two input (V_p, V_n) and two output (V_z, I_{xn}) ports of high impedances. Equation 1, indicates the difference of the input voltages of ports p and n with transconductance gain g_{m1} which produces current at port z. The voltage drop at port z with transconductance parameter (g_{m2}) is transformed into output current at port x_n. The first stage of the proposed DD-VDTA block comprises two Arbel–Goldminz transconductance-based OTAs [23] and their transconductance (g_{m1} and g_{m2}) can be mathematically obtained as:

$$g_{m1} = \frac{g_3 + g_4}{2}, \tag{5}$$

$$g_{m1} = \frac{g_7 + g_8}{2}, \tag{6}$$

$$g_i = \sqrt{I_{bi} \mu_i C_{ox} \frac{W_i}{L_i}} \tag{7}$$

where g_i can be defined as the transconductance value of the ith transistor as a function of MOSFET fundamental electrical parameters (μ_i, C_{ox}, W/L) and bias current (I_{bi}). Second stage is DDA that consists of transistors M_9 to M_{20}. It consists of two high impedance input port (Y_1, Y_2) and one low impedance output port (V_w). There are two differential stages (M_{14}, M_{15}) and (M_{12}, M_{13}) with a PMOS (p-type metal oxide semiconductor) current mirror employing M_9 and M_{10}. This current mirror converts the differential current to output voltage (V_w) at the drain terminal of M_{11}. This part gives the difference of input voltages present at ports y_1 and y_2 through the output port z. The transistors M_{19} and M_{20} makes a current mirror pair in which the current flowing through w port and z port are same (Eq. 3). The block diagram of DD-VDTA is given in Fig. 1b.

3 Applications of DD-VDTA

The proposed DD-VDTA is useful for the design of lossy integrator, lossy integration subtraction, and lossless integrator circuit as shown in Fig. 2a–c, respectively. The property of DD-VDTA block supports the differential difference of two voltage inputs with characteristics of a VDTA circuit. Hence, it can be easily used to synthesize integrators with two outputs, i.e. inverting and non-inverting output through the same design. These designs are purely resistor-less and require only one passive capacitor.
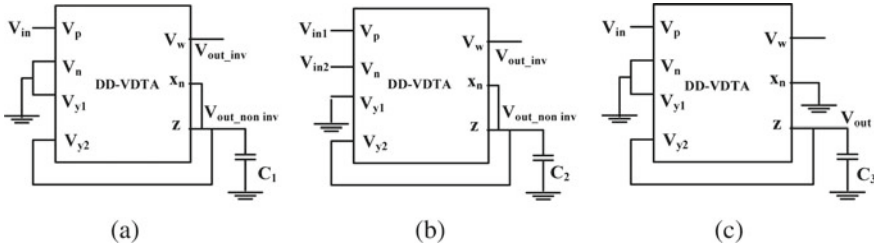
Fig. 2 Integrator designs using DD-VDTA **a** lossy integrator **b** lossy subtraction integrator **c** lossless ntegrator

For lossy integrator circuit, the ports, V_{y2}, z, and x_n are shorted to each other and a grounded capacitor is connected across it. In this circuit, input is given at only one port (V_p). The inverted and non-inverted outputs are taken from w and z ports respectively and it yields;

$$v_z = \frac{v_p}{1 + sC/g_m} \tag{8}$$

and,

$$v_w = -\frac{v_p}{1 + sC/g_m} \tag{9}$$

For lossy subtraction integration circuit, the circuit design is similar to the lossy integrator circuit. The only difference is inputs at both the ports V_p and $\underline{V}_n$ and the output is differential difference integration output as:

$$v_z = \frac{v_p - v_n}{1 + sC/g_m} \tag{10}$$

and,

$$v_w = -\frac{v_p - v_n}{1 + sC/g_m} \tag{11}$$

For lossless integrator, V_{y2} and z port are shorted together and a grounded capacitor is connected across it. The input is given at port V_p and output is taken from V_z. It results the following voltage expression for V_z and V_w respectively as:

$$v_z = \frac{v_p}{sC/g_m} \tag{12}$$

and,

$$v_{\mathrm{w}} = -\frac{v_{\mathrm{p}}}{sC/g_m} \tag{13}$$

4 Simulation Results

To check the working behaviour of DD-VDTA, the CMOS structure of DD-VDTA is simulated in the PSPICE simulator using 0.18 µm technology parameter. The aspect ratios of the transistors are given in Table 1 with supply voltages ± 0.9 V and the bias current (I_{b1}, I_{b2}, $I_{\mathrm{b3,}}$ and I_{b4} are chosen 150 µA). The transconductance values of g_{m1} and g_{m2} are 627 µA/V.

The DC characteristic of the proposed DD-VDTA is shown in Fig. 3. To plot $\left(I_z = g_{m1}\left(V_{\mathrm{p}} - V_n\right)\right)$, a resistor is connected across the z port and the respective voltages of ports p, n, and z are plotted in Fig. 3a and b. In similar manner Eq. 2 is also plotted and given in Fig. 4c. Equation 3, is represented in Fig. 4d by giving an input current at port z. and in the similar way Eq. 4 is shown in Fig. 4e and f, which are simply plotted by giving an input voltage at y_1 and y_2 port.

In addition to DC characteristics, the AC transfer characteristic of DD-VDTA is also observed in Fig. 4. The frequency response of transconductance gain g_m from port z to v_{p}, from port z to v_n, and from z to v_x is given in Fig. 4a–c respectively followed by V_{y1}/V_w and V_{y2}/V_w is shown in Fig. 4d and e, respectively. The points X and Y in the graphs shows the parasitic gain and the bandwidth of the block.

The performance parameters and their numerical value of the proposed block DD-VDTA is also well observed in Table 2.

Finally, the simulated output of the proposed integrator circuits Fig. 2a–c is shown in Fig. 5a–c, respectively. The capacitor values taken for lossy integrator simulation C_1, C_2 is 250 nF and for lossless integrator simulation C_3 is 180 nF. For the lossy integrator circuit, one square shape input of 200 mV amplitude is applied at the input port and an integrated triangular shape output of 200 mV amplitude is obtained. On the other side, lossy subtractor integrator, uses two square signals of amplitude 300 and 200 mV at two input ports (V_{p} and V_n) which results subtraction integration output of 1 mV output is at V_{w} and V_{z} ports.

Table 1 Aspect ratios (W/L) for DD-VDTA

S. no.	MOSFETs	W/L (in µm)
1	M_1, M_2, M_5, M_6	3.6/0.36
2	M_3, M_4, M_7, M_8	16.64/0.36
3	$M_9, M_{10}, M_{11}, M_{19}$	3.6/0.18
4	$M_{12}, M_{13}, M_{14}, M_{15}$	0.72/1.08
5	$M_{16}, M_{17}, M_{18}, M_{20}$	2.16/1.08
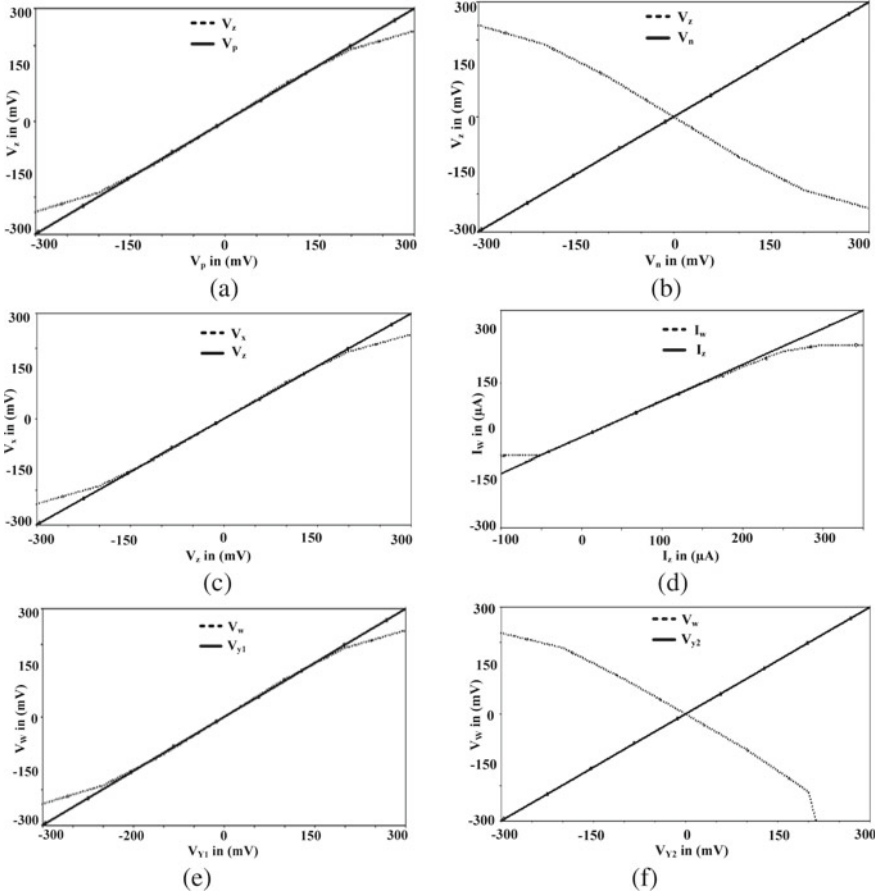
Fig. 3 DC voltage variation behaviour of DD-VDTA at different port **a** voltage z to p **b** voltage z to n **c** voltage x to z **d** current w to z **e** voltage w to y_1 **f** voltage w to y_2
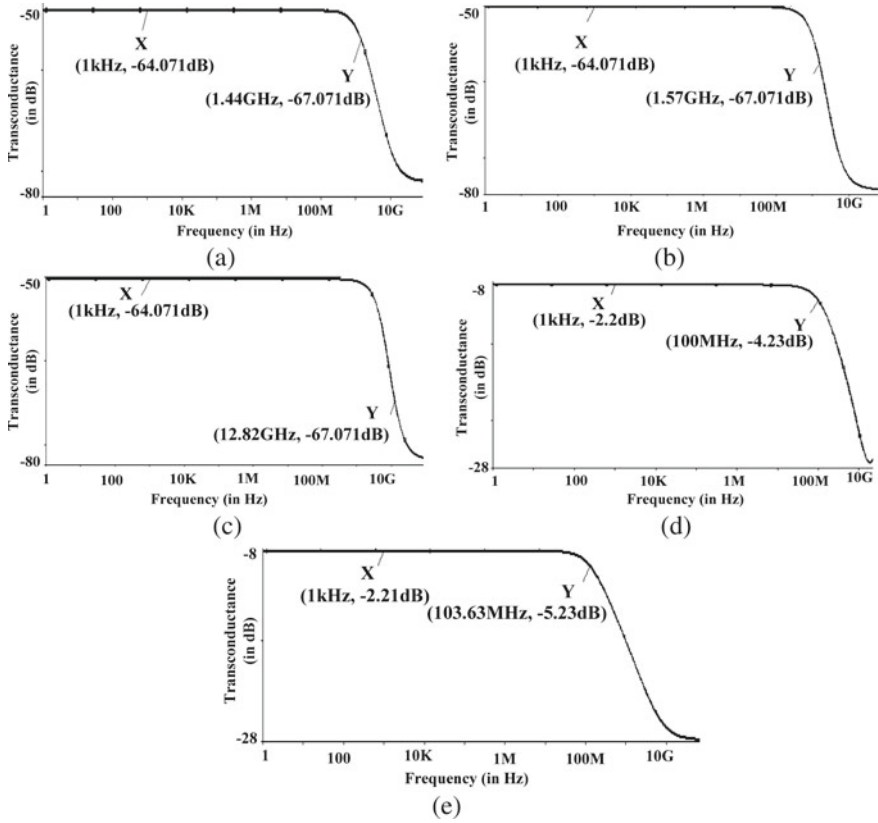
Fig. 4 AC characteristics of DD-VDTA **a** frequency response of g_m from port z to v_p **b** frequency response of g_m from port z to v_n **c** frequency response of g_m from z to v_x **d** frequency response of V_{y1}/V_w **e** frequency response of V_{y2}/V_w

Table 2 Performance parameters of DD-VDTA

S. no.	Parameter	Value
1	Dynamic range of input currents (I_z)	$-80\ \mu$ to $250\ \mu A$
2	Dynamic range of input voltage (V_{y1}, V_{y2})	-250 m to 250 m V
3	Parasitic gain (V_{y1}/V_w, V_{y2}/V_w)	0.77, 0.54
4	Voltage tracking error	0.23, 0.46
5	Transconductance gain (g_m)	627 $\mu A/V$
6	Bandwidth (V_{y1}/V_w), (V_{y2}/V_w)	100 and 103.63 MHz
7	Bandwidth (I_z/V_p), (I_z/V_n), (I_x/V_z)	1.44, 1.57, and 12.82 GHz
8	Power dissipation	2.81 mW
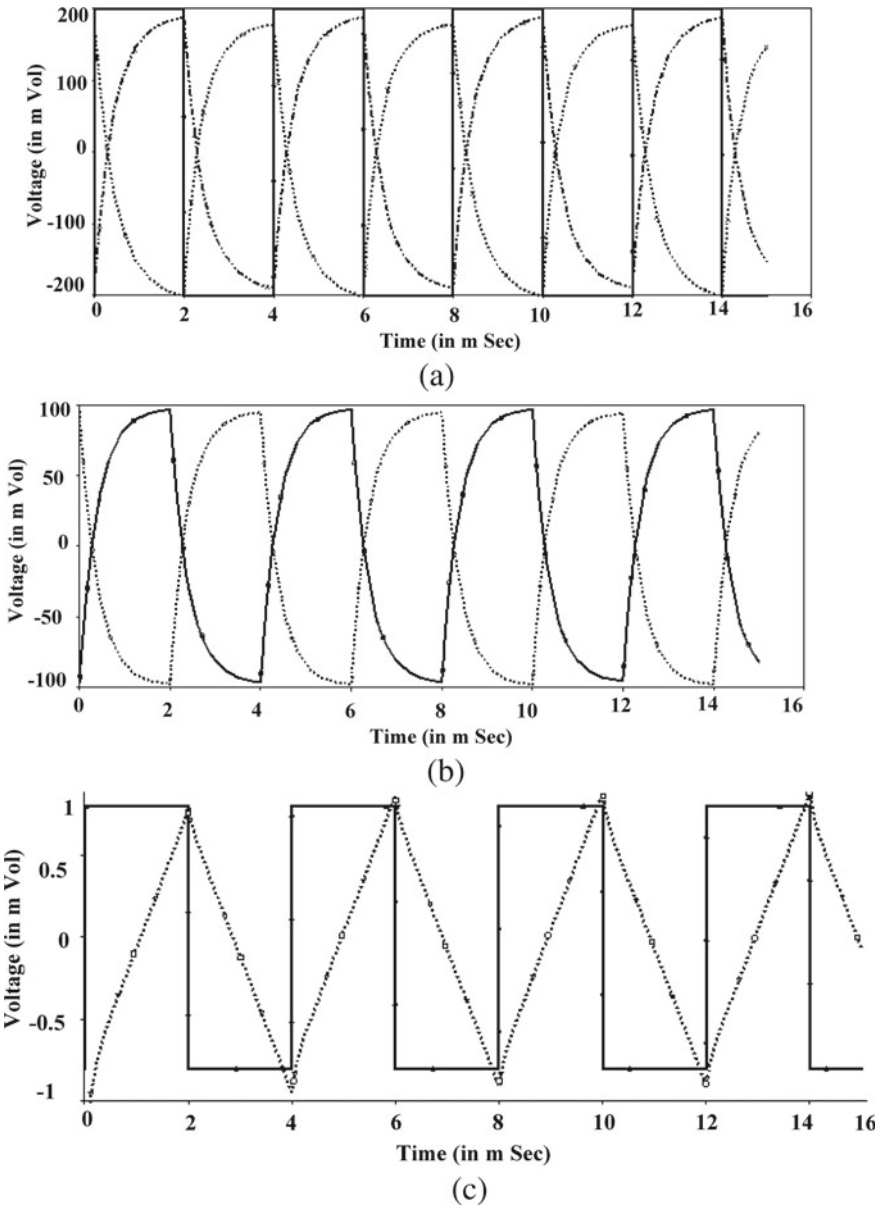
Fig. 5 Output response of integrator circuits **a** lossy integrator **b** lossy subtractor integrator **c** lossless integrator

5 Conclusion

A new active block named Differential Difference VDTA (DD-VDTA, an extension VDTA is proposed. The DC and AC characteristics of this block is done and found that this block can operate for higher frequency range applications. The applications of this block is shown for the lossy integrator, lossy subtractor integrator, and lossless integrator. It is designed using only one DD-VDTA block and one capacitor, having two outputs, viz., inverted and non-inverted at a time. The proposed configurations enjoys advantageous feature like less components are used, dual output at a time, low power dissipation and can be electronically tuned by changing its bias currents and easy design. These applications of proposed block has been proposed and verified. All the mathematical results have been verified by SPICE simulations with TSMC 0.18 μm CMOS process parameters.

References

1. Biolek D, Senani R, Biolkova V, Kolka Z (2008) Active elements for analog signal processing; classification, review and new proposals. Radioengineering 17(4):15–32
2. Moraveji F (1996) A tiny, high-speed, wide-band, voltage-feedback amplifier stable with all capacitive load. IEEE J Solid-State Circuits 31(10):1511–1516
3. Hou CL, Huang CC, Lan YS, Shaw JJ, Chang CM (1999) Current-mode and voltage-mode universal biquads using a single current-feedback amplifier. Int J Electron 86(8):929–932
4. Geiger RL, Sanchez-Sinencio E (1985) Active filter design using operational transconductance amplifiers: a tutorial. IEEE Circ Dev Magazine 1(2):20–32
5. Smith KC, Sedra A (1968) The current conveyor—a new circuit building block. Proc IEEE 56(8):1368–1369
6. Toumazou C, Lidgey FJ, Haigh D (eds) (1990) Analogue IC design: the current-mode approach, vol 2. Presbyterian Publishing Corp
7. Ferri G, Guerrini NC (2003) Low-voltage low-power CMOS current conveyors. Springer Science and Business Media
8. Tangsrirat W, Unhavanich S (2014) Voltage differencing transconductance amplifier-based floating simulators with a single grounded capacitor. IJPAP 52(6):423–428
9. Sedra A, Smith K (1970) A second-generation current conveyor and its applications. IEEE Trans Circ Theory 17(1):132–134
10. Fabre A (1995) Third-generation current conveyor: a new helpful active element. Electron Lett 31(5):338–339
11. Ibrahim MA, Minaei S, Yuce E, Herencsar N, Koton J (2012) Lossy/lossless floating/grounded inductance simulation using one DDCC. Radioengineering 21(1)
12. Elwan HO, Soliman AM (1997) Novel CMOS differential voltage current conveyor and its applications. IEE Proc Circ Dev Syst 144(3):195–200
13. Pathak JK, Singh AK, Senani R (2016) New canonic lossy inductor using a single CDBA and its application. Int J Electron 103(1):1–13
14. Zeki A, Toker A (2003) The dual-X current conveyor (DXCCII): a new active device for tunable continuous-time filters. Int J Electron 89(12):913–923
15. Srivastava M, Chitranshi G, Sengar P, Prasad D (2015) Novel current mode electronically tunable lossless integrator employing VDTA. In: 2015 Annual IEEE India conference (INDICON). IEEE, pp 1–4

16. Santhoshini KM, Musala S, Avireni S (2019) An integrator circuit using voltage difference transconductance amplifier. Solid State Electron Lett 1(1):10–14
17. Kumawat M, Mathur A, Sharma J, Srinivasulu A (2021) An Integrator using voltage differencing transconductance amplifier. In: Proceedings of 6th international conference on recent trends in computing. Springer, Singapore, pp 1–8
18. Kumar P, Pandey N, Paul SK (2017) Operational simulation of LC ladder filter using VDTA. Active Passive Electron Comp
19. Singh H, Arora K, Prasad D (2010) VDTA-based wave active filter. Circ Syst 5:124–131
20. Pandey N, Kumar P, Paul SK (2015) Voltage differencing trans-conductance amplifier based resistorless and electronically tunable wave active filter. Analog Integr Circ Sign Process 84:107–117
21. Singh G (2020) CMOS realization of VDTA based electronically tunable wave active filter with minimum power consumption at low supply voltage ±0.82 V. Circ Syst 11(02):11
22. Rana P, Ranjan A (2021) Odd-and even-order electronically controlled wave active filter employing differential difference trans-conductance amplifier (DDTA). Int J Electron 108(10):1623–1651
23. Rana P, Vista J, Ranjan A (2019) Three-input single-output (TISO) multifunction bi-quad filter implementation using VDTA. In: 2019 3rd international conference on I-SMAC (IoT in social, mobile, analytics and cloud) (I-SMAC). IEEE, pp 431–433

A Compact Dual-Element MIMO Antenna with High Isolation for Wideband Applications

Mahd Azharuddin, Deepak Kumar Barik, Kalyan Mondal, Lakhindar Murmu, and Tapan Mandal

1 Introduction

In the contemporary wireless communications system, there is a need for high channel bandwidth, fast data connections, smart multimedia and improved spectrum performance. These basic needs could well be provided by MIMO antenna systems. In MIMO antenna system, the transmitter and the receiver are equipped with multiple antennas provided there is minimum coupling to achieve among the ports of antenna elements. However, it is a great challenge among the researchers to achieve very low coupling among the ports of antenna elements. Therefore, maintaining low coupling is one of the major parameters in MIMO system [1].

Nowadays, there are many techniques available to mitigate mutual coupling between antenna elements, such as etching slot on the ground structure [2], using metal strip reflector [3], adopting orthogonal mode [4], adding self-curing decoupling with capacitive loads [5] and inserting parasitic stubs between antenna elements [6]. Moreover, isolation can also be increased by incorporating defected ground plane [7], electromagnetic band gap structure [8] and maintaining suitable gaps between the antenna elements.

In this paper, a wideband MIMO antenna consisting of two antenna elements has been proposed and analyzed. The antenna occupies a compact size of 28×32 mm^2 and offers broad frequency bandwidth ranging from 3.15 GHz to 8.5 GHz with good isolation.

M. Azharuddin (✉) · D. K. Barik · K. Mondal
National Institute of Technology Manipur, Imphal, Manipur, India
e-mail: azharuddinmohd.124@gmail.com

L. Murmu
International Institute of Information Technology, Naya Raipur, Chhattisgarh, India

T. Mandal
Government College of Engineering and Textile Technology, Serampore, West Bengal, India

2 Antenna Design

Figure 1 illustrates the configuration of MIMO antenna structure. The MIMO antenna has got FR4 as substrate whose dielectric constraint is $\varepsilon_r = 4.4$. It occupies a compact size of 28×32 mm^2 whose thickness is 1.6 mm. The proposed MIMO antenna contains two similar antenna elements separated by an I-shaped ground stub with partial ground plan. The numerical analysis of the antenna structure was performed by using Ansys electromagnetic simulation software HFSS. The final optimized numerical constraints for the proposed MIMO antenna are recorded as follows (in millimeters): $W = 32$ mm, $L = 28$ mm, $L_g = 10.5$ mm, $w_p = 3$ mm, $w_1 = 14$ mm, $w_2 = 2$ mm, $w_3 = 5$ mm, $w_4 = 1$ mm, $w_5 = 2$ mm, $w_6 = 6.5$ mm, $l_1 = 5$ mm and $l_2 = 14.75$ mm.

To study the effect of I-shaped ground stub on isolation, the current distributions are analyzed at 3.45 GHz by exciting only port 1 as depicted in Fig. 2. Figure 2a demonstrates that a strong surface current is transmitted from port 1 to port 2, due to which there is very low isolation in this case. Further, Fig. 2b shows the surface current coupling from port 1 to port 2 is found to be reduced when the I-shaped ground stub is introduced. The consequence is that high isolation has been provided between the antenna elements. The corresponding S parameters are also studied as shown in Fig. 3 in which the impedance bandwidth is found to be improved as well as the isolation (S_{21}) is enhanced by employing the I-shaped stub.
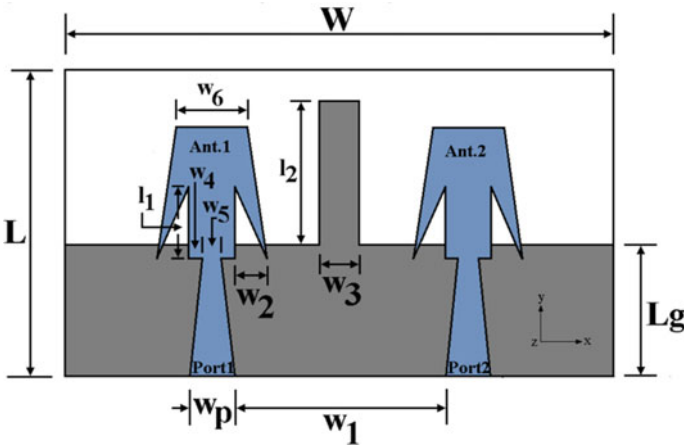
Fig. 1 Geometry of the proposed wideband MIMO antenna
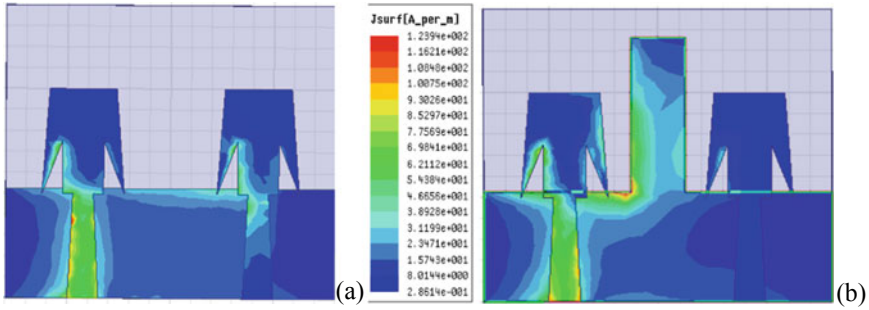
Fig. 2 Current distributions for antennas at 3.45 GHz **a** without I-shaped stub and **b** with I-shaped stub

Fig. 3 S parameters of the proposed antenna with and without I-shaped ground stub
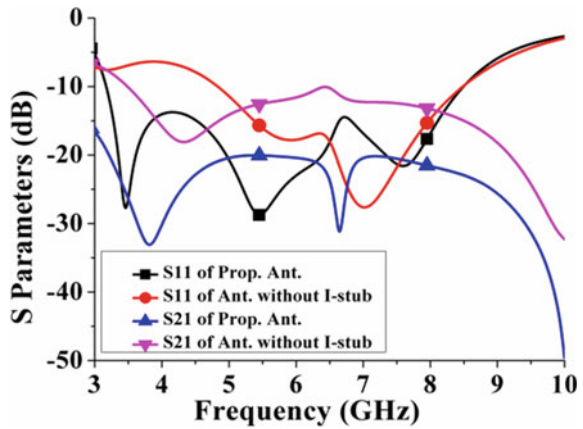
3 Results and Discussion

3.1 S Parameters

Figure 4 describes the simulated S parameters of the proposed MIMO antenna. It is seen that the MIMO antenna offers an operation bandwidth from 3.15 to 8.5 GHz for $S_{11} < -10$ dB. A high isolation, which is more than 20 dB over the entire frequency spectrum, is achieved.

3.2 Diversity Performance

The simulated ECC and DG curves are shown Fig. 5a, b, respectively.

Fig. 4 S parameters of the proposed antenna
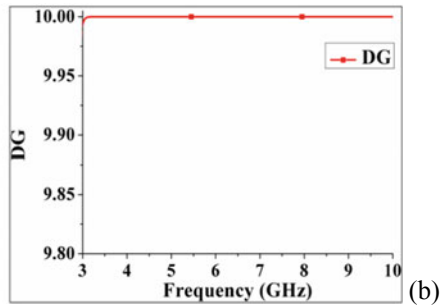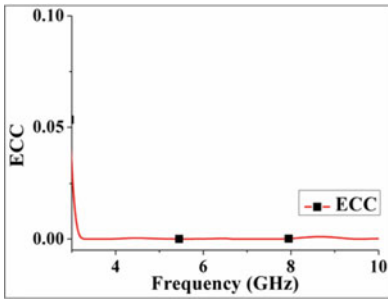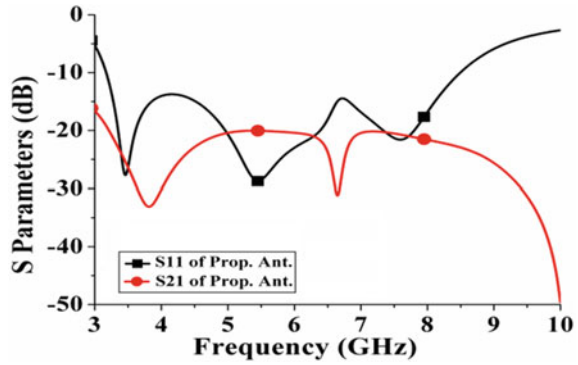
Fig. 5 Simulated **a** envelop correlation coefficient, **b** diversity gain of the proposed antenna

The figures also show that the simulated ECC value is found to be below 0.02 which is less than the maximum acceptable ECC value of less than or equal to 0.5. Moreover, diversity gain is found to be greater than 9.95 dBi throughout the entire frequency spectrum which ranges from 3.15 to 8.5 GHz.

3.3 Radiation Characteristics

The simulated radiation patterns of the proposed MIMO antenna are analyzed by exciting only port 1. Figure 6a, b depict the radiation patterns under *xoz*- and *yoz*-planes at 3.45 GHz and 5.45 GHz, respectively. The far field radiation pattern shows that field distribution is omnidirectional at 3.45 GHz and 5.45 GHz, respectively.
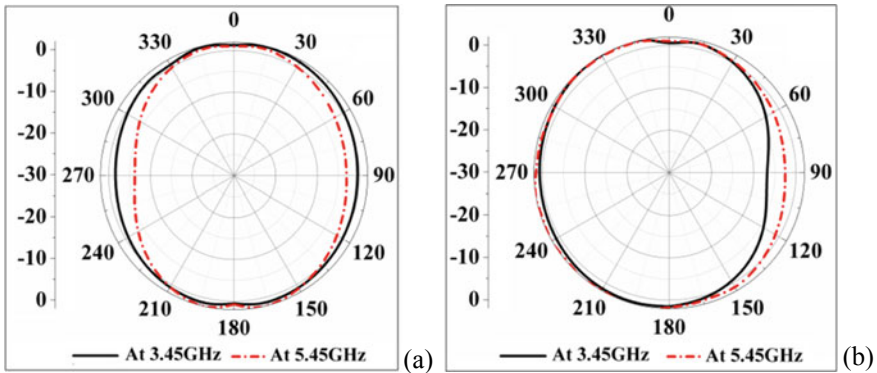
Fig. 6 Simulated radiation patterns of the MIMO antenna on **a** XOZ- and **b** YOZ-planes

4 Conclusion

The proposed work presents a compact MIMO antenna having a size of $0.32 \lambda_0 \times 0.36 \lambda_0$ for application in wideband spectrum. The operation frequency band of the antenna ranges from 3.15 to 8.5 GHz. The isolation is less than -20 dB throughout the operating band which confirms very good isolation in the MIMO system. The ECC value is found to be less than 0.02 and diversity gain is greater than 9.95 dBi throughout the band.

References

1. Sharawi MS (2013) Printed multi-band MIMO antenna systems and their performance metrics [wireless corner]. IEEE Anten Propag Mag 55(5):218–232
2. Ren J, Hu W, Yin Y, Fan R (2014) Compact printed MIMO antenna for UWB applications. IEEE Anten Wirel Propag Lett 13:1517–1520
3. Roshna TK, Deepak U, Sajitha VR, Vasudevan K, Mohanan P (2015) A compact UWB MIMO antenna with reflector to enhance isolation. IEEE Trans Anten Propag 63(4):1873–1877
4. Sun L, Feng H, Li Y, Zhang Z (2018) Compact 5G MIMO mobile phone antennas with tightly arranged orthogonal-mode pairs. IEEE Trans Anten Propag 66(11):6364–6369
5. Sui J, Wu KL (2018) Self-curing decoupling technique for two inverted-F antennas with capacitive loads. IEEE Trans Anten Propag 66(3):1093–1101
6. Liu L, Cheung SW, Yuk TI (2015) Compact MIMO antenna for portable UWB applications with band-notched characteristic. IEEE Trans Anten Propag 63(5):1917–1924
7. Wei K, Li JY, Wang L, Xing Z, Xu R (2016) Mutual coupling reduction of microstrip antenna array by periodic defected ground structures. In: IEEE conference on antennas and propagation (APCAP), Kaohsiung, pp 389–390
8. Wu W, Yuan B, Wu A (2018) A quad-element UWB-MIMO antenna with band-notch and reduced mutual coupling based on EBG structures. Int J Antennas Propag

Design of GWO-MBIMC Controller to Stabilize the Frequency of Microgrid on Real-Time Simulation [OPAL-RT OP4510] Platform

Badal Kumar, Shuma Adhikari, and Nidul Sinha

1 Introduction

Distribution systems can transform existing passive distribution networks into active distribution networks that typically operate in grid connected channel by incorporating power generation on small scale. These small power plants however need to be isolated from the main power grid because vibration, voltage drop, swelling, faults are susceptible to the main grid frequency. Therefore, it is necessary to run them as a separate grid, called MG [1]. The modelling of MG and its frequency stabilization using the Ziegler–Nichols-based controller was presented in [2] and the LDR-based PID controller, in which the Ziegler–Nichols open loop approach is used as the tuning algorithm, was presented in [3]. It should be noted that with time delay, the Ziegler–Nichols closed loop cannot be extended to first-order or second-order systems. The tuning algorithm will automatically be shifted to the Ziegler–Nichols open loop if the Ziegler–Nichols closed loop approach is chosen for these situations. It is also found that under more stressed situation (worst situations), the scheme suffers from high peak variance. In order to achieve good dynamic frequency responses, a secondary controller [4] oriented MB has been documented, but scenario of high stress situation have not been considered. Even if it had been considered, its setting time would

B. Kumar (✉)
Electrical Engineering Department, K. K. Wagh Institute of Engineering Education & Research, Nashik, Maharashtra 422003, India
e-mail: kumarbadal89@gmail.com

S. Adhikari
Electrical Engineering Department, National Institute of Technology Manipur, Langol, Imphal 795004, India

N. Sinha
Electrical Engineering Department, National Institute of Technology Silchar, Assam 788010, India

have been much higher. A modified bias type LDR (MB-LDR) controller has been suggested in [5] to obtain optimum PID controllers and frequency bias parameters for controlling the plant frequency under various operating conditions, including worst-case scenario. In reality, however, the LDR only considers open loop strategies because it has more settling time, peak overshoot, and more oscillations. The reliability of the open loop plant is, however, weaker than that of the closed loop plant. Thus, it is preferable to use a closed loop type approach for tuning PID gain values, such as the IMC technique [6, 7]. IMC approach is based on pole-zero cancellation. A low pass filter choice is very necessary in the IMC controller system design because it affects the device output. Low pass filter choices are suggested in [6] for several common plant models. In different fields of sciences and technology [7–9], this approach is often used. After selection of low pass filter, selection of closed loop time constant or filter coefficient (λ) is important for each process model to optimize PID performance. So, by using the trial and error approach to achieve optimum performance, it involves tuning by the designer, making it a time-consuming operation, and it is not guaranteed that it can have the universal solution. Meta-heuristic implementations, however, have proven to tackle these kinds of optimization concerns very effectively. In addition, meta-heuristic optimization efficiency relies greatly on how fertile the fields they are for searching. An Arbitrary collection of control variables ranges can result in inefficient search, and a variety of insecure (infeasible) search points may also be available. However, if MB & IMC related gain values are taken for choice of suitable ranges for the meta-heuristic search algorithm control variable, this would ensure that the search algorithm produces secure (feasible) search points and as a result, the algorithm performs better. MB & IMC focused gain values are the guiding parameter for selecting the limit for efficient search of the control variables in this context alone.

A contemporary meta-heuristic method, known as GWO, is proposed in [10] and Particle swarm Optimization (PSO). PSO method is old but very well-known algorithms so, authors have not given their detailed descriptions. It has been used for the purposes of comparison with GWO only. GWO is based on the social cooperation and hunting behaviours of grey wolves to locate and pursue a prey (solution) [11, 12]. GWO implementation is quite effective compared to other optimization techniques since particular input parameters are not needed. In GWO, regardless of their capabilities, three better wolf categories have been given the same weight in order to estimate the next position of the wolf in the iterative process.

In this paper, a recent meta-heuristic technique known as GWO is taken into consideration to calculate the best gain values of the proposed controller. The suggested controller is an MBIMC-based controller to control the frequency of a self-reliant MG system, as shown in Fig. 1. Additionally, OPAL-RT LAB (OP4510), a digital version of OPAL-RT, was employed to validate the MG system on a real-time platform. The steps for the OPAL-RT lab-based real-time Simulink are described in [13].

The following are the main contributions of this work:

Fig. 1 Basic MG islanded diagram
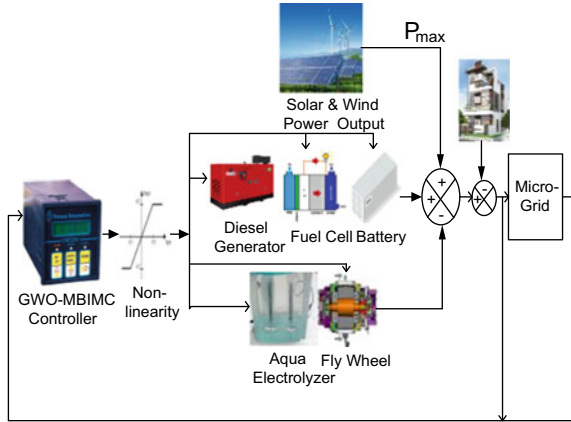
i. Design the proposed controller (i.e. GWO-MBIMC Controller), with use of GWO algorithms to optimize its gain and frequency bias parameters.

ii. On the real-time simulator OPAL-RT LAB (OP4510) platform, the suggested controller's realistic performance is verified.

The structure of this article is as follows: Sect. 2 provides the theoretical foundation for MG modelling, variation in power and frequency, and modified bias (MB) type controller design. The design of an IMC-based gain is described in Sect. 3. Section 4 presents the use of the GWO type MBIMC controller and its functionality in MG. A comparison of the tuned controller parameter using conventional and evolutionary methods (GWO and PSO) is shown in Sect. 5. Section 6 presents the findings of the real-time simulation. In Sect. 7, the results are summarized.

2   Modelling and P-F Droop Design for MG Plant

The self-governing power source's average rating is the same as the load rating, considering a 100% independent isolated microgrid. The Detailed Description about Modelling of Uncontrollable source, Controllable source, and Designing of p-f droop for all the controllable sources using bode plots, relation between power and frequency deviation, etc. are depicted in [3, 5].

3   Design of IMC Controller Parameters

In this article, a third-order for the DEG and second order for remaining secondary controllable sources with plant are shown in Eq. (1). The PID controller design procedures using IMC tested by the real-time simulator using MATLAB. The Detailed

Description about IMC parameters are depicted in [6–9, 14].

$$
\left.
\begin{aligned}
G_{\text{deg}} &= \frac{K_{\text{hps}}}{(1 + sT_{\text{deg}\,t})(1 + sT_{\text{deg}\,g})(1 + sT_{\text{hps}})}; \ G_{\text{ae}} = \frac{K_{\text{hps}}}{(1 + sT_E)(1 + sT_{\text{hps}})} \\
G_{\text{bess}} &= \frac{K_{\text{hps}}}{(1 + sT_{\text{bess}})(1 + sT_{\text{hps}})}; \ G_{\text{fc}} = \frac{K_{\text{hps}}}{(1 + sT_{\text{fc}})(1 + sT_{\text{hps}})}; \\
G_{\text{fess}} &= \frac{K_{\text{hps}}}{(1 + sf_{\text{bess}})(1 + sT_{\text{hps}})}
\end{aligned}
\right\}
\tag{1}
$$

where $T_{\text{deg}\,t}$, $T_{\text{deg}\,g}$, T_{ae}, T_{fw}, T_{fess}, and T_{bess} are time constants of turbine and governor of DEG, AE, FC, FESS, and BESS, respectively, whereas K_{hps} and T_{hps} are the gain and time constant of Hybrid Power System. The detailed description of all system parameters is given in appendix.

4 Application of GWO on MBIMC-Based Controllers

If the controller is required in an electrical network, all conditions, including plant robustness, dynamics and so on, should be optimum. The output of the controller is thus verified using different objective functions. The objective function allows producing a transient response with minimum numbers of oscillations, overshooting, and settling time from an optimization perspective. Keeping this in mind, the ISE is considered as an objective function in Eq. (2) for determining transient responses based on minimum time for settling. Number of oscillations and peak overshoot in this research.

$$
\text{ISE} = \int_0^T \Delta f^2 \mathrm{d}t
\tag{2}
$$

Here, Δf and T are the microgrid frequency deviation and simulation time, respectively.

Implementation of GWO-MBIMC technique, the GWO Method & Regular flow chart in Figs. 3 and 4 depicts the stages involved in developing the proposed controller-based GWO technique and the reader might refer to [10–12] for a more in-depth discussion of the GWO approach. The GWO-based MB & IMC technique has used to conduct this research. Figure 3b, also demonstrates how the parameters of the proposed controller-based ISE objective function have been chosen. Table 1 also shows the maximum and minimum ranges of the controller settings.

Table 1 Parameter limit for GWO-based MBIMC technique

GWO based MBIMC parameters	
Number of Iteration	50
Number of search agents	30
Range of P	[0.0001 10]
Range of I	[0.0001 1]
Range of D	[0.0001 15]
Range of R	[0.0001 410]
No. of variables/dimensions	20

5   Tuned Controller Parameters for Comparative Analysis

The key aim of the controller is to mitigate the frequency deviation by adjusting the secondary sources supply's power output and also producing an acceptable control signal to boost the efficiency of the system. There is a possibility of adverse contact between these regulators, in the presence of several secondary sources, contributing to the disruption of MG frequency. There is not a single description so far for applying optimum tuning techniques to all loops. Consequently, to prevent adverse interactions [4], the individual PID controller must be appropriately tuned.

In PSO-MBIMC and GWO-MBIMC techniques we have calculated all values of controller parameter from the MATLAB programming code. Table 2 provide the frequency regulation parameters, i.e. K_f and gain values for PID controllers, i.e. K_p, K_i, and K_d obtained by the MB-LDR [13], MBIMC [14], PSO-MBIMC [14], and Proposed GWO-MBIMC methods.

The value obtained from the above table has been used in the Fig. 5 of [12], and the results were compared. The GWO-MBIMC and PSO-MBIMC convergence characteristics are seen in Fig. 2. This graph is showing that the GWO algorithm is more superior to PSO. MB & IMC focused gain values are the guiding parameter for selecting the limit for efficient search of the control variables for GWO & PSO.

6   Real-Time Simulation Results and Discussion

In the MATLAB (SIMULINK) (R2018b) system, Fig. 5 of [13] is designed and applied to real-time simulation in the OPAL-RT environment [13]. In order to research the output of the microgrid plant and evaluate the effects in depth by subjecting the plant to power fluctuations at both source and load sides, the microgrid system is run for 70 s. It is considered that the plant has wind power, solar power, and load at the beginning.

Solar and wind systems provide constant powers of 0.4 pu and 0.3 pu, respectively, whereas load demand is estimated to be 0.7 pu prior to any type of disturbance. On Intel, 4 GB RAM and 3.4 GHz core i-7 C.P.U computer, the simulation of the research

Table 2 Adjusting controller parameters through GWO-based MBIMC technique with comparisons

Microgrid components	MB-LDR method [13]				MBIMC method [14]				GWO-MBIMC method			
	MB K_f	LDR K_p	LDR K_i	LDR K_d	MB K_f	IMC K_p	IMC K_i	IMC K_d	GWO-MB K_f	GWO-IMC K_p	GWO-IMC K_i	GWO-IMC K_d
DEG	20.4918	0.1025	0.0073	0.3013	20.4918	0.058	0.0015	0.61	250	0.495	0.0092	9.725
AE	0.002	1.498	4.54	0.1044	0.002	0.0253	0.0015	0.005	0.0037	0.1322	0.0003	0.003
FC	40.9836	0.1858	0.0464	0.156	40.9836	0.031	0.0015	0.10	151	0.496	0.0055	0.069
BESS	0.001	3.38	0.1251	0.098	0.001	0.0252	0.0015	0.0025	8	0.5	0.01	0.0372
FWSS	0.001	3.38	0.1251	0.098	0.001	0.0252	0.0015	0.0025	0.0011	2.49	1	0.0006

Fig. 2 Convergence characteristics
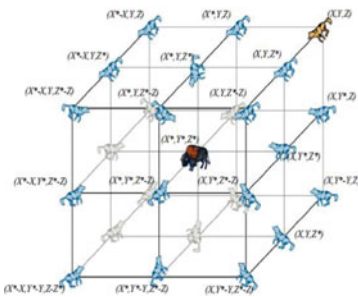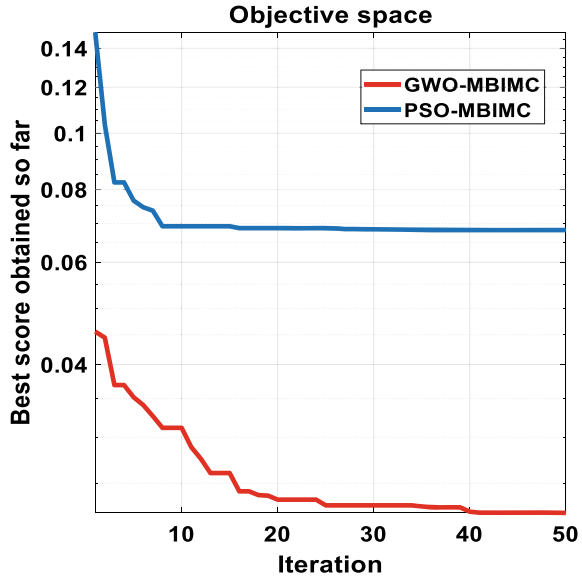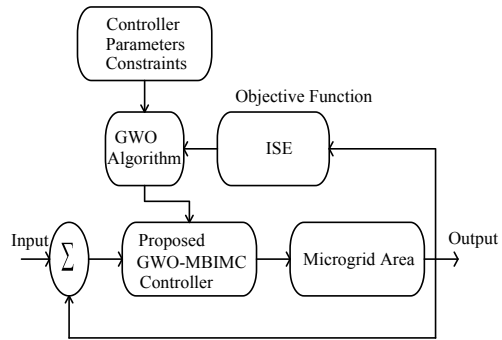
Fig. 3 **a** 3D position of GWO scheme, **b** Process of determining optimal constraints of the proposed GWO-MBIMC controller

model is carried out through optimization programmes. For the GWO process, 30 population sizes and 50 iteration numbers are considered in this article. The optimization technique runs individually 3 times, and as controller parameters, the best optimal results are used. The GWO-MBIMC method obtained results are compared to those obtained from the PSO-MBIMC, MBIMC, and MB-LDR controllers, and it is observed that out of all four methods, the proposed method produces the best responses. System efficiency is considered for 4 separate scenarios, which are defined in Figs. 5, 6, 7 and 8, respectively.
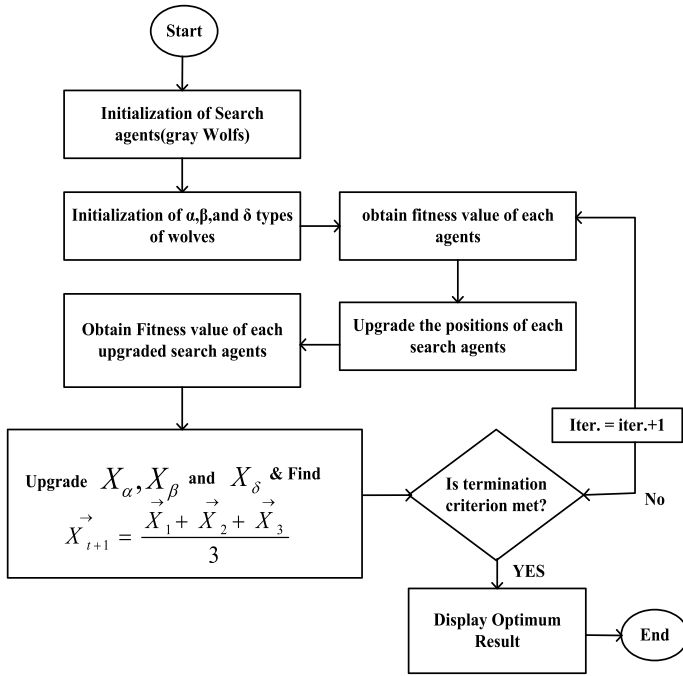
Fig. 4 GWO algorithm flowchart [10, 15].
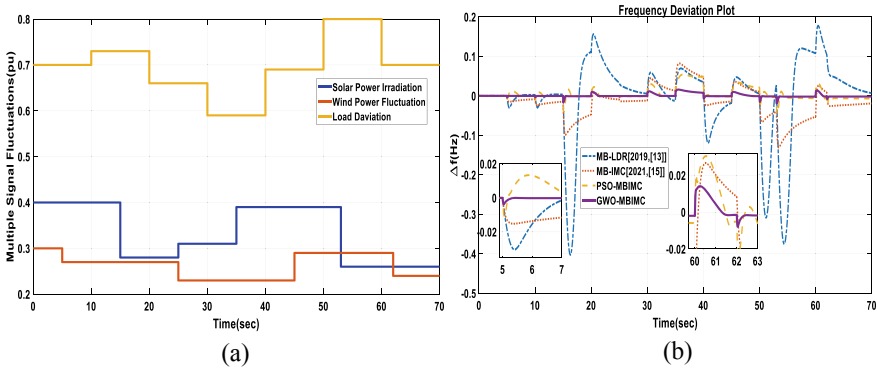
Fig. 5 Scenario-1 result, **a** Multiple signal fluctuation versus **b** frequency deviation

Scenario-1: Realistic Situation (Real-Time Fluctuation in Wind Energy, Solar Energy, and Load):

Here, simultaneous disruptions are considered in this case in terms of wind capacity, sun irradiation strength, and load. The combined variations in wind energy,
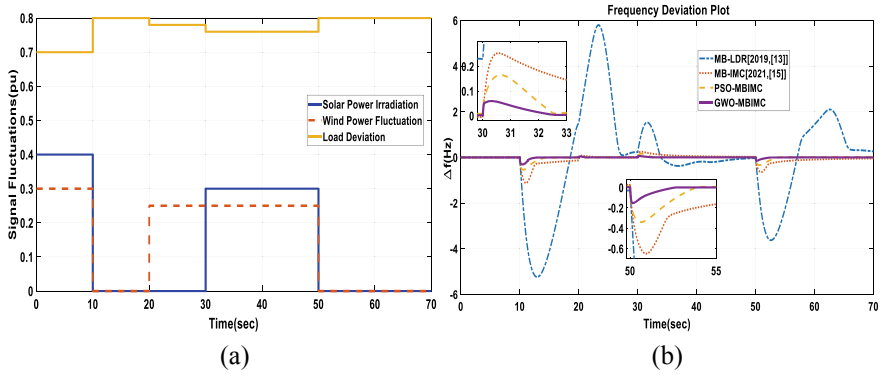
Fig. 6 Scenario-2 result, **a** signals fluctuation versus **b** frequency deviation
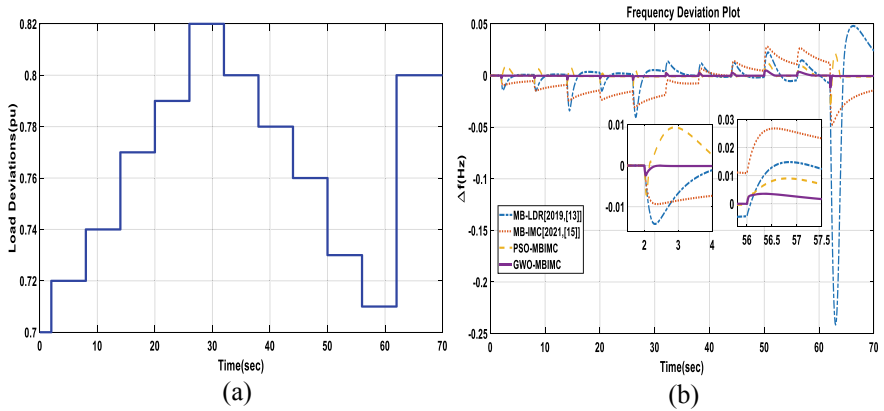
Fig. 7 Scenario-3 result, **a** load signal deviations versus **b** frequency deviation
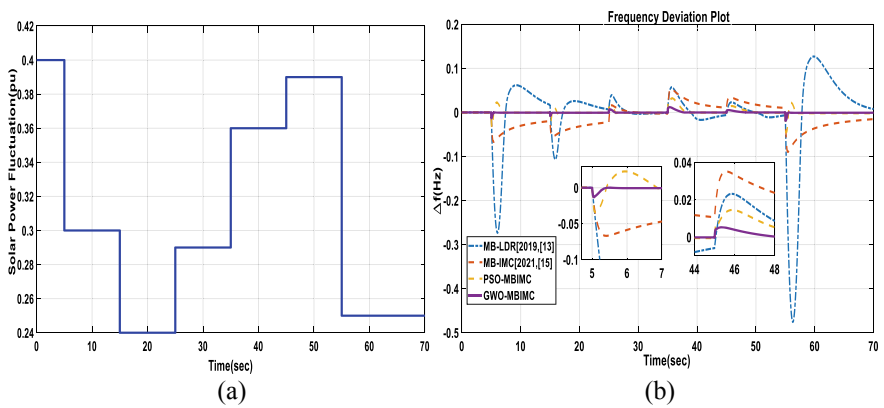
Fig. 8 Scenario-4 result, **a** Solar signals fluctuation versus **b** frequency deviation

solar energy, and load are seen in Fig. 5a. Figure 5b demonstrate the action of Frequency Response.

Scenario-2: More Stressed Situation (Parallel Fluctuation in Wind Energy and Solar Energy and Load):

In this case, the load demand is increased from 0.7 pu to 0.8 pu while the wind and solar power are varied from 0.3 pu & 0.4 pu respectively to 0 pu (worst/stressed situation) for a duration of 70 s as seen in Fig. 6a. The results of the frequency response are displayed in Fig. 6b.

The microgrid faces rapid oscillations due to immediate variation in the load. This large fluctuation in the microgrid frequency is due to a quick major change in both generation and load. All of the secondary sources in the microgrid have their output responses tuned in order to reduce the imbalance between power demand and frequency deviation.

Scenario-3: Wind and solar power sources are held constant while the load is increased from 0.7 pu to 0.82 pu for 70 s. (time domain analysis)

As illustrated in Fig. 7a, the only load on the MG system is changed in this part in a step pattern across a time range of 2–62 s. Using the MB-LDR, MBIMC, PSO-MBIMC, and suggested GWO-MBIMC-based controllers, the frequency responses of the MG system under the modification of step pattern-based load are obtained as shown in Fig. 7b.

Scenario-4: For a period of 70 s, the solar is varied from 0.4 pu (highest range) to 0.24 pu (lowest range), while the load (0.7 pu) is maintained constant. (Time Domain Analysis)

As shown in Fig. 8a, the amount of sun irradiation is gradually changed from time 5 to 55 s. The frequency responses of the system's controllers to variations in solar sources are shown in Fig. 8b.

For this all 4-test scenario, both MB-LDR and MBIMC controllers failed and could not stabilize the microgrid frequency although the suggested PSO-MBIMC methods produce better responses in all situations. The results of the simulation indicate that the dynamic output response of the proposed GWO-MBIMC controller produces best results in terms of settling time, peak transient variance, and number of oscillations in all of the above-mentioned scenarios.

7   Conclusion

The regulation of the load frequency of microgrid associated with non-conventional sources is being investigated. An approach is made to improve the performance of microgrid system where a MBIMC-based controller is designed and its values are optimized using GWO-based modern meta-heuristic technique. In an OPAL-RT digital simulator-based environment for real-time simulation, the new proposed

GWO-MBIMC controller is also applied to verify the output of microgrid in a number of scenarios.

The simulation results on real-time platform obtained using proposed GWO-MBIMC techniques are compared with that of other methods like PSO-MBIMC, MBIMC, and MB-LDR. It is observed, from simulation results, that GWO-MBIMC controller gives better results over other methods in terms of smaller number of oscillations, settling time and peak deviation and stability of microgrid during transients, thus shows that the proposed controller provides optimal stable operation.

Appendix

A. **Nominal parameters of the MG**.

f_{sys}= 50 Hz, P_{base}= 1MVA, $D = 0.012$ MW/Hz, $M = 0.2$ s, $T_{\deg g} = 2$ s, $T_{\deg t} = 20$ s, T_{fc}= 4 s, T_{ae}= 0.2 s, T_{bess}= 0.1 s, T_{fess}= 0.1, $K_{\deg g} = K_{\deg t} = K_{ae} = K_{fc} = K_{bess} = K_{fess} = 1$, $K_{hps} = 83.33$, T_{hps}=16.67.

B. **Nominal MG parameters**.

Wind Power Source (300 kW), Solar Power Source (400 kW), Load (700 kW), Diesel Generator (500 kW), Fuel Cell (200 kW), Aqua Electrolyzer (100 kW), Battery (30 kWh) and Fly Wheel (30 kWh).

References

1. Bevrani H (2009) Robust power system frequency control, 2nd edn. Springer, Berlin
2. Mallesham G, Mishra S, Jha AN (2011) Ziegler-Nichols based controller parameters tuning for load frequency control in a microgrid. In: 2011 International conference on energy, automation and signal, Bhubaneswar, Odisha, pp 1–8. https://doi.org/10.1109/ICEAS.2011.6147128
3. Kumar B, Bhongade S (2016) Load disturbance rejection based PID controller for frequency regulation of a microgrid. In: 2016 Biennial international conference on power and energy systems: towards sustainable energy (PESTSE), Bengaluru, India, pp 1–6. https://doi.org/10.1109/PESTSE.2016.7516459
4. Mishra S, Mallesham G, Sekhar PC (2013) Biogeography based optimal state feedback controller for frequency regulation of a smart microgrid. IEEE Trans Smart Grid 4(1):628–637. https://doi.org/10.1109/TSG.2012.2236894
5. Kumar B, Bhongade S (2017) Load disturbance rejection based PID controller for frequency regulation of a microgrid. Ind J Electr Eng Comput Sci 7(3):625–642. https://doi.org/10.11591/ijeecs.v7.i3.pp625-642
6. Rivera DE, Morari M, Skogestad S (1986) Internal model control: PID controller design. Ind Eng Chem Process Des Dev 25(1):252–265

7. Suvilath S, Khongsomboun K, Benjanarasuth T, Komine N (2011) IMC-based PID controllers design for a two-links SCARA robot. In: TENCON 2011—2011 IEEE region 10 conference, Bali, pp 1030–1034. https://doi.org/10.1109/TENCON.2011.6129267

8. Yadav AK, Gaur P (2015) Intelligent modified internal model control for speed control of nonlinear uncertain heavy duty vehicles. ISA Trans 56:288–298

9. Touati N, Saidi I, Dhahri A et al (2019) Internal Multimodel control for nonlinear overactuated systems. Arab J Sci Eng 44:2369–2377. https://doi.org/10.1007/s13369-018-3515-5

10. Mirjalili S, Mirjalili SM, Lewis A (2014) Grey wolf optimizer. Advances in Engg. Software 69:46–61

11. Padhy S, Panda S, Mahapatra S (2017) A modified GWO technique based cascade PI-PD controller for AGC of power systems in presence of plug in electric vehicles. Eng Sci Technol Int J 20(2):427–442

12. Guha D, Roy PK, Banerjee S (2016) Load frequency control of interconnected power system using grey wolf optimization. Swarm Evol Comput 27:97–115

13. Kumar B, Adhikari S, Datta S, Sinha N (2019) Real Time simulation of modified bias based load disturbance rejection controller for frequency regulation of islanded micro-grid. Int J Emerg Electr Pow Syst 20(5):1–13

14. Kumar B, Adhikari S, Sinha N (2021) Real time simulation of an intelligently optimized controller to control the frequency of an islanded AC microgrid: MBA-MBIMC tuning approach. In: 2020 3rd International conference on energy, power and environment: towards clean energy technologies, pp 1–6. https://doi.org/10.1109/ICEPE50861.2021.9404465

15. Kumar B, Adhikari S, Datta S, Sinha N (2021) Real time simulation for load frequency control of multisource microgrid system using grey wolf optimization based modified bias coefficient diagram method (GWO-MBCDM) controller. J Electr Eng Technol 16:205–221. https://doi.org/10.1007/s42835-020-00596-2

Improved Polar Extensions of an Inequality for a Complex Polynomial with All Zeros on a Circle

Kshetrimayum Krishnadas and Barchand Chanam

1 Introduction

Let $f(w) = \sum_{j=0}^{m} c_j w^j$ be a polynomial of degree m and let

$$\text{Max}(f, r) = \max_{|w|=r} |f(w)|.$$

If $f(w)$ has no zero in $|w| < \rho$, $\rho \geq 1$, Malik [8] proved that

$$\text{Max}(f', 1) \leq \frac{m}{1 + \rho} \text{Max}(f, 1), \tag{1}$$

for which the equality holds for the polynomial $(w + \rho)^m$.

A natural question that arise is whether there exists an analogous inequality of (1) for $f(w)$ having no zero in $|w| < \rho$, $\rho \leq 1$. In this regard, Govil [5, 6] proved the following two results.

Theorem 1 [5] *If $f(w)$ is a polynomial of degree m having no zero in $|w| < \rho$, $\rho \leq 1$, then*

$$\text{Max}(f', 1) \leq \frac{m}{1 + \rho^m} \text{Max}(f, 1), \tag{2}$$

K. Krishnadas (✉)
Department of Mathematics, Shaheed Bhagat Singh College, University of Delhi, New Delhi 110017, India
e-mail: kshetrimayum.krishnadas@sbs.du.ac.in

B. Chanam
Department of Mathematics, National Institute of Technology Manipur, Imphal, Manipur 795004, India

provided that $|f'(w)|$ *and* $|F'(w)|$ *attain their maxima at the same point on the circle* $|w| = 1$, *where*

$$F(w) = w^m \overline{f\left(\frac{1}{\overline{w}}\right)}.$$

Theorem 2 [6] *If* $f(w)$ *is a polynomial of degree m having all its zeros on* $|w| = \rho, \rho \leq 1$,

$$\text{Max}(f', 1) \leq \frac{m}{\rho^m + \rho^{m-1}} \text{Max}(f, 1). \tag{3}$$

In literature we find refinements, generalizations and extensions of Theorems 1 and 2 by involving some coefficients of the polynomial $f(w)$ (see [2–4]).

Definition 1 *Let* $f(w)$ *be a polynomial of degree m and let* β *be any complex number. The polar derivative of* $f(w)$ *with respect to the point* β, *denoted by* $D_\beta f(w)$, *is defined as*

$$D_\beta f(w) = mf(w) + (\beta - w)f'(w).$$

$D_\beta f(w)$ is a polynomial of degree at most $m - 1$. It can be considered as a generalized form of the ordinary derivative of $f(w)$ with respect to w due to the fact that

$$\lim_{\beta \to \infty} \frac{D_\beta f(w)}{\beta} = f'(w).$$

Polar derivative extension of (1) was proved by Aziz [1], who under the same hypothesis on $f(w)$ proved that

$$\text{Max}(D_\beta f(w), 1) \leq m\left(\frac{\rho + |\beta|}{1 + \rho}\right) \text{Max}(f, 1), \quad \text{where } |\beta| \geq 1. \tag{4}$$

2 Lemmas

We need the following results to prove our results.
    The following lemma is due to Govil and Rahman [7].

Lemma 1 [7] *If* $f(w)$ *is a polynomial of degree m having all its zeros on* $|w| = \rho$, $\rho \leq 1$, *then on* $|w| = 1$.

$$|f'(w)| + |F'(w)| \leq m \, \text{Max}(f, 1). \tag{5}$$

The next two lemmas are due to Barchand et al. [3].

Lemma 2 [3] *If* $f(w) = \sum_{j=0}^{n} c_j w^j$ *is a polynomial of degree m having all its zeros in $|w| \leq \rho$, $\rho \leq 1$, then*

$$\frac{1}{\rho m}\left(\frac{c_{m-j}}{c_m}\right) \leq 1. \tag{6}$$

Lemma 3 [3] *If* $f(w) = \sum_{j=0}^{n} c_j w^j$ *is a polynomial of degree m having all its zeros on $|w| = \rho$, $\rho \leq 1$, then on $|w| = 1$.*

$$\text{Max}(f', 1) \leq \frac{m}{\rho^m + \rho^{m-1}} E_\rho \text{ Max}(f, 1), \tag{7}$$

where

$$E_\rho = \frac{(1 + |t|)(\rho^2 + |t|) + \rho(m-1)|s - t^2|}{(1 - |t|)(1 - \rho + \rho^2 + \rho|t|) + \rho(m-1)|s - t^2|},$$

$$t = \frac{1}{\rho m}\left(\frac{\overline{c}_{m-1}}{\overline{c}_m}\right)$$

$$s = \frac{2}{\rho^2 m(m-1)}\left(\frac{\overline{c}_{m-2}}{\overline{c}_m}\right).$$

Lemma 3 is, in fact, a refinement of Theorem 2 due to Govil [6] (see Remark 3).

Remark 1

Under the hypothesis of Lemma 3, we have (see [3, Lemma 2.6]).

$$|t| = \frac{1}{\rho m}\left|\frac{\overline{c}_{m-1}}{\overline{c}_m}\right| \leq 1.$$

The following lemma was proved by Malik [8].

Lemma 4 [8] *If $f(w)$ is a polynomial of degree m having no zero in $|w| < \rho$, $\rho \geq 1$, then*

$$\rho|f'(w)| \leq |F'(w)|, \tag{8}$$

where $F(w) = w^m \overline{f\left(\frac{1}{\overline{w}}\right)}$.

Lemma 5 *If $f(w)$ is a polynomial of degree m having all its zeros on $|w| = \rho$, $\rho \leq 1$, then on $|w| = 1$.*

$$|F'(w)| \leq \rho|f'(w)|. \tag{9}$$

Proof

Since $f(w)$ has all its zeros on $|w| = \rho$, $\rho \leq 1$, then $F(w)$ has all its zeros on $|w| = \frac{1}{\rho}$, $\frac{1}{\rho} \geq 1$. This implies that $F(w)$ has no zeros in $|w| < \frac{1}{\rho}$, $\frac{1}{\rho} \geq 1$. Thus, applying Lemma 1 to the polynomial $F(w)$ we get, on $|w| = 1$.

$$\frac{1}{\rho}|F'(w)| \leq |f'(w)|.$$

$$\therefore |F'(w)| \leq \rho|f'(w)|.$$

3 Main Results

In this paper, we prove polar extensions of Theorem 2. Precisely, we prove the following result.

Theorem 3 *If $f(w) = \sum_{j=0}^{n} c_j w^j$ is a polynomial of degree m having all its zeros on $|w| = \rho$, $\rho \leq 1$, and $\beta \in \mathbb{C}$ with $|\beta| \geq 1$, then on $|w| = 1$.*

(a)

$$|D_\beta f(w)| \leq m\left(1 + \frac{|\beta| - 1}{\rho^m + \rho^{m-1}} E_\rho\right) \text{Max}(f, 1). \tag{10}$$

(b)

$$|D_\beta f(w)| \leq m\left(\frac{|\beta| + \rho}{\rho^m + \rho^{m-1}} E_\rho\right) \text{Max}(f, 1). \tag{11}$$

where

$$E_\rho = \frac{(1 - |t|)(\rho^2 + |t|) + \rho(m - 1)|s - t^2|}{(1 - |t|)(1 - \rho + \rho^2 + \rho|t|) + \rho(m - 1)|s - t^2|}, \tag{12}$$

$$t = \frac{1}{\rho m}\left(\frac{\overline{c}_{m-1}}{\overline{c}_m}\right) \tag{13}$$

$$s = \frac{2}{\rho^2 m(m - 1)}\left(\frac{\overline{c}_{m-2}}{\overline{c}_m}\right). \tag{14}$$

Proof

Let $F(w) = w^m \overline{f\left(\frac{1}{\overline{w}}\right)}$. Then on $|w| = 1$.

$$\left|F'(w)\right| = \left|mf(w) - wf'(w)\right|. \tag{15}$$

For $\beta \in \mathbb{C}$, polar derivative of $f(w)$ with respect to β is

$$D_\beta f(w) = mf(w) + (\beta - w)f'(w).$$

Therefore,

$$\left|D_\beta f(w)\right| = \left|mf(w) + (\beta - w)f'(w)\right|$$

$$\leq \left|mf(w) - wf'(w)\right| + |\beta|\left|f'(w)\right| \tag{16}$$

$$= \left|F'(w)\right| + |\beta|\left|f'(w)\right| \tag{17}$$

$$= \left|F'(w)\right| + \left|f'(w)\right| + (|\beta| - 1)\left|f'(w)\right|. \tag{18}$$

Using (5) of Lemma 1 in (18) we have on $|w| = 1$

$$\left|D_\beta f(w)\right| \leq m\mathrm{Max}(f, 1) + (|\beta| - 1)\mathrm{Max}\left(f', 1\right). \tag{19}$$

Now using (7) of Lemma 3 in (19) we have for $|w| = 1$

$$\left|D_\beta f(w)\right| \leq m\left(1 + \frac{|\beta| - 1}{\rho^m + \rho^{m-1}}E_\rho\right)\mathrm{Max}(f, 1),$$

which proves (a).
We now prove (b).
Using (9) of Lemma 5 in (17), we get for $|w| = 1$

$$\left|D_\beta f(w)\right|$$

$$\leq \rho\left|f'(w)\right| + |\beta|\left|f'(w)\right| \tag{20}$$

$$\leq (|\beta| + \rho)\mathrm{Max}\left(f', 1\right). \tag{21}$$

Using (7) of Lemma 3, we obtain for $|w| = 1$.

$$\left|D_\beta f(w)\right| \leq m\left(\frac{|\beta| + \rho}{\rho^m + \rho^{m-1}}E_\rho\right)\mathrm{Max}(f, 1),$$

which proves *(b)*.

Remark 2

Dividing inequalities (10) and (11) by $|\beta|$ and taking $|\beta| \to \infty$, both reduce to

$$\text{Max}\left(f', 1\right) \leq \frac{m}{\rho^m + \rho^{m-1}} E_\rho \text{Max}(f, 1),$$

which is the conclusion of Lemma 3, where E_ρ is given by (12).

Remark 3

Inequalities (10) and (11) are improved extensions of inequality (3) to polar derivative. In other words, the ordinary form of (10) and (11) obtained in Remark 2 is an improvement of (3). To see this, it is sufficient to show that

$$E_\rho \leq 1.$$

That is,

$$\frac{(1 - |t|)\left(\rho^2 + |t|\right) + \rho(m - 1)\left|s - t^2\right|}{(1 - |t|)\left(1 - \rho + \rho^2 + \rho|t|\right) + \rho(m - 1)\left|s - t^2\right|} \leq 1.$$

i.e., $\rho^2 + |t| \leq 1 - \rho + \rho^2 + \rho|t|$

which holds as $\rho \leq 1$ and $|t| \leq 1$ (by Remark 1).

References

1. Aziz A (1988) Inequalities for the polar derivative of a polynomial. J Approx Theory 55:183–193
2. Chanam B, Babina Devi KH, Krishnadas K, Birkramjit Singh T (2020) Improved version of an inequality for the derivative of a polynomial. Far East J Math Sci 127(1):61–70
3. Chanam B, Devi KB, Krishnadas K, Devi MT, Ngamchui R, Singh TB (2021) on an inequality of S. Bernstein. Nonlinear Funct Anal Appl 26(2):373–380
4. Dewan KK, Mir A (2007) Note on a Theorem of S. Bernstein. Southeast Asian Bull Math 31:691–695
5. Govil NK (1980) On a Theorem of S. Bernstein. Proc Nat Acad Sci 50:50–52
6. Govil NK (1980) On a Theorem of S. Bernstein. J Math Phy Sci 14(2):183–187
7. Govil NK, Rahman QI (1969) Functions of exponential type not vanishing in a half-plane and related polynomials. Trans Amer Math Soc 137:501–517
8. Malik MA (1969) On a derivative of a polynomial. J London Math Soc 1:57–60

Universe with Power Law Expansion

Check for updates

S. Surendra Singh and Nikhil Swami

1 Introduction

Notion of inflation was established in the study of the very early cosmos to solve numerous cosmological problems such as flatness problem, entropy problem, horizon problem, monopole problem, and so on. Even though there are numerous competing solutions to the above-mentioned difficulties of hot Big Bang (BB) model, we don't possess a fully functional inflationary model (IM). If an IM can reheat the cosmos, it is termed a feasible IM. When the inflationary phase has finished, reheating begins, raising the temperature of a very cold cosmos, making this period extremely important for our universe. No feasible model could tackle the challenges of hot BB model and the graceful exit issue of old IM until 1982.

An IM was proposed by Linde [1] known as the "New" IM. This IM provided answers to the difficulties of a hot BB and elegant exit. In contrast to the old IM, which showed the cosmos to be inhomogeneous, the new IM depicts it as homogeneous. The Cosmic Microwave Background power spectrum has recently been observed to be identical to the order of 10^{-5} [2, 3] demonstrating greater success for the new IM than the old IM. Many model observations show that the cosmos is presently going through a period of accelerated expansion. Following the discovery that cosmic expansion is speeding up [4, 5] subsequent Balloon-born experiments such as Boomergang [6] and Maxima [7] have identified the anisotropic spectrum of the CMBR observation of a flat universe. This evidence suggests that present mainstream model of cosmology is influenced by dark energy, an unclustered fluid with a huge -ve pressure that is reason for the universe's expansion. Spergel et al. [8] also discovered that the cosmos is spatially flat, which accounts for 70% of dark energy. Alternative theories to be found such as $f(R)$ gravity [9–14] and $f(T)$ gravity [15, 16]. The Einstein-Hilbert (E–H) action has

S. S. Singh · N. Swami (✉)
Department of Mathematics, NIT Manipur, Imphal 795004, India
e-mail: nikhilswami6828@gmail.com

demonstrated the modifications of general relativity. Harko et al. [17] proved the greatest continuation of E–H action by adopting gravitational Lagrangian in the form of an arbitrary function of the R and matter Lagrangian Lm. $f(R, T)$ gravity is a generalised form of $f(R)$ gravity [17]. Trace dependency must be caused by exotic imperfect fluid or quantum processes. They showed three different variations of arbitrary function $f(R, T)$. Ibotombi et al. [18, 19] provided power law and exponential law based on bulk viscous cosmological models in Lyra's manifold and scale covariant theory of gravity. Adhav [20] studied the anisotropic perfect fluid cosmological model within the context of this theory and Ahmed et al. [21] explored the BT-V model for particular form $f(R, T) = f1(R) + f2(T)$, where they took cosmological constant Λ as a function of T. Sahoo et al. [22] investigated Locally Rotationally Symmetric BT-I model in the context of this theory with variable $\Lambda(T)$ and got many surprising solutions. Sahoo et al. [23] examined the physical and geometrical solutions to the variable deceleration parameter in anisotropic cosmological models under this theory. Singh et al. [24] studied power law inflation on Lyra's manifold with an anisotropic fluid and discovered that cosmos is non-isotropic at the beginning of universe and becomes isotropic afterwards. Singh et al. [25] examined the dynamical properties of non-isotropic dark energy in gravity theory and discovered that values of matter and dark energy densities Ω_m and Ω_Λ are in complete conciliation with WMAP statistics over the previous five years. For the very first time in this gravity theory, S. Bhattacharjee et al. [26] offered a modelling of inflationary scenarios. Singh [27] looked into the theory in a 5D universe and established that dark energy is important in the Kaluza–Klein world as wet dark fluid, as well as the fact that anisotropic and new isotropic models of the Kaluza–Klein universe can be developed. Another work by Singh [28] examined dark energy in the context of this modified theory from Locally Rotationally Symmetric BT-I metric.

Current work was motivated by the previous work in order to investigate power law inflation in $f(R, T)$ theory and organised as follows: We derived FE of this theory in Sect. 2. In Sect. 3 model with power law has been discussed. Energy conditions and model's observational parameters with power law are then discussed. The energy conditions and other model observational parameters are then discussed in Sect. 4. We examine our findings and concluded in Sect. 5.

1.1 $f(R, T)$Gravity and Its FE

Anisotropic Locally Rotationally Symmetric BT-I model is defined by the metric as in an orthogonal frame.

$$ds^2 = dt^2 - A^2 dx^2 - B^2(dy^2 + dz^2) \tag{1}$$

The cosmic scale factors are denoted by A and B. This metric exhibits symmetry about z-axis and has a symmetric plane in conjunction with xy-plane. Tensor of matter's energy momentum is given as

$$T^i_j = \text{diagonal}\big[\rho, -\rho_x, -\rho_y, -\rho_z\big]. \tag{2}$$

And can be parametrized as

$$
\begin{aligned}
T^i_j &= \text{diagonal}\big[\rho, -p_x, -p_y, -p_z\big]\\
&= \text{diagonal}\big[1, -\omega_x, -\omega_y, -\omega_z\big]\rho\\
&= \text{diagonal}[1, -\omega, -(\omega + \delta), -(\omega + \delta)]\rho
\end{aligned} \tag{3}
$$

Energy density is denoted by ρ, and the symbols $p_x.$, p_y, and p_z respectively, stand for pressures across the x, y,. and z axes. ω_x, ω_y, and ω_z, respectively, are fluid's directional EoS parameters along x, y, and z axes. By establishing $\omega_x = \omega_y = \omega_z$, we can now parameterize the deviation from isotropy. The divergence from ω on the y and z axes is the skewness term δ, which is introduced after that. δ and ω ain't necessary constants in this case, and they can be considered as functions of t (cosmic time). FE of this theory are determined using E–H variational principle. For this theory, Harko et al. [17] utilise subsequent action

$$
\begin{aligned}
S &= \frac{1}{16\pi}\int f(R, T)\sqrt{-g}\,d^4x + \int L_m\sqrt{-g}\,d^4x\\
&= \int \sqrt{-g}\left(\frac{1}{16}f(R, T) + L_m\right)d^4x
\end{aligned} \tag{4}
$$

with T as trace of energy momentum tensor T_{ij}, R as scalar curvature, L_m is matter Lagrangian density and g denotes metric determinant. Taking $f(R, T) = R + 2f(T)$, we vary action into Eq. (4) w.r.t g_{ij}, and obtain the FE of the theory as

$$R_{ij} - \frac{1}{2}g_{ij}R = 8\pi G T_{ij} + 2f_T T_{ij} + [2pf_T + f(T)]g_{ij} \tag{5}$$

here G denotes gravitational constant, R_{ij} is Ricci scalar, and $g_{ij}u^i u^j = 1$. We let function $f(T) = \mu T$ with μ as constant, then for metric (1), obtain the FE as

$$\left(\frac{\dot{B}}{B}\right)^2 + 2\left(\frac{\dot{A}\dot{B}}{AB}\right) = -\rho[8\pi G + 2\mu + 1 - 3\omega - 2\delta] - 2\mu p \tag{6}$$

$$\left(\frac{\dot{B}}{B}\right)^2 + 2\frac{\ddot{B}}{B} = \rho[(8\pi G + 2\mu)(\omega) - (1 - 3\omega - 2\delta)] - 2\mu p \tag{7}$$

$$\frac{\ddot{A}}{A} + \frac{\ddot{B}}{B} + \frac{\dot{A}\dot{B}}{AB} = \rho[(8\pi G + 2\mu)(\omega + \delta) - (1 - 3\omega - 2\delta)] - 2\mu p \tag{8}$$

here the overhead indicates the differentiation w.r.t t. Spatial volume (V) can be calculated as follows:

$$V = a^3 = AB^2$$

$$a = (AB^2)^{\frac{1}{3}} \tag{9}$$

a stands for the universe's scalar factor.

The following formula is used to compute average Hubble constant or parameter (H):

$$H = \frac{\dot{a}}{a} = \frac{1}{3}\left(\frac{\dot{A}}{A} + 2\frac{\dot{B}}{B}\right) \tag{10}$$

In x, y, and z axes' directions, directional H can be described as

$$H_x = \frac{\dot{A}}{A},$$

$$H_y = H_z = \frac{\dot{B}}{B} \tag{11}$$

The shear term σ^2 and the expansion term θ are provided by

$$\theta = \frac{\dot{A}}{A} + 2\frac{\dot{B}}{B} \tag{12}$$

and

$$\sigma^2 = \frac{1}{2}\left[\Sigma H_i^2 - 3H^2\right] \tag{13}$$

When we deduct (7) from (8), we obtain

$$\frac{d}{dt}\left(\frac{\dot{A}}{A} - \frac{\dot{B}}{B}\right) + \left(\frac{\dot{A}}{A} - \frac{\dot{B}}{B}\right)\frac{\dot{V}}{V} = -(8\pi G + 2\mu)\delta\rho \tag{14}$$

When we integrate the equation above, we get

$$\frac{\dot{A}}{A} - \frac{\dot{B}}{B} = \frac{\lambda}{V}\exp\int\frac{(8\pi G + 2\mu)}{\left(\frac{\dot{B}}{B} - \frac{\dot{A}}{A}\right)}\rho\delta dt \tag{15}$$

where the integrating constant is λ. We assume the following form to determine the exact solution to Eq. (15):

$$\frac{\rho\delta}{\frac{\dot{B}}{B} - \frac{\dot{A}}{A}} = \frac{1}{t} \tag{16}$$

from (15) and (16), we achieve

$$\frac{\dot{A}}{A} - \frac{\dot{B}}{B} = \frac{\lambda}{V} t^{(8\pi G + 2\mu)} \tag{17}$$

2 Model with Power Law

To thoroughly solve FE of $f(R, T)$, we assumed the following Power Law. $a = a_0 t^n$ so using value of a in (9), we get

$$
\begin{aligned}
V &= a^3 \\
&= AB^2 \\
&= a_0^3 t^{3n}
\end{aligned}
\tag{18}
$$

where $a_0 > 0$ and $n \geq 0$ are constants. Using (18) in (17) and on integrating, it gives

$$A = BC_1 \exp\left[\frac{\lambda t^{(8\pi G + 2\mu - 3n + 1)}}{a_0^3 (8\pi + 2\mu - 3n + 1)}\right] \tag{19}$$

where integration constant is C_1. From above expression of A and B, we see that their rates of expansion are different in the different directions. Using (18) and (19), we now get A and B as follows:

$$A = a_0 t^n C_1^{\frac{2}{3}} \exp\left[\frac{2\lambda t^{(8\pi G + 2\mu - 3n + 1)}}{3a_0^3 (8\pi G + 2\mu - 3n + 1)}\right] \tag{20}$$

$$B = \frac{a_0 t^n}{C_1^{\frac{1}{3}}} \exp\left[\frac{-\lambda t^{(8\pi G + 2\mu - 3n + 1)}}{3a_0^3 (8\pi G + 2\mu - 3n + 1)}\right] \tag{21}$$

here integration constant is C_1. The following are the directional H for this model:

$$H_x = \frac{\dot{A}}{A} \text{ and } H_y = \frac{\dot{B}}{B}.$$

On solving we get,

$$H_x = \frac{n}{t} + \frac{2\lambda t^{(8\pi G + 2\mu - 3n + 1)}}{3a_0^3} \tag{22}$$

and

$$H_y = \frac{n}{t} - \frac{\lambda t^{(8\pi G + 2\mu - 3n + 1)}}{3a_0^3} \tag{23}$$

$$H = \frac{1}{3}\left(\frac{\dot{A}}{A} + \frac{2\dot{B}}{B}\right)$$

$$= \frac{1}{3}\left[H_x + 2H_y\right] \tag{24}$$

$$= \frac{n}{t} \tag{25}$$

For $n > 0$, $a > 0$, H remains positive. This demonstrates that universe is expanding as it evolves. This observation is in accordance with latest observational data. The scalar expansion is denoted by the symbol θ where $\theta = 3H$. So,

$$\theta = \left(\frac{\dot{A}}{A} + \frac{2\dot{B}}{B}\right)$$

$$= \frac{3n}{t} \tag{26}$$

It suggests that in the beginning, the universe expands at an unlimited rate and then expands and returns into the phase of initial singularity in later periods. The shear scalar σ^2 is written like this:

$$\sigma^2 = \frac{1}{2}\left[\Sigma H_i^2 - 3H^2\right]$$

$$= \frac{1}{2}\left[H_x^2 + 2H_y^2 - 3H^2\right] \tag{27}$$

Now in our model we are using the condition,

$$\omega + \delta = 0 \tag{28}$$

Now using above condition and Eq. (22) and (23) in (6), (7), and (8), we get

$$\rho = \frac{-1}{(8\pi G + 2\mu)}\left[\frac{2n}{t^2} - \frac{\lambda t^{(8\pi G + 2\mu - 3n - 1)}(8\pi G + 2\mu)}{3a_0^3} - 6\left(\frac{\lambda t^{(8\pi G + 2\mu - 3n)}}{3a_0^3}\right)^2\right] \tag{29}$$

$$p = \frac{-1}{8\pi G + 2\mu}\left[\frac{\lambda t^{(8\pi G + 2\mu - 3n - 1)}[9n + 3(8\pi G + 2\mu - 3n)]}{3a_0^3}\right] \tag{30}$$

Now using the relations of equation of state, $\omega = \frac{p}{\rho}$ using this on dividing p by ρ we get our ω as

$$\omega = \cfrac{\left[\cfrac{\lambda t^{(8\pi G+2\mu-3n-1)}[9n+3(8\pi G+2\mu-3n)]}{3a_0^3}\right]}{\left[\cfrac{2n}{t^2} - \cfrac{\lambda t^{(8\pi G+2\mu-3n-1)}(8\pi G+2\mu)}{3a_0^3} - 6\left(\cfrac{\lambda t^{(8\pi G+2\mu-3n)}}{3a_0^3}\right)^2\right]} \tag{31}$$

From this mathematical expression, we see that plot of ω shifts from +ve quadrant to −ve quadrant. Thus shifting from deceleration to acceleration phase of universe is witnessed in this model. There are numerous options for obtaining values for a_0, λ, μ, and n. Finding suitable values for these parameters is all that is required to develop physically viable cosmological models. In Fig. 1, time variation of directional parameters are plotted and are decreasing functions of time in positive domain. Plot of energy density versus time is displayed in Fig. 2 and is decreasing as the universe evolves. Pressure (p) with time is shown in Fig. 3 and is always negative which implies universe's expansion. We see plot of EOS parameter against time in Fig. 4 and there is the phase transition of deceleration to acceleration.
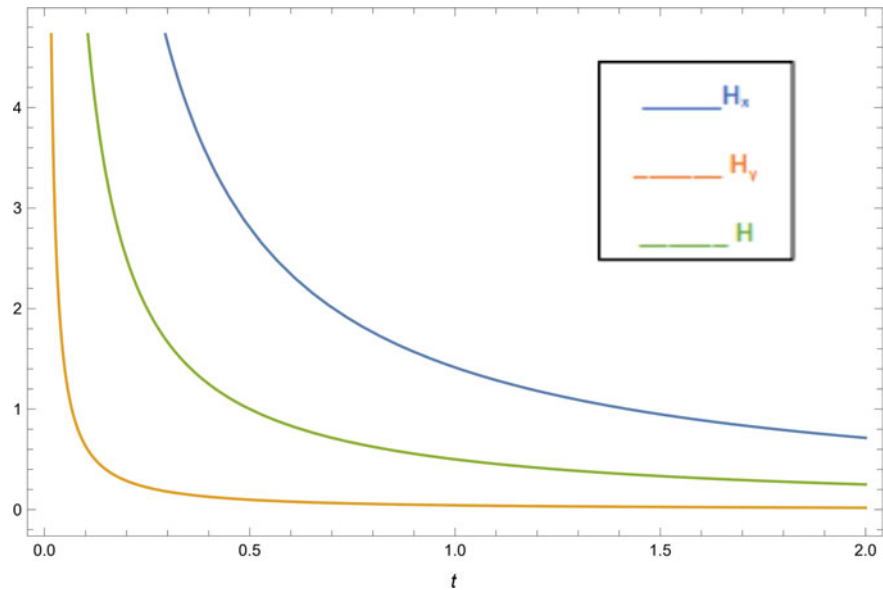
Fig. 1 Time variation of H_x, H_y, and H for $n = 0.5$, $\lambda = 1$, $\mu = 0.01$, $8\pi G = 0.5$, $a_0 = 0.9$
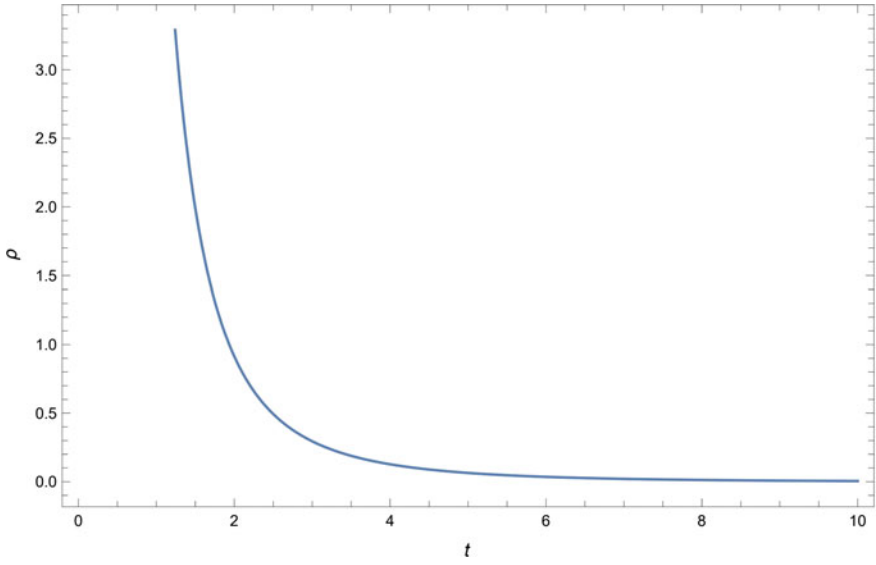
Fig. 2 Time (t) versus Density (ρ) for $n = 0.5$, $\lambda = 1$, $\mu = 0.1$, $8\pi G = 0.1$, $a_0 = 0.8$
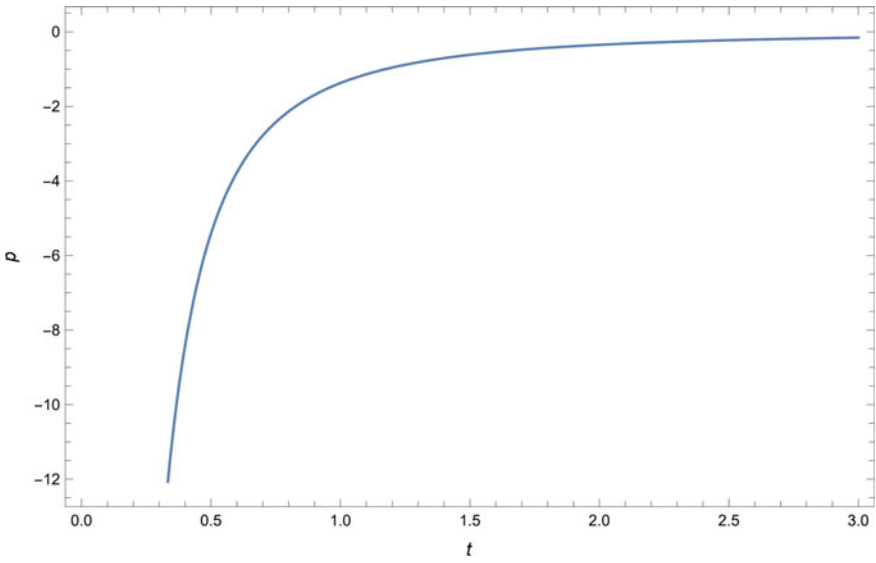
Fig. 3 Time (t) versus Pressure (p) for $n = 0.5$, $\lambda = 1$, $\mu = 0.01$, $8\pi G = 0.5$, $a_0 = 0.9$
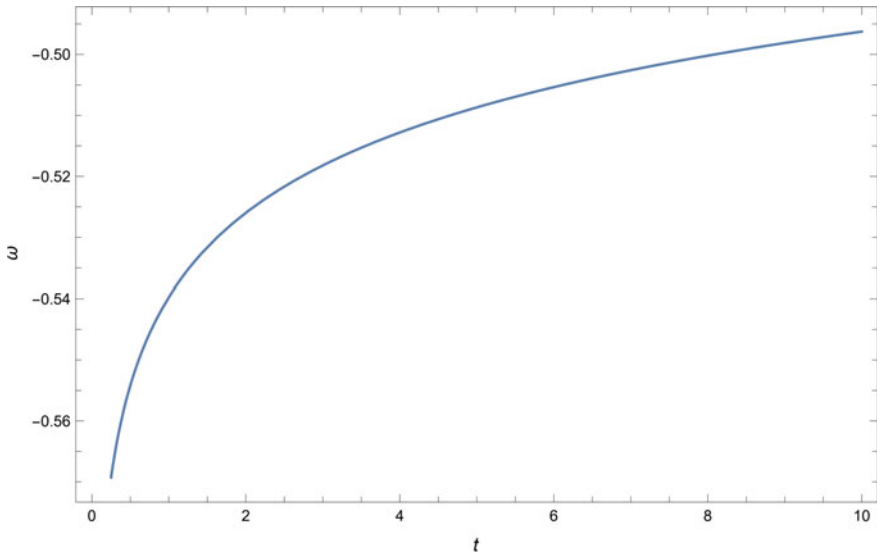
Fig. 4 EOS Parameter (ω) versus time (t) for $n = 0.5$, $\lambda = 1$, $\mu = 0.01$, $8\pi G = 0.5$, $a_0 = 0.9$

3 Energy Conditions and Some Observational Parameters

$a(t) = \frac{1}{1+z}$ is the observational setup, and the time-redshift relationship is stated as

$$t = \frac{n}{\alpha} W \left[\frac{\alpha}{n} \left(\frac{1}{a_0(1+z)} \right)^{\frac{1}{n}} \right] \tag{32}$$

The Lambert W function, commonly called the product logarithm or omega function, is denoted by W. Using the above relationship, Redshift can be used to represent the parameters of the derived model. This kind of relationship is useful for putting the model to the test with real-world data. In general relativity, energy conditions are classified into four types: weak (WEC), null (NEC), strong (SEC), and dominant (DEC) and respectively defined by

$$\text{NEC} \Leftarrow \rho + p \geq 0 \tag{33}$$

$$\text{WEC} \Leftarrow \text{NEC} \text{ and } \rho \geq 0 \tag{34}$$

$$\text{SEC} \Leftarrow \rho + 3p \geq 0 \tag{35}$$

$$\text{DEC} \Leftarrow \rho - p \geq 0 \tag{36}$$

The density remains positive, as shown in Fig. 2 at both early and late times. NEC > 0, WEK > 0, DEC > 0, and SEC < 0 were found in Figs. 5, 6, and 7. SEC is failed, whereas NEC, DEC, and WEC are all fulfilled.
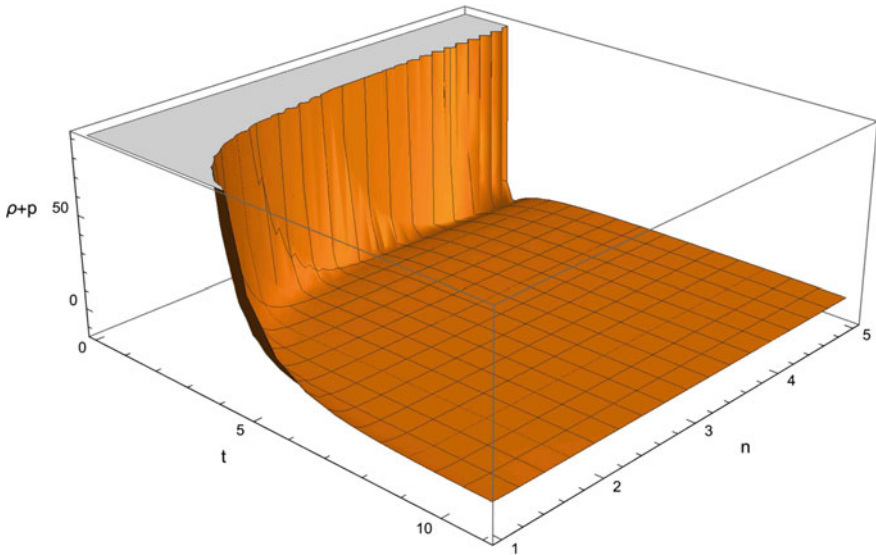
Fig. 5 3D plot of EC ($\rho + p$) versus time (t) for $n = 0.5$, $\lambda = 1$, $\mu = 0.1$, $8\pi G = 0.1$, $a_0 = 0.15$
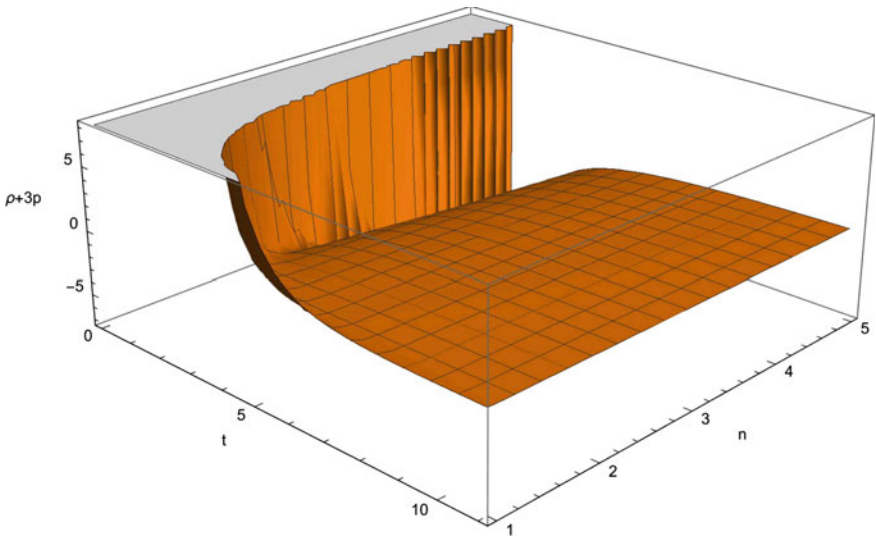
Fig. 6 3D plot of EC ($\rho + 3p$) versus time (t) for $n = 0.5$, $\lambda = 1$, $\mu = 0.1$, $8\pi G = 0.1$, $a_0 = 0.2$
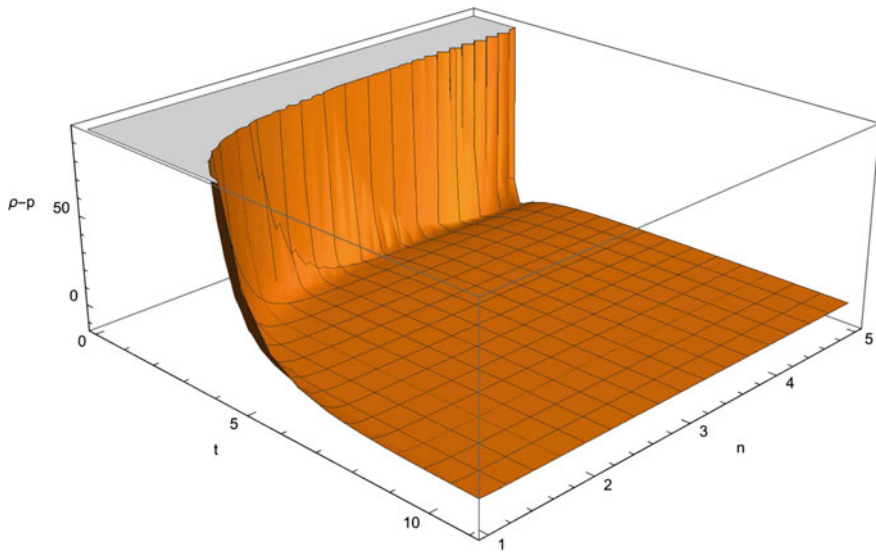
Fig. 7 3D plot of EC ($\rho - p$) versus time (t) for $n = 0.5$, $\lambda = 1$, $8\pi G = 0.1$, $\mu = 0.1$, $a_0 = 0.15$

4 Conclusion

We explored a generalised method of finding the exact solutions of Locally Rotationally Symmetric BT-I space time in this theory by using power law cosmology. Here, we assume $f(R, T) = 2\mu T + R$. In figures, energy density of universe is decreasing as ages of universe progress, and it demonstrates a positive condition that favours observation. Pressure is always negative. In Figs. 1 and 2, the parameters p and ρ becomes infinite at $t \to 0$ which suggests that universe starts from Big Bang and these parameters becomes extremely small at $t \to \infty$ which are consistent with observations. The derived model shows the characteristics of dark energy model as ω approaches to -1 with the evolution of time which is in agreement with present universe that is assumed to be dominated by dark energy. WEC, SEC, DEC, and NEC of model are found to be satisfied. The present model may be able to highlight behaviours of universe from the anisotropic behaviours at early universe to accelerated expansion at late epoch. Although this model is simple, this investigation may lead to the cosmologists for further research in modified cosmology.

References

1. Linde AD (1987) AdSAC 3:149
2. Penzias AA, Wilson RW (1965) ApJ 142:419
3. Boggess NW, Mather JC, Weiss R et al (1992) ApJ 397:420
4. Perlmutter S et al (1999) Measurements of and from 42 high-redshift. Astrophys J 517, 565

5. Reiss AG et al (1998) Observational evidence from supernovae for an accelerating universe and a cosmological constant. Astrophys J 116:1009
6. De Bernardis P et al (2000) Nature 404:955
7. Stompor R et al (2001) Astrophys J 561:L7
8. Spergel DN et al (2007) Astrophys J Suppl Ser 170:377
9. Carroll SM, Duvvuri V, Trodden M, Turner MS (2004) Is cosmic speed up due to new gravitational physics. Phys Rev D 70:043528
10. Nojiri S, Odintsov SD, Tsujikawa S (2006) Modified f(R) gravity consistent with realistic cosmology from a matter dominated epoch to a dark energy universe. Phys Rev D 74:086005
11. Nojiri S, Odintsov SD (2007) Introduction to modified gravity and gravitational alternative for dark energy. Int J Geom Methods Mod Phus 4:115
12. Nojiri S, Odintsov SD (2011) Unified cosmic history gravity from f(R) theory to lorentzn non-invariant models. Phys Rep 505:59
13. Bertolami O, Bohmer CG, Harko T, Lobo FSN (2007) Extra force in f(T) modified theories of gravity. Phys Rev D 75:104016
14. Sotiriou TP, Faraoni V (2010) f(R)Theories of gravity. Rev Mod Phys 82:451
15. Bengocheu GR (2009) Dark torsion as the cosmic speed up. Phys Rev D 79:124019
16. Linder EV (2010) Einsteins other gravity and the accelerations of the universe. Phys Rev D 81:127301
17. Harko T, Lobo FS, Nojiri SI, Odintsov SD (2011) f (R, T) gravity. Phys Rev D 84(2):024020
18. Ibotombi Singh N, Romaleima Devi S, Surendra Singh S, Sumati Devi A (2009) Bulk viscous cosmological models of universe with variable deceleration parameter in Lyra's Manifold. Astrophys Space Sci 321:233–239
19. Ibotombi Singh N, Surendra Singh S, Romaleima Devi S (2010) A new class of bulk viscous cosmological models in a scale covariant theory of gravitation. Astrophys Space Sci 326:293–297
20. Adhav KS (2012) LRS Bianchi type-i cosmological model in f(R, T) theory of gravity. Astrophys Space Sci 339:365
21. Ahmed N, Pradhan A (2014) Bianchi type-v cosmology in f(R, T) gravity with. Int J Theor Phys 53:289306
22. Sahoo PK, Sivakumar M (2015) LRS Bianchi type-I cosmological model in f (R, T) theory of gravity with Λ (T). Astrophys Space Sci 357:1–2
23. Sahoo PK, Sahoo P, Binaya KB (2017) Anisotropic cosmological models in f(R, T) gravity with variable deceleration parameter. Int J Geom Meth Mod Phys 14(1750097):25
24. Singh SS, Devi YB, Singh MS (2017) Power law inflation with anisotropic fluid in Lyra's manifold. https://doi.org/10.1139/cjp-2016-0897
25. Singh MS, Singh SS (2019) Cosmological dynamics of anisotropic dark energy in f (R, T) gravity. New Astron 72:36–41
26. Bhattacharjee S, Santos JR, Moraes PH, Sahoo PK (2020) Inflation in f (R, T) gravity. Eur Phys J Plus 135(7):576
27. Surendra Singh S, Manihar Singh K, Kumrah L (2021) Kaluza–Klein Universe interacting with wet dark fluid in f (R, T) gravity. Int J Mod Phys A 36(07):2150043
28. Alam MK, Singh SS, Devi LA (2022) Adv High Energy Phys 2022(Article ID5820222). https://doi.org/10.1155/2022/5820222

An Optimal Fourth-Order Iterative Method for Multiple Roots of Nonlinear Equations

Waikhom Henarita Chanu, Sunil Panday, Shubham Kumar Mittal, and G Thangkhenpau

1 Introduction

The problems of real world natural phenomena are usually express by nonlinear equation which exact roots are infeasible due to inherent complexities. Analytical methods for solving equations are not applicable for such type of equations. Due to this inconvenience, we use iterative methods for solving nonlinear equation. Most of researchers in Numerical Analysis are trying to construct iterative methods for solving nonlinear equations.

Newton's method is one popular iterative method for solving nonlinear equations, which quadraticaly converges for simple roots but linear for multiple roots. For a nonlinear equation $\zeta(\upsilon) = 0$ having multiple roots with multiplicity $\omega \geq 1$ modified Newton's methods is given as:

$$\upsilon_{n+1} = \upsilon_n - \omega \frac{\zeta(\upsilon_n)}{\zeta'(\upsilon_n)}, \tag{1}$$

This iterative method (1) is quadraticaly convergence [1, 2].

Chebyshev's Method for multiple is given as follows:

$$\upsilon_{n+1} = \upsilon_n - \frac{\omega(3-\omega)}{2} \frac{\zeta(\upsilon_n)}{\zeta'(\upsilon_n)} - \frac{\omega^2}{2} \frac{\left(\zeta(\upsilon_n)^2\right)\zeta'(\upsilon_n)}{\zeta'(\upsilon_n)^3}, \tag{2}$$

which is third order of convergence [1]. However, these are one point iteration functions. The one point iteration functions required higher-order derivative to increase

W. H. Chanu (✉) · S. Panday · S. K. Mittal · G. Thangkhenpau
Department of Mathematics, National Institute of Technology Manipur, Langol, Imphal, Manipur 795004, India
e-mail: henaritawai@gmail.com

the order of convergence (see in Traub [1] and Ostrowski [2]). To reduce this diffi-
culties, several research scholar whose are in the field of Numerical Analysis, are
trying to get multipoint iteration function of higher-order [3].

In the year 2009, Shengguo et al. [4] construct an iterative scheme for multiple
root which is fourth-order of convergence.

$$y_n = v_n - \frac{2\omega}{\omega + 1} \frac{\zeta(v_n)}{\zeta'(v_n)}$$

$$v_{n+1} = v_n - \frac{\frac{1}{2}\omega(\omega - 2)\left(\frac{\omega}{\omega+2}\right)^{\omega}\zeta'(y_n) - \frac{\omega^2}{2}\zeta'(v_n)}{\zeta'(v_n) - \left(\frac{\omega}{\omega+2}\right)^{\omega}\zeta'(y_n)} \tag{3}$$

Li et al. [5] introduced a fourth-order scheme for multiple roots, in 2010.

$$y_n = v_n - \frac{2\omega}{\omega + 2} \frac{\zeta(v_n)}{\zeta'(v_n)}$$

$$v_{n+1} = v_n - a_3\frac{\zeta(v_n)}{\zeta'(y_n)} - \frac{\zeta(v_n)}{b_1\zeta'(v_n) + b_2\zeta'(y_n)} \tag{4}$$

where

$$a_3 = -\frac{(\omega - 2)\omega\left(\frac{\omega}{\omega+2}\right)^{\omega}(\omega + 2)^3}{2(\omega^3 - 4\omega + 8)}$$

$$b_1 = -\frac{(\omega^3 - 4\omega + 8)^2}{\omega(\omega^2 + 2\omega - 4)(\omega^4 + 4\omega^3 - 4\omega^2 - 16\omega + 16)}$$

$$b_2 = \frac{\omega^2\left(\frac{\omega}{\omega+2}\right)^{-\omega}(\omega^3 - 4\omega + 8)}{(\omega^2 + 2\omega - 4)(\omega^4 + 4\omega^3 - 4\omega^2 - 16\omega + 16)}$$

In 2019, Bhel and Al-Hamadan [5] presented a fourth-order method for multiple
roots which is optimal:

$$y_n = v_n - \omega\frac{\zeta(v_n)}{\zeta'(v_n)}$$

$$z_n = v_n - \omega\frac{\zeta(y_n)}{\zeta'(v_n)}\left(\frac{1 - \mu}{1 - 2\mu}\right)Q(\mu). \tag{5}$$

where $\mu = \left(\frac{\zeta(y_n)}{\zeta(v_n)}\right)^{\frac{1}{\omega}}$ and $Q(\mu)$ is weight function.

In this article, we present an optimal fourth-order iterative scheme for multiple
roots using weight function. We check the behaviour of the developed scheme a
using numbers nonlinear examples. Form the result, it is notice that the developed

method perform better as compare to other standard iterative schemes available in the literature. The remaining part of this article is sort out as follows. In the second section, we are presenting two fourth-order iterative schemes function using weight function and the proof is also provided. In the third section, we present the comparison result of existing methods with the new methods. Conclusion is presented in the last section.

2 Development of the Method

In the year 2019, Francisco et al. [6] present an iterative scheme which is written as follows:

$$y_n = \upsilon_n - \frac{\zeta(\upsilon_n)}{\zeta(\upsilon_n)}$$

$$\upsilon_{n+1} = \upsilon_n - \frac{\zeta^{(2)}(\upsilon_n) + \zeta(\upsilon_n)\zeta(y_n) + 2\zeta^{(2)}(y_n)}{\{\zeta(\upsilon_n)\zeta'(\upsilon_n)\}} \tag{6}$$

The order of convergence of the scheme defined in Eq. (6) is four for the nonlinear functions having simple roots. We are trying to improve the method in Eq. (6) to an iterative scheme for solving nonlinear equations having multiple roots. The scheme is given as follows:

$$y_n = \upsilon_n - \omega\frac{\zeta(\upsilon_n)}{\zeta'(\upsilon_n)}$$

$$\upsilon_{n+1} = \upsilon_n - \omega\Gamma(t_n)\frac{\zeta^2(\upsilon_n) + \zeta(\upsilon_n)\zeta(y_n) + 2\zeta^2(y_n)}{\zeta(\upsilon_n)\zeta'(\upsilon_n)} \tag{7}$$

where $t_n = \left(\frac{\zeta(y_n)}{\zeta(\upsilon_n)}\right)^{\frac{1}{\omega}}$

Theorem 1 If $\zeta : R \rightarrow R$ has a multiple zero β with multiplicity $\omega = 2$ and is sufficiently differentiable function in the neighbourhood of the roots β. . Then, the order of convergence of iterative methods defined in Eq. (7) is four, if $\Gamma(t_n)$ satisfies $\Gamma(0) = 1$, $\Gamma'(0) = 1$, and $\Gamma''(0) = 2$, then (7) has the following error equation

$$\upsilon_{n+1} - \beta = \left(-\frac{1}{48}k_1^3\big(\Gamma^{(3)}(0) - 27\big) - \frac{k_2k_1}{4}\right)e_n^4 + O\big(e_n^5\big) \tag{8}$$

Proof
Since β is a root of $\zeta(\upsilon)$ with multiplicity $\omega = 2$, let $e_n = \upsilon_n - \beta$ be the error at the nth iteration. Using Taylor's series expansion on $\zeta(\upsilon_n)$ and $\zeta'(\upsilon_n)$ we get:

$$\zeta(\upsilon_n) = \frac{\zeta''(\beta)}{2!}e_n^2\left(1 + e_n k_1 + e_n^2 k_2 + e_n^3 k_3 + e_n^4 k_4 + O(e_n)^5\right) \tag{9}$$

And

$$\zeta'(\upsilon_n) = \frac{\zeta^{(2)}(\beta)}{2!}e_n^2(2 + 3k_1 e_n + 4k_2 e_n^2 + 5k_3 e_n^3 + 6k_4 e_n^4 + O(e_n)^5 \tag{10}$$

where $k_i = \frac{2!}{(2+i)!}\frac{\zeta^{(2+i)}(\beta)}{\zeta^{(2)}(\beta)}, i = 1, 2, 3, \ldots$
Using (9) and (10) in the second step of (7)

$$y_n - \beta = \frac{k_1 e_n^2}{2} + \left(k_2 - \frac{3k_1^2}{4}\right)e_n^3 + \left(\frac{9k_1^3}{8} - \frac{5k_2 k_1}{2} + \frac{3k_3}{2}\right)e_n^4 + O(e^5) \tag{11}$$

Taylor's expansion of $\zeta(y_n)$ about β, we have

$$\zeta(y_n) = \frac{\zeta^{(2)}(\beta)}{2!}e_n^2\left(\frac{1}{4}k_1^2 e_n^2 + k_1\left(k_2 - \frac{3k_1^2}{4}\right)e_n^3 + \left(\frac{29k_1^4}{16} - 4k_2 k_1^2 + \frac{3}{2}k_3 k_1 + k_2^2\right)e_n^4\right)$$
$$+ O(e_n^5) \tag{12}$$

Using (9) and (12), we get

$$\upsilon_{n+1} - \beta = (1 - \Gamma(0))e_n + \frac{1}{2}k_1\left(\Gamma(0) - \Gamma'(0)\right)e_n^2$$
$$+ \left(k_2(\Gamma(0) - \Gamma'(0)) - \frac{1}{8}k_1^2\left(-10\,\Gamma'(0) + \Gamma''(0) + 8\Gamma(0)\right)\right)e_n^3$$
$$+ \frac{1}{48}\left(\begin{array}{c}k_1^3\left(27\left(-5\,\Gamma'(0) + \Gamma''(0) + 4\,\Gamma(0)\right) - \Gamma^{(3)}(0)\right)\\ - 12\,k_2\,k_1\left(-17\,\Gamma'(0) + 2\,\Gamma''(0) + 14\,\Gamma(0)\right) + 72\,k_3\left(\Gamma(0) - \Gamma'(0)\right)\end{array}\right)e_n^4$$
$$+ O(e_n^5) \tag{13}$$

Using (9) (12) and (13) in the last step of (7).

$$\upsilon_{n+1} - \beta = (1 - \Gamma(0))e_n + \frac{1}{2}k_1\left(\Gamma(0) - \Gamma'(0)\right)e_n^2$$
$$+ \left(k_2(\Gamma(0) - \Gamma'(0)) - \frac{1}{8}k_1^2\left(-10\Gamma'(0) + \Gamma''(0) + 8\Gamma(0)\right)\right)e_n^3$$
$$+ \frac{1}{48}\left(\begin{array}{c}k_1^3\left(27\left(-5\Gamma'(0) + \Gamma''(0) + 4\Gamma(0)\right) - \Gamma^{(3)}(0)\right)\\ - 12k_2 k_1\left(-17\Gamma'(0) + 2\Gamma''(0) + 14\Gamma(0)\right) + 72k_3\left(\Gamma(0) - \Gamma'(0)\right)\end{array}\right)e_n^4$$
$$+ O(e_n^5) \tag{14}$$

If we are putting $\Gamma(0) = 1$, $\Gamma'(0) = 1$ and $\Gamma''(0) = 2$, error equation become

$$\upsilon_{n+1} - \beta = \left(-\frac{1}{48}k_1^3\left(\Gamma^{(3)}(0) - 27\right) - \frac{k_2k_1}{4}\right)e_n^4 + O\left(e_n^5\right) \tag{15}$$

Thus the proof is completed.

Theorem 2 If $\zeta : R \to R$ has a multiple zero β with multiplicity $\omega \geq 3$ and is sufficiently differentiable function in the neighbourhood of the roots β. Then, the order of convergence of iterative methods defined in Eq. (7) is four, if $\Gamma(t_n)$ satisfies $\Gamma(0) = 1$, $\Gamma'(0) = 1$, and $\Gamma''(0) = 4$, then (7) has the following error equation

$$\upsilon_{n+1} - \beta = \left(\frac{k_1^3\left(\Gamma^{(3)}(0) + 3\omega + 27\right) - 6k_2k_1\omega}{6\omega^3}\right)e_n^4 + O\left(e_n^5\right) \tag{16}$$

Proof
The proof is same as the proof of theorem 1.

3 Numerical Results

In this section, we deal with computational aspects of the proposed scheme with other existing methods such as fourth-order methods given in Eq. (3) denoted as (LM), method in Eq. (4) denoted as (LCM) and Behl Method (BM) given in Eq. (5) by applying on various nonlinear examples. In Tables 1, 2, 3 and 4 we have presented $|\zeta(\upsilon_n)|$, absolute of difference between the successive iterations $|\upsilon_n - \upsilon_{n-1}|$. Approximate roots. (υ_n). obtained after completion of 4 iterations and the computational order of convergence (COC) for each example are also presented. The COC is obtained by using the following formula [1]:

$$\rho \approx \frac{Log\left|\frac{\upsilon_{n+1}-\beta}{(\upsilon_n-\beta)}\right|}{Log\left|\frac{\upsilon_n-\beta}{\upsilon_{n-1}-\beta}\right|}$$

Table 1 Convergence behaviour for $\zeta_1(\upsilon)$

Method	υ_4	$|\upsilon_n - \upsilon_{n-1}|$	$|\zeta(\upsilon)|$	COC
LM	1.2917332924436028	5.5927×10^{-40}	1.1200×10^{-312}	4.0000
LCM	1.2917332924436028	6.1797×10^{-36}	1.1200×10^{-312}	4.0000
BM	1.2917332924436028	1.1567×10^{-38}	2.0891×10^{-301}	4.0000
NMM	1.2917332924436028	7.0179×10^{-41}	7.664×10^{-319}	4.0000

Table 2 Convergence behaviour for $\zeta_2(v)$

Method	v_4	$|v_n - v_{n-1}|$	$|\zeta(v)|$	COC
LM	2.0000000000000000	6.0530×10^{-57}	$3.5493 \times 10^{-11259}$	4.0000
LCM	2.0000000000000000	2.9024×10^{-57}	$3.3971 \times 10^{-11259}$	4.0000
BM	2.000000000000000	4.8282×10^{-59}	$8.1432 \times 10^{-11619}$	4.0000
NMM	2.0000000000000000	7.0179×10^{-61}	$7.6647 \times 10^{-11719}$	4.0000

Table 3 Convergence behaviour for $\zeta_3(v)$

Method	v_4	$. |v_n - v_{n-1}|$	$|\zeta(v)|$	COC
LM	2.8500000000000000	6.3883×10^{-84}	7.5412×10^{-370}	4.0000
LCM	2.8500000000000000	6.3883×10^{-84}	7.5412×10^{-370}	4.0000
BM	2.8500000000000000	1.3949×10^{-49}	9.8347×10^{-396}	4.0000
NMM	2.8500000000000000	1.4067×10^{-49}	1.0569×10^{-397}	4.0000

Table 4 Convergence behaviour for $\zeta_4(v)$

Method	v_4	$|v_n - v_{n-1}|$	$|\zeta(v)|$	COC
LM	2.4905398276083051	1.0312×10^{-19}	2.3983×10^{-1505}	4.0000
LCM	2.4905398276083051	4.0539×10^{-20}	8.8944×10^{-1538}	4.0000
BM	2.4905398276083051	1.9627×10^{-19}	1.7767×10^{-1477}	4.0000
NMM	2.4905398276083051	1.9627×10^{-19}	2.1011×10^{-1667}	4.0000

The numerical result has been carried out with Mathematica 12 software, * denotes for divergence.

Example 1
$\zeta_1(v) = \left(9 - 2v - 2v^4 + cos2v\right)\left(5 - v^4 - sin^2v\right)$, $v_0 = 1.5$, $\omega = 2$.

Example 2
$\zeta_2(v) = \left((v - 1)^3 - 1\right)^{50}$, $v_0 = 2.1$, $\omega = 50$

Example 3
$\zeta_3(v) = v^4 + 11.50v^3 + 47.49v^2 + 83.06325v + 51.23266875$, $v_0 = -2.7$ *and* $\omega = 2$.

Example 4
$\zeta_4(v) = \left(e^{-v^2+v+3} - v + 2\right)^{20}$, $v_0 = 2.1$ and $\omega = 20$,

4 Conclusion

We construct a new fourth-order iterative method base on Newton's and Francisco's methods using weight functional approach. It attends its optimal order. We test, by comparing the newly developed methods with other methods having the same convergence are order using several nonlinear equations having multiple roots. The results obtained from the comparison tables illustrate the superiority of the method over the existing methods, despite choosing the same test problem and the same initial guess. Tables confirm that our iterative method has smaller value of $|\zeta(v)|$ and $|v_n - v_{n-1}|$.

References

1. Traub JF (1964) Iterative methods for the solution of equations. Prentice-Hall, Englewood Cliffs, NJ, USA
2. Ostrowski AM (1960) Solution of equations and system of equation. Academic Press, New Yowk
3. Henarita Chanu W, Panday S, Dwivedi M (2020) New fifth order iterative method for finding multiple root of nonlinear function. Eng Lett 29(3):942–947
4. Shengguo L, Xiangke I, Lizhi C (2009) A new fourth order iterative method for finding multiple roots of nonlinear equations. Appl Math Comput 215:1288–1292
5. Li S, Cheng L, Neta B (2010) Some fourth order non lininera solvers with closed formulae for multiple roots. Comput Math Appl 59:126–135
6. Behl R, Al-Hamdan WM (2019) A 4th-order optimal extension of Ostrowski's method for multiple zeros of univariate nonlinear functions. Math MDPI 7(9):803

A Dual-Band Circularly Polarized with Large Impedance Bandwidth Planar Monopole Antenna for Wireless Application

Deepak Kumar Barik, Mahd Azharuddin, Kalyan Mondal, Lakhindar Murmu, and Tapan Mandal

1 Introduction

Printed microstrip patch radiator is largely used in different communication application over the last decades of years due to easy fabrication process, its compact size, and lightweight in structure. However, for multiband application, planar monopole antenna made its own space in couple of years during this modern era of communication [1]. Nevertheless, circularly polarized planar monopole antenna shows a great participation in today's communication system because of its advantages like overcoming from polarization mismatch, multipath reflection problem, and weak weather penetration power [2]. However, the problems in circularly polarized planar monopole antenna (CPMA) are large in size and small axial ratio bandwidth. Therefore, various types of technique are used [3–14] to improve size and ARBW. In this paper, all above factors are aimed to improve. A dual-band ARBW under a broad IBW with an average gain circularly polarized monopole antenna is implemented. A simple half-moon and full moon slotted along with triangular etched rectangular radiator is proposed. A simple microstrip line feeding is used asymmetrically to the radiator to improve ARBW. The details description of proposed structure, design procedures, and response are portrayed.

D. K. Barik (✉) · M. Azharuddin · K. Mondal
National Institute of Technology Manipur, Imphal, Manipur, India
e-mail: deepakkumarbarik1985@gmail.com

L. Murmu
International Institute of Information Technology, Naya Raipur, Atal Nagar, Chhattisgarh, India

T. Mandal
Government College of Engineering and Textile Technology, Murshidabad, West Bengal, India

2   Antenna Design and Analysis

The design details and the values of length and width of physical parameter are mentioned in Fig. 1. The physical antenna consists of a thin cupper radiator with a microstrip feed, substrate, and partial ground plane from top to down with respect to Z-axis as shown in Fig. 1. Here, substrate is taken as FR-4 dielectric material whose loss tangent is 0.024 and permittivity is 4.4. In this design, length (L) and width (W) are taken in Y-axis and X-axis, respectively. The physical dimension of antenna is $29 \times 31 \times 1.6$ ($L \times W \times h$) in millimeter. A full moon and half-moon-shaped slots are inserted inside the rectangular patch along with triangular-shaped cut at the right side of the radiator to improve IBW and ARBW. The position of microstrip feed line is asymmetrically connected and optimized to improve ARBW. By optimizing different parameters, ARBW and IBW are resulted. Final parameters are summarized in Table 1. The ANSYS HFSS (ver.15) is used to design and simulation of the designed antenna.

Fig. 1 Simplified diagram with all dimensions of proposed antenna
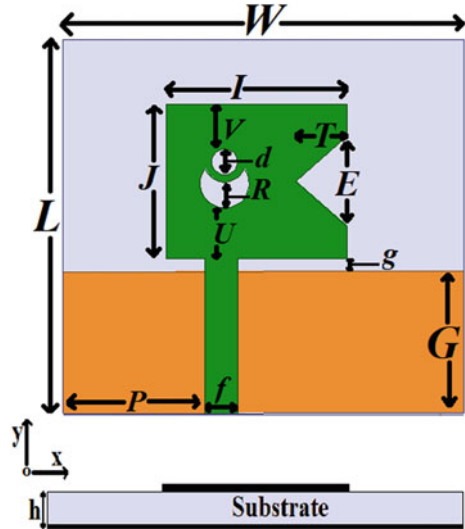
Table 1 Physical parameters of the proposed antenna (unit in millimeter)

Param.	Value	Param.	Value	Param.	Value	Param.	Value
W	31	J	12	d	2	G	11
L	29	E	7	R	2	f	2.5
h	1.6	T	4	U	3.95	P	11
I	14	V	3.5	g	1		
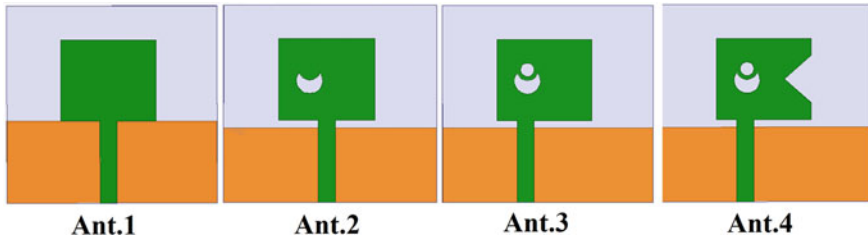
Fig. 2 Design procedure of proposed antenna
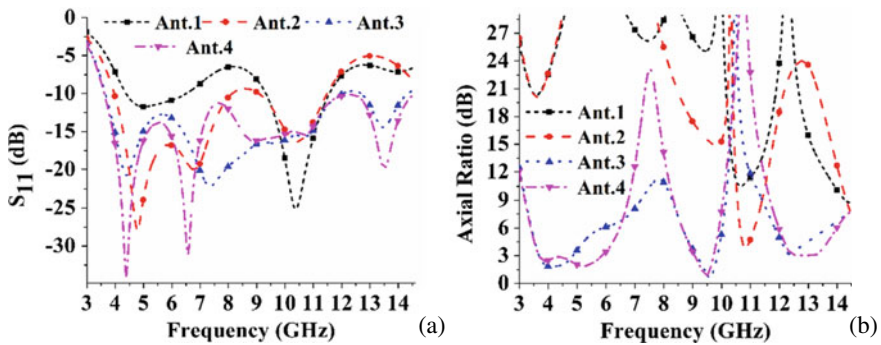
Fig. 3 Responses of Ant.1 to Ant.4: **a** S_{11} and **b** axial ratio

To get the final proposed antenna, four design steps are introduced from Ant.1 to Ant.4 (proposed) which are explained in Fig. 2. The details response of each steps are analyzed in Fig. 3.

To generate CP radiation, the horizontal and vertical component of electrical field should have equal magnitude and 90° phase difference. In Ant.1, very narrow −10 dB IBW is found at a resonant frequency of 10.46 GHz without any CP radiation in Fig. 3. However, IBW is improved largely up to 11.34 GHz and AR approaches toward 3 dB. In next step (Ant.3), IBW is further improved along with dual bands of CP radiation in Fig. 3. Finally, optimized dual band is found in proposed antenna (Ant.4) along with broad IBW. The 1st resonant frequency (f_0) of proposed antenna is found at 4.39 GHz, which is shown in Fig. 3, and its corresponding wavelength is λ_0 (1st resonant wavelength).

2.1 Surface Current Analysis to Interpret CP Radiation

The resultant current direction can be determined from the figure of surface current distribution. From that resultant current direction, circular polarization characteristic can be verified by orthogonal orientation of phase angles. Figure 4 shows that surface
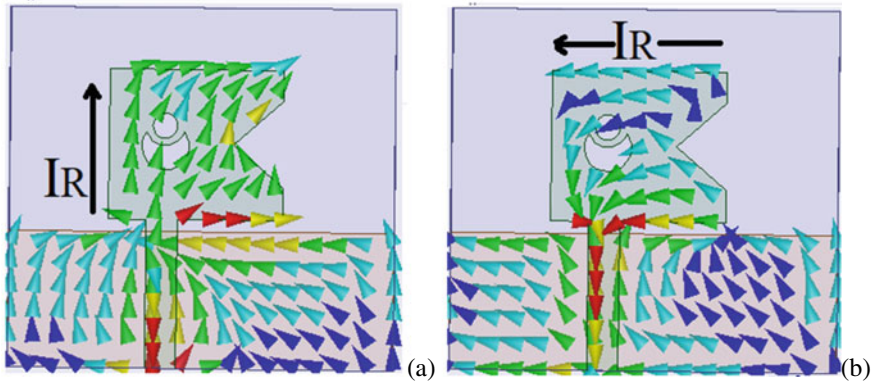
Fig. 4 Proposed antenna surface current distribution effect at 4.39 GHz for phase angle: **a** 0° and **b** 90°

current directions at 4.39 GHz for orthogonal phase angles 0° and 90°. By summing all the current vectors, the resultant current direction is found in a particular direction. All horizontal currents are canceled due to opposite direction of currents and only current is flowing in vertical up direction as shown in Fig. 4a. However, Fig. 4b shows that all verticals currents are canceled due to opposite direction of currents and only current is flowing in horizontal (right to left) direction. From Fig. 4, it is clearly shown that the current distribution oriented counter clockwise direction. From which, it is confirmed that it provides the right hand circularly polarized (RHCP) radiation in broadside direction and left hand circularly polarized (LHCP) radiation in backside direction.

3 Results and Discussion

The results of different characteristic parameters of designed antenna are displayed in Figs. 5 and 6. The IBW of CPMA is from 3.59 to 14.6 GHz confirmed in Fig. 5a. The dual band of ARBW is from 3.65 to 5.89 GHz and from 9.05 to 9.75 GHz are verified in Fig. 5b. The peak gain is found as 4.3 dBi, and average efficiency is more than 90%, which are displayed in Fig. 5c and d.

The far field radiation patterns of the proposed antenna are illustrated in Fig. 6. The YOZ plane and XOZ plane radiation patterns are plotted for variation of angle θ at 4.39, 9.45 GHz to verify the circular polarization in two different CP bands. In both frequency, RHCP and LHCP waves are radiated toward broadside and backside directions, respectively.
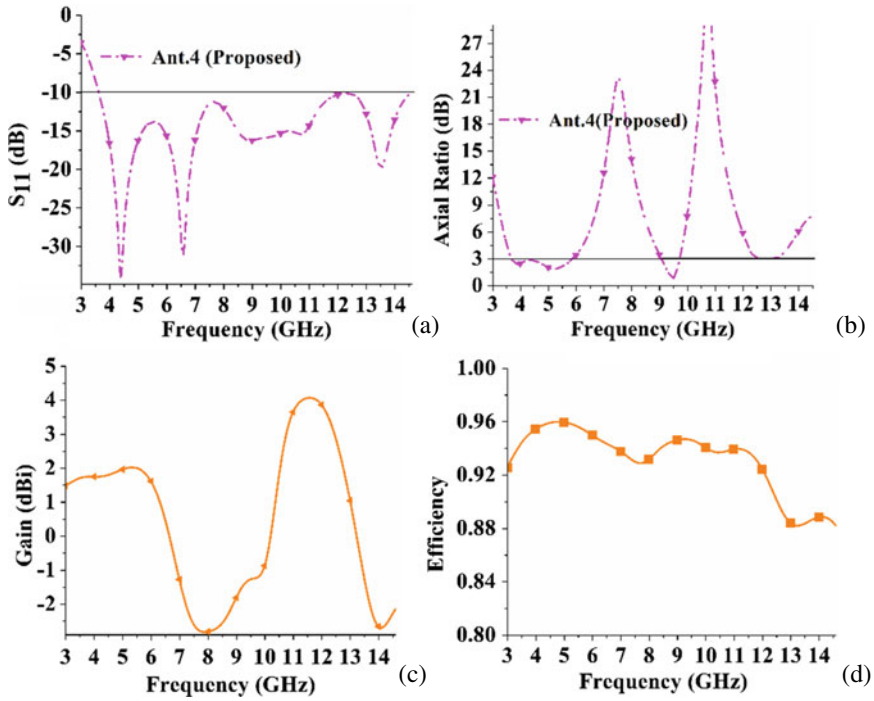
Fig. 5 Results of proposed antenna: **a** return loss, **b** AR, **c** gain, and **d** efficiency
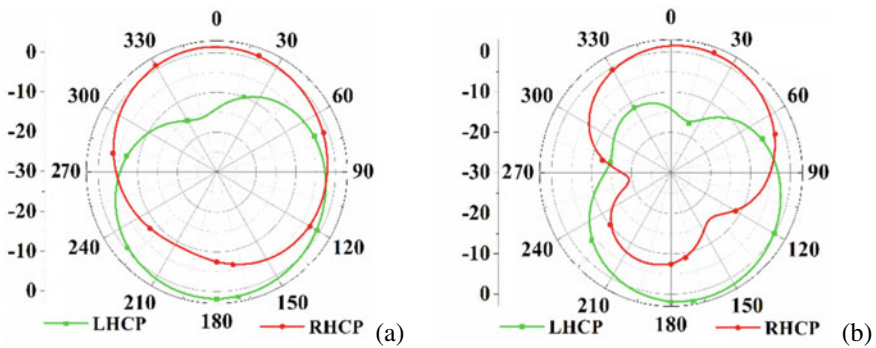
Fig. 6 Radiation pattern: **a** 4.39 GHz in XOZ plane, **b** 4.39 GHz in YOZ plane, **c** 9.45 GHz in XOZ plane, and **b** 9.45 GHz in YOZ plane
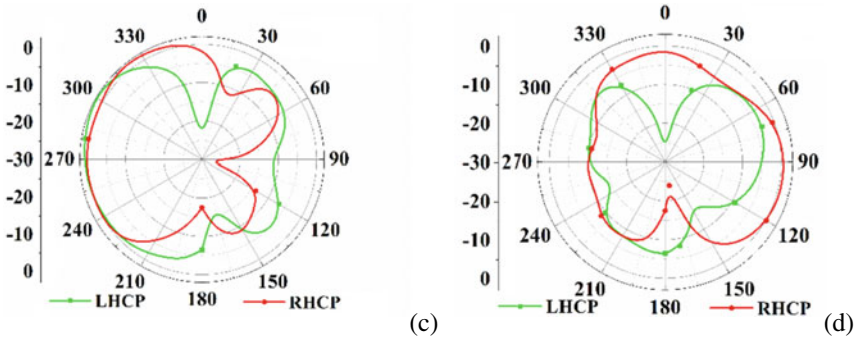
Fig. 6 (continued)

4 Conclusion

A dual-band circularly polarized planar monopole antenna is implemented with very large IBW. Multiple number of slots are inserted to the rectangular radiator to optimize the ARBW and IBW. The IBW of CPMA is 121.05% (3.59–14.6 GHz) achieved. The dual ARBW of 2.24 GHz (3.65–5.89 GHz) and 0.7 GHz (9.05–9.75 GHz) are achieved. The proposed antenna can be used for different wireless communication like 5G application, S-band radar, and maritime radio navigation.

References

1. Lin CI, Wong KL (2007) Printed monopole slot antenna for internal multiband mobile phone antenna. IEEE Trans Antennas Propag 55(12):3690–3697
2. Gao SS, Luo Q, Zhu F (2013) Circularly polarized antennas. Wiley-IEEE Press, New York
3. Liu WC, Kao PC (2007) Design of a probe-fed H-shaped microstrip antenna for circular polarization. J Electromagn Waves Appl 21(7):857–864
4. Ghobadi A, Dehmollaian M (2012) Printed circularly polarized y-shaped monopole antenna. IEEE Antennas Wirel Propag Lett 11:22–25
5. Zhou SW, Li PH, Wang Y, Feng WH, Liu ZQ (2011) A CPW-fed broadband circularly polarized regular-hexagonal slot antenna with l-shape monopole. IEEE Antennas Wirel Propag Lett 10:1182–1185
6. Ding K, Gao C, Yu T, Qu D (2014) Broadband C-shaped circularly polarized monopole antenna. IEEE Trans Antennas Propag 63(2):785–790
7. Samsuzzaman M, Islam MT, Singh MJ (2018) A compact printed monopole antenna with wideband circular polarization. IEEE Access 6:54713–54725
8. Han RC, Zhong SS (2016) Broadband circularly polarized chifre-shaped monopole antenna with asymmetric feed. Electron Lett 52(4):256–258
9. Wu JW, Ke JY, Jou CF, Wang CJ (2010) Microstrip-fed broadband circularly polarized monopole antenna. IET Microwaves Antennas Propag 4(4):518–525
10. Le TT, Park HC (2014) Very simple circularly polarised printed patch antenna with enhanced bandwidth. Electron Lett 50(25):1896–1898
11. Nasimuddin BH, Shen Z (2015) Broadband circularly polarized moon-shaped monopole antenna. Microwave Opt Technol Lett 57(5):1135–1139

12. Cai YM, Li K, Yin YZY, Hu W (2014) Broadband circularly polarized Printed antenna with branched microstrip feed. IEEE Antennas Wirel Propag Lett 13:674–677
13. Ullah U, Koziel S (2019) A novel coplanar-strip-based excitation technique for design of broadband circularly polarization antennas with wide 3 dB axial ratio beam width. IEEE Trans Antennas Propag 67(6):4224–4229
14. Omar AA, Shen Z (2019) A compact and wideband vertically polarized monopole antenna. IEEE Trans Antennas Propag 67:626–631

FPGA-Based True Random Number Generator Architecture Using 15-Bit LFSR and ADPLL

Huirem Bharat Meitei and Manoj Kumar

1 Introduction

Random numbers are required in virtually all cryptographic operations. Initialization vectors, block padding, nonces, and keys are all encrypted structures that require a randomly generated sequence of numbers. Since the majority of the data created by the random number generator (RNG) are communicated in the open domains, a passive hacker has enough opportunity to evaluate the RNG's findings and exploit any flaws discovered. Consequently, RNG employed in cryptographic operations must constantly be regarded as a crucial element of the encrypted algorithm. A defect or malfunction in the RNG could cause the entire system to fail [1]. The infamous Netscape V2.0 website compromise is a prime example of an effective assault on a vulnerable RNG architecture [2]. Despite an increasing reliance on data collected from different applications, digital phones, and devices, communication infrastructure security becomes crucial. Consumers' privacy must be appropriately secured by implementing an RNG that is both secure and reliable, such as the TRNG. RNGs are essential parts of every cryptographic algorithm, as used in block ciphers, digital signatures [3] as well as in one-time padding. The field of RNG also addressed the deterministic random bit generators (DRBGs) built on a hash algorithm and the SHA-256 encryption technique [4]. Our proposal has two significant advantages over previous concepts: It is vendor-independent and eliminates the need for human installation and channeling throughout the production process, making the generator more compact. A TRNG's development shouldn't be dependent on any specific technology. Recently, Cherkaoui et al. [5], inspired by Sunar et al., devised a novel design in which ROs were replaced with a Self-Timed Ring (STR). Cret et al. expand on the core concept of utilizing only 2 ROs [6]. In this method, the writer uses

H. B. Meitei (✉) · M. Kumar
Department of ECE, NIT Manipur, Imphal 795004, India
e-mail: thinktank453@gmail.com

275

a multiplexer to split the sampling pulse. All digital phase lock loops (ADPLL)-based TRNGs have a number of advantages over PLL-based TRNGs, including low-power consumption, reduced area requirements, simple synthesizability, as well as the capability to be quickly modified. In contrast, LFSR has the most widespread application in communications and cybersecurity.

LFSRs are typically composed of a D flip-flop and two input XOR gates. It could be done in two steps: using either the Fibonacci algorithm or the D flip-flop algorithm. The LFSR uses a feedback loop to shift the bits of binary data rely on the current state of the register and a predetermined set of feedback rules. This allows for the generation of a sequence of random numbers that are difficult to predict or reproduce Murali Krishna et al. [7] Outcome assessment suggests that envisaged LFSR with and without seed value provides superior results, reduced power intake, as well as increased unpredictability in runtime with Partial Reconfiguration (PR).TRNG based on an LFSR and an ADPLL is a hardware-based random number generator that uses the principles of both LFSR and an ADPLL to generate random numbers. It offers a higher degree of randomness and unpredictability than an LFSR alone, making it useful for applications where true randomness is important, such as cryptography. Therefore, TRNGs employing FPGA digital logic design have additional freedom, performance, and convenience than TRNGs employing analog circuitry [8]. Because of the different sources of unpredictability utilized in the production of stochastic randomized sequence, such as acquired Jitter [9], metastability [10], and transitional impacts [11] from different resources like PLL, ROs, and FFs which all greatly influence the TRNG's speed. The strong level of security provided by the cryptographic system is based on generations of random and unique digital key sequences [12]. This work aims to establish and implement a 15-bit LFSR with ADPLL-based TRNG (15-LAT) architecture on the FPGA platform, as explained in the sections below: Sect. 1: Introduction, Sect. 2: Summary of an ADPLL used in proposed TRNG (15-LAT) Architecture, Sect. 3: Proposed design for implementing TRNG based on ADPLL with LFSR, Sect. 4: FPGA realization of TRNG design centered on ADPLL with 15-bit LFSR, Sect. 5: Experiment results, Sect. 6: Comparison among different TRNGs, Sect. 7: Conclusion.

1.1 Random Number Generators Type

RNGs are typically divided into 2 different categories.

Pseudo-Random Number Generators (PRNGs)
A PRNG is a probabilistic algorithm that generates an unpredictable bitstream of data from a non-repetitive sequence of random bits. It is generated via specific software instructions and is activated by a function that generates a predictable key stream [13]. As a consequence, a continuous, cyclical, and repeated sequence of a random sequence is generated that approximates the properties of earlier generated arbitrary sequences. PRNGs employ a preset technique to generate a series of outcomes based

on the previous entropy seed [14]. It takes an initial number (the "seed"), performs a series of calculations on it, and then outputs a sequence of numbers that seem to be random. The sequence of numbers is dependent on the seed, so if the seed thus generated is known, then the generated numbers can be reproduced.

True Random Number Generators (TRNGs)
Whereas TRNG is a method of producing unexpected randomized numbers that rely on both physical processes and a non-deterministic source. These physical processes range from measuring thermal noise in resistors to observing radioactive decay. Additionally, the absence of similarity between the recently generated bitstream as well as the most recently created data sequence [10] provides TRNG with high robustness over PRNG. This architecture receives its output entirely from an asynchronous physical process occurring under the surface. Because no intermediate data is preserved in the generator, the result is entirely dependent on the physical operation and never upon any predefined information. TRNGs provide a guarantee of unpredictability that PRNGs cannot provide.

2 A Summary of an ADPLL Used in the Proposed TRNG (15-LAT) Architecture

We combine our LFSR using an ADPLL layout comprising ring oscillators and flip-flops to provide the requisite total entropy seed for the stochastic sequence generation of the proposed TRNG. ADPLL is an electronic circuitry technique that permits the functional reproduction of the basic digital block on the Field programming Gate Array board. ADPLL implements phase-locked loops (PLLs) entirely digitally [15]. It is consist of 3 basics element, i.e., Phase detectors (PDs), Loop filters (LFs) as well as Digital Control Oscillators (DCOs). XOR-Gate act as the phase detector (PD) [16]. ADPLL makes use of PD to reduce the amount of difference that exists between the 2 streams.

To remove undesired frequency components, K counter is utilized as a loop filter [18]. ID counters operate in the same way to DCOs such that they modify the frequencies based on the LF input result. Figure 1 depicts the overall circuit diagram of an ADPLL. Mfo, the clock frequency driving the K counter, is the same as the K clock. The ID counter's clock signal is 2Nfo, while M and N represent the K counter and ID counter's modulus controls, correspondingly. M is often set to 8, 16, 32, …, with $M = 2N$ used to fix the N values. The XOR gate's XOR out signal is fed into the K counter, which generates a carry signal (ca) [15].
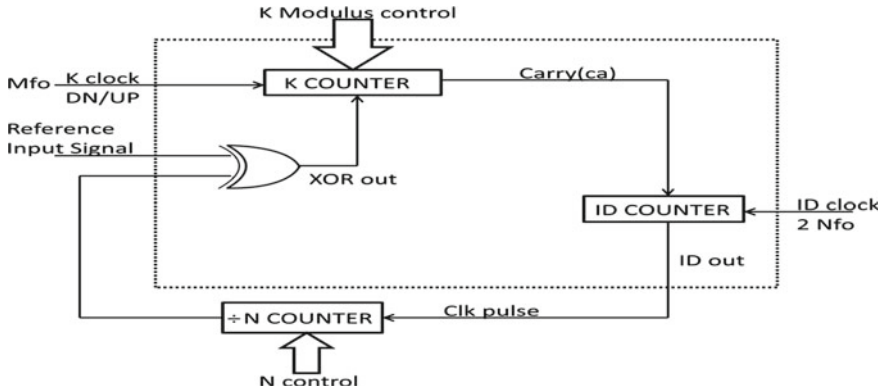
Fig. 1 Circuit diagram of an ADPLL used in ACT architecture [17]

Fig. 2 Block diagrammatic form of *K* counter used as loop filters in our proposed (15-LAT) architecture [16]
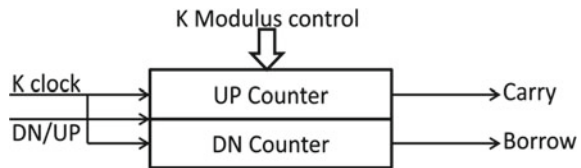
2.1 K Counter as Loop Filter (LF) Used for ADPLL

A *K* counter, as shown in Fig. 2, is a type of LF or integrator that operates in tandem with an EXOR or JK phase detector. It is made up of 2 distinct counters, an up counter as well as a down counter, that both count upward. The *K*-modulus counter has a *k* range of 0 to $k - 1$, and its value equals *M* times the center frequency. For the down counter to be enabled, the DN/UP condition must always be high, while for the up counter to be enabled, the DN/UP condition must always be low [19].

2.2 Digital Control Oscillator (DCO) Used in ADPLL

DCOs are a form of adaptive oscillator that use the output of the loop filter to alter the input signal's frequency [19]. Figure 3 depicts the DCO network configuration of the ADPLL. The final output of DCO is known as id out [15].
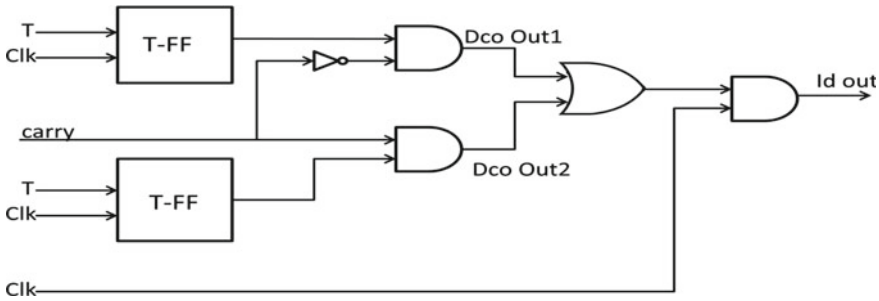
Fig. 3 DCO circuit diagram [17]

3 Proposed Design for Implementing TRNG Architecture Based on ADPLL with 15-LFSR

A LFSR is a form of a shift register that generates a sequence of binary numbers via feedback. The sequence is decided by the feedback function, which is a Boolean function that accepts the current state of the shift register as an input and returns a new bit that is shifted into the register. It is possible to develop an LFSR capable of generating a stream of random numbers by carefully configuring the feedback function. Theoretically, an LFSR is a series of connected FFs, with each flip-input flop's being the output of the previous one [20]. It is produced by combining XOR gates within the feedback of a series of flip-flops. The initial number of the LFSR, also known as the seed number, consists of both 1s and 0s. Even if the seed number influences following random variables, it is essential to choose a number with minimal energy usage. Figures 4a and b exhibit the architectures of LFSR1 and LFSR2, respectively. Exclusive-OR (XOR) with a single bit is used as a linear function. LFSR is composed of XOR gates with DFF. Typical LFSR polynomials are characterized by XOR positions. $P(x) = x5 + x4 + x + 1$. An LFSR with a properly selected feedback mechanism can generate a seemingly unpredictable bitstream. The starting value of the LFSR is referred to as the seed; the sequence of integers generated by the shift register depends on its prior or present configuration. Due to the register's limited number of stages, it subsequently reaches a cycle. The maximum size of an LFSR series is $2n$ minus one, producing a randomized periodic pattern [21]. The feedback function of an LFSR is often represented by a simple polynomial. The maximum feasible LFSR generates the greatest number of PRPG configurations with a design count of $2n - 1$, where n is the number of register components in the LFSR.

Figure 4 illustrates the fundamental structural design of the 15-bit-LFSR utilized in our suggested architecture. Additionally, TRNG shouldn't be overly dependent on technology. Nevertheless, because the PLL is not available across all FPGA types, it is challenging for engineers. In addition, PLL-based TRNG needs additional energy and occupies greater storage [21] over ADPLL-based TRNG. Due to its ADPLL-based digitized structure, our Noval 15-LAT design may be created as well as improved in a short time.
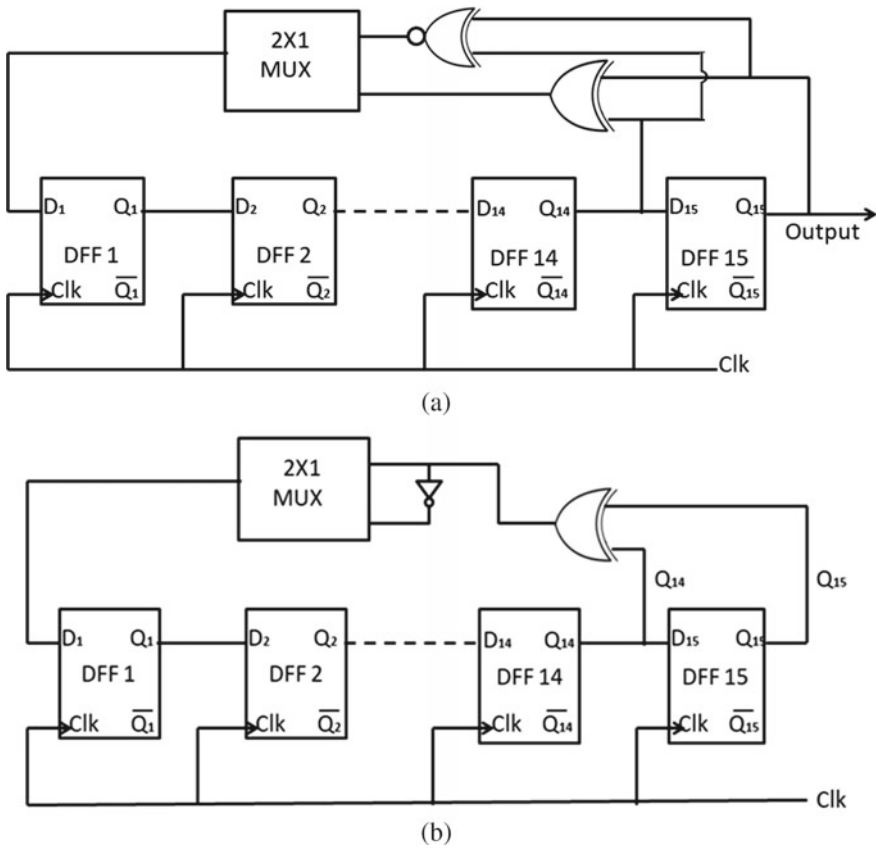
Fig. 4 a 15-bit linear feedback sift register 1 (15-Bit LFSR1) [7]. **b** 15-bit linear feedback sift register 2 (15-Bit LFSR2) [7]

ADPLLs configured separately using dual 15-bit LFSRs connected in series using the ring oscillator comprise our design proposal. In our RO architecture, we implemented a pulse generator consisting of 51 inverters in order to generate the clock signal for the suggested 15-LAT structure. Using VHDL, 15-bit LFSR and ADPLL-based TRNG implementations are constructed. In additional to the flip-flop and metastability criterion, explore all important types of entropy, such as disruption by ADPLLs [22] and proposed rings oscillators [17]. Figure 5 depicts the suggested TRNG design-based ADPLL with 15-bit LFSR. The system clock is set to 100 MHz and is supplied to the divide by two counters on the pulse generator circuit.

Here, 100 MHz is used as the operating system CLK and is sent to the pulse generator circuit's divide by two counter. Now, a 50 MHz pulse is received by a circuit in which a pulsing signal oscillates amongst 2 voltage states signifying true and false. XORing the jitter output produced by two cascading ring-oscillators and a 15-bit LFSR using the 400 MHz ID output pulse (DCO output) of the ADPLL as
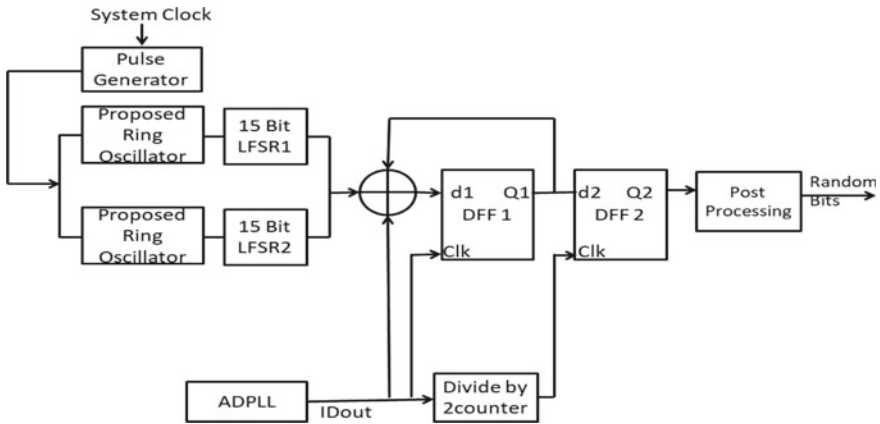
Fig. 5 Proposed TRNG architecture with 15-bit LFSR based on ADPLL (15-LAT)

well as the Q1 of DFF1 feedback loop. The outcome of Q1 of DFF1 is therefore supplied to d2 of DFF2 together with the counter's produced CLK pulse. By using an ADPLL in conjunction with an LFSR, it is possible to create a TRNG that has a higher degree of randomness and unpredictability than an LFSR alone. Q2 of DFF2 yields a random bitstream that is transmitted for post-processing to verify that the sampler's outputs random numbers are unbiased.

4 FPGA Realization of TRNG Design Based on ADPLL with 15-Bit LFSR

The inquiry utilizes an Artrix-7 FPGA system (XC7A35T CPG236-1) and an oscilloscope (DSO-X3012A) to record the pattern. Table 1 lists the FPGA pinouts for the TRNG solution based on ADPLL with CS phenomenon. Figure 6 shows the experimental setup for an ACT-TRNG. The total system clock is generated utilizing the W5 input mode as well as the V17 T-FF inputs. The outcomes are connected to JB1:A14, which serves as the DSO's live probe, and JB5:GND, which serves as the ground probe.

Table 1 Pins information for the integrated proposed TRNG architecture centered on ADPLL with 15-bit LFSR

Symbol	Details	Mode	Pinned mode
CLK	Operating-clock	Input	W5
t	T-FF	Input	V17
rst	Reset	Input	V16
q3	Resulted random data	Output	A14
Vcc	Power supply		USB port
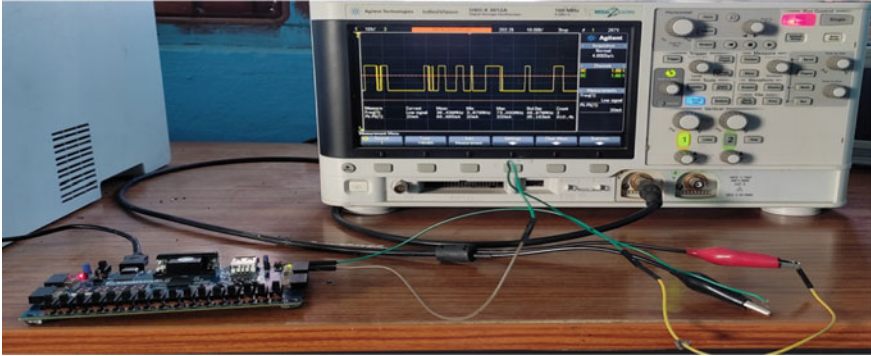
Fig. 6 Experimental configuration of the FPGA-DSO interface for the 15-LAT architecture

5   Experiment Results

The suggested TRNG centered on ADPLL with 15-bit LFSR is schematically depicted in Fig. 7. All schematic designs were generated in Vivado v.2015.2, and simulations were run on an Artrix-7 FPGA board with the xc7a35tcpg236-1 module. Figure 8 and Fig. 9 illustrates the TRNG output waveforms along with the FFT waveform respectively collected by DSO of the proposed 15-LAT architecture (Table 2).
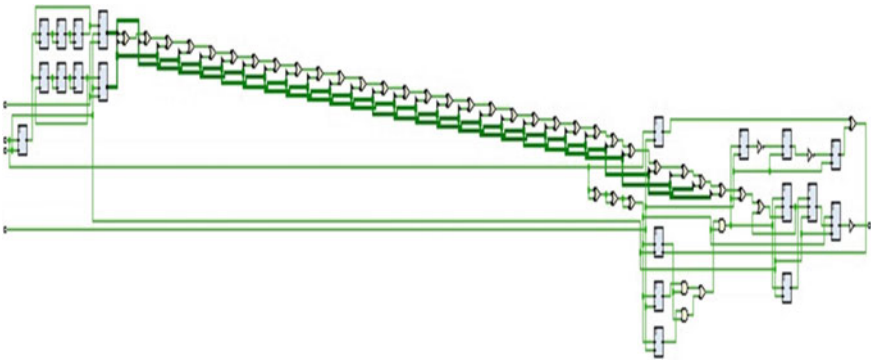
Fig. 7 Schematic diagram for the proposed 15-LAT architecture

Fig. 8 The envisioned TRNG's output waveform is constructed upon ADPLL and LFSR (15-LFSR)

Fig. 9 FFT output waveform of TRNG build on ADPLL along with LFSR (15-LFSR) architecture

Table 2 NIST (SP 800-22) test result

NIST-test	p-value	Result
Frequencies	0.9990	Passed
Block-frequencies	0.9999	Passed
Run	0.9990	Passed
Rank	0.0000*	Failed
DFT	0.o218	Passed
Serial test	0.9936	Passed
Linear-complexities test	0.0000*	Failed
Longest-run test	0.0500	Passed
Approximate-entropy test	0.0000*	Failed
Cumulative-sum test	0.9990	Passed

[*] NIST failed whenever the p value is 0.000

6   Comparison Between Existing TRNG with Proposed TRNG Based on ADPLL with 15-Bit LFSR

Tables 3 and 4 compares the new 15-LAT architecture to the previous TRNGs architectures. The performance of numerous TRNGs is compared in the Table 3. Our presented TRNG architecture-based ADPLL with 15-bit LFSR made better use of existing hardware capacity even though utilizing minimum power.

7   Conclusion

Our method makes a substantial contribution toward the growing utilization of FPGAs in encryption methods. The system's overall security is strengthened by having the capacity to completely enclose a TRNG architecture with the FPGA. In this paper, we introduced a unique entropy source for the 15-LAT architecture that significantly reduces engineering complexity. We empirically confirmed the architecture's viability and showed that achieving the stochastic model's entropy needs is always possible, especially whenever placing constraints are removed or the architecture is ported to a different generation of FPGAs. This characteristic provides the TRNG a good candidate for inclusion into wider crypto methodologies. Randomized bits are generated and are statistically valid using NIST 800-22 test. By using 15-bit LFSR along with ADPLL in the TRNG design, we can reduce hardware resources while increasing the efficacy of the FPGA chips, as shown in Table 3. With this study, the potential of cybersecurity via ADPLL-based TRNG with 15-LFSR appears to be optimistic, giving it a more dependable and secure solution for a variety of applications.

Table 3 A comparative study of different TRNGs architectural efficiency

Paper	Entropy	Device	Hardware-resource	Post-processing
15-LAT	Jitter and metastability	Artrix-7	0 LUT, 4 FFs	Yes
[23]	Latches-oscillatory and metastability	Artrix-7	40 LUTs	No
[24]	Self-time ring-oscillator	Artrix-6	56 LUTs	No
[25]	Jitters and metastability	Artrix-7	50 LUTs	No
[10]	Jitters or metastability	XCKU040	1 PLL, 5 Primitives, 5 Slices	Yes
[14]	Jitters or metastability	XC7A35T CPG236-1	1 LUT	Yes
[26]	Chaotic ring-oscillator	XC6SLX16	44 LUTs	Yes
[27]	Metastability	Virtex5	0 LUT	No
[28]	Ring-oscillator	Spartan-6	10 LUTs, 5 FFs	No
[17]	Jitter or metastability	XC7A35T	2 LUTs, 1 Slice	No

Table 4 Synthesis results in comparison among different TRNG's

Parameter	Paper [17]	Paper [29]	Paper [30]	Paper [31]	15-LAT
Platform	FPGA	FPGA	FPGA	0.35 μm CMOS	FPGA
Power consumption (W)	0.076	0.125	0.095	0.026	0.072
Post processing	No	No	No	Yes	Yes
Testing	NIST-SP800-22	NIST-SP800-22	NIST-SP800-22	NIST-SP800-22	NIST-SP800-22

References

1. Eastlake D, Crocker S, Schiller J (1994) Randomness recommendations for security—RFC 1750. Available at http://www.faqs.org
2. Goldberg I, Wagner D (1996) Randomness and the Netscape Browser. Dr. Dobb's Journal, January 1996
3. Buchmann J, Dahmen E, Szydlo M (2009) Hash-based digital signature schemes. Post-Quantum Cryptogr, 35–93. https://doi.org/10.1007/978-3-540-88702-7_3
4. Loza S, Matuszewski L (2014) A true random number generator using ring oscillators and SHA-256 as post-processing. In: 2014 international conference on signals and electronic systems (ICSES). https://doi.org/10.1109/icses.2014.6948739
5. Cherkaoui A, Fischer V, Fesquet L, Aubert A (2013) A very high speed true random number generator with entropy assessment. In: Proceedings of international workshop on cryptographic hardware and embedded systems (CHES'13). Lecture notes in computer science, vol 8086. Springer, pp 179–196
6. Cret O, Suciu A, Gyrfi T (2008) Practical issues in implementing TRNGs in FPGAs based on the ring oscillator sampling method. In: 10th international symposium on symbolic and numeric algorithms for scientific computing (SYNASC'08). IEEE Computer Society, pp 433–438
7. Murali Krishna B, Madhumati GL, Khan H (2019) FPGA based pseudo random sequence generator using XOR/XNOR for Communication cryptography and VLSI testing applications. Int J Innov Technol Exploring Eng 8(4):485–494
8. Gupta R, Pandey A, Baghel RK (2018). Efficient design of chaos based 4 bit true random number generator on FPGA. Int J Eng Technol 7(3):1783. https://doi.org/10.14419/ijet.v7i3.16586
9. FIPS 140-1 (n.d.) Security requirements for cryptographic modules. Csrc.nist.rip. Retrieved 13 April 2021, from https://csrc.nist.rip/publications/fips/fips140-1/fips1401.htm
10. Stanchieri G, De Marcellis A, Palange E, Faccio M (2019) A true random number generator architecture based on a reduced number of FPGA primitives. AEU Int J Electron Commun 105. https://doi.org/10.1016/j.aeue.2019.03.006
11. Kohlbrenner P, Gaj K (2004) An embedded true random number generator for FPGAs. In: Proceeding of the 2004 ACM/SIGDA 12th international symposium on field programmable gate arrays—FPGA'04. https://doi.org/10.1145/968280.968292
12. Vasyltsov I, Hambardzumyan E, Kim Y-S, Karpinskyy B (n.d.) Fast digital TRNG based on metastable ring oscillator. In: Cryptographic hardware and embedded systems—CHES 2008, pp 164–180. https://doi.org/10.1007/978-3-540-85053-3_11
13. Internet resource: AMD random number generator library. https://developer.amd.com/wordpress/media/2013/12/AMD-Number-Generator-User-Guide.pdf. Accessed Sept 2018
14. Meitei H, Kumar M (2022) FPGA implantations of TRNG architecture using ADPLL based on FIR filter as a loop filter. SN Appl Sci 4:96. https://doi.org/10.1007/s42452-022-04981-6
15. Chaudhary AK, Kumar M (2017) Design and implementation of ADPLL for digital communication applications. IEEE Xplore. https://doi.org/10.1109/I2CT.2017.8226159
16. Kumar M, Lata K (2012) All digital phase locked loop (ADPLL): a survey. In: Proceeding the 4th IEEE international conference on electronics computer technology (ICECT 2012), 6–8 April 2012, Kanyakumari, India

17. Meitei HB, Kumar M (2021) FPGA implementation of true random number generator architecture using all digital phaselocked loop. IETE J Res. https://doi.org/10.1080/03772063.2021.1963333
18. Lata K, Kumar M (2013) ADPLL design and implementation on FPGA. IEEE Xplore. https://doi.org/10.1109/ISSP.2013.6526917
19. Lata K, Kumar M (2013) All digital phase-locked loop (ADPLL): a Survey. Int J Future Comput Commun, 551–554. https://doi.org/10.7763/ijfcc.2013.v2.225
20. Hollmann H (1990) Design of test sequences for VLSI self-testing using LFSR. IEEE Trans Inf Theory 36(2):386–392
21. Arnault F, Berger T, Minier M, Pousse B (2011) Revisiting LFSRs for cryptographic applications. IEEE Trans Inf Theory 57(12):8095–8113
22. Radhapuram S, Yoshihara T, Matsuoka T (2019) Design and emulation of all-digital phase-locked loop on FPGA. Electronics (Basel) 8:1307
23. Fujieda N (2020) On the feasibility of TERO-based true random number generator on Xilinx FPGAs. In: Proceedings of the 2020 30th international conference on field-programmable logic and applications (FPL), Gothenburg, Sweden, 31 August–4 September 2020, pp 103–108
24. Wang X, Liang H, Wang Y, Yao L, Guo Y, Yi M, Huang Z, Qi H, Lu Y (2021) High-throughput portable true random number generator based on jitter-latch structure. IEEE Trans Circ Syst I Regul, Pap. 2021 68:741–750
25. Lin J, Wang Y, Zhao Z, Hui C, Song Z (2020) A new method of true random number generation based on Galois ring oscillator with event sampling architecture in FPGA. In: Proceedings of the 2020 IEEE international instrumentation and measurement technology conference (I2MTC), Dubrovnik, Croatia, 25–29 May 2020, pp 1–6
26. Yang Y et al (2017) A reliable true random number generator based on novel chaotic ring oscillator. In: 2017 IEEE international symposium on circuits and systems (ISCAS). www.semanticscholar.org/paper/A-reliable-true-random-number-generator based-on-Yang-Jia/c3e65e27fd09968934977d250f0ead2c13e60b35. https://doi.org/10.1109/ISCAS.2017.8050843
27. Ben-Romdhane M, Graba T, Danger J-L (2013) Stochastic model of a metastability-based true random number generator. Trust Trust Comput. https://doi.org/10.1007/978-3-642-38908-5_7
28. Yang B, Rožic V, Grujic M, Mentens N, Verbauwhede I (2018) ESTRNG: a high-throughput, low-area true random number generator based on edge sampling
29. Cicek I, Pusane AE, Dundar G (2014) A new dual entropy core true random number generator. Analog Integr Circ Sig Process 81(1):61–70
30. Petrie CS, Connelly JA (2000) A noise-based IC random number generator for applications in cryptography. IEEE Trans Circ Syst I: Fund Theory Appl 47(5):615–621. https://doi.org/10.1109/81.847868
31. Ergun S, Ozoguz S (2010) Truly random number generators based on non-autonomous continuous-time chaos. Int J Circuit Theory Appl 38(1):1–24

IoT Based LPG, Smoke, and Alcohol Detection System with Automatic Power Cut-off

Dwarakanath Dey, Saikat Datta, Subhojit Datta, Souptik Das, and Tanusree Dutta

1 Introduction

Nowadays, gas leakage [1, 2] has become a serious concern, affecting human lives and properties. Liquefied petroleum gas (LPG), a flammable mixture of hydrocarbon gas is used as a fuel in a variety of applications such as in homes, industries, and automobiles due to its beneficial features. LPG is extremely combustible and can even burn far away from the site of the leakage. The majority of this energy source is composed of propane and butane, which are extremely flammable chemical compounds. These gases are easily combustible. In our country, approximately 30 crore people use LPG, constituting almost 40% of the population. As a result, there is a true requirement for the gas leakage detection system [3, 4] to be developed. Fires in homes and industries are becoming more common, causing greater risk to human lives and property in recent years. Because of physical characteristics such as toxicity, flammability, and asphyxia, they pose a risk of explosion, fire, and suffocation. The fatality rate from gas cylinder explosions has been rising tremendously. Substandard cylinders, old valves, no frequent checking of gas cylinders, wiped out regulators, and a lack of information on how to handle gas cylinders [5] are the main causes of such explosions. As a result, gas leakage should be controlled to protect people from severe accidents [6]. In this paper, we have proposed a low-cost advanced sensor-based [7] gas leakage detector, warning, and control system. The system is incredibly efficient, simple to use, and cost-effective. It is an IoT-based [8] LPG, smoke, and Alcohol detection system [9] which can cut the power supply automatically to ensure better safety and security. The entire simulation was carried out using Proteus [10] software. When gas leakage occurs, the concentration in the air reaches the required level, and

D. Dey (✉) · S. Datta · S. Datta · S. Das · T. Dutta
Department of Electronics & Communication Engineering, St. Thomas' College of
Engineering & Technology, Kolkata, India
e-mail: dwarakanathdey@gmail.com

a smart alert message is sent via an LCD screen, a buzzer, and a Wi-Fi module, which is capable of delivering alert notifications to the customers' mobile phones. Our system will also cut-off the power as soon as the gas leakage is detected. The gas leakage detection system can be used in a variety of places, including residential homes, hotels, and industries, where gas leakage is a serious concern.

2 Related Works

2.1 Literature Survey

A number of reviews on research papers or technical reports on the subject of Gas Detection system which were done in the past are presented below.

The paper proposed a design for a gas leakage detection system that could automatically detect LPG leakage. A MQ6 semiconductor sensor was used, which could detect gas concentrations anywhere from 200 to 10,000 ppm. When the system senses gas in the environment, it outputs a digital value of 1; if no gas is detected, the sensor outputs a value of 0. The sensor output is used as a digital input by Arduino. The buzzer will start to tune if the sensor output is high, and the LCD will display "Gas detected: Yes." The buzzer will not tune if the sensor output is low, and the LCD will display "Gas detected: No" [7].

The paper aims to design a PCB model of a gas monitoring system. The model uses an MQ-5 gas sensor to detect gas leakage and an Arduino Uno as a central processing unit. When the gas concentration exceeds a particular limit, the Arduino sends a signal to turn on the buzzer and also the exhaust fan to extract the gas. A Wi-Fi module was used to deliver messages to the dashboard, which displays the location of the gas leakage. A step-down transformer, bridge rectifier, and capacitive filter are used to convert line voltage (AC) to the DC voltage required to drive the model. The system is ideal for industrial applications but might be a bit expensive for household kitchen applications [8].

The paper proposes a design for a gas leakage detection system that can detect and alert users to LPG leakage. The PIC16F877A microcontroller and MQ6 gas sensor were used in the system. The leak is detected by the system sensor, which sends an SMS to the housemates and activates the alarm and exhaust fan. In addition, the system continuously monitors the cylinder's LPG level using a load sensor and automatically books the cylinder [5].

The objective of this paper was to detect and alert the leakage of LPG. The system was designed using a PIC18F1320 microcontroller, a MQ-5 gas sensor, buzzers, LEDs, and accessory circuitry for sensor sensitivity adjustment. When the sensor voltage lies below 1.2 V, then no gas leakage occurs. If the sensor voltage ranges between 1.2 V and 4 V, then a low-level warning signal gets generated, which activates the green and yellow LEDs along with a buzzer. If the voltage increases above 4 V, then a high-level dangerous warning signal gets activated, which turns on the red LED

along with both the buzzers. The above process continues until the sensor voltage drops within the specified limits [4].

3 Design and Implementation

Gas leakage-related accidents have become a big threat in our daily lives. As the world becomes more technologically advanced, it is vital to use technology whenever possible [11]. So, we have designed an IoT-based gas leakage detecting system [12] that is both efficient and cost-effective.

The entire project simulation was carried out in Proteus software along with Arduino IDE.

3.1 Block Diagram

In this paper, we have used Arduino Uno as the main control unit, gas sensors (MQ-2 and MQ-3) to detect the concentration of flammable gases in air, buzzer as an smart alerting unit, LCD to display the present state of gas leakage, LEDs to indicate gas leakage, relay unit for power cut-off, COMPIM as a Wi-Fi module to establish a serial communication between the Arduino and the computer's virtual USB port and the Blynk [13] app to receive the notifications into the phones as an when gas leakage occurs (Fig. 1).

3.2 Component List

The sensitive material of MQ-2 and MQ-3 gas sensors is tin dioxide (SnO_2), which has lower conductivity in clean air. With an increase in combustible gas concentration, the conductivity of the sensor also increases, helping in the early detection of gas leakage (Tables 1, 2 and 3; Figs. 2, 3, 4 and 5).

4 Methodology

STEP 1: The gas sensor MQ-2 and MQ-3 periodically sends a voltage signal to the Arduino Uno board; the signal is maintained 0 V (LOW) in case of no detection. The signal becomes 5 V (HIGH) whenever it detects smoke or LPG or Alcohol vapors.

STEP 2: If HIGH voltage signal is read by the Arduino board, it activates the LCD display, buzzer, LED, relay circuit, and Blynk notification request.
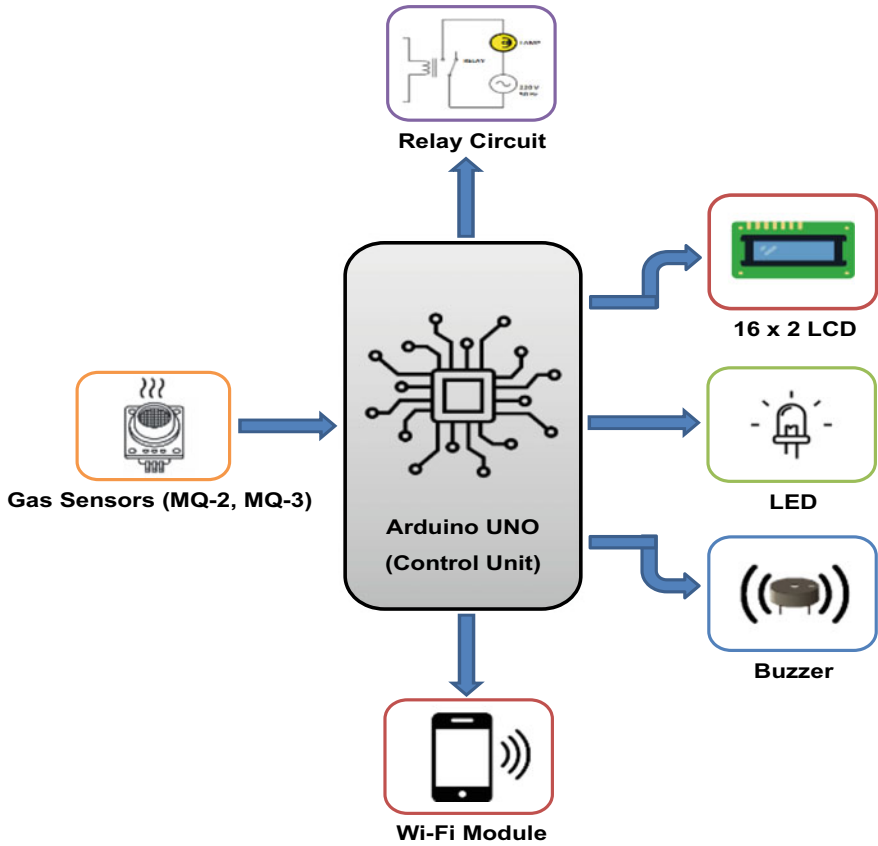
Fig. 1 Block diagram of the proposed system

Table 1 Components used

S. No.	Components	Specifications	Quantities
1	MQ-2 sensor	LPG gas sensor	1
2	MQ-3 sensor	Alcohol and other hazardous gas sensor	1
3	Arduino board	Arduino Uno Rev 3, Clock Freq. 16 MHz	1
4	SPDT relay	Component value: 5 V, Coil resistance: 240 Ω	1
5	Buzzer	–	1
6	LCD (16 × 2)	LM016L	1
7	LEDs	Fwd. volt: 2.2 V (at 10 mA), Full drive current: 20 mA	1
8	LAMP	Nominal voltage: 12 V	1
9	COMPIM	Baud rate: 9600 bps	1
10	Resistance	10 kΩ	1

Table 2 Technical specifications of MQ-2

Model No.	MQ-2
Sensor type	Semiconductor gas sensor
Sensing material	Tin dioxide (SnO_2)
Detection gas	LPG, propane, butane, hydrogen
Concentration	300–10000 ppm (combustible gas)
Loop voltage (V_c)	$\leq$24 V DC
Heater voltage (V_H)	5.0 V $\pm$ 0.2 V AC or DC
Load resistance (R_L)	Adjustable
Sensing resistance (R_S)	2–20 kΩ (in 2000 ppm C_3H_8)
Sensitivity (S)	R_S (in air)/R_S (1000 ppm isobutane) $\geq$ 5
Standard test circuit	V_c: 5.0 V $\pm$ 0.1 V; V_H: 5.0 V $\pm$ 0.1 V

Table 3 Technical specifications of MQ-3

Model No.	MQ-3
Sensor type	Semiconductor gas sensor
Sensing material	Tin dioxide (SnO_2)
Detection gas	Alcohol, smoke
Concentration	0.04–4 mg/l alcohol
Loop voltage (V_c)	$\leq$24 V DC
Heater voltage (V_H)	5.0 V $\pm$ 0.2 V AC or DC
Load resistance (R_L)	Adjustable
Sensing resistance (R_S)	2–20 kΩ (in 0.4 mg/l alcohol)
Sensitivity (S)	R_s (in air)/R_s (0.4 mg/L alcohol) $\geq$ 5
Standard test circuit	V_c: 5.0 V $\pm$ 0.1 V; V_H: 5.0 V $\pm$ 0.1 V

Fig. 2 Sensitivity characteristics of MQ-2 sensor
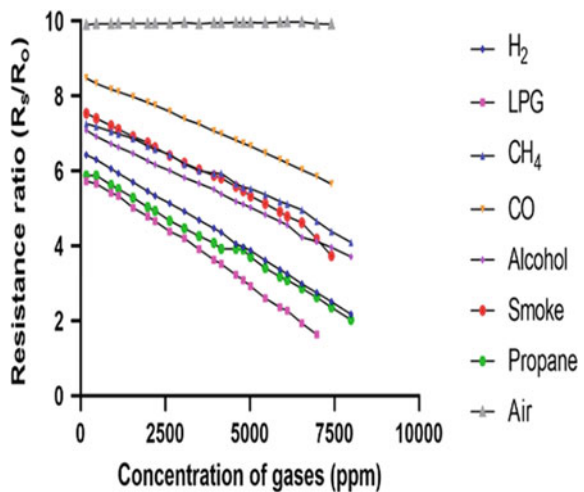
Fig. 3 Temperature and
humidity characteristics of
MQ-2 sensor
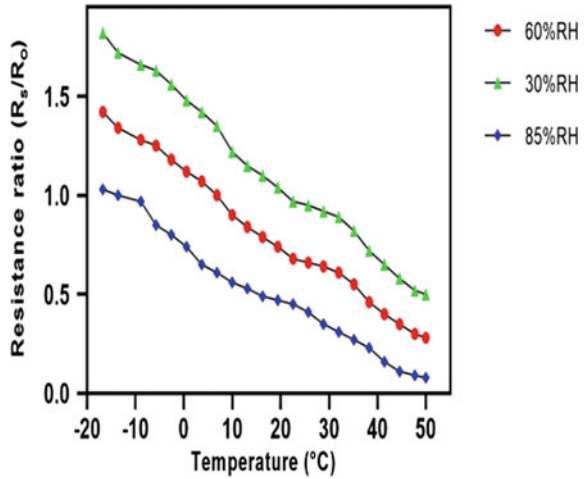
Fig. 4 Sensitivity
characteristics of MQ-3
sensor
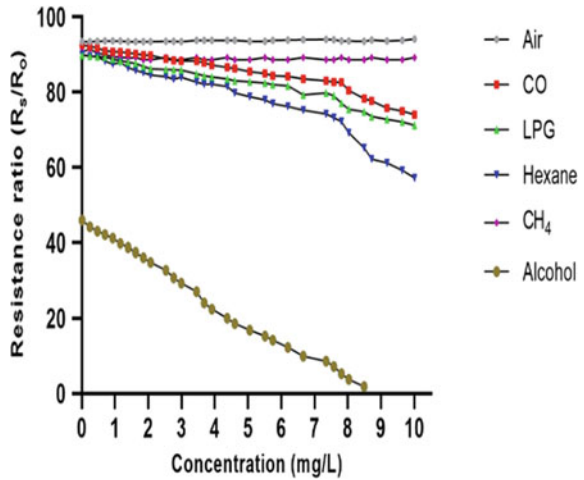
(a) The LCD displays "LPG Detected" in case of smoke or LPG detection.
(b) The LCD displays "Alcohol Detected" in case of Alcohol vapor detection.
(c) The LCD displays "Fire Hazard" in case both LPG and Alcohol vapor is
    detected.
(d) The LCD displays "NO Smoke" in case of no detection.

STEP 3: The relay circuit when triggered by the Arduino disconnects the main
power supply, the connection is maintained in case of no detection.
STEP 4: The buzzer when triggered by the Arduino outputs a high-pitched audio
signal to alert.

STEP 5: The Wi-Fi module replicated using the COMPIM tool sends notification alerts to the Blynk server which is redirected to the smartphones of the respective users to alert them further.

5 Result and Discussions

Initially, the gas sensors send signals to the Arduino. If the output of the gas sensor is HIGH (i.e., 5 V), then the Arduino will activate the LCD display, buzzer, LED, and relay circuit followed by the Blynk notification request to indicate the present state of gas leakage. When the gas sensor output is LOW (i.e., ~0 V), then the Arduino will only notify that there is no leakage of flammable gases in the atmosphere (Fig. 6 and Table 4).

Therefore, there can be four scenarios:

1. When there is no leakage, both sensors are OFF (Fig. 7)
2. When there is LPG leakage, MQ-2 sensor turns ON A0 pin, but MQ-3 is OFF, so A1 is low (Fig. 8a, b)
3. When there is Alcohol leakage, MQ-3 sensor turns ON A1 pin, but MQ-2 is OFF, so A0 is low (Fig. 9a, b)
4. When there is both LPG and Alcohol leakage, then both the sensors are ON (Fig. 10a, b).

The Arduino is programmed in such a way that in case '1' it only sends a signal to the LCD display to show "No Smoke." In case '2' or '3' or '4', it turns ON pin no 6, 7, and 8. As a consequence, the relay secondary circuit is open so power supply is OFF, LED and buzzer are turned ON to alert people inside the room. Also, the LCD shows a message like "LPG Detected" or "Alcohol Detected" or "Fire Hazard," respectively.
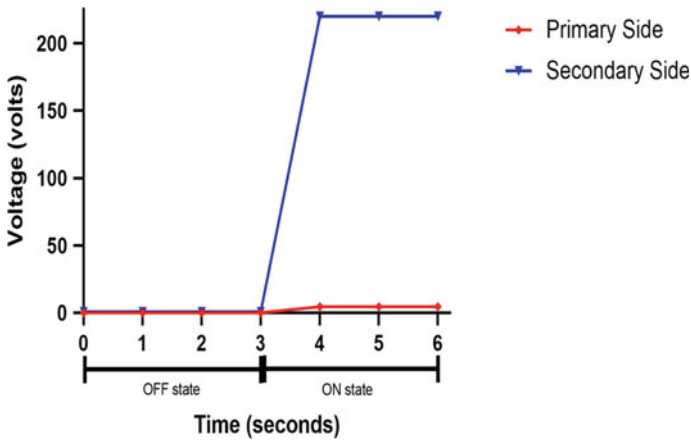
Fig. 6 Relay voltage characteristics (for $t = 3$ s transition instance)

Table 4 Voltage and power across different components

Component	Voltage (ON state) (V)	Voltage (OFF state) (V)	Power (ON state)	Power (OFF state) (W)
Gas sensor (MQ-2)	5	0	0.25 μW	0
Gas sensor (MQ-3)	5	0	0.25 μW	0
Relay (primary side)	4.61	0	87 mW	0
Relay (secondary side)	220	0.9	0 W	8.3

Moreover, Arduino does serial communication through COMPIM which sends serial information from Arduino to COM1 port of the computer. We have used a software called VSPE [14] which pairs COM1 and COM2 port of our computer. We run the Blynk server through command prompt by COM2 port. Hence, a connection is established between our circuit in Proteus and the Blynk server. By this path a notification is sent to the verified owner of the house.

The main advantages of our proposed system are

- IoT enabled
- User friendly and cost-effective
- Automatic power cut-off to prevent the spread of electrical fires
- Prevents wastage of gases
- Lower power consumption.
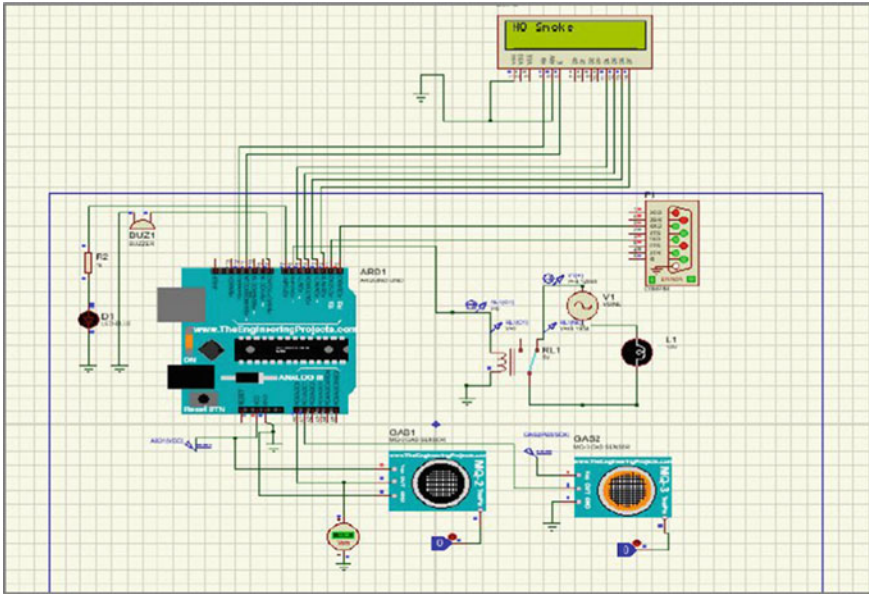
Fig. 7 Output when no smoke is detected (no notification)
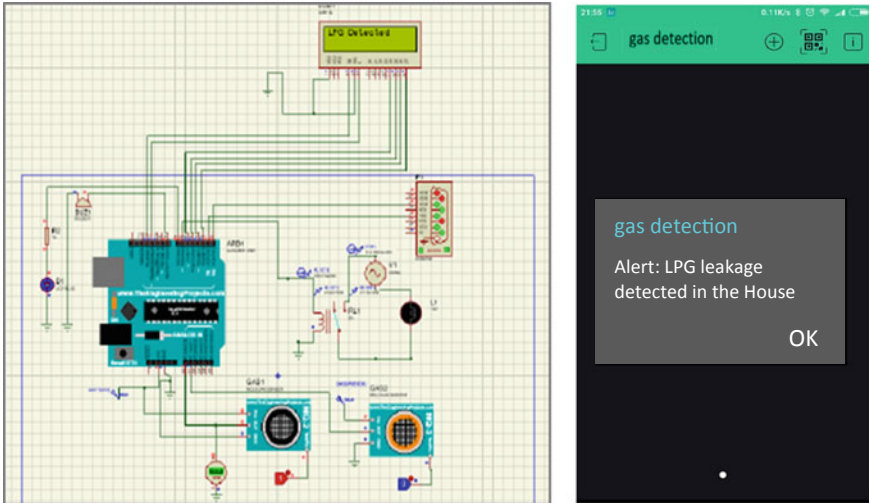
(a)	**(b)**

Fig. 8 **(a)** When MQ-2 detects LPG (LCD shows **LPG Detected**, buzzer, and LED turns ON) and **(b)** notification at users mobile
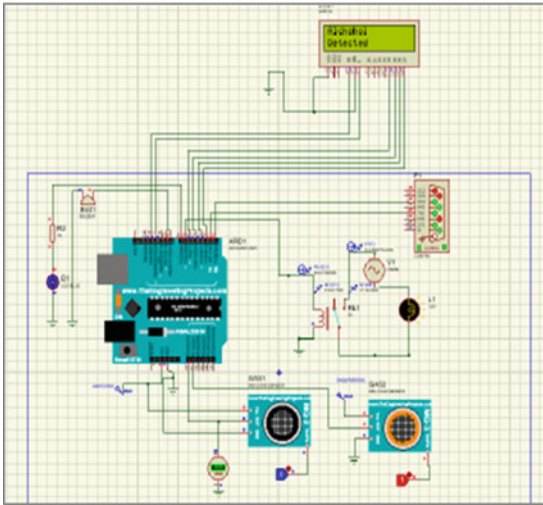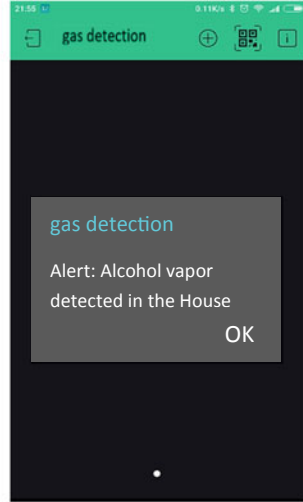
(a)   (b)

Fig. 9 **(a)** When MQ-3 detects Alcohol (LCD shows **Alcohol Detected**, buzzer, and LED turns ON) and **(b)** notification at users mobile
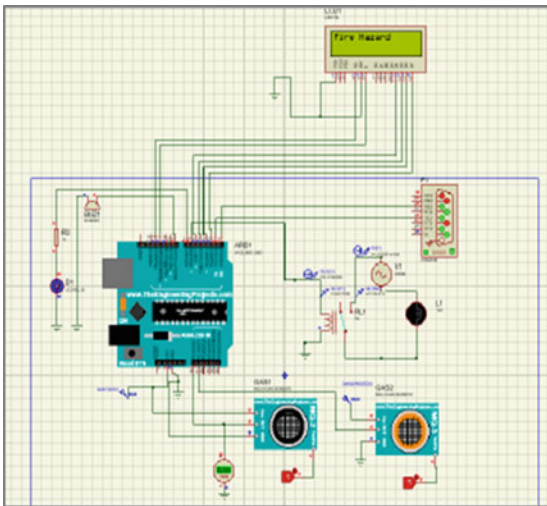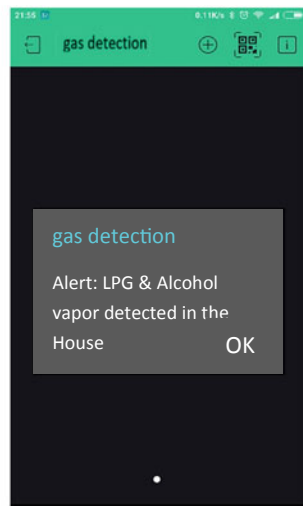
(a)   (b)

Fig. 10 **(a)** When MQ-2 and MQ-3 both detect LPG and Alcohol (LCD shows **Fire Hazard**, buzzer, and LED turns ON) and **(b)** notification at users mobile

6 Conclusion

The design and implementation of the proposed IoT-based Gas Detection and Alert System was carried out in Proteus, a set of proprietary software tools mostly used for electronic design automation. Hence, the circuit model is software in nature, but the output results were both virtual and real. Devoid of any hardware, the implemented model shows high efficiency and effectiveness, and the applications of buzzer and notification alerts can be exhibited in the real world. The main application of the model includes the automatic power cut-off system, which is extremely beneficial to prevent further electrical fires. The design is user friendly and economical, and its applications are numerous from safely monitoring gas leakage in households to effective industrial applications for sensing and alerting the presence of toxic gases and preventing excessive wastage of valuable resources.

7 Future Prospects

The future advancement of the proposed design, being software in nature, produces a lot of prospects, the main being its real-time implementation. The proposed circuit can easily be put together into a simple and cost-effective design while maintaining its efficiency. Further sensors can be added to increase the range of detection of the toxic and hazardous gases and make it more effective. Also, applications such as alerting the responsible authorities in case of household units will make it one step closer to a wide range of commercial uses. In the modern world, this will serve as an important tool of surveillance and prevent further damage to natural and human resources.

References

1. Imade S, Rajmanes P, Gavali A, Nayakwadi VN (2018) Gas leakage detection and smart alerting system using IoT. IJIRS 8(2)
2. Fraiwan L, Lweesy K, Bani-Salma A, Mani N (2011) A wireless home safety gas leakage detection system. In: Proceedings of 1st Middle East conference on biomedical engineering. IEEE, Sharjah, UAE, pp 11–14
3. Nkem NF (2021) Gas leakage detector system with SMS alert. Global Sci J 9(10)
4. Mahalingam A, Naayagi RT, Mastorakis NE (2012) Design and implementation of an economic gas leakage detector. In: Proceedings of recent researches in applications of electrical and computer engineering, pp 20–24
5. Raj A, Viswanathan A, Athul T (2015) LPG gas monitoring system. IJITR 3(2):1957–1960
6. Manichandana K, UmmeRuman S, Biderkota H, Anisha P, Murthy BVR, Kumar CK (2018) Gas leakage detection and smart alerting system. GJESR (ICITAIC-2018)
7. Khan MM (2020) Sensor-based gas leakage detector system. In: Proceedings of 7th international electronic conference on sensors and applications

8. Alshammari BF, Chughtai MT (2020) IoT gas leakage detector and warning generator. Eng Technol Appl Sci Res 10(4)
9. Siddika A, Hossain I (2020) LPG gas leakage monitoring and alert system using arduino. IJSR 9(1)
10. Proteus Homepage. https://www.labcenter.com/simulation
11. CircuitsToday. https://www.circuitstoday.com/lpg-sensor-using-arduino
12. Pandey RC, Verma M, Sahu LK (2017) Internet of Things (IoT) based gas leakage monitoring and alerting system with MQ-2 sensor. IJEDR 5(2)
13. Blynk Homepage. https://blynk.io
14. VSPE Download. https://lab4sys.com/en/download-vspe-virtual-serial-ports-emulator-software

A Dual-Band Dual-Polarized Coupled Asymmetric T-Shaped Monopole Antenna for Linear and Circular Polarization Applications

Abinash Kumar Singh, Reshmi Dhara, and Sanoj Mahato

1 Introduction

Nowadays, dual-band dual-polarization (DBDP) antennas fulfill the multiple demand for wireless communication network. On the other side, antennas that have omnidirectional radiation pattern pose very useful for global positioning systems (GPSs) and for personal mobile system. On the contradictory side, antennas having unidirectional radiation pattern had become a very pivotal for the wireless communication technology in the likes of point-to-point communication, satellite systems.

Dual-band antennas having linear polarization and circular polarization properties are considered to be very popular among antenna technologies, and to fulfill the two bands design, distinct antenna designs are considered [1–6]. Researchers have been working on important methods that guarantee the linear polarization requirement [1–5]. And for dual-band dual-polarization (DBDP), the various research papers are mentioned in [1–5].

Inspired by the above research work and acknowledging the scope of DBDP, a coupled asymmetric T-shaped antenna is proposed in this paper and has a dimension of $50 \times 50 \times 50$ mm^3. The antenna is engraved into an FR4 epoxy material.

Here, the proposed antenna poses both circular polarization and linear polarization. And to achieve this polarization, a gap of 0.5 mm is introduced in the coupled asymmetric T-shaped radiating patch.

A. K. Singh · R. Dhara (✉) · S. Mahato
Department of Electronics and Communication Engineering, National Institute of Technology, Sikkim, India
e-mail: reshmidhara@nitsikkim.ac.in

A. K. Singh
e-mail: m210028@nitsikkim.ac.in

S. Mahato
e-mail: phec220008@nitsikkim.ac.in

The proposed design in this paper has applications like it can be used for WiMAX, WLAN, and the higher resonating frequency can be used for C-band like broadcasting and fixed mobile communication application. The T-shaped antenna shows right-hand circular polarization (RHCP) characteristics as well as linear polarization. Given that the suggested antenna has dual-band dual-polarization, its lower resonating band is linearly polarized (LP), while its higher resonating band is right-handed circularly polarized (RHCP).

As a result, the suggested antenna is simulated and optimized with electromagnetic simulation software Ansys HFSS software, and the simulated results are observed and examined further in the software.

2 Antenna Design

The top and the bottom view of the suggested coupled asymmetric T-shaped antenna are shown in Fig. 1b, c. The simulated overview dimensions of the suggested antenna are being shown in Fig. 1a. The coupled asymmetric T-shaped antenna is excited with 50 Ω lumped port in the Ansys HFSS simulator. It is implemented on FR4 substrate. The antenna is resonating at lower band 2.2 GHz, having dielectric constant (ε_r) 4.4 and loss tangent (tan δ)0.02 (Figs. 2 and 3; Table 1).

A 50 Ω microstrip line is used to feed the proposed linked asymmetric T-shaped antenna for signal transmission. Antenna operates on inset microstrip line feedline in this instance. The antenna is designed to have a coupled asymmetric T-like shape to achieve the desired dual-band dual-polarization (DBDP) characteristics. The coupled line T-shape, $L_1 = 20$ mm and $L_2 = 20$ mm, is placed 0.5 mm apart to achieve DBDP, and $L_2 = 20$ mm is shifted a little bit left side to get the desired result.

The ground plane and the substrate are truncated with triangle of length $P = 31.2$ mm at two of the corners and also a square shape length of $S = 34$ mm which

(a)   (b)   (c)
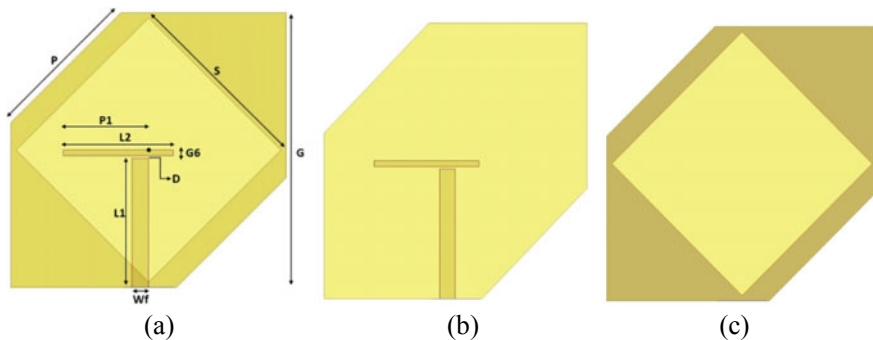
Fig. 1 Suggested antennas dimension, top view and bottom view in Ansys HFSS (**a**) views of the planned antenna are shown in (**b**), respectively (**c**)
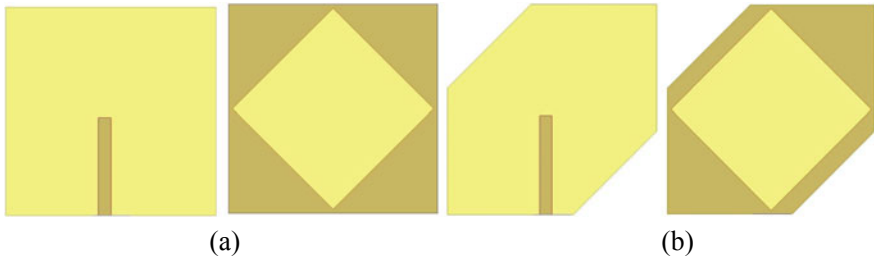
Fig. 2 Top and bottom view of antenna1 (**a**), top and bottom view of antenna2 (**b**)
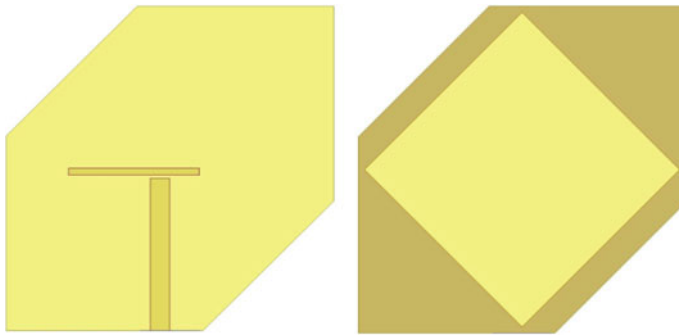
Fig. 3 Top and bottom view of antenna3

Table 1 Suggested antenna's size

Parameter	Value (mm)	Parameter	Value (mm)
G	50	L_1	20
P	31.2	G_6	1
S	34	P_1	15.5
W_f	3	L_2	20
L_{off}	1.5	D	0.5

is inclined at 45° is etched in the middle of the ground plane and doing so the result of the antenna improved gradually.

Figure 4 compares the three antennas' axial ratios and reflection coefficients (a). For all the three antennas the difference in result in impedance bandwidth is minimal but for axial ratio bandwidth for antenna1 the result is not good. After improvement in antenna2, the axial ratio bandwidth is slightly improved but still not satisfy the 3db requirement. In antenna3, the 3db requirement is met as it can be seen in Fig. 4b.

The desired result of the proposed antenna we got by optimizing three antenna design one after another. In antenna1, the radiating patch of $L_1 = 20$ mm was only considered without the coupled $L_2 = 20$ mm, and for the ground plane, the two opposite corners were not etched as is shown in Fig. 2a.
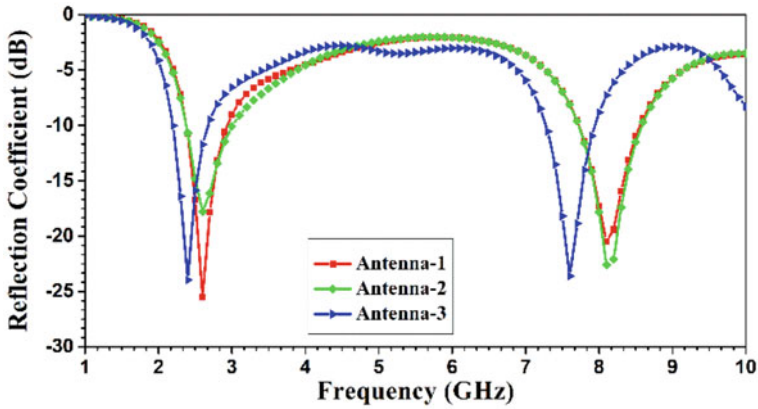
(a)
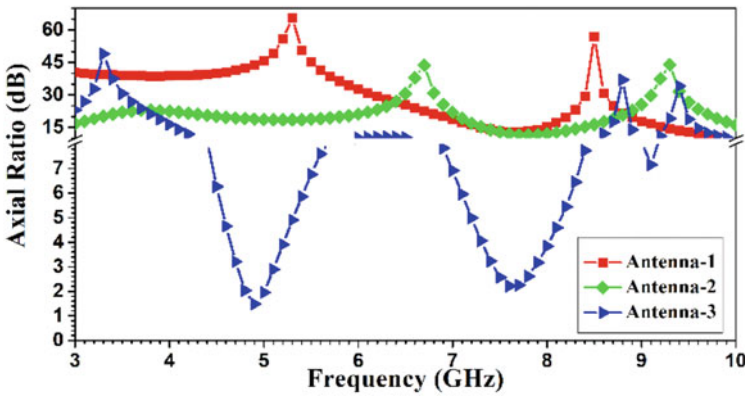
(b)

Fig. 4 Simulated impedance bandwidth (**a**), and ARBW (**b**) of antennas 1–3

For antenna2, the radiating patch was as same as antenna1, but the only change was in the ground plane, where two opposite edges of $P = 31.2$ mm are etched as it is shown in Fig. 2b. To meet the required DBDP parameter, the length P_1 and the width G_6 are varied. After setting $P_1 = 15.5$ mm and $G_6 = 1$ mm, the desired result for DBDP is observed. As dual-band dual-polarization characteristics are still no achievable in antenna1 and antenna2, as a result in the third antenna the coupled asymmetric T-shaped radiating patch was designed, and the desired results have been observed.

3 Parametric Optimization

The variation of width (G_6) of the coupled T-shaped patch is informed below. Five different values for G_6 are taken. From the result, it can be understood that there is minimal variation in impedance bandwidth but a huge change in axial ratio bandwidth. And from the result, it can be found that $G_6 = 1$ gives the best result in Fig. 5.

Similarly, for P_1, five values are taken, and it is found that $P_1 = 15.5$ gives the best result as it is seen in Fig. 6. When $P_1 = 15.5$, it creates asymmetric in geometry. For that reason, *Ex/Ey* magnitude is nearly equal to 90 which is desired for creation for circular polarization at higher frequency region.
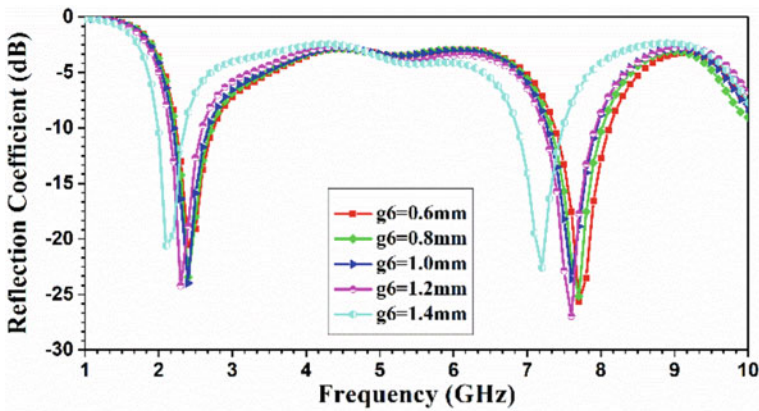
(a)
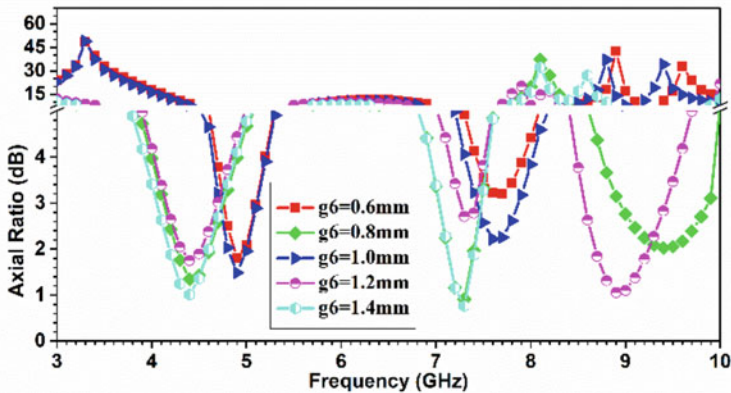
(b)

Fig. 5 Comparison of different values of G_6 for reflection coefficient (**a**), ARBW (**b**)
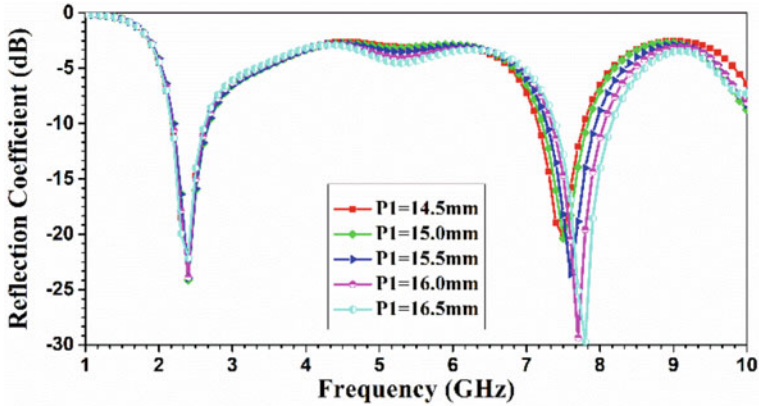
(a)
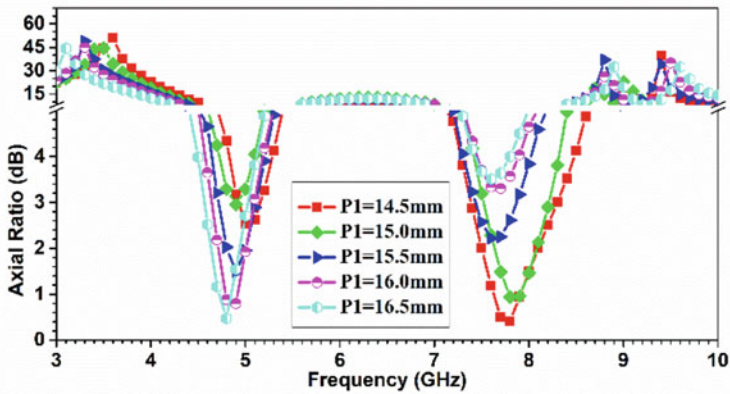
(b)

Fig. 6 Comparison of different values of P_1 for reflection coefficient (**a**), ARBW (**b**)

4 Result and Discussion

Optimization and simulation are done in Ansys HFSS. The antenna poses dual-band dual-polarization characteristics. Here for the lower resonating band, the antenna is linearly polarizing, and for the higher resonating band, the antenna is circularly polarizing. The simulated IBW at low frequency and higher frequency region are 2.17–2.66 GHz and 7.15–7.78 GHz which are resonating at 2.41 GHz and 7.46 GHz frequency. And they are 20.28% and 8.43%. The simulated 3 dB axial ratio bandwidth at higher resonating frequency is 7.38–8.3 GHz which is 11.73% at center resonating frequency 7.84 GHz.

Figure 8a depicts the simulated gain, with the two modes' greatest gains equaling 1.73 dB (linear polarized mode) and 4.06 dB (circular polarized mode). Figure 7 displays the simulated radiation efficiency (b). This antenna's maximum radiation efficiency is 95%, and its radiation efficiency is higher than 85% across the whole impedance bandwidth range.

Figure 9 depicts the surface current distribution of the proposed antenna at 7.88 GHz for Phase 0°, 90°, 180°, and 270°. Here, the surface current rotates counterclockwise. It is for this reason that the polarization is described as having a right-hand circular polarization (RHCP).

| (a) | (b) |
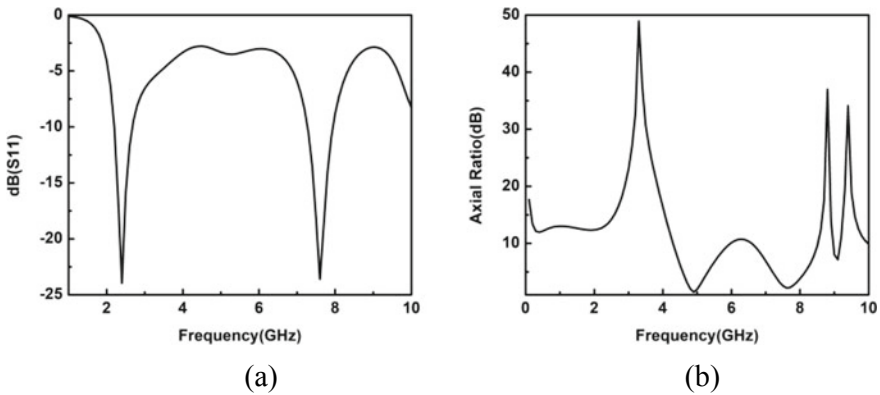
Fig. 7 Impedance bandwidth (**a**) and ARBW (**b**) of the proposed antenna

| (a) | (b) |
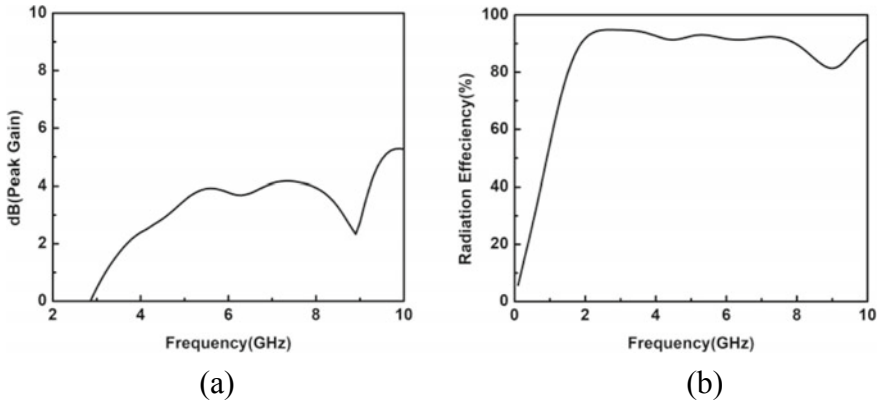
Fig. 8 Peak gain (**a**) and radiation efficiency (**b**) of the proposed antenna
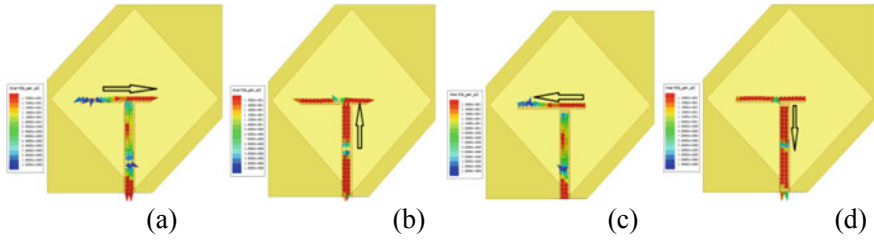
Fig. 9 Simulated surface current distribution at 7.84 GHz for Phase 0° (**a**), Phase 90° (**b**), Phase 180° (**c**), Phase 270° (**d**)

Figure 10 depicts the simulated electric field radiation pattern for Eco (elevation plane, *XZ* plane, = 90°) and Ecross (azimuthal plane, *XY* plane, = 0°) for the linked asymmetric T-shaped antenna operating at the frequency of 7.84 GHz. (a) It suggests that, like a traditional dipole mode antenna, the radiation pattern for the *E*-field is omnidirectional. Figure 10 depicts the suggested antenna's predicted magnetic field radiation pattern for the H_{co} plane (azimuthal plane, *XY* plane, = 0°) and H_{cross} plane (elevation plane, *XY* plane, = 90°). (b) Cross-polarization radiations are similarly strong in this area, but they are 15 dB less intense than co-polarization radiations.

Figures 11a and b show the radiation patterns for the suggested antenna for LHCP and RHCP at frequencies of 7.84 GHz and = 0° (*XZ* plane) and = 90° (*YZ* plane), respectively. Cross-polarization levels are 17 dB smaller than co-polarization levels in the broadside direction at the given frequency, and the steady radiation pattern at that frequency is obtained.

The simulated IBW at low frequency and higher frequency region are 2.17–2.66 GHz and 7.15–7.78 GHz which are resonating at 2.41 GHz and 7.46 GHz frequency. And they are 20.28% and 8.43%. The simulated 3 dB axial ratio bandwidth at higher resonating frequency is 7.38–8.3 GHz which is 11.73% at center resonating frequency 7.84 GHz. RHCP characteristics are exhibits by using this design.

The highest peak gain is 4.31 dBi at 7.4 GHz. The maximum radiation efficiency of this implemented antenna is 95%, and throughout the impedance bandwidth region, the radiation efficiency is greater than 85%.

The proposed antenna in this work has the highest axial ratio bandwidth (ARBW) when compared to other relevant papers, as can be seen in Table 2. The proposed antenna was also found to have an excellent impedance bandwidth (IBW).

(a)

(b)
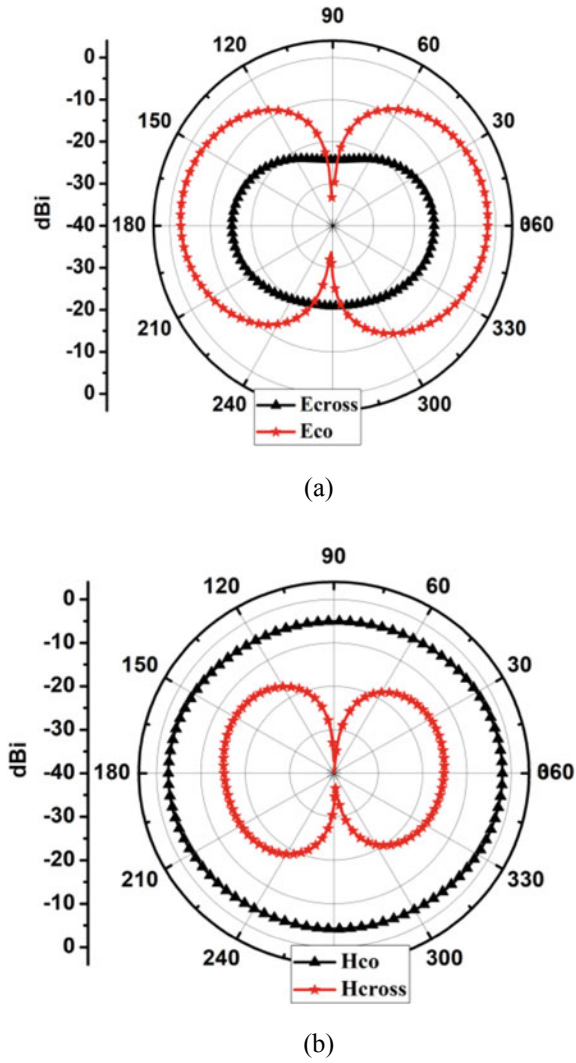
Fig. 10 Simulated radiation pattern of suggested antenna for E_{co} ($\varphi = 90°$) plane and E_{cross} ($\varphi = 0°$) plane (**a**) and H_{co} ($\varphi = 0°$) plane and H_{cross} ($\varphi = 90°$) plane (**b**)
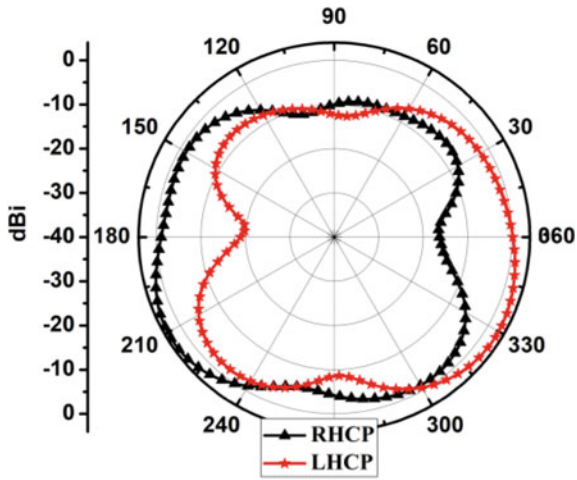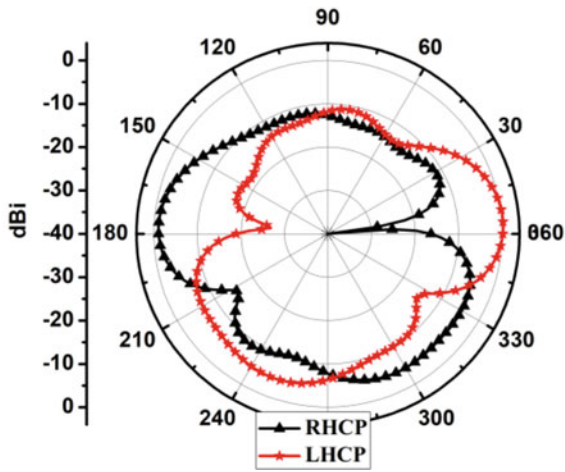
Fig. 11 Simulated radiation patterns (LHCP and RHCP) in *XZ* ($\varphi = 0°$) (**a**) and *YZ* ($\varphi = 90°$) (**b**) planes

Table 2 Comparison between recommended antennas with related DBDP antenna

Ref. (year)	Antenna size (mm^3)	Polarization	IBW (%)	Freq. (GHz)	ARBW (%)	fc (GHz)
1 (2016)	40 × 40 × 3	LP, RHCP	5.1, 14.9	4.42, 5.74	8.8	5.815
2 (2017)	R = 23.5, H = 6.508	LP, RHCP	2, 1	1.38, 1.57	1.39	1.576
3 (2017)	40 × 40 × 3	LP, LHCP	4.81, 23	4.72, 5.78	8.81	5.815
4 (2019)	50 × 50 × 2.325	LP, LHCP	8, 5.8	2.5, 5.92	5.91	5.75
5 (2021)	25 × 20 × 1.6	LP, LHCP	7.72, 14.26	9.45, 11.56	10.74	12.13
Proposed work	50 × 50 × 1.6	LP, LHCP	20.28, 8.43	2.4, 7.78	11.73	7.84

5 Conclusion

To accomplish the dual-band dual-polarized (DBDP) characteristics, the coupled asymmetric T-shaped antenna is fed with an inset microstrip feed line in this paper. A coupled asymmetric T-shaped microstrip feeding line which is having asymmetric characteristics is connected to a 50 Ω microstrip line. A quadratic groove is cut out of the ground plane and the substrate to provide the dipole mode's bandwidth. The circular polarization (CP) is produced by introducing linked asymmetric T-shaped design, while the dipole's linear polarization (LP) properties are unaltered. A small gap is introduced in the coupled T-shaped to attain impedance matching in two of the frequency bands. The simulated IBW at lower and higher frequency regions is 2.17–2.66 GHz, 20.28% and 7.15–7.78 GHz, 8.43% which are resonating at 2.41 GHz and 7.46 GHz frequencies. The simulated 3 dB axial ratio bandwidth at higher resonating frequency is 7.38–8.3 GHz, 11.73% at center resonating frequency 7.84 GHz. Right-hand circular polarization (RHCP) characteristics are exhibited by using this design. The highest peak gain is 4.31 dBi at 7.4 GHz. The maximum radiation efficiency of this implemented antenna is 95%, and throughout the impedance bandwidth region, the radiation efficiency is greater than 85%. The proposed design in this paper has multiple applications like it can be used for at lower resonating frequency linearly polarized band WiMAX, WLAN application, and higher resonating frequency band can be used for C-band: broadcasting and fixed mobile communication application.

References

1. Cao WQ (2016) Compact dual-band dual-mode circular patch antenna with broadband unidirectional linearly polarised and omnidirectional circularly polarised characteristics. IET Microwaves Antennas Propag 10(2):223–229

2. Lin J, Qian Z, Cao W, Shi S, Wang Q, Zhong W (2017) A low-profile dual-band dual-mode and dual-polarized antenna based on AMC. IEEE Antennas Wirel Propag Lett 16:2473–2476
3. Ding K, Gao C, Wu Y, Qu D, Zhang B, Wang Y (2017) Dual-band and dual-polarized antenna with endfire radiation. IET Microwaves Antennas Propag 11(13):1823–1828
4. Mao CX, Werner DH, Zhang Y, Zhang XY (2019) Compact dual-band dual-mode antenna with omni-/unidirectional radiation characteristics. IEEE Antennas Wirel Propag Lett 18(12):2657–2660
5. Dhara R, Kundu T (2021) Dual-band dual-mode antenna with polarization diversity. Radioelectron Commun Syst 64(5):266–280
6. Balanis CA (2005) Antenna theory analysis and design, 3rd edn. Wiley, New York

New Derivative-Free Families of Four-Parametric with and Without Memory Iterative Methods for Nonlinear Equations

G Thangkhenpau, Sunil Panday, and Shubham Kumar Mittal

1 Introduction

In the realm of numerical computations, a vast number of problems are expressed using nonlinear equations of the following form:

$$\Omega(s) = 0, \tag{1}$$

where $\Omega : D \subseteq \mathbb{R} \to \mathbb{R}$ is a real function defined on D, an open interval. Finding the numerical solutions of these problems expressed by (1) has always been a challenging task but at the same time of great importance due to its numerous applications in various branches of science and engineering. Iterative methods are extensively used for solving these problems in order to get approximate solutions of (1) but with high accuracy. The following is one such iterative method, called the Newton's method [1], which is widely used for finding the simple roots of (1):

$$s_{n+1} = s_n - \frac{\Omega(s_n)}{\Omega'(s_n)}, \quad n = 0, 1, 2, \ldots \tag{2}$$

It is a classical optimal one-point without memory method with quadratic order of convergence. However, due to the requirement of derivative evaluation and its low convergence order, the Newton's method (2) is not suitable for many practical uses. As a result, various multi-point without memory methods have been developed

G Thangkhenpau (✉) · S. Panday · S. K. Mittal
Department of Mathematics, National Institute of Technology Manipur, Langol, Imphal, Manipur 795004, India
e-mail: tkpguite92@gmail.com

S. Panday
e-mail: sunilpanday@hotmail.co.in

and studied in literature which have higher convergence order with higher efficiency [2–4]. An iterative method is called optimal if it satisfies the unproved Kung–Traub's conjecture [5] according to which an iterative without memory method requiring k function evaluations per iteration is optimal if it has the convergence order of 2^{k-1}.

In the past decade, various multi-point with memory methods for finding simple roots of nonlinear equations using accelerating parameters have gained much attention among researchers [6–10]. To theoretically determine the efficiency of an iterative method, Ostrowski [11] introduced the efficiency index (EI) = $p^{\frac{1}{k}}$, where k is the number of function evaluations at each iteration and p is the order of convergence. In fact, it was Traub who first introduces the with memory method, known as the Traub–Steffensen method [1], using a parameter as accelerating parameter. The method is given below:

$$w_n = s_n + \alpha_n \Omega(s_n), \alpha_n \neq 0$$
$$s_{n+1} = s_n - \frac{\Omega(s_n)}{\Omega[s_n, w_n]}, \tag{3}$$

where $\Omega[s_n, w_n] = \frac{\Omega(s_n) - \Omega(w_n)}{s_n - w_n}$ and α_n is the accelerating parameter calculated as follows:

$$\alpha_n = -\frac{1}{N_1'(s_n)}; \ N_1 = \Omega(s_n) + (s - s_n)\Omega[s_n, w_n], \quad n \geq 0.$$

The method (3) has convergence order of 2.41 which is higher than the quadratic convergence order of Newton's method. This is achieved without any additional function evaluation. Also, unlike Newton's method, Traub–Steffensen method does not require any evaluation of the derivatives and is derivative-free. This has motivated us to develop new multi-point with and without memory iterative methods containing more number of accelerating parameters with increased order of convergence having high efficiency index of almost 2.

In this paper, we introduce new derivative-free four-parametric families of four-point with and without memory iterative methods for computing simple roots of nonlinear equations. Formulation of the family of with memory methods is based on the extension of the new family of without memory methods by using accelerating parameters without any extra function evaluations. As a result, the convergence order increases from 8 to 15.5156. The accelerating parameters are approximated by Newton's interpolatory polynomials through the best-saved points so as to obtain highly efficient family of with memory methods.

The remaining content of the paper has been structured as follows. In Sect. 2 development of the new derivative-free family of without memory methods is discussed, and the theoretical convergence properties are fully investigated. Section 3 deals with the development and convergence analysis of the new derivative-free family of with memory methods. Section 4 covers the numerical results and the comparison of the proposed families of with and without memory methods with other existing methods

on some test functions. Some real-world problems have been included in this section to confirm the applicability of the proposed families of with and without memory methods. Finally, Sect. 5 presents some concluding remarks.

2   Optimal Four-Parametric Family of Four-Point Without Memory Methods

Let us first consider the following non-optimal three-point Newton steps of eighth-order containing the first-order derivative.

$$
\begin{aligned}
y_n &= s_n - \frac{\Omega(s_n)}{\Omega'(s_n)} \\
z_n &= y_n - \frac{\Omega(y_n)}{\Omega'(y_n)} \\
s_{n+1} &= z_n - \frac{\Omega(z_n)}{\Omega'(z_n)}
\end{aligned}
\tag{4}
$$

To minimize the number of function evaluations from the above Eq. (4), we first approximate $\Omega'(y_n)$ using the following expression:

$$
\Omega'(y_n) \approx \frac{\Omega'(s_n)}{1 + \frac{\Omega(y_n)}{\Omega(s_n)} \frac{\Omega(s_n) + (\lambda - 1)\Omega(y_n)}{\Omega(s_n) - \Omega(y_n)}}, \quad \lambda \in \mathbb{R}
\tag{5}
$$

Then, we approximate $\Omega'(s_n)$ from the first two steps of the above Eq. (4) as follows:

$$
\begin{aligned}
y_n &= s_n - \frac{\Omega(s_n)}{\Omega[s_n, w_n] + \beta\Omega(w_n)}, \quad w_n = s_n + \alpha\Omega(s_n) \\
z_n &= y_n - \left(1 + \frac{\Omega(y_n)}{\Omega(s_n)} \frac{\Omega(s_n) + (\lambda - 1)\Omega(y_n)}{\Omega(s_n) - \Omega(y_n)}\right) \frac{\Omega(y_n)}{\Omega[y_n, w_n] + \beta\Omega(w_n)} \\
s_{n+1} &= z_n - \frac{\Omega(z_n)}{\Omega'(z_n)},
\end{aligned}
\tag{6}
$$

where λ is any real parameter and $\alpha, \beta \in \mathbb{R} - \{0\}$.

We want to make this Eq. (6) optimal as well as derivative-free. So, we approximate $\Omega'(z_n)$ in the last step of (6) by the following polynomial:

$$
Q(v) = l_0 + l_1(v - z_n) + l_2(v - z_n)^2 + l_3(v - z_n)^3,
\tag{7}
$$

where l_0, l_1, l_2 and l_3 are some unknowns to be determined by means of the following conditions:

$$Q(s_n) = \Omega(s_n), \ Q(y_n) = \Omega(y_n), \ Q(w_n) = \Omega(w_n), \ Q(z_n) = \Omega(z_n).$$

Now, solving Eq. (7) under the above conditions and simplifying, we obtain the values of l_0, l_1, l_2 and l_3 as follows:

$$l_0 = \Omega(z_n) \tag{8}$$

$$l_3 = \Omega[s_n, y_n, z_n, w_n] \tag{9}$$

$$l_2 = \Omega[y_n, z_n, w_n] - l_3(y_n + w_n - 2z_n) \tag{10}$$

$$l_1 = \Omega[z_n, w_n] - l_2(w_n - z_n) - l_3(w_n - z_n)^2, \tag{11}$$

where $\Omega[x, y, z] = \frac{\Omega[x,y] - \Omega[y,z]}{x-z}$ and $\Omega[x, y, z, v] = \frac{\Omega[x,y,z] - \Omega[y,z,v]}{x-v}$ are second and third divided differences, respectively.

Using (9), (10) and (11), the approximation of $\Omega'(z_n)$ from Eq. (7) is obtained as follows:

$$\Omega'(z_n) \approx Q'(z_n) = l_1 = \Omega[z_n, w_n] - l_2(w_n - z_n) - l_3(w_n - z_n)^2. \tag{12}$$

Now, substituting (12) in (6) and adding two new parameters $\gamma, \delta \in \mathbb{R} - \{0\}$ in the last two steps, we obtain a new optimal derivative-free family of four-point without memory methods which are presented below. We shall denote it by FM8.

$$w_n = s_n + \alpha\Omega(s_n)$$
$$y_n = s_n - \frac{\Omega(s_n)}{\Omega[s_n, w_n] + \beta\Omega(w_n)}$$
$$z_n = y_n - \left(1 + \frac{\Omega(y_n)}{\Omega(s_n)} \frac{\Omega(s_n) + (\lambda - 1)\Omega(y_n)}{\Omega(s_n) - \Omega(y_n)}\right) \times$$
$$\frac{\Omega(y_n)}{\Omega[y_n, w_n] + \beta\Omega(w_n) + \gamma(y_n - s_n)(y_n - w_n)}$$
$$s_{n+1} = z_n - \frac{\Omega(z_n)}{Q'(z_n) + \delta(z_n - s_n)(z_n - y_n)(z_n - w_n)} \tag{13}$$

It is evident that the new family of without memory methods (13) consumes only four function evaluations per full iteration and is completely derivative-free. Also, it preserves the optimal convergence order eighth with the efficiency index $8^{\frac{1}{4}} \approx 1.682$. Now, we present the following theorem through which the convergence criteria of (13) are theoretically discussed.

Theorem 1 *Let $\xi \in D$ be a simple root of a sufficiently differentiable real function $\Omega : D \subseteq \mathbb{R} \to \mathbb{R}$, where D is an open interval. If an initial guess s_0 is close enough to*

ξ, *then the family of proposed methods defined by* (13) *has eighth-order convergence for any* $\lambda \in \mathbb{R}$ *and* $\alpha, \beta, \gamma, \delta \in \mathbb{R} - \{0\}$. *And, it has the error equation given by:*

$$
\begin{aligned}
\varepsilon_{n+1} = {} & \frac{1}{\Omega'(\xi)^2}\left(1 + \Omega'(\xi)\alpha\right)^4(\beta + d_2)^2 \\
& \left(-\gamma + \Omega'(\xi)\left(1 + \Omega'(\xi)\alpha\right)\beta^2(-1 + \lambda)\right. \\
& + \Omega'(\xi)\left(2\beta\left(-2 + \Omega'(\xi)\alpha(-1 + \lambda) + \lambda\right)d_2\right. \\
& \left.+\left(-3 + \Omega'(\xi)\alpha(-1 + \lambda) + \lambda\right)d_2^2 + d_3\right)\right) \\
& \left(\delta + d_2\left(-\gamma + \Omega'(\xi)\left(1 + \Omega'(\xi)\alpha\right)\beta^2(-1 + \lambda)\right.\right. \\
& + \Omega'(\xi)\left(2\beta\left(-2\Omega'(\xi)\alpha(-1 + \lambda) + \lambda\right)d_2\right) \\
& \left.\left.+\left(-3 + \Omega'(\xi)\alpha(-1 + \lambda) + \lambda\right)d_2^2 + d_3\right)\right) \\
& -\Omega'(\xi)d_4\Big)\varepsilon_n^8 + O\left(\varepsilon_n^9\right)
\end{aligned}
\tag{14}
$$

where $d_j = \frac{1}{j!}\frac{\Omega^j(\xi)}{\Omega'(\xi)}$, $j = 2, 3, ...$, and $\varepsilon_n = s_n - \xi$ is the error at n^{th} iteration.

Proof We construct and apply the following self-explained Mathematica code for the proof of the optimal eighth order of convergence of (13).

$$\Omega\big[\varepsilon_-\big] := \Omega'(\xi)\left(\varepsilon + d_2\varepsilon^2 + d_3\varepsilon^3 + d_4\varepsilon^4 + d_5\varepsilon^5 + d_6\varepsilon^6 + d_7\varepsilon^7 + d_8\varepsilon^8\right);$$

$$\varepsilon_w = \varepsilon + \alpha\Omega[\varepsilon] \quad (*\varepsilon_w = w - \xi*)$$

$$\text{Out[1]} : \left(1 + \Omega'(\xi)\alpha\right)\varepsilon + O[\varepsilon]^2$$

$$\Omega\big[x_-, y_-\big] := \frac{\Omega[x] - \Omega[y]}{x - y};$$

$$\Omega\big[x_-, y_-, z_-\big] := \frac{\Omega[x, y] - \Omega[y, z]}{x - z};$$

$$\varepsilon_y = \text{Series}\left[\varepsilon - \frac{\Omega[\varepsilon]}{\Omega[\varepsilon, \varepsilon_w] + \beta\Omega[\varepsilon_w]}, \{\varepsilon, 0, 8\}\right] /\!/\text{Simplify}$$

$$\text{Out[2]} : \left(1 + \Omega'(\xi)\alpha\right)(\beta + d_2)\varepsilon^2 + O[\varepsilon]^3$$

$$\varepsilon_z = \varepsilon_y - \left(1 + \frac{\Omega\big[\varepsilon_y\big]}{\Omega[\varepsilon]}\frac{\Omega[\varepsilon] + (\lambda - 1)\Omega\big[\varepsilon_y\big]}{\Omega[\varepsilon] - \Omega\big[\varepsilon_y\big]}\right) \times$$

$$\frac{\Omega\big[\varepsilon_y\big]}{\Omega\big[\varepsilon_y, \varepsilon_w\big] + \beta\Omega[\varepsilon_w] + \gamma\left(\varepsilon_y - \varepsilon\right)\left(\varepsilon_y - \varepsilon_w\right)}/\!/\text{Simplify}$$

$$\text{Out[3]}: \frac{-1}{\Omega'(\xi)}\big(1 + \Omega'(\xi)\alpha\big)^2(\beta + d_2)$$

$$\big(-\Omega'(\xi)\beta^2 - \Omega'(\xi)^2\alpha\beta^2 - \gamma + \Omega'(\xi)\beta^2\lambda$$

$$+\ \Omega'(\xi)^2\alpha\beta^2\lambda + 2\Omega'(\xi)\beta\big(-2 + \Omega'(\xi)\alpha(-1 + \lambda) + \lambda\big)d_2$$

$$+\Omega'(\xi)\big(-3 + \Omega'(\xi)\alpha(-1 + \lambda) + \lambda\big)d_2^2 + \Omega'(\xi)d_3\big)\varepsilon^4 + O[\varepsilon]^5$$

$$l_3 = \Omega\big[\varepsilon, \varepsilon_y, \varepsilon_z, \varepsilon_w\big];$$

$$l_2 = \text{Series}\big[\Omega\big[\varepsilon_y, \varepsilon_z, \varepsilon_w\big] - l_3\big(\varepsilon_y + \varepsilon_w - 2\varepsilon_z\big), \{\varepsilon, 0, 8\}\big]//\text{Simplify};$$

$$Q'(\varepsilon_z) = \text{Series}\big[\Omega\big[\varepsilon_z, \varepsilon_w\big] - l_2(\varepsilon_w - \varepsilon_z) - l_3(\varepsilon_w - \varepsilon_z)^2, \{\varepsilon, 0, 8\}\big]//\text{Simplify};$$

$$\varepsilon_{n+1} = \text{Series}\left[\varepsilon_z - \frac{\Omega[\varepsilon_z]}{Q'(\varepsilon_z) + \delta(\varepsilon_z - \varepsilon)(\varepsilon_z - \varepsilon_y)(\varepsilon_z - \varepsilon_w)}, \{\varepsilon, 0, 8\}\right]//\text{FullSimplify}$$

$$\text{Out[4]}: \frac{1}{\Omega'(\xi)^2}\big(1 + \Omega'(\xi)\alpha\big)^4(\beta + d_2)^2$$

$$\big(-\gamma + \Omega'(\xi)\big(1 + \Omega'(\xi)\alpha\big)\beta^2(-1 + \lambda) + \Omega'(\xi)$$

$$\big(2\beta\big(-2 + \Omega'(\xi)\alpha(-1 + \lambda) + \lambda\big)d_2$$

$$+\big(-3 + \Omega'(\xi)\alpha(-1 + \lambda) + \lambda\big)d_2^2 + d_3\big)$$

$$\big(\delta + d_2\big(-\gamma + \Omega'(\xi)\big(1 + \Omega'(\xi)\alpha\big)\beta^2(-1 + \lambda)\big) - \Omega'(\xi)d_4\big)$$

$$+\Omega'(\xi)\big(2\beta\big(-2\Omega'(\xi)\alpha(-1 + \lambda) + \lambda\big)d_2\big)\big)$$

$$+\big(-3 + \Omega'(\xi)\alpha(-1 + \lambda) + \lambda\big)d_2^2 + d_3\big)\big)\varepsilon_n^8 + O\big(\varepsilon_n^9\big)$$

which shows that Eq. (13) is of optimal order eighth.

3   Four-Parametric Family of Four-Point with Memory Methods

From the error Eq. (14), the convergence order can be increased from 8 to 16 for the method (13) if $\alpha = -\frac{1}{\Omega'(\xi)}$, $\beta = -d_2$, $\gamma = \Omega'(\xi)d_3$ and $\delta = \Omega'(\xi)d_4$. However, the exact value of ξ is not available to us. As such, we shall use $\alpha = \alpha_n$, $\beta = \beta_n$, $\gamma = \gamma_n$ and $\delta = \delta_n$, where α_n, β_n, γ_n and δ_n are accelerating parameters which will be computed using the available information from the current and the previous iterations.

Now, to approximate the accelerating parameters α_n, β_n, γ_n and δ_n, we use interpolation as follows.

$$\alpha_n = -\frac{1}{\mathbb{N}_4'(s_n)}, \ \beta_n = -\frac{\mathbb{N}_5''(w_n)}{2\mathbb{N}_5'(w_n)}, \ \gamma_n = \frac{\mathbb{N}_6'''(y_n)}{6}, \ \delta_n = \frac{\mathbb{N}_7^{iv}(z_n)}{24}, \ n = 0, 1, 2, \ldots$$

(15)

where $\mathbb{N}_j(t)$, $j = 4, 5, 6, 7$ are Newton's interpolatory polynomials of j degrees set through the points, i.e., $s_n, y_n, z_n, w_n, s_{n-1}, y_{n-1}, w_{n-1}, z_{n-1}$.

Now, by using (15) in the method (13), we obtain the following new derivative-free family of with memory methods. We shall denote it by FWM8.

For a given $s_0, \alpha_0, \beta_0, \gamma_0, \delta_0$, we have $w_0 = s_0 + \alpha_0 \Omega(s_0)$. Then,

$$\alpha_n = -\frac{1}{\mathbb{N}_4'(s_n)}, \beta_n = -\frac{\mathbb{N}_5''(w_n)}{2\mathbb{N}_5'(w_n)}, \gamma_n = \frac{\mathbb{N}_6'''(y_n)}{6}, \delta_n = \frac{\mathbb{N}_7^{iv}(z_n)}{24}$$

$$w_n = s_n + \alpha_n \Omega(s_n)$$

$$y_n = s_n - \frac{\Omega(s_n)}{\Omega[s_n, w_n] + \beta_n \Omega(w_n)}$$

$$z_n = y_n - \left(1 + \frac{\Omega(y_n)}{\Omega(s_n)} \frac{\Omega(s_n) + (\lambda - 1)\Omega(y_n)}{\Omega(s_n) - \Omega(y_n)}\right) \times$$

$$\frac{\Omega(y_n)}{\Omega[y_n, w_n] + \beta_n \Omega(w_n) + \gamma_n(y_n - s_n)(y_n - w_n)}$$

$$s_{n+1} = z_n - \frac{\Omega(z_n)}{Q'(z_n) + \delta_n(z_n - s_n)(z_n - y_n)(z_n - w_n)}$$

(16)

Lemma 1 *If* $\alpha_n = -\frac{1}{\mathbb{N}_4'(s_n)}$, $\beta_n = -\frac{\mathbb{N}_5''(w_n)}{2\mathbb{N}_5'(w_n)}$, $\gamma_n = \frac{\mathbb{N}_6'''(y_n)}{6}$ *and* $\delta_n = \frac{\mathbb{N}_7^{iv}(z_n)}{24}$, $n = 0, 1, 2, \ldots$, *then the following estimates*

$$1 + \alpha_n \Omega'(\alpha) \sim \varepsilon_{n-1}\varepsilon_{n-1,y}\varepsilon_{n-1,w}\varepsilon_{n-1,z}$$

(17)

$$\beta_n + d_2 \sim \varepsilon_{n-1}\varepsilon_{n-1,y}\varepsilon_{n-1,w}\varepsilon_{n-1,z}$$

(18)

$$K_n \sim \varepsilon_{n-1}\varepsilon_{n-1,y}\varepsilon_{n-1,w}\varepsilon_{n-1,z}$$

(19)

$$L_n \sim \varepsilon_{n-1}\varepsilon_{n-1,y}\varepsilon_{n-1,w}\varepsilon_{n-1,z}$$

(20)

hold, where

$$K_n = -\gamma_n + \Omega'(\xi)\left(1 + \Omega'(\xi)\alpha_n\right)\beta_n^2(-1 + \lambda)$$
$$+ \Omega'(\xi)\left(2\beta_n\left(-2 + \Omega'(\xi)\alpha_n(-1 + \lambda) + \lambda\right)d_2\right.$$
$$+ \left(-3 + \Omega'(\xi)\alpha_n(-1 + \lambda) + \lambda\right)d_2^2 + d_3\right),$$

$$L_n = \delta_n + d_2\left(-\gamma_n + \Omega'(\xi)\left(1 + \Omega'(\xi)\alpha_n\right)\beta_n^2(-1 + \lambda)\right.$$

$$+ \Omega'(\xi)\big(2\beta_n\big(-2 + \Omega'(\xi)\alpha_n(-1 + \lambda) + \lambda\big)d_2$$
$$+\big(-3 + \Omega'(\xi)\alpha_n(-1 + \lambda) + \lambda\big)d_2^2 + d_3\big)\big) - \Omega'(\xi)d_4,$$

$$\varepsilon_n = s_n - \xi, \ \varepsilon_{n,y} = y_n - \xi, \ \varepsilon_{n,w} = w_n - \xi, \ \varepsilon_{n,z} = z_n - \xi.$$

Proof As for the proof, see Lemma 1 of [12].

Now, we present the following theorem for analyzing the R-order of convergence [13] of the derivative-free four-parametric family of four-point with memory methods (16).

Theorem 2 *If an initial guess s_0 is sufficiently close to a root ξ of $\Omega(s) = 0$, the parameters α_n, β_n, γ_n and δ_n are calculated by the expressions in (15), then the R-order of convergence of the methods (16) is at least* 15.5156.

Proof Let the sequence of approximations $\{s_n\}$ produced by (16) converges to the root ξ with the order r. Then, we can write

$$\varepsilon_{n+1} \sim \varepsilon_n^r, \quad \text{where } \varepsilon_n = s_n - \xi. \tag{21}$$

Then,

$$\varepsilon_n \sim \varepsilon_{n-1}^r \tag{22}$$

Thus,

$$\varepsilon_{n+1} \sim \varepsilon_n^r = \big(\varepsilon_{n-1}^r\big)^r = \varepsilon_{n-1}^{r^2} \tag{23}$$

Let the iterative sequences $\{w_n\}$, $\{y_n\}$ and $\{z_n\}$ have orders r_1, r_2 and r_3, respectively. Then, using (21) and (22) gives

$$\varepsilon_{n,w} \sim \varepsilon_n^{r_1} = \big(\varepsilon_{n-1}^r\big)^{r_1} = \varepsilon_{n-1}^{rr_1} \tag{24}$$

$$\varepsilon_{n,y} \sim \varepsilon_n^{r_2} = \big(\varepsilon_{n-1}^r\big)^{r_2} = \varepsilon_{n-1}^{rr_2} \tag{25}$$

$$\varepsilon_{n,z} \sim \varepsilon_n^{r_3} = \big(\varepsilon_{n-1}^r\big)^{r_3} = \varepsilon_{n-1}^{rr_3} \tag{26}$$

From Theorem 1, we have

$$\varepsilon_{n,w} \sim \big(1 + \alpha_n\Omega'(\xi)\big)\varepsilon_n \tag{27}$$

$$\varepsilon_{n,y} \sim \big(1 + \alpha_n\Omega'(\xi)\big)(\beta_n + d_2)\varepsilon_n^2 \tag{28}$$

$$\varepsilon_{n,z} \sim \left(1 + \alpha_n \Omega'(\xi)\right)^2 (\beta_n + d_2) K_n \varepsilon_n^4 \tag{29}$$

$$\varepsilon_{n+1} \sim \left(1 + \alpha_n \Omega'(\xi)\right)^4 (\beta_n + d_2)^2 K_n L_n \varepsilon_n^8 \tag{30}$$

Using the above Lemma 1 and (27)–(30), we get

$$\varepsilon_{n,w} \sim \left(1 + \alpha_n \Omega'(\xi)\right)\varepsilon_n = \varepsilon_{n-1}^{r+r_1+r_2+r_3+1} \tag{31}$$

$$\varepsilon_{n,y} \sim \left(1 + \alpha_n \Omega'(\xi)\right)(\beta_n + d_2)\varepsilon_n^2 = \varepsilon_{n-1}^{2r+2r_1+2r_2+2r_3+2} \tag{32}$$

$$\varepsilon_{n,z} \sim \left(1 + \alpha_n \Omega'(\xi)\right)^2 (\beta_n + d_2) K_n \varepsilon_n^4 = \varepsilon_{n-1}^{4r+4r_1+4r_2+4r_3+4} \tag{33}$$

$$\varepsilon_{n+1} \sim \left(1 + \alpha_n \Omega'(\xi)\right)^4 (\beta_n + d_2)^2 K_n L_n \varepsilon_n^8 = \varepsilon_{n-1}^{8r+8r_1+8r_2+8r_3+8} \tag{34}$$

Now, comparing the corresponding powers of ε_{n-1} on the right sides of (24) and (31), (25) and (32), (26) and (33), (23) and (34), we get

$$
\begin{aligned}
rr_1 - r - r_1 - r_2 - r_3 - 1 &= 0 \\
rr_2 - 2r - 2r_1 - 2r_2 - 2r_3 - 2 &= 0 \\
rr_3 - 4r - 4r_1 - 4r_2 - 4r_3 - 4 &= 0 \\
r^2 - 8r - 8r_1 - 8r_2 - 8r_3 - 8 &= 0
\end{aligned}
\tag{35}
$$

This system has the non-trivial solution $r_1 = 1.9394$, $r_2 = 3.8789$, $r_3 = 7.7578$ and $r = 15.5156$. Hence, the R-order of convergence of the proposed family of methods (16) is at least 15.5156. The proof is complete.

4 Numerical Results

In this section, we examine the performance and the computational efficiency of the newly developed with and without memory methods discussed in Sects. 2 and 3 and compare with some methods of similar nature available in literature. In particular, we have considered for the comparison, the following four-parametric methods: LAM8(3.31) [8], ZM8 (ZR1 from [9]) and ACM8 (M1 from [10]). All numerical tests are executed using the programming software Mathematica 12.2. Throughout the whole computation, we have chosen the same values of the parameters $\alpha_0 = \beta_0 = \gamma_0 = \delta_0 = -1$ and $\lambda = 2$ in all the test functions in order to start the initial iteration. These same values are used for the corresponding parameters of all the compared methods in order to have fair comparison. Numerical test functions which comprise

some standard academic examples and real-life chemical engineering problems along with their simple roots (ξ) and initial guesses (s_0) are presented below.

Example 1 A standard academic test function given by:

$$\Omega_1(s) = e^{-s^2}(1 + s^3 + s^6)(s - 2). \tag{36}$$

It has a simple root $\xi = 2$. We start with the initial guess $s_0 = 2.3$.

Example 2 A standard academic test function given by:

$$\Omega_2(s) = \log(s^2 + s + 2) - s + 1. \tag{37}$$

It has a simple root $\xi \approx 4.1525907367571583$. We start with the initial guess $s_0 = 4.5$.

Example 3 A standard academic test function given by

$$\Omega_3(s) = \sin^2 s + s \tag{38}$$

It has a simple root $\xi = 0$. We start with the initial guess $s_0 = 0.6$.

Example 4 The azeotropic point of a binary solution problem given by the following nonlinear equation (for details see [14]).

$$\Omega_4(s) = \frac{FG(G(1-s)^2 - Fs^2)}{(s(F-G) + G)^2} + 0.14845, \tag{39}$$

where $F = 0.38969$ and $G = 0.55954$.

It has a simple root $\xi \approx 0.69147373574714142$. We start with the initial guess $s_0 = 1.1$.

In Tables 1 and 2, we have displayed the absolute residual errors $|\Omega(s_n)|$ at the first three iterations obtained by the compared methods. We also include the computational convergence order (COC) of each compared method which is computed by the following formula [15]:

$$\text{COC} = \frac{\log|\Omega(s_n)/\Omega(s_{n-1})|}{\log|\Omega(s_{n-1})/\Omega(s_{n-2})|}. \tag{40}$$

From the two Tables 1 and 2, the numerical results affirm the robust performance and high efficiency of the proposed with and without memory methods thus reaffirming their theoretical results. The proposed methods give better accuracy with high efficiency in terms of minimal residual errors after three iterations as compared to the other methods. Further, the COC supports the theoretical convergence order of the new proposed with and without memory methods in the test functions.

Table 1 Comparison of the without memory methods on the test functions

Methods	$\Omega(s)$	$|\Omega(s_1)|$	$|\Omega(s_2)|$	$|\Omega(s_3)|$	COC
ZM8	$\Omega_1(s)$	0.00083523	8.4618×10^{-26}	9.4210×10^{-202}	8.0000
ACM8	$\Omega_1(s)$	0.0061646	1.3517×10^{-20}	7.6376×10^{-162}	7.9986
LAM8	$\Omega_1(s)$	0.000015809	7.9235×10^{-41}	3.1542×10^{-323}	8.0000
FM8	$\Omega_1(s)$	9.6344×10^{-7}	2.6067×10^{-49}	7.4864×10^{-390}	8.0000
ZM8	$\Omega_2(s)$	0.19523	0.076326	3.9752×10^{-6}	10.502
ACM8	$\Omega_2(s)$	0.90214	0.13906	0.000042430	4.3292
LAM8	$\Omega_2(s)$	0.014457	8.6044×10^{-13}	1.5302×10^{-94}	7.9948
FM8	$\Omega_2(s)$	0.011023	7.2266×10^{-15}	3.4512×10^{-112}	7.9880
ZM8	$\Omega_3(s)$	6.7546×10^{-7}	2.7816×10^{-100}	1.9034×10^{-1594}	16.000
ACM8	$\Omega_3(s)$	5.8728×10^{-7}	2.9666×10^{-101}	5.3325×10^{-1610}	16.000
LAM8	$\Omega_3(s)$	1.4810×10^{-6}	7.9355×10^{-95}	3.6632×10^{-1507}	16.000
FM8	$\Omega_3(s)$	1.9346×10^{-7}	5.7032×10^{-109}	1.8559×10^{-1733}	16.000
ZM8	$\Omega_4(s)$	0.074840	1.3522×10^{-10}	1.4582×10^{-80}	8.002
ACM8	$\Omega_4(s)$	Divergent	Divergent	Divergent	–
LAM8	$\Omega_4(s)$	0.012310	1.9635×10^{-16}	8.0818×10^{-127}	8.000
FM8	$\Omega_4(s)$	0.0069383	2.2151×10^{-18}	2.3818×10^{-142}	8.001

Table 2 Comparison of the with memory methods on the test functions

Methods	$\Omega(s)$	$|\Omega(s_1)|$	$|\Omega(s_2)|$	$|\Omega(s_3)|$	COC
ZM8	$\Omega_1(s)$	0.00083523	4.9114×10^{-50}	4.7977×10^{-749}	15.120
ACM8	$\Omega_1(s)$	0.0061646	2.1885×10^{-39}	1.3451×10^{-598}	15.342
LAM8	$\Omega_1(s)$	0.000015809	4.4940×10^{-80}	5.0217×10^{-1237}	15.520
FWM8	$\Omega_1(s)$	9.6344×10^{-7}	1.3319×10^{-96}	5.0044×10^{-1498}	15.596
ZM8	$\Omega_2(s)$	0.19523	4.8315×10^{-25}	3.1255×10^{-375}	14.834
ACM8	$\Omega_2(s)$	0.90214	3.6884×10^{-11}	9.1383×10^{-174}	15.653
LAM8	$\Omega_2(s)$	0.014457	5.3706×10^{-43}	8.9914×10^{-669}	15.478
FWM8	$\Omega_2(s)$	0.011023	1.1528×10^{-44}	3.5410×10^{-695}	15.496
ZM8	$\Omega_3(s)$	6.7546×10^{-7}	1.3098×10^{-95}	5.0428×10^{-1467}	15.459
ACM8	$\Omega_3(s)$	5.872×10^{-7}	8.5805×10^{-99}	3.2362×10^{-1551}	15.816
LAM8	$\Omega_3(s)$	1.4810×10^{-6}	2.2038×10^{-95}	1.5131×10^{-1494}	15.751
FWM8	$\Omega_3(s)$	1.9346×10^{-7}	7.6149×10^{-104}	3.6249×10^{-1632}	15.853
ZM8	$\Omega_4(s)$	0.074840	4.6224×10^{-22}	3.3998×10^{-330}	15.247
ACM8	$\Omega_4(s)$	Divergent	Divergent	Divergent	–
LAM8	$\Omega_4(s)$	0.012310	4.0863×10^{-37}	3.1559×10^{-569}	15.433
FWM8	$\Omega_4(s)$	0.0069383	1.3673×10^{-39}	2.4610×10^{-609}	15.522

5 Concluding Remarks

We have presented in this paper new derivative-free families of with and without memory methods for finding the solutions of nonlinear equations. The use of four accelerating parameters in the with memory methods has enabled us to increase the convergence order of the without memory methods from 8 to 15.5156 and obtain very high-efficiency index of $15.5156^{\frac{1}{4}} \approx 1.9847$ without extra function evaluations. The numerical results further confirm the good performance, validity and applicability of the proposed with and without memory methods. They are found to be more efficient with better accuracy as compared to the existing methods in comparison in terms of minimal residual errors after three iterations for convergence toward the required simple roots.

References

1. Traub JF (1982) Iterative methods for the solution of equations. Am Math Soc 312
2. Singh A, Jaiswal JP (2016) A class of optimal eighth-order Steffensen-type iterative methods for solving nonlinear equations and their basins of attraction. Appl Math Inf Sci 10(1):251–257. https://doi.org/10.18576/amis/100125
3. Panday S, Sharma A, Thangkhenpau G (2023) Optimal fourth and eighth-order iterative methods for non-linear equations. J Appl Math Comput 69(1):953–971. https://doi.org/10.1007/s12190-022-01775-2
4. Behl R, Alshomrani AS, Chun C (2020) A general class of optimal eighth-order derivative free methods for nonlinear equations. J Math Chem 58:854–867
5. Kung HT, Traub JF (1974) Optimal order of one-point and multipoint iteration. J ACM (JACM) 21(4):643–651. https://doi.org/10.1145/321850.321860
6. Zafara F, Yasmina N, Kutbib MA, Zeshana M (2016) Construction of tri-parametric derivative free fourth order with and without memory iterative method. J Nonlinear Sci Appl 9:1410–1423
7. Chanu WH, Panday S, Thangkhenpau G (2022) Development of optimal iterative methods with their applications and basins of attraction. Symmetry 14(10):2020. https://doi.org/10.3390/sym14102020
8. Lotfi T, Assari P (2015) Two new three and four parametric with memory methods for solving nonlinear equations. Int J Ind Math 7(3):269–276
9. Zafar F, Cordero A, Torregrosa JR, Rafi A (2019) A class of four parametric with-and without-memory root finding methods. Comp and Math Methods 1:e1024
10. Cordero A, Janjua M, Torregrosa JR, Yasmin N, Zafar F (2018) Efficient four parametric with and without-memory iterative methods possessing high efficiency indices. Math Prob Eng 2018:1–12. https://doi.org/10.1155/2018/8093673
11. Ostrowski AM (1966) Solution of equations and systems of equations. Academic Press, New York-London
12. Džunić J (2013) On efficient two-parameter methods for solving nonlinear equations. Numer Algor 63:549–569. https://doi.org/10.1007/s11075-012-9641-3
13. Ortega JM, Rheinboldt WG (1970) Iterative solution of nonlinear equations in several variables. Academic Press, New York
14. Solaiman OS, Hashim I (2019) Efficacy of optimal methods for nonlinear equations with chemical engineering applications. Math Probl Eng 2019, Article ID 1728965, 11 p. https://doi.org/10.1155/2019/1728965
15. Petković MS (2011) Remarks on '"On a general class of multipoint root-finding methods of high computational efficiency."' SIAM J Numer Math 49:1317–1319

Computer Science and Engineering

Meetei Mayek, Hindi, and English Text Detection from Natural Scene Images Using YOLO

Chingakham Neeta Devi, Nella Kartheek, Bokka Purna Manikanta, Motha Yasaswini Saisree, and Manjeet

1 Introduction

Due to the presence of complex things in natural scene images, including variable text fonts and sizes, inclined lines, lighting conditions, and complex backdrops, detecting text is a critical task in computer vision research problems. For tasks like licence plate recognition, text-to-speech translation for those with visual impairments and various others, it is important to detect text in these images before recognizing them.

Manipuri is the mother tongue for people residing in Manipur located in northeastern India. There are many literary works on text detection for the English and Hindi languages, but there exists only finger count research works on Manipuri, which is an Indian language having two scripts—Meetei Mayek and Bengali. However, Bengali script was borrowed and used from the eighteenth century but Meetei Mayek script has been recently restored. Currently, Meetei Mayek is the official script in use, and it is employed for all writing; hence, signboards and others are written using the script in Manipur. In a nation like India, where there are numerous languages and many of them have similar scripts. According to a widely held belief, any image may only contain a maximum of three languages: the mother tongue of the particular geographical area, Hindi, and English. Also, there is a necessity in detecting the language as well. Additionally, no work has been reported yet on detecting text in natural scene images containing Meetei Mayek, Hindi, and English (Fig. 1).

The writing on the signs boards won't make sense to a visitor from another state or nation who doesn't speak the native language. The major goal of this effort is the creation of a system for precise text detection, and go for further recognizing it, and make use of language translation that may be used to translate into a tourist-friendly language. Some other applications include licence plate detection, helping autonomous vehicles to retrieve text, helping visually impaired people by text-to-

C. N. Devi · N. Kartheek (✉) · B. P. Manikanta · M. Y. Saisree · Manjeet
National Institute of Technology Manipur, Imphal 795004, India
e-mail: ai07kartheek@gmail.com

Fig. 1 Natural scene images consisting of Meetei Mayek, Hindi, English text

speech translation. This paper presents the first stage of the development. This paper explains the processes in detail for the Meetei Mayek, English, and Hindi text detection from natural scene images using YOLO V7 [1].

2 Literature Survey

Computer vision researchers found that multilingual text detection from images is a daunting and challenging task. Tian et al. [2] had proposed text detection system in natural scene images. Their work includes extraction of text line from images by character candidate detection method comprising of connected components and sequential processing. Lundgren et al. [3] had proposed a octShuffleMLT which make use of convolutional neural networks, with a less number of layers and parameters, that can precisely recognize multilingual scene text and can be used for issues having hardware constraints. A work on multilingual scene text recognition and language identification had been suggested by Saha et al. [4]. Their model finds the text using maximally stable extremal regions and followed by stroke width transform and then refines the ideas using a generative adversarial network. In order to identify the language of the detected scene texts for English, Hindi, Bangla, and Korean, a CNN-based architecture is utilized.

As English and Hindi languages had a wide range of research work been reported. But there are very few in the perspective of Meetei Mayek script. With the ability to accept an input of clipped text image of a printed document, Ghosh et al. [5] had presented an OCR system for hard copy of Meetei Mayek script. The data had undergone pre-processing, segmentation before classification. In order to perform classification using extracted features, a multi-staged SVM had been employed and gives the output in Unicode format. A Meetei Mayek pattern recognition system was designed by Hijam et al. [6]. Leveraging their manually written dataset with various sets of features and classification strategies, authors have examined their suggested CNN.

Algorithm presented in [7] makes use of edge-enriched MSER and SWT-based slightly modified text detection algorithm for Meetei Mayek script and achieved appreciable outcomes approximately 0.7 for f-measure, precision, and recall.

3 YOLO V7 Text Detection Process

This section describes the process of Meetei Mayek, Hindi, English text detection from natural scene images as it has. Among text detection methods [8], YOLO version 7 (YOLO V7) algorithm has been chosen for performing robust detection of Meetei Mayek, Hindi, and English text from a naturally scene image.

For each input image, the output will be a bounding box for each word of text present in the images and classify them with their respective language, namely Manipuri (Meetei Mayek), Hindi, or English. In order to achieve the accurate detection of words, the text detection system has been implemented using YOLO V7 and all the requirements of YOLO V7 for object detection have been followed. So the first step of the process is the collection of natural scene images that contain Meetei Mayek, Hindi, and English text. Every image may not have combination of all three languages but at least one of the languages is a must.

3.1 Word Annotation in Images

The next step after image data collection will be annotation of each word in the collected natural scene images. Here, in total we have collected around 609 images. Annotation of them manually is a time-consuming task; hence, a free image annotation tool has been used. In fact, there are several image annotation tools available for free, but we have chosen Roboflow software for our task because it is easy to use, collaborate with our teammates, and smoothen our annotation process. Each image is annotated in such a way that a tight bonding box is given around each word and label with their respective class names (Manipuri, Hindi, English). A sample annotation of an image is shown in Fig. 2 where blue bounding box represents English word, pink bounding box represents Manipuri word, and red bounding box represents Hindi word.

3.2 Image Preprocessing and Database Creation

The total database is split into 500 images for training, 88 validation images, and 21 testing images. As the data is large enough, this paper doesn't make use of image augmentation techniques and finally export the image data from roboflow in YOLO V7 dataset format stated below (Fig. 3).

Fig. 2 Word annotation for Meetei Mayek, Hindi, and English words
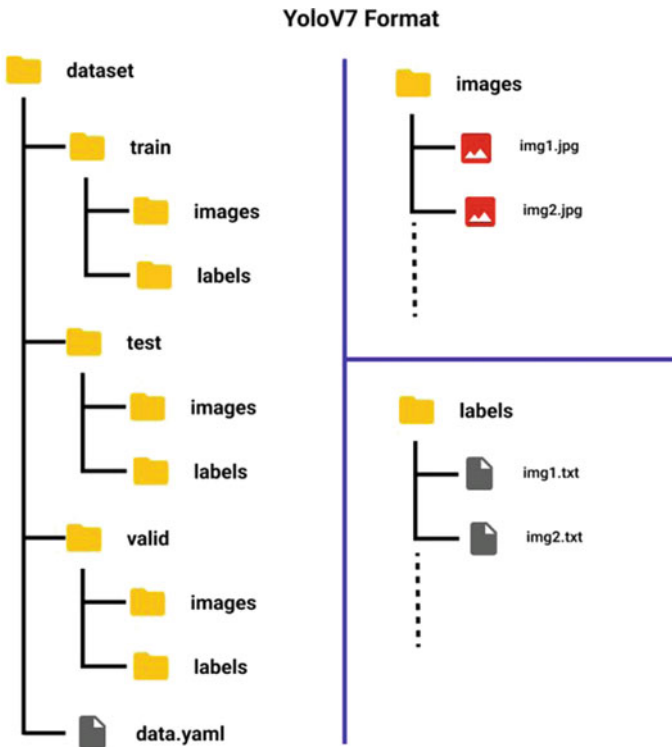
Fig. 3 Visualization of COCO dataset format accepted by YOLO V7

Fig. 4 Splitting of image into equally sized $M \times M$ grid

The *data.yaml* file contains the information about classes and path for train, test, and validation data. Each text file in labels folder has an information about class, coordinates for a box's centre, width, and height of a box.

3.3 Word Detection Parsing in YOLO

According to the You Only Look Once (YOLO) a deep learning architecture, detecting objects only takes one forward propagation via a neural network. It's gaining popularity over the years because of its speed and accuracy with minimal background errors. The newest version of YOLO is version 7 which has 106 layers in total. Convolutional layers in YOLO extract the features from the image to predict the probability of learned classes.

In YOLO, object detection is carried out as a regression problem by providing bounding boxes for detected objects and class probabilities for the same. Initially, the input image is split into equal sized $M \times M$ grid cells.

If the centre of any object is present in any grid cell, then that cell is in full control of identifying that object in an image. In Fig. 4, two red dots determine centres of two English words and three green dots represent box centres for three Meetei Mayek words. These dotted cells are responsible for detecting words.

Single grid cell may have the presence of multiple objects. So each grid predicts some bounding boxes and three class probabilities each for English, Hindi, and Manipuri. Additionally every bounding box has five key values (c_x, c_y, *width, height, conf*),

Fig. 5 Illustration of calculation of IOU

$$\text{IOU} \;=\; \frac{\text{area of overlap}}{\text{area of union}} \;=\; \frac{}{}$$

where

(c_x, c_y) = *box's centre coordinates*

(width, height) = box's width and height

conf = indicates the confidence score for the presence of any word.

Bounding Box Equation YOLO V7 by default employs nine anchor boxes [9], from which it derives the actual width and height of anticipated bounding boxes. The underlying equation for predicting the bounding box is shown below

$$B_x = \sigma(O_x) + T_x$$
$$B_y = \sigma(O_y) + T_y$$
$$B_w = A_w * e^{O_w}$$
$$B_h = A_h * e^{O_h}$$

where

(O_x, O_y, O_w, O_h) = outputs of neural network.

(A_w, A_h) = anchor box's width and height.

(T_x, T_y) = grid cell's top-left coordinates.

(B_x, B_y) = bounding box's centre.

(B_w, B_h) = bounding box's width and height.

Localization metrics

Intersection over Union (IOU) For the evaluation of object detection, IOU is a well-liked statistic to assess localization precision and compute localization flaws.

In Fig. 5, green bounding represents ground truth box, whereas predicted one is the red bounding box in the same area. Area of overlap is the intersection area, whereas area of union gives the whole outer area. This ratio provides a solid idea of how well the projected bounding box resembling the actual box. IOU ranges between 0 and 1 (both inclusive).
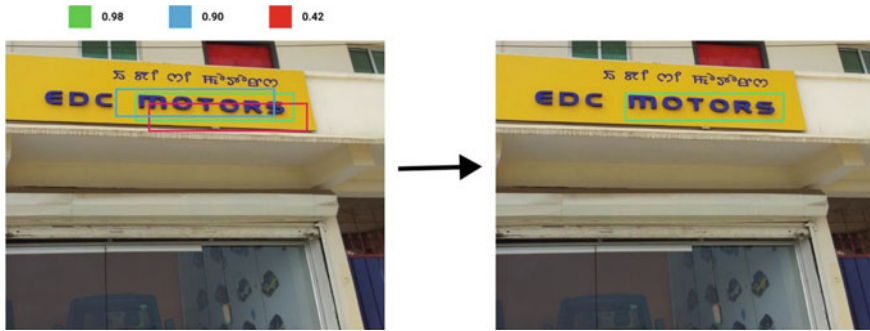
Fig. 6 Illustration of non-maximal suppression on a single word

Precision The ratio of predicted true positives (TP) to all predicted positives $(TP + FP)$.

Recall The proportion of true positive predictions (TP) to all of the ground truth positives $(TP + FN)$ is known as recall. Recall and precision are tested for boundary box predictions instead of predicting the classes when examining the relevance of outcomes in object detection. To be more specific, a prediction is judged positive if the IOU level exceeds 0.5 and negative if it is below 0.5.

Average Precision and Mean Average Precision(mAP) area covered by the precision–recall curve is said to be average precision. The mean of the average precision determined for all classes is known as the mAP. It's value ranges between 0 and 1 (higher is better).

Non-Maximal Suppression Each grid cell predicts the objects that appear within them. So due to several cells anticipating the same word with various bounding box predictions results in a lot of redundant predictions.

Non-maximal suppression (NMS) technique used by YOLO to suppress all the bounding boxes which has less IOU score. Figure 6 shows predicted boxes with their IOU scores for the same word. NMS suppresses all other boxes except green one. This applies for all other words, and finally we end up with a single bounding box for each object.

4 Experimental Results

The experiment carried out with 500 images for training, 88 images for validation, 21 images for testing. The model uses *yolov7.pt* (medium-sized weights) as initial weights, with a batch size of 16, and trained for 200 epochs, the values of the performance metrics have been given.

Table 1 Confusion matrix for Meetei Mayek, Hindi, English classes

Predicted	True			
	English	Hindi	Manipuri	Background FP
English	**0.92**	0.01	0.01	0.7
Hindi	0	**0.98**	0	0.11
Manipuri	0.01	0	**0.96**	0.19
Background FN	0.07	0	0.03	**0**
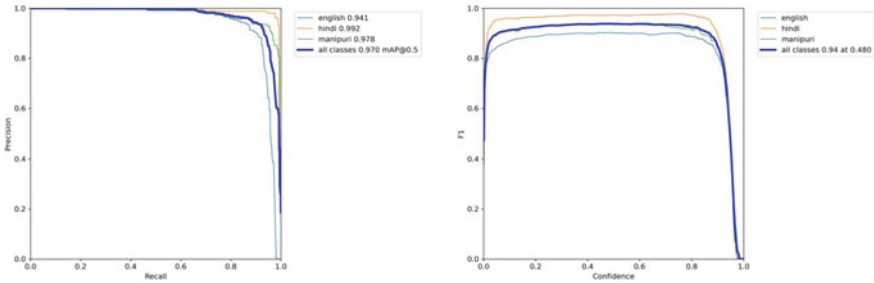
Fig. 7 Precision–recall curve on left side and $F1$ curve on right side

4.1 Confusion Matrix

In Table 1, background false positive (FP) refers to the background objects that do not belong to any of the classes but are detected as one of them. Background false negative (FN) refers to the appropriate words missed by the detector and considered as some other background objects.

4.2 Precision–Recall Curve and F1 Curve

From Fig. 7, it can be seen how well the model is trained and its promising results on unseen data.

4.3 Output

The output images are showing the detection of English and Meetei Mayek text along with a confidence score as seen in Figs. 8 and 9.

Fig. 8 Output of an image containing detection of English and Meetei Mayek text along with a confidence score

Fig. 9 Output of an image containing detection of English and Meetei Mayek text along with a confidence score

5 Conclusion and Future Work

Due to the intricate backgrounds of natural scene photographs, text identification is a significant, and difficult area of study. Text detection is a difficult task in any country, but it becomes considerably more difficult in a place like India where there are many different languages and scripts. The most frequent finding is that any image contains no more than three languages—mother tongue, Hindi, and English, forcing the adoption of script recognition as well. This paper has presented a robust Meetei Mayek, Hindi, English text detection system on natural scene images using YOLO V7 model which is well known as the most powerful object detection algorithm.

Further, the detected words will be extracted by recognizing them accurately and can be used for translating from one language to another to better assist a tourist. The recognized text will also be useful for text-to-speech translation for visually impaired people. Some other applications include licence plate detection, helping autonomous vehicles to retrieve text from sign boards.

References

1. Wang C, Bochkovskiy A, Liao HYM (2022) YOLO V7: trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. ArXiv Preprint ArXiv:2207.02696
2. Tian S, Pan Y, Huang C, Lu S, Yu K, Tan C (2015) Text flow: a unified text detection system in natural scene images. In: Proceedings of The IEEE international conference on computer vision, pp 4651–4659
3. Lundgren A, Castro D, Lima E, Bezerra B (2019) OctShuffleMLT: a compact octave based neural network for end-to-end multilingual text detection and recognition. In: 2019 international conference on document analysis and recognition workshops (ICDARW), vol 4. IEEE, pp 37–42
4. Saha S, Chakraborty N, Kundu S, Paul S, Mollah A, Basu S, Sarkar R (2020) Multi-lingual scene text detection and language identification. Pattern Recogn Lett 138:16–22
5. Ghosh S, Barman U, Bora P, Singh T, Chaudhuri B (2013) An OCR system for the Meetei Mayek script. In: 2013 fourth national conference on computer vision, pattern recognition, image processing and graphics (NCVPRIPG). IEEE, pp 1–4
6. Hijam D, Saharia S (2018) Convolutional neural network based Meitei Mayek handwritten character recognition. In: International conference on intelligent human computer interaction. Springer, pp 207–219
7. Devi CN, Devi HM, Das D (2015) Text detection from natural scene images for Manipuri Meetei Mayek script. In: 2015 IEEE international conference on computer graphics, vision and information security (CGVIS). IEEE, pp 248–251
8. Yang L, Ergu D, Cai Y, Liu F, Ma B (2022) A review of natural scene text detection methods. Procedia Comput Sci 199:1458–1465
9. Gao M, Du Y, Yang Y, Zhang J (2019) Adaptive anchor box mechanism to improve the accuracy in the object detection system. Multimed Tools Appl 78:27383–27402

Digit Recognition of Hand Gesture Images in Sign Language Using Convolution Neural Network Classification Algorithm

M. Navyasri and G. Jaya Suma

1   Introduction

A static sign language is a hand gesture images used for communicating instead of spoken words. Every person in the world wants to convey the information or share their emotion to other person. Communication plays an important role in day-to-day life; not only a person who knows the language will communicate with words but also through his body language or through his facial expression he will communicate to others. The person who does not know the same language, but if he can able to speak will communicate to others. The person who could not able to speak or hear, also communicates with others through hand gestures which is called as sign language. The person who knows the nonverbal language using hand gestures can communicate to the person who understand the sign language. One deaf and impaired person will face many daily life challenges. A translation process is required which can be used to interpret static nonverbal language which is a hand gesture images to text and then to voice conversion can fill the bridge gap of communication among deaf impaired person and normal person who does not know sign language. There are many sign languages based on country urban, rural and tribal areas. Machine translation which is used to translate sign language to text and voice, and voice to sign language is used in the field of education to teach and train special abled people. The remaining paper is structured as follows. Section 2 discusses the literature reviews and related works of sign language recognition. In Sect. 3, the proposed methodology using convolution

M. Navyasri (✉)
JNTUK, Kakinada, Andhra Pradesh, India
e-mail: navyasrimullapudi@gmail.com

KCCITM, Greater Noida, Uttar Pradesh, India

G. J. Suma
IIIT, Department of IT, JTUGV, Vizianagaram, Andhra Pradesh, India

neural network classification algorithm is illustrated followed by evaluation metrics in Sect. 4. Lastly, the conclusion and summary work are presented in Sect. 5.

2 Literature Survey

Kohlon and Singh, in the text to sign language machine translation review [1], deal with the state of the art in the advanced deep learning technologies to build the translation optimal. Farooq et al., in the sign language translation, challenging and limitations [2], provide a systematic review in all the aspects of sign language in multidisciplinary subjects. Halvardsson et al. and Rastgoo et al. [3, 4], the hand gesture recognition applying deep learning approaches for spatio temporal information using the deep learning, LSTM, SSD. And the dynamic sign language recognition using 2D convolution neural networks, 3D hand key points, SVD. Halvardsson et al. [3] use the transfer learning and convolution neural network system to provide the image perception and the mini-batch gradient algorithm during the pre-training and the accuracy model. In Sharma and Singh [5], Indian sign language recognition has created 65 different uncontrolled environments and shows the encouraging performance. In Adeyanju et al. [6], a review of hand gesture image recognition with the expert system learning methods demonstrates the remarkable success to achieve good results. In Sharma and Anand [7], Indian sign language recognition deep models perform a systematic evaluation and statistical deep models to pre-train using gradient-based optimization hyperparameters only a few ISL recognition. In Imran et al. [8], the communication gap has less understood based on the country sign and visual gesture language to communication among deaf people.

3 Proposed Methodology

The dataset covers 5000 raw image files from sign 0 to sign 9 (500 files of each sign) and 5000 corresponding output image files (applying Media Pipe) downloaded from Kaggle. This is an American Sign Language Digits Dataset, from sign 0 to sign 9. This dataset uses depth information for generating hand key points (using Media Pipe), which enriches the dataset and enhances the accuracy during classification. This is an American Sign Language Digits Dataset, using Media Pipe framework, which accurately detects the hand and 21 hand key points from a raw RGB image, and stores the co-ordinate values of these key points. The dataset contains 5000 such raw image files from sign 0 to sign 9 (500 files of each sign) and 5000 corresponding output image files (applying Media Pipe) (Fig. 1).

The graphical representation of digit images is displayed in Fig. 2, where Horizontal axis represents the label of respective digit and Vertical axis represents the no of image files.

Fig. 1 Sample digit image 0 (raw image and image with media pipe)
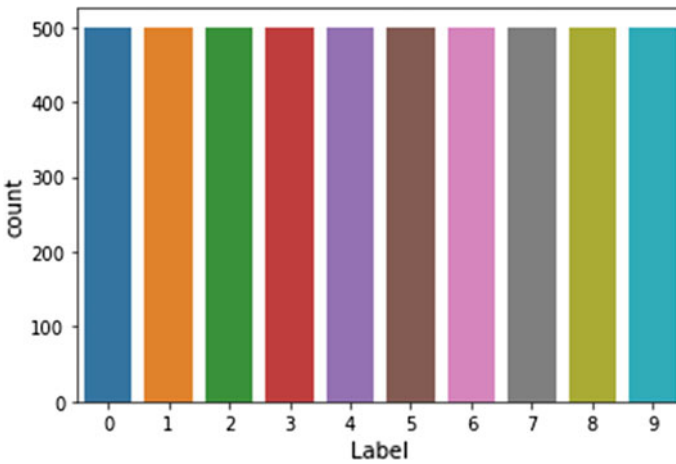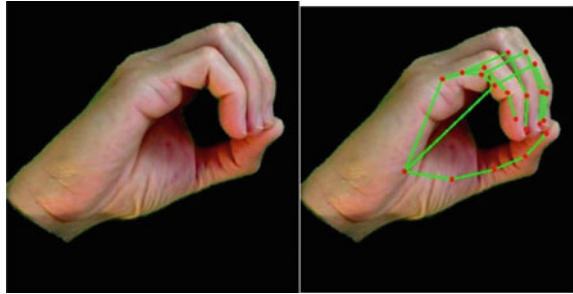
Fig. 2 Graphical representation of digit dataset

Convolution neural networks is a deep learning algorithm and an important classification algorithm that can take an input image, assign biases and weights to numerus objects in the input image [9, 10]. Compared to remaining classification algorithms, requirement of pre-processing is extensively lesser in Convolution Net. Like the pattern connectivity of neurons in the brain of human, so the architecture is analogues and inspired by the arrangement of Visual cortex in Convolution Net. The respective field is the visual field, a region where individual neurons respond (Table 1).

In the dataset, 75% of data is used for training and 25% of data used to test by the proposed model. The diagrammatic representation of convolution neural network algorithm which applied on digit dataset is displayed in Fig. 2. A convolution neural network algorithm will go through several steps for training the input data or input hand gesture images [9]. The algorithm has four layers which are rectified linear, convolution layer, pooling layer, unit layer, and fully connected layer. The first step is convolution layer. Any image is measured in the form of pixel values which is a matrix. In convolution layer, important features are extracted from an image which is in the form of matrix of pixels. They are operation of convolution phase which is

Table 1 Model summary for digit recognition using CNN algorithm

Layer	Shape of the Output	Param
CONV 2D (1ST)	(None,150,150,32)	2432
MAX POOLING 2D	(None, 75, 75, 32)	0
CONV 2D (2ND)	(None, 75, 75, 64)	18,496
MAX POOLING 2D	(None, 37, 37, 64)	0
CONV 2D (3RD)	(None, 37, 37, 96)	55,392
MAX POOLING 2D	(None, 18, 18, 96)	0
CONV 2D (3RD)	(None, 18, 18, 96)	83,040
MAX POOLING 2D	(None, 9, 9, 96)	0
FLATTEN	(None, 7776)	0
DENSE	(None, 512)	3,981,824
ACTIVATION	(None, 512)	0
DENSE	(None, 10)	5130

Model Sequential

focused on feature detectors, basically works as neural network's filters. The second phase is ReLU layer or rectified linear unit, which will reconnoitre the functions of linearity in the framework of convolution neural network. The third step in the algorithm is pooling [10] that can provide a method for sampling feature maps. The next step is Soft Max, a method of smoothen the combined feature map into a consecutive long vector and the last phase is Fully Connected layer, which is simply a feed forward network and it is an output of ending pooling.

4 Evaluation Metrics

The assessment metrics that are considered in this analysis are Accuracy, Precision, Recall and F1-score. The proposed measures play a significant role in the comparative analysis of different classification algorithms.

$$\text{Accuracy} = \frac{\text{TrPos} + \text{TrNeg}}{\text{TrPos} + \text{TrNeg} + \text{FalPos} + \text{FalNeg}}$$

$$\text{Precision} = \frac{\text{TrPos}}{\text{TrPos} + \text{FalPos}}$$

$$\text{Recall} = \frac{\text{TrPos}}{\text{TrPos} + \text{FalNeg}}$$

$$\text{F1} - \text{Score} = 2 * \frac{\text{Recall} * \text{Precision}}{\text{Recall} + \text{Precision}}$$

Table 2 Classification report for digit recognition

Digit	Pr	Re	f1-sc	Sup
0	1	1	1	146
1	1	1	1	123
2	0.98	0.99	0.99	126
3	0.99	0.98	0.99	116
4	1	0.97	0.99	118
5	1	0.98	0.99	118
6	1	0.92	0.96	123
7	0.89	0.95	0.92	141
8	0.92	0.92	0.92	106
9	0.94	0.99	0.96	136
Accuracy			0.97	1250

Where Pr = Precision, Re = Recall, f1-sc = F1-score, sup = Support

where

TrPos   True (Correctly labelled) Positive
TrNeg   True (Correctly labelled) Negative
FalPos   False (In Correctly labelled) Positive
FalNeg   False (In Correctly labelled) Negative

The experiments are done with the accuracy of 97.12 using the modules Tensor-Flow with keras in Python. The results are analysed using F1-score, Precision and Recall which are listed in the classification report of digit recognition in the Table 2. The output graph of confusion matrix is also analysed, and digit 0,1,4,5 and 6 hand gesture images are recognized correctly compared to other digits (Figs. 3, 4).

The correctly classified digits and incorrectly classified digits are displayed in Figs. 5 and 6. It is observed that the hand gesture image of digit 9 is classified as 5 and 7. And the hand gesture image of 7 is identified as 8 because of the hand positions.
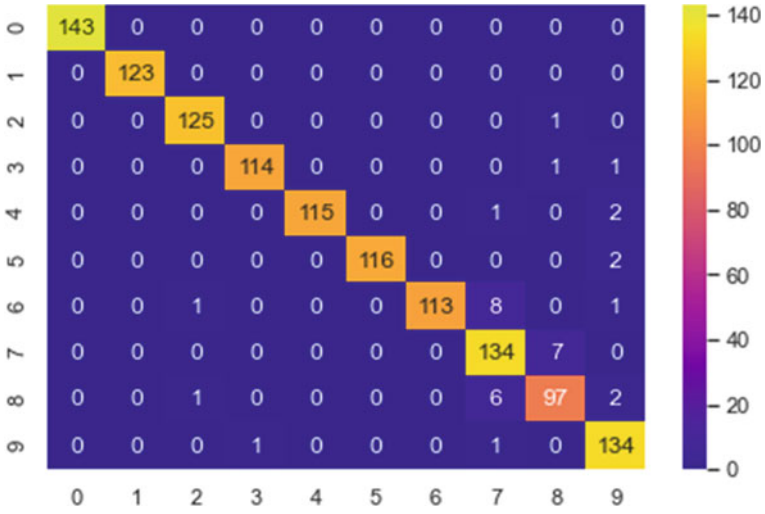
Fig. 3 Confusion matrix output graph for digit recognition
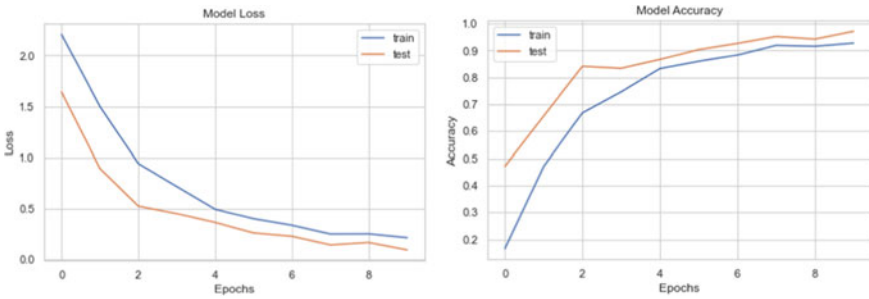
Fig. 4 Model loss and model accuracy output graph for digit recognition

Fig. 5 Correctly classified digits
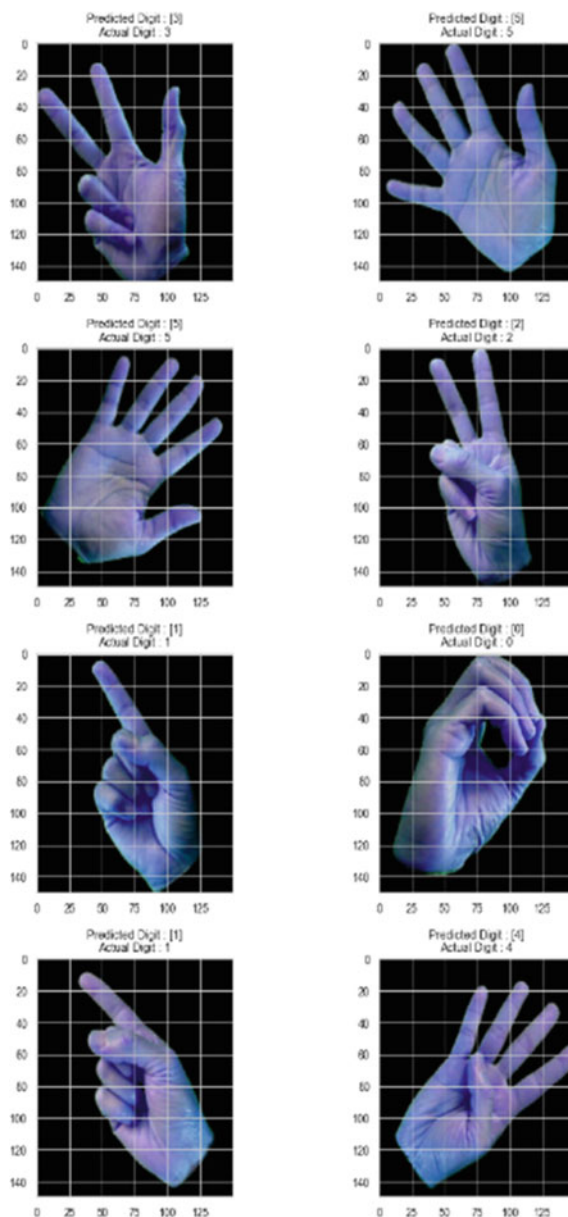
Fig. 6 Incorrectly classified digit
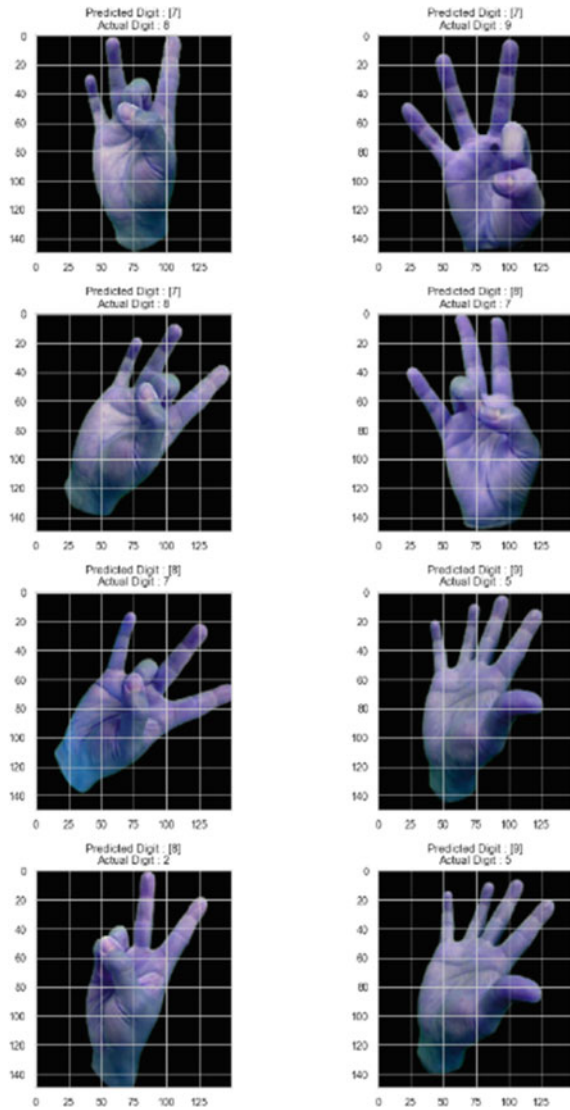
5 Conclusion

Digit recognition using hand gestures in sign language for deaf and muted people is a helpful way to fill the gap among the deaf muted person and common person. The manuscript presents the performance of convolutional neural network on digit hand gesture static images with Media Pipe. 97% of accuracy is obtained in the experiment analysis. In future, the partial voice of muted people will be added with dynamic video of sign language to recognize the sentence and emotion of that person

and the recurrent neural network with long short-term memory will be applied on hand gesture images to get more accuracy compared to convolution neural network.

References

1. Kahlon NK, Singh W (2023) Machine translation from text to sign language: a systematic review. Univ Access Inf Soc 22:1–35
2. Farooq U, Rahim MSM, Sabir N et al (2021) Advances in machine translation for sign language: approaches, limitations, and challenges. Neural Comput Appl 33:14357–14399
3. Halvardsson G, Peterson J, Soto-Valero C et al (2021) Interpretation of Swedish sign language using convolutional neural networks and transfer learning. SN Comput Sci 2(207):1–15
4. Rastgoo R, Kiani K, Escalera S (2022) Real-time isolated hand sign language recognition using deep networks and SVD. J Ambient Intell Human Comput 13:591–611
5. Sharma S, Singh S (2022) Recognition of Indian sign language (ISL) using deep learning model. Wireless Pers Commun 123:671–692
6. Adeyanju IA, Bello OO, Adegboye MA (2021) Machine learning methods for sign language recognition: a critical review and analysis. Intell Syst Appl 12:200056
7. Sharma P, Anand RS (2021) A comprehensive evaluation of deep models and optimizers for Indian sign language recognition. Graph Vis Comput 5:200032
8. Imran A, Razzaq A et al (2021) Dataset of Pakistan sign language and automatic recognition of hand configuration of Urdu alphabet through machine learning. Data Brief 36:107021
9. Yim J, Ju J, Jung H, Kim J (2015) Image classification using convolutional neural networks with multi-stage feature. In: Kim JH, Yang W, Jo J, Sincak P, Myung H (eds) Robot intelligence technology and applications 3. Advances in intelligent systems and computing, vol 345. Springer, Cham, pp 587–594
10. Suma J, Oguri S (2020) A multi-biometric iris recognition system using convolution neural network. i-manager's J Pattern Recognit 7(1):1–7

Design and Evaluation of Speech Processing Systems for Meetei/Meitei Mayek

Hoomexsun Pangsatabam, Yambem Jina Chanu, and Naorem Karline Singh

1 Introduction

Speech technologies are evolving at a rapid pace, and processing systems are becoming more intelligent than ever. Manipuri is the language of Manipur, a small state in India. Although the number of speakers is low in comparison with other languages, it has its own script and Unicode symbols to represent the Unicode alphabet [1]. As a result, various researchers are looking for ways to improve the technological aspects of the language. Speech processing is made up of various systems and models, the most important of which are speech recognition and speech synthesis. Researchers have adopted various strategies as technology has advanced over the years [2–5]. Concatenative synthesis, which stitches together small units of previously recorded waveforms, was the state-of-the-art TTS for a long time until the introduction of end-to-end models [2]. Similarly, it also revolutionizes ASR [6].

Although text is part of natural language processing, modelling a language acoustically from a linguistic standpoint is difficult. As a result, before developing a robust system, signal processing, linguistic features, and speech features must be studied [7]. Language model and acoustic model are two components of a speech system that deal with the textual and spoken parts, respectively. Furthermore, before introducing the data, it must be normalized and prepared [8].

H. Pangsatabam (✉) · Y. J. Chanu · N. K. Singh
Department of Computer Science and Engineering, National Institute of Technology Manipur, Imphal, Manipur 795004, India
e-mail: humex.pangz@gmail.com

Y. J. Chanu
e-mail: jina.yambem@gmail.com

N. K. Singh
e-mail: naoremkarline@gmail.com

Section 2 of this paper describes the literature survey used for various languages, including both highly resourced and low resourced languages. Section 3 details the modules of the speech processing architecture and provides a brief overview of the system. Section 4 discusses the various evaluation metrics used to test products or systems for comparative study, while Sect. 5 summarizes our findings.

2   Related Works

Research for Manipuri can be found in a number of areas, including [3, 9–11]. Singh et al. have emphasized the benefits of choosing Meitei Mayek script over Bengali script when it comes to phonemes and graphemes for creating the language model [12]. Patel et al. employed Bengali script when developing automatic speech transcription in Manipuri since Bengali scripts provide a large amount of textual material [3]. Meetei et al. experimented with Speech-To-Text translation, with the result being in English [11], with a focus on the low resource features of the Manipuri language. With the aid of unit selection synthesis, Patil et al. [13] created speech synthesis for Manipuri and another 12 languages. Syllables serve as the foundation for the unit selection synthesis, and work on syllables can be found in applications like the syllable labelling carried out by Nandakishor et al. [9]. According to a rule-based method, the syllabification identifies the characters as the nucleus first, followed by the onset and coda based on the letter combination [9]. However, speech technologies have significantly advanced for high resourced languages generally, and end-to-end models are increasingly favoured for a variety of applications, including speech recognition (ASR), speech synthesis (TTS), speech translation (STS), and others [6, 7, 14, 15]. Transformer and Generative Adversarial Network-based models, which are also enhanced, are used to evolve attention-based models. Such works are found in [5, 16]. A thorough study of neural voice synthesis reveals that, except from normalizations, end-to-end models are elegantly conceived, consume enormous amounts of data, and require little to no preparation. It is clear from the discussion of the problems with voice synthesis using neural models in [2] that both language models and acoustic models have problems. When wavenets were launched, Tacotron, which was created by Wang et al. and Shen et al. for voice recognition and later upgraded, became known as Tacotron 2 [7, 14]. Arik et al. [17] also created DeepVoice, a real-time neural speech synthesis system. End-to-end models also extend beyond text synthesis, as demonstrated by Sperber et al.'s work on translation [15] and Li mention of ASR [6].

3   Design of Speech Processing Systems

An acoustic model and a linguistic model are the two main parts of speech processing systems. The type of pipeline will depend on the components chosen. While TTS generates speech from the text, ASR recognizes speech as the appropriate text. Both contain a linguistic model and an audio model. MFCC is a well-liked option for acoustic features. Vocoders are only necessary in TTS applications when speech needs to be produced from input features such as speech features produced by another acoustic model or a language model as an output of some neural-based model. Typically, a seq2seq model is used for the acoustic component of TTS, which means that an input sequence is converted into an output sequence. Later, flow modelling is employed in Flow-TTS, an autoregressive model that generates spectrograms and its variants, while Tacotron uses Variational Autoencoder (VAE) [2]. The following subsections discuss the various component in design of speech processing systems.

3.1   Speech Database Recording

Professional speakers with sharp skills for speech are preferred for recording [13]. Recording is usually done in a controlled environment using a good recording device to remove all unwanted noise. The speech is sampled at 44 kHz, but these results in huge amounts of frames and later requires more time to process, so others prefer to record at 16 kHz [9, 13]. Wave file format is a lossless audio format which is preferred. Although the speech data is recorded in different sampling frequency, the higher audio files are down-sampled to provide uniform data for processing [3, 13]. However, in cases of noise reduction areas, the speech is recorded in noisy environment or noise is introduced to the noiseless audio, which is found in Speech Enhancement [18].

3.2   Text Analysis

Many speech processing approaches based on machine learning rely heavily on training and test data. Datasets with raw speech corpora or with annotations come in a variety of shapes and sizes. Annotations are also done manually or automatically. The relationships between each word are also important resource for speech processors. These relationships may be semantic, evolutionary, morphological, etc. [2]. In text data, abbreviations, numbers, and acronyms need to be normalized. Acronyms are sometime read out character by character or as words. Text normalization is accomplished by numerous techniques. It is possible to utilize an independent dictionary listing of the most popular acronyms or algorithms based on restrictions placed on the letter sequences in actual words [2, 8].

3.3 Phonetic Analysis and Language Modelling

In phonetic analysis, the Pronunciation Dictionary breaks down the phoneme sequence of each word of the language. It was designed to help with quick speech synthesis because delayed TTS is caused by a lack of vocabulary terms [13]. Grapheme to Phoneme (G2P) module is used for this task. A one-to-one mapping of the word and the sequence of phoneme are built manually from linguistic point of view. The words are written in Unicode if it exists or using a borrowed script as found in different works [3]. The phonemes are usually written in IPA while some researchers adopted their own sets of phonemes. In contrast to what is intended to be uttered, IPA depicts what is actually said. Since International Phonetic Alphabet (IPA) retains all distinguishing acoustic phonetic properties of the speech stream, it is also termed as an Acoustic Phonetic Segment (APS) [9, 10]. Meitei Mayek which is the script of Manipuri has 38 phonemes in total. The best approach is to hard code the dictionary for the lexicons by professional linguists. However, the linguistic view is often different from the technical view which we adopt in our speech processing systems and tools. For language modelling, n-gram is widely used [3].

3.4 Acoustic Modelling and Vocoders

For each speech unit in a language, such as a phone or a word, the statistics of speech attributes are modelled using an acoustic model. DNN- and HMM-based models are two examples of common acoustic models. A vocoder is used in TTS models to create audio from speech features [4]. A general basic block of ASR is also shown in Fig. 1 and TTS in Fig. 2. In vocoder analysis, the speech is analysed to obtain acoustic characteristics like F0, band aperiodicity, and mel-cepstral coefficients. These acoustic properties are used to create speech waveforms in vocoder synthesis [2].
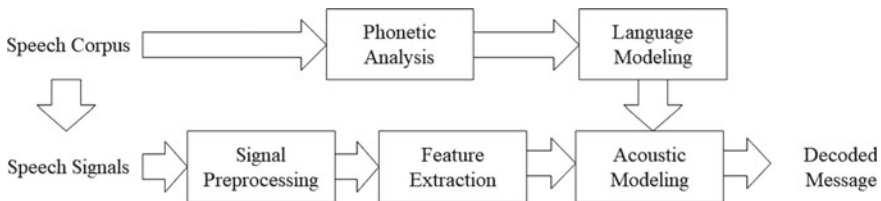
Fig. 1 Basic block diagram of ASR
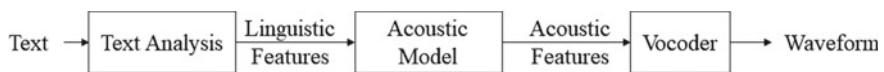
Fig. 2 Neural TTS key components

3.5 End-to-End Models

End-to-end models are a robust single model system different from the traditional multi stage models. With end-to-end systems, tedious feature engineering which highly depends on the choice of design is no longer necessary. Rich conditioning on many factors, like language or speaker, or high-level properties like sentiment, is easier to implement. However, for a certain input, such model must be able to handle significant signal level changes [6, 7, 14, 15, 19].

End-to-end speech processing models are trained on pairs of text and audio with minimal human annotation. In Tacotron and Tacotron 2 TTS, a character sequence is taken as input, and the corresponding spectrograms are generated. Tacotron can be trained from scratch with only text normalization in text preprocessing. It is also faster than the sample level autoregressive model [7]. It is also evident that E2E models surpasses hybrid models which are abundant in speech recognition. The most popular end-to-end (E2E) ASR are RNN-T and Attention-based Encoder-Decoder (AED). The encoder structure, which is the most important module, is shifting between LSTM and Transformers [6]. There are also E2E translation models [15] and Speaker Diarization models [19]. This shows that the interest is being shifted towards it.

4 Evaluation Metrics

Evaluation metrics vary according to various models. In ASR, objective evaluation is done using error rates such as WER and CER. TTS cannot be evaluated objectively and is therefore based on scores which are explained in the subsections.

4.1 Error Rate

Word error rate (WER) and character error rate (CER) are popular accuracy measurement for measuring performance of Speech Recognizer. WER recognized three kinds of error. When a speech segment/frame is not recognized, it is termed as deletion. When a word is recognized but is not in the test data, it is termed as insertion. If a word or character is recognized but as a different word from the test data, it is termed as substitution. Such errors are counted and the number of errors to the total number of words gives the word error rate [13] as in Eq. (1). WER is always used to measure the accuracy of speech recognition system. The calculation is also same for CER, but the characters are compared in this case. CER is always less than WER for Automatic Speech Recognizer (ASRs).

$$\text{WER} = \frac{\text{Deletion } + \text{ Insertion } + \text{ Substitution}}{\text{Total number of words}} \times 100 \qquad (1)$$

Table 1 Evaluation of end-to-end MOS test

System	MOS
Parametric	3.492 ± 0.096
Concatenative	4.166 ± 0.091
Tacotron (Griffin-Lim)	4.001 ± 0.087
Wavenet (Linguistic)	4.341 ± 0.051
Ground truth	4.582 ± 0.063
Tacotron 2	4.526 ± 0.066

4.2 Subjective Evaluation

Subjective evaluation measures the naturalness of speech. In speech synthesis, mean opinion score is used since objective evaluations is not enough to determine the quality of the output [7, 13, 20, 21].

Using a Likert scale of 5 points, people or subjects were asked to rate the output speech in speech synthesis systems to rate the speech's naturalness. This is called Mean Opinion Score Test or MOS tests where crowdsourcing is done for rating. Multiple output speeches are rated by native speakers and only ratings where headphones are used are included for calculating MOS [7, 13]. The 5 scale MOS evaluation by Shen et al. for their work on end-to-end Tacotron and Tacotron 2 with comparisons from [14] are given in Table 1, respectively. The models used the LJ Speech Dataset which consists of 24 h of audio and text data spoken by a professional speaker in English [22].

5 Conclusion

It may be concluded that both TTS and ASR have a variety of design options. Sets of data are essential to the system's design decision. The choice of evaluation metrics is crucial, though. In the majority of cases, a universal answer is needed during the various stages of development. End-to-end models are discovered to be more preferable because they involve a large amount of data yet require less data preparation. While all other applications have objective evaluations with error rates, TTS is evaluated by crowdsourcing. The collection of data that are appropriate for modelling is of great concern, and preparing the data is another time-consuming activity that is frequently challenging for a low resource language like Manipuri. Although E2E models are flourishing currently, it will be a challenge for low resource language. Since there are multitude of models, many factors including speed, naturalness, robustness, energy efficient, data efficient, and parameter efficient need to be thoroughly discussed.

References

1. Achom A, Basu A (2015) Design and evaluation of Unicode compliance Meitei/Meetei Mayek keyboard layout. In: 2015 International symposium on advanced computing and communication (ISACC). IEEE, pp 90–97. https://doi.org/10.1109/ISACC.2015.7377322

2. Tan X, Qin T, Soong F, Liu T-Y (2021) A survey on neural speech synthesis. arXiv:2106.15561 [eess.AS], arXiv:2106.15561v3 [eess.AS], https://doi.org/10.48550/arXiv.2106.15561

3. Patel T et al (2018) An automatic speech transcription system for Manipuri language. In: INTERSPEECH 2018, pp 2388–2389

4. Sarma K, Sarma KK (2015) Chapter 12—Acoustic modeling of speech signal using artificial neural network: a review of techniques and current trends. In: Intelligent applications for heterogeneous system modeling and design, pp 282–299. https://doi.org/10.4018/978-1-4666-8493-5.ch012

5. Li N, Liu S, Liu Y, Zhao S, Liu M (2019) Neural speech synthesis with transformer network. In: Proceedings of the AAAI conference on artificial intelligence, vol 33, issue 1, pp 6706–6713

6. Li J (2021) Recent advances in end-to-end automatic speech recognition. arXiv:2111.01690 [eess.AS], arXiv:2111.01690v2 [eess.AS], https://doi.org/10.48550/arXiv.2111.01690

7. Wang Y, Skerry-Ryan RJ, Stanton D, Wu Y, Weiss RJ, Jaitly N, Yang Z, Xiao Y, Chen Z, Bengio S, Le Q, Agiomyrgiannakis Y, Clark R, Saurous RA (2017) Tacotron: towards end-to-end speech synthesis. arXiv:1703.10135 [cs.CL], arXiv:1703.10135v2 [cs.CL], https://doi.org/10.48550/arXiv.1703.10135

8. Macchi M (1998) Issues in text-to-speech synthesis. In: Proceedings of the IEEE international joint symposia on intelligence and systems (Cat. No.98EX174). IEEE, pp 318–325. https://doi.org/10.1109/IJSIS.1998.685467

9. Nandakishor S, Dutta SK, Singh LJ (2015) An HMM based semi-automatic syllable labeling system for Manipuri language. In: International Conference on computing, communication and automation. IEEE, pp 1044–1047. https://doi.org/10.1109/CCAA.2015.7148548

10. Bhaskararao P (2011) Salient phonetic features of Indian languages in speech technology. Sadhana 36:587–599

11. Meetei LS et al (2021) An experiment on speech-to-text translation systems for Manipuri to English on low resource setting. In: Proceedings of the 18th international conference on natural language processing (ICON), pp 54–63. National Institute of Technology Silchar, Silchar, India. NLP Association of India (NLPAI)

12. Singh LS, Thaoroijam K, Das PK (2007) Written Manipuri (meiteiron)—from phoneme to grapheme. Lang India 7(6):2–22

13. Patil HA et al (2013) A syllable-based framework for unit selection synthesis in 13 Indian languages. In: 2013 International conference oriental COCOSDA held jointly with 2013 conference on Asian spoken language research and evaluation (O-COCOSDA/CASLRE). IEEE, pp 1–8. https://doi.org/10.1109/ICSDA.2013.6709851

14. Shen J, Pang R, Weiss RJ, Schuster M, Jaitly N, Yang Z, Chen Z, Zhang Y, Wang Y, Skerrv-Ryan R, Saurous RA, Agiomvrgiannakis Y, Wu Y (2017) Natural TTS synthesis by conditioning wavenet on MEL spectrogram predictions. In: 2018 IEEE International conference on acoustics, speech and signal processing (ICASSP). IEEE, pp 4779–4783. https://doi.org/10.1109/ICASSP.2018.8461368

15. Sperber M et al (2019) Attention-passing models for robust and data-efficient end-to-end speech translation. Trans Assoc Comput Linguist 7:313–325

16. Wei K, Guo P, Jiang N (2022) Improving transformer-based conversational ASR by inter-sentential attention mechanism. arXiv:2207.00883 [cs.SD], arXiv:2207.00883v1 [cs.SD], https://doi.org/10.48550/arXiv.2207.00883

17. Arik SÖ, Chrzanowski M, Coates A, Diamos G, Gibiansky A, Kang Y, Li X, Miller J, Ng A, Raiman J, Sengupta S, Shoeybi M (2017) Deep voice: real-time neural text-to-speech. In: Proceedings of the 34th international conference on machine learning, PMLR, vol 70, pp 195–204

18. Valentini-Botinhao C, Wang X, Takaki S, Yamagishi J (2016) Speech enhancement for a noise-robust text-to-speech synthesis system using deep recurrent neural networks. In: INTERSPEECH 2016, pp 352–356. http://dx.doi.org/10.21437/Interspeech.2016-159
19. Fujita Y, Kanda N, Horiguchi S, Nagamatsu K, Watanabe S (2019) End-to-end neural speaker diarization with permutation-free objectives. arXiv:1909.05952 [eess.AS], arXiv:1909.05952v1 [eess.AS], https://doi.org/10.48550/arXiv.1909.05952
20. Johnston RD (1996) Beyond intelligibility: the performance of text-to-speech synthesisers. BT Technol J 14(1):100–111
21. Salza PL, Foti E, Nebbia L, Oreglia M (1996) MOS and pair comparison combined methods for quality evaluation of text-to-speech systems. Acta Acustica Unit Acustica 82(4):650–656
22. Ito K, Johnson L (2017) The LJ speech dataset. Retrieved from https://keihito.com/LJ-Speech-Dataset/. Accessed on 12 Nov 2022

Congestive Heart Failure Prediction Using Artificial Intelligence

M. Sheetal Singh, Khelchandra Thongam, and Prakash Choudhary

1 Introduction

Congestive heart failure, or heart failure, is a serious medical condition when the heart muscle is unable to pump enough blood to fulfill our body's needs. As the heart weakens over time, the muscle wall becomes less able to contract, which limits the ability to fill blood [1]. Due to less contraction of the heart, two problem arises: (i) Fluid will build up inside the heart and (ii) not enough blood flow in our organs, which can cause serious problem. Certain heat state, such as high blood pressure or narrowed arteries, leads to the heart failing to efficiently fill and pump enough blood around the body [2]. The body tries to make up for the condition either by beating faster or enlarging itself, which causes breathing problems, and the kidney also begins to retain sodium and water. With time, the condition gets worse. Heart disease is one of the most common reasons for death worldwide, and according to WHO and the target patient are mostly over age 65 [3]. However, with proper treatment, the sign and symptoms of heart failure can be improved, and one may recover to good health. Some changes in lifestyle, such as a low sodium diet, exercising, and managing stress, can improve the risk of CHF. CHF is a complex heart disease, and if not handled carefully, it may cause premature death [4]. With the help of data mining technologies and medical science, we can predict CHF. In this work, we have studied various machine learning (ML) techniques which predict heart failure using different

M. S. Singh (✉) · K. Thongam
National Institute of Technology Manipur, Imphal West, Manipur, India
e-mail: msheetalsingh@live.com

K. Thongam
e-mail: thongam@gmail.com

P. Choudhary
Central University of Rajasthan, Ajmer, Rajasthan, India
e-mail: choudharyprakash87@gmail.com

355

types of knowledge abstraction methods. We use the CHS dataset in our work which contains the record of 5888 patients with 13 different cardiovascular disease classes. In pre-processing, we impute missing data using PMM and select the best feature using the information gain ratio for our classifiers. Mainly in this study, we compare five different ML algorithms (K-nearest neighbor (KNN), Support Vector Machine (SVM), Random Forest (RF), Artificial Neural Network (ANN), and decision tree (DT)) to predict CHF. After comparing the performance of all the classifiers, ANN gives optimal results than other classifiers.

The structure of the paper is divided into the following sections: Related works, techniques, and existing methods are briefly discussed in Sect. 2. Section 3 discusses the dataset, methods used in pre-processing, and different classification algorithms used in this work. Section 4 discusses the comparative results obtained from all the classification algorithms. Finally, Sect. 5 concludes the current work by discussing the future.

2 Related Work

Gjoreski et al. [5] used a set of classification algorithms to detect CHF from the sound of the heart, which is collected from 122 subjects using an electronic stethoscope. Of 122 subjects, 23 people were diagnosed CHF, and 99 were healthy. A combination of five machine learning techniques (J48, Naïve Bayes, KNN, SVM, and Boosting) achieved the optimal performance in the segment-based machine learning phase and RF as meta-learner. The experimental method achieves an accuracy of 96% using leave-one-subject-out (LOSO) cross-validation.

Sharanyaa et al. [6] predict heart disease by developing a hybrid model by combining k-nearest neighbor and fuzzy logic characteristics in a public dataset from the UCI repository. The dataset consists of 13 features and compared four classifiers to compare the performance. The hybrid algorithm obtained the highest accuracy of 94%.

Ishaq et al. [7] suggested a supervised machine learning technique to predict heart failure patient survival using data mining techniques. The experiment uses heart failure clinical records dataset from the UCI repository, which has 13 features and 299 medical information. The synthetic minority oversampling technique (SMOTE) is used to address the dataset's imbalance class problem. High-ranked features for the experiment are obtained using the Random Forest algorithm. Furthermore, the extra tree classifier achieves 92.62% accuracy, which outperformed the other eight classifiers.

Plati et al. [8] developed a machine learning-based method to diagnose heart failure, which uses a rotation forest and logistic model tree classifier for classification. They collected more than 400 subjects each from two different healthcare universities. After classification, their approach achieved 84.12 and 91.23% accuracy using a clinical feature and all combined features, respectively.

Pecchia et al. [9] proposed a data mining technique for an effective remote health monitoring platform for early heart failure detection based on classification and the CART method. Their method achieved an accuracy of 96.39 and 100.00% precision in detecting heart failure.

Aljaaf et al. [10] presented a predictive algorithm that uses a decision tree classifier to predict five categories of heart failure risk. UCI repository heart disease dataset, which has 297 instances with 13 attributes, is employed by the model. After classification, they achieved 86.53 accuracy, 95.5% specificity, and 86.5% sensitivity which overcomes other models in their study.

Melillo et al. [13] proposed an automatic CHF patient risk assessment classifier which uses standard long-term heart rate variability measures to separate high and low risk patients. They use 41 records with 30 high risk and 11 low risk patients. For classification in this study, classification and regression tree (CART) is used and achieved 63.60% specificity and 93.30% sensitivity.

Gjoreski et al. [11] developed a technique based on heart sounds to identify CHF. They integrate machine learning with the deep learning method and evaluate on preprocessed PhysioNet dataset (947 records), which is a publically available online. The model had an accuracy of 93.2%, which is higher than baseline methods.

Sandhya [14] used a SVM to predict heart disease and archived 85% Accuracy.

Singh et al. [12] exploited the Cleveland heart disease dataset and applied the Random Forest classifier and achieved an accuracy of 85.81%. More than three hundred data instances and 13 attributes are trained using tenfold cross-validation.

A comprehensive review of the literature shows that the heart failure prediction capabilities of existing research are effective across a variety of datasets. Various optimization techniques were applied to boost the precision and accuracy of a number of measurements. Table 1 shows the comparison of existing work. The goal of our work is to highlight the best suitable machine learning technique for CHF prediction after comparing different machine learning techniques.

3 Methodology

In this work, we use the CHS dataset, which can be downloaded online [15]. The dataset contains the total record of 5888 patients diagnosed with 12 different types of cardiovascular diseases and more than 400 features (clinical, demographic, and personal). The dataset is inconsistent and noisy, with more than 30% missing inputs. So pre-processing is an important stage in our work where the dataset is cleaned and imputed, and the best feature is selected from the raw dataset. The selected feature transforms the raw dataset into a meaningful set of data which is further used for classification in different classification methods. Figure 1 shows the proposed methodology.

Table 1 Comparison of existing works

Study	Classifier	Dataset	Measures
Ishaq et al. [7]	Extra tree classifier (ETC)	Heart failure clinical records dataset 299 patients, 13 features	92.62% Accuracy
Plati et al. [8]	Rotation forest (RO) and logistic model tree (LMT)	UCD and DCUH 422 subjects, 19 features	91.23% Accuracy 93.83% Sensitivity 89.62% Specificity
Sharanyaa et al. [6]	K-nearest neighbor	UCI repository 303 records, 13 features	94.00% Accuracy
Gjoreski et al. [11]	Fully connected neural network (FCNN) and CNN	Phonocardiogram (PCG) recording 161 patient, 32 features	93.20% Accuracy
Singh et al. [12]	Random forest	Cleveland heart disease dataset 303 records, 13 features	85.81% Accuracy
Gjoreski et al. [5]	Leave-one-subject-out (LOSO) evaluation	Phonocardiogram (PCG) data 122 subjects, 21 features	96.00% Accuracy
Aljaaf et al. [10]	C4.5 decision tree classifier	Cleveland heart disease dataset 297 patient records, 13 features	86.5% Sensitivity 95.5% Specificity 86.53% Accuracy
Melillo et al. [13]	Classification and regression tree (CART)	41 patients with 13 features	78.5% Accuracy 93.3% Sensitivity 63.6% Specificity

3.1 Dataset

In our study, we use the CHS dataset, which is a population-based longitudinal study of individuals 65 and older with coronary heart disease and stroke. It's accessible online at [12]. The dataset consists of 5888 patient record with more than 400 attributes, and more than 60% of the information were not related to our study or left empty.

3.2 Data Pre-processing

To make the prediction, data pre-processing is one of the principal milestones in machine learning techniques. The raw data is often created and stored by humans; as a result, it may contain missing values, input errors, or duplicate data. And to train machine learning algorithms, the input data needs to be transformed into an efficient and useful format to provide a precise, reliable, and robust outcome. In our study, data pre-processing is divided into the following three stages:

Fig. 1 Experiment workflow
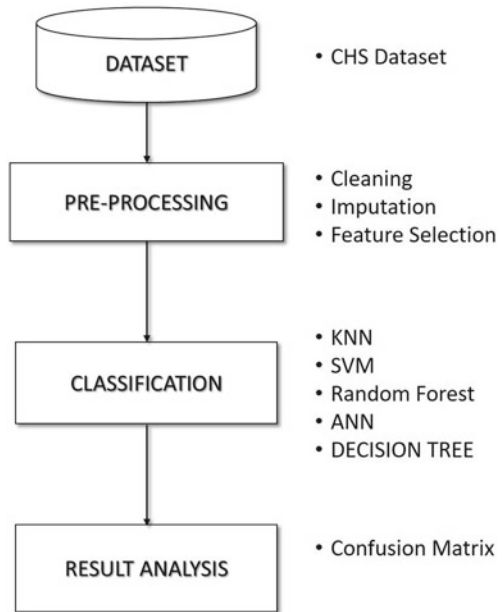
1. Data cleaning
2. Missing data imputation
3. Feature selection.

Data Cleaning CHS dataset is a very complicated dataset that is very challenging to use directly. In this section, we remove all the columns with empty values of more than 25 percent and all the records not related to Congestive heart failure. Further, we remove all the inconsistent, noisy, and duplicate records.

Missing Data Imputation In this stage, we impute the missing value from the known values in the dataset using an iterative method called Predictive Mean Matching (PMM) imputation available at the MICE package in R. It employs a Bayesian linear regression to determine the similarities between each data value and all full cases and missing values in the dataset. Then for each missing value, the most similar complete case donor candidate is selected [16, 17].

Feature Selection For selecting the best feature for classification, we calculate the gain ratio of all attributes using C4.5 algorithm [18]. After calculating the gain ratio of all the attributes, we select 14 attributes and remove other attributes with a gain ratio of zero or close to zero. And the selected attributes form a subset with low intercorrelation but with high correlation with the class [19]. Finally, the new dataset consists of 1080 patient records (500 diagnosed with CHF and 580 with no CHF) and 14 features. Table 2 shows the selected features, and Fig. 2 shows the gain ratio of the selected features.

Table 2 Description of selected features

Sl. No	Attributes	Description
1	AGE	Age of subject
2	GENDER	Gender
3	WEIGHT	Current weight
4	CHSTPN	Ever had pain in your chest
5	HIBP	High blood pressure
6	DIABETES	Calc. diabetes status
7	SMOKE	Smoking status
8	EXINTEN	Exercise intensity
9	OVRWT	Obesity, 120%
10	HRATE	Heart rate
11	ALCOH	Weekly alcohol consumption (drinks)
12	HIBSUG	High blood sugar status
13	BMI	Body mass index
14	HYPER	Calc. hypertension status
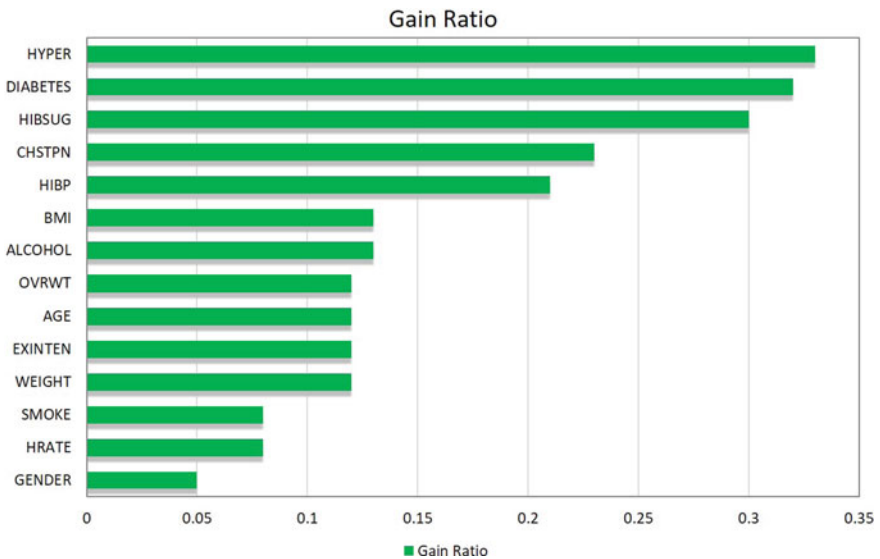
Fig. 2 Gain ratio of selected features

3.3 Classification Modeling

The classification is done using the pre-processed new dataset to evaluate the performance of each classifiers. And the best performing model is identified according to

evaluation matrices (such as accuracy, F1-score, sensitivity, specificity, and precision) to evaluate the effectiveness of each classifier.

K-Nearest Neighbor KNN is a simple and easy-to-understand supervised learning algorithm. It calculates the Euclidian distance of k neighbor to identify the closest neighbor to determine the category of an unknown sample [20, 21].

Support Vector Machine SVM is a supervised machine learning technique used primarily for classification and regression. SVM finds the hyper-plane, which is also known as an optimal solution. To find the hyper-plane following mathematical computation is done [21, 22].

Random Forest Random Forest is a bagging technique that classifies using multiple decision trees to find an optimal result. Every decision tree indicates the class of the object [14, 23].

Artificial Neural Network An ANN is a machine learning algorithm based on a neural network found in the human brain, also known as a multi-level perceptron, and each perceptron has its weight, which is used to compute the output. It has three layers: an input layer with neurons where input data is applied, a hidden layer where value is predictable, and an output layer where the output is hold [24, 25].

Decision Tree A DT is a simple classification algorithm based on the divide and conquer technique and constructed in a top-down approach. The structure of the classifier is like a tree, where there is an internal node, branch, and leaf node, which represent features, decision rules, and outcomes of a dataset, respectively [21].

3.4 Evaluation Matrices

In this study, the performance for the difference in classifiers is examined in terms of sensitivity, specificity, accuracy, F1-score, and precision with the help of confusion matrix [26]. Figure 3 represents the confusion matrix and the measures of performance are as follows:

$$\text{Accuracy} = \frac{\text{TN} + \text{TP}}{\text{TN} + \text{FN} + \text{TP} + \text{FP}}$$

$$\text{F1 Score} = \frac{2\text{TP}}{2\text{TP} + \text{FP} + \text{FN}}$$

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}}$$

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

Fig. 3 Representation of confusion matrix

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

where

TP predicted true and the actual class is also true.
FP predicted true but the actual class is false.
TN predicted false and the actual class is also false.
FN predicted false but the actual class is true.

4 Evaluation Results

The dataset we used for classification had 1080 instances with 14 features, and we considered a tenfold cross-validation technique to compare the results. The confusion matrix of all methods used for classification is shown in Fig. 4. The result comparison is given in Table 3. The result compares the F1-score, sensitivity, specificity, precision, and accuracy of all algorithm used for classification. After comparison, it was observed that the ANN classification technique achieved the highest accuracy than the other models, as shown in Fig. 5.

5 Conclusion and Future Work

A comprehensive model for the prediction of CHF is presented in this paper. The CHS dataset used for the training purpose has some missing values and is noisy. Hence, the proposed model has noise removal and missing data imputation capabilities. The trained model is able to predict CHF accurately based on the patient's medical health history. In pre-processing, the model imputes missing data and selecting the best features for classifiers. The model used the Predictive Mean Matching (PMM) method for missing data imputation and DT C4.5 algorithm for feature selection. Mainly, in this paper, we compared five different machine learning algorithms to

91	9
24	100

SVM

95	15
18	96

DT

112	10
20	82

KNN

124	9
10	81

RF

119	8
6	87

ANN
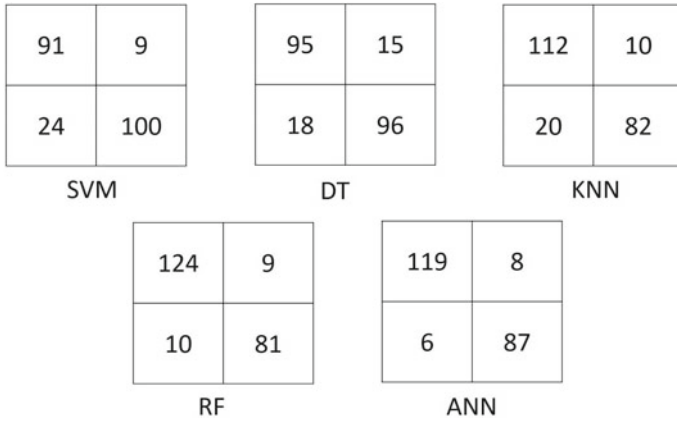
Fig. 4 Confusion matrix of all classifier

Table 3 Result comparison tables

Classifier	F1-score	Sensitivity	Specificity	Precision	Accuracy
RF	0.9288	0.9254	0.9000	0.9323	0.9152
DT	0.8520	0.8407	0.8649	0.8636	0.8527
KNN	0.8819	0.8485	0.8913	0.9180	0.8661
SVM	0.8465	0.7913	0.9174	0.9100	0.8521
ANN	**0.9444**	**0.9520**	**0.9158**	**0.9370**	**0.9364**

The bold values indicate the effectiveness of the proposed model
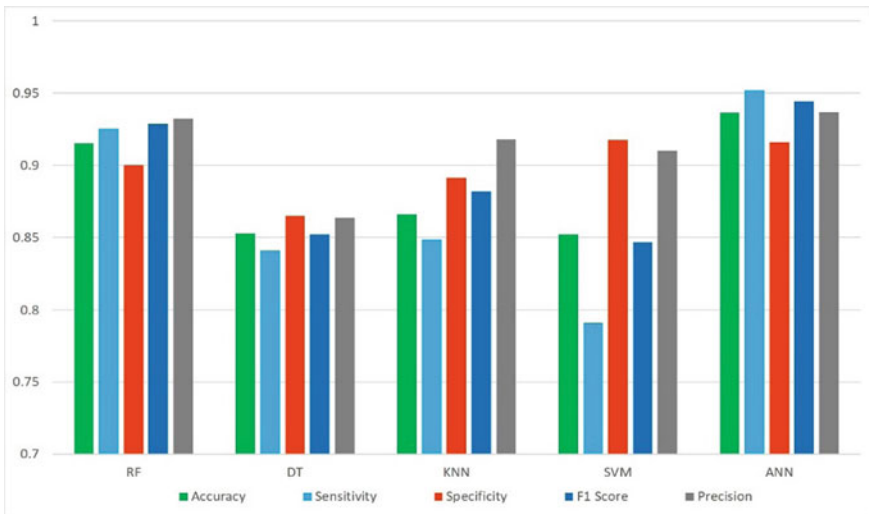
Fig. 5 Comparison graph

predict CHF. After comparing the performance of all the classifiers, the ANN classifier obtained the optimal result. There are a large number of life threatening diseases, such as stroke, Angioplasty, Claudication, Angina, and many more which can be predicted by exploiting the patient's medical history. In the future, we will try to apply the proposed model to predict other diseases, which can be predicted by analyzing the patient's medical history.

References

1. Braunwald E, Zipes DP, Libby P, Bonow RO (2004) Heart disease: a textbook of cardiovascular medicine, 7th edn. Elsevier Science, Philadelphia, PA, USA
2. European Society of Cardiology (2014) Heart failure: preventing disease and death worldwide. Retrieved from https://www.escardio.org/static-file/Escardio/Subspecialty/HFA/WHFA-whitepaper-15-May-14.pdf. Accessed on 5 Mar 2015
3. World Health Organization (2014) Global status report on non-communicable diseases 2014
4. Medical professionals: cardiovascular diseases and cardiac surgery. In: Mayo clinic. Retrieved from https://www.mayoclinic.org/medical-professionals/cardiovascular%20diseases. Accessed on 29 Dec 2020
5. Gjoreski M, Simjanoska M, Gradišek A, Peterlin A, Gams M, Poglajen G (2017) Chronic heart failure detection from heart sounds using a stack of machine-learning classifiers. In: 2017 International conference on intelligent environments (IE). IEEE, pp 14–19. https://doi.org/10.1109/IE.2017.19
6. Sharanyaa S, Lavanya S, Chandhini MR, Bharathi R, Madhulekha K (2020) Hybrid machine learning techniques for heart disease prediction. Int J Adv Eng Res Sci 7(3):44–48. https://doi.org/10.22161/ijaers.73.7
7. Ishaq A, Sadiq S, Umer M, Ullah S, Mirjalili S, Rupapara V, Nappi M (2021) Improving the prediction of heart failure patients' survival using SMOTE and effective data mining techniques. IEEE Access 9:39707–39716. https://doi.org/10.1109/ACCESS.2021.3064084
8. Plati DK et al (2021) A machine learning approach for chronic heart failure diagnosis. Diagnostics 11(10):1863. https://doi.org/10.3390/diagnostics11101863
9. Pecchia L, Melillo P, Bracale M (2011) Remote health monitoring of heart failure with data mining via CART method on HRV features. IEEE Trans Biomed Eng 58(3):800–804. https://doi.org/10.1109/TBME.2010.2092776
10. Aljaaf AJ, Al-Jumeily D, Hussain AJ, Dawson T, Fergus P, Al-Jumaily M (2015) Predicting the likelihood of heart failure with a multi level risk assessment using decision tree. In: 2015 Third international conference on technological advances in electrical, electronics and computer engineering (TAEECE). IEEE, pp 101–106. https://doi.org/10.1109/TAEECE.2015.7113608
11. Gjoreski M, Gradišek A, Budna B, Gams M, Poglajen G (2020) Machine learning and end-to-end deep learning for the detection of chronic heart failure from heart sounds. IEEE Access 8:20313–20324. https://doi.org/10.1109/ACCESS.2020.2968900
12. Singh YK, Sinha N, Singh SK (2017) Heart disease prediction system using random forest. In: Advances in computing and data sciences. ICACDS 2016. Communications in computer and information science, vol 721. Springer, Singapore, pp 613–623. https://doi.org/10.1007/978-981-10-5427-3_63
13. Melillo P, De Luca N, Bracale M, Pecchia L (2013) Classification tree for risk assessment in patients suffering from congestive heart failure via long-term heart rate variability. IEEE J Biomed Health Inform 17(3):727–733. https://doi.org/10.1109/JBHI.2013.2244902
14. Sandhya Y (2020) Prediction of heart diseases using support vector machine. Int J Res Appl Sci Eng Technol (IJRASET) 8(II):126–135. ISSN: 2321–9653; IC Value: 45.98; SJ Impact Factor: 7.177

15. Cardiovascular Health Study (CHS). In: National heart, lung and blood institute. Retrieved from https://biolincc.nhlbi.nih.gov/studies/chs/. Accessed on 1 Apr 2020
16. Bertsimas D, Pawlowski C, Zhuo YD (2018) From predictive methods to missing data imputation: an optimization approach. J Mach Learn Res 18:1–39
17. Bräm DS, Nahum U, Atkinson A, Koch G, Pfister M (2022) Evaluation of machine learning methods for covariate data imputation in pharmacometrics. CPT Pharmacometrics Syst Pharmacol 11(12):1638–1648
18. Lestari A, Alamsyah (2020) Increasing accuracy of C4.5 algorithm using information gain ratio and adaboost for classification of chronic kidney disease. J Soft Comput Explor 1(1):32–38
19. Singh MS, Choudhary P (2017) Stroke prediction using artificial intelligence. In: 2017 8th Annual industrial automation and electromechanical engineering conference (IEMECON). https://doi.org/10.1109/IEMECON.2017.8079581
20. Kramer O (2013) K-nearest neighbors. In: Dimensionality reduction with unsupervised nearest neighbors. Intelligent systems reference library, vol 51. Springer, Berlin, Heidelberg, pp 13–23. https://doi.org/10.1007/978-3-642-38652-7_2
21. Mohan S, Thirumalai C, Srivastava G (2019) Effective heart disease prediction using hybrid machine learning techniques. IEEE Access 7:81542–81554. https://doi.org/10.1109/ACCESS.2019.2923707
22. Hearst MA, Dumais ST, Osuna E, Platt J, Scholkopf B (1998) Support vector machines. IEEE Intell Syst Appl 13(4):18–28
23. Pal M, Parija S (2021) Prediction of heart diseases using random forest. J Phys Conf Ser 1817(012009):1–8
24. Awan SM, Riaz MU, Khan AG (2018) Prediction of heart disease using artificial neural network. VFAST Trans Softw Eng 6(1):51–61
25. Yang GR, Wang X-J (2020) Artificial neural networks for neuroscientists: a primer. Neuron 107(4):1048–1070
26. Hay AM (1988) The derivation of global estimates from a confusion matrix. Int J Remote Sens 9(8):1395–1398

Intelligent Speaker Identification System Under Multi-Variability Speech Conditions

Banala Saritha, Tungala Thiru Venkata Naga Manoj, Sachin Kumar Sharma, Rabul Hussain Laskar, Madhuchhanda Choudhury, and K. Anish Monsley

1 Introduction

Speech is a robust mode of communication that transmits rich and essential information such as gender, accent, emotion, and other distinctive traits of a speaker. These unique features help identify the speaker and find the applications across biometrics, forensics, and access control systems [1]. Finding an unknown speaker from a group of well-known speakers involves extensive research in speaker identification [2]. Researchers' attention has recently been to recent developments in deep learning techniques [3]. Speech is more unreliable than other forms of personal identification like the iris, face, and fingerprint. Age, gender, emotion, language, and speaking style concerns can affect speech characteristics. The speech samples used for training and testing are one of the main variables limiting speaker identification accuracy in addition to these inconsistencies. The downsides of using human speech for identification include variability within a speaker and mismatch in testing and training situations [4]. A speaker's voice can change over time and may not always sound the same, referred to as intra-speaker variability. A mismatch occurs when speech acquires from the same person in diverse contexts. The extensive usage of several sensing devices (such as a microphone on mobile phones) and contextual factors may change the voice quality of the original speech. Also, multi-language and conversational styles affect the performance of speaker identification. To work around these shortcomings, we used the multi-variability (IITG-MV) speaker identification database [5]. We also introduced a novel structure for an end-to-end intelligent speaker identification system relying on a VGG-like network. The rest of the paper is organized as

B. Saritha (✉) · T. T. V. N. Manoj · S. K. Sharma · R. H. Laskar · M. Choudhury · K. A. Monsley
Department of Electronics and Communication Engineering, National Institute of Technology Silchar, Silchar, Assam 788010, India
e-mail: banalasarita@gmail.com

follows. The related works are described in Sect. 2. The analysis and comparison of the experimental work with the acquired results are presented in Sect. 4.

2 Related Works

Due to advances in computing power and the emergence of massive databases, deep learning techniques for speaker identification have garnered much attention from researchers. Acoustic models based on Deep Neural Networks (DNN) have performed much better than those based on Gaussian mixture models over the past few years (GMM). The authors of [6] compared speaker identification algorithms that use DTW, GMM, and SVM. The authors [7] used a variety of normalizing strategies to overcome the discrepancy between the training and testing settings. The researchers employed feature mapping approaches in [8]. The authors introduce speaker model synthesis [9] in mismatched conditions. Additionally, factor analysis by authors [10] and nuisance attribute projection (NAP) [11] approaches are proposed. In their research, the effectiveness of short time cepstral characteristics generated by filter banks with Bark and ERB rate warping is examined and the robustness of speaker identification in mismatch situations. Chougule and Chavan [12] used Warped Filter Bank Features to accomplish the speaker identification task in mismatch contexts. They used databases with mismatched sensors for this. In [13], Haris et al. created the IITG multi-variability speaker recognition database in the context of India. These authors used adapted Gaussian mixture modeling to assess the speaker identification and verification tasks. We also compared the findings of this research using the IITG-MV database. Moreover, we presented a deep CNN architecture for an end-to-end speaker identification system built on a VGG-like network.

3 Methodology

This study investigates the speaker identification task in an uncontrolled setting using input speech collected from various sensor devices. Phase III of the IITG multi-variability speaker recognition database uses to evaluate our work. The Indian Institute of Technology Guwahati prepared the IITG-MV speech database. In the context of India, where speech is highly varied and diverse, the database promotes the reliability of speaker identification. Here, the third phase of the database is described briefly.

3.1 Multi-Variability Speech Data Base

IITG-MV Phase III Database In the third recording phase, a telephone network was utilized, keeping in mind the possibility of remote person authentication in speech mode. At the same time, the Phase I dataset is used in the sensor mismatch scenario. Phase II database includes a multi-environment like laboratories, student rooms, corridors, etc. In contrast to Phases I and II, a facilitator connects a conference call between two people in Phase III during their leisure time. The Phase III database contains four of the variabilities that are described below.

Multi-Environment Conversations were captured in various real-world settings, including coffee shops, offices, living spaces, laboratories, etc.

Multi-Sensor A sample frequency of 8 kHz is used to record speech data from several mobile phones.

Multi-Lingual Every speaker used their mother tongue or English (favorite language).

Conversation Style Over a conference call, each speaker spoke conversationally time.

The recording process is depicted in Fig. 1. The individuals were asked to provide a meeting time or a specified time slot during which they and their friends would be available to speak. The facilitator then used his mobile handset to call the individual at the appointed time. In response, the individual provided the contact number of the person he wanted to speak with. After putting the individual on hold, the facilitator called the other party. The facilitator then connected the two people on a conference call. There were no restrictions on the location or language of the recording. The entire time, subjects were discussed openly, with frequent language and subject matter shifts.

3.2 Proposed Architecture

This section presents a novel architecture for end-to-end intelligent speaker identification tasks built on a network similar to a VGG. The exceptional performance of the VGG architecture in speaker identification and computer vision is the inspiration for the proposed framework. In this study, we modified the layers and filters to redesign the fundamental structure of VGG-13. Figure 2 depicts the proposed architecture.

There are ten convolutional layers in this model. Each layer utilizes a rectified linear unit (ReLU) activation function and is a 2D convolutional layer with a kernel size of 3×3. After each convolutional layer, a batch normalizing layer is added on top. At the same time, the average pooling layer of stride 2×2 is added after every few convolutional layers. After a global average pooling (GAP) layer, a fully

Fig. 1 Phase III data collecting scenario for a conference call [13]

connected layer with 4096 neurons is implanted. A dropout of 0.5 is positioned on either side of the FC layer to prevent generalization issues. The softmax classification layer is eventually connected to the output convolutional layer. The 100 nodes of the output convolutional layer represent the number of speakers that can be identified using the softmax loss function.

4 Experimental Setup and Results

An intelligent end-to-end speaker identification is the main objective of this research. 20–30 ms of conversational style 1 speech data are used for training, while 8–15 ms of conversational style 2 are used for testing. Log-Mel spectrogram is generated from each raw audio sample as shown in Fig. 3. The Librosa library is used in Python to carry out this operation for all voice recordings. These spectrograms are scaled to 224×224 before being fed as input into the proposed architecture. Additionally, the presented model is trained on the Kaggle GPU using the TensorFlow Python library with a 256-batch size. The Adam optimizer was used for training, with the following learning rates: $lr = 0.001, \alpha = 0.95, \epsilon = 10.5$. The speaker identification accuracy of the proposed model is measured as a%age, which is the ratio of correctly predicted classes of test samples to the total positive prediction classes of the test samples.

Fig. 2 Architecture of the proposed end-to-end intelligent speaker identification system
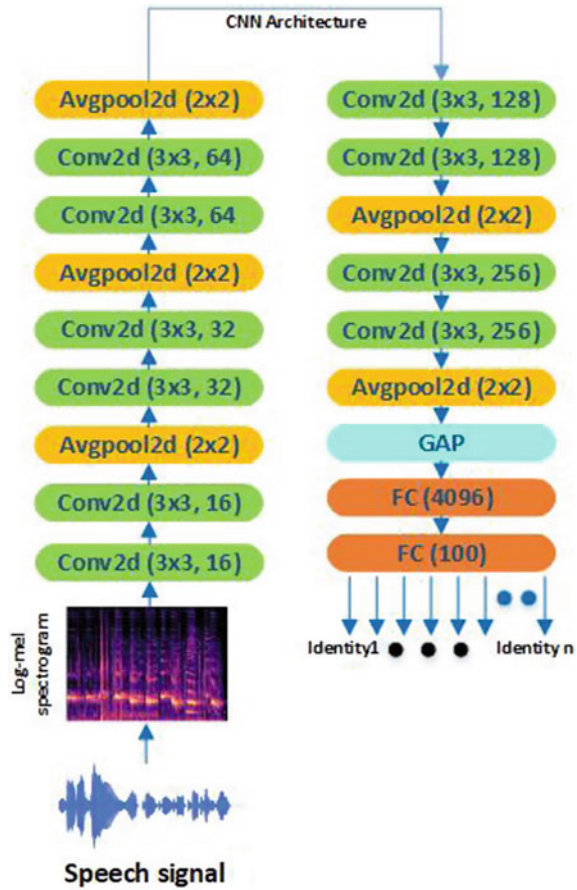
Mismatch Effects in Conversational Style The four variabilities of the Phase III database are described in Sect. 3.1. This research considers conversational styles 1 and 2 speech data for training and testing purposes. Speaker identification accuracy was 98.05 and 95.62% for matched and mismatched styles, respectively. Speaker identification accuracy of 88.1% is achieved using traditional approaches [13].

Figure 4 displays the speaker identification accuracy of the proposed architecture under conversational style mismatch conditions.
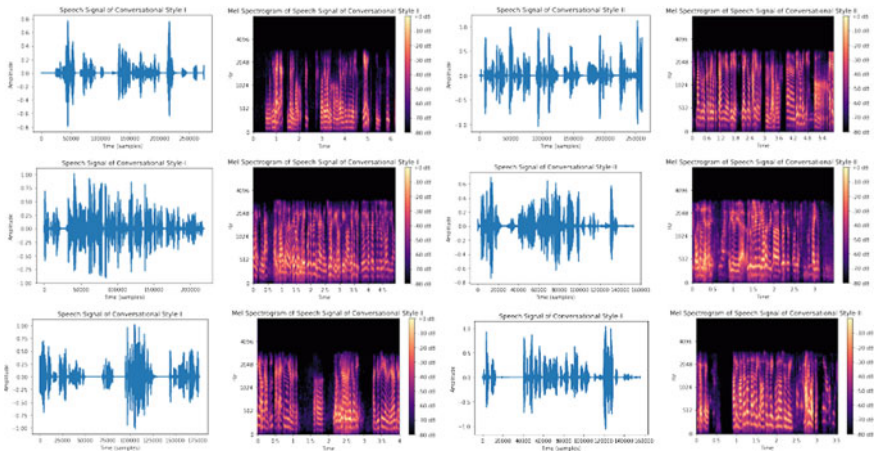
Fig. 3 Time domain and Mel spectrogram representation of speech data with different conversational style

Fig. 4 Accuracy (%) comparison for conversational mismatch cases
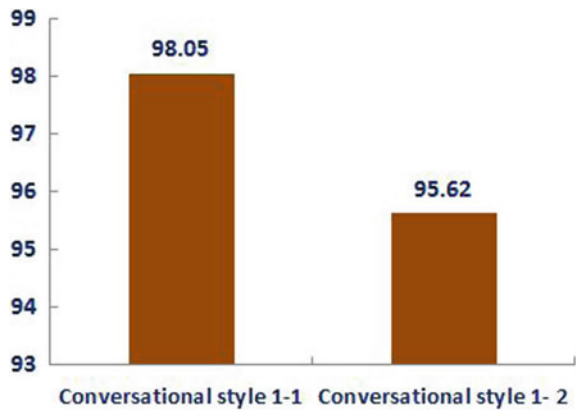
5 Conclusion

The robustness of speaker identification systems in various circumstances and diverse situations is challenging for real-world applications. This research examined how speaker identification performance is affected by degraded speech. This analysis considers mismatch effects in telephonic speech data collected during casual conference conversations. We utilized Phase III of the IIT-G database to address the issues. The presented VGG-like architecture provides robust speech features for an end-to-end intelligent speaker identification system, in contrast to existing methods. It is evident from the results that the identification accuracy for conversational style mismatch in conference call scenarios has enhanced by 7%. Future research includes

ResNet architectures to assess the speaker identification system performed when faced with language mismatch.

References

1. Saritha B, Shome N, Laskar RH, Choudhury M (2022) Enhancement in speaker recognition using SincNet through optimal window and frame shift. In: 2022 2nd International conference on intelligent technologies (CONIT). IEEE, pp 1–6
2. Saritha B, Laskar MA, Laskar RH (2023) A comprehensive review on speaker recognition. In: Biswas A, Wennekes E, Wieczorkowska A, Laskar RH (eds) Advances in speech and music technology. Signals and communication technology. Springer, Cham, pp 3–23. https://doi.org/10.1007/978-3-031-18444-4_1
3. Saritha B, Laskar MA, Laskar RH, Choudhury M (2022) Raw waveform based speaker identification using deep neural networks. In: 2022 IEEE Silchar subsection conference (SILCON). IEEE, pp 1–4. https://doi.org/10.1109/SILCON55242.2022.10028890
4. Jahangir R, Teh YW, Nweke HF, Mujtaba G, Al-Garadi MA, Ali I (2021) Speaker identification through artificial intelligence techniques: a comprehensive review and research challenges. Expert Syst Appl 171:114591
5. Haris BC, Pradhan G, Misra A, Shukla S, Sinha R, Prasanna SRM (2011) Multi-variability speech database for robust speaker recognition. In: 2011 National conference on communications (NCC). IEEE, pp 1–5
6. Ding I-J, Yen C-T, Ou D-C (2014) A method to integrate GMM, SVM and DTW for speaker recognition. Int J Eng Technol Innov 4(1):38–47
7. Auckenthaler R, Carey M, Lloyd-Thomas H (2000) Score normalization for text-independent speaker verification systems. Digit Signal Process 10(1–3):42–54
8. Reynolds DA (2003) Channel robust speaker verification via feature mapping. In: Proceeding of the 2003 IEEE international conference on acoustics, speech, and signal processing (ICASSP'03). IEEE, pp II–53
9. Teunen R, Shahshahani B, Heck L (2000) A model-based transformational approach to robust speaker recognition. In: 6th International conference on spoken language processing (ICSLP 2000)
10. Yin S-C, Rose R, Kenny P (2007) A joint factor analysis approach to progressive model adaptation in text-independent speaker verification. IEEE Trans Audio Speech Lang Process 15(7):1999–2010
11. Solomonoff A, Campbell WM, Boardman I (2005) Advances in channel compensation for SVM speaker recognition. In: Proceedings of the IEEE International conference on acoustics, speech, and signal processing (ICASSP'05), vol 1. IEEE, pp I/629–I/632
12. Chougule SV, Chavan MS (2017) Speaker recognition in mismatch conditions: a feature level approach. Int J Image, Graph Sign Process 4:37–43
13. Haris BC, Pradhan G, Misra A et al (2012) Multivariability speaker recognition database in Indian scenario. Int J Speech Technol 15:441–453

Artificial Neural Networks and Enhanced Adam Optimization for Effective Wi-Fi Intrusion Detection

Lenin Narengbam and **Shouvik Dey**

1 Introduction

The trend in recent years has been toward portable computing devices, such as smartphones, laptops, wearable, and tablets, overtaking traditional computers. A paradigm shift in computing has occurred as more and more software apps are rendered as cloud-based services. Traffic on the Internet has increased significantly since mobile devices and cloud-based services have become more prevalent. In addition, Wi-Fi networks have become ubiquitous and significant and become vulnerable to attacks and intrusions as a result. In Wi-Fi networks, there are growing security concerns. An immediate need exists for the creation of security methods and tools that cater to Wi-Fi networks.

The goal of an intrusion detection system (IDS) is to accurately demarcate between normal and abnormal network traffic while minimizing the number of false alarms (FAR). This is achieved through the use of a classifier that has been trained to accurately categorize network traffic. It is solely the training data and classification method that determine the classifier's accuracy and minimum FAR. A high-performance intrusion detection system is preferable to security experts. A wide range of intrusion detection solutions can be found in the literature [1–3], but the majority use traditional paradigms, while Wi-Fi intrusion detection methods are less widely used. A stunning range of applications of artificial intelligence (AI) has emerged in the last few years across a huge range of fields. It has sparked the interest of researchers in developing supervised learning frameworks to anticipate malicious attacks. These frameworks are developed on datasets that contain a variety of attacks. One of the biggest challenges faced by artificial neurons is the optimization process, specifically

L. Narengbam (✉) · S. Dey
Department of Computer Science and Engineering, National Institute of Technology Nagaland, Chumukedima, Dimapur, Nagaland 797103, India
e-mail: lenin.narengbam@gmail.com

selecting the right weights during training [4]. There may be a tendency for training procedures to become stuck at local minimums, and methods may fail to find a global optimal solution [5]. In the middle of behavior construction, the training algorithm takes an excessive amount of time to reach convergence.

An optimizer in machine learning helps find a low value in a complex function with multiple local minima and saddle points. The optimization process can be difficult, especially when the input has many dimensions. Gradient descent is a common optimization method, but it can be limited and may not always work effectively [6]. This is because it only moves in one direction at a fast pace and may not explore other areas where the derivative stays negative [7, 8]. Adam is a machine learning optimization algorithm that updates network weights using training data instead of traditional gradient descent [9]. It includes momentum and adjusts the gradient estimates to improve convergence. However, it may overshoot the optimal solution and oscillate before converging to the global minimum. Additionally, using a high learning rate may cause the classification model to crash for multidimensional datasets [3, 10].

In this study, we propose an enhanced adaptive moment estimation (Adam) optimizer (EA0) technique which is an adaptive learning rate coupled with the look-ahead mechanism for the Wi-Fi intrusion detection system. In this technique, the current velocity must be used to determine the gradient. A temporary position value depends on the movement in the direction of history. EAO advances in the direction of the gradient at the temporary point, rather than moving along the gradient. As a result of this look-ahead mechanism, the EAO reduces minimum valley oscillations. The EAO employs adaptive learning rates as well. The notion is that the parameters' learning rate should decay proportionally to their revised history. It may be appropriate to utilize a unique learning rate for each parameter when the sensitivity direction aligns to some extent. This will allow us to automatically adjust these learning rates as we progress through the learning process [11]. When it is close to convergence, the adaptive learning rate inhibits the EAO from remaining in local minima. Unlike Adam, the proposed method focuses on the global minimum and does not collapse even at increased learning rates.

The effectiveness of the method is measured using statistical metrics such as precision, accuracy, recall, false positive rate, and F1-score. The performance of the EAO is evaluated using traditional methods including Nesterov Accelerated Gradient Descent (NAG) [12], Root Mean Square Propagation (RMSProp) [13], Adaptive Gradient (AdaGrad) [14], Adaptive Moment Estimation (Adam) [15], and Stochastic Gradient Descent (SGD) [16]. The primary contribution of this research is as follows:

- The study utilized a binary classification method for Wi-Fi IDS that blends a dynamic learning rate with a look-ahead mechanism.
- The AWID dataset was used to assess the performance of the proposed model against traditional learning algorithms
- Existing literature approaches were utilized as a benchmark to validate the reliability of the proposed model.

The subsequent sections of this paper are organized as follows: Sect. 2 provides an overview of the Wi-Fi intrusion detection system using ANN and EAO. Section 3

delves into the simulation setup and evaluation metrics. The AWID dataset details are given in Sect. 4. The result analysis is shown in Sect. 5. Section 6 concludes with directions for future work.

2 Wi-Fi IDS Using ANN and EAO

The Wi-Fi IDS proposed in this study consists of three components: the input section and two main sections, which are the neural network section and the section that employs the enhanced Adam optimization algorithm. The overall framework of the Wi-Fi IDS using artificial neural networks (ANN) and enhanced Adam optimization (EAO) is summarized and depicted in a corresponding Fig. 1.

The proposed Wi-Fi intrusion detection system starts with an input phase, responsible for extracting important information from the flow of Wi-Fi network traffic. It then filters and prepares the data for processing. The crucial aspect of the input phase is to scale the incoming data to a range of [0,1] so that it can be effectively utilized by the neuron section.
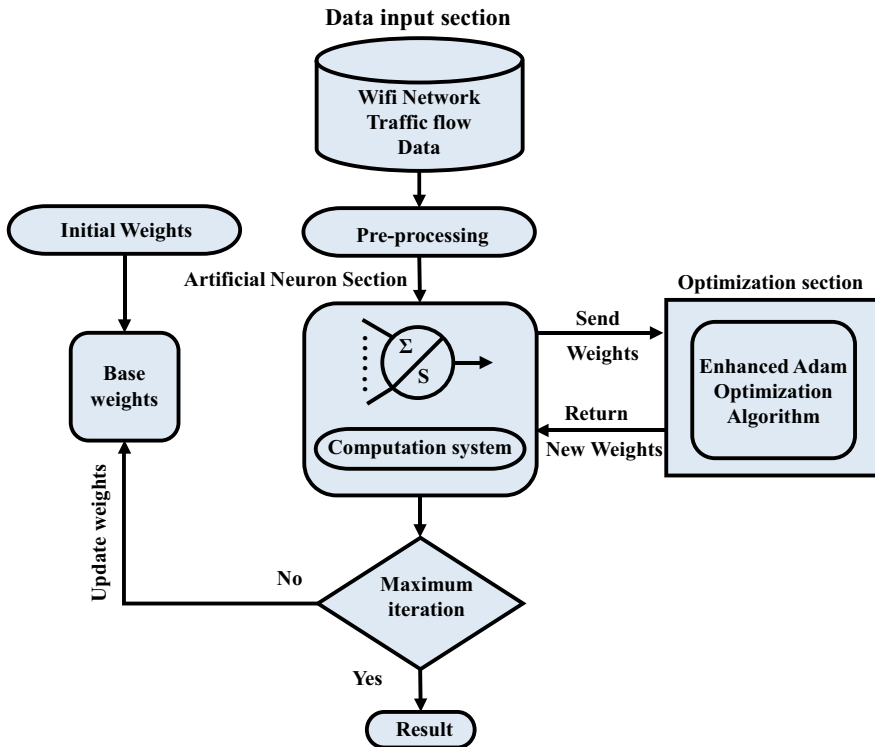
Fig. 1 Wi-Fi IDS using EAO

The second process involves the neuron section acquiring N features from the input dataset dimension. The training of the Wi-Fi network's behavior is carried out by utilizing the training dataset from the initial input section. The EAO section is designed to work on its own, and with every iteration, it will update the connection weights until the maximum is reached. A logistic sigmoid function is used to classify the data samples, and the sigmoid function's parameters are optimized by adopting the EAO. Using the current velocity, the slope is evaluated at a temporary spot after some distance has been covered. The gradient is calculated at the transition point, allowing the parameters' velocities to be updated. An exponentially decaying average is used to update the history of the gradient. In the early stages of training, the correction of biases is done to ensure a smooth training process and keep track of updates as they come in. Finally, the objective function parameters are changed. The global minima are reached by reducing oscillation and converging to the model using this technique. By canceling each other, reverse negative movement and positive movement impede loss function from traveling beyond the minima region. Both exponential decay rates were set to 0.9 throughout the 100 iterations and the learning rate at 0.001.

In the final step, the neuron, having undergone training with the training dataset, is then put to use in making predictions with the testing datasets. The accuracy of the predictions is measured through the computation of the Mean Squared Error (MSE). The comparison of the predicted results with the actual results enables the assessment of the efficiency of the neuron testing process.

3   Simulation Setup

Implementation and evaluation of the proposed method have been carried out using a computer running 64-bit Windows 10 Pro and featuring an Intel(R) Core (TM) i3-4010U CPU clocked at 1.70 GHz and 8 GB RAM. The Python programming language is used in the Google Colab to simulate all approaches. The simulation was performed with the help of the machine learning library scikit-learn. The learning rate is set at 0.01, 0.0001, 0.0001, 0.01, and 0.0001 for NAG, RMSProp, AdaGrad, SGD, and Adam, respectively.

The performance of the Wi-Fi IDS was assessed using various evaluation metrics, including accuracy (ACC), false positive rate (FPR), recall rate (RCR), precision rate (PCR), and F1-score (F1S). The results were analyzed using statistical measures such as average (AVE), standard deviation (SDE), median (MDI), and the lowest MSE achieved over 100 iterations. The goal of this evaluation was to achieve high accuracy, detection rate, and low false alarm rate, which were calculated through the use of a confusion matrix.

4 Dataset

In this work, we use the AWID dataset, created by Kolias et al. [17], which is publicly accessible, for evaluating the Wi-Fi IDS framework. This dataset features two attack class types, "CLS" and "ATK". Our study makes use of the reduced version of the "CLS" attack class, combining impersonation, injection, and flooding into one class, with the other class being designated as normal. The training set, AWID-CLS-R-Trn, has 1,633,190 normal instances and 162,385 attack instances, while the test set, AWID-CLS-R-Tst, includes 530,785 normal instances and 44,858 attack instances. To maintain balance, the normal and attack instances in the training and test datasets were proportioned at a 1:1 ratio, resulting in a reduction of normal instances to 163,320 in the training set and 44,934 in the test set. The dataset includes 154 features extracted from Wi-Fi network frames for each record, along with its corresponding class label. After preprocessing, only 34 features were selected from the dataset.

5 Result

The effectiveness of the proposed Wi-Fi IDS framework, which uses the EAO trainer for the neuron, was evaluated by comparing its accuracy rate, detection rate, false positive rate, precision, F1-measure, and MSE against ML and other learning methods on the AWID dataset. The top results were highlighted in bold text. A summary of the statistical results of the MSE, including average, median, standard deviation, and the lowest MSE, is presented in Table 1. The results show that the EAO trainer outperforms other learning methods in terms of MSE.

Table 2 presents the comparison of the performance of NAG, RMSProp, AdaGrad, SGD, Adam, and EAO using metrics such as ACC, RCR, FPR, PCR, and F1S. The proposed EAO model recorded the highest accuracy rate of 99.46%, surpassing Adam which had 98.34%. On the other hand, NAG had the lowest rate of 92.54%. The EAO model showed a better detection rate and precision rate compared to other methods, with a detection rate of 99.76% and a precision rate of 98.37%, respectively. The proposed framework achieved 99.06%, a better result than NAG (92.98%), RMSProp

Table 1 MSE for NAG, RMSProp, AdaGrad, SGD, Adam, and EAO

Algorithms	AVE MSE	MDI MSE	SDE MSE	Best MSE
NAG	0.2378	0.1986	0.1712	0.1855
RMSProp	0.2136	0.1462	0.1085	0.1467
AdaGrad	0.2023	0.1591	0.1156	0.1628
SGD	0.1896	0.1143	0.0927	0.1318
Adam	0.1824	0.1263	0.0908	0.1307
EAO	**0.1681**	**0.1089**	**0.0836**	**0.1052**

(95.34%), AdaGrad (93.01%), SGD (96.33%), and Adam (96.95%). The results of the comparison between the proposed EAO trainer and other machine learning techniques are presented in Table 3. This comparison takes into account the ACC, RCR, FPR, PCR, and F1S. The proposed Wi-Fi IDS framework was found to surpass all other techniques in terms of false positive rate, precision, accuracy, detection rate, and F1-measure. Furthermore, Fig. 2 displays the convergence curves of NAG, RMSProp, AdaGrad, SGD, Adam, and EAO based on the MSE. The results show that the convergence of the EAO trainer is faster compared to the other methods.

Table 2 Performance evaluation of Wi-Fi IDS

Methods	ACC (%)	RCR (%)	FPR (%)	PCR (%)	F1S (%)
NAG	92.54	95.27	17.56	91.21	92.98
RMSProp	94.17	94.72	14.21	92.64	93.66
AdaGrad	93.89	95.08	13.78	91.73	93.01
SGD	97.75	98.16	11.19	94.56	96.33
Adam	98.34	98.78	11.54	95.18	96.95
EAO	**99.46**	**99.76**	**01.16**	**98.37**	**99.06**

Table 3 Comparison of proposed EAO trainer with traditional machine learning algorithms

Method	ACC (%)	RCR (%)	FPR (%)	PCR (%)	F1S (%)
LR	88.47	93.62	20.43	89.61	89.66
NB	79.42	88.24	16.27	81.69	80.53
DT	83.53	92.19	17.45	91.06	87.13
SVM	96.48	96.87	08.34	94.71	95.59
EAO	**99.46**	**99.76**	**01.16**	**98.37**	**99.06**

Fig. 2 AWID dataset convergence curve for algorithms
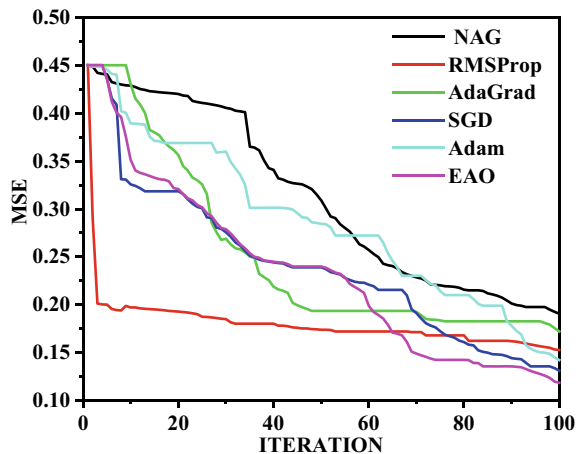
Table 4 Evaluation of the proposed method in comparison to other techniques for the AWID dataset

Year	References	Methods	ACC (%)	FPR (%)
2015	Kolias et al. [17]	AdaBoost	92.21	–
		Hyperpipes	92.21	–
		Random tree	91.43	–
2017	Aminanto and Kim [18]	FU DL	94.81	–
2020	Lopez-Martin et al. [19]	DDQN	95.70	–
		Policy gradient	92.21	–
		Actor critic	92.21	–
2020	Lee et al. [20]	IMPACT	94.72	4.61
Proposed model		N-HHO	**99.46**	**1.16**

The results presented in Table 4 indicate that the proposed EAO model surpasses other models in terms of accuracy when applied to the AWID dataset. The proposed technique demonstrates its ability to correctly classify instances in the dataset, outperforming other methods. The data presented in the tables and figures show that the use of the EAO trainer in conjunction with the AWID dataset leads to better results in terms of avoiding local minimum traps. Furthermore, the best accuracy achieved by the EAO model shows significant improvement when compared to the results from other algorithms.

6   Conclusion

In this study, a framework is introduced for creating an efficient Wi-Fi IDS using EAO. By integrating the look-ahead method while updating parameters, Adam's traditional optimization process was enhanced. The proposed EAO framework combines the look-ahead mechanism with an adaptive learning rate to reduce error and prevent oscillations in updating parameters. Models such as AdaGrad, Adam, and RMSProp crash after numerous iterations when the learning rate for a high-dimensional dataset is high. In contrast, the EAO algorithm demonstrates exceptional results even when using high learning rates on large datasets and consistently reaches global minimum values.

Consequently, the Wi-Fi IDS framework based on EAO shows potential as a viable solution for creating a high-performing intrusion detection system for Wi-Fi networks. In the future work, the EAO framework can be improved with an additional self-adaptive mechanism. Currently, the features in the dataset are selected manually. To enhance the system's performance, an advanced and adaptive feature selection method can be employed.

References

1. Ahmad Z, Khan AS, Shiang CW, Abdullah J, Ahmad F (2021) Network intrusion detection system: a systematic study of machine learning and deep learning approaches. Trans Emerg Telecommun Technol 32(1, e4150):1–29. https://doi.org/10.1002/ett.4150
2. Aldweesh A, Derhab A, Emam AZ (2020) Deep learning approaches for anomaly-based intrusion detection systems: a survey, taxonomy, and open issues. Knowl-Based Syst 189:105124. https://doi.org/10.1016/j.knosys.2019.105124
3. Khaire UM, Dhanalakshmi R (2020) High-dimensional microarray dataset classification using an improved adam optimizer (iAdam). J Ambient Intell Humaniz Comput 11:5187–5204
4. Ozturk C, Karaboga D (2011) Hybrid artificial bee colony algorithm for neural network training. In: 2011 IEEE Congress of evolutionary computation (CEC). IEEE, pp 84–88. https://doi.org/10.1109/CEC.2011.5949602
5. Hassim YMM, Ghazali R (2014) Optimizing functional link neural network learning using modified bee colony on multi-class classifications. In: Advances in computer science and its applications. Lecture notes in electrical engineering, vol 279. Springer, Berlin, Heidelberg, pp 153–159
6. Kolossoski O, Monteiro RDC (2017) An accelerated non-Euclidean hybrid proximal extragradient-type algorithm for convex–concave saddle-point problems. Optim Methods Softw 32(6):1244–1272
7. Le T-T-H, Kim J, Kim H (2017) An effective intrusion detection classifier using long short-term memory with gradient descent optimization. In: 2017 International conference on platform technology and service (PlatCon). IEEE, pp 1–6
8. Takase T, Oyama S, Kurihara M (2018) Effective neural network training with adaptive learning rate based on training loss. Neural Netw 101:68–78
9. Kingma DP, Ba J (2014) Adam: a method for stochastic optimization. arXiv:1412.6980 [cs.LG], arXiv:1412.6980v9 [cs.LG], https://doi.org/10.48550/arXiv.1412.6980
10. Goodfellow I, Bengio Y, Courville A (2016) Deep learning. The MIT Press, London, England
11. Smith LN, Topin N (2019) Super-convergence: very fast training of neural networks using large learning rates. In: Proceedings of the artificial intelligence and machine learning for multi-domain operations applications, vol 11006, pp 369–386. SPIE
12. Giselsson P, Doan MD, Keviczky T, De Schutter B, Rantzer A (2013) Accelerated gradient methods and dual decomposition in distributed model predictive control. Automatica 49(3):829–833
13. Mukkamala MC, Hein M (2017) Variants of RMSProp and Adagrad with logarithmic regret bounds. In: Proceedings of the 34th international conference on machine learning, vol 70. PMLR, pp 2545–2553
14. Ward R, Wu X, Bottou L (2020) AdaGrad stepsizes: sharp convergence over nonconvex landscapes. J Mach Learn Res 21(1, Article No.: 219):9047–9076
15. Yu Y, Liu F (2019) Effective neural network training with a new weighting mechanism-based optimization algorithm. IEEE Access 7:72403–72410
16. Chen Y, Chi Y, Fan J, Ma C (2019) Gradient descent with random initialization: fast global convergence for nonconvex phase retrieval. Math Program 176:5–37
17. Kolias C, Kambourakis G, Stavrou A, Gritzalis S (2016) Intrusion detection in 802.11 networks: empirical evaluation of threats and a public dataset. IEEE Commun Surv Tutorials 18(1):184-208. https://doi.org/10.1109/COMST.2015.2402161
18. Aminanto ME, Kim K (2018) Improving detection of Wi-Fi impersonation by fully unsupervised deep learning. In: Information security applications. WISA 2017. Lecture notes in computer science, vol 10763. Springer, Cham, pp 212–223
19. Lopez-Martin M, Carro B, Sanchez-Esguevillas A (2020) Application of deep reinforcement learning to intrusion detection for supervised problems. Expert Syst Appl 141:112963
20. Lee SJ, Yoo PD, Asyhari AT, Jhi Y, Chermak L, Yeun CY, Taha K (2020) IMPACT: impersonation attack detection via edge computing using deep autoencoder and feature abstraction. IEEE Access 8:65520–65529

Computerized Sensing of Diabetes Retinopathy with Fundus Images Using CNN

Waseem Khan and Khundrakpam Johnson Singh

1 Introduction

In the current scenario, we are all aware that diabetes is one of the worst and most deadly diseases not only in India as well as in the world. A large number of people are diagnosed with this disease, and from this, the chance of other problems or other medical problems is also there [1, 2]. Diabetes may lead to a higher chance of many eye diseases; one such is diabetic retinopathy. Screening and testing of eyes are not as cheaper as it is also a time and cost-effective procedure. Apart from screening and testing of eyes, a clinical person needs a required lot of experience so that they could detect whether either person is suffering from diabetic retinopathy or not. Several models and techniques have been introduced using DL models to discover DR.

However, DR is irreversible blindness, and it is mostly in the mid-age population; DR is one of the major causes of unrepairable visual deficiency. Figures 1 and 2 show the normal and diabetic retina images. As diabetes is a chronic disease, severe and systemic disease, it affects different organs like kidneys, eyes, and systems never still as the heart. Diabetic retinopathy may create no side effects or fair minor visual anomalies at, to begin with.

However, it can cause blindness. Unquenchable diabetes and its complexity led to DR [3]. According to a report, there are 103.12 million people are suffering from diabetic retinopathy, and by the year 2045, it will increase to 160.50 million. It is one of the driving causes of visual insufficiency in people.

DR might happen in type-1 and type-2 diabetes patients; it happened since of the high level of sugar inside the blood [4].

In diabetic retinopathy, the retina's blood vessels are harmed due to the increased level of sugar in the blood. Diabetic retinopathy is broadly separated into parts

W. Khan (✉) · K. J. Singh
Department of Computer Science and Engineering, NIT Manipur, Imphal, Manipur, India
e-mail: waseemk49@gmail.com
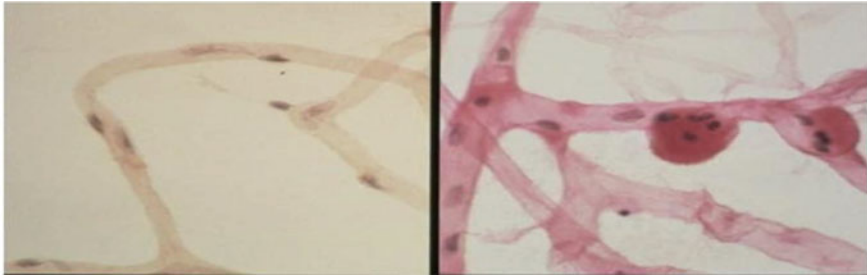
Fig. 1 Normal and diabetic retina veins image
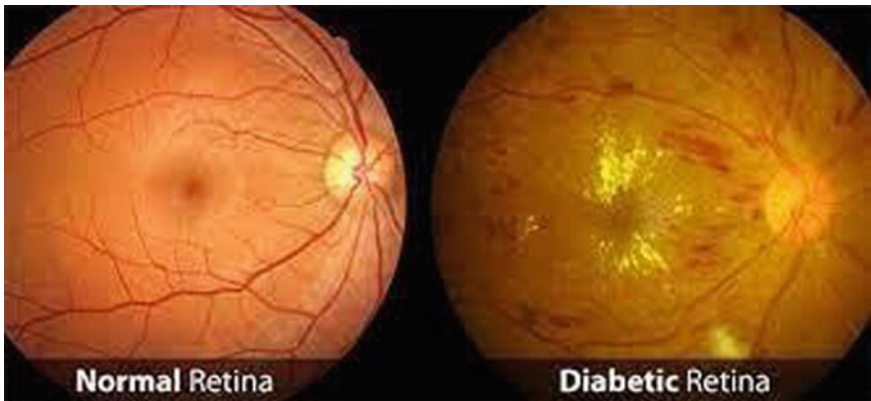
Fig. 2 Normal retina and diabetic retina

named pre-diabetic retinopathy and post-diabetic retinopathy. Diabetic retinopathy is additionally alluded to as non-proliferating diabetic retinopathy (NPDR), and post-diabetic retinopathy is alluded to as proliferative diabetic. Table 1 represents the different severity scales of diabetic retinopathy. In the first category of diabetic retinopathy, the new blood vessels are not developed which is called proliferating. In this, the walls of the veins in our retina get weakened [5]. Diminutive bulges are obtruded from the walls of the tiny vessels. Often it seems that leaking of blood and fluid into the retina. Broad retinal vessels can begin to widen and may lead to unpredictable shapes as well. NPDR can be taking place from mild to severe as a large number of blood vessels become blocked.

Though inside the post-diabetic retinopathy can reach out to this more serious sort. In this, impeded veins are cut off, which bring practically the improvement of present-day, untypical veins inside the retina [4].

These as of late made veins are fragile and can drain into the reasonable, jam-like substance that fills the focal point of our eye (glassy). In case the as of late made veins barge in into the standard stream of fluid out of the eye, the oblige can build

Table 1 International clinical DR disease severity scale (ICDRDSS)

Disease severity level	Finding upon dilated ophthalmoscopy
0 = No DR	No abnormality
1 = Mild DR	Micro-aneurysms only
2 = Moderate DR	More than micro-aneurysms but less than NPDR
3 = Severe DR	Any of the following and no sign of proliferative retinopathy: more than 20 intraretinal hemorrhages in each of four quadrants Definite venous beading in two or more quadrants Prominent IRMA in one or more quadrants
4 = Proliferative DR	One or both of the following Neovascularization vitreous/pre-retinal hemorrhage

inside the eyeball. It can harm the nerve that takes a picture from the eye to our brain (optic nerve), which comes about in glaucoma.

In spite of the fact that screening and rescreening of an individual experiencing DR is an exorbitant and exceptionally tedious cycle.

In this manner, numerous pseudoscientists have explored different avenues regarding scientific methods and the programmed discovery of DR [1]. As the screening of DR is the time taken to handle it likewise expected a thoroughly prepared clinician to check and confirm the computerized fundus image of the retina. Figure 3 shows the various instances of DR, by the time human per user presents their survey, which might prompt the postponed results lost to follow-up, miscommunication, and deferred treatment [4].

An expert ophthalmologist can find DR by the closeness of wounds related with the vascular combinations from the standard brought around by sickness. As this approach is sensible, its resource demands are particularly tall. The expert ophthalmologist and contraption are regularly lost in the locale where how much diabetes inside the local people is tall and DR exposure is required. As how much of people with diabetes increments bit by bit, the system is expected to stay missing people from visual needs since DR will get the opportunity to insulant be for unquestionably more [6].

The rate of increment of diabetes among adults from 4.7% in 1980 to 8.5% in 2014. The experienced clinician is very less and is unevenly distributed. The general number of experienced ophthalmologists all over the planet in the year 2012 was 210,730 which proposes that 29 ophthalmologists. The differentiation between the general population of diabetes patients and the expert clinical can be greater, which mirrors the expectations for a squeezing structure for the area of diabetic retinopathy thusly [2]. Figure 4 represents the clinical protocol chart.

Deep learning methods rapidly become the most interesting as well as a promising technique for pseudoscientists and are acquiring acceptance in countless practical applications in engineering [5].
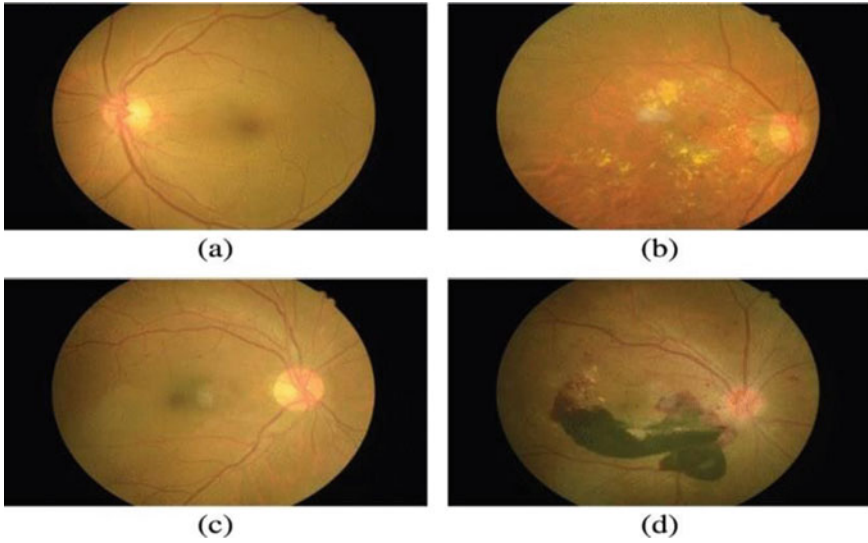
Fig. 3 Different cases of diabetic retinopathy

Fig. 4 Clinical protocol chart
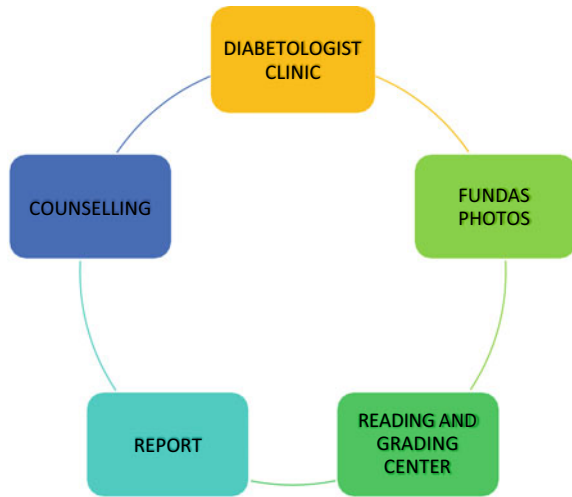
So much work is done in computers to detect automatically diabetic retinopathy diagnoses. Earlier models removed different features to at first remove significant information from the fundus pictures. From that point forward, these features are sustained to specific kinds of classifiers. What is more, these hand-planned highlight-based procedures are troublesome and routinely miss the mark to abandon far-off better; much better; a higher; a stronger; an improved a far-off improved outcome.

Inside the seem decade, fabricated pieces of data have achieved directly arrives at the place of understanding in various zones. Artificial neural networks and deep learning are ideally and structurally influenced by traditional neural systems; they become an interesting and promising technique for pseudoscientists in different fields including analysis of medical imaging [7]. Deep learning means learning the representation of data. Figure 5 shows what is deep learning?

Most deep neural networks are based on neural network architecture that is so often called DNN. The word significant suggests there are a small bunch of disguised layers inside the mindset up. The timetable mindset up s so to talk dealing with 2–3 up layers, disregarding the way that it can have as various as 150. Figure 6 seems the neural frameworks, which are worked with in layers containing a lot of interconnected focus centers. Frameworks can have tens or various covered layers [5].

Fig. 5 What is deep learning?
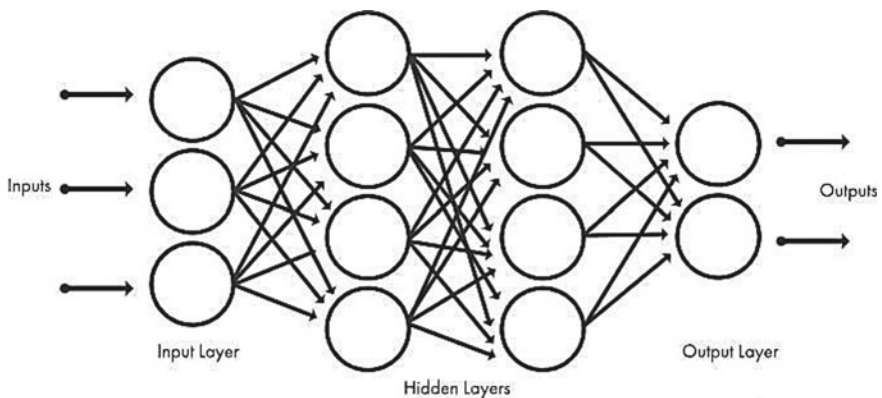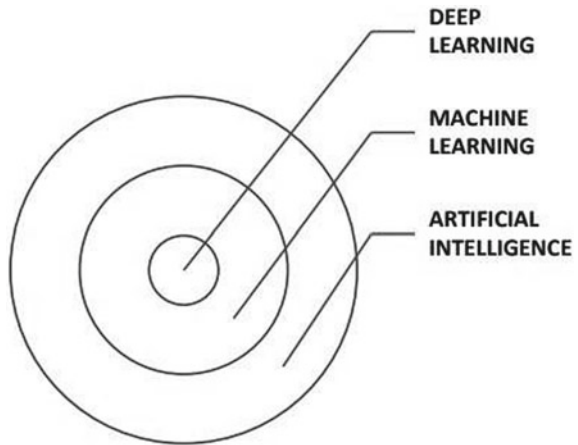
Fig. 6 What is a neural network?

Countless named datasets are used for the getting ready of significant learning models. The brain coordinate plan will clearly gain from the data with no manual feature extraction [5]. It has performed mostly well as a classifier for image-processing applications and as a function estimator for linear as well as nonlinear applications.

In the recent decade, the deep neural network has found revolutionary outcomes in different fields. In the 1970s, coordinate models were arranged that work on picture data with important applications and different ways to deal with testing tasks like physically composed character affirmations. It brings around leap forwards in defy affirmation, talk affirmation, PC vision, trademark lingo taking care of, and various more [5, 7].

The different application of deep neural networks surprises humans such as COVID detection, face recognition, and wide-scale visual recognition. The utilization of significant brain frameworks inside the assurance of diabetic retinopathy has in addition gotten much captivated, and much development has been made. Notwithstanding the way that various impels have been made, clinical utilization of customized diabetic retinopathy end structures stays blocked off; thus, much work is as yet expected to be finished [7]. Figure 7 represents the various ways for DR detection.

A branch of a DNN, i.e., known as a convolutional neural network. A convolutional brain orchestrate basically has a spot to the feed-forward made brain organize, which has a spot to significant brain frameworks. The CNN might be an incredibly notable significant learning plan that can get familiar with a food chain of features that is used for picture grouping [2]. As a show learns more complicated features, understanding, and other mutilation features in advanced layers, the precision of the exhibit can be higher. Moreover, in light of this suspicion, we explore the working of the convolutional brain put together for diabetic retinopathy [8]. Regardless, an essential multi-facet CNN show is framed, and tests are taken a stab at interesting retina data.
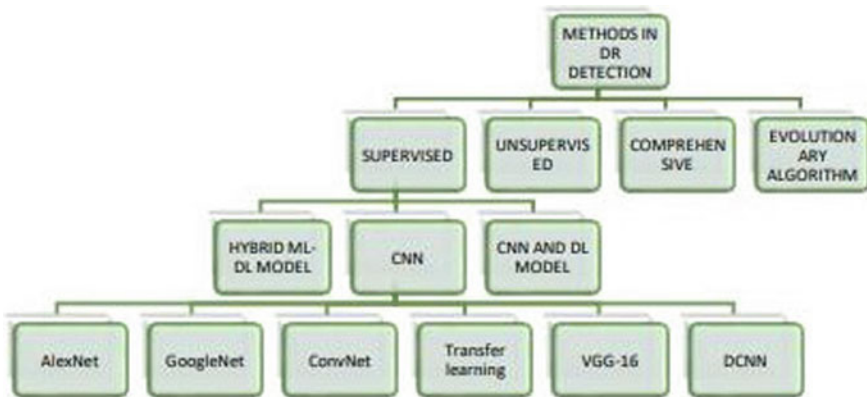
Fig. 7 Different methods for DR detection

Inside the ongoing circumstance, especially broad CNN was used to actually deal with really complex picture affirmation issues with unmistakable dissent classes to an imperative norm. It is used in various present statuses of craftsmanship picture characterization work, for example, COCO challenges and the yearly ImageNet.

There are a few issues within the automated grading and particularly with CNN. The first one is to achieve a required offset insensitivity which means the person is correctly identified as either it going through DR or specificity which means a person is not going through DR. That is strikingly convenient for public measures which might be a five-name issue from common, delicate, immediate, outrageous, proliferative diabetic retinopathy classes [5]. In addition, the overfitting of the dataset is one more issue in brain organizations. Slanted datasets may prompt the brain organization to overfit the class which is most conspicuous in the dataset [3]. Large datasets are frequently hugely slanted. In this paper, we take a significant learning-based convolutional brain coordinate show to recognize particular kinds of diabetic retinopathy with the help of fundus pictures. Typically, a remedial imaging handle with creating indicative importance, as we discussed earlier, and one that has been exposed to various things about inside the past.

At particular phases of the CNN procedure, it contains an extensive number of teachable boundaries which are used to find essential features from retina pictures at unmistakable reflection levels. The most downright horrendous piece of the CNN-based technique is that it needs an epic dataset to get ready [9]. Isolated from this, it required speedy registering resources for planning and tuning with the hyperparameters.

It could produce a good result in different disease diagnoses like the categorization of pneumonia as well as COVID-19 patients with the help of analysis of chest X-ray image, blind patients' differentiation with analysis of retina image, classification of brain tumor with MRI image analysis, and many more. CNN's methods are also very helpful to find the classes of DR severity from the original retina images [10].

2 Related Work

This part discusses CNN-based models for DR severity classification.

An exhaustive examination has been performed on the model for double separation of DR with a promising end.

Gardner et al. used brain networks as well as pixel values power to achieve responsiveness and particularity aftereffects of 88.4 and 83.5% individually for 0 or 1 arrangement of DR.

They utilized the restricted dataset of pictures and parted each picture into patches and afterward expected an ophthalmologist to confirm the patches for the component before the execution of SVM.

Within the three-class classification of DR, neural systems have moreover been connected. The range of exudates and the zone of blood vessels, as well as textural

parameters, were utilized by Nayak et al. To classify pictures into typical, non-proliferative retinopathy, and proliferative retinopathy, highlights are embedded into the neural arrangement. These characteristics were utilized as categorization input by the neural organization [8].

The detection results were confirmed by comparing them to professional ophthalmologist grading. They achieved 93% classification accuracy, 90% sensitivity, and 100% specificity. This was done on a dataset of 140 photographs, and highlight extraction was required on all of them in both preparing and testing, which took a long time [6].

Support vector machines have been employed in the great bulk of studies on the five-class categorization (SVMs) [11]. The five classifications have been identified using an automated approach developed by Acharya et al. The SVM classifier businesses feature taken from the unrefined data utilizing a higher-request range way to deal with catching the data.

The precision, affectability, and explicitness of this SVM approach were all over 80% typical. Acharya et al. made a five-class order strategy by evaluating the zones of a couple of qualities like hemorrhages, micro-aneurysms, exudate, and veins [12]. The principal basic attributes, veins, miniature aneurysms, exudates, and hemorrhages, were recuperated from crude pictures with the help of picture-getting-ready techniques.

These were by then set into the SVM to be arranged. This approach achieved an affectability of 82%, particularity of 86%, and precision of 85.9%. These methodologies were associated with astoundingly little datasets, and the reduction in affectability and explicitness was in all probability connected with the intricacy of the five-class issue [4].

Adarsh et al. used pictures dealing with gadgets to supply a robotized determination for DR by perceiving retinal veins, exudate, miniature aneurysms, and textural attributes. The incorporate vector for the multivariate data SVM was constructed using the zone of wounds and textural attributes. On average citizens 89 and 130 picture datasets DIARETDB0 and DIARETDB1, are made precision paces of 96 and 94.6%, independently [8].

Every one of the first five class methods required picture highlight extraction prior to being taken care of into an SVM classifier, and they were just confirmed on minuscule test sets of around 100 photographs. These methodologies are less material continuously than a CNN [13].

The researchers proposed a graph neural network (GNN)-based approach for DR severity classification in their paper. The method initially determines the ROI from imagery which target the regions that identify lesions [8]. Following that, the technique employs the GNN for fundus image categorization.

A deep learning approach for detecting and classifying DR fundus was proposed by the researcher. The procedure began by reducing the excess noise that emerges on the edges. The image's important sections were then retrieved using histogram-based segmentation. Finally, in the DR fundus pictures, the synergic deep learning (SDL) approach is employed to identify severity classes [8].

The pseudoscientist proposed a procedure for diagnosing DR in view of two specific variables: (1) dim level concentrated and (2) surface features removed from fundus pictures. These qualities are by then used to order data utilizing a choice tree-based outfit learning approach. Subsequently, the show achieves 94.20% order accuracy and an F-proportion of 93.51%.

The KNN classifier was used by the maker to recognize DR illnesses by removing attributes of the optic plate, veins, and exudates from retina pictures. The DR-impacted picture order showed by the expert depended on the stationary wavelet change and the discrete wavelet change coefficient. The model's precision was 94.17% [8].

The authors created a classification approach that combined SVM and neural network models [12]. They investigated feature extraction and segmentation processes before applying model pre-processing. The model has an accuracy of 80%.

The scientists outlined a methodology for recognizing drusen, cotton-fleece fixes, and exudates in retina pictures [4]. The estimation gets 0.95 inside the Recipient Working Characteristics (ROC) score and performs so likewise to an expert. The makers showed a constant procedure for fundus picture order in view of an AM-FM modified assessing structure [8]. Sometimes as of late applying the retinal features to the illustration, the makers removed them regardless.

Mehedi et al. [12] demonstrated the approach with area under the receiver operating characteristics (AUROC) which is used to classify the true and false cases. Figure 8 represents the AUROC graph of the model. The methods represent the fair performance in comparison between the different cases of DR [8].

3 Methodology

3.1 Dataset, Hardware, and Software

The dataset of retina pictures that we used during the testing and getting ready for the show was taken from the Kaggle people group. Until later times, sets of remarked on pictures of diabetic retinopathy were confined. We had the option to get ready in the general dataset by resizing these photos and running our CNN on a top-of-the-line GPU, the NVIDIA GTX. The NVIDIA GTX contains 1344 CUDA focuses and integrates the NVIDIA CUDA Profound Neural Organize (cuDNN) library for GPU learning.

About 15,000 photos were traded onto the GPU memory out of nowhere by using this program. Keras's (http://keras.io/) significant learning PC program was joined with the Theano (http://deeplearning.net/programming/theano/) AI back end. This was picked since it has good documentation and a quick calculation time. An image might be arranged in 0.04 s, allowing for continuous contribution to the calm.
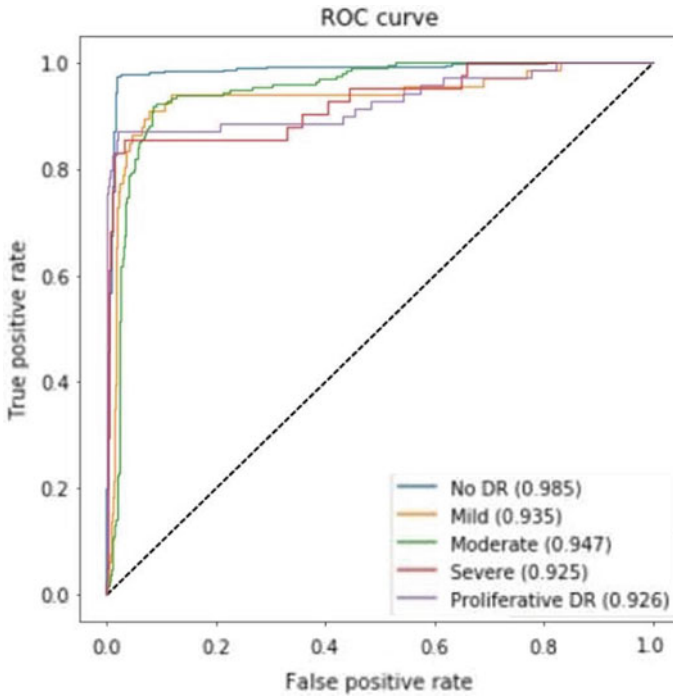
Fig. 8 AUROC score of the proposed method for variety of DR classes

3.2 Preprocessing

The collection comprises photos from patients of multiple ethnicities, ages, and illumination levels in fundus photography. This has an effect on the gray levels in the photos, causing additional variance unrelated to categorization levels. To overcome this issue, we rescale the images. Using rescaling, we treat all the images in a similar way because some images are a high pixel range, while others are a moderate pixel range. These images are sharing the same method.

3.3 Training and Proposed Model

In this study, CNN is employed for feature studying of referable DR. Four convolutional layers are used for each channel, with the number of filters increasing in succeeding layers to 32, 64, 120, and 256. To keep away from overfitting, most limit pooling, revised direct unit authorization work, and dropout are used.

The completely connected layers are coupled after flattening from the two channels to statistically finding the detection of referable diabetic retinopathy. TensorFlow

software and Python are used to code the suggested referable DR detection technique. For training the network, the binaryCross_Entropy loss function and the ReLu algorithm with a learning rate of 0.001 are used. The CNN was pre-trained on the number of retina images at first until it achieved a substantial level.

This was required in order to produce a pretty speedy classification result without spending a significant amount of training time. After 120 epochs of preparing on the first photographs, the organization was prepared for one more 40 epochs on the total 5000 preparation pictures. Neural networks suffer from significant overfitting, especially in a dataset like ours, where the bulk of the photos are categorized in one class that displays no evidence of retinopathy.

To deal with this issue, we joined ongoing course loads inside the network. The class weights were updated using a ratio proportional to the number of images in the training batch and were differentiated as having no evidence of DR for each batch applied for back-propagation. This considerably lowered the possibility of over-fitting to a specific class. A modest learning rate of 0.001 was utilized for three epochs to stabilize the weights.

A couple of specific CNN models have been suggested and assessed in our tests. The significance of the attempted brain arrangement runs from 9–18, though the convolution part measure goes from 1 to 5. To meet the information gauge of the CNN, we scale the image measure to $200 \times 200 \times 3$. Figure 9 seems a definitive coordinate arrangement used in our research [2]. The sort-out returns two probabilities for each piece of information that incorporates dependent upon one, one for each class (our issue might be a twofold grouping issue). In our attempt, we used various marked photos which are used to prepare the brain coordinate, while a couple of the photos are utilized to evaluate the execution of the arranged brain sort out.
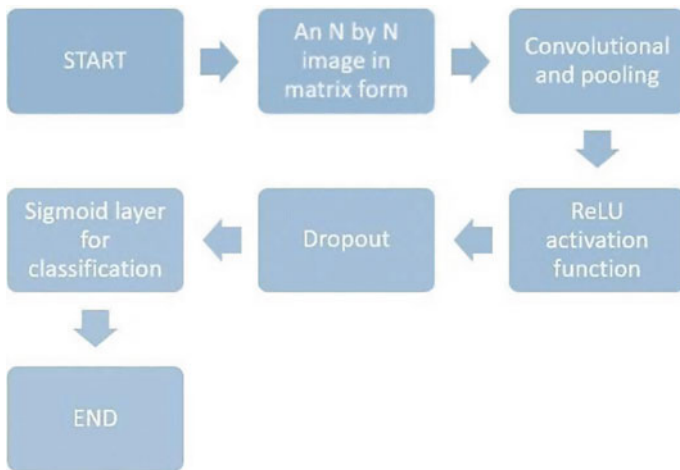
Fig. 9 CNN architecture used in our experiment
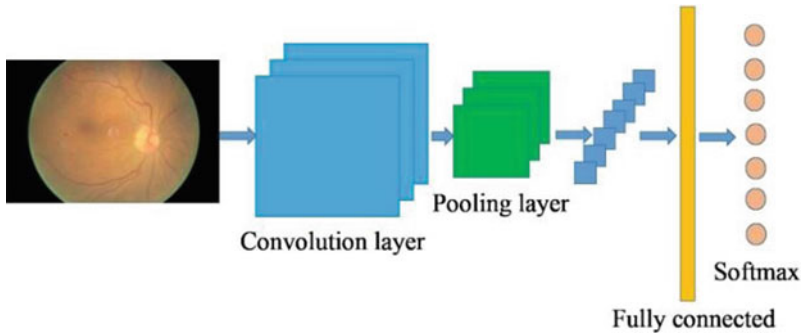
Fig. 10 Exemplary Architecture of convolutional neural network

The model was then trained with a low studying rate on the whole training set of pictures. Within a few huge epochs of the whole dataset, the network's accuracy had grown to almost 98%. When compared to other approaches, the CNN-based models outperform them, supporting the premise made in the introductory section.

The entropy picture of the brightness of fundus picture indicates the complexity of the actual retinal picture and aids in the training of the convolutional neural network-based DL. To quantify the heterogeneity, the merit in the entropy picture is computed locally from n × n blocks. Entropy is decided by the probability distribution of the local intensity.

In this paper, a lightweight CNN model for identifying the severity of DR from retina pictures. Figure 9 depicts the suggested custom model. The model contains four 2D convolutional layers. In each\s layer, the filter size is 3 × 3, and a similar padding is used. A "ReLu" activation function is used by each layer.

In extension, the show uses the MaxPooling2D approach after the second and fourth layers, as well as dropout regularization after the occasion, fourth, and completely associated layers. The softmax activation work is associated with the grouping layer due to the staggered order.

In the model, the (RMSprop) optimizer is used. RMSprop eliminates oscillations and automatically adjusts learning rate. The studying rate in RMSprop is divided by an exponentially decaying average of squared gradients. RMS propagation prevents you from searching in the direction of oscillations. RMS propagation additionally chooses a unique learning rate for each parameter. Figure 10 represents the basic model of the CNN [2].

3.4 Result

A large dataset of retina images was saved for the validation and training process. Moreover, the training dataset of retina images is divided into 80% to train the network and rest 20% for the validation of the network. Working on the validation images

dataset on the network took a few seconds. For this, we define two-class specificity either the patient is suffering from diabetic retinopathy or the patient is not suffering from diabetic retinopathy. The final trained model acquired almost 98% accuracy, i.e., shown, in Fig. 11 and model loss which shown in Fig. 12.

Our trained convolutional neural network has the potential benefit of being able to categorize hundreds of pictures per minute, allowing it to be utilized in real-time anytime a new image is collected. Pictures are given to specialists for exploring in sharpen; however, they are not appropriately assessed when the tranquil come in for screening. The trained CNN allows for a speedy diagnosis and immediate reaction to
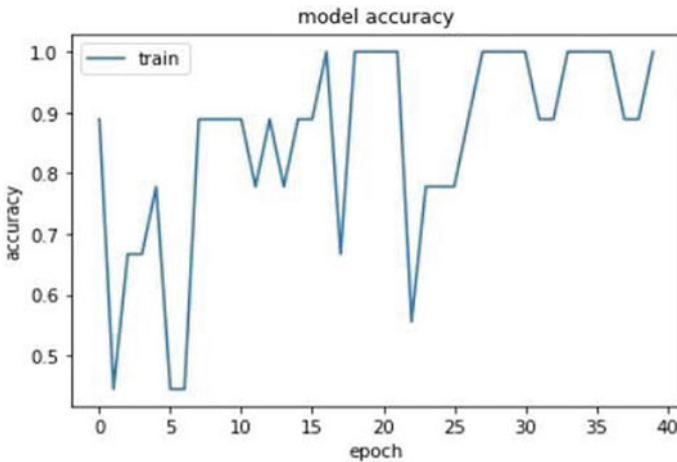
Fig. 11 Model accuracy graph
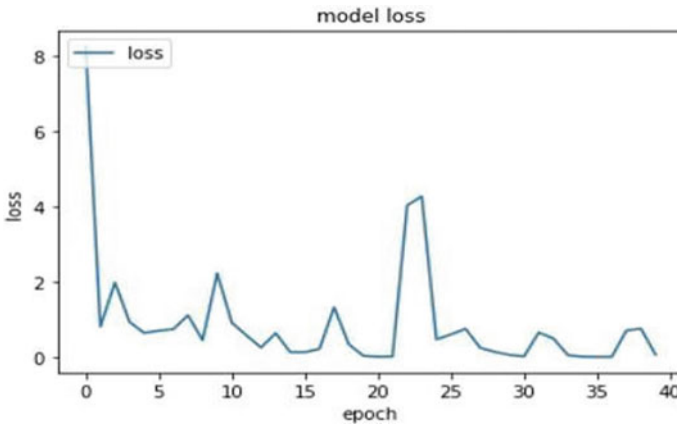
Fig. 12 Model losses graph

Table 2 Performance differentiation with different approaches

Methods	Accuracy (%)
Hard exudates + GBM	89.4
Red lesions + GBM	88.7
Micro-aneurysms + GBM	86.2
Blood vessel detection + GBM	79.1
CNN without data augmentation	91.5
CNN with data augmentation	94.5

a patient. These results were likewise attained by the network with only one picture per eye.

Pictures are given to doctors for reviewing in hone, but they are not suitably evaluated when the quiet comes in the continuous advancements in CNNs enable for far deeper networks that can learn the nuanced properties that this network was unable to understand. The findings of our network from an orthodox network topology are highly encouraging. Unlike earlier techniques, nothing explicitly connected to the properties of our fundus photos, such as vessels, exudate, and so on, has been utilized. This makes the CNN results outstanding, but we have plans to tailor our network to this specific purpose in the future in order to learn the more nuanced categorization criteria screening.

The system has no difficulty learning to recognize a picture of a normal eye. This is most likely because the sample contains a significant proportion of healthy eyes. The training required to categorize photos at the extreme ends of the spectrum was substantially less during training. The difficulties arose while attempting to train the network to discriminate among moderate, mild, and severe cases of diabetic retinopathy.

DR is a serious medical health issue that causes blindness, and DL approaches can play a more vital role in its diagnosis and early identification than traditional procedures. This document covers DR in detail, including its symptoms, characteristics, form, size, and location, as well as how DR causes blindness. It also explains numerous ML and DL strategies for detecting aberrant RBV and OD behavior in order to diagnose DR lesions. Table 2 represents the performance differentiation with different approaches.

4   Conclusion

With a certain number of medical personnel, an automation model can considerably reduce the time-consuming human work needed in diagnosing large numbers of retinal pictures. In prior investigations, feature extraction-based diabetic retinopathy diagnosis was dominating. Our model has shown good evidence of learning the characteristics needed to classify fundus pictures, correctly distinguishing the bulk

of proliferative cases and patients with no DR. Our research has demonstrated that the challenge for nationwide diabetic retinopathy screening may be handled using a convolutional neural network method. As in other large dataset research, increased specificity has come at the expense of poorer sensitivity. Our strategy achieves an equivalent that happens to these earlier methodologies notwithstanding the non-attendance of element explicit conspicuous evidence and the use of a stunningly greater dataset.

The results are positive when compared to human grading reports; hence, a clinical study will be conducted in order to include the given method into a tool for diagnosing diabetic retinopathy. To sum up, we represented that CNNs can be told to perceive diabetic retinopathy attributes in fundus pictures. CNNs have the capacity to enhance beneficial to diabetic retinopathy physicians in the future when models and datasets improve and they can provide genuine-time categorization.

References

1. Pao S-I, Lin H-Z, Chien K-H, Tai M-C, Chen J-T, Lin G-M (2020) Detection of diabetic retinopathy using bichannel convolutional neural network. J Ophthalmol 2020(9139713):1–7
2. Kele X, Feng D, Mi H (2017) Deep convolutional neural network-based early automated detection of diabetic retinopathy using fundus images. Molecules 22(12):2054
3. Kolla M, Venugopal T (2021) Efficient classification of diabetic retinopathy using binary CNN. In: 2021 International conference on computational intelligence and knowledge economy (ICCIKE). IEEE, pp 244–247
4. Bodapati JD, Naralasetti V, Shareef SN, Hakak S, Bilal M, Maddikunta PKR, Jo O (2020) Blended multi-modal deep ConvNet features for diabetic retinopathy severity prediction. Electronics 9(6):914
5. Tsiknakis N, Theodoropoulos D, Manikis G, Ktistakis E, Boutsora O, Berto A, Scarpa F, Scarpa A, Fotiadis DI, Marias K (2021) Deep learning for diabetic retinopathy detection and classification based on fundas images: a review. Comput Biol Med 135:104599
6. Gao Z, Li J, Guo J, Chen Y, Yi Z, Zhong J (2018) Diagnosis of diabetic retinopathy using deep neural networks. IEEE Access 7:3360–3370
7. Chudzik P, Majundar S, Calivá F, Al-Diri B, Hunter A (2018) Microaneurysm detection using fully convolutional neural networks. Comput Methods Programs Biomed 158:185–192
8. Pratt H, Coenen F, Broadbent DM, Harding SP, Zheng Y (2016) Convolutional neural networks for diabetic retinopathy. Procedia Comput Sci 90:200–205
9. Maniadakis N, Konstantakopoulou E (2019) Cost effectiveness of treatments for diabetic retinopathy: a systematic literature review. PharmacoEconomics37:995–1010
10. Kollias AN, Ulbig MW (2010) Diabetic retinopathy: early diagnosis and effective treatment. Dtsch Arztebl Int 107(5):75–84
11. Ghosh R, Ghosh K, Maitra S (2017) Automatic detection and classification of diabetic retinopathy stages using CNN. In: 2017 4th International conference on signal processing and integrated networks (SPIN). IEEE, pp 550–554
12. Masud M, Alhamid MF, Zhang Y (2022) A convolutional neural network model using weighted loss function to detect diabetic retinopathy. ACM Trans Multimedia Comput Commun Appl 18(1s, Article 40):1–16. Association for Computing Machinery, New York, U.S
13. Elswah DK, Elnakib AA, El-din Moustafa H (2020) Automated diabetic retinopathy grading using Resnet. In: 2020 37th National radio science conference (NRSC). IEEE, pp 248–254

Intrusion Detection System Using Supervised Machine Learning

Shubham Kumar and Khundrakpam Johnson Singh

1 Introduction

The Internet is a very important part of our life; we use the Internet every time. The data we use to travel is in the form of packets from the source to the destination [1]. Large-scale organization has their cybersecurity experts to dealing with crackers for protecting their organization's data but those organizations or businesses which are on small scale and simple people who are using the Internet and sharing their data on the Internet via a network through their systems, not all of them are aware of attacks caused by a person or group of persons. These attacks provide harm economically or mentally or both.

In 1999 [2], a fifteen-year-old boy with his own PC accessed and hacked the NASA system. His name is Jonathan James; he was showed himself as a software engineer to hinder data from the Defense Threat Reduction Agency. He approached more than 3,300 messages from DTRA representatives by accessing their usernames and passwords. This incident caused the loss of USD 41,000 for repair and made NASA shutdown their systems for 21 days. In 2004, according to the information or data information given by the global cybersecurity market, the 3.50 billion (https://www.wired.co.uk/article/job-security-cybersecurity-alec-ross) dollars was stolen by attackers. Ransomware casualties gave about 600 million dollars to cybercriminals or attackers in 2021. Many big-scale companies are also attacked by attackers like Amazon, Twitter, Yahoo, many banking sectors also, and so on. According to Global Cybercrime Damage Cost, their cybersecurity experts says that it will grow by fifteen % per year for the upcoming five years, reaching USD of 10.50 trillion dollars by 2025, its soaring from 3 trillion dollar in 2015 (https://www.microsoft.com/security/blog/2016/01/27/the-emerging-era-of-cyber-defense-and-cybercrime/) [2].

S. Kumar (✉) · K. J. Singh
Department of CSE, National Institute of Technology Manipur, Imphal, Manipur, India
e-mail: shubhamkumarnitmn@gmail.com
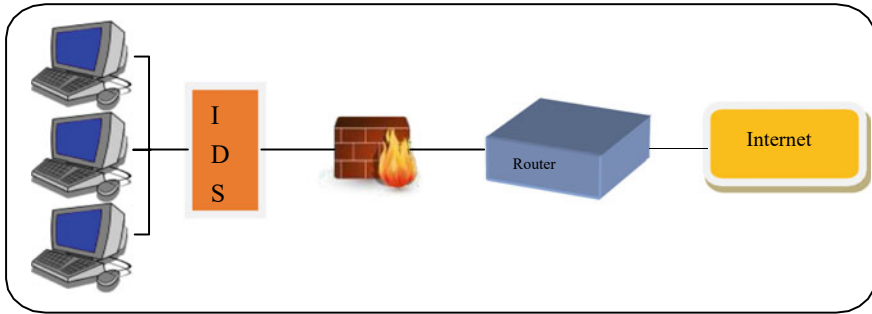
Fig. 1 Intrusion detection system

Attackers are capable from stealing money to launch a nuclear war.

Previously, firewalls and different strategies have been utilized for the security of information and were viewed as great however they are not sufficient to stop the attackers from taking the information [3]. "*Those persons or groups, who miss using the Internet for criminality, are called Crackers, and those persons who bypass the security for good intentions are known as a Hacker*".

In this paper, we learn about the system which is capable of detecting intrusion into the network, and this system is known as the intrusion detection system. So, before understanding the intrusion detection system, first, we must understand that intrusion. "An intrusion is an effort to bypass the defensive or protection methods of a computer system or the network" and "an intrusion detection system is the discovery or search system of the events passed in the network or the computer system" [3].

Intrusion detection system is shown in Fig. 1.

Intrusion detection systems are mainly separated in two types: signature-based intrusion detection and anomaly-based intrusion detection.

I used the KDD'99 cup dataset [4] because it is an effective benchmark for scientists to look at changed kinds of intrusion detection system (IDS) strategies, assemble an interruption discovery framework (host-based or network-based), and accomplish certain trials in the digital protection regions likewise, there are countless benefits. We used Pearson's correlation method for data filtering; we use because the KDD'99 dataset has non-numerical attributes also. So, this method is good for dealing with these types of datasets. Decision Trees, Support Vector Machine, and Logistic Regression are the classification algorithms used to make the model and make a comparison between them after performing actions to find the accuracy of the model.

2   Related Work

Many research happened after the 1980s, and few research that improves the detection of intrusion in a network are discussed.

Denning [5] proposed a prototype of a real-time intrusion detection expert system. The model is grounded on the postulation that safety violations can be distinguished by covering a framework's review records for strange patterns of framework activity. Her model contains data for showing the kind of subjects as for the objects in conditions of metrics and statistical model, and rules for acquiring information about this nature from test records and for identifying different behavior. In her hypothesis, the model is autonomous of any certain computer system, submission environment, system liability, or type of intrusion or attack, thereby requiring a framework for a general-purpose intrusion detection expert system.

Sommer and Paxson [6] examine and propose that there is a wide volume of research happened on anomaly-based detection in the theoretical intrusion detection group, but operational deployments of such systems happen in the real world is very less. They gave certain tasks that come with deploying the intrusion detection systems by using machine learning are

- Outlier revealing is needed, while machine learning rather performs better at tracking down similitudes.
- High-level costs for classification mistakes, which consolidate mistake rates as experienced in different areas far-fetched.
- Gap concerning detection results and their operation.
- Giant inconsistency of nonthreatening traffic, making it hard to look through stable ideas of ordinariness.
- Noteworthy tasks with performing evaluation.
- Execute in argumentative situation.

Phadke et al. [7] found in their works that static detection techniques only capture the known pre-founds attack. They also require continuous updates in signature-based databases. So, they used machine learning methods for developing the intrusion detection system, which is capable of detecting intrusion in the network. Based on their work and evaluations, they made it, and they can be affirmed that the ability of machine learning could have great potential in Network Intrusion Detection. In their paper, the best methods were intended by Chowdhury et al. [8], the design part of the Artificial Neural Network was done by Shenfield et al. [9]. By using the UNSW-NB15, dataset accuracies were observed on the machine learning models trained was best in their case.

Narayan and Chatterjee [10] proposed an examination of various kinds of intrusion detection systems is done close with the help of many examination papers which have used AI, DNA grouping, design coordinating, and information digging as a technique for learning assaults and taking preventive exercises when practically identical kinds of assaults are capable later. Assessment of these papers has given a piece of huge data

to likewise explore the connected techniques in the field of interruption recognition frameworks.

Moradi and Zulkernine [11] introduced IDS that proposes ANN for successful interruption discovery. One of the hindrances of their approach is that it grows the planning time. Sarasamma et al. [12] proposed in their paper the multilevel hierarchical Kohonen networks that are used to find intrusions in networks. In their work, they subjectively choose some pieces of information from KDD Cup 99 to prepare and test the classifier.

Bivens et al. [13] proposed in their work network-based intrusion detection using a neural network methodology. They gave the idea measured network-based interruption location framework that dissects information, and a portion of its parts should be prepared.

3 Feature Extraction

The main aim of this paper is to make a comparison between numbers of classification algorithms and apply them to the dataset to predict which model provides us with the best accuracy with low error. There are a number of steps involved in executing a machine learning algorithm based on the classification procedure using a dataset.

3.1 Dataset

In this paper, we used the KDD'99 dataset [13] because this dataset has a good benchmark for the research point of view.

We add one more column in the dataset, which holds the record of attacks according to the different types of attacks. We have 21 types of attacks in this dataset; these attacks fall into four different categories [14] such as

- Denial of Service (DOS): A DoS attack is a type of Internet attack in which the attacker restricts access to the approved individual into the framework or organization. DoS assault accomplishes this by flooding the objective framework with traffic or sending it information that sets off a framework disappointment or framework crash, e.g., syn flood, smurf.
- Remote to Local (R2L): It is a type of computer attack in which attackers transfer some set of packets to another server or computer network where the attacker does not have authentication to access as a local system user, e.g., guessing passwords.
- User to Root (U2R): In this type of attack, attackers are double-dealings in which the programmer gets going on the framework with an ordinary client record and endeavors to abuse shortcomings in the structure to secure super client respects (Gain root access), e.g., various "buffer overflow" attacks.

- Probe: A test is an assault wherein the programmer channels a machine or an organization to assemble data for shortcomings or track down known weaknesses, e.g., port scanning.

The training set contains 41 columns that are used to train the data, and the test set contains 1 column for showing the test result data. Steps for making our input data ready for modeling.

Step: (1) The KDD'99 dataset stores the data in a common separated value file format. To read this type of file and perform analysis on it, we must import the "Pandas" library of Python because it gives ready-to-use high-performance data structure.

Step: (2) We get the KDD'99 dataset in a CSV file, which has 4,898,430 data in rows and 41 attributes in columns.

Step: (3) We assign attribute names to each column and add one column at last and the name that the last column is the target.

Step: (4) We use the target column to create or add another attribute or column, which is called attacks_types, whose value only contains four distinct attack names, i.e., DoS, Remote to Local, User to Root, and Probe.

3.2 Feature Selection

This dataset is still impure. All these columns are not necessary to create a model therefore; we must remove or drop those columns which are not useful for our model development. To create a better model, we must first apply some processes to the data to remove unnecessary things from the dataset. Figure 2 represents the feature selection way by removing noise from the dataset. So, we check that the dataset contains a NULL value; if it has a null value, then we drop that null value. In our paper, filter method is best fitted.
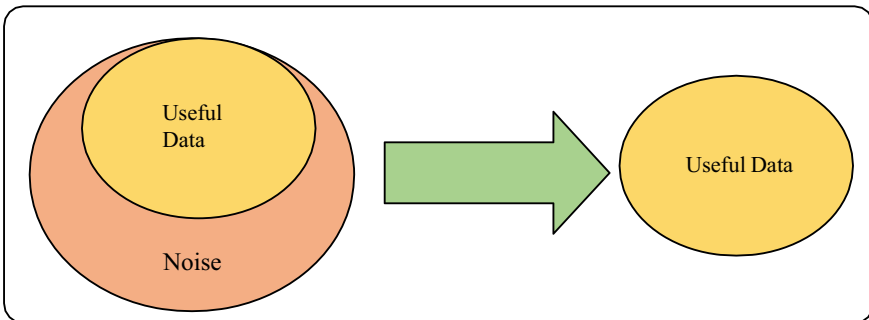
Fig. 2 Feature selection

In the filter method, features or attributes are removed by knowing the output result's relation with the attributes column values, or we say how they are important to the output result. We use to check the relation to know that whether the features are positively or negatively correlated to the result labels and according to that, we drop those features accordingly. Some feature selection methods are Pearson Correlation, information gain, Fisher's Score, etc.

The next step is to decide which feature selection technique is best fit for our model. This step is easy to do, by knowing that what the output desire is.

Basically, the dataset variable is found in two types. If the variable or attribute values of the dataset are integers, float, or in numeric form, then it is known as Numerical Variables, and if the attribute values include strings, Boolean variables than it is known as Categorical Variables. Table 1 represents the models for feature selection according to the input values we have and the output which we want.

We use supervised learning in this project, and the data is the KDD'99 dataset which has features that contain numerical values and categorical values and linear values also, but the number of numerical features is more than categorical values. So, in filtering the attribute in the best way that takes the model perfectly, we use Pearson's correlation coefficient to filter the dataset because the dataset is in linear form.

Relation or we say connection between individual attributes to the output is known as correlation of that attribute and output result. Utilizing this relationship, we can foresee one component from the other element. The idea driving involving a correlation coefficient for feature selection is that the helpful elements profoundly correspond with the objective. Also, highlights ought to be related to the result target however ought to be uncorrelated alongside themselves. In the event that two features of the dataset are related to one another, we can foresee one feature or attribute from the other feature. Hence, in the event that two factors are connected, the model just actually needs one of them to feature, as the subsequent one contains no more essential data which will help in making the model.

Pearson's correlation coefficient is also referred to as "Pearson's r". Pearson's correlation coefficient has utilized the figure out the relationship or association between two elements or information or amounts. It brings about the proportion of the strength of the collision between two factors. The value ranges from -1 to $+1$ of Pearson's correlation coefficient.

After finding the highly correlated data from two features, we just drop one of the features because both give the same target output. To find the high correlation

Table 1 Feature selection model

Input variable	Output variable	Model
Numerical	Numerical	Pearson's correlation, Spearman's rank
Numerical	Categorical	ANOVA (linear), Kendall's (nonlinear) correlation
Categorical	Numerical	ANOVA (nonlinear), Kendall's (linear) correlation
Categorical	Categorical	Chi-Squared test, mutual information

coefficient, we take some threshold values like 0.99 and 0.98 in positive and negative both are treated as $+1$ and -1, respectively.

Our dataset contains some non-numerical features or data in qualitative forms, i.e., the dataset contains text features also. And the computer system cannot recognize a feature in text format so it cannot process that data. So, it is compulsory to convert categorical feature data into a numerical data format. We have assigned some default values which must be numerical like 0 and further as the variance of 1 to all the attributes in the dataset. In our dataset, two features, i.e., "protocol_type" and "flag" are non-numeric attributes. The values of "protocol_type" are "icmp", "tcp", "udp" mapped by using the map function of Python as 0, 1, 2, respectively. And the values of the "flag" feature are SF, S0, REJ, RSTR, RSTO, SH, S1, S2, RSTOS0, S3, OTH assigned by the value range from 0 to 10, respectively.

4 Classification Models

For classification models, we need to separate the dataset into two sections, one section is used for training the model, and the other section is for testing the dataset for calculating the accuracy. To split the data, we must use Python's machine learning library, i.e., "Scikit-learn". In this library, there is a predefined test train split function that helps our model for training and testing in some ratio or we say in percentage.

In this paper, three different classifications are used for creating models: Decision Tree, Support Vector Machines, and Logistic Regression.

4.1 Decision Tree

A Decision Tree is a supervised learning which is used for classification as well as regression problems, but it is mostly good for classification problems. Idea behind this method is similar like human mind at the time of taking a decision. So it is very easy to understand the logic of this method. In this method, firstly the dataset is provided to the root node for the prediction. After that, the algorithm or method starts comparing the feature or attribute value of the root node with the record attribute. According to this process, comparisons happen and the process moves to the next node. The significant piece of the choice tree is to choose the trait to obtain a precise outcome. This can perform by using two popular techniques, which are information gain and Gini Index.

Information gain is the quantity of change in entropy after the dissemination of a dataset in view of a characteristic. It gauges how much information a component gives us about a class. Then, we start building the Decision Tree by splitting the node according to the class value. Information gain is a difference of entropy of complete dataset and entropy of that particular feature. It can be determined utilizing the Eq. (1).

$$IG = \text{Ent(Sp)} - \big[(\text{Weighted Avg}) * \text{Ent(per feature)}\big] \qquad (1)$$

where IG stands for information gain, Ent stands for entropy, and Sp stands for samples of data.

4.2   Support Vector Machine

Support Vector Machines are widely used supervised machine learning algorithm because it is very powerful and performed much better for regression and classification problems. At starting time of SVM, it is only used to accept and solve classification problems. SVM tries to find a hyper plain between the data to differentiate the data into some parts. And after that, it creates two parallel lines on both sides of hyper plains known as margins. There are two types of data we get, linear separable data and nonlinear separable data. Linearly separable are those data which can be separated easily by using the straight line on the graph.

4.3   Logistic Regression

Logistic Regression belongs to the category of supervised learning technique. It is used for anticipating the clear-cut subordinate's variable utilizing a given dataset of free factors. The result or consequence of this calculation is a bool or Boolean sort like Yes or No, 0 or 1, true or False, and so forth however rather than giving the specific worth in mathematical structure as 0 and 1, it gives the probabilistic kind qualities which values go from numeric 0 to numeric 1.

It is a substantial machine learning algorithm because it has the competence to assign probabilities value and organize new data using continuous and discrete values of datasets.

Logistic Regression can be applied to group the perceptions utilizing various kinds of information and can undoubtedly decide the best factors utilized for the order. Sigmoid function is a very important term in Logistic Regression; it basically maps different types of values into the range of numerical 0 to 1. If the value lies below the threshold value, then it shifts that to 0 and those values which occur above the threshold value, basically shifted to 1.

5 Experimental Results and Discussion

In this section, the experiment is performed to validate the performance of all the algorithms. In the first phase, the KDD'99 dataset was provided to all three machine learning algorithms, and in the second phase, we analyzed which machine learning algorithm has the best accuracy for our dataset.

In the experiment, we take 10% dataset from the actual dataset so 10% of actual data, i.e., 494,021 rows of data in the experiment set, and then we split the dataset into percentages, i.e., 67% for training and 33% for testing. 67% of 31 feature data were used for the training of the model and named "X train" and 67% of 1 feature, i.e., attacks_types data used for the training of the model named "X test". Now 33% of 32 feature data were used for the testing of the model and named "Y train" and 33% of 1 feature, i.e., attacks_types data used for the testing of the model is named "Y test".

Around 330,994 rows of data are used for training and 163,027 rows of data are used for testing.

5.1 Decision Tree

- In the beginning, we give the whole training set of data to the root.
- For set the criterion by entropy method and give the maximum_depth of the tree is 4.
- Based on attribute values present in the top node (root node), elements of the dataset split.
- We utilize measurable philosophy for requesting credits esteem as a root node or the nodes present in internal.
- Our main aim is to get the final decision which is present in the leaf node of the tree. We perform multiple time operations to get the leaf node of every branch.

5.2 Support Vector Machine

During this calculation, the value of the boundary of gamma is scale. The gamma boundary portrays how detached the effect of one preparation accomplishes, with low qualities addressing "far" and high qualities addressing "close". The contrary of the compass of influence by the support vectors so that we use gamma for a boundary.
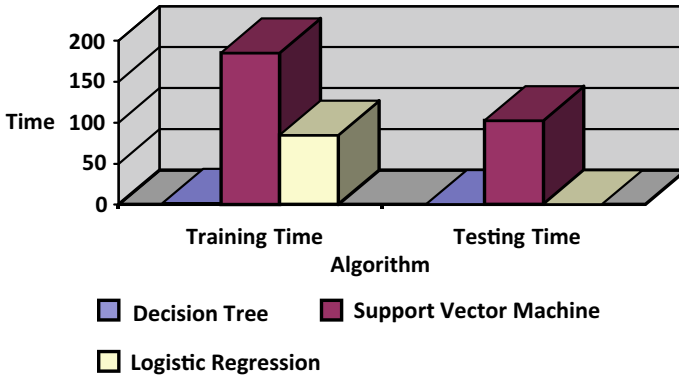
Fig. 3 Computational train–test time graph

Table 2 Comparison of various approaches for intrusion detection system using machine learning

Model	Training time	Testing time	Train score (%)	Test score (%)
Decision tree	1.8086	0.0927	99.0582	99.0523
Support vector machine	184.7094	102.1106	99.8755	99.8791
Logistic regression	84.9530	0.0688	99.352858	99.352867

5.3 Logistic Regression

Logistic Regression first scales the value of the attribute. After creating a class of Standard Scaler (https://scikit-learn.org/stable/modules/generated/sklearn.preprocessing.StandardScaler.html), we calculate (fit) the mean and standard deviation for scaling.

Figure 3 represents the time taken for training and testing of all three models.

Table 2 shows the records of each machine learning algorithm with their training time, testing time, train score, and test score data.

6 Conclusion

Intrusions in networks are a major problem because it violates security and privacy. Since 1980 a lot of work has been done in the field of IDS and it's still active since crackers continue to develop new methods of network attack. Due to this, the victims are an individual persons or large- or small-scale industries facing the problems of economic or mental loss of their money.

In this paper, we work on an anomaly-based intrusion detection system in the network by using three different machine learning techniques. We discuss all three machine learning algorithms briefly. Each algorithm has its own methodology for the

improvement of IDS, Decision Tree takes 1.8086 for training and 0.0927 testing time and gives 99.0582% train score with 99.0523% test score. Support Vector Machine takes 184.7094 for training and 102.1106 for testing time and gives 99.8755% train score with 99.8791% test score. And Logistic Regression takes 84.9530 for training and 0.0688 testing time and gives 99.352858% train score with 99.352867% test score. When we do a comparison between these algorithms, we found that the Decision Tree takes less training and testing time. Support Vector Machine has the highest accuracy as we compared it with Decision Tree and Logistic Regression.

For our KDD'99 dataset, Decision Tree is good since it takes less time, and Support Vector Machine is better because it is more accurate.

References

1. Senthilnayaki B, Venkatalakshmi K, Kannan A (2014) Intrusion detection system using feature selection and classification technique. Int J Comput Sci Appl (IJCSA) 3(4):145–151
2. Cybercrime to cost the world $10.5 trillion annually by 2025. Retrieved from https://cybersecurityventures.com/hackerpocalypse-cybercrime-report-2016/. Accessed on 20 Mar 2022
3. Haripriya L, Jabbar MA (2018) Role of machine learning in intrusion detection system: review. In: 2018 Second international conference on electronics, communication and aerospace technology (ICECA). IEEE, pp 925–929
4. KDD cup 1999 data. Retrieved from http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html. Accessed on 7 Nov 2021
5. Denning DE (1987) An intrusion-detection model. IEEE Trans Softw Eng SE-13(2):222–232
6. Sommer R, Paxson V (2010) Outside the closed world: on using machine learning for network intrusion detection. In: 2010 IEEE Symposium on security and privacy. IEEE, pp 305–316
7. Phadke A, Kulkarni M, Bhawalkar P, Bhattad R (2019) A review of machine learning methodologies for network intrusion detection. In: 2019 3rd International conference on computing methodologies and communication (ICCMC). IEEE, pp 272–275
8. Chowdhury, Nasimuzzaman M, Ken F, Ferens M (2016) Network intrusion detection using machine learning. In: Proceedings of the international conference on security and management (SAM). The steering committee of the world congress in computer science, computer engineering and applied computing (WorldComp), pp 30–35
9. Shenfield A, Day D, Ayesh A (2018) Intelligent intrusion detection systems using artificial neural networks. ICT Express 4(2):95–99
10. Narayan G, Chatterjee JM (2018) A review on intrusion detection system and various attacks on network. Int J Res Appl Sci Eng Technol (IJRASET) 6(VI):46–50
11. Moradi M, Zulkernine M (2004) A neural network-based system for intrusion detection and classification of attacks. In: Proceedings of the IEEE international conference on advances in intelligent systems—theory and applications, vol 148. IEEE, Luxembourg, pp 1–6
12. Sarasamma ST, Zhu QA, Huff J (2005) Hierarchical kohonen net for anomaly detection in network security. IEEE Trans Syst Man Cybern Part B Cybern 35(2):302–312
13. Bivens A, Palagiri C, Smith R, Szymanski B, Embrechts M (2002) Network-based intrusion detection using neural networks. In: Proceedings of the intelligent engineering systems through artificial neural networks ANNIE-2002, St. Louis, MO, vol 12, ASME Press, New York, NY, pp 579–584
14. Singh KJ, De T (2017) Analysis of application layer DDoS attack detection parameters using statistical classifiers. Internetworking Indonesia 9(2):23–31

Examining Bioactivity of Medicines in Twenty-First Century Smart Society 4.0: An Approach with ML and DS

Rohit Rastogi, Yash Rastogi, Saurav Kumar Rathaur, and Vaibhav Srivastava

1 Introduction

The development of new medications is a tedious and exorbitant cycle. For sure, to guarantee both the patients' security and medication viability, planned drugs should go through a cutthroat and long system. Drug advancement is generally parted into four significant stages, called stages. Basic research, drug disclosure, and preclinical tests to assess the effectiveness and physiological effects of the rival drug make up stage 0. The last three stages are known as clinical primers, and they are Stage I (which includes analyzing part harm, transitory secondary effects, and dynamic associations), Stage II (which confirms that the prescription is being taken as recommended), and Stage III (which looks at the part to the norm of care) (Stage III). Pharmacological publicizing that feature enduring unfriendly impacts and medication collaborations with different treatments may alternatively incorporate a Phase IV.

2 Literature Review

This section finishes up for certain broad points of view on contemplations and requirements for additional exploration in light of this report.

Pati et al. (2020) and the group utilized the Medication Bank and Super Target Dataset to make an AI model. The researchers anticipated that their model would outperform the existing calculations in its ability to predict test information findings with high precision. The model's long-term use is that whenever an expert mixes

R. Rastogi (✉) · Y. Rastogi · S. K. Rathaur · V. Srivastava
Department of CSE, ABES Engineering College, Ghaziabad, Uttar Pradesh 201009, India
e-mail: rohitrastogi.shantikunj@gmail.com

up another tiny heterocyclic particle, they may use our calculation to predict the likelihood that the atom will be a drug [9].

The author discovered a lab bunch database on E. Coli (which causes cholecystitis, urinary tract infection (UTI), loose bowels, and pneumonia) and HIV apart from SCRIPPS and Wide Organization public databases. The author and his colleagues discovered a tool developed by a team at UC Irvine that simply converted a 3D particle addressed in its Grin depiction into the number of different valuable clusters it includes. We have to really make it happen as the author and their team never intended for it to coordinate in our code [3, 5, 10].

Author and their team discovered two datasets comprising the Grin's of various atoms and their motion in preventing HIV infection and E. Coli microscopic organisms. The device couldn't instantly handle the large number of atoms; therefore, the author transformed all of the Grin particles into their corresponding number of useful gatherings in bunches. The author and his colleagues implemented common clustering strategies like support vector machines for the machine learning model (SVM), ANNs, as well as deep learning algorithms [4, 7, 9].

Manne, R., researched the finding of different ml approaches tools, and techniques in drug discovery, by analyzing the importance of clinical value and their combination of compounds in drug. Different seven phases can be used in designing the phase of a drug compound. Identifying the target can be the very first approach to understanding misfolded proteins, potential diseases, and DNA mutations. It is essential that proteins are present in any creatures and are responsible for critical functionality in any being. Protein capabilities are subject to 3D design, the usefulness of protein can be changed, and this is a significant calculate drug improvement matter. As per the exploration, the creator has put out a framework that utilizes profound figuring out how to find and gauge drug target co-operations. This profound learning technique for DTI expectation utilizes primary information on the two synthetic compounds and proteins [6].

In general, the goal is to find the most effective arms with the highest average reward earned through repeated arm selection. Because the CGA can be used to create a serial suggesting model, the recommender model and CGAs are inextricably linked [2, 8].

3   Methodology and Setup of Experiment

Author and his gathering team used the ChEMBL data base which contains different line and portion regards (the certifiable condition of the informational collection is 4695 lines, 882 areas),g but it has taken the relevant qualities values like raising ligands and protein. In model creation, backslide computations were used like irregular backwoods, decision tree, QSAR, YSquare, and significantly more AI estimations. While differentiating these computations and checking the precision of the model, the QSAR model gave a respectable accuracy which is perfect for the proposed assumption model.
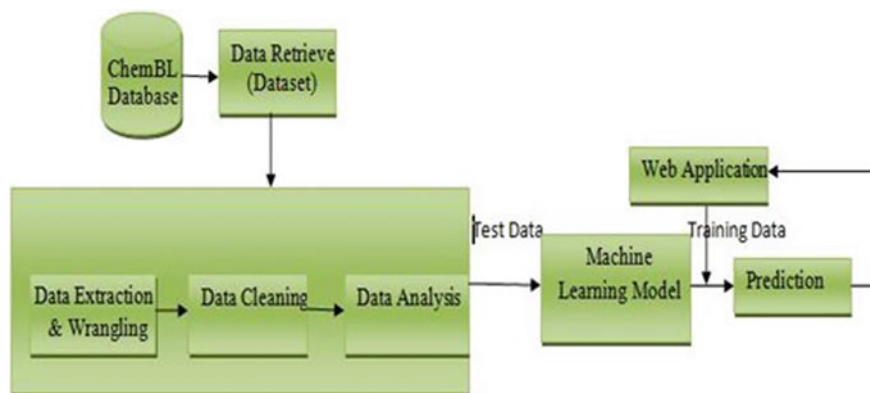
Fig. 1 Drug recognition interaction and model

Similarly, the creators recently cleaned the data, found the missing characteristics and stayed aware of these missing characteristics and played out some Exploratory Information Investigation (EDA) on it, analyzed the dataset, appreciated the dataset, found some connection between the data credits, and subsequently made the model (acc to Fig. 1).

3.1 Name of Algorithm Used

In the given model, the author tries to do some critical analysis and apply linear regression as well as clustering algorithms like Random Forest, Lasso Regression, Ridge Regression, Decision Tree in regression, and classification or Support vector machine in clustering algorithm.

3.2 Different Branches of Database

This substance dataset contains the data of synthetic mixtures utilized in Medication Identification is ChEMBL. The ChEMBL assortment contains microscopic particles that look like bioactive medications, that include 2D plans, decided properties (as Lipinski factors, Molar Mass, LogP, etc.), and engrossed bioactivities (for instance limiting constants, pharmacology, and ADMET information). It is a gathering of truly coordinated natural particles with drug-like properties. To help in the change of genomic information into usable restorative drugs, it consolidates compound, bioactivity, and hereditary information. The ChEMBL dataset contains information coordinated bioactivity data of 2 million combinations.

3.3   Dataset

The expert gathering through searched in the web and found different game plans of datasets some of them are referred to beneath:

1. ChEMBL Information base
2. UCI ML Medication Audit Dataset
3. QM9 drug disclosure dataset (QM9 gives quantum synthetic properties).
4. DrugBank
5. SuperTarget.

However, the assessment bunch included the ChEMBL Informational collection of drug compound information as it is allowed to download. Maker filtered the informational index and helped their appropriate information through this. So they made an amount of 2818 sections and four segments of educational assortment, which the maker called Bioactivity dataset.

3.3.1   ChEMBL Origin and URL

! pip install chembl_webresource_client

Using above line of Python code, the Creator presented the ChEMBL web organization pack.

from chembl_webresource_client.new_client import new_client

Getting the above library, the Creator used the ChEMBL informational collection and isolated the dataset according to their model.

3.3.2   Flowchart

The workflow for bioactivity drug acknowledgment of combinations is shown in the above sketch. The creator of the offered image illustrated how each and every task uses the informational index and brought the important example data through this. In the aforementioned stream chart, the creator illustrated each stage of their work. (as per Fig. 2).

4   Result and Discussions

4.1   Implementation of Model and Analysis

Examination among bioactivity and sub-atomic load of synthetic compound.
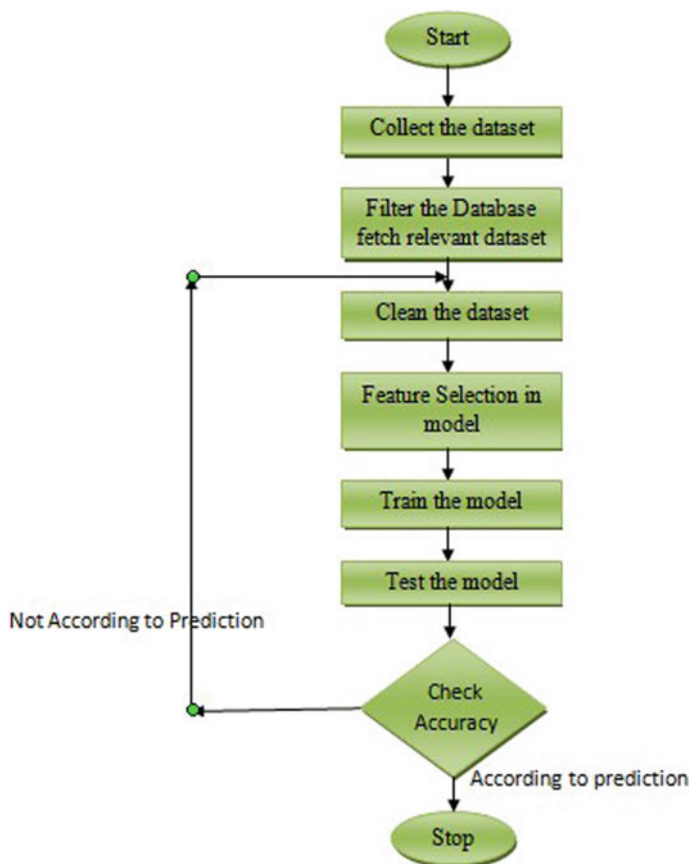
Fig. 2 Flowchart for bioactivity identification of medication compound

In the above given outline, we are taking apart the bio-movement of a substance compound to its nuclear weight. Maker brings the sub-nuclear heap of Harmed from the rdkit library. The blue and orange tone exclusively showed dynamic and Dormant IC50 values. It might just be seen that the two bioactivity classes are crossing tantamount manufactured spaces as clear by the scattered plot of MW versus LogP (bioactive IC50 regard) (as per Fig. 3).

4.2 Lasso Regression

See Fig. 4.
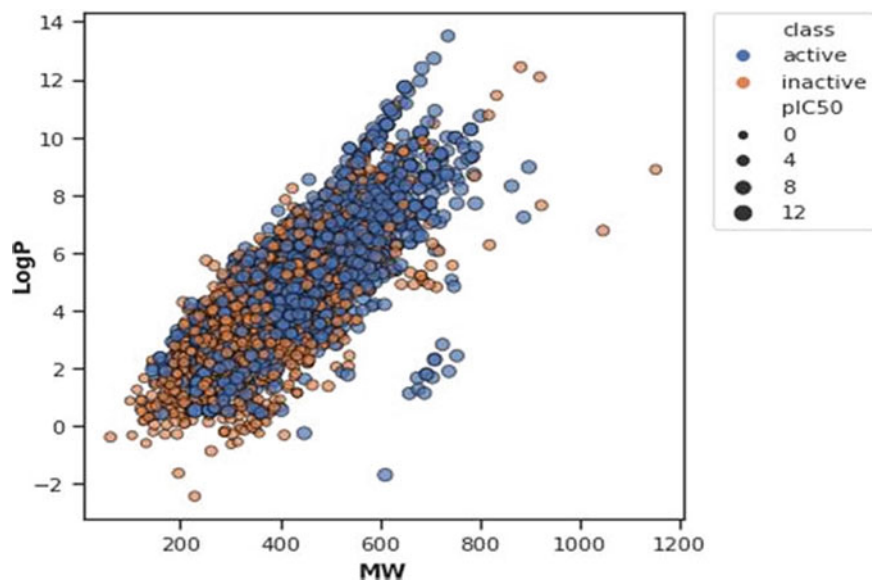
Fig.3 Comparison between LogP and MW
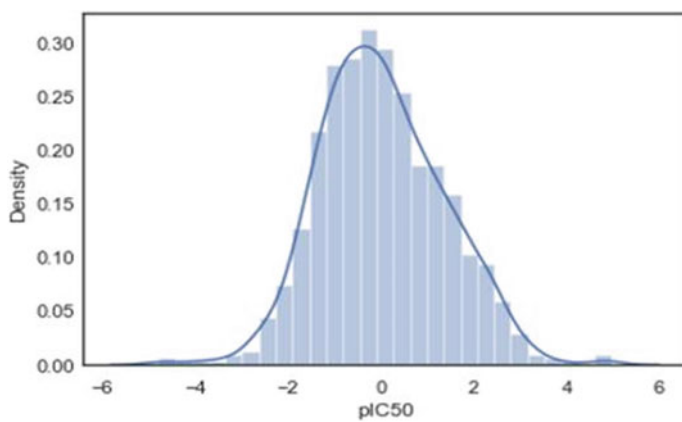
Fig.4 Implementation of Lasso Regression

4.3 Ridge Regression

See Fig. 5.
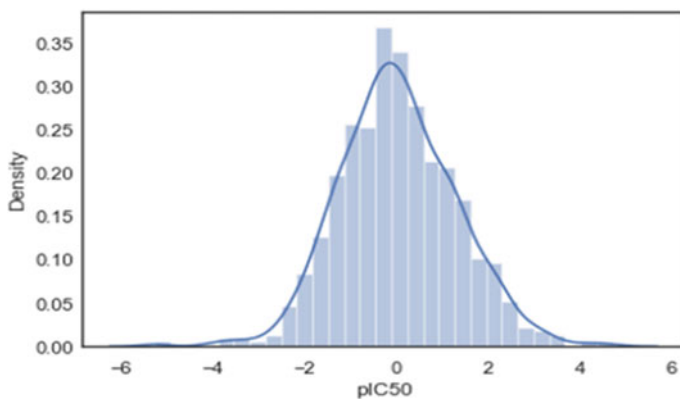
Fig.5 Implementation of Ridge Regression
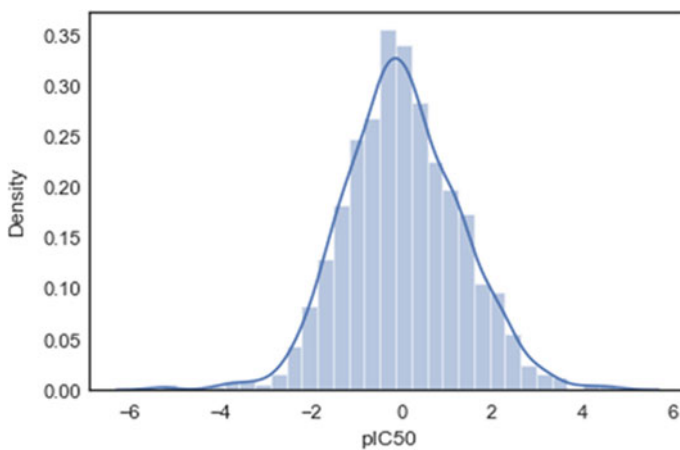
Fig. 6 Implementation of Linear Regression

4.4   Linear Regression

See Fig. 6.

4.5   Random Forest

See Fig. 7.
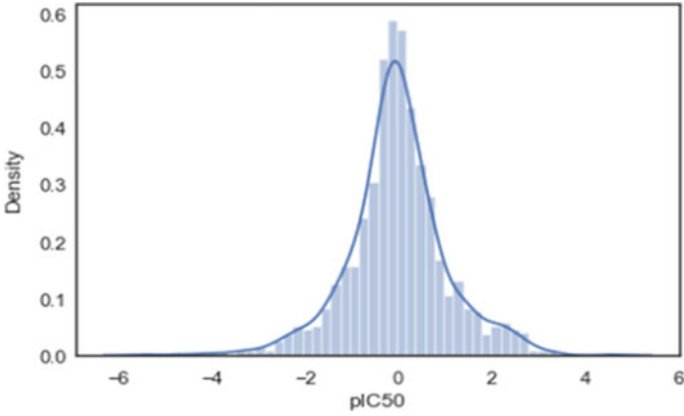
Fig. 7 Implementation of Random Forest
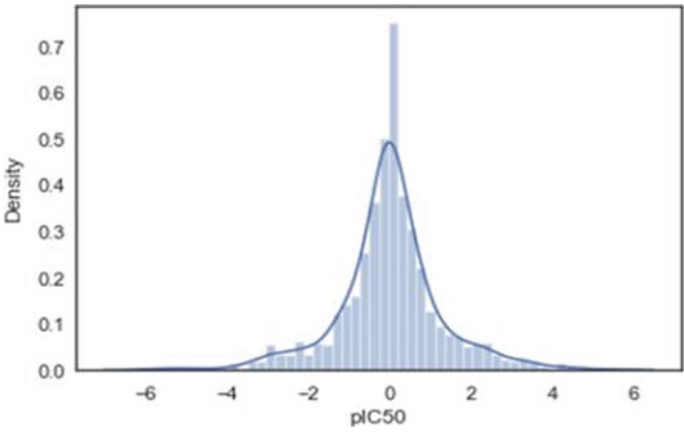
Fig.8 Implementation of Decision Tree

4.6 Decision Tree

See Fig. 8.

4.7 Knn

See Fig. 9.
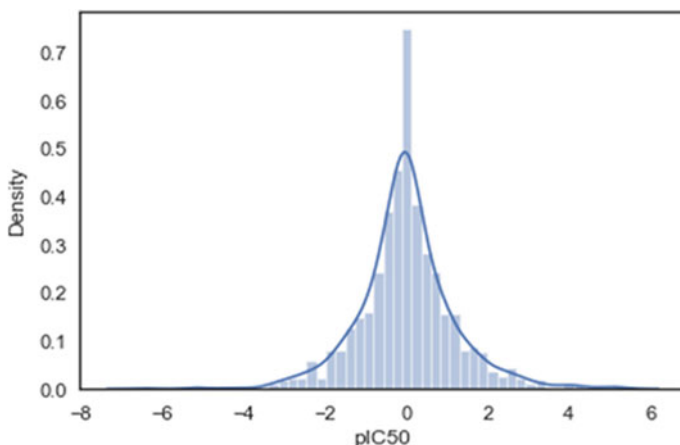
Fig.9 Implementation of KNN

4.8 Algorithm Comparison

To improve QSAR models, a dataset collection with 2570 combinations was used. To discover the top performing set, twelve game plans of extraordinary finger impression descriptors were benchmarked. Feature assurance was used before the presentation to do away with collinear descriptors. Then, all twelve models were created utilizing an 80/20 data split, where 80% of the illuminating list was used for the internal set and 20% for the exterior set. All 100 independent data points from the technique were utilized to predict the course of events in iterative fashion, and the findings are presented here together with the mean and standard deviation figures that were determined from these iterations (Table 1).

A combinational classifier called irregular woods (RF) is created from a couple of Choice trees. The principal idea driving RF is that various trees are developed to restrict change as opposed to supporting precision, rather than building a critical Choice tree with a solid making number of focuses, which might be at risk for overfitting and overtraining of the information. At the point when results stray from an impeccably coordinated decision tree in that cutoff, the outcomes will be noisier,

Table 1 Data of implemented model accuracy

	r2_score	name
0	0.523966	Random
1	0.418336	KNN
2	0.295332	Ridge
3	0.270133	Linear
4	0.266675	Dtree
5	0.240358	Lasso
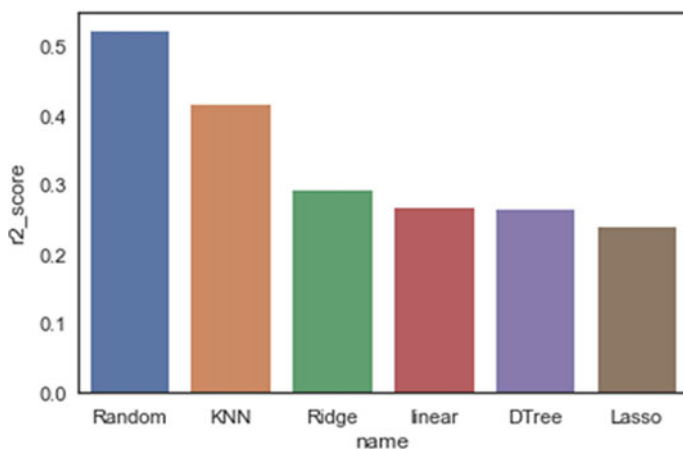
Fig.10 Implemented model accuracy comparison

yet they are ordinarily powerful and generous. This is a fast execution of the RF assessment that was utilized in the model development (as per Fig. 10).

5 Novelties

In regard to the disclosure of small molecule drugs, artificial intelligence can do so in four different ways: by opening doors to new, better, or novel fields of knowledge, by improving achievement rates, and by facilitating quicker and less expensive disclosure procedures. Many challenges and requirements in conventional research and development can be addressed by the advancement. Each application offers additional information to calm down revelation groups, and occasionally it may even be possible to reclassify deeply ingrained work routines. Understanding and differentiating between use cases is essential because these advances are applicable to a variety of revelation settings and typical targets.

6 Future Directions

We accept that the results of our model can be used to the models created in our reference papers, increasing the accuracy of discovering a potentially effective new drug. In the event that we expand the dataset, we will then be able to obtain a better-prepared model.

7 Conclusion

The formation of QSAR models used a dataset of 2570 blends. Twelve single unique finger impressions descriptors specifically have been adjusted to perceive the most effect on the fruitful of 80/20, where 80% of the enlightening list was utilized as the internal set and 20% as the external set. The standard deviation and mean qualities acquired from these runs are utilized to give the show results. Every one of the 100 independent snippets of data was used separately in this strategy's iterative advancement of models.

References

1. Balganesh T, Kundu TK, Chakraborty TK, Roy S (2014) Drug discovery research in India: current state and future prospects. ACS Med Chem Lett 5(7):724–726. https://doi.org/10.1021/ml500183c
2. Deng J, Zhibo Y, Ojima I, Samaras D (2011) Artificial intelligence in drug discovery: applications and techniques. Elsevier Pub Health Emer Collection 26(1). https://doi.org/10.1093/bib/bbab430
3. Ding X, Zhang B (2022) Contrastive learning of coarse-grained force fields. J Chem Theor Comput 18(10):6334–6344. https://doi.org/10.1021/acs.jctc.2c00616
4. El-Behery H, Attia AF, El-Fishawy N, Torkey H (2021) Efficient machine learning model for predicting drug-target interactions with case study for Covid-19. Comput Biol Chem 93:107536. https://doi.org/10.1016/j.compbiolchem.2021.107536
5. Henstock VP (2019) Artificial intelligence for pharma: time for internal investment. Sci Soc Spec Issue Rise Mach Med 40(8). https://doi.org/10.1016/j.tips.2019.05.003
6. Manne R (2021) Machine learning techniques in drug discovery and development. Int J Appl Res 7(4, Part A). https://doi.org/10.22271/allresearch.2021.v7.i4a.8455
7. Mohapatra S, Nath P, Chatterjee M, Das N, Kalita D, Roy P, Sethapathi S (2020) Repurposing therapeutics for COVID-19: rapid prediction of commercially available drugs through machine learning and docking. https://doi.org/10.1371/journal.pone.0241543
8. Patel V (2022)., Artificial intelligence and machine learning in drug discovery and development. Intell Med 2:134–140. https://doi.org/10.1016/j.imed.2021.10.001
9. Réda C, Kaufmann E, Delahaye-Duriez A (2020) Machine learning applications in drug development. Comput Struct Biotechnol J 18:241–252.https://doi.org/10.1016/j.csbj.2019.12.006
10. Schmid EF, James K, Smith AD (2001) The impact of technological advances on drug discovery today. Drug Inf J/Drug Inf Assoc 35:41–45. https://doi.org/10.1177/00928615010350010

A Review on Automatic Assessment and Detection of Pathological Speech

Ashita Batra and Pradip K. Das

1 Introduction

Speech is a highly complex process that necessitates synchronizing several muscle groups involved in articulation, airflow direction, respiration and laryngeal function, which results in speech hindrance in various ways due to illnesses induced by hereditary factors, physical abnormalities, loss of hearing or neurological malfunction. PS usually indicates in the context of speech distortion resulting from abnormalities in voice or in the articulatory structure as reason of physical injury, illness or other biological injury to the production system. In order to provide accurate clinical diagnoses of speech pathologies and therapeutic feedback, it is essential to monitor speech changes as they provide important information about the pathology and its severity. Furthermore, speech analysis can provide an early indication of neurological disease progression [1]. There are many types of alterations that may occur in PS due to vocal disorders and disruptions of vocal processes caused by illness and disease. Different locations and sizes of tumours in the head and neck affect speech signals in different ways. Speech pathology refers to the study of speech production abnormalities that influence aeromechanical, fluency, phonation and intonation components of respiration. The American Speech-Language-Hearing Association is a professional organization dedicated to the advancement of communication abilities (ASHA) [2] classifies pathological disorders into five categories namely (i) language disorder, (ii) speech disorder, (iii) social communication disorders, (iv) cognitive communication disorders and (v) fluency disorder. For instance, dysarthria or dysphonia (a motor speech disorder) affect a subject's speech, reducing its understandability by others

A. Batra (✉) · P. K. Das
Indian Institute of Technology Guwahati, Guwahati, Assam 781039, India
e-mail: b.ashita@iitg.ac.in

P. K. Das
e-mail: pkdas@iitg.ac.in

[3]. As a result, speech therapies are often conducted for the subject. Speech therapy frequently helps to restore lost speech intelligibility. But the outcomes of intense therapy aren't always noticeable. However, intelligibility assessments are carried out before and after-therapies where "Intelligibility refers to how well a speaker's acoustic signal can be accurately recovered by a listener a.k.a determined by high ratio of number of words correctly interpreted". These assessments are done by SLPs, which is a subjective evaluation task and hence dependent on therapist's perspective as well as time-taking. Hence, the need of automatic assessment of pathological speech arises. We shall discuss the following (i) challenges in pathological speech, (ii) feature selection methods and (iii) approaches used for automatic assessment of pathological speech. By reviewing the following topics, a basis of the assessment of pathological speech shall be formed from the various state of the arts.

1.1 Background

Whenever we want to track the severity of the disease in case of PD, multiple times Unified Parkinson's Disease Rating Scale (UPDRS) is exploited. But the major drawback being that it requires that patient to be present in clinic physically, which is a time-taking task. Therefore, Tsanas et al. [1] developed a remote application by performing subjective tests at home using a LCD screen, afterwards sending the collected data using Internet to the clinic. He exploited LASSO for feature selection mapping them with linear and nonlinear regression techniques. A study done by Karan et al. [4] which helps finding the severity of the PD along with the level of dysarthric speech in the subject. He first exploits hilbert spectrum (HS) which extracts voice impairments then applying Instantaneous energy deviation cepstral coefficient (IEDCC) for feature extraction. Hence, mapping these features with regression techniques. A similar work done by Bocklet et al. [5] but he checked the severity of PD in a subject instead of dysarthric level speech. He separately made it as 2-class problem (whether a subject has PD or not) and 3-class problem (severity of PD in a patient). It is always beneficial if any disease is caught in its early stages only. A work based on this is done by Mallela et al. [6] by performing early diagnosis of PD and Amyotrophic Lateral Sclerosis (ALS) and making it a 3-class problem (ALS/PD, ALS/HC, PD/HC). The concept of data-driven approach is used. 1-D CNN is exploited for retrieving the representations from raw speech and then using BLSTM for classification purposes. There is always a scarcity of databases when it comes to pathological speech irrespective of diseases. On this basis, a work has been proposed by Gillespie et al. [7] who has used the concept of cross databases modelling in which testing is done on one dataset whereas training is done on another dataset. This helps in getting good results as every dataset has its own characterstics according to subject. Much work has been in intelligibility assessment as evaluation and assessment of pathological speech go hand in hand. Wallen et al. [8] performed a screening test by focusing only on those subjects who couldn't produce high quality speech. He used HMM for recognition. Usually long speech tends to have extra parameters like noise. Hence,

a technique based on short time intelligibility assessment has been addressed by Taal et al. [9] which works for both noisy and TF-weighted noisy speech. He made four different models for assessment namely covariance-based STI (CSTI), sophisticated perceptual model (DAU), normalized subband envelope correlation (NSEC) and proposed method (fusion of all these). Table 1 has been included to provide an overview of the reviewed techniques. In the forthcoming sessions, Sect. 2 discusses the challenges in the field of pathological speech. Sections 3 and 4 comprehensively discusses the current approaches used to study the pathological conditions. At last, Sect. 5 confers the conclusion.

2 Challenges in Pathological Speech

Speech impairments can affect the breathing, phonation and articulation processes necessary for human speech production [10]. Speech therapy frequently helps to restore lost speech intelligibility. But the outcomes of intense therapy aren't always noticeable. Patients with this sort of diseases experience communication challenges in daily life due to improper pronunciation and voicing. As a result, intelligibility evaluation is frequently used by SLPs to describe the severity of symptoms and to offer follow-up remedies. As a result, the most widely utilised technique for determining intelligibility is perceptual evaluation carried out through subjective listening in medical and therapeutic places [11]. Speech utterances must be recorded by human listeners for a precise and effective evaluation. However, this approach necessitates a great deal of work, exorbitant and cannot be replicated. Additionally, certain listening tests' linguistic background information and the listener's familiarity with the subject's speech have an impact on this evaluation [12]. The necessity for automated intelligibility metrics to support therapists in their assessments is highlighted by these disadvantages [13]. With many origins and types of variability sometimes make it a highly difficult computational processing challenge, despite the fact that autonomous evaluation of speech intelligibility and quality could be helpful in these situations to aid specialists in diagnosing and treating the problem.

2.1 Subjective Assessments

Subjective evaluations of intelligibility are not only time-taking and costly, but it is also labour-intensive. They are influenced, however, by the clinicians' familiarity mostly with the person's speech pathology as well as the contextual/linguistic cues usable in speech sound [14]. Using automated detection of speech methods based on machine learning and signal processing can assist in the clinical diagnosis of SP and prevent the pitfall associated with clinical assessments [15]. Furthermore, APS intelligibility measures have been proposed as an economical and efficient alternative to subjective intelligibility assessments [5, 16–20]. Automatic analysis of PS pro-

Table 1 Overview of reviewed techniques in pathological speech

Sl. No.	Year	Title	Technique	Dataset corpus
1	1996	A screening test for speech pathology assessment using objective quality measures	HMM	21 [11 female (F) and 10 male (M)] preschool aged native speakers of English
2	2008	Suitability of dysphonia measurements for tele-monitoring of Parkinson's disease	Feature selection exploiting PPE	Consisting of 195 sustained vowel phonations from 31 M & F subjects, in which 23 are PWP
3	2010	A Short-time objective intelligibility measure for time-frequency weighted noisy speech	DFT-based TF-decomposition	15 normal-hearing native Danish speaking subjects
4	2012	Automated dysarthria severity classification for improved objective intelligibility assessment of spastic dysarthric speech	Mahalanobis-distance discriminant function	UA-speech
5	2017	Detection of different voice diseases based on the nonlinear characterization of speech signals	DHMM	LP: 72 recordings. 36 speakers with impairments (organic and traumatic disorders), and other 36 speakers are HC. PD: 50 having PD and 50 HC. CLP: 5 Spanish vowels by 65 children with repaired CLP and 54 HC
6	2017	Cross-database models for the classification of dysarthia presence	HMM	15 speakers (4 F, 11 M) of cerebral palsy-related spastic dysarthria mostly, plus 13 controls without dysarthria (4 F, 9 M)
7	2019	Super-Gaussianity of speech spectral coefficients as a potential biomrker for dysarthic speech detection	French recordings of 10 PD patients (6 M, 4 F)	None
8	2020	Hilbert spectrum analysis for automatic detection and evaluation of Parkinson's speech	Feature selection using Hilbert Huang Transformation	PC-GITA corpus, which has 50 PD and 50 HC and another DB consisting of 20 PD (10 M and 10 F) and 20 HC (10 M and 10 F)
9	2020	Spectro-temporal sparsity characterization for dysarthric speech detection	SVM	Spanish recordings of 45 HC and 45 PD

vides frequent, efficient, economical and objective tools for evaluating intelligibility and detection of PS, as opposed to clinicians' auditory-perceptual assessments. In addition to improving diagnostic and management capabilities, these techniques can also be used in speech therapy. These techniques can be delivered remotely, to help with early diagnosis and disease management [8, 15].

3   Feature Selection Methods

Wallen et al. [8] performs a screening test by combining objective quality measures that examine excitation and vocal tract characteristics. Pitch detector, pitch perturbation, amplitude perturbation, main cepstral peak and log-likelihood measures are used as objective speech quality measures and HMM is used for pathology assessment. Tanas et al. [21] addressed tremors, general lack of muscle control and rigidity as well as cognitive impairment as symptoms of PD. In order to extract features, dysphonia measures are used and four algorithmic features are selected: namely least absolute shrinkage and selection operator (LASSO) [22], minimum redundancy maximum relevance (mRMR) [23], RELIEF [24] and local learning-based feature selection (LLBFS) [25]. Then, in the subsequent mapping phase, these features are fed as input into the classifier, i.e. random forest and SVM, based on the dysphonia measures for binary classification, the final predictions will be HC/PD. A work done by Little et al. [26] proposes a new dysphonia measurement, Pitch Period Entropy (PPE), it is robust to a number of uncontrollable confounding factors inclusive of acoustically noisy environments as well as normal, healthy different versions in speech frequency. 31 sustained phonations were collected out of which 23 were PWP, as a result, it was discovered that non-conventional methods combined with traditional HNR ratios are the best at distinguishing subjects as HC/PD. The following non-conventional methods [correlation dimension (D2), recurrence period density entropy (RPDE), detrended fluctuation analysis (DFA)] are robust to many uncontrollable variations in acoustic environment and individual subjects and are thus well-suited to tele-monitoring applications. Advantage being that it uses a logarithmic pitch scale and is robust to confounding factors such as smooth vibrato which is present in healthy voices as well as dysphonic voices. Tsanas et al. [27] uses RPDE, DFA and PPE to implement a robust feature selection procedure. It was discovered that these measures provide valuable information about healthy and PWP. There were 16 dysphonia measures used in total which are statistically mapped using linear and nonlinear regression.

4   Approaches Used for Automatic Detection of Pathological Speech

4.1   Classification Approaches

Generally, acoustic features are designed to parameterize a short segment of speech (short-time feature) and depending on the classifier, the representations of short-time features need to be gathered while pronouncing or at speaker level prior to giving input into the classifier. To aggregate short-time features throughout an utterance length, statistical functions such as kurtosis, standard deviation and mean are calculated to obtain a vector of lengths for each utterance. The scarcity of pathological

speech training data increases the risk of overfitting by retaining a large number of features after such feature aggregation. As a consequence, algorithms which use linear discriminant analysis (LDA), principal component analysis (PCA) and correlation filtering are often used prior to training classifiers. Furthermore, GMM and iVector-based modelling is also exploited for this task.

4.2 Data-Driven Approaches

It has been observed that handcrafted features work well when applied to machine learning algorithms but when we apply those features to deep learning models the features may fail to satisfactorily capture pathological speech characterstics by omitting important acoustic cues present in pathological speech. Therefore, data-driven deep learning approaches have been used. Though deep learning models improves the performance but due to limited data availability they haven't yet dominated classical machine learning approaches. In order to increase the training samples, speech signals are split into short segments thereby labelling each sample as HC or PS and then CNN is trained on these samples for PS detection. An identical approach is practised in Mallela et al. [6], where the speech components are classified using not only cascaded CNN but also long short-term memory (LSTM) layers, with the features from the raw speech waveform. A cascaded architecture with a 1-DCNN layer and BLSTM layers for classifying subjects with PD, ALS and HC exploiting raw speech waveform. A 256 filters magnitude response for ALS, PD & HC without pre-emphasis is used. A similar work done by Bhati et al. [28] used Siamese networks based on LSTM to detect PS and were trained on sets of data which is fed with same prosodic composition. Pairwise training has the benefit of extracting discriminant features of PS. Because input data must contain the equivalent phonetic content, unique LSTM networks must be trained on independent utterances.

5   Approaches Used for Automatic Assessment of Pathological Speech

The assessment of speech intelligibility is essential for evaluating the functional outcomes of rehabilitation, speech therapies, surgical and non-surgical treatment. Berisha et al. [29] aims to identify specific acoustic cues associated with how SLPs judge PS similarity. Using these features, we hope to create listening models capable of evaluating speech along various perceptual dimensions. In this case, feature selection is accomplished through the use of relative dissimilarities between vectors, which correspond to speakers in the form "A is more like B than C". It is demonstrated that using the features chosen by the algorithm, we can create predictive models that reliably evaluate PS. Kim et al. [29] produced a sentence-level feature capturing

technique. A smoothing scheme applied after classification, which refines the sample. Thereafter, performing feature-level fusion and subsystem decision fusion to get an intelligibility score. A new technique proposed by Kim et al. [30] in which multi-resolution sinusoidal transform scheme was used to extract spectral features from multiple acoustic features, allowing for a better representation of spectral and harmonic characteristics. Regression methods were used to predict utterance interval-scaled intelligibility scores, allowing binary intelligibility classification. Similarly, Fletcher et al. [31], Hummel et al. [32] used multiple acoustic features for handcrafting and combining feature selection/reduction methods before training regression models to estimate speech intelligibility. Approach to use iVectors, a set of measures that capture many aspects of a person's speech, including intelligibility for intelligibility assessment has also been exploited by Martinez et al. [19] when there is an existing set of words annotated for intelligibility from the speaker to be evaluated, which can be used to train our system and then perform intelligibility rating for words not seen during training using iVectors. After that, the intelligibility score is calculated by training a regression model on the GMM-based iVector. Imed et al. [20] used a Gaussian posteriogram representation instead of an iVector in a similar manner.

6 Conclusions and Future Directions

In this paper, an impression of work going on in the field of assessment and detection of PS has been addressed. Here, the focus was majorly on the following outlooks; challenges faced during the assessment of PS; feature selection techniques for obtaining the best set of features from the speech signal; approaches used for automatic detection and assessment of speech independently. Moreover, as it been proved by various state of the arts that neural architectures show promising results for pathological speech detection tasks, but they haven't yet significantly overcome the machine learning-based approaches because of lack of availability of large training data. Hence, the use of data-driven approaches is encouraged or cross-database modelling is also performed to obtain better results. As we know the area of PS is evolving as more and more feature selection technique are being introduced. Hopefully the advancement will be helpful in terms of physically developing a tool which can be used by the people suffering from PS disorders.

References

1. Duffy JR (2000) Motor speech disorders: clues to neurologic diagnosis. In: Parkinson's disease and movement disorders. Curr Clin Pract.https://doi.org/10.1007/978-1-59259-410-8_2
2. American Speech-Language-Hearing Association (2019) Council for clinical certification in audiology and speech-language pathology. https://www.asha.org/

3. Enderby P (2013) Disorders of communication: dysarthria, pp 273–281. https://doi.org/10.1016/B978-0-444-52901-5.00022-8
4. Karan B et al (2020) Hilbert spectrum analysis for automatic detection and evaluation of Parkinson's speech. Biomed Sign Process Control 6:102050. https://doi.org/10.1016/j.bspc.2020.102050
5. Bocklet T et al (2013) Automatic evaluation of parkinson's speech-acoustic, prosodic and voice related cues, pp 1149–1153
6. Mallela J et al (2020) Raw speech waveform based classification of patients with ALS, Parkinson's disease and healthy controls using CNN-BLSTM. https://doi.org/10.21437/Interspeech.2020-2221
7. Gillespie S et al (2017) Cross-database models for the classification of dysarthria presence, pp 3127–3131. https://doi.org/10.21437/Interspeech.2017-216
8. Wallen EJ et al (1996) A screening test for speech pathology assessment using objective quality measures. https://doi.org/10.1109/ICSLP.1996.607478
9. Taal CH et al (2010) A short-time objective intelligibility measure for time-frequency weighted noisy speech, pp 4214–4217. https://doi.org/10.1109/ICASSP.2010.5495701
10. Hickok G (2012) Computational neuroanatomy of speech production, pp 135–145. https://doi.org/10.1038/nrn3158
11. Fex S (1992) Perceptual evaluation, pp 155–158. https://doi.org/10.1016/S0892-1997(05)80130-4
12. McHenry M (2011) An exploration of listener variability in intelligibility judgments. https://doi.org/10.1044/1058-0360(2010/10-0059)
13. Middag C (2012) Automatic analysis of pathological speech. http://hdl.handle.net/1854/LU-3007443
14. Landa S et al (2014) Association between objective measurement of the speech intelligibility of young people with dysarthria and listener ratings of ease of understanding, pp 408–416. https://doi.org/10.3109/17549507.2014.927922
15. Baghai-Ravary L et al (2012) Automatic speech signal analysis for clinical diagnosis and assessment of speech disorders. https://doi.org/10.1007/978-1-4614-4574-6
16. Maier A et al (2009) PEAKS-A system for the automatic evaluation of voice and speech disorders, pp 425–437. https://doi.org/10.1016/j.specom.2009.01.004
17. Middag C et al (2009) Automated intelligibility assessment of pathological speech using phonological features, pp 1–9. https://doi.org/10.1155/2009/629030
18. Kalita S et al (2018) Intelligibility assessment of cleft lip and palate speech using Gaussian posteriograms based on joint spectro-temporal features, pp 2413–2423. https://doi.org/10.1121/1.5064463
19. Martínez D et al (2015) Intelligibility assessment and speech recognizer word accuracy rate prediction for dysarthric speakers in a factor analysis subspace, pp 1–21. https://doi.org/10.1145/2746405
20. Laaridh I et al (2017) Automatic prediction of speech evaluation metrics for dysarthric speech. https://hal.archives-ouvertes.fr/hal-01771613
21. Tsanas A et al (2012) Novel speech signal processing algorithms for high-accuracy classification of Parkinson's disease, pp 1264–1271. https://doi.org/10.1109/TBME.2012.2183367
22. Tibshirani R et al (1996) Regression shrinkage and selection via the lasso, pp 267–288. https://doi.org/10.1111/j.2517-6161.1996.tb02080.x
23. Peng H et al (2005) Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy, pp 1226–1238. https://doi.org/10.1109/TPAMI.2005.159
24. Kira K et al (1992) A practical approach to feature selection, pp 249–256. https://doi.org/10.1016/B978-1-55860-247-2.50037-1
25. Sun Y et al (2009) Local-learning-based feature selection for high-dimensional data analysis, pp 1610–1626. https://doi.org/10.1109/TPAMI.2009.190
26. Little M et al, Suitability of dysphonia measurements for telemonitoring of Parkinson's disease. https://doi.org/10.1038/npre.2008.2298.1

27. Tsanas A et al (2009) Accurate telemonitoring of Parkinson's disease progression by non-invasive speech tests, pp 1–1. https://doi.org/10.1038/npre.2009.3920.1
28. Bhati S et al (2019) LSTM Siamese network for Parkinson's disease detection from speech. https://doi.org/10.1109/GlobalSIP45357.2019.8969430
29. Kim J et al. (2015) Automatic intelligibility classification of sentence-level pathological speech, pp 132–144. https://doi.org/10.1016/j.csl.2014.02.001
30. Kim JC et al (2014) Speech intelligibility estimation using multi-resolution spectral features for speakers undergoing cancer treatment, pp EL315–EL321. https://doi.org/10.1121/1.4896410
31. Fletcher AR et al (2017) Predicting intelligibility gains in dysarthria through automated speech feature analysis, pp 3058–3068. https://doi.org/10.1044/2017_JSLHR-S-16-0453
32. Hummel R et al (2011) Spectral features for automatic blind intelligibility estimation of spastic dysarthric speech
33. Berisha V et al (2014) Modeling pathological speech perception from data with similarity labels. https://doi.org/10.1109/ICASSP.2014.6853730
34. Lang AE, Lozano AM (1998) Parkinson's disease. First of two parts, pp 1044–1053. https://doi.org/10.1056/NEJM199810083391506
35. Rajput AH et al (1991) Accuracy of clinical diagnosis in parkinsonism-a prospective study, pp 275–278. https://doi.org/10.1017/S0317167100031814
36. Hughes AJ et al (1993) A clinicopathologic study of 100 cases of Parkinson's disease, pp 140–148. https://doi.org/10.1001/archneur.1993.00540020018011

Proposing ML Approach for Detection of Diabetes

Vaibhav Kant Singh and **Nageshwar Dev Yadav**

1 Introduction

Diabetes mellitus is generally speaking a chronic disease which is basically long lasting issue related to health where the consequence is on the way food is converted into the form of energy. There are generally three variants of it, namely type1also called juvenile, type2 also called as adult-onset, and gestational. In case of diabetes, there is discrepancy in the way how insulin which is a hormone gets into action. Diabetes is a major reason of deaths in developed countries like USA. Some generally observed symptoms are slow healing of sores, urination problem, blurry vision, weak and tired feeling, feeling thirsty, mood change, weight loss, and having infection such as skin. You can prevent diabetes by having maintained body weight, taking healthy diet, doing physical activity as much as possible, and non-intake of tobacco. The major problems related to health which a person may encounter because of diabetes are cardiovascular disease, damage of eye, Alzheimer, damage of nerve, depression, kidney damage, hearing impairment, damage of foot and skin and mouth related problems. The Management of Diabetes could be done via a set of activities like having a proper lifestyle that means to say waking up on time, doing exercise, taking meal on time, and so on. Medication is also a solution, which is basically for controlling glucose and trying to level blood pressure. We can go for surgery to lower down the body weight. The traces of the disease are identified in Egyptian Scripts.

V. K. Singh (✉)
Department of Computer Science and Engineering, School of Studies of Engineering and Technology, Guru Ghasidas Vishwavidyalaya, Central University Bilaspur, Chhattisgarh 495001, India
e-mail: vibhu200427@gmail.com

N. D. Yadav
Informatica Solution, Banglore, India

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2024

In the past few years, we have seen a lot of development in various fields of engineering like chemical engineering, industrial and production engineering, mechanical engineering, information technology and engineering, civil engineering, and many more areas of engineering. The current work is a point to make in the field of engineering and technology.

This paper is to address an issue of medical science called diabetes. Recently an incident happened in the locality near Nehru nagar of Bilaspur, Chhattisgarh, India. A man in the morning was walking and suddenly got some problem because of which he fell down and lost his conciseness. Some of the people walking in the street saw him and took him to a nice place and then tried to take him out of his unconsciousness. After some time, the person got conscious and told that he is suffering from the problem of Sugar and Blood Pressure may be because of that he fell down. This incident and many of this type we see in our surrounding every day. In the case mentioned above, the person was safely handled because of the kind people around. But this may not happen with everyone and even with the same person every time. Thus, it is very important to take care of diseases like this with care and shear attention.

In the current work, we will be having a combination of the flexible approach of Computer Science called Machine Learning and in that too we will be enjoying the Vehicle called Python as the Journey material. Our Journey in the Paper will comprise sections which includes problem statement, literature survey, implementation and results, conclusion, and at last references.

2 Problem Statement

One of the authors of the paper has lost his close relative in the near past because of the non-diagnosis of the disease that the relative was having at an appropriate time even by the best doctor in the Town. We every now and then are losing our near and dear ones because of the non-timely detection of a disease. As already conveyed above there are several complication which may arise because of diabetes like cardiovascular disease and so on. The diseases are sometimes catastrophic which may even cause the death of a Person. Life is a precious entity. Every life is important. In the current work, we are focused on diabetes which is a very general disease found nowadays.

3 Literature Survey

In [1–13], the authors used Machine Learning for detection of diseases like diabetes, Lung Cancer, Breast Cancer, and Thyroid. In the papers, we will observe an insider about the aforesaid diseases. At the same time, we will have an exploration on the various different types of datasets and the various interfaces designed to tackle

the problem discussed in the mentioned papers. In [14–22], the authors gave ANN models for XOR and Ex-NOR problems. Artificial Neural Network is one of the most versatile and outstanding field of engineering that is in demand nowadays. In the above mentioned papers, we will observe a very basic understanding of how we can do modeling to take care of various types of problems. In [23], simulator named TRIVENI is conveyed. TRIVENI is a model that could be a very handy one for the generation to come which is eager to work on ANN and VLSI. In [24–28], the authors published patents in the field of ANN, etc. The patents published are very nice models for VLSI and computer designing. In [29], Vector Model is discussed. The paper is a highly cited paper for its implementation point of view. In [30–35], RSTDB and other FPM techniques are discussed. There is a great capability inculcated in we human beings that is what is exploited in the papers above. In [36–46], the authors made discussion on techniques that come out of the fields like AI and SC. Artificial Intelligence and Soft Computing are the fields of current attraction in terms of implement ability and its long lasting results. The papers are path finders for the current research. In [47–55], the authors gave demonstration of systems implemented on the test bed of Python. Python language exploration and how to take its utilization in research objective is what you can take out of the above papers. In [56], Dual method for Security is mentioned. Security is the major concern and is what is covered in the above paper. In [57, 58], various SC and AI technologies are covered. Soft Computing Technologies on its account of its usage in the current paper are cited again. In [59], how to make systems using C and Web-based Programming is discussed. It is an apparent fact that we can go for development using any vehicle. How to go for development is the idea of the literature cited. In [60], multitier architecture for VLSI is proposed. VLSI is the bed on which we will be going to run our programs and thus the idea of our concern. In [61], ANN concept for Perceptron is discussed. Perceptron landscape is what is explored in the cited work. In [62], a Python-based natural language processing implementation is carried out with a shear intention to brought in implementation the Machine Language capability and the test bed into one platform. In [63], we saw a Python implementation with an idea to tackle the problem of forgery in images. In [64], a Python-based search engine is implemented; this system is going to be utilized for a number of datasets on which training will be done, and then the traversal it goes in a fourfold way. It is a novel and great way for the purpose held.

4 Proposed Implementation and Result

The implementation of the Code is done in Python. The Algorithm comprises around 17 steps. The first step is importing of the necessary libraries. Step two is importing and reading the dataset. Step three is preprocessing of the data. Step four is analysis using Seaborn and Matplotlib. Step five is column checking for features that are numeric. Step six is setting columns for training. Step seven is testing various ML models for performance. Step eight is testing ML for result identification. Step nine
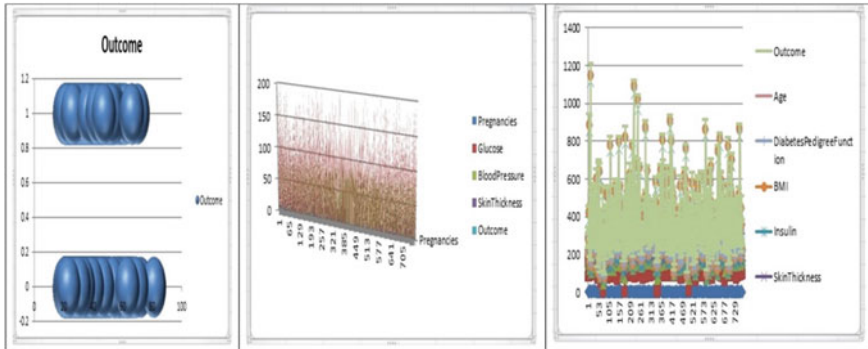
Fig. 1 Dataset visualization for data preprocessing and column selection

is searching for best fold by repetitive analysis. Step ten is finding the best approach. Step eleven is comparing the approach that is best with all other ones. Step twelve is hyper tuning the model chosen. Step thirteen is reading model from local directory. Step fourteen is splitting dataset into training and testing. Step fifteen is using the best model for testing. Step sixteen is predicting the new case. Step seventeen is finding out the accuracy measure. We in our implementation went with the approach where we made the implementation of the various approaches for the current dataset. The dataset is in under procession taken from Kaggle. If you see Fig. 2 of the current paper, you can see that Logistic Regression gives accuracy of around 77%, Linear Discriminant Analysis gives an accuracy of around 77%, Ridge Classifier gives an accuracy of around 77%, Random Forest Classifier gives an accuracy of around 75%, Naïve Bayes gives an accuracy of 75%, Gradient Boosting Classifier gives an accuracy of 75%, Extra Trees Classifier gives accuracy of 75%, Light Gradient Boosting Machine gives accuracy of around 73%, Ada Boost Classifier gives accuracy of around 72%, Quadratic Discriminant Analysis gives accuracy of around 72%, K Neighbors Classifier gives accuracy of around 71%, Decision Tree Classifier gives accuracy of around 67%, Dummy Classifier gives accuracy of around 65%, and SVM taking linear as the Kernel gives an accuracy of around 56%. Figure 2 gives a very good idea of the various approaches on the given dataset from Kaggle. Not only accuracy Fig. 2 of the current paper gives a number of other results also that would be utilized for algorithm assessment. We are doing the assessment of the algorithm so that we can go with an approach that would be the best for our purpose (Figs. 1 and 3).

	Model	Accuracy	AUC	Recall	Prec.	F1	Kappa	MCC	TT (Sec)
lr	Logistic Regression	0.7748	0.8297	0.5403	0.7228	0.6100	0.4600	0.4729	0.7957
lda	Linear Discriminant Analysis	0.7747	0.8306	0.5405	0.7307	0.6128	0.4617	0.4760	0.0114
ridge	Ridge Classifier	0.7710	0.0000	0.5295	0.7270	0.6045	0.4515	0.4665	0.0071
rf	Random Forest Classifier	0.7597	0.8069	0.5678	0.6706	0.6105	0.4405	0.4456	0.0871
nb	Naive Bayes	0.7578	0.8090	0.5899	0.6645	0.6232	0.4460	0.4489	0.0071
gbc	Gradient Boosting Classifier	0.7524	0.7998	0.5845	0.6504	0.6127	0.4325	0.4354	0.0486
et	Extra Trees Classifier	0.7522	0.8013	0.5132	0.6749	0.5802	0.4107	0.4196	0.1014
lightgbm	Light Gradient Boosting Machine	0.7374	0.7850	0.6172	0.6175	0.6131	0.4155	0.4186	0.2814
ada	Ada Boost Classifier	0.7244	0.7624	0.5570	0.5989	0.5752	0.3725	0.3739	0.0414
qda	Quadratic Discriminant Analysis	0.7206	0.7991	0.5026	0.6181	0.5504	0.3513	0.3578	0.0086
knn	K Neighbors Classifier	0.7131	0.7090	0.5189	0.5964	0.5512	0.3426	0.3469	0.4300
dt	Decision Tree Classifier	0.6742	0.6407	0.5354	0.5248	0.5276	0.2800	0.2814	0.0071
dummy	Dummy Classifier	0.6592	0.5000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0071
svm	SVM - Linear Kernel	0.5680	0.0000	0.5657	0.3382	0.3919	0.1163	0.1406	0.0086

Fig. 2 Comparison chart of the various algorithm on the dataset

	Accuracy	AUC	Recall	Prec.	F1	Kappa	MCC
Fold							
0	0.6296	0.6698	0.1667	0.3750	0.2308	0.0323	0.0369
1	0.7778	0.8688	0.4444	0.8000	0.5714	0.4375	0.4719
2	0.7963	0.8225	0.4444	0.8889	0.5926	0.4762	0.5270
3	0.7407	0.8349	0.5000	0.6429	0.5625	0.3824	0.3884
4	0.7778	0.8331	0.5263	0.7692	0.6250	0.4749	0.4921
5	0.7963	0.8286	0.6842	0.7222	0.7027	0.5479	0.5484
6	0.7963	0.9083	0.6316	0.7500	0.6857	0.5367	0.5410
7	0.8491	0.8571	0.7778	0.7778	0.7778	0.6635	0.6635
8	0.8302	0.8746	0.7222	0.7647	0.7429	0.6163	0.6168
9	0.7358	0.7619	0.5556	0.6250	0.5882	0.3948	0.3963
Mean	0.7730	0.8260	0.5453	0.7116	0.6080	0.4562	0.4682
Std	0.0581	0.0637	0.1668	0.1330	0.1444	0.1654	0.1654

```
              precision   recall  f1-score  support

        0        0.81      0.85     0.83       123
        1        0.71      0.64     0.67        69

    accuracy                        0.78       192
   macro avg     0.76      0.75     0.75       192
weighted avg     0.77      0.78     0.77       192


#FINAL ACCURACY OF THE MODEL
print("Accuracy Score is : ",lr_model.score(x_test,y_test)*100,"%")

Accuracy Score is : 77.60416666666666 %
```

Fig. 3 Classification report and tuned report for lr the model

5 Conclusion

The algorithm that which is selected as the model for prediction of diabetes is the Logistic Regression model (LR). LR model is essentially a classification approach. The basic idea of the LR model is generally to find out a mapping between the dependent and the independent variable. The accuracy of around 77.60 is obtained for the given dataset. The model accuracy obtained is not a very good one and leaves scope for future improvement. The research objective however was to derive a computer-based system that can solve the detection problem in a time effective manner, that objective is achieved with a high degree. The authors made a comparative chart for the various algorithms operated on the dataset. The author from their work in

the ML field can make a statement that a much better efficiency in terms of accuracy can be obtained. One reason for the less accurate measure could be unavailability of huge amount of data for experiment purpose. The non-extensive nature of the dataset would have been a barrier for the level of accuracy obtained by the authors in the experimental work. Python is a very successful programming language for the preparation of the ML systems that can perform the detection activity in a variety of problem domains. The Comparative chart comprises 14 different ML techniques which include svm, qda, lda, light gbm, knn, ridge, et, nb, ada, gbc, dt, rf, dummy, and lr. In our approach out of the various, we made a selection of the lr model.

References

1. Singh VK (2021) SVM using rbf as kernel for diagnosis of breast cancer. In: Conference proceeding ICIRSMT 2021. AIMTCP. Bilaspur, India, pp 343–348. ISSN: 2769-5093
2. Singh VK (2021) Support vector machine using rbf, polynomial, linear and sigmoid as kernel to detect diabetes cases and to make a comparative analysis of the models. In: Conference proceeding ICIRSMT 2021. AIMTCP. Bilaspur, India, p 25. ISSN: 2769-5093
3. Singh VK, Yadav ND, Singh RK, Sahu M (2022) Detection of thyroid using machine learning approach. NIU Int J Hum Rights (UGC Care Listed Journal) 3(II):65–80. India, ISSN: 2394-0298
4. Yadav ND, Singh VK, Singh RK, Sahu M (2022) A comparative analysis of SVM kernels for detection of diabetes. Anvesak (UGC Care Listed Journal) 52(5):61–67, India. ISSN: 0378-4568
5. Yadav ND, Singh VK, Singh RK, Sahu M (2022) ML based SVM taking RBF as kernel for detection of breast cancer. J Educ Rabindra Bharati University (UGC Care Listed Journal) XXV(5):72–79, India. ISSN: 0972-7175
6. Singh VK (2022) Support vector machine an ML approach for breast cancer. In: 6th International Izmir economics congress, Izmir, Turkey
7. Singh VK (2022) SVM an ML approach for diabetes. In: 6th International Izmir economics congress, Izmir, Turkey
8. Kumari P, Gupta R, Kumar S, Singh VK (2021) ML approach for detection of lung cancer. In: ISPEC 8th international conference on agriculture, animal science and rural development. Bingol, Turkey
9. Singh VK (2022) Pycaret a python implementation for thyroid disease. In: IArcSAS 2nd international architectural sciences and applications symposium. Baku, Azerbaijan
10. Singh VK (2022) SVM an ML approach for breast cancer. In: IArcSAS 2nd international architectural sciences and applications symposium. Baku, Azerbaijan
11. Singh VK (2022) An implementation in python for diagnosis of thyroid using machine learning approach. In: 9 International GAP summit scientific research congress, Adiyaman, Turkey
12. Singh VK (2022) ML based python implementation for detection of lung cancer. In: IV International Halich congress on multidisciplinary scientific research. Istanbul, Turkey
13. Singh VK (2022) A computerized approach for thyroid disease having the usage of ML technique. In: IV International Halich congress on multidisciplinary scientific research. Istanbul, Turkey
14. Singh VK (2016) Proposing solution to XOR problem using minimum configuration MLP. In: Science direct international conference on computational modeling and security (CMS 2016), vol 85, pp 263–270. Procedia Computer Science, Bangalore, India
15. Singh VK, Pandey S (2016) Minimum configuration MLP for solving XOR problem. In: Proceeding of the 10th INDIAcom IEEE conference ID: 37465. IEEE Explore, BVICAM Delhi, India, pp 168–173. ISBN: 978-9-3805-4421-2

16. Singh VK, Pandey S (2016) Proposing an Ex-NOR solution using ANN. In: Proceeding international conference on information communication and computing technology (ICICCT-2016). JIMS, Delhi, India, pp 277–284. ISBN: 978-93-85777-66-0

17. Singh VK (2016) Mathematical explanation to solution for Ex-NOR problem using MLFFN. Int J Inf Sci Tech 6:105–122

18. Singh VK (2016) Proposing a new ANN model for solving XNOR problem. In: IEEE international conference system modeling and advancement in research trends (SMART). IEEE Explore, Moradabad, India. ISBN: 978-1-5050-3543-4

19. Singh VK, Singh DK (2015) Multilayer perceptron for XOR problem. In: IRSMT-2015, Bilaspur, India, p 33

20. Singh VK (2015) Two solutions to the XOR problem using minimum configuration MLP. Int J Adv Eng Sci Technol Res (IJAESTR) 3(3):16–20. ISSN: 2321-1202

21. Singh VK (2015) One solution to XOR problem using multilayer perceptron having minimum configuration. Int J Sci Eng 3(2):32–41. ISSN: 2347-2200

22. Singh V. (2016) ANN implementation of construction logic gates focuses on Ex-NOR. Res J Comput Inf Technol Sci 4(6):1–11. ISSN: 2320-6527

23. Singh VK (2017) Designing simulators for various VLSI designs using the proposed artificial neural network model TRIVENI. In: IEEE international conference on information, communication, instrumentation and control (ICICIC), IEEE Explore, pp 1–6, Indore, India. ISBN: 978-1-5090-6313-0

24. Singh VK, Singh AK, Deo Y, Reddy DS (2022) A dual authentication implementation using biometric properties of fingerprint following covid protocol and face matching age and gender. In: Intellectual property of India, Application Number: 202221063732

25. Singh VK, Yadav ND (2022) SFG for showing the solution to XOR problem which is a non linearly separable problem. In: Intellectual property of India, Application Number: 202221041900

26. Singh VK, Yadav ND, Singh RK (2022) BDR for solving linearly separable problems like AND, OR, NAND & NOR. In: Intellectual property of India, Application Number: 202221047404

27. Singh VK, Yadav ND, Singh RK (2022) SFG for solving non-linearly separable problem that is Ex-NOR problem. In: Intellectual property of India, Application Number: 202221041911

28. Pippal RK, Singh VK, Gupta P, Gangil M, Bux S, Sadiwala R, Hande JY, Dorave JD (2022) Method for migration of sessions among cipher machines using Lagrange interpolation. In: Intellectual property of India, Application Number: 202221000273

29. Singh VK, Singh VK (2015) Vector space model: an information retrieval system. Int J Adv Eng Res Stud 4(2):141–143

30. Singh VK (201*) Proposing pattern growth methods for frequent pattern mining on account of its comparison made with the candidate generation and test approach for a given data set. In: Software engineering. Springer Nature, Singapore, pp 203–209

31. Singh VK (2010) RSTDB & cache conscious techniques for frequent pattern mining. In: Proceeding 4th international conference on computer application in electrical engineering recent advances (CERA-09), Indian Institute of Technology, Roorkee, India

32. Singh VK, Thakur VS, Pandey NK (2010) Proposing data mining as an efficient technique for rural development. In: Proceeding of international multi conference on intelligent system and nanotechnology, Haryana, India, pp 110–112

33. Singh VK, Singh VK (2010) RSTDB a new candidate generation and test algorithm for frequent pattern mining. In: Proceeding of international conference on advances in communication, network and computing. IEEE Computer Society, ACM DL Digital Library pp 416–418, Calicut, Kerala, India. ISBN: 978-0-7695-4209-6

34. Singh VK, Singh VK (2009) Minimizing space time complexity by RSTDB a new method for frequent pattern mining. In: Proceeding of the first international conference on human computer interaction. Springer, Indian Institute of Information Technology, Allahabad, India, pp 361–371

35. Singh VK, Dubey V, Singh AK (2010) Proposing data mining as an efficient technique for solving frauds in digital data. In: Proceeding international conference on intelligent information systems and management, India , pp 1–4. ISBN: 978-1-4507+2041-0

36. Singh VK (2022) A short report making a comparison between artificial intelligence and artificial neural network. In: IArcSAS 2nd international architectural sciences and applications symposium. Baku, Azerbaijan

37. Singh VK (2016) Mathematical analysis for training ANNs using basic learning algorithms. Res J Comput Inf Technol Sci 4(7):6–13. ISSN: 2320-652

38. Singh VK, Shah V, Jain YK, Shukla A, Thoke AS, Singh VK, Dule C, Parganih V (2008) Proposing an efficient method for frequent pattern mining. In: Proceeding world academy of science engineering and technology, Bangkok, Thailand, pp 1184–1189. ISSN: 2070-3740

39. Singh VK (2022) A short report on the basic artificial neuron model. In: IArcSAS 2nd international architectural sciences and applications symposium. Baku, Azerbaijan

40. Singh VK (2022) Fuzzy associative memory for truck backer system. In: International conference on global practice of multidisciplinary scientific studies. Turkey

41. Singh VK (2022) Backpropagation algorithm of ANN. In: International conference on global practice of multidisciplinary scientific studies. Turkey

42. Singh VK (2022) A report on principal component analysis. In: International conference on global practice of multidisciplinary scientific studies. Turkey

43. Singh VK (2022) Historical development in the field of artificial intelligence. In: 2nd International archeology, art history and cultural heritage congress. Sanliurfa, Turkey

44. Singh VK (2022) A study on elementary neurophysiology of human brain. In: IV International Halich congress on multidisciplinary scientific research. Istanbul, Turkey

45. Singh VK (2022) A study on feedback in an artificial neural system. In: IV International Halich congress on multidisciplinary scientific research. Istanbul, Turkey

46. Singh VK (2022) A study on artificial neural network when viewed as a directed graph. In: IV International Halich congress on multidisciplinary scientific research. Istanbul, Turkey

47. Singh VK (2021) Colorization of old gray scale images and videos using deep learning. J Oriental Res Madras (UGC Care Listed Journal), India, pp 44–49. ISSN: 0022-3301

48. Chandrashekhar RC, Singh VK (2021) Twitter sentiment analysis. In: ISPEC 8th international conference on agriculture, animal science and rural development. Bingol, Turkey

49. Sailokesh P, Jupudi S, Vamsi IK, Singh VK (2021) Automatic number plate recognition. In: ISPEC 8th international conference on agriculture, animal science and rural development. Bingol, Turkey

50. Reddy YK, Yadav KM, Singh VK (2021) Human activity recognition. In: ISPEC 8th international conference on agriculture, animal science and rural development. Bingol, Turkey

51. Prasad RNRK, Ram PSSR, Dinesh S, Singh VK (2021) Text summarization. In: ISPEC 8th international conference on agriculture, animal science and rural development. Bingol, Turkey

52. Singh VK (2022) An image based search engine system using python. In: 9 International GAP summit scientific research congress, Adiyaman, Turkey

53. Singh VK (2022) A system for news classification in regional language using python. In: 9 International GAP summit scientific research congress, Adiyaman, Turkey

54. Singh VK (2022) A system for cartoonifying an image using python. In: 9 International GAP summit scientific research congress, Adiyaman, Turkey

55. Singh VK (2022) A sign language recognition system. In: 9 International GAP summit scientific research congress, Adiyaman, Turkey

56. Singh VK, Singh AK (2010) Dual level digital watermarking for images. In: Proceeding international conference on methods and models in science and technology, AIP conference proceeding, Chandigarh, India, pp 284–287. ISBN: 978-0-7354-0879-1

57. Singh VK (2022) Soft computing, 1st edn. Book Rivers. India. ISBN: 9789355154026

58. Singh VK (2022) Advanced artificial intelligence, 1st edn. Book Rivers. India. ISBN: 9789355154019

59. Singh VK, Singh A, Singh A, Yadav ND (2022) Lab manual of web technology and C-programming. Ideal International E-Publication Private Limited. ISBN: 978-93-89817-67-6

60. Singh VK, Singh VK (2008) VLSI architecture for multitier wireless systems. In: International conference on: interdisciplinary approaches in physical science: growing trends and recent advances CONIAPS-X. Bilaspur, India, pp 166

61. Singh VK (2017) Analysis of stability and convergence on perceptron convergence algorithm. In: Proceeding ICICCT. Delhi, India, pp 149–161. ISBN: 978-93-86647-85-6
62. Singh VK, Soni SK, Yadav ND, Chandra P (2022) A system for news classification in regional language python. Shodh Prabha Referred Peer Rev Q Res J 47(4):169–182. ISSN: 0974-8946
63. Singh VK, Soni SK, Yadav ND, Chandra P (2022) A system for cartoonifying an image using python. SAMRIDDHI J Phys Sci Eng Technol 14(5):84–100. ISSN: 2229-7111
64. Yadav ND, Singh VK (2022) An image based search engine system using python. IJFANS Int J Food Nutrition Sci 11(3):2821–2836. ISSN: 2320-7876

Comparative and Preventive Analysis of Dictionary Attacks

Sanat Shourya, Ilayaraja Venkatachalam, Harpal Patel, and Manit Mittal

1 Introduction

With the shift to cloud and increasing usage of social media platforms that are existing and those that are newly being released every day, risk quotient is also shooting up. Smallest of services providing apps or websites for their usage, as the first step, requires you to go through the Registration/Sign-Up process, putting your personal info at risk, only secured by one element, passwords. Hackers are increasing who are vigilantly waiting for any breakthrough they can find and hence exploit for personal gains. In contrast to this, good length passwords (10–12 characters long, for instance), having a mix of lowercase, uppercase numbers, special symbols, etc., would prove to be much more resistant to these types of attacks. Consequently, a lot of apps and websites have started using Password Strength Classifiers on the Registration/Sign-Up pages, but these classifiers are prone and can easily succumb to dictionary attacks that are personalized and targeted in their approach. Users tend to keep their passwords using personal details. For example, passwords like 'Jacob43John' or 'MaryJane1998' are highly vulnerable to targeted dictionary attacks. Even variations in such passwords are easily breakable using personalized dictionary attacks. So, even though the password may be given a classification as 'strong' that you may be using can easily be argued.

S. Shourya · I. Venkatachalam (✉) · H. Patel · M. Mittal
School of Computer Science and Engineering, Vellore Institute of Technology, Vellore, India
e-mail: ilayaraja.v@vit.ac.in

2 Literature Overview

The potential causes of insufficient security in academic institutions and the effectiveness of security solutions in shielding online applications from a variety of known threats are revealed. Each targeted website is scanned during the process, and any vulnerabilities discovered are listed and cross-referenced with the top ten OWASP vulnerabilities list [1]. SQL injection attacks databases that could be accessed through a web front end by taking advantage of input validation rules in Web components like CGI scripts. To stop data leaking, a proxy for de-randomization is required [2]. Designing a template and outlining how to use a functional tool are required for starting a new project. However, automated machinery needs the significant capital investment [3]. The complexity of cracking a password and how tough it is to create a strong password determine the security level. Sowmya et al. [4] suggested experiments that were done to assess how well passwords of various strengths stood up to brute force assaults. The account is susceptible if the passwords are kept the same with little variation. All additional passwords may be simply recovered once the password is known. A number of high security provisioning techniques are investigated. The length and complexity of a password determine its strength [5]. While working on the servers, multiple dangers like privilege to unauthorized users, password cracking, knowing the root of machine happens that can demolish the system and gain server access [6]. The S-Hash function's behavior has been evaluated using the Czech and English dictionaries. It suggested contrasting S-Hash hashing techniques for textual data. A typical approach to textual and geometrical data is provided by the S-Hashing algorithm [7]. SQLi attacks are carried out by inserting malicious statements into your feedback boxes. Using these statements, the entire system can be accessed which can cause serious harm to the client [8]. The proposed scheme assesses the vulnerability of speaker check systems to word reference attacks. With the Any10 method, an MV may imitate 80% of ladies and 65% of men in the most insecure situation. A MV can generally imitate 20% of females and 10% of men in the most reliable design [9]. Subangan and Senthooran [10] introduces freestyle, a novel randomized figure with the ability to generate up to 2^{256} distinct ciphertexts for a given key, nonce, and message. It shows KGP > 1, which allows it to operate more quickly on a low-fuel machine. To counter the kind of covert word attacks, a workable and beneficial secure grid-based verification instrument and area-based confirmation component is presented [11]. A novel method for ensuring password security in the face of word reference attacks is suggested. It uses a word reference that is stored as a person tree to check the security of the user's passwords. Its shortcomings include the lack of a backup facility [12]. To prevent passageway from being recognized as an attack, the MAC address of the PC changes often. However, it is possible to utilize side-channel leaks and downgrade attacks to obtain the Wi-Fi network's password. The information that WPA3 was supposed to securely encrypt would thus be accessible to any determined adversary [13]. In addition to the normal adversary capabilities, we also take into account additional cryptographic

advantages provided to the enemy while taking into account the secret key convention being established with certain encryption schemes [14]. Integrity and privacy, CGI script, access control, data transmission over TCP/IP, and other mechanisms are among those presented. The security and privacy of data and information sent, stored, and processed in real time is one of the key challenges [15].

3 Methodology

We tried to mimic the real-world scenario of Sign-Up/ Registration pages, on how secure a password-protected data/file/pdf is, by exploiting all the possible ways we can guess the correct password and even check its vulnerability, if we know the target user's: personal details, family information, have access to his personal folder drive, etc.

3.1 Different Dictionaries (Text Files) Created:

3.1.1 Manual Dictionary

In this, we created dictionary of passwords from our own knowledge or those which are extremely common in usage. Additionally, suppose you get to see some part of the target's password, we can think of possible passwords from it.

3.1.2 Crunch Dictionary

In this, we tried all the possible combination of characters by using the length of password. Using this inbuilt Linux tool, complete wordlist can be generated including alphanumeric characters for an orthodox, full-fledged brute force attack.

3.1.3 Personalized Dictionary

If we know our target's personal details or get these details through Sign-Up pages of important search engines or for that case, be it any other site, we can easily create a hyper-personalized dictionary by trying all the possible password combinations from the supplied data.

3.1.4 Word Harvest Dictionary

It reads word from nearby .text and .txt files. It creates the dictionary searching the filesystem for files and extract alphanumeric words from them, saving all unique occurrences.

3.2 Description of Modules

3.2.1 Bruteforce.Py

Here, we specify our dictionary names and password-protected pdf file name as well as their directory. Thereby it reads each word from each line of the dictionary and tries it as a password in the given pdf. If it gets the right password, then it displays the password and also creates a dummy text file containing the content present in pdf file. It does the brute force dictionary attack to find the password.

3.2.2 Wordharvest.c

On executing this code, it reads all the '.text' and '.txt' supported files present in the same directory as this and stores each word (unique occurrences) from those files as dictionary passwords. It reads each line and separates words on the basis of '\n' spaces. It uses Regex function to extract all the alpha numeric words. The Hash 5381 and linked list codes help us to remove repetition of words inside the dictionaries.

3.2.3 Target.py

This Python program will ask you a series of questions about your target and his personal details, and it generates a list of possible passwords using the answers user has provided. The text file can then be added to an existing dictionary list or used by itself to try to crack passwords as shown in Fig. 2. It exploits the fact that the general public tends to keep their passwords fairly simple and easy, making use of their personal details, giving themselves a false sense of security. However, these can be easily tracked down with hyper-personalized dictionaries. Even the top websites don't give feedbacks whether the password is strong or not in such terms.

3.2.4 Crunch

Crunch is a highly important tool for generating a list of potential words for any password. The crunch utility is free to use under Kali Linux. This application supports Unicode, numbers, symbols, different upper and lowercase characters. If we get to
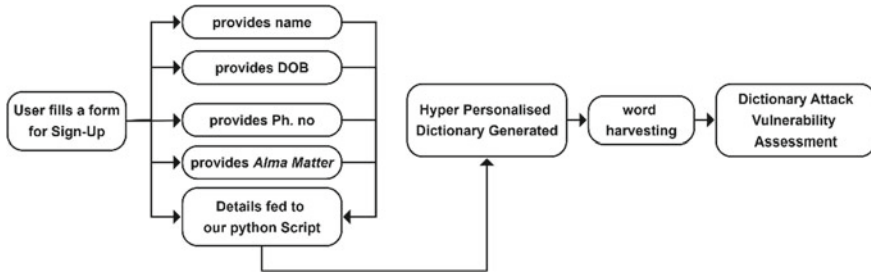
Fig. 1 Generation of hyper-personalized dictionary from sign-up/ registration pages in apps, websites, etc.

know the characters present or length of our target's password, we can generate all the possible combinations and use in a brute force attack. Its output can be piped to another program, displayed on the screen, or even saved to a file.

3.3 Description of Algorithm

3.3.1 XOR-Version DJB2 Algorithm

Every algorithm has different hash numbers. This algorithm specifically uses 5381 as its hash. It is one of the best hash functions for utf-8 strings that returns 32bit or 64bit integer, both considering performance and 'minimal collisions.'

4 System Architecture

See Fig. 1.

5 Results and Discussion

For execution, we need to open the terminal as a root login in the same directory where the dictionaries, modules are present (required path is necessary for proper working of the individual modules). Figure 3 shows the password-protected pdf file, made using the 'pdftk' tool, which is present in the same folder is tried on with the different dictionaries for vulnerability assessment. If we get the right password, we create dummy text document having the content of protected pdf file and the password gets displayed.

Fig. 2 Password decryption using word harvest dictionary

Fig. 3 Password decryption using crunch dictionary

5.1 Attacking the Password-Protected Pdf File—Target2.Pdf from All the Dictionaries Created

5.1.1 Using Word Harvest Dictionary—harvest2.txt

See Fig. 2.

5.1.2 Crunch Dictionary—Crunch2.txt

This takes a huge amount of time as combinations of words exceed 1,000,000 words, unlike other dictionaries (Fig. 3).

Fig. 4 Password decryption using personalized dictionary

5.1.3 Personalized Dictionary—Rahulpassword.txt (Optimal Case)

Among all the dictionaries, as shown in Fig. 4, personalized dictionary takes the least amount of time (found the result by scanning only 28% of the document), followed by word harvest dictionary shown in Fig. 2, which takes more time as it makes use of all files in the system so the search space is increased and finally crunch dictionary is the slowest taking up approx. 4 h to complete execution as shown in Fig. 3. Table 1 provides time taken to crack passwords using crunch dictionary, serving as a basis for relative calculation of time taken by other dictionaries. For example, 'Rahul2001' which take approx. 4 h to crack, take only few minutes when using personalized dictionaries. Hence, *completely randomized passwords* of sufficient characters (10+), having either no or disguised personal references, which would take millions of years to crack such as 'Sasnowym@2!3' given in Table 2, can safely be classified as 'Strong.'

6 Conclusion and Future Work

The Sign-Up page encountered on every website is of little significance if the entry point itself is compromised. Therefore, to keep oneself safe is by keeping strong passwords.

Many websites use Password Strength Classifier on the Sign-Up page, but they are still prone to dictionary attacks and even more to hyper-personalized dictionaries as shown in this paper. Our research can serve as the foundation of an intelligent Password Strength Classifier, which will also take into account the possibilities of dictionary attacks before classifying any passphrase as strong enough. A classifier, able to prevent people from making up passwords based on their personal info. Itself in a mixed-up manner and propelling them into forming a randomized password.

Table 1 Time comparison of breaking different types of passwords using crunch dictionary

Char. length	Only numbers	Only uppercase/ lowercase letters	Uppercase + lowercase letters both	Numbers + uppercase + lowercase letters	Numbers + uppercase + lowercase + symbols
3	Instantly	3 s	6 s	11 s	30 s
4	3 s	8 s	3 min	10 min	38 min
5	5 s	5 min	4 h	9 h	3 days
6	20 s	3 h	5 days	10 days	4 weeks
7	2.5 min	2 days	4 months	1 year	6 years
8	24 min	1 week	17 years	56 years	400 years
9	4 h	10 months	800 years	5 K years	32 K years
10	1.5 days	2 years	43 K years	255 K years	4 M years
11	3 weeks	41 years	2 M years	18 M years	166 M years

Table 2 Strength comparison of different types of passwords against different dictionary attacks

Passwords	Hyper-personalized, targeted dictionary	Word harvest dictionary	Crunch dictionary
Password008	Weak	Weak	Moderate
apoorva1	Weak	Weak	Weak
JaneMary1997	Weak	Moderate	Strong
03Stella	Moderate	Moderate	Moderate
Rahul2001	Weak	Moderate	Strong
Shikha12	Weak	Weak	Moderate
Ananshaaryan02	Moderate	Moderate	Strong
John43jacob	Weak	Strong	Strong
Verma.Abhi19	Moderate	Moderate	Strong
Sasnowym@2!3	Strong	Strong	Strong

The Preventive Analysis Data presented can be used to build a Machine Learning-based prediction model, which can accurately predict the time, in which a particular passphrase made up of jumbled details, can be guessed by the hackers. These insights will then be used to suggest stronger passwords to the users that practically take forever to be guessed and can practically be classified as 'strong.'

References

1. Mohammed MA, Degadzor AF, Effrim BF, Appiah KA (2017) Brute force attack detection and prevention on a network using Wireshark analysis. Int Jo Eng Sci Res Technol 6(6):26–37
2. Boyd SW, Keromytis AD (2004) SQLrand: preventing SQL injection attacks. In: International conference on applied cryptography and network security, ACNS 2004
3. Stiawan D et al (2019) Investigating brute force attack patterns in IoT network. J Electr Comput Eng
4. Sowmya G, Jamuna D, Venkata Krishna Reddy M (2012) Blocking of brute force attack. Int J Eng Res Technol (IJERT) 1(6). ISSN: 2278-0181
5. Grover V, Gagandeep (2020) An efficient brute force attack handling techniques for server virtualization. In: Proceedings of the international conference on innovative computing & communications (ICICC)
6. Chanda K (2016) Password security: an analysis of password strengths and vulnerabilities. Int J Comput Netw Inf Secur 7:23–30
7. Patel D, Dhamdhere N, Choudhary P, Pawar M (2020) A system for prevention of SQLi attacks. In: International conference on smart electronics and communication
8. Marras M, Korus P, Memon ND, Fenu G (2019) Adversarial optimization for dictionary attacks on speaker verification. In: 20th Annual conference of the international speech communication association: crossroads of speech and language, INTERSPEECH 2019
9. Puthuparambil AB, Thomas JJ (2019) Freestyle, a randomized version of ChaCha for resisting offline brute-force and dictionary attacks. J Inf Secur Appl 49:102396
10. Subangan S, Senthooran V (2019) Secure authentication mechanism for resistance to password attacks. In: 2019, 19th International conference on advances in ICT for emerging regions (ICTer). IEEE
11. Jose J, Tomy TT, Karunakaran V et al (2016) Securing passwords from dictionary attack with character tree. In: 2016 International conference on wireless communications, signal processing and networking (WiSPNET). IEEE
12. Patel M, Amritha PP, Sam Jasper R (2021) Active dictionary attack on WPA3-SAE. In: Advances in computing and network communications, pp 633–641
13. Corin R, Doumen J, Etalle S (2005) Analysing password protocol security against off-line dictionary attacks. Electron Notes Theoret Comput Sci 121(4):47–63
14. Wang P, Kim Y, Kher V, Kwon T (2005) Strengthening password-based authentication protocols against online dictionary attacks. In: 3rd International conference on applied cryptography and network security, ACNS 2005
15. Bernstein DJ (2007) What output size resists collisions in a xor of independent expansions? In: ECRYPT hash workshop

Task Decomposing Optimization in Wireless Sensor Network

Arpana Mishra , **Rashmi Priyadarshini** , **and R. M. Mehra**

1 Introduction

An efficient line ups of these constellation methods can empress both the total estimation time and the power consumption [1–3]. Evolutionary optimization ways can manipulate this matter successfully due to their qualification to contend nonlinear cases with numerous scheme variables [4, 5]. In this document, an evolutionary algorithm newly evolved, titled Social Network Optimization (SNO), has subsisted referred to the trouble of chore grant in a WSN [6]. The optimization results on two test cases have been anatomized in the first one, no energy checks enjoy subsisted tacked to the optimization, while in the extra bone, a minimal piece of biography cycles is charged [7, 8].

According to new updates in present days smart, tiny and very less cost sensing points are gathering information from various ad hocking positions and aggregating respectively [9]. All sensing points are helping to other for combing and processing the collected information in sensor area network [10–12]. Information flow from collecting points to head supervisor through the multi-hopping [13, 14]. A sensor mode is made up by using sensing units, processing units, communicating units, and energy unit [5] (Figs. 1 and 2).

Communication is taking place among the scattered nodes through the radio frequencies by using a kind of transmitting and receiving modules [16, 17]. Sensed data may be the type of light, audio, video, picture, and text. Several application of WSN broadening load on batteries [18]. Wild life monitoring, health monitoring, traffic monitoring, on ground and underground environmental monitoring are very typical tasks in deployed urban area of wireless sensor network [19].

A. Mishra (✉) · R. Priyadarshini · R. M. Mehra
Sharda University, Greater Noida 201306, India
e-mail: shubhparna20@gmail.com
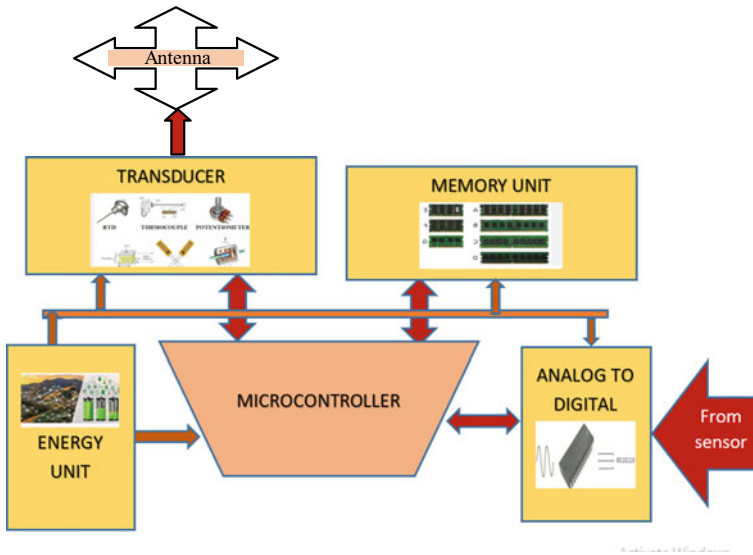
Fig. 1 Basic block diagram of WSN

Required threshold of energy is increasing day by day in wireless sensor network tremendous increments in work load of typical task [20]. Military purpose application of WSN is highly sensitive in all perspective of security [21]. Maximum throughput, high accuracy, less redundancy and maximum life time [22].

Multiple layers in communication model of WSN are an easy approach to fulfill the optimized aspects. Battery life in WSN can be increased by many years by suitable approach [23, 24]. Minimax sampling utilizes for better memory constraint when using the large scale network [25].

Clusterization is the process in which a vast area of cells divided into number of groups. All cells consist the sensing points. From among nodes of group, one node is selected as a cluster head at one time. Selected cluster head (CH) is working as group monitor and responsible to communicate with other clusters [25]. The communication may be multiple or single hop. The simple node member of a cluster is responsible only for the collection of data from their surroundings. Aggregation, processing, and transmissions to corresponding CH and base station are performed by the cluster head in particular cluster [26]. Some ordinary sensing points performing the task of passing the route as gate-way to the next cluster.

Many clusterization algorithms are not concerning about the energy optimization. Since the WSN is ad-hoc network, energy constraint is the peak focusing point. Several algorithms are proving a better quality services in their clusterization process. Low Energy Adaptive Clustering Hierarchy (LEACH) is consisting the hierarchical routing instructions for energy saving in WSN [22].

Decomposition of the task is also named as task analytics. After identifying the typical task, the subdivisions are done for easy accomplishment in wireless sensor
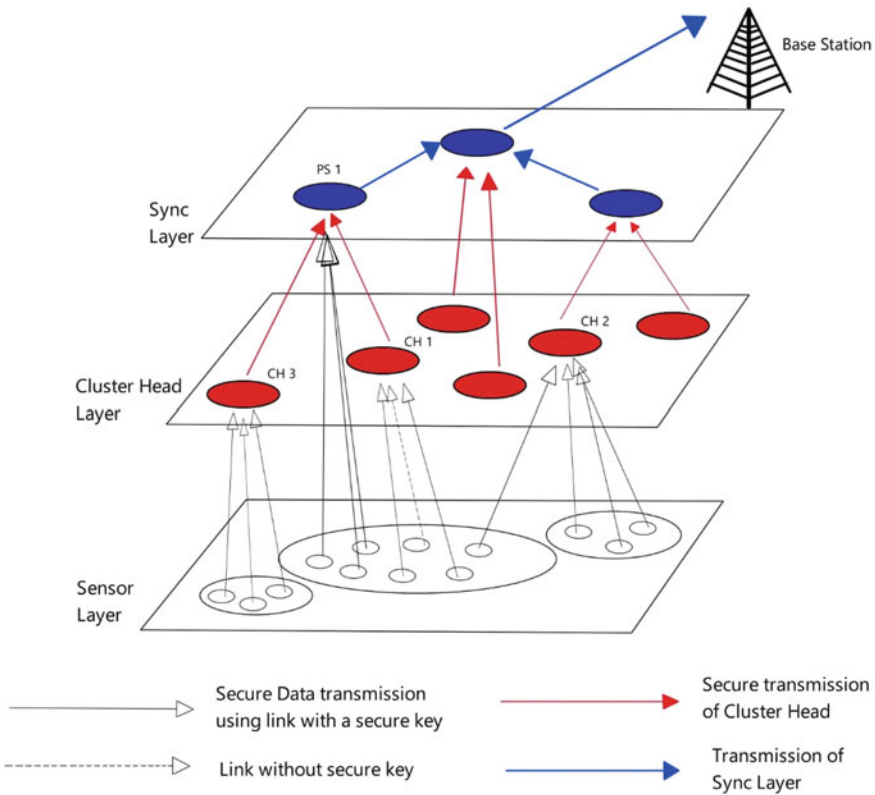
Fig. 2 Layered structure of WSN

network. Selected Energy Protocol (SEP) is one more algorithm in heterogeneous network. More levels of energy can be saved using SEP algorithm [21]. This method is different from other algorithms in selection procedure of cluster supervisor [20].

Layered supervisor and sub-supervisor are decided in random manner. A uniform energy is required among all distributed sensing points in SEP algorithm [19]. Communication distance among the nodes is an important parameter for the energy saving purpose. Decreased communication distance increases the battery life time and prolonging the network [8, 17].

For complex task in wireless sensor network, this distance required more basic resources. In this article, we proposed an algorithm to reducing the distance among the nodes while performing the typical task in wireless sensor network. This paper is contributing improved energy efficiency by shorting the communication distance among the nodes in wireless sensor network. Simulation work is conducted to see the performance of proposed algorithm. This paper is properly organized to represent the flow of work. With the introduction in starting this proposed algorithm represents

flow diagram, mathematical modeling, and simulation. Conclusion and future scope are represented at the end of paper.

2 Reduced Typical Task Distancing Algorithm in WSN (RTDW)

Clusterization, in optimized manner, is very effective way to energy saving and ignoring the long distance traveling in network while performing the typical task. Aggregated data move toward the sub-layer from primary layer of nodes. Resources are offered by the clusters to minimizing the number of sensing points while transmitting at large distance. In performance of typical task, clusterization involves several numbers of rounds. Transmission of typical task load is one of the primary concerns to minimizing the distance and energy both.

Some assumptions for deciding the cluster supervisor.

- Residual energy at the starting are equal for are sensing points
- Total number of sensing points are N
- Total number of clusters are K.

Average number of sensing points in each cluster is P

$$P = N/K \tag{1}$$

A cluster head consists a supervisor at the end of after P phrases/rounds. Phrases represent the cycles or round for performing the task in wireless sensor network.

P_i represents the ith average node, where $i = 1, 2, 3, 4, (N - 1)$.

In starting no cluster head in cluster $j = 0$, where j is the CHs counter. A node r is selected and runs all the following steps.

Firstly, r increments i by 1 and check if i is even, if yes that node is selected as supervisor for same phrase and acknowledges its new positions to all nodes. Else if i odd it will not become cluster supervisor for same phrase.

Probability of sensing point to become new cluster supervisor is represented as

$$T_i = \left(E_{\text{remaining}}(i) * K\right) / \left(E_{\text{average}}(i) * P\right) \tag{2}$$

where $E_{\text{remaining}}(i)$ represents the energy left in ith node and $E_{\text{average}}(i)$ represents the average energy in ith node.

3   Reduced Typical Task Distancing Cluster Formation

For cluster formation in typical task performance hierarchal steps are followed in both cases of intra-cluster and inter-cluster as shown in diagram. Resources requirement in well doing tasks decides the communication distance and cluster head selection. Heterogeneity in cluster's resources extends the network life time and distance among sensing points. Hybrid TDMA allocation increases the bandwidth optimization.

Difficult task decomposition results more numbers of hopping; thus, multi-hopping procedure needs more powers if using some ordinary algorithms. TPCAHH (Typical Task protocol allocation in heterogeneous hierarchy) is better suited for distance perspective to decompose the load in wireless sensor network. Latency and localization for power management in typical tasking are better controllable using GPSA (Greedy particle swarm algorithms). Dynamic programming (DP), NP-hard, integer linear programming (ILP), and Discrete Dynamic voltage scaling (DVS) are proving better decompositions by using acute computations.

```
┌────────────────────────────────────────────────┐
│ Primary cluster head send the message to non-supervisor │
│ nodes for further communication using network protocol. │
└────────────────────────────────────────────────┘
                        │
                        ▼
┌────────────────────────────────────────────────┐
│ Sensing points decides the cluster to forward their typical │
│ task decomposed data at minimum distant point │
└────────────────────────────────────────────────┘
                        │
                        ▼
┌────────────────────────────────────────────────┐
│ Non supervisor acknowledge to the Cluster head │
└────────────────────────────────────────────────┘
                        │
                        ▼
┌────────────────────────────────────────────────┐
│ Cluster head received Hybrid TDMA for decomposed forwarded data │
└────────────────────────────────────────────────┘
                        │
                        ▼
┌────────────────────────────────────────────────┐
│ All sensing points uses hybrid TDMA to transmit the │
│ data using multi hop │
└────────────────────────────────────────────────┘
```

4 Simulation Work and Results

The primary goal of the proposed article is to minimize the distance for the energy saving prospection in typical task performance in wireless sensor network. Simulation work has done on MATLAB. Execution results are compared with MLEACH (Modified Low Energy Adaptive Clustering Hierarchy) and DLLEACH (Distributed Load Low Energy Adaptive Clustering Hierarchy) on MATLAB platform.

To observe the extended saved battery time, 200 sensing points are deployed randomly as represented in Fig. 3. Axis measures between (0, 0) and (200, 200) m. Base station sets at coordinates (70, 380).

4.1 Energy Dissipation Model

Simple radio transmitter model utilizes the k bit data to transmit at distance d.

$$
\begin{aligned}
E_{tx}(k, d) &= E_{\text{elec}} * k + E_{\text{amp}} * k \\
&= E_{\text{elec}} * k + \varepsilon_{fs} + d^2 * k \quad \text{where } d < d_o \\
&= E_{\text{elec}} * k + \varepsilon_{fs} + d^4 * k \quad \text{where } d \geq d_o
\end{aligned} \tag{3}
$$

To receive k bit of data energy is calculated as

$$
E_{rx}(k) = E_{\text{elec}} * k \tag{4}
$$
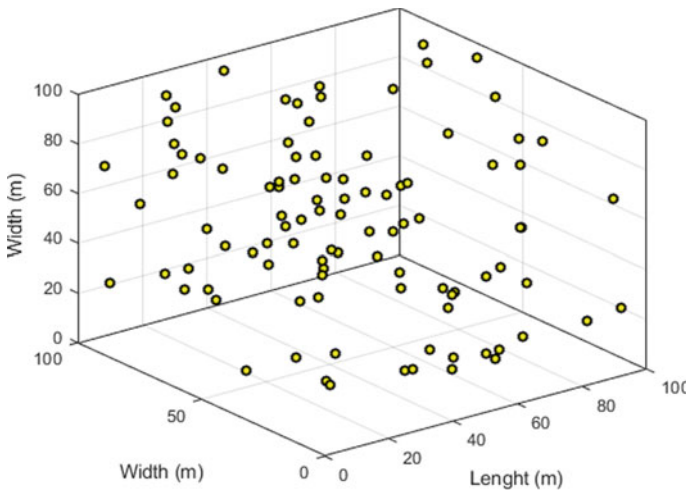
Fig. 3 Random deployment of sensing points

Table 1 Simulation features

Features	Dimensions
Network dimension	100 m × 100 m
Node energy at starting (K)	0.6 J
Electric power dissipation at sender	60 nJ/bit
Multipath amplification dissipation	0.0016 pJ/bit/m^4
Packet size of data	100 bytes
Space state amplifier dissipation(ε_{fs})	10 pJ/bit/m^2

E_{elec} is the energy expended per bit to operate the transmitter and receiver circuit as shown in Eqs. 3 and 4. In this work, ordinary and multipath fading power loss are considered. Threshold value of d_o is expressing the distance between transmitter and cluster supervisor receiver.

The values of simulation features are considered as given in Table 1. Simulated model has been tested by using these values with MATLAB.

4.2 Performance Evaluation, Analysis, and Comparisons

100 nodes are clusturized in five clusters as shown Fig. 4. Each cluster consists a head for supervisions of all activities in cluster. Defining the task as typical or ordinary also performed by cluster head. Proposed algorithm (RTDW) is decomposing the typical task as per minimum selected path to transmit.

Comparisons of performance evaluate better marking points. We compare with other protocols DLLEAC and MLEACH in the manner of number of phrases.

● Cluster head   ● Sensing points
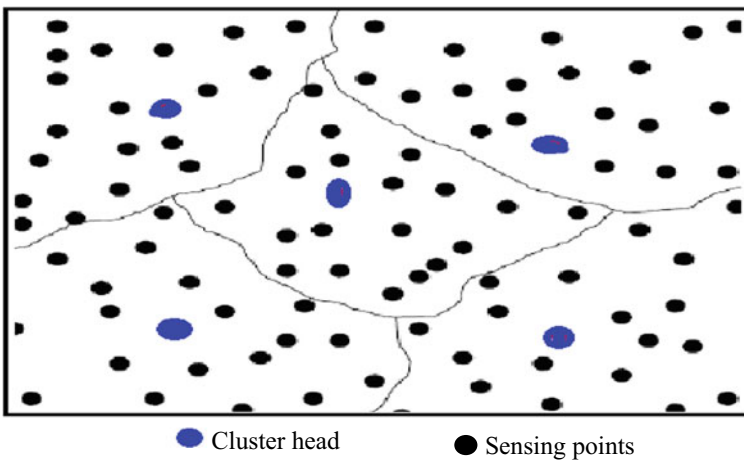
Fig. 4 Cluster formation

Phrases time taken from first node death to last node death. As shown in Figs. 5 and 6, RTDW is providing better results in number of surviving nodes and quantity of transmission data.
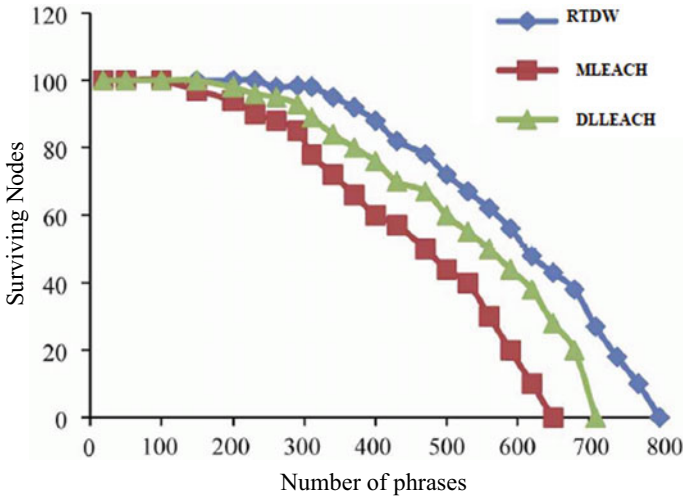
Fig. 5 Comparative graphs of surviving nodes after applying phrases
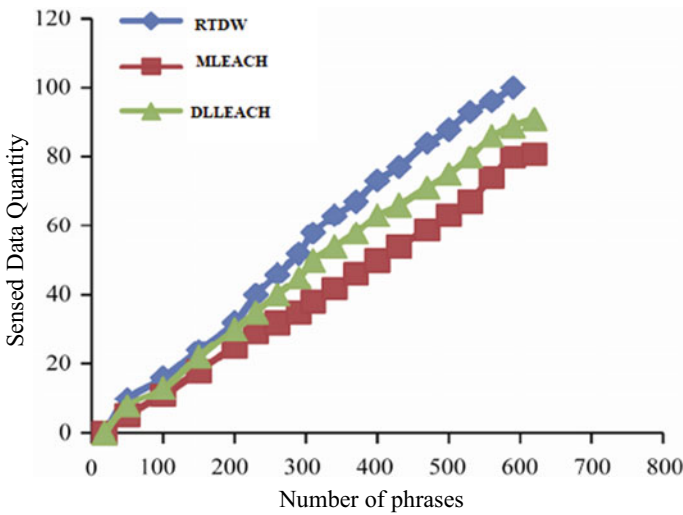
Fig. 6 Comparative graphs of sensed data quantity after applying phrases

5 Conclusion and Future Work

In this article, proposed algorithm RTDW is providing a mark able solution for typical task performance in WSN. Minimization of the distance is providing the better life time. Firstly distribute the network in clusters then selection procedure has been done of head on the basis minimum distance. This algorithm saving the energy even task is typical in WSN. In the future, this distance can further minimize if considering hybrid transmission model. In progressive version, we can compare with other more algorithms.

References

1. Pilloni V, Atzori L (2011) Energy-efficient task allocation for distributed applications in wireless sensor networks. In: Proceedings of the GLOBECOM workshops, Houston, TX. IEEE, New York, pp 321–326
2. Zhang J, Tang J, Wang Z (2020) Load-balancing rendezvous approach for mobility-enabled adaptive energy efficient data collection in WSNs. KSII T Internet Inf 14:1204–1227
3. Sun Z, Liao G, Zeng C, Lan L, Zhao G (2022) GLBR: A novel global load balancing routing scheme based on intelligent computing in partially disconnected wireless sensor networks. Int J Distrib Sens Netw 18(4):15501329221090458
4. Wang J, Liu Y, Niu S (2021) Lightweight block chain assisted secure routing of swarm UAS networking. Comput Commun 16:131–140
5. Zhang K, Han Q, Cai Z (2019) DOAMI: a distributed on-line algorithm to minimize interference for routing in wireless sensor networks. Theor Comput Sci 07:1–13
6. Kavitha A, Guravaiah K, Velusamy RL (2020) CS-CGMP: clustering scheme using Canada Geese Migration Principle for routing in wireless sensor networks. Wirel Pers Commun 15:1363–1384
7. Dargie W, Poellabauer C (2010) Fundamentals of wireless sensor networks: theory and practice. Wiley
8. Mishra A, Priyadarshini R, Choudhary S, Mehra RM (2021) Qualitative analysis of intra-class and inter-class clustering routing and clusterization in wireless sensor network. EVERGREEN Joint J Novel Carbon Resour Sci Green Asia Strat 08(02):358–373
9. Dietrich I, Dressler F (2009) On the lifetime of wireless sensor networks. ACM/ IEEE Trans Sens Netw 5:1–39
10. Peng, Zhang K, Han Q, Cai Z (2019) DOAMI: a distributed on-line algorithm to minimize interference for routing in wireless sensor networks. Theor Comput Sci 774:1–13
11. Jiang Y, Zhou Y, Wang W (2013) Task allocation for undependable multi agent systems in social networks. IEEE Trans Parallel Distrib Syst 4:1671–1681. IEEE
12. Wang J, Ma T, Cho J (2011) An energy efficient and load balancing routing algorithm for wireless sensor networks. Comput Sci Inf Syst 8:991–1007
13. Giannecchini S, Caccamo M, Shih CS (2004) Collaborative resource allocation in wireless sensor networks. In: Proceedings of the 16th Euro micro conference on real-time systems (ECRTS 04), Catania, IEEE, New York, pp 35–44
14. Xiao W, Low S, Tham C (2009) Prediction based energy-efficient task allocation for delay-constrained wireless sensor networks. In: Proceedings of the 6th IEEE annual communications society conference on sensor, mesh and ad hoc communications and networks, Rome, IEEE, New York, pp 1–3
15. Li W, Delicato FC, Pires PF (2014) Efficient allocation of resources in multiple heterogeneous wireless sensor networks. J Parallel Distr Com 74(1):1775–1788

16. Wang J, Li B, Xia F (2014) An energy efficient distance-aware routing algorithm with multiple mobile sinks for wireless sensor networks. Sensors 14:15163–15181
17. Wang J, Cao Y, Ji S (2017) Energy-efficient cluster-based dynamic routes adjustment approach for wireless sensor networks with mobile sinks. J Super Comput 73(7):3277–3290
18. Shih E, Cho S, Ickes N (2001) Physical layer driven protocol and algorithm design for energy-efficient wireless sensor networks. In: Proceedings of 7th annual international conference on mobile computing and network. ACM, Rome, New York, pp 272–286
19. Wang A, Chandrakasan A (2002) Energy-efficient DSPs for wireless sensor networks. IEEE Signal Proc Mag 19(4):68–78
20. McMahan HB, Likhachev M, Gordon GJ (2005) Bounded real-time dynamic programming: RTDP with monotone upper bounds and performance guarantees. In: Proceedings of the 22nd international conference on machine learning, Bonn. ACM, New York, pp 569–576
21. Li W, Delicato FC, Zomaya AY (2013) Adaptive energy-efficient scheduling for hierarchical wireless sensor networks. ACM T Sens Netw 9(3):1–34
22. Jin Y, Jin J, Gluhak A (2012) An intelligent task allocation scheme for multihop wireless networks. IEEE T Parall Distr 23(3):444–451
23. Guo W, Xiong N, Chao HC (2011) Design and analysis of self-adapted task scheduling strategies in wireless sensor networks. Sensors 11:6533–6554
24. Guo W, Li J, Chen G (2015) A PSO-optimized real-time fault-tolerant task allocation algorithm in wireless sensor networks. IEEE T Parall Distr 26(12):3236–3249
25. Luo T, Tan HP, Quek T (2012) Sensor Open Flow: enabling software-defined wireless sensor networks. IEEE Commun Lett 6(11):1896–1899
26. Zeng D, Li P, Guo S (2015) Energy minimization in multi-task software-defined sensor networks. IEEE T Comput 64(11):3128–3139
27. More A, Raisinghani V (2017) A survey on energy efficient coverage protocols in wireless sensor networks, Science Direct. J King Saud Univ 29(4):428–448
28. Elshrkawey M, Samiha M, Elsherif M, Wahed E (2018) An enhancement approach for reducing the energy consumption in wireless sensor networks, Science Direct. J King Saud Univ Comput Inf Sci 30:259–267
29. Sikandar A, Kumar S, Singh P, Tyagi MK, Kumar D (2018) Energy efficient transmission in the presence of interference for wireless sensor networks. In: ICACCT 2018, CCIS 899. Springer Nature Singapore, pp 55–64

LMSF: Lightweight Minimal Scheduling Function for 6TiSCH Networks

Karnish N. A. Tapadar, Priyanshu Singh, and Manas Khatua

1 Introduction

Recently, the number of Internet-connected devices has risen significantly and given rise to a new domain called Internet of Things (IoT). Internet Engineering Task Force (IETF) has standardized several network protocols to empower IoT. The IETF's 6TiSCH working group aims to provide IPv6 interoperability with the Time Slotted Channel Hopping (TSCH) Medium Access Control (MAC) layer. Apart from addressing interoperability challenges, one more crucial task is TSCH cell scheduling. TSCH communications occur within a slotframe, a collection of slots that automatically repeats over time. A slotframe consists of cells, represented by a pair of (*slotOffset*, *channelOffset*), where *slotOffset* indicates the position of the timeslot, and *channelOffset* maps to a real physical channel. A TSCH schedule instructs a node when to transmit or receive a packet. However, the TSCH standard does not define how to create or manage a communication schedule. Considering that, the 6TiSCH architecture presents a scheduling function whose task is to specify the scheduling policies to follow for communication. Several scheduling algorithms were proposed to satisfy different application requirements. However, only some of them adapt to traffic demand. The 6TiSCH working group recently introduced the Minimal Scheduling Function (MSF) [2] as a reference scheduling standard for the 6TiSCH network. MSF decides when and how many cells to add or delete from the schedule, while the 6P protocol is responsible for node-to-node negotiation for cell

K. N. A. Tapadar (✉) · P. Singh · M. Khatua
Indian Institute of Technology Guwahati, Guwahati, Assam 781039, India
e-mail: karnish_nasrin@iitg.ac.in

P. Singh
e-mail: spriyanshu@iitg.ac.in

M. Khatua
e-mail: manaskhatua@iitg.ac.in

addition or deletion. It is observed that MSF incurs a huge controlling overhead on 6P, as MSF requires multiple consecutive 6P transactions to bring the cell usage in the stable zone, i.e. within the 25 and 75%. Further, MSF suffers in two ways: it does not proactively act on traffic demand and updates (adds or deletes) its scheduled cells by unity, regardless of the heavy changes in traffic demand.

Therefore, this paper proposes a Lightweight Minimal Scheduling Function (LMSF) to overcome the limitations of standard MSF. The LMSF uses well-established Poisson distribution for future traffic estimation for a fixed duration. Then, based on its rate, a demand factor is decided to keep cell utilization approximately in the middle of the stable region after adding (or deletion) cells in (or from) the schedule. Accordingly, the 6P negotiation is adjusted so that a node can directly adapt to high changes in traffic demand. The contributions of this paper can be summarized as follows:

- A traffic demand factor is calculated depending on the Poisson distribution-based traffic estimation.
- Using the demand factor, Lightweight Minimal Scheduling Function (LMSF) is proposed to adjust traffic demand changes dynamically while the 6P negotiation overhead remains low.
- Finally, simulation-based performance evaluation is carried out using the 6TiSCH simulator to evaluate the proposed scheme.

The rest of the paper is structured as follows. Section 2 provides an overview of the 6TiSCH architecture and MSF traffic adaptation. Section 3 presents the summary of existing works on link scheduling in 6TiSCH networks. Section 4 presents our proposed scheme. Section 5 offers a simulation-based performance evaluation of the proposed approach. Finally, Sect. 6 concludes this paper.

2 Overview of 6TiSCH Architecture

2.1 6TiSCH

A 6TiSCH architecture has been proposed to empower the Industrial Internet of Things (IIoT). It enables the connection of existing IPv6 networks with low-power TSCH-enabled devices. TSCH fails to maintain global synchronization and does not specifies the protocols for the allocation of communication slots. Considering this, the working group has included a new intermediate layer called 6top on top of TSCH to manage TSCH cell scheduling. A more detailed discussion on 6TiSCH is presented in [12].

2.2 6top Protocol (6P)

In addition to 6TiSCH-MC, 6TiSCH carries dynamic scheduling. TSCH cells are added or removed dynamically to match the traffic requirements. The 6top Protocol (6P) works along the scheduling function to manage dynamic scheduling. A scheduling function decides when to trigger 6P negotiation, and 6P protocol performs cell negotiation to add/delete/relocate cells between two neighbours. This negotiation can be conducted in 2 or 3 step transactions. The node initiating the 6P will select the candidate cells in a 2 step transaction, while in the latter case, the sender node allows the other node to choose candidate cells.

2.3 Traffic Adaptation in MSF

MSF runs on top of 6TiSCH-MC [13] for dynamic scheduling. MSF maintains autonomous and negotiated cells. The primary purpose of autonomous cells is to perform 6P communication, while negotiated cells transmit actual data packets. MSF adds/deletes/relocate cells to match traffic with link-layer resources, handles parent changes and manages cell collision. However, this work only focuses on the traffic adaptation part. Before discussing the remaining part, note that Table 1 summarizes the frequently used abbreviations in this paper.

Adaptive traffic means that sometimes the traffic load is high and sometimes low, so static allocation of cells does not work because packets will be dropped frequently

Table 1 Important notations

NumCellUsed	The number of negotiated cells utilized for transmission and reception.
NumCellElapsed	Total number of negotiated cells, regardless of whether they are used as Tx or Rx cells
Cell utilization	$NumCellUsed/NumCellElapsed$
LIM_NUMCELLUSED_HIGH	Upper boundary of the stable zone (75%)
LIM_NUMCELLUSED_LOW	Lower boundary of the stable zone (25%)
Tx_CELLS	Negotiated Tx cells per slotframe
len(TX_CELLS)	Returns the entire number of Tx_CELLS
MAX_NUMCELLS	The length of running window to calculate cell utilization. Here, it is 16
EstAvgCellUsed	It is the predicted value of NumCellUsed per slotframe for the next cycle
demand (d)	Demand is a factor to make cell utilization 50% approximately, $demand(d) = 2 * EstAvgCellUsed$

if traffic is high. Similarly, if the traffic demand is low and the allocated cells are high, there will be a waste of TSCH cells. Further, it will be less energy efficient.

MSF maintains two counters, namely *NumCellElapsed* and *NumCellUsed*. When a new cell is added through negotiation, it increments *NumCellElapsed* by 1. Again, if a negotiated cell is used to transmit or receive frames, it increments *NumCellUsed* by 1. Cell usage is calculated by the ratio between *NumCellUsed* and *NumCellElapsed*. It is calculated only when *NumCellElapsed* reaches the predefined threshold value of *MAX_NUMCELL*. If the cell utilization goes above the upper threshold *LIM_NUMCELLSUSED_HIGH*, MSF triggers a 6P request and negotiates to add a new cell. On the other hand, if the ratio is less than *LIM_NUMCELLSUSED_LOW*, it triggers a 6P request to remove a cell. Every time it estimates cell usage, the counters are reset to zero and run the same process.

3   Related Work

Designing a TSCH schedule is application-specific, and thus, several scheduling approaches were designed to schedule TSCH communication. Scheduling approaches can be autonomous, distributed and centralized.In centralized schemes, the central node takes charge of creating the cell schedule for all other nodes in the network. Authors in TASA [8] use recursive matching and colouring to determine the centralized TSCH link schedule. Their main goal was to find a minimum length conflict-free schedule. Most centralized algorithms have huge control overhead for to and forth information flow from all the nodes to the coordinator and vice versa. Further, they fail for large and dynamic networks as every time a change happens; the schedule needs to be recomputed and redistributed.

In distributed schemes, all nodes decide their wake-up and sleep time autonomously by negotiating with their neighbours based on some local, partial information exchanged between them. Here, it is challenging to find interfering links until a collision happens. Therefore, computing a distributed schedule is more challenging than centralized methods. DeTAS [1] is an example of this approach for converge cast networks. Under autonomous category, Orchestra [4] relies on RPL structure to compute each node's schedule. ALICE [7] and OST [6] were proposed as an extension of Orchestra, which follow link-based hashing for resource allocation.

There are many scheduling algorithms explicitly designed for 6TiSCH network architecture. Palattela et al. [9] allocates TSCH resources based on the traffic demand. However, it did not mention how traffic estimation should be carried out. Later, Righetti et al. [10] conducted a performance analysis on OTF and proposed a revised OTF, namely EOTF. Daneels et al. [3] focuses on minimizing the latency of recurrent traffic to build a minimal latency path and activate recurrent cells whenever recurrent traffic encounters. However, it is difficult to maintain such a low latency path. Recently, Tapadar et al. [11] proposed IMSF, where the authors estimate the traffic demand of a node. Accordingly, based on the estimation, they decide the amount of cells to be added/removed from the schedule. MSF [2] follows a reactive mechanism

on changes in traffic demand; thus, it may dynamically cope to traffic variations, but it does not utilize traffic estimation.

In brief, all the scheduling functions discussed above improve different QoS requirements. The majority of them, however, are not adaptable to traffic demand. Further, none guarantee minimum 6P control overhead and minimum latency while increasing reliability.

4 Proposed Methodology

As discussed, the addition of one cell at a time is not always sufficient to reach the stable zone because it may take several 6P transactions, resulting in a significant 6P control overhead. To overcome the limitations, we propose Lightweight Minimal Scheduling Function (LMSF). LMSF works as follows:

- Estimates the expected traffic demand (i.e. *NumCellUsed*) for the next cycle
- Demand factor (d) is computed based on the value of calculated *NumCellUsed*, which indicates the amount of cells to be added or deleted in the schedule to reach the stable zone.
- Finally, the 6P transaction is carried out according to the current cell usage and demand factor.

4.1 Prediction of NumCellUsed and Compute Demand Factor

We formulate the prediction of the future traffic demand using the Poisson process as the system satisfies the following two properties:

- Property 1: Transmission cells (*Tx_CELLS*) used by every node in the network is an independent event.
- Property 2: Time between each slotframe is finite.

Based on the properties 1 and 2, we use the Poisson distribution model in which the parameter λ is the mean number of events that occurs per unit of time. Here, we consider λ as the average *NumCellUsed* in the previous cycle. The probability mass function is given by formulae:

$$P\left(X=k\right)=\frac{e^{-\lambda}\lambda^{k}}{k!} \tag{1}$$

where e represents Euler's number and k indicates the number of occurrences. Using Eq. (1), we get the number of occurrences k at which the distribution gives highest probability for a given value of λ. This index k represents the expected traffic

demand (*EstAvgCellUsed*) per slotframe for the next cycle. Now, the demand factor is computed as:

$$demand(d) = 2 * EstAvgCellUsed \qquad (2)$$

Justification of demand factor calculation is explained at the end of this section (Sect. 4.3).

4.2 Decision on Cell Addition/Deletion

Based on the value of d, LMSF decides whether to add, delete or do nothing for the cells as follows:

- If the predicted value of d is more than $len(Tx\_CELLS)$ and the cell utilization exceeds *LIM_NUMCELLUSED_HIGH*, then LMSF decides to have cells addition for the next slotframe window.
- If the predicted value of d is less than the value of $len(Tx\_CELLS)$ and cell utilization falls below *LIM_NUMCELLUSED_LOW*, then LMSF decides to remove certain cells from the current cell schedule.
- If the predicted value of d is smaller than the value of $len(Tx\_CELLS)$ or the cell utilization lies between the two thresholds, then LMSF will take no action.

The objective of cell addition/deletion in LMSF is to bring the cell utilization approximately in the middle of the stable zone, i.e. *(LIM_NUMCELLUSED_HIGH + LIM_NUMCELLUSED_LOW)/2*. The reason for using the 50% idea is that if the predicted value slightly deviates from the actual value, the cell utilization will still be in the stable zone, i.e. within 25% and 75%. Algorithm 1 summarizes the above steps.

Algorithm 1 Lightweight MSF (LMSF)

λ = **ceil**(NumCellUsed / 16)
EstAvgCellUsed = $\underset{k}{\text{argmax}} \, P(X = k)$
d = 2*EstAvgCellUsed
if d > len(Tx_CELLS) **and**
  LIM_NUMCELLUSED_HIGH < cell_utilization **then**
    add(d - len(Tx_CELLS))
end if
if d < len(Tx_CELLS) **and**
  LIM_NUMCELLUSED_LOW > cell_utilization
 then
   if d > 1 **then**
      delete(len(Tx_CELLS) - d)
   end if
end if

4.3 Explanation with Example

To understand LMSF, let us consider the below scenario. We assume that there is x number of *Tx_CELLS*. From the prediction, we get *EstAvgCellUsed*, which is the expected average value of *NumCellUsed* for the next 16 slotframes. To maintain cell utilization in the 50% range, we consider $demand(d) = 2 * EstAvgCellUsed$. Then we get the count of cells to be added as follows:

$$y = d - x \quad \text{(as per \textbf{add} equation in LMSF)} \tag{3}$$

After adding this y cells with the already scheduled negotiated cells, i.e. x, we get

$$
\begin{aligned}
t &= y + x \quad \text{(total number of TX\_CELLS)} \\
&= d - x + x \quad \text{(from Equation 3)} \\
&= d
\end{aligned}
$$

Finally, cell utilization will become $EstAvgCellUsed/d$. As $d = 2 * EstAvgCellUsed$, it will result in $cell\_utilization = 0.5$.

Similarly, for delete operation, if we delete y number of cells, then

$$y = x - d \quad \text{(as per \textbf{delete} equation in LMSF)} \tag{4}$$

After deleting this y cells from the already scheduled negotiated cells, i.e. x, we get,

$$
\begin{aligned}
t &= x - y \quad \text{(Remaining number of TX\_CELLS)} \\
&= x - (x - d) \quad \text{(from Eq. 4)} \\
&= d
\end{aligned}
$$

Again after deletion, the total number of *Tx_CELLS* becomes d. Thus, cell utilization will be $EstAvgCellUsed/d$. Since $d = 2 * EstAvgCellUsed$, $cell\_utilization$ will be 0.5.

5 Simulation Results

This section discusses the operational performance of the proposed LMSF using an open-source 6TiSCH simulator for 6TiSCH networks. 6TiSCH simulator is Python based and follows a discrete-event simulation model and thus models the task or operation as a sequence of events. It supports all the 6TiSCH Standard Protocol stack. Precisely, it can monitor the scheduling function, routing protocol, link-layer 6TiSCH packet drops and response from the application. However, it does not affect

Table 2 Simulation parameters

Slotframe iteration	3600
Number of runs	30
Slotframe Size	101
Timeslot duration	10 ms
Network deployment	Random
Tx queue buffer	10 packets
Number of nodes	3–25
Packet length	90 bytes
MAC max retransmission	5

the parent's selection, as each node can hear only from its two nearest neighbours. For our experiment, we used random topology. This aim avoids the influence of a specific topology in the network. It implements RPL MRHOF [5] objective function. The simulation parameters and their respective values are shown in Table 2. Each simulation is conducted 30 times, with 3600 slotframes in each iteration.

5.1 Results and Discussions

Figure 1a represents the 6P overhead in the network to exchange 6P negotiation messages. We have compared the 6p overhead against the network size. It has been seen that the 6P overhead rises as nodes increase. As the simulation is run multiple times, we took into consideration the average 6P packets sent by each network node. Results show that LMSF outperforms the MSF. The reason is that MSF adds one cell at a time, irrespective of the traffic demand. Thus, MSF may need many consecutive 6P transactions to reach the stable zone. However, LMSF can add the required cells with a single 6P transaction, thus reducing the 6P overhead significantly over the network lifetime.

Latency is calculated from the timestamp when a packet gets generated from the application layer of the source node to the time until it reaches the application layer of the destination node. We have calculated the average latency for different network sizes as shown in Fig. 1b. We observe that even for a higher number of nodes and heavy traffic, the average latency in LMSF does not cross 0.3 s. This improvement is due to the proposed scheduling mechanism that successfully improves resource allocation.

Reliability is determined by dividing the total number of data packets which are successfully received at the destination node by those packets that were actually sent by all the nodes in the network. Figure 1c shows that LMSF performs better than MSF due to its ability to reduce packet losses. When the queue gets full due to the

(a) 6p signalling overhead

(b) End-to-end latency

(c) Reliability

(d) Energy consumption
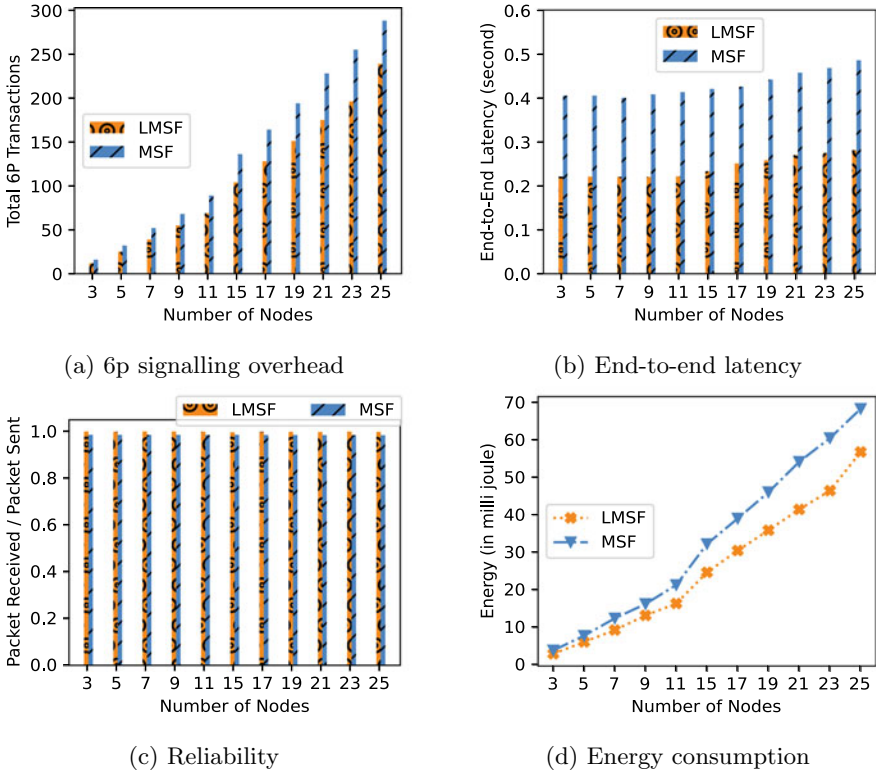
Fig. 1 Comparison of 6P overhead, end-to-end latency, reliability and energy consumption under the varying network size

unavailability of TSCH cells, packets get dropped. The difference in PDR is almost constant for different network sizes between MSF and LMSF scheduling functions.

The number of 6P packets is reduced by a fair amount which saved the session time effectively. The proposed scheme can fulfil the same purpose by sending fewer 6p packets, therefore decreasing the redundant processing time of packets by each node. Because of that, the active time of the mote also decreases, thereby decreasing the overall energy consumption. Figure 1d represents the comparison of energy consumed between MSF and LMSF, where it can be observed that, with increasing the number of nodes, LMSF performs better than MSF in terms of energy consumption.

6 Conclusion

This paper proposes a Lightweight Minimal Scheduling Function (LMSF) for 6TiSCH networks. First, we have discussed MSF's limitations in traffic adaptation. To overcome this problem, we introduced LMSF, which uses Poisson distribution to

predict the future required number of cells. LMSF can add and delete single or multiple cells at a time depending on how close or far the cell usage is from the stable region (25–75%). The 6TiSCH simulator is used to validate the proposed scheme, and the obtained results are compared with the standard MSF. Achieved results proved that the proposed LMSF considerably reduces 6P negotiation overhead with improved latency, reliability and energy consumption.

References

1. Accettura N, Vogli E, Palattella MR, Grieco LA, Boggia G, Dohler M (2015) Decentralized traffic aware scheduling in 6TiSCH networks: design and experimental evaluation. IEEE Internet Things J 2(6):455–470. https://doi.org/10.1109/JIOT.2015.2476915
2. Chang T, Vucinic M, Vilajosana X, Duquennoy S, Dujovne D (2021) 6TiSCH Minimal scheduling function (MSF). Internet Engineering Task Force, RFC, p 9033
3. Daneels G, Spinnewyn B, Latré S, Famaey J (2018) ReSF: recurrent low-latency scheduling in IEEE 802.15.4e TSCH networks. Ad Hoc Netw 69:100–114
4. Duquennoy S, Nahas BA, Landsiedel O, Watteyne T (2015) Orchestra: robust mesh networks through autonomously scheduled TSCH. In: Proceedings of of the 13th ACM conference on embedded networked sensor systems (SenSys), NY, USA, pp 337–350
5. Gnawali O, Levis P (2012) The minimum rank with hysteresis objective function. Internet Engineering Task Force, RFC6719
6. Jeong S, Kim HS, Paek J, Bahk S (2020) OST: On-demand TSCH scheduling with traffic-awareness. In: Proceedings of IEEE INFOCOM, pp 69–78
7. Kim S, Kim HS, Kim C (2019) ALICE: autonomous link-based cell scheduling for TSCH. In: Proceedings of the 18th international conference on information processing in sensor networks, pp 121–132
8. Palattella MR, Accettura N, Dohler M, Grieco LA, Boggia G (2012) Traffic aware scheduling algorithm for reliable low-power multi-hop IEEE 802.15.4e networks. In: Proceedings of 23rd international symposium on personal, indoor and mobile radio communications—(PIMRC), pp 327–332. https://doi.org/10.1109/PIMRC.2012.6362805
9. Palattella MR, Watteyne T, Wang Q, Muraoka K, Accettura N, Dujovne D, Grieco LA, Engel T (2016) On-the-fly bandwidth reservation for 6TiSCH wireless industrial networks. IEEE Sens J 16(2):550–560
10. Righetti F, Vallati C, Anastasi G, Das SK (2018) Analysis and improvement of the on-the-fly bandwidth reservation algorithm for 6TiSCH. In: Proceeding of IEEE 19th international symposium on a world of wireless, mobile and multimedia networks (WoWMoM), pp 1–9
11. Tapadar KN, Khatua M, Tamarapalli V (2022) Imsf: improved minimal scheduling function for link scheduling in 6tisch networks. In: Proceedings of 23rd international conference on distributed computing and networking, pp 124–127
12. Vilajosana X, Watteyne T, Chang T, Vuini M, Duquennoy S, Thubert P (2019) IETF 6TiSCH: a tutorial. IEEE Commun Surv Tutor, pp 1–21
13. Wang Q, Vilajosana X, Watteyne T (2018) 6TiSCH operation sublayer (6top) protocol (6P). Internet Engineering Task Force, RFC8480

An Efficient Content-Based Image Retrieval Using Threefold Technique

Nepoleon Keisham and Arambam Neelima

1 Introduction

Digital images have become more accessible to the public with the development of multimedia content available over the Internet and on digital devices. A significant role is played by digital images in a variety of disciplines, including entertainment, medicine, news, and safety. Furthermore, processing digital images of huge amounts become a serious problem. To retrieve digital images from a huge database, an effective and efficient retrieval system is required. Similar image search capabilities can be enhanced by a CBIR system, particularly for images with multiple languages tagged and annotated. CBIR system is composed of two phases. The initial phase, referred to as indexing, involves the extraction of features from the dataset, and these features are saved as feature vectors. In the next phase, the features are obtained from the image query and match up to the feature vectors from the previous step. Many CBIR techniques have been developed utilizing a single feature; however, the results have been disappointing. After that, numerous CBIR techniques came up using the combination of multiple features like shape, texture, color, spatial information, etc. The results are quite promising although it takes a lot of time while retrieving similar images. Various image retrieval systems have been introduced for minimizing computational costs by filtering irrelevant images using layered approaches. This paper suggests a TfCBIR system that uses the first and second layers for filtering images and the third layer for retrieving them. In this paper, the following major contributions are made:

N. Keisham (✉) · A. Neelima
Department of Computer Science and Technology Nagaland, Dimapur, Nagaland 797103, India
e-mail: nepoleonk@gmail.com

N. Keisham
Department of Computer Science, Tetso College, Sovima, 797115 Dimapur, India

473

1. Two modules make up the proposed TfCBIR system: The first extracts color, texture, and shape information from dataset images. The second module consists of three parts: First, the color feature space of all the images and the image query is analyzed to obtain the index of the most relevant S images to the image query. A second step involves matching S images with query images for shape feature space and retrieving the indexes of the closest $S1$ images to the image query. Finally, the texture feature space of the $S1$ images acquired from the second step is compared with the query image, and the top R most relevant images are obtained as the output.
2. In addition to reducing the number of comparisons, the proposed system also improves the accuracy of the results.
3. COREL 1 K and GHIM 1 K datasets are used for comparison with the existing models.

This article is elucidated as follows. In Sect. 2, a short state of the art is discussed on CBIR. Section 3 discusses the proposed TfCBIR model. The results and implementation details of TfCBIR are furnished in Sect. 4. Lastly, Sect. 5 discusses future directions and conclusions.

2   Related Works

Recently, many effective CBIR systems, including QIBC, FIDS, and others, have been developed. In CBIR, low-level features such as color, texture, and shape are commonly utilized. To make rotation, scaling, translation invariant, features have been extracted from the images and compared [1]. Bag of visual words (BoVWs) are also a popular technique in the field of CBIR. Although, it fails due to the lack of spatial information [2]. CBIR system based on layer of features also became a popular technique. Recently, a bi-layer CBIR system [3] has been developed for retrieving similar images which consist of two layers. Initially, features from both texture and shape are utilized, and in the subsequent stage, features from color and shape are employed. This model adheres to a set number of images to be obtained from the initial stage. And, it needs to be check with various combinations. To overcome this problem, TfCBIR system has been developed in this article which consist threefold to decrease the time taken during retrieval, and the number of images to be retrieved in each layer has been checked with various combination to find the best result.

3   Proposed TfCBIR Method

In this study, a novel threefold content-based image retrieval system (TfCBIR) is presented. The system employs the extraction of color, shape, and texture features to perform the retrieval task. The TfCBIR system comprises of two components, with
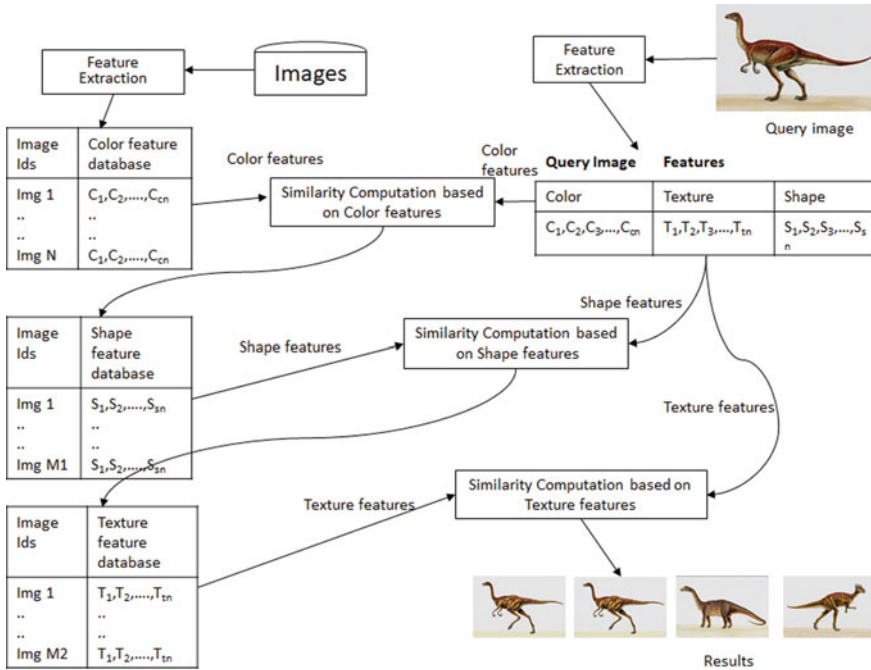
Fig. 1 Schematic representation of the TfCBIR system design

the first component responsible for obtaining the features from the images in the dataset. The second component, which executes the retrieval task, is separated into three stages. The system architecture is outlined in the accompanying Fig. 1.

3.1 Feature Extraction

The images in the dataset are denoted by $I_{\text{DaBa}} = \{I_1, I_2, I_3, \ldots, I_N\}$, which contains N number of images, and the features observed in TfCBIR for each image I are I^c color, I^t texture, and I^s shape along with fixed size feature vector. Color features are extracted using histogram on HSV space and have 81 vector length. A Gabor filter which consists of 5 scales and 6 orientations, 60 vector lengths in total, is used to extract texture features. Using Zernike moments with a vector length of 21, shape features are extracted. After extracting the features from the image dataset, the extracted features are stored into feature database. For each image I, individual feature vectors are generated to represent color IFeat^c, texture IFeat^t, and shape IFeat^s features.

3.1.1 HSV for Color Feature

To extract color features, histograms of quantized color values are employed using the hue (H), saturation (S), and value (V) color spaces. The HSV color space is favored over the RGB color space as it is considered to be more aligned with human perception, as demonstrated in [1]. The RGB image is first transformed into the HSV color space, and then quantized uniformly. A feature vector of 81 bins is constructed by considering 9 values for Hue (H), 3 values for saturation (S), and 3 values for value (V).

$$\text{IFeat}^c = \left\{ \text{IFeat}^c_1, \text{IFeat}^c_2, \text{IFeat}^c_3, \ldots, \text{IFeat}^c_{cn} \right\} \tag{1}$$

3.1.2 Gabor Filter for Texture Feature

In order to obtain the texture features of an image, the Gabor filter is utilized. A complex sinusoidal signal is modulated with a Gaussian function in this filter, as stated in prior studies [4]. By using a combination of 5 scales and 6 orientations, the texture features of an image are obtained, resulting in a feature vector with a length of 60. This feature vector is created by considering the mean and standard deviation of the various combinations of scale and orientation. The texture feature vector is constructed by considering the 2D Gabor function (Gf (x, y)) and its Fourier transform (FT (u, v)). The texture feature descriptor is then formulated as follows:

$$\begin{aligned} \text{IFeat}^t = \big\{ &\text{IFeat}^m_{11}, \text{IFeat}^m_{12}, \ldots, \text{IFeat}^m_{s \times o}, \\ &\text{IFeat}^s d_{11}, \text{IFeat}^s d_{12}, \ldots, \text{IFeat}^s d_{s \times o} \big\} \end{aligned} \tag{2}$$

3.1.3 Zernike Moments for Shape Feature

The extraction of shape features in this work is achieved through the use of Zernike moments (ZM) [5]. These moments utilize Zernike polynomials as a means of forming feature vectors that are invariant to rotation. The Zernike polynomials serve as projections of the function f(x, y) onto a set of polynomials. To represent the images, 21 foundational Zernike moments are employed, resulting in the formation of the shape feature vector.

$$\text{IFeat}^s = \left\{ \text{IFeat}^s_1, \text{IFeat}^s_2, \text{IFeat}^s_3, \ldots, \text{IFeat}^s_{sn} \right\} \tag{3}$$

3.2 Procedure of Retrieving the Images

Every individual type of feature (color, texture, and shape) of the image query Qu_i is depicted by a separate feature vector. Feature vectors are represented in (1) for the query image.

$$
\begin{aligned}
Qu^c &= \left\{ Qu_1^c, \ Qu_2^c, \ Qu_3^c, \ \ldots, \ Qu_{cn}^c \right\} \\
Qu^t &= \left\{ Qu_1^t, \ Qu_2^t, \ Qu_3^t, \ \ldots, \ Qu_{tn}^t \right\} \\
Qu^s &= \left\{ Qu_1^s, \ Qu_2^s, \ Qu_3^s, \ \ldots, \ Qu_{sn}^s \right\}
\end{aligned}
\tag{4}
$$

The threefold approach is used to retrieve images that have similar visual characteristics to Qu_i. In the initial stage of the algorithm, a color feature is employed to assess the likeness between Qu_i and I_{DaBa}. This leads to the elimination of non-pertinent images, resulting in the production of the S most similar images, referred to as I_{DaBaS}. The second stage involves evaluating Qu_i and I_{DaBaS} by utilizing shape features. The final stage matches the texture features of Qu_i and I_{DaBaS1}, producing the indexes of the R highest relevance, which are stored in $I_{DaBaOut}$.

3.3 Analysis of Similarity

It is necessary to use similarity functions to calculate the likeliness between Qu_i and I_{DaBa}. This research computes the similarity of color (C_{ss}) and texture (T_{ss}) as well as shape (S_{ss}) features using Euclidean distance (ED).

$$
\mathrm{ED} = \sum_{i=1}^{|f|} \sqrt{\left(I^f - Qu^f \right)^2}
\tag{5}
$$

4 Results and Discussion

A detailed description of the proposed TfCBIR system as well as the datasets used to test its performance can be found in this section. MATLAB 2017b software is used to develop the proposed system. This work is conducted on a Ryzen 5, 3.0 GHz system with 24 GB of RAM and Windows 11.

4.1 Dataset

Two image datasets are used to evaluate the proposed system. There are 1000 images in COREL dataset containing images of African people, Beaches, Buildings, Buses, Dinosaurs, Elephants, Flowers, Horses, Mountains, and Foods. There are a total of 1000 images in the GHIM-1 k dataset, each containing a hundred images from one of 10 categories (Buildings, Walls, Cars, Fireworks, Bees, Mountains, Flowers, Trees, Fields, and Beaches).

4.2 Parameters Used for Evaluation

The efficiency of the TfCBIR system is measured using precision (Pr), recall (Re), and f-score (Fs) based on retrieving 20 numbers of images. The evaluation metrics used are defined in parameters (6), (7), and (8).

$$Pr = \frac{|I(\text{Retrieved}) \cap I(\text{Relevant})|}{|I(\text{Retrieved})|} \tag{6}$$

$$Re = \frac{|I(\text{Retrieved}) \cap I(\text{Relevant})|}{|I(\text{Relevant})|} \tag{7}$$

$$Fs = \frac{2 \times Pr \times Re}{Pr + Re} \tag{8}$$

The term "I(Retrieved)" refers to the images that have been retrieved, while "I(Relevant)" signifies the relevant images.

4.2.1 Selection and Tuning of Parameters

There are two factors that give impact on the results. First, the order of features on each layer (color (C), shape (S), and texture (T). The possible order of features that can be considered is CST, TCS, TSC, STC, and SCT. Second, the amount of images that is retrieved in each layer. As shown in Table 1, some possible combinations of images retrieved in each layer are given. It should follow the given condition $N \gg S > S1 > R$.

Table 1 Combinations of retrieved images per layer

S	500	400	300	200	500	400	300	200
S1	100	100	100	100	50	50	50	50
R	20	20	20	20	20	20	20	20

Table 2 Comparison of TfCBIR with state of the art

Dataset		Corel			Ghim		
Parameter		Zeng et al. [6]	Bi-layer [3]	Proposed	Parameter	Bi-Layer [3]	Proposed
Africa	Pr	0.725	0.95	0.85	Fireworks	0.95	0.95
	Re	0.145	0.19	0.17		0.19	0.19
Beach	Pr	0.652	0.75	0.95	Building	0.9	0.95
	Re	0.13	0.15	0.19		0.18	0.17
Building	Pr	0.706	0.85	0.95	Wall	0.75	0.85
	Re	0.141	0.17	0.19		0.15	0.17
Bus	Pr	0.892	1	0.95	Cars	0.9	0.9
	Re	0.178	0.2	0.19		0.19	0.19
Dinosaur	Pr	1	1	1	Bees	0.95	0.9
	Re	0.2	0.2	0.2		0.19	0.19
Elephant	Pr	0.705	0.9	1	Mountains	1	0.9
	Re	0.141	0.18	0.2		0.2	0.19
Flower	Pr	0.948	1	1	Flowers	0.9	1
	Re	0.19	0.2	0.2		0.18	0.2
Horse	Pr	0.918	1	1	Trees	0.9	0.9
	Re	0.184	0.2	0.2		0.18	0.18
Mountain	Pr	0.723	0.8	0.9	Fields	0.9	0.95
	Re	0.145	0.16	0.18		0.18	0.19
Food	Pr	0.788	0.95	0.9	Beaches	0.75	0.85
	Re	0.158	0.19	0.18		0.15	0.17
Average	Pr	0.806	0.92	0.94	Average	0.9	0.92
	Re	0.161	0.18	0.19		0.18	0.18
	F	0.269	0.3	0.31		0.3	0.3

Testing all the possible combinations, it has been found that CST with $S = 500$, $S1 = 100$ gives the most accurate results for both datasets. The proposed model has been analyzed with the state of the art and gives better results with less time taken to retrieve the image. This is shown in Table 2.

5   Conclusions

In this study, a new CBIR system, referred to as TfCBIR, is presented. In this system, both queries and datasets are represented by features which include color, texture, and shape. The TfCBIR system involves a three-step retrieval process. Firstly, color

features are compared to identify the most similar images, which are then passed on to the next step. Secondly, shape features are evaluated, and the images with the highest similarity are progressed to the final step. In the third step, texture features are compared, and the images that are most relevant to the query are obtained. The proposed system demonstrates higher accuracy and faster performance based on experimental results. To determine the optimal feature sequence, experiments were conducted, and it was found that the retrieval rate is highest when color features are considered in the first step, shape features in the second, and texture features in the third. The proposed TfCBIR system is a triple-step image retrieval method that can be applied to all types and sizes of images. In future work, the incorporation of a deep neural network-based image feature may further enhance the performance of the system.

References

1. Swain MJ, Ballard DH (1991) Color indexing. Int J Comput Vis 7(1):11–32
2. Yousuf M, Mehmood Z, Habib HA, Mahmood T, Saba T, Rehman A, Rashid M (2018) A novel technique based on visual words fusion analysis of sparse features for effective content-based image retrieval. Math Probl Eng 2018(2134395):1–13
3. Singh S, Batra S (2020) An efficient bi-layer content based image retrieval system. Multimedia Tools Appl 79(25):17731–17759
4. Tou JY, Tay YH, Lau PY (2007) Gabor filters and grey-level co-occurrence matrices in texture classification. In: MMU International symposium on information and communications technologies, pp 197–202
5. Teague MR (1980) Image analysis via the general theory of moments. JOSA (J Opt Soc Am) 70(8):920–930
6. Zeng S, Huang R, Wang H, Kang Z (2016) Image retrieval using spatiograms of colors quantized by Gaussian mixture models. Neurocomputing 171:673–684

A Study of Various Audio Augmentation Methods and Their Impact on Automatic Speech Recognition

Naorem Karline Singh, Yambem Jina Chanu, and Hoomexsun Pangsatabam

1 Introduction

Data augmentation is a popular method in machine learning, where the training data has a significant impact on the model efficiency. It is widely utilized in computer vision and image processing. However, it is not frequently utilized in audio processing. In speech recognition, it is used to alleviate the problem of data scarcity in low-resource languages. This technique increases the number of samples a machine learning algorithm sees during training. It derives new samples from the original data by using transforms that preserve class to cover more problem space as much as possible. It is an effective practice used to improve the model's robustness and accuracy, prevent over-fitting, and reduce mismatch. In one way, it saves resources to collect and label data. But, the performance of a model doesn't always correlate to data augmentation. The type of data augmentation depends on the nature of the problem that we wished to solve. For automatic speech recognition, adding noise data augmentation won't improve accuracy that much since the variability of the training data is more important than noise robustness. Likewise for music instrument classification adding pitch shifting or impulse response is more suitable.

With the emergence of deep neural networks over the past few years, data augmentation has become an essential part of many speech recognition systems. It has been shown to be very helpful for large datasets in end-to-end systems. In [1, 2], for

N. K. Singh (✉) · Y. J. Chanu · H. Pangsatabam
Department of Computer Science and Engineering, National Institute of Technology Manipur, Imphal West, Manipur 795004, India
e-mail: naoremkarline@gmail.com

Y. J. Chanu
e-mail: jina.yambem@gmail.com

H. Pangsatabam
e-mail: humex.pangz@gmail.com

low-resource speech recognition tasks, synthetic data was generated using different distortion methods. In [3], vocal tract length normalization has been modified for data augmentation. For LVSCR tasks, raw audio has been subjected to speed perturbation [4]. In [5, 6], feature drop-outs have been used to train several ASR systems. Pitch adjustment through spectral smoothing was explored in [7]. Recently, SpecAugment [8] was introduced, and it works by masking bands of frequency or time and achieve good improvements in recognition accuracy. Moreover, for end-to-end systems, a new mixup technique called mixSpeech was introduced [9].

In this paper, we examine several currently used techniques, ranging from frequency domain data augmentation to raw data augmentation. We described their benefits and drawbacks in terms of a few key factors such as computational accuracy and feasibility. The structure of this paper is as follows. First, we discuss the different modes of data augmentation based on the configuration and its two types based on the domain of transformation. Next, in Sect. 3, we briefly describe various existing audio data augmentation tools. In Sect. 4, we analyze a few selected popular techniques for speech recognition. In Sect. 5, we present list of problems and future directions for audio augmentation. In Sect. 6, we conclude with a description of the challenges and future scope of improvement.

2   Audio Data Augmentation

According to the configuration and timing of applying data augmentation, there are two modes of data augmentation—online and offline data augmentation.

1. **Offline Data Augmentation**—In addition to pre-calculating the transformation prior to training, it can be used as a pre-processing step. When applied to smaller datasets, augmentation is advantageous, but disc space must be taken into account when doing so for bigger datasets. Computation is done on the CPU, and it requires more disk space. Moreover, it saves computation in the long run.
2. **Online Data Augmentation**—In this case, augmentation occurs in real time. Because the augmented data do not need to be saved to disc, this method is typically used for larger datasets. Here, alterations are carried out in small groups before being provided to the model. With online data augmentation, the model sees additional examples during training, which aids in generalization. Since it uses GPU, it is computationally faster but more expensive in the long run.

Audio data augmentation is discovered to be employed in music processing and other areas of audio processing outside speech recognition. It is mainly used in three subdomains of audio processing, namely music instrument classification, speech recognition, and environmental sound classification. Even though audio data augmentation improves speech recognition accuracy, we have to note that augmented data is not as good as additional data. Additional data is qualitatively better since

it has more variety of data samples. Audio data augmentation can be broadly classified into two groups—raw audio augmentation transformation and spectrogram augmentation transformation.

2.1 Raw Audio Augmentation Transformation

1. **Clipping**—The audio sample is normalized to ensure that a certain number of points are saturated. The samples that are out of range are then clipped.
2. **Compression**—By utilizing an encoder, this function lowers the audio's quality through compression.
3. **Filter**—It filters the incoming audio in a variety of ways. High-pass, high-shelf, low-pass, low-shelf, band-pass, band-stop, and peaking filters are a few examples of common filters.
4. **Harmonic Distortion**—An effect of saturation is produced by repeatedly applying the transformation $\sin(x)$.
5. **Inversion**—The audio can be multiplied by -1 on the y-axis, or it can be inverted along the x-axis.
6. **Noise**—It adds noise to the original samples. Background noise is created by combining the original audio with background noise. In general, the original audio is mixed with noises from street workers, traffic, people on the street, and parks. By combining audio and synthetic noise, such as white noise, synthetic noise generates new samples.
7. **Peak Normalization**—When a digital audio stream is subjected to peak normalization, each sample's volume is increased by the same amount, ensuring that the loudest sample reaches a predetermined level. The maximum audio signal level is set to 0 dBFS. The loudest volume must be set to $[-1, 1]$.
8. **Pitch Shifting**—Each audio signal in the datasets has its pitch adjusted by a factor, while the duration remains constant. A pitch shifter or a harmonizer is used to achieve pitch shifting.
9. **Time Stretching**—It either slows down or accelerates an audio sample by a preset ratio without altering the pitch drastically.
10. **Volume Adjustment**—It is done by varying the loudness of the audio file. It can be achieved in two ways loudness and dynamic range compression. In loudness, the volume of all samples is increased and decreased at a random or fixed rate. And in the case of dynamic range compression, it distorts samples by altering the loudness of the original sample using different noises.
11. **Wow Resampling**—The resampling frequency oscillates with a given frequency and amplitude around the original sampling rate but with the intensity changing over time.

2.2 Spectrogram Augmentation Transformation

1. **Equalized Mixture**—Using randomly chosen timings, equalized mixture data augmentation (EMDA) combines two sounds belonging to the same class [10]. It is calculated using the weighted average of two spectrograms with the same label that were chosen at random. Along with combining sounds, EMDA also tampers with them by enhancing or reducing specific frequency bands.
2. **Spectrogram Sound Mix**—The two random spectrograms from the same class are added together by SSM to produce a new image.
3. **Vocal Tract Length Normalization**—It's an ASR technique that distorts the spectrum in the direction of a medium-level vocal treatment. In ASR, this technique is used to remove the variability that exists between the two vocal tract lengths from each speaker [3].
4. **Spectrogram Channel Shuffle**—It shuffles the channels of the spectrogram. By using this function, bias may be mitigated.
5. **SpecAugment or Spectrogram Random Mask**—It consists of removing portions of spectrograms [8].

   a. *Frequency Mask*—on the spectrogram, add horizontal bars to randomly mask out a range of successive frequencies. It cuts out or masks certain frequencies horizontally. We can replace the areas with zeros or mean values of the energy across the whole spectrogram. The frequency masking is shown in Fig. 1c.
   b. *Time Mask*—comparable to frequency masks, except instead we use vertical bars to arbitrarily block out time intervals from the spectrogram. Essentially, it cuts out little pieces of the spectrogram in the time domain, i.e., to add a blind spot which improves the robustness of the model. It is shown in Fig. 1d.

6. **Spectrogram Time Shift**—It is a change in time that consists of dividing the spectrum into two parts and later restoring them in reverse order.

3 Data Augmentation Tools

1. **MUDA**—Musical data augmentation [11] is a software framework that provides an easy way to regularly apply perturbations to music annotation data in order to construct statistical models. Internally, it uses NumPy to represent the audio, and the annotation is stored in JAMS format (JSON annotated music specification).
2. **Audiogmenter**—The MATLAB package Audiogmenter can be used to enhance audio data [12]. Eight approaches are used to enhance spectrogram data, while five algorithms are used to enhance raw audio data.
3. **Sox**—Sox is regarded as the Swiss Army knife of sound processing. It is a cross-platform command line audio editing tool [13]. It has various effects and features like noise removal, adjustment of speed, pitch, sampling rate, etc. It is found to be used in Kaldi to perform audio augmentation like speed perturbation.
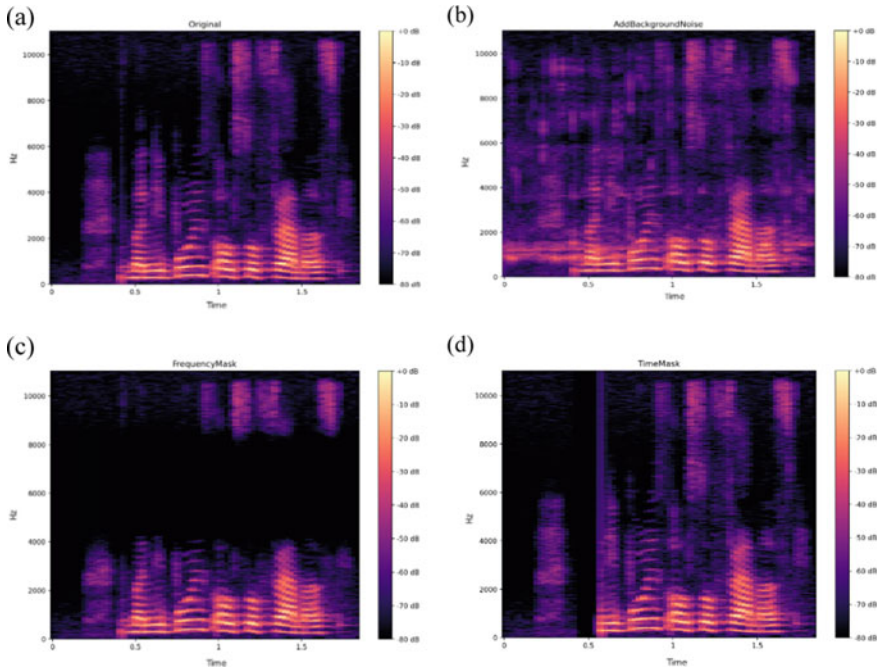
Fig. 1 Transformation of a sample audio signal. Spectrogram images of some of the transformation techniques—**a** original audio, **b** background noise addition, **c** frequency masking **d** time masking

4. **Keras**—Keras is an open-source Python library for implementing neural networks [14]. It acts as an interface for the TensorFlow library. The TensorFlow-io package provides a fast and easy way of converting waveform to spectrogram and spectrogram augmentations such as time masking and frequency masking.
5. **Audiomentations**—Deep learning can benefit greatly from the audio data augmentation capabilities of the Python package called audiomentations [15]. It runs only on the CPU. It can be used along with Keras or PyTorch. Also, it supports mono and multichannel audio. Both waveform and spectrogram transform.
6. **Torch-audiomentations**—It is a quick audio data augmentation tool in PyTorch that was motivated by audiomentations. Currently, it's in the early development stage [16]. It provides most of the waveform transforms available in the time domain. It provides support for both CPU and GPU (CUDA).
7. **TorchAudio**—For working with audio and signal processing, TorchAudio is a machine learning package for data transformation and manipulation [17]. It was released as an extension of PyTorch in their latest updated framework. It provides strong GPU acceleration while training models. It supports various audio types. Moreover, it provides transforms both in time and frequency domains, namely frequency masking, time masking, and time stretching.

8. **Librosa**—Librosa is a Python toolkit for manipulation and analysis of audio signal [18]. It is popularly used in music generation and automatic speech recognition to perform feature extraction and data augmentation using different techniques. It gives an easy way to handle and visualize sounds. Additionally, it offers the components required to construct music information retrieval systems.

4 Audio Augmentation in Speech Recognition

In this section, some of the popular audio augmentation techniques that are currently used in speech recognition are discussed in detail.

4.1 Vocal Tract Length Perturbation (VTLP)

This approach shows one way of augmenting speech datasets by applying a random linear warping transform in a spectrogram [3]. This method's major goal is to reverse the normalizing of speakers to generate variations in the input data. Instead of using a canonical mean, this is accomplished by normalizing inputs to random targets. Vocal tract length normalization is typically employed in speech recognition to eliminate speaker variability brought on by variations in vocal tract length.

In VTLP, a random wrap factor α is generated for each utterance. This wrapping factor converts a frequency f to a new frequency f' as shown in Eq. (1):

$$f' = \begin{cases} f\alpha & f \leq F_{\text{hi}}\frac{\min(\alpha,\,1)}{\alpha} \\ \frac{S}{2} - \frac{\frac{S}{2} - F_{\text{hi}}\min(\alpha,\,1)}{\frac{S}{2} - F_{\text{hi}}\frac{\min(\alpha,\,1)}{\alpha}}\left(\frac{S}{2} - f\right) & \text{otherwise} \end{cases} \tag{1}$$

In VTLP, to not create distorting utterances, a small range is chosen for the wrap factor between 0.9 and 1.1. Values that fell outside of the 0.9–1.1 range were constrained by the boundaries. The filter banks are subjected to the warping process directly. Therefore, new triangular filter banks were generated by switching the central frequencies around $f(i)$ of the $1 \leq i \leq N$ filter banks by applying the formula (1) to new frequencies $f(i)'$. A normal distribution with a center value of 1 and a standard deviation of 0.1 was used to obtain the α values. Acceleration parameters and deltas were calculated during training and added to the data. For decoding, a model averaging method is used. Using deep neural networks with various hidden layer depths, it has been claimed that this technology can reduce the phone error rate by an average of 0.65%. Additionally, convolutional neural networks showed a 1.0% improvement. Both of these experiments were carried out without increasing the number of training epochs.

4.2 Speed Perturbation

A temporal signal is twisted as a result of speed perturbation [4]. An audio signal $x(t)$ can be time-warped by a factor α to produce the signal $x(\alpha t)$. According to the Fourier transform of $x(\alpha t)$, $\alpha^{-1}\hat{x}(\alpha^{-1}\omega)$, that a shift in the frequency components of the $\hat{x}(\omega)$ caused by the warping factor is proportional to frequency ω. The duration of the signal is altered by speed perturbation [1], which also changes the number of frames in the utterance. Moreover, when the speed of the signal is reduced, this leads to small energies of the higher Mel bins. In practice, this doesn't create any issues.

Two additional training sets were created by resampling the original signal. The speed was modified to 90 and 110% of the original rate using the speed function of the Sox program. This technique has low implementation cost since it directly processes the raw signal. Moreover, this technique experimented on 4 different LVCSR tasks, and it shows an average relative improvement of 4.3% word error rate (WER).

4.3 Spectral Smoothing

Spectral smoothing is used to resolve the problem of pitch-induced distortions [7]. A dataset that contains high-pitched child speakers leads to poor performance in WER. In this method, pitch variations in speech are reduced by smoothing the spectrum. Initially, modified empirical mode decomposition (MEMD) is used to divide the short-time magnitude spectrum into a number of components. Then, the spectrum is reconstructed using the higher-order components by discarding the lowest-order components. Next, the generated spectrum is used in the standard MFCC extraction.

Due to pitch-harmonics, the high-pitched signal spectrum exhibits ripples. In the case of ASR, the ripples need to be removed since only the spectral envelope is important. Therefore, using MEMD the magnitude spectra are first to split into their corresponding 8 intrinsic mode functions. The short-time magnitude spectrum is reconstructed by completely eliminating the first few IMFs in order to reduce the pitch-harmonics. Since Mel-scale warping and low-time liftering will perform additional smoothing, dropping two modes could result in over-smoothing. Through the use of the resulting smoothed spectrum, the cepstral coefficients are calculated. Error rates are further decreased by combining the spectral smoothing method with linear frequency warping through VTLN.

4.4 Additive Noise

This approach shows a way of increasing the robustness of an ASR system to additive noise [6]. It eliminates less trustworthy voice signal spectrum components. It is predicated on the idea that additive noise can affect spectral peaks. Additionally, it was

noted that dropout and data augmentation are two related strategies. Instead of random dropout, two simple heuristics are utilized. The first one retains the spectro-temporal maximum, and the second is based on a crude estimation of spectral dominance. This method was implemented on convolutional neural network-based speech recognizers. The input is a standard mel-spectral time–frequency representation called filter banks.

A perceptually driven heuristic controls the choice of the dropped pixels. This heuristic assigns a vulnerability score to each element of the input block itself. The first strategy retains the spectral peaks (high amplitude) since it carries information required for speech intelligibility. It was reported that preserving just 10% of the components gives better results than preserving 20% of the components. The second strategy is based on spectral masking. By only maintaining the component with the highest amplitude in each column of the 9×9 Spectro-temporal windows, it imitated the spectral dominance effect. This method allowed for the accurate retention of 9 spectral components at each time instance in the 9 windows that were employed to span the whole frequency range.

4.5 SpecAugment

This approach was developed based on the assumption that speech features should be resistant to time domain and frequency domain deformation or partial loss in small segments of speech. SpecAugment [8] relies on frequency masking, time warping, and time masking of spectrogram audio augmentation methods. It directly applies the transformation to the input features (filter bank coefficients) of a neural network. The transformation procedure is as follows

- **Time warp**—We regard a given log Mel spectrogram with τ time steps to be an image. Within the time steps $(W, \tau - W)$, a random point along the horizontal line passing through the center of the image is to be warped to the left or right by a distance w picked from a uniform distribution from $[0, W]$, where W is the time warp parameter along that line.
- **Frequency masking** is used to mask f consecutive Mel frequency channels $[f_0, f_0 + f)$, where f is selected initially from a uniform distribution from $[0, F]$, and f_0 is selected from $[0, v - f)$. The amount of mel frequency channels is v.
- **Time masking** is utilized to mask t successive time steps $[t_0, t_0 + t)$. T is the time mask parameter, and t_0 is taken from $[0, \tau - t)$, with t being picked from a uniform distribution starting at $[0, T]$. Due to an upper constraint on the time mask, its breadth cannot exceed p times the number of time steps.

For end-to-end voice recognition tasks, this technique was used with the listen, attend, and spell networks. It was tested on the Librespeech 960h tasks and achieved an improvement of 6.8% WER without using the language model and 5.8% WER with a shallow fusion language model. Previously, the benchmark was 7.5% WER.

4.6 MixSpeech

MixSpeech is a new data augmentation technique for end-to-end ASR models [9]. It is based on the mixup technique, which introduces a contrasting signal to force the ASR model to predict the correct text accurately. Mel spectrograms or MFCCs are used to train an ASR model by combining the weights of two separate speech sequences. To determine the recognition loss, each label is used. The same weight as in the input speech is used to aggregate the two losses. It follows the following formulae

$$
\begin{aligned}
X_{\text{mix}} &= \lambda X_i + (1 - \lambda) X_j, \\
\mathcal{L}_i &= \mathcal{L}(X_{\text{mix}}, Y_i) \\
\mathcal{L}_j &= \mathcal{L}(X_{\text{mix}}, Y_j), \\
\mathcal{L}_{\text{mix}} &= \lambda \mathcal{L}_i + (1 - \lambda) \mathcal{L}_j,
\end{aligned}
\tag{2}
$$

where X_i is the input speech sequence, Y_i is the target text sequence, and X_{mix} is the mixup speech sequence created by adding the two speech sequences frame-by-frame with weight λ. Calculated using $\mathcal{L}(., .)$, the ASR loss (which also accounts for the recognition process) is then combined with $\mathcal{L}_{\text{mix}}$ using the same weights λ as the voice input used in the training phase. Following the initial blunder, we opt for $\lambda \sim \text{Beta}(\alpha, \alpha)$ where $\alpha \in (0, \infty)$ following the original mixup.

This method was applied to two end-to-end models listen, attend, and spell, and transformers models using three datasets TIMIT, WSJ, and HKUST. And it was compared with the data augmentation method specAugment. And it outperforms specAugment with an improvement of 10.6% PER on the TIMIT dataset. Moreover, it achieves a strong WER of 4.7% and 22.9% on WSJ and HKUST datasets, respectively.

5 Issues and Future Directions

List of issues and challenges:

- There is no concrete mechanism to assess the quality of the augmented data.
- Requires tuning of different parameters and components for different types of augmented audio.
- Errors and biases are propagated to the newly generated synthetic data if the original data contains the same.
- Advanced applications require new methods and strategies for augmentation.

List of future directions for improvement:

- Concatenate several speech segments with their appropriate text sequences to create a new utterance.
- Generate code-switched speech data from monolingual dataset.
- Develop an optimal data augmentation strategy for any tasks.
- Explore new techniques, e.g., generative adversarial networks.

6 Conclusion

This paper presents the different types of audio augmentation techniques and some of the tools to perform audio augmentation. Moreover, six popular audio augmentation techniques of speech recognition are thoroughly discussed and can conclude that adding more data does not always improve WER performance. We found out that no specific augmentation method is superior to other methods. The transformation function requires specific parameters and timing, to make the transformation effective. In general, data augmentation converts over-fit problems into under-fit problems and also improves the robustness and accuracy of the systems.

Future study can be focused on increasing the variety of data augmentation approaches and examining whether WER performance can be enhanced using various data augmentation strategies as described in this paper. Other newly introduced audio augmentation techniques could be explored for speech recognition.

References

1. Kanda N, Takeda R, Obuchi Y (2013) Elastic spectral distortion for low resource speech recognition with deep neural networks. In: 2013 IEEE Workshop on automatic speech recognition and understanding. IEEE, pp 309–314
2. Ragni A, Knill KM, Rath SP, Gales MJF (2014) Data augmentation for low resource languages. In: INTERSPEECH 2014: 15th Annual conference of the international speech communication association, Singapore. International Speech Communication Association (ISCA), pp 810–814
3. Jaitly N, Hinton GE (2013) Vocal tract length perturbation (VTLP) improves speech recognition. In: Proceedings of the 30th international conference on machine learning. JMLR:W&CP, vol 28, pp 1–5
4. Ko T, Peddinti V, Povey D, Khudanpur S (2015) Audio augmentation for speech recognition. In: INTERSPEECH 2015, pp 3586–3589
5. Mallidi SH, Hermansky H (2016) Novel neural network based fusion for multistream ASR. In: 2016 IEEE International conference on acoustics, speech and signal processing (ICASSP). IEEE, pp 5680–5684
6. Tóth L, Kovács G, Van Compernolle D (2018) A perceptually inspired data augmentation method for noise robust CNN acoustic models. In: Speech and computer. SPECOM 2018. Lecture notes in computer science, vol 11096. Springer, Cham, pp 697–706
7. Sai BT, Yadav IC, Shahnawazuddin S, Pradhan G (2018) Enhancing pitch robustness of speech recognition system through spectral smoothing. In: 2018 International conference on signal processing and communications (SPCOM). IEEE, pp 242–246

8. Park DS, Chan W, Zhang Y, Chiu C-C, Zoph B, Cubuk ED, Le QV (2019) SpecAugment: a simple data augmentation method for automatic speech recognition. arXiv:1904.08779 [eess.AS], arXiv:1904.08779v3 [eess.AS], https://doi.org/10.48550/arXiv.1904.08779

9. Meng L, Xu J, Tan X, Wang J, Qin T, Xu B (2021) MixSpeech: data augmentation for low-resource automatic speech recognition. In: ICASSP 2021—2021 IEEE International conference on acoustics, speech and signal processing (ICASSP). IEEE, pp 7008–7012. https://doi.org/10.1109/ICASSP39728.2021.9414483

10. Takahashi N, Gygli M, Pfister B, Van Gool L (2016) Deep convolutional neural networks and data augmentation for acoustic event detection. arXiv:1604.07160 [cs.SD], arXiv:1604.07160v2 [cs.SD], https://doi.org/10.48550/arXiv.1604.07160

11. McFee B, Humphrey EJ, Bello JP (2015) A software framework for musical data augmentation. In: 16th International society for music information retrieval conference (ISMIR), pp 248–254

12. Maguolo G, Paci M, Nanni L, Bonan L (2021) Audiogmenter: a MATLAB toolbox for audio data augmentation. Appl Comput Inform (ahead-of-print). https://doi.org/10.1108/ACI-03-2021-0064

13. SoX—Sound eXchange [Computer software]. Retrieved from https://sox.sourceforge.net/

14. Chollet F et al (2015) Keras. In: GitHub. Retrieved from https://github.com/fchollet/keras

15. Jordal I Audiomentations. Retrieved from https://github.com/iver56/audiomentations. Accessed on 13 Jun 2022

16. Jordal I Torch-audiomentations. Retrieved from https://github.com/asteroid-team/torch-audiomentations

17. Yang Y-Y, Hira M, Ni Z, Astafurov A, Chen C, Puhrsch C, Pollack D, Genzel D, Greenberg D, Yang EZ, Lian J, Hwang J, Chen J, Goldsborough P, Narenthiran S, Watanabe S, Chintala S, Quenneville-Bélair V (2022) Torchaudio: building blocks for audio and speech processing. In: ICASSP 2022—2022 IEEE International conference on acoustics, speech and signal processing (ICASSP). IEEE, pp 6982–6986

18. McFee B, Raffel C, Liang D, Ellis DPW, McVicar M, Battenberg E, Nieto O (2015) librosa: audio and music signal analysis in python. In: Proceedings of the 14th python in science conference (SCIPY 2015), pp 18–25

A Hybrid Federated Reinforcement Learning Approach for Networked Robots

Gayathri Rangu, Divya D. Kulkarni, Jayprakash S. Nair, and Shivashankar B. Nair

1 Introduction

Smartphones, tablets, and other web-enabled devices have emerged as indispensable tools for communication, information, and entertainment worldwide. Each one of these numerous devices produces data every second. This data is used to analyze and learn user patterns, enhancing user experience by sending appropriate recommendations. Conventional edge-cloud Machine Learning (ML) techniques require this humongous amount of data to be uploaded onto a server and are used to train the model centrally. This method suffers from privacy issues since user-sensitive data lands up at the server. McMahan et al. [1] have proposed a way out by using what they call federated learning (FL). In this method, the models are generated locally on each of the devices using the data within. These local models are then sent to the server where they are aggregated and returned to all the connected devices, thereby preserving privacy [2]. This centralized approach, however, suffers from various drawbacks such as single-node dependency, maintenance costs, and scalability issues. To overcome these, decentralized approaches of FL have been proposed [3] wherein all the devices are connected over a network and there is no global server that aggregates the models. In these approaches, the learned model weights are exchanged with all

G. Rangu (✉) · D. D. Kulkarni · S. B. Nair
Indian Institute of Technology Guwahati, Guwahati, India
e-mail: gayathri.rangu@iitg.ac.in

D. D. Kulkarni
e-mail: divyadk@iitg.ac.in

S. B. Nair
e-mail: sbnair@iitg.ac.in

J. S. Nair
Federal Institute of Science and Technology, Angamaly, India
e-mail: jsnair.hi@gmail.com

the other devices. This approach can circumvent the issues of centralized FL but at the cost of increased learning times.

Of late, FL has also been used in networked robotics [4]. Robots may inhabit the same or different environments. Every robot generates its dataset. In such cases, using FL can train models specific to a robot's environment. Time-critical devices such as robots, drones, and other IoT devices usually do not have a predefined dataset to be trained on. In tasks like exploration, obstacle avoidance, etc., where robots need to evolve on their own, most often, Reinforcement Learning (RL) [5] is used. RL focuses on how agents interact with their environment and learn to evolve in a trial-and-error manner. A policy is learned based on which the agent takes action and in turn, gets rewarded or penalized for the action taken. Overall, the agent focuses on maximizing the cumulative rewards in the long run. RL can also be implemented in a federated manner [6].

Federated Reinforcement Learning (FRL) can be implemented either as a client-server model or a peer-to-peer model. The former may not be always suitable in the case of robotics because when the server goes down, the whole network could fail. Although a peer-to-peer approach can handle this drawback, it is prone to delays in model updates. Since both centralized and decentralized approaches of FL have their own drawbacks, a combination of both these approaches could be explored. This paper presents such a Hybrid FRL (*HyFRL*) wherein the learned models are aggregated and shared among a set of connected robots inhabiting different environments. The approach uses mobile agents that move from robot to robot to aggregate and share the models they carry on the go. Unlike the decentralized approaches, [7] aggregation is performed, not just by a mobile agent but also by a server, thus making it partly centralized and partly decentralized.

The following are the significant contributions of the work presented in this paper:

i. A Hybrid Federated Reinforcement Learning (*HyFRL*) model which is a combination of both centralized FRL and decentralized FRL.
ii. A comparison of the performances of centralized, decentralized, and Hybrid FRL approaches.
iii. Use of multiple agents to hasten convergence.

2 Related Work

Zhou et al. [8] have proposed a real-time data processing architecture for multi-robots based on FL for recognition tasks using MNIST and CIFAR10 datasets for training. Agrawal et al. [3] described a decentralized FL approach using mobile agents using Deep Learning. In robotic applications where such datasets are not available *a priori*, the use of RL [9] can be a better option. In standard RL models such as Q-learning, an agent learns to take actions by continuously perceiving and interacting with its environment.

FRL is a relatively new research direction that has come up to address the problems in FL and RL. A method to decentralize FRL in a multi-robot scenario has been proposed in [7]. Though the need for a central server was eliminated, this model suffers from delayed updates. On a larger scale, when the number of nodes increases, a single mobile agent going around and sending updates results in slowing the entire system. This paper thus proposes a Hybrid of centralized and decentralized FRL to mitigate the issue.

3 Methodology

An *agent* is a piece of software residing in a device that can either be static or mobile. A *mobile agent* can carry information, program, or data as a payload from one device to another. Mobile agents, thus, can knit through connected devices, sharing information thereby keeping the models in these up to date. Agent platforms need to run on individual devices to facilitate the execution of agents. This work uses a Python-based variant of an open-source mobile agent platform, viz. *Tartarus* [10]. This platform allows features like agent programming, execution, mobility, etc. Mobile agents have been used to aggregate and share locally learned models in the network to realize Hybrid FRL (*HyFRL*), which uses both centralized and decentralized approaches.

A mobile agent, A_i, traverses a set of networked N robots, R. Each robot, $R_i \in R$, uses RL to churn out its local model. On visiting a robot, the mobile agent, A_i, aggregates the model which it carries as payload with the one locally available within R_i, to generate a new model, M_i. It then proceeds to the next connected robot, R_{i+1}, and performs the same aggregation to produce, M_{i+1}. By doing so, A_i, it tends to aggregate and share the models locally learned by the robots in the set, R, achieving decentralized FRL as in [7]. In the proposed *HyFRL*, after every τ hops of such aggregation in the network, A_i sends the currently aggregated model it carries, to a central server which in turn aggregates the received model with the one it already has, and updates all the connected robots with this global aggregate. This *aggregation* and subsequent *updation* constitutes a *round* and augments the decentralized FRL with a centralized flavor.

4 Algorithm

Algorithm 1 depicts the three sub-parts of the proposed *HyFRL* algorithm, viz. *(a)* Robot side, *(b)* Mobile Agent side, and *(c)* Server side. The algorithm takes in the number of robots, N, the number of hops after which the updates are sent to the server, τ, and the Q-learning algorithm, and outputs the aggregated Q-table, Q_{agg}. The algorithm is made to run till the learning is saturated. On the robot side, the robots learn to avoid obstacles using Q-learning and update their respective Q-tables

based on Eq. (1):

$$Q_n(s, a) = Q(s, a) + \alpha * (R_n + (\gamma * Q_{\max}(s', a')) - Q(s, a)) \qquad (1)$$

where $Q_n(s, a)$ and $Q(s, a)$ are the new and old Q-values of the state, s. $Q_{max}(s, a)$ is the maximum of the Q-values considering the current state of the agent (after performing the action a), R_n is the reward that the agent received by performing a in s, α is the learning rate, and γ is the discount factor. The agent goes through the episodes several times, each time adjusting its policy based on its newly acquired knowledge (*state*, *action*, *new state*, *reward*). As a human would learn how to play a game to get a good score, the agent eventually learns the policy that produces a good reward over the episode.

The robots also replace their Q-tables with the global updates, Q_g, as and when they receive them from the server. On the Mobile Agent side, after the creation of the agent within R_0, it carries Q_{agg}, as a payload to the next robot, R_{j+1} where it again performs a similar aggregation of Q_{agg} with the locally available R_{j+1}^Q. It carries this new aggregate to another robot and repeats this procedure. After every τ such migrations or hops, it sends the Q_{agg} to the server. At the server side, Q_{agg}s received from the mobile agent are aggregated and sent back to all the robots. It may be observed that while the server aids in centralized FL, the mobile agent facilitates a decentralized version. The three sub-parts thus, constitute the *HyFRL* approach.

Algorithm 1: The proposed *HyFRL* algorithm for a single agent.

Input: N, τ, Q-learning algorithm

Output: Learned Q-table, Q_g

(a) Robot Side:	(b) Mobile Agent Side:	(c) Server Side:
1 for $i = 0$ to $(N - 1)$:	1 $Q_{agg} = R_0^Q$	1 initialize Q_g
2 Initialize R_i^Q	2 while(True):	2 while(True):
3 while(True):	3 for $i = 0$ to $(N - 1)$:	3 receive(Q_{agg})
4 for $i = 0$ to $(N - 1)$:	4 for $j = 1$ to τ:	4 $Q_g = aggr(Q_g, Q_{agg})$
5 avoidObstacle()	5 Payload(Agent, Q_{agg})	5 for $i = 0$ to $(N - 1)$:
6 update R_i^Q	6 R_{j+1} ←move(Agent)	6 R_i ← send(Q_g)
7 if $received(Q_g)$:	7 Q_{agg}=aggr(R_{j+1}^Q, Q_{agg})	
8 $R_i^Q = Q_g$	8 server← send(Q_{agg})	

5   Experiments and Results

We have used *Webots*® [11] and *Tartarus* platforms on computers connected via a LAN to create a network of robots that can facilitate agent mobility. Each *Webots*® platform had a robot performing obstacle avoidance using Q-learning which is a form

State/Action	A0	A1	A2	A3
S0	-6.61	-0.83	-0.16	10.51
S1	-6.03	-0.29	0.55	11.05
S2	-5.18	0.32	1.19	12.07
S3	-4.61	0.95	1.82	12.83
S4	-4.12	1.65	2.31	12.56
S5	-4.47	2.17	2.81	14.23
S6	-2.49	2.43	3.65	15.33
S7	-1.81	3.46	4.55	15.91
S8	-1.32	2.95	4.33	17.01
S9	-7.53	-1.58	-0.57	9.84
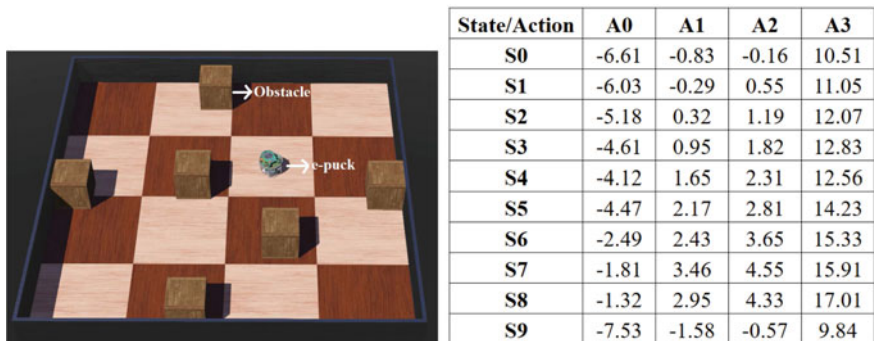
(a) A $Webots^{®}$ arena with an e-puck robot and blocks constituting the obstacles (b) Structure of a Q-table with 10 states (rows) and 4 actions (columns)

Fig. 1 A Sample arena and a Q-table

of RL. The environments of all robots were kept different. In the *HyFRL* implementation, a mobile agent within one of these robots was made to carry the locally generated Q-table (model) as its payload and migrate to the neighboring robot where it aggregated this Q-table with the local one available in this robot and shared the same with it. After visiting τ robots, it was made to send the aggregated Q-table it carries, to the central server hosted on a separate computer. The server aggregated this with the model it already had and sent the aggregated global Q-table to all the robots, thereby constituting a *round*. While the movement and local aggregations facilitated decentralized FRL, the aggregation at the server constituted its centralized counterpart. After several such migrations, the aggregated Q-table formed the optimal policy. The agent was made to continue to do this until the performance saturates.

We have compared the performance of proposed *HyFRL* with those of the fully centralized and fully decentralized versions of FRL, using five connected robots (*Webots®*), each with a single robot. These robots were situated in distinct environments, E_0, E_1, E_2, E_3, and E_4, where the number of obstacles and their respective positions were different. Each robot needed to learn a model suited to its environment. A robot along with its environment is shown in Fig. 1a, and the corresponding Q-table structure is shown in Fig. 1b. Our experiments are conducted by taking 10 states ($S_0,...S_9$) and 4 actions (A_0, ..., A_4). These four actions are for the robot to move Forward, Backward, Stop, and Left.

We have investigated the performance of two aggregating techniques—*Maximum*, where the maximum of the entries within the Q-tables was taken for aggregation and *Average*, where the mean of the entries formed the aggregated Q-table entries. The 5-robot network was used to compare the performances of the centralized, decentralized, and the proposed *HyFRL* versions. The numbers of agents and of hops were taken to be 1 and 2, respectively. In order to find the effect of using multiple mobile agents to aid in the aggregation process, we conducted experiments on a 10-robot

(a) Performance of C_AVG, C_MAX

(b) Performance of D_AVG, D_MAX

(c) Performance of H_AVG, H_MAX

(d) Performance of $HyFRL\_MAX$ using 10 Robots and Multiple Agents
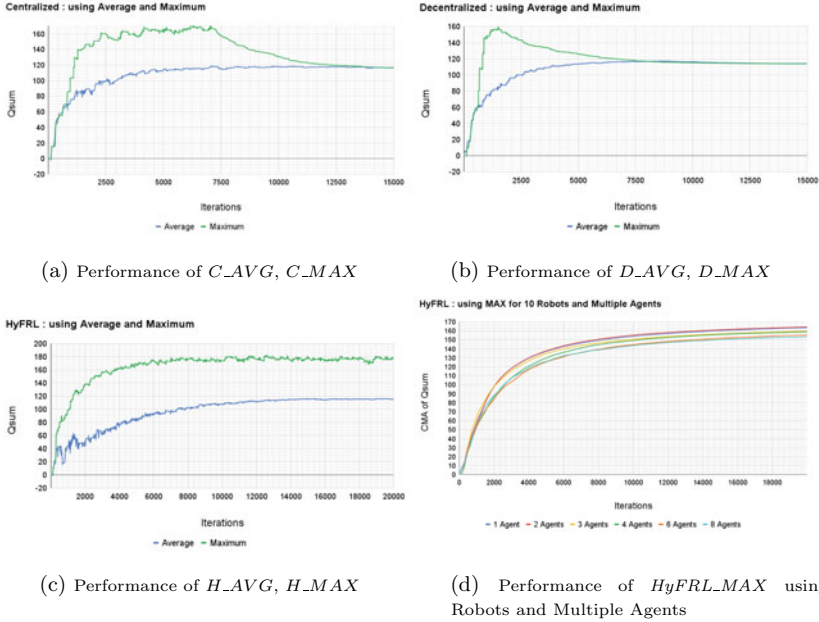
Fig. 2 Performances of centralized, decentralized, and *HyFRL* approaches

network using the *HyFRL* approach. Each agent was made to start from a different node so as to hasten the process of learning. The performance of an approach was measured based on the *Qsum* values obtained at every iteration. These values were calculated using Eq. (2).

$$Qsum = \frac{\sum_{n=0}^{N-1} \sum_{s=0}^{S-1} \sum_{a=0}^{A-1} Q_{s,a}^n}{N} \qquad (2)$$

As can be observed, *Qsum* is the average of the sum of all the values, $Q_{s,a}^n$, within the Q-tables, of all the robots where *N*, *S*, and *A* stand for the number of robots, states, and actions, respectively. Each experiment was performed 5 times with the robots positioned at different starting locations, and the average of the *Qsum* values obtained from these experiments was used to plot the relevant graphs.

For clarity, the curves in the graphs are labeled as *A_M* where *A* could be either *C*, *D*, or *H* signifying Centralized, Decentralized, or Hybrid FRL, respectively. *M* could be either *AVG* or *MAX* indicating the aggregation methods—*Average* or *Maximum*. Figures 2a–c depict the performances of the Centralized, Decentralized, and *HyFRL* in a 5-robot network using both *Average* and *Maximum* for aggregation. Figure 2a shows the performance of the centralized FRL version using the aggregation methods. As can be seen in Fig. 2a, though the C_MAX performs better in between, both C_AVG and C_MAX eventually saturate to the same range of 115–118. The sudden

initial surge in the performance in the case of C_MAX is mainly due to the selection of the maximum values from across the Q-tables. The maximum of one robot may not always be suitable to the others, consequently increasing penalties and lowering the curve to join C_AVG. Another reason for this could be the frequent updates sent to the server by each robot which gives lesser time for the robot(s) to adapt to the updates received. In the decentralized FRL, Fig. 2b, both D_AVG and D_MAX saturate to *Qsum* values in the range of 113–114 over time, indicating a slightly lower performance than their centralized counterparts. This could be due to the delayed updates received by all the robots. In the case of the *HyFRL* version in Fig. 2c, there is a noticeable difference in the performance of both H_AVG and H_MAX. Unlike the centralized and decentralized approaches, both curves do not converge over time. H_AVG saturates to around 115–117, whereas H_MAX saturates to 171–182. This is observed because the frequency of updates is less compared to the centralized version which allows the robots to adapt to the updates received. *HyFRL* using *maximum* as the aggregation method outperforms all the other models. In the subsequent experiment, we, therefore, investigated the effect of increasing the number of mobile agents in the case of *HyFRL* using the *maximum* aggregation method.

Using a network of ten robots, six separate experiments were performed using 1,2,3,4,6, and 8 agents, respectively. Figure 2d shows the comparison of the performances of each experiment wherein the average of all the *Qsum* values within the Q-tables of all the robots per iteration, in an experiment, was taken into consideration. Thus, each curve denotes the cumulative performance of all the robots in an experiment. In each experiment, the average of all the *Qsum* values within the Q-tables of all the robots per iteration was logged. Figure 2d shows the Cumulative Moving Average (CMA) of these values for each of the six experiments. It can be observed that when multiple agents are introduced into the network, the overall performances of the robots vary. For lower ratios of the number of agents μ to the number of robots N, (viz. μ equal to 1,2), the performances seem to increase due to a proportionate increase in updates. However, for larger μ:N ratios, (viz. μ equal to 3,4,6,8), the performances deteriorate because of a drastic increase in the frequency of updates. This gives lesser time for the robots to learn and modify the Q-tables, thereby resulting in redundant models being aggregated. It may thus be inferred that this ratio needs to be determined empirically to increase the performance of the system.

6   Conclusions and Future Work

In this paper, we have proposed a combination of centralized and decentralized FRL, *HyFRL*, to eliminate the problems of single-node dependency and delayed updates. Experiments conducted reveal that the *HyFRL* approach using *maximum* as the aggregation method outperforms centralized and decentralized methods irrespective of the aggregation methods used in the latter two. Learning can be further accelerated by increasing the number of mobile agents in the network. However, this needs to be

done with caution. An appropriate μ:N ratio may need to be found to enhance the performance. In the future, we plan to work on experimenting *HyFRL* using more robots to converge on a prescription for μ:N ratio. In addition, we also intend to investigate the effect of agent hops, τ, on the performance of the system. Heterogeneous learning mechanisms, where robots learn using different learning algorithms (such as State-Action-Reward-State-Action, (SARSA) and Deep Learning in addition to Q-learning), in conjunction with other aggregation methods, in a federated manner, will also aid in hastening convergence and increasing the performance.

References

1. McMahan B, Moore E, Ramage D, Hampson S, Arcas BA (2017) Communication-efficient learning of deep networks from decentralized data. Artific Intell Stat 1273–1282. PMLR
2. AbdulRahman Sawsan, Tout Hanine, Ould-Slimane Hakima, Mourad Azzam, Talhi Chamseddine, Guizani Mohsen (2020) A survey on federated learning: the journey from centralized to distributed on-site learning and beyond. IEEE Internet of Things J 8(7):5476–5497
3. Agrawal A, Kulkarni DD, Nair SB (2020) On Decentralizing Federated Learning. In: 2020 IEEE international conference on systems, man, and cybernetics (SMC)
4. Xianjia Y, Queralta JP, Heikkonen J, Westerlund T (2021) Federated learning in robotic and autonomous systems. Proc Comput Sci 191:135–142
5. Kober J, Bagnell JA, Peters J (2013) Reinforcement learning in robotics: a survey. Int J Robot Res 32(11):1238–1274
6. Liu Boyi, Wang Lujia, Liu Ming (2019) Lifelong federated reinforcement learning: a learning architecture for navigation in cloud robotic systems. IEEE Robot Autom Lett 4(4):4555–4562
7. Nair JS, Kulkarni DD, Joshi A, Suresh S (2022) On decentralizing federated reinforcement learning in multi-robot scenarios. In: 2022 7th South–East Europe design automation, computer engineering, computer networks and social media conference (SEEDA-CECNSM), 1–8. IEEE
8. Zhou W, Li Y, Chen S, Ding B (2018) Real-time data processing architecture for multi-robots based on differential federated learning. In: 2018 IEEE smartWorld, ubiquitous intelligence and computing, advanced and trusted computing, scalable computing and communications, cloud and big data computing, internet of people and smart city innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI), 462–471
9. Sutton RS, Barto AG (2018) Reinforcement learning: an introduction. MIT press
10. Semwal T, Bode M, Singh V, Jha SS, Nair SB (2015) Tartarus: a multi-agent platform for integrating cyber-physical systems and robots. In: Proceedings of the 2015 conference on advances in robotics, pp 1–6
11. Michel O (2004) Cyberbotics Ltd. Webots™: professional mobile robot simulation. Int J Adv Robotic Syst 1(1):5

Proposing ML Approach for Detection of Lung Cancer

Vaibhav Kant Singh and **Nageshwar Dev Yadav**

1 Introduction

Lung cancer an asymptomatic disease at early stages is a major fatal cancer type causing thousands of deaths every year. Detection of this at an early stage is very prominent from the point of saving life. Also one more important point to ponder is the increasing medication price. Cancer is a type of problem in which the cells present in the body grow in an uncontrollable way and spread from one place to another place. This may originate at any body organ; we know that human body organs are made up of a number of cells of different type which depend on the different organ. The uncontrolled cell growth gives rise to tumours which are cancerous or sometimes non-cancerous. A process called metastasis is occurring during the cancerous tumours creation which is often called malignant. Another type which is not fatal is called benign. Cancer is generally a result of a genetic change. Genetic changes happen with an error probably in division of cell and so on. These are several cases of tissue change that may terminate as a cancer tissue which are hyperplasia, carcinoma in situ, and dysplasia. Generally speaking, around 100 types of cancer exist. Some major types include prostate, breast, kidney, testicular, thyroid, leukaemia, bladder melanoma, sarcoma, pancreatic, nasopharyngeal, lymphoma, carcinoma, and lung cancer. In the current work, we are focussed on lung cancer. When the infected organ is lung in terms of cancer infection, we term the type as lung cancer. Lung cancer starts at lung and can infect brain and lymph nodes. Lung cancer is of small cell and non-small cell

V. K. Singh (✉)

Department of Computer Science and Engineering, School of Studies of Engineering and Technology, Guru Ghasidas Vishwavidyalaya, Central University Bilaspur, Chhattisgarh 495001, India

e-mail: vibhu200427@gmail.com

N. D. Yadav

Informatica Solution, Banglore, India

of the two types; the latter is more commonly observed. Smoking is the most general feature observed in the person having this problem, although sometimes this is found in other people also. Sub-types of non-small are adenocarcinoma, squamous cell, and large-cell carcinoma. The major factors causing lung cancer are smoking, radon, second hand smoking, history of parents, emission of substances such as asbestos. The major symptoms are headaches, hoarseness, cough that is not stopping, blood in cough, bone pain, chest pain, and weight loss and breathe shortness. There are stages, namely 0, 1, 2, 3, 3A, 3B, and 4, of non-small type and limited and extensive of small type. The survival rates are very poor for both the cases. The major treatments available are surgery, chemotherapy, radiation therapy, and targeted therapy. In the current work looking to the positives that computer industry gives to the people we have developed a system that makes usage of both the hardware and software privileges available for the mankind.

The paper is a result of various types of incidents that we face in our day-to-day life. Recently in the city Bilaspur, one case was encountered where the person observed his death because of this disease that is lung cancer. The disease has taken a lot of lives and is in the urge of taking many more. It is our duty that is the duty of the people working in the field of science and technology to take care of problems like this and create an environment where people feel safe and secure of the disease like this. Although we and many more people who are working in the field of science and technology are working hard tackle this problem. The current paper is just a small beginning towards the area to help people to look after the problem in a sophisticated way that is described in the current paper.

2 Problem Statement

One author of the paper lost his parent because of cancer and the reason was the non-diagnosis of the disease at a correct time. A number of cases not only of cancer but of different types of diseases are observed where a proper diagnosis or treatment at an appropriate time would have saved a life. Quoting from our religious books the life of human is a gift that a soul gets after getting through a number of species. Humans are the best known species on earth. The cost of life could be understood by any of the person who enjoys it. We are having sense organs which make us to enjoy experience of various types from the environment we are in. Life is simply precious. In the current work, we are majorly focussed on a disease called lung cancer which is causing death of a number of people around the globe. The idea in this work is to design a computer-oriented system that can save life from lung cancer.

3 Literature Survey

In [1–13], the authors made implementation of ML for diseases like thyroid, diabetes, lung cancer, and breast cancer. The papers addressed are having a note on diseases like thyroid, diabetes, lung cancer, and breast cancer. The catastrophic results obtained in the aforesaid conditions are elaborated in the above papers. The papers cited above are the fundamental work for the current paper. In [14–22], the authors propose model to solve non-linearly separable problems. Artificial neural network is capable to solve a variety problems which we may in fundamental classify between LSP and NLSP. The papers above are engaged in the aforementioned issues faced in ANN problem domain. In [23], TRIVENI is being proposed. TRIVENI is a computer oriented model that is going to provide the researchers working in the field of electronics and engineering, electrical and engineering, and computer science people to think of the test bed that is going to solve the problem of the mankind in the best possible way. In [24–28], the authors made specification for ANN, Lagrange, and ML models for solving complex problems. The above cited works are collection of patents that are having the modelling that could be used to cover the ANN landscape to solve the more general problems. In [29], space model is proposed. The model is a great model to go for the search. In today's time, we are looking for efficiency and the paper speaks of that only. In [30–35], data mining techniques are discussed. We human beings especially top notch is going to be looking for the approaches discussed in the papers cited above. In [36–46], various concepts related with paradigms like soft computing, artificial intelligence, etc., are discussed. The concepts mentioned in the above citations are the fundamental in the drafting of the current objective. In [47–55], Python is utilized to solve various real life problems. The list of the above citation covers the wide spectrum of the libraries and the features there in Python that could be proved handy in the current work as well. In [56], watermarking is discussed. Security is a very essential feature you want to have in any application and that is what is dispensed in the above citation. In [57, 58], the authors gave insight about Soft computing and AI. Again, we cite papers that spoke on the fields of AI and SC that are the implementation work beds. In [59], implementation of code is discussed using C. How to make implementation and how to go for programming is dealt in the above citation. In [60], VLSI architecture is demonstrated. No discussion in computer is covered without the test bed, and VLSI architecture is the base. In [61] Stability and convergence is discussed. Perceptron and its features dealt in the above citation. In [62], we observe a Python-based implementation that goes on with the concepts of NLP and also makes utilization of various basic Python libraries to have training and then gives a good accuracy on news classification. In [63] we observe a system that makes cartoonification with an objective to trace forgeries. This paper also makes utilization of ML and Python like what we will be doing in the current work. In [64], we observe a search engine that is having ML and Python embedding to do the task in the best possible way.

4   Proposed Implementation and Result

The implementation of the code is done in Python. The algorithm starts with importing of the libraries, importing of the dataset, preprocessing of the data which comprise of checking information in the dataset and data cleaning, exploration of the data using the exploratory libraries, checking of the columns for numeric features, setting up the columns such that they can go for training, testing of the ML models for results, searching the best model, best model chosen for implementation, perform hyper tuning on the best model. The tuned model is saved for future recall, split the dataset into training and testing, use the best model for testing and compare the predicted and true values obtained and at last find the accuracy measure.

We made a utilization of a dataset taken from Kaggle. The dataset is a standard dataset and is meant for the experimental purpose. We made all the work in the field to make the best utilization of the dataset for the purpose of drilling to get our work done. We have made a very good utilization of the language to go up the scene to have our work done.

The dataset comprises of seven features all in all. The size of the dataset is around 59 records. The size is 1.66 KB. Although the number of records is not that much but the records are such nicely diced as we can go up in the panel to get good results of the experiment done by us in the current paper.

The Python code used is a landmark for a lot of purposes. As it sets a benchmark to think of which approach to go with when we are having the drilling operation. We had presented some of the figures (Figs. 1, 2, 3) showing the relationship between the various approaches, and we also will see several parameters that are used to make the assessment of the algorithms to go through with the point that which we are interested in. The work is a very good work in for all of us. The steps carried out are already described in the section above. The implementation is very useful for the mankind as this is going to make automation in the current system and the approach is going to utilize the ML approach in totality.
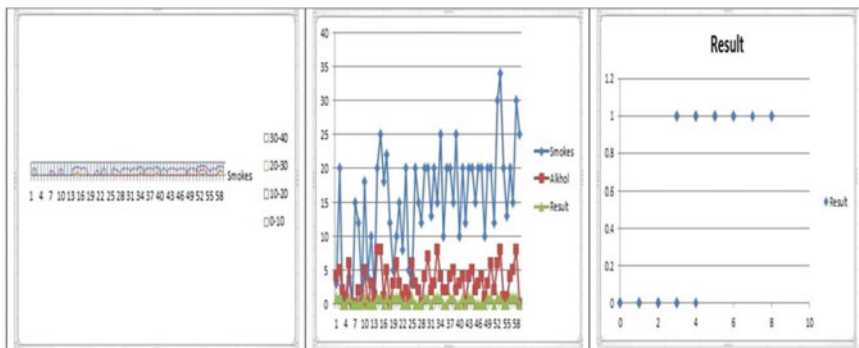
Fig. 1 Dataset visualization for data preprocessing and column selection

	Model	Accuracy	AUC	Recall	Prec.	F1	Kappa	MCC	TT (Sec)
lr	Logistic Regression	0.9762	0.9683	0.9524	1.0000	0.9714	0.9524	0.9582	0.7914
nb	Naive Bayes	0.9762	1.0000	0.9524	1.0000	0.9714	0.9524	0.9582	0.0071
ridge	Ridge Classifier	0.9762	0.0000	0.9524	1.0000	0.9714	0.9524	0.9582	0.0071
qda	Quadratic Discriminant Analysis	0.9762	1.0000	0.9524	1.0000	0.9714	0.9524	0.9582	0.0086
lda	Linear Discriminant Analysis	0.9762	1.0000	0.9524	1.0000	0.9714	0.9524	0.9582	0.0143
et	Extra Trees Classifier	0.9762	1.0000	0.9524	1.0000	0.9714	0.9524	0.9582	0.0671
rf	Random Forest Classifier	0.9524	1.0000	0.9524	0.9643	0.9510	0.9048	0.9163	0.0829
svm	SVM - Linear Kernel	0.9476	0.0000	0.9524	0.9643	0.9510	0.8874	0.9028	0.0057
ada	Ada Boost Classifier	0.9286	0.9921	0.9524	0.9357	0.9351	0.8435	0.8638	0.0343
knn	K Neighbors Classifier	0.9238	0.9365	0.9048	0.9643	0.9224	0.8498	0.8687	0.4457
dt	Decision Tree Classifier	0.8810	0.8690	0.9048	0.9000	0.8862	0.7483	0.7801	0.0086
gbc	Gradient Boosting Classifier	0.8810	0.8770	0.9048	0.9000	0.8862	0.7483	0.7801	0.0200
lightgbm	Light Gradient Boosting Machine	0.5381	0.5000	1.0000	0.5381	0.6976	0.0000	0.0000	0.2514
dummy	Dummy Classifier	0.5381	0.5000	1.0000	0.5381	0.6976	0.0000	0.0000	0.0071

Fig. 2 Comparison chart of the various algorithms on the dataset

Fig. 3 Classification report and tuned rf report for the model

5 Conclusion

The final model chosen for implementation is the rf model. The accuracy obtained is of around 93.33% which is a good accuracy measure, and thus, we can go for the model for prediction of lung cancer so that medication can start at proper time.

The pycaret-based implementation gives a comparison chart shown in Fig. 2. The accuracy value of logistic regression is 97%. The accuracy value of Naïve Bayes is 97%. The accuracy value of ridge classifier is 97%. The accuracy value of quadratic discriminant analysis is 97%. The accuracy value of linear discriminant analysis is 97%. The accuracy value of extra tree classifier is 97%. The accuracy value of

random forest classifier is 95%. The accuracy value of SVM-linear kernel is 94%. The accuracy value of AdaBoost classifier is 92%. The accuracy value of K-neighbour classifier is 92%. The decision tree classifier gives accuracy of 88%. The gradient boosting classifier gives accuracy of 88%. The light gradient boosting machine gives accuracy of 53%, and dummy classifier gives accuracy of 53%.

References

1. Singh VK (2021) SVM using rbf as kernel for diagnosis of breast cancer. In: International conference on innovative research in science, management and technology (ICIRSMT 2021). Department of Computer Science and Application, Atal Bihari Vajpayee University, Bilaspur (CG), India in association with American Institute of Management and Technology (AIMT), USA, pp 343–348. ISSN: 2769-5093
2. Singh VK (2021) Support vector machine using rbf, polynomial, linear and sigmoid as kernel to detect diabetes cases and to make a comparative analysis of the models. In: International conference on innovative research in science, management and technology (ICIRSMT 2021). Department of Computer Science and Application, Atal Bihari Vajpayee University, Bilaspur (CG), India in association with American Institute of Management and Technology (AIMT), USA, p 25. ISSN: 2769-5093
3. Singh VK, Yadav ND, Singh RK, Sahu M (2022) Detection of thyroid using machine learning approach. NIU Int J Hum Rights India (UGC Care Listed J) 3(II):65–80. ISSN: 2394–0298
4. Yadav ND, Singh VK, Singh RK, Sahu M (2022) A comparative analysis of SVM kernels for detection of diabetes. Anvesak India (UGC Care Listed J) 52(5):61–67. ISSN: 0378–4568
5. Yadav ND, Singh VK, Singh RK, Sahu M (2022) ML based SVM taking RBF as kernel for detection of breast cancer. J Educ Rabindra Bharati Univ India (UGC Care Listed J) XXV(5):72–79. ISSN: 0972-7175
6. Singh VK (2022) Support vector machine an ML approach for breast cancer. In: 6th International Izmir economics congress, Izmir, Turkey
7. Singh VK (2022) SVM an ML approach for diabetes. In: 6th International Izmir economics congress, Izmir, Turkey
8. Kumari P, Gupta R, Kumar S, Singh VK (2021) ML approach for detection of lung cancer. In: ISPEC 8th International conference on agriculture, animal science and rural development, Bingol, Turkey
9. Singh VK (2022) Pycaret a python implementation for thyroid disease. In: IArcSAS 2nd International architectural sciences and applications symposium, Baku, Azerbaijan
10. Singh VK (2022) SVM an ML approach for breast cancer. In: IArcSAS 2nd International architectural sciences and applications symposium, Baku, Azerbaijan
11. Singh VK (2022) An implementation in python for diagnosis of thyroid using machine learning approach. In: 9 International GAP summit scientific research congress, Adiyaman, Turkey
12. Singh VK (2022) ML based python implementation for detection of lung cancer. In: IV International halich congress on multidisciplinary scientific research, Istanbul, Turkey
13. Singh VK (2022) A computerized approach for thyroid disease having the usage of ML technique. In: IV International halich congress on multidisciplinary scientific research, Istanbul, Turkey
14. Singh VK (2016) Proposing solution to XOR problem using minimum configuration MLP. In: Science direct international conference on computational modeling and security (CMS 2016), vol 85. Procedia Computer Science, Bangalore, India, pp 263–270
15. Singh VK, Pandey S (2016) Minimum configuration MLP for solving XOR problem. In: Proceedings of the 2016 international conference on computing for sustainable global development (INDIAcom), New Delhi, India. IEEE, pp 168–173. IEEE Conference ID: 37465. ISBN: 978–9–3805–4421–2

16. Singh VK, Pandey S (2016) Proposing an Ex-NOR solution using ANN. In: Proceedings of the International conference on information communication and computing technology (ICICCT-2016). JIMS, Delhi, India, pp 277–284. ISBN: 978–93–85777–66–0
17. Singh VK (2016) Mathematical explanation to solution for Ex-NOR problem using MLFFN. Int J Inf Sci Tech 6(1/2):105–122
18. Singh VK (2016) Proposing a new ANN model for solving XNOR problem. In: 2016 International conference system modeling and advancement in research trends (SMART), Moradabad, India. IEEE, pp 32–36. ISBN: 978–1–5050–3543–4
19. Singh VK, Singh DK (2015) Multilayer perceptron for XOR problem. In: IRSMT-2015, Bilaspur, India, p 33
20. Singh VK (2015) Two solutions to the XOR problem using minimum configuration MLP. Int J Adv Eng Sci Technol Res (IJAESTR) 3(3):16–20. ISSN: 2321–1202
21. Singh VK (2015) One solution to XOR problem using multilayer perceptron having minimum configuration. Int J Sci Eng 3(2):32–41. ISSN: 2347–2200
22. Singh VK (2016) ANN Implementation of construction logic gates focuses on Ex-NOR. Res J Comput Inf Technol Sci 4(6):1–11. E-ISSN: 2320–6527
23. Singh VK (2017) Designing simulators for various VLSI designs using the proposed artificial neural network model TRIVENI. In: 2017 International conference on information, communication, instrumentation and control (ICICIC), Indore, India. IEEE, pp 1–6. ISBN: 978–1–5090–6313–0
24. Singh VK, Singh AK, Deo Y, Reddy DS (2022) A dual authentication implementation using biometric properties of fingerprint following COVID protocol and face matching age and gender. In: Intellectual Property of India, Application Number: 202221063732
25. Singh VK, Yadav ND (2022) SFG for showing the solution to XOR problem which is a non linearly separable problem. In: Intellectual Property of India, Application Number: 202221041900
26. Singh VK, Yadav ND, Singh RK (2022) BDR for solving linearly separable problems like AND, OR, NAND and NOR. In: Intellectual Property of India, Application Number: 202221047404
27. Singh VK, Yadav ND, Singh RK (2022) SFG for solving non-linearly separable problem that is Ex-NOR problem. In: Intellectual Property of India, Application Number: 202221041911
28. Pippal RKS, Singh VK, Gupta P, Gangil M, Bux S, Sadiwala R, Hande JY, Dorave JD (2022) Method for migration of sessions among cipher machines using Lagrange interpolation. In: Intellectual property of India, Application Number: 202221000273
29. Singh VK, Singh VK (2015) Vector space model: an information retrieval system. Int J Adv Eng Res Stud 4(2):141–143
30. Singh VK (2019) Proposing pattern growth methods for frequent pattern mining on account of its comparison made with the candidate generation and test approach for a given data set. In: Software engineering. Springer Nature, Singapore, pp 203–209
31. Singh VK (2010) RSTDB and cache conscious techniques for frequent pattern mining. In: Proceedings of the 4th international conference on computer application in electrical engineering recent advances (CERA-09). Indian Institute of Technology, Roorkee, India
32. Singh VK, Thakur VS, Pandey NK (2010) Proposing data mining as an efficient technique for rural development. In: Proceedings of the international multi conference on intelligent system and nanotechnology, Haryana, India, pp 110–112
33. Singh VK, Singh VK (2009) RSTDB a new candidate generation and test algorithm for frequent pattern mining. In: Proceedings of the international conference on advances in communication, network and computing, Calicut, Kerala, India. CNC-2010, ACEEE and IEEE, IEEE Communication Society, Washington, DC. ACM DL Digital Library, pp 416–418. ISBN: 978–0–7695–4209–6
34. Singh VK, Singh VK Minimizing space time complexity by RSTDB a new method for frequent pattern mining. In: Proceedings of the first international conference on human computer interaction (IHCI 2009). Indian Institute of Information Technology, Allahabad, India. Springer, pp 361–371

35. Singh VK, Dubey V, Singh AK (2010) Proposing data mining as an efficient technique for solving frauds in digital data. In: Proceedings of the international conference on intelligent information systems and management (IISM'2010), India, pp 1–4. ISBN: 978–1–4507–2041–0

36. Singh VK (2022) A short report making a comparison between artificial intelligence and artificial neural network. In: IArcSAS 2nd International architectural science and applications symposium, Baku, Azerbaijan

37. Singh VK (2016) Mathematical analysis for training ANNs using basic learning algorithms. Res J Comput Inf Technol Sci 4(7):6–13. E-ISSN: 2320–6527

38. Singh VK, Shah V, Jain YK, Shukla A, Thoke AS, Singh VK, Dule C, Parganiha V (2008) Proposing an efficient method for frequent pattern mining. In: Proceedings of the world academy of science engineering and technology, Bangkok, Thailand, pp 1184–1189. ISSN: 2070–3740

39. Singh VK (2022) A short report on the basic artificial neuron model. In: IArcSAS 2nd International architectural science and applications symposium, Baku, Azerbaijan

40. Singh VK (2022) Fuzzy associative memory for truck backer system. In: International conference on global practice of multidisciplinary scientific studies, Turkey

41. Singh VK (2022) Backpropagation algorithm of ANN. In: International conference on global practice of multidisciplinary scientific studies, Turkey

42. Singh VK (2022) A report on principal component analysis. In: International conference on global practice of multidisciplinary scientific studies, Turkey

43. Singh VK (2022) Historical development in the field of artificial intelligence. In: 2nd International archeology, art history and cultural heritage congress, Sanliurfa, Turkey

44. Singh VK (2022) A study on elementary neurophysiology of human brain. In: IV International halich congress on multidisciplinary scientific research, Istanbul, Turkey

45. Singh VK (2022) A study on feedback in an artificial neural system. In: IV International halich congress on multidisciplinary scientific research, Istanbul, Turkey

46. Singh VK (2022) A study on artificial neural network when viewed as a directed graph. In: IV International halich congress on multidisciplinary scientific research, Istanbul, Turkey

47. Singh VK (2021) Colorization of old gray scale images and videos using deep learning. J Orient Res Madras (UGC Care Listed J), India, pp 44–49. ISSN: 0022–3301

48. Chandrashekhar RC, Singh VK (2021) Twitter sentiment analysis. In: ISPEC 8th International conference on agricultural, animal science and rural development, Bingol, Turkey

49. Sailokesh P, Jupudi S, Vamsi IK, Singh VK (2021) Automatic number plate recognition. In: ISPEC 8th International conference on agriculture, animal science and rural development, Bingol, Turkey

50. Reddy YK, Yadav KM, Singh VK (2021) Human activity recognition. In: ISPEC 8th International conference on agriculture, animal science and rural development, Bingol, Turkey

51. Prasad RNRK, Ram PSSR, Dinesh S, Singh VK (2021) Text summarization. In: ISPEC 8th International conference on agricultural, animal science and rural development, Bingol, Turkey

52. Singh VK (2022) An image based search engine system using python. In: 9th International GAP summit scientific research congress, Adiyaman, Turkey

53. Singh VK (2022) A system for news classification in regional language using python. In: 9th International GAP summit scientific research congress, Adiyaman, Turkey

54. Singh VK (2022) A system for cartoonifying an image using python. In: 9th International GAP summit scientific research congress, Adiyaman, Turkey

55. Singh VK (2022) A sign language recognition system. In: 9th International GAP summit scientific research congress, Adiyaman, Turkey

56. Singh VK, Singh AK (2010) Dual level digital watermarking for images. In: Proceeding of the CP1324, international conference on methods and models in science and technology (ICM2ST-10), vol 1324. AIP Conference Proceeding, Chandigarh, India, pp 284–287. ISBN: 978–0–7354–0879–1

57. Singh VK (2022) Soft computing, 1st edn. Book Rivers, India. ISBN: 9789355154026

58. Singh VK (2022) Advanced artificial intelligence, 1st edn. Book Rivers, India. ISBN: 9789355154019

59. Singh VK, Singh A, Singh A, Yadav ND (2022) Lab manual of web technology and C-programming. Ideal International E-Publication Private Limited. ISBN: 978–93–89817–67–6
60. Singh VK, Singh VK (2008) VLSI architecture for multitier wireless systems. In: International conference on: interdisciplinary approaches in physical science: growing trends and recent advances CONIAPS-X, Bilaspur, India, p 166
61. Singh VK Analysis of stability and convergence on perceptron convergence algorithm. In: Proceeding of international conference on information, communication and computing technology (ICICCT 2017), Delhi, India, pp. 149–161. ISBN: 978–93–86647–85–6
62. Singh VK, Soni SK, Yadav ND, Chandra P (2022) A system for news classification in regional language using python. Shodh Prabha Referred and Peer Rev Q Res J 47(4):169–182. ISSN: 0974–8946
63. Singh VK, Soni SK, Yadav ND, Chandra P A System for Cartoonifying an Image Using Python. SAMRIDDHI J Physical Sci Eng Technol 14(5):84–100. ISSN: 2229–7111 (Print); ISSN: 2454-5767 (Online)
64. Yadav ND, Singh VK (2022) An image based search engine system using python. IJFANS Int J Food Nutr Sci 11(3):2821–2836. ISSN: 2319–1775 (Print); ISSN: 2320–7876 (Online)

Predicting the Heart Attacks Risk Using Artificial Neural Networks

Rayi Naveen Kumar and Mullapudi Navyasri

1 Introduction

The heart assault is a typical issue in every single individual through the oldness over thirty years. The fat quantity is additional serious issue which prompts heart assault. The information disclosure in records is an appropriately procedure holding a few unmistakable strides to get the ideal precision [1]. Intelligence extracting is the core advance, results in the disclosure of shrouded data with valuable learning. The found learning will be applied through the societal coverage overseers to anticipate a portion of the infections and issues like heart assaults [2]. Anticipating patient's conduct later on is the fundamental use of information mining procedures. A proper significance of learning acknowledgement in records is specified as pursues: "Intelligence extracting is the non-unimportant drawing out of certain previously ambiguous in addition feasibly important data about knowledge" [3]. Restorative conclusion is a significant yet confused undertaking that should be done precisely and effectively. The robotization of this framework is particularly expected to assist the doctors with doing better finding and treatment [4]. Underprivileged medical adoptions be able to rapid terrible outcomes which remain available this mode inadequate. The significant test of the human services framework these days is to foresee the infections in quality way. The clinical choices made by specialists might be mistake inclined and prompts a few issues for patients [5]. This framework ought to be computerized here the way that to foresee the maladies in the exact way. The main principle issues inside human services framework remains to foresee the atria and ventricles chambers assaults of the patient role already. To hand certain systems accessible to foresee these mechanisms in an exact way. The accessible Intelligence extracting procedures remain not

R. N. Kumar · M. Navyasri (✉)
Department of CSE, KCC Institute of Technology and Management, Greater Noida, Uttar Pradesh, India
e-mail: navyasrimullapudi@gmail.com

utilized appropriately to anticipate the maladies in the social insurance frameworks [6, 7]. Sixty years is the normal period of atria and ventricles chambers assaults patients in India contrary to sixty-three years to sixty-eight in created nations. In India Andhra Pradesh state is in danger of more passing due to CHD. Thus, there is a need to battle the coronary illness. Conclusion of the coronary illness in the early stage tackles numerous lives [8].

2 Literature Survey

There are numerous methodologies and calculations have remained and utilized in the direction of anticipate the atria and ventricles chambers assaults. Anbarasi et al. [9] described an improved forecast of coronary illness with highlight subset determination utilizing hereditary calculation. They anticipated all the more precisely the nearness of coronary illness with diminished number of qualities. They utilized Innocent Bayes, Grouping, and Choice Tree techniques to anticipate the analysis of patients with a similar precision as acquired before the decrease of characteristics. They inferred that the choice tree technique outflanks the other two strategies [10]. Contrasting with the works talked about over, this work is distinctive with the utilization of neural system and choice tree calculation in an integrative way to foresee heart assaults with high measure of precision [11]. Neural centred knowledge classifier framework for characterizing Intelligence extracting undertakings is projected in the paper Dam et al. [12]. They directed trials on 13 diverse record sets as of the College of California, one fake dataset and Irvine store. They demonstrated that Neural centred knowledge classifier framework accomplishes identically to managed learning classifier framework on top of five datasets, essentially great execution lying on six record sets and altogether terrible showing on top of three record sets [13]. Patil and Kumaraswamy [14], proposed a smart as well as powerful atria and ventricles chambers assault forecast framework utilizing information mining and fake neural system. They additionally proposed removing noteworthy examples for coronary illness expectation. They utilized K-implies bunching to remove the information proper to atria and ventricles chambers assault as of the storage warehouse. They utilized MAFIA calculation to extract the successive patterns [9, 15]. Guru et al. [16] anticipated a choice emotionally supportive network for coronary illness determination utilizing neural system. They prepared their framework with 78 patient records and the mistakes made by people are maintained a strategic distance from in this framework [17].

3 Methodology

Artificial neural system (ANN) is one of the best classification algorithms to in the prediction of heart diseases. It is the re-enactment of human mind and is being connected to an undeniably number of genuine issues. Neural systems as an instrument we can mine learned information from information product house. ANN are prepared to perceive, store, and recover examples to take care of combinatorial streamlining issues [12]. The artificial network has multiplayer perceptron algorithm which can have 3 layers. They are layer of inputs, hidden, and output layers (Fig. 1).

Example acknowledgement and capacity estimation capacities make ANN pervasive utility in information mining. Their principle bit of leeway is that they can take care of issues that are unreasonably mind boggling for regular innovations. Neural systems are appropriate to issue like example acknowledgment and forecasting. ANN are utilized to extricate valuable examples from the information and derive rules from them. These are valuable in giving data on affiliations, groupings, and clustering [14]. In this paper we present a characterization approach which consolidates multi-layer discernment with back spread learning calculation and highlight determination for arrangement of coronary illness with decreased no of characteristics. Our methodology him demonstrates characterization and decides the characteristics which contribute more towards the predication of coronary illness which in a roundabout way Lessens no. of finding test which are required cylinder taken by a patient. The list of parameters is given below

Instances: 270

Attributes: 14

1. 1. Patient's Age—oldness or age in years
2. Patient's Sex—man or woman
3. Rib cage—patient's chest pain type
4. Relaxing—Haemoglobin/blood pressure
5. Blood serum—cholesterol

Fig. 1 Neural network model
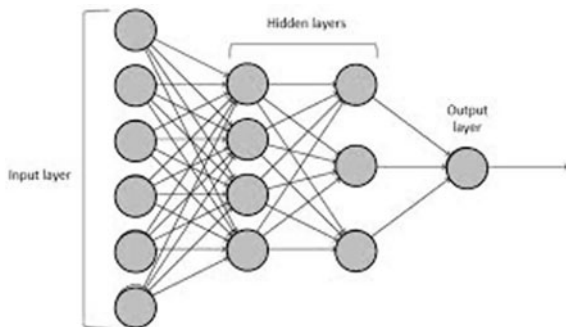
6. Blood—sugar while fasting
7. ECG—(Resting electrocardiographic results)
8. Extreme—achieved heart rate
9. Exercise—induced angina
10. Old peak—ST depression by exercise related to rest
11. Slope—slope of the highest exercise ST segment
12. Amount of main vessels—coloured by fluoroscope
13. Defect of Thal type.

4 Experimental Results

See Tables 1, 2, 3, 4 and Figs. 2 and 3.

Table 1 Prediction of the risk of heart disease by neural network with cross validation-5

Rate of TP	Rate of FP	Accuracy	Recall	F-Evaluation	MCC	Area-ROC	Area-PRC	Labels or classes
0.800	0.242	0.805	0.800	0.803	0.558	0.856	0.865	Absent
0.758	0.200	0.752	0.758	0.755	0.558	0.856	0.852	Present

Table 2 Confusion matrix with cross validation-5

Class	a (Patients are having heart disease)	b (Patients are Not having heart disease)
a (Patients are having heart disease)	True (Correctly classified instances) positive 120	False (In correctly classified instances) negative 29
b (Patients are not having heart disease)	False (In correctly classified instances) positive 30	True (Correctly classified instances) negative 91

TP (True Positive): While the number of instances are actually true, true positive represents the same that means the number of records classified as true

FN (False Negative): While the number of instances are actually true, false positive represents the number of records classified as false

FP (False Positive): While the number of instances are actually false, false positive represents the same that means the number of records classified as false

TN (True Negative): While the number of instances are actually true, true positive represents the same that means the number of records classified as true

Table 3 Prediction of the risk of heart disease by using neural network with cross validation-10

Rate of TP	Rate of FP	Accuracy	Recall	F-Evaluation	MCC	ROC area	PRC area	Label or classes
0.780	0.217	0.818	0.780	0.799	0.561	0.839	0.826	Absent
0.783	0.220	0.740	0.783	0.761	0.561	0.839	0.824	Present

Table 4 Confusion matrix with cross validation-10

Class	A (has heart disease)	B (no heart disease)
A (has heart disease)	117	33
B (no heart disease)	26	99
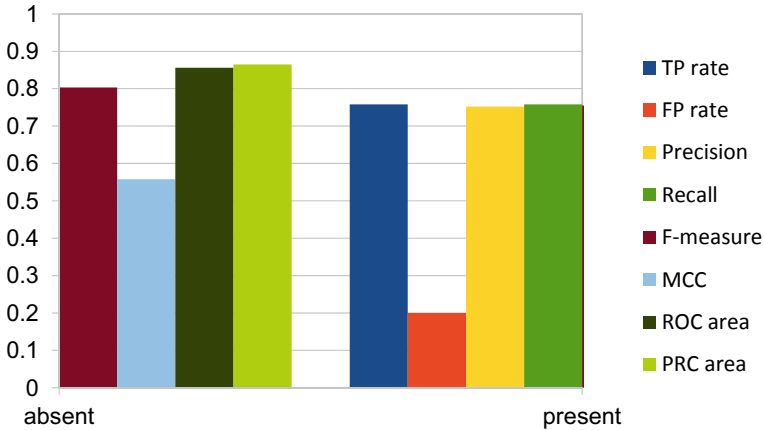
Fig. 2 Pictorial representation of heart disease prediction system using neural network with cross validation-5
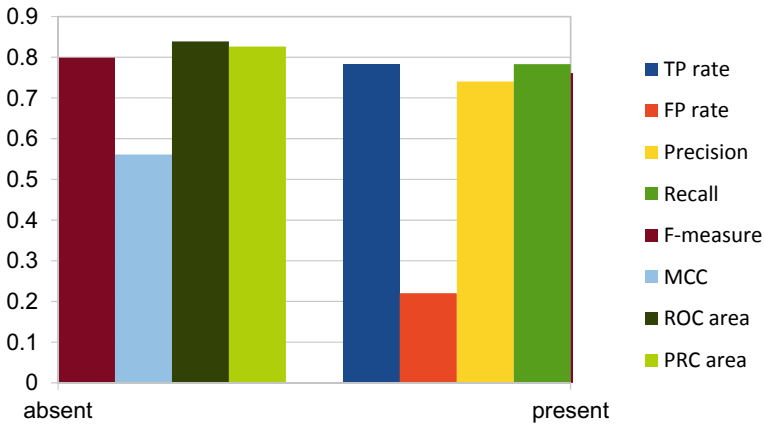
Fig. 3 Pictorial representation of heart disease prediction system using neural network with cross validation-10

Assessments were led with Weka 3.6.0 instrument. Informational collection of 270 records with 13 qualities is utilized. All qualities are made all out and irregularities are settled for effortlessness. The age of new populace proceeds until it develops
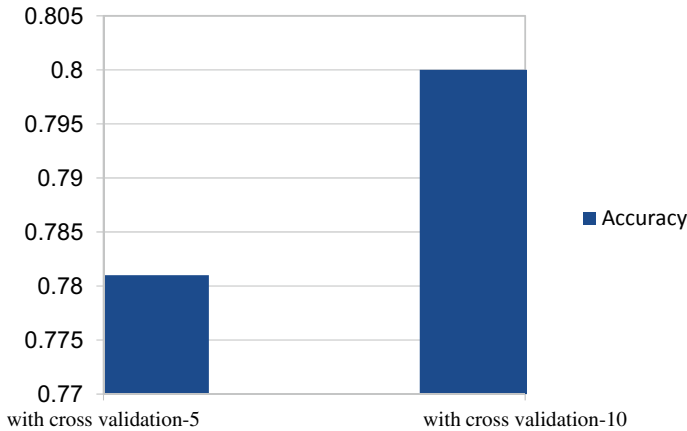
Fig. 4 Graph shows the accuracy for cross validation-5 and cross validation-10

a populace where each principle is fulfilled by the populace. With 0.6 traverse likelihood and 0.033 transformation likelihood, which contributes more towards the finding of the heart sickness. Results are appeared Table 1 and Table 3. The test are conducted with cross validation-5 and cross validation-10. 80% accuracy obtained with cross validation-10. The graphical representation is shown in Fig. 4.

5   Conclusion

We believed the subject of obliging plus abridging various calculations of intelligence extracting. We focused on applying various results for anticipating combinations of a few target qualities. In this manuscript, we develop a canny plus compelling heart assault expectation strategies utilizing information mining. In our future work, this can additionally improve and extended. For anticipating heart assault fundamentally 13 properties are recorded. Other than the 13 recorded in medicinal writing we can likewise consolidate other information mining strategies, such that Time Arrangement, Grouping, and Affiliation Principles.

References

1. Mert A, Kiliç N, Akan A (2014) Evaluation of bagging ensemble method with time-domain feature extraction for diagnosing of arrhythmia beats. Neural Comput Appl 24(2):317–326
2. Nahar J, Imam T, Tickle KS, Chen Y-PP (2013) Association rule mining to detect factors which contribute to heart disease in males and females. Expert Syst Appl 40(4):1086–1093. https://doi.org/10.1016/j.eswa.2012.08.028

3. Frawley WJ, Piatetsky-Shapiro G (1996) Knowledge discovery in databases: an overview. The AAAI/MIT Press, Menlo Park, CA
4. Srinivas K, Rani BK, Govrdhan A (2010) Applications of data mining techniques in healthcare and prediction of heart attacks. Int J Comput Sci Eng (IJCSE) 2(2):250–255
5. Mourão A, Martins F, Magalhães J (2015) Multimodal medical information retrieval with unsupervised rank fusion. Comput Med Imaging Graph 39:35–45
6. Chaurasia V, Pal S (2013) Data mining approach to detect heart diseases. Int J Adv Comput Sci Inf Technol (IJACSIT) 2(4):56–66
7. Chaurasia V, Pal S (2013) Early prediction of heart diseases using data mining techniques. Carib J Sci Technol 1:208–217
8. Krishnapuram B, Harternink AJ, Carin L, Figueiredo MAT (2004) A Bayesian approach to joint feature selection and classifier design. IEEE Trans Pattern Anal Mach Intell 26(9):1105–1111
9. Anbarasi M, Anupriya E, Iyengar NCSN (2010) Enhanced prediction of heart disease with feature subset selection using genetic algorithm. Int J Eng Sci Technol 2(10):5370–5376
10. Palaniappan S, Awang R (2009) Intelligent heart disease prediction system using data mining techniques. Int J Comput Sci Netw Secur 8(2):235–241
11. Moftah RA, Maatuk AM, White R (2016) Methods to access structured and semi-structured data in bioinformatics databases: a perspective. In: 2016 International conference on engineering and MIS (ICEMIS). IEEE, pp 1–5
12. Dam HH, Abbass HA, Lokan C, Yao X (2008) Neural-based learning classifier systems. IEEE Trans Knowl Data Eng 20(1):26–39
13. Khan YD, Ahmad F, Anwar MW (2012) A neuro-cognitive approach for iris recognition using back propagation. World Appl Sci J 16(5):678–685
14. Patil SB, Kumaraswamy YS (2009) Extraction of significant patterns from heart disease warehouses for heart attack prediction. Int J Comput Sci Netw Secur 9(2):228–235
15. Chowdhury DR, Chatterjee M, Samanta RK () An artificial neural network model for neonatal disease diagnosis. Int J Artif Intell Expert Syst (IJAE) 2(3):96–106
16. Guru N, Dahiya A, Rajpal N (2007) Decision support system for heart disease diagnosis using neural network. 8(1):99–101
17. Patil SB, Kumaraswamy YS (2009) Intelligent and effective heart attack prediction system using data mining and artificial neural network. Eur J Sci Res 31(4):642–656

A Printed Character Recognition System for Meetei-Mayek Script Using Transfer Learning

Vishwaksena Vishnu Simha Dingari, Ganapathi Kosanam, Devi Sri Shankar Chavatapalli, and Chingakham Neeta Devi

1 Introduction

A text in English can be easily identified but when it comes to northeastern India, many languages are there which are less developed. Manipuri is the official language written using Meetei-Mayek and has another script - Bengali which is borrowed from Bengali language [1]. It is a challenge for everyone to understand and works must be carried out for development of such languages. Only a handful of works have been found in literature for Meetei-Mayek Character Recognition [2–5]

In research, character recognition is applicable in a wide range of fields that can help to save time. However, for detecting printed characters in regional scripts, researchers still need to come up with a more trustworthy method or, more precisely, a system. This would require collecting images containing Meetei-Mayek text before proceeding for processes for recognition purpose. To recognize the characters in the image, a deep learning model can be used, such as VGG-16, VGG-19, and ResNet152-V2, which will be trained on this dataset.

Following the training of the model, it can be used to recognize printed Meetei-Mayek characters from images. By pre-processing the image, an image format can be created that can be input to the model. Based on these predictions, the model is able to recognize the Meetei-Mayek characters from the input image.

V. V. S. Dingari (✉) · G. Kosanam · D. S. S. Chavatapalli · C. N. Devi
National Institute of Technology Manipur,Imphal795004, India
e-mail: dingarivishwaksenavishnusimha@gmail.com

G. Kosanam
e-mail: kosanam.ganapathi@gmail.com

D. S. S. Chavatapalli
e-mail: dsshankar2017@gmail.com

C. N. Devi
e-mail: neeta.ch@nitmanipur.ac.in; neeta.chingakham@gmail.com

The characters of Meetei-Mayek are substantially more complicated than those in standard English because of the addition of modifiers, forms, and structures. In order to efficiently classify these characters, a sophisticated pattern recognition algorithm will need to be developed. Furthermore, as of now no research has been published to date on text recognition based on printed Meetei-Mayek character recognition using transfer learning.

Through optical character recognition (OCR) [6–8], an image of text can be converted into the machine-readable Unicode format. This can be achieved through segmentation and CNN. However, while segmentation, the extraction of text is difficult and not so accurate (in some cases).

Convolutional neural networks (CNNs) are a variant of neural networks that are frequently used to classify images. Text classification is being done with CNN, where images with text is converted into each character as the down-sampling technique has been used in convolutional neural networks.

Transfer learning involves the use of a pre-trained model as it is or with some added layers to analyze an image. This can be implemented when there is a lot of data for each character which can be used to train a new model. In this case, VGG-16, VGG-19, and ResNet152-V2 are used for analysis of classification of characters.

1.1 Brief on CNN and Transfer Learning

The primary motive is to recognize printed Meetei-Mayek characters from an image and get it as an editable text. So first segmentation method is used, which includes finding contours and extracting each segment (character) from the image. Then, train the pre-trained CNN is re-trained which is known as transfer learning with the segregated segmented images. Now the trained CNN can be used to identify images with individual characters.

Convolutional Neural Networks: Convolutional neural networks, often known as CNN or ConvNet, are a subtype of neural networks [9, 10] that are particularly skilled in processing data with a grid-like design. Convolutional, pooling, and fully connected layers make up the majority of a CNN.

Convolutional layers extracts high-level features from the input image. The starting convolutional layers extract low-level features. As one propagate through the neural network, the later convolutional layers learn high-level features similar to the way humans perceive images.

After applying convolution functions, the output spatial size is decreased by pooling layers. Through dimensionality reduction, this aids in lowering the amount of computing power needed to process the data.

Fully-connected layer is used for classification of the image using the high-level features extracted using the previous convolutional layers and pooling layers.

There are several variants of convolutional neural networks (CNNs) that are commonly used for character recognition. Some examples include LeNet, AlexNet, VGGNet, ResNet, Inception.

These are just a few examples of CNN architectures that have been used for character recognition. It is important to note that the best architecture for a specific task will depend on the dataset, the complexity of the task, and the resources available for training. For the Meetei-Mayek character recognition using transfer learning, the following CNN's has been compared: VGG-16, VGG-19, and ResNet152-V2.

Transfer Learning: Algorithms for classical machine learning and deep learning have, up to now, often been created to operate independently. Transfer learning is the concept of eschewing the isolated learning paradigm and applying the information obtained to tackle connected problems.

Transfer learning is a technique where a model that has been trained for one task is utilized as the foundation for another related task. When it comes to character recognition using CNNs, the process of transfer learning can be broken down into pre-training, freezing layers, fine-tuning, training, evaluation.

In summary, transfer learning allows us to use a pre-trained CNN model as a starting point for a new model, which can save time and resources when training the new model, and also improve its accuracy. The pre-trained model is used to extract features that are useful for character recognition and these features are used to train a new model on a target dataset. Here, transfer learning is implemented using VGG-16 [11], VGG-19 [12], and ResNet152-V2 [13] from which a model can be selected as per the requirements based on the results obtained.

VGG-16, VGG-19 and ResNet152-V2: VGG-16, VGG-19 (VGG: Visual Geometry Group) [14], and ResNet152-V2 (ResNet: Residual Neural Network) [15] are convolutional neural networks that are 16, 19, and 152 layers deep, respectively. These CNN's are pre-trained models using the ImageNet database which contains more than a million images. The trained network can classify images into more than a thousand different object categories, including keyboard, mouse, pencil, and several animals. This pre-trained model is used for classification of printed Meetei-Mayek characters using transfer learning.

2 Related Works

In literature, there have been few reports on Meetei-Mayek Character Recognition System. However, no work has been found to recognize characters using transfer learning of printed Meetei-Mayek characters to date. Therefore, a broad scope of research needs to be done associated with Manipuri script recognition. There have been several studies on the use of deep learning for Meetei-Mayek character recognition. Some examples include

Thokchom et al. [3] used a connected component analysis and extracted probabilistic and fuzzy features from the resized character matrix. Their model was com-

pared with other machine learning algorithms, and the neural network was found to have achieved the highest accuracy. Inunganbi et al. [4] used background directional distribution, projection histogram, and uniform local binary pattern features. They extracted distinct features for recognition using KNN classifier and achieved the highest accuracy on the dataset. Meetei et al. [5] used several algorithms, including EAST, MSER, and SWT, to detect Meetei-Mayek and Mizo text in natural images and document images. Optical character recognition was used to analyze the detected text in both languages, and the EAST algorithm was found to be the most accurate on the dataset they used.

As a whole, these studies suggest that deep learning can be an effective approach for extracting Meetei-Mayek characters from text. Meetei-Mayek OCR using transfer learning has not yet been implemented yet. In the process of creating an OCR for Meetei-Mayek, the database of characters using color images will be generated, which will be converted into three channeled black and white images for neural network training.

3   Meetei-Mayek Printed Character Segmentation

A description of the process of extracting characters and creating database is provided in this section. As input images, only text was separated from newspaper articles, and the present algorithm is designed to detect the text and divide each word into a character (Fig. 1).
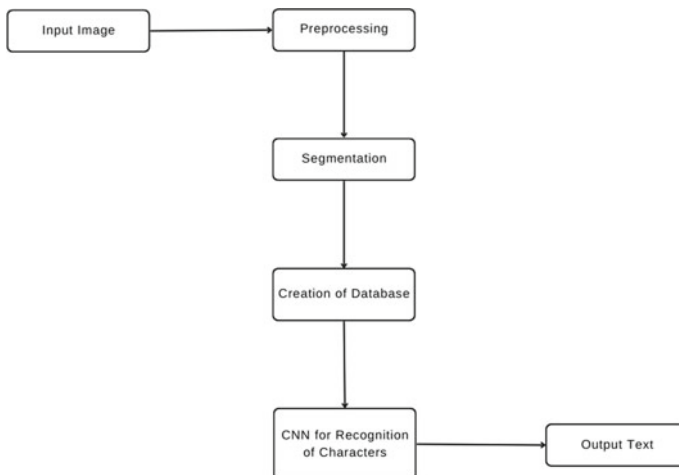
Fig. 1 Flowchart that depicts the process of text extraction from images
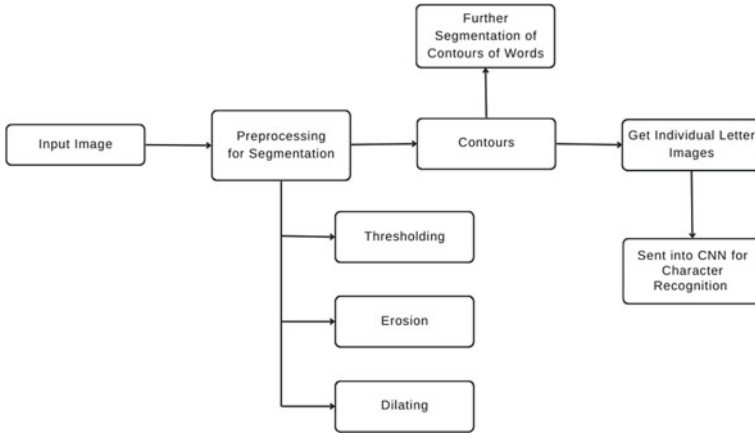
Fig. 2 Flowchart that depicts segmentation process
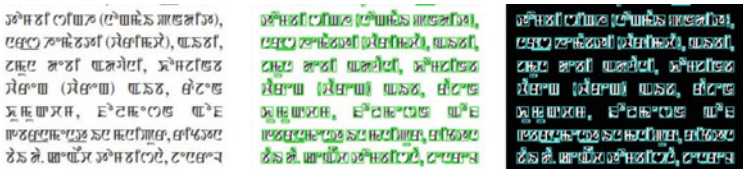
Fig. 3 Figure showing contours before the segmentation process

3.1 Segmentation

All the collected images (collected images are scanned copies of newspaper or other articles) are segmented using a segmentation process which includes a preprocessing step that is thresholding, simplifies the visual data for analysis which convert each pixel to black or white based on a threshold value. When an image is dilated, the objects in the image expand. The borders of items in a picture are increased in size by the amount of pixels. It is under the structural component's control. A matrix of 1's and 0's serves as the structural component. The two processes are very similar. The difference is because the pixel value was determined using the lowest dilation rather than the maximum. The determined minimum pixel replaces the picture underneath the anchor point. The deeper shaded areas expand, contrary to expectations. As it gets lighter or less white, it gets darker (Fig. 2). After this preprocessing step the characters are extracted using contours.

The characters have been bounded with a rectangle using a contour concept which is nothing but identifying a continuous object and bounding it, in this case the objects are characters. These contours are stored as separate images into a categorized folder. In this case, the contours are characters; these character images be identified using a trained CNN which can be used to recognize characters (Fig. 3).
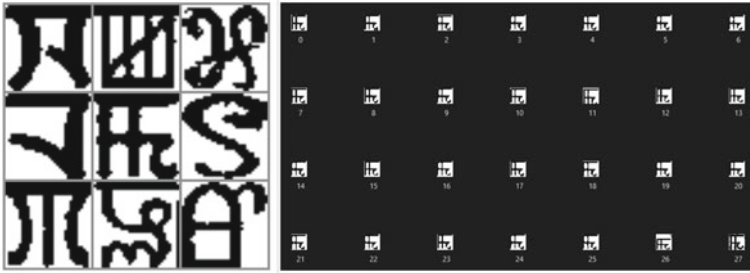
Fig. 4 Some letters from database after segmentation (left image), a part of database of a character (right image)

4 Database Creation

In the process of database creation, a folder for each character is created (to sort images of characters systematically, which is done manually). Now the segmented images are segregated into different folders where each segmented image, i.e., each letter is placed into its corresponding folder which is used for training of the CNN (Fig. 4).

Some images from the segmented images folder can have words/multiple letters which are cropped manually and the further cropped images (final letters) are placed into their corresponding letter folder.Now all the images are converted into black and white images with three channels and then resized to 32×32 sized images which will be used as the database for training, validation and testing.

Also added character images to the database from [16] to the database and used in training, validation, and testing.

5 Classification Using Transfer Learning—Training and Testing

The data or the images have been read from the folders of each character and stored into a numpy array, while reading the data the labels are also assigned to the images into another numpy array. After storing images into a numpy array, the images and labels have been shuffled without changing the one-one relation between the images and labels.

Now all the images in the numpy array undergo pre-processing (pre-processing for each model is different) which will convert the image data into some format where this final data can be sent into the model for training. While reading images, the test set is also created and undergoes the same process as training dataset.

5.1 Training Process

In the training process, a pre-trained model is trained with the training dataset, and it is tested with the test dataset. The preprocessed training dataset is now sent into the modified pre-trained model which is then trained with some freeze and unfreeze layers. The freeze layers are where the feature extraction is done, and the unfreeze layers are the last fully connected layers, where the weights get updated and the images are classified. After the training of the model, the trained model is used to predict the images in the test dataset to verify the models accuracy. The used models for the Meetei-Mayek character classification are VGG-16, VGG-19, and ResNet152-V2. Using multiple models gives the option to opt out of a best model based on the needs.

5.2 Testing Process

Once the model is trained, the test dataset is used to assess its performance on the specific task. This will give an idea of how well the model can generalize to unseen data. As it is important to evaluate the model on the test set because it gives an unbiased estimate of the model performance on new data. The trained model is used to predict each image in the test dataset which classifies the image into a character and gives the predicted character.

5.3 Images to Text

Recognizing text from the images is now a simple task using the above-trained model and the segmentation process. The given image is segmented into letters and classified using the trained model which can be converted into text using the labels.

6 Experimental Results:

The models that have been used for transfer learning are VGG-16, VGG-19, and ResNet152-V2.The section presents the benchmark scores of the respective models that are trained and tested.The total size of the database of the characters is 58482 images (which includes the database from [16]), in which 58372 images are used for training (of which 10% is used for validation) and 110 images are used for testing (Fig. 5 and Table 1).
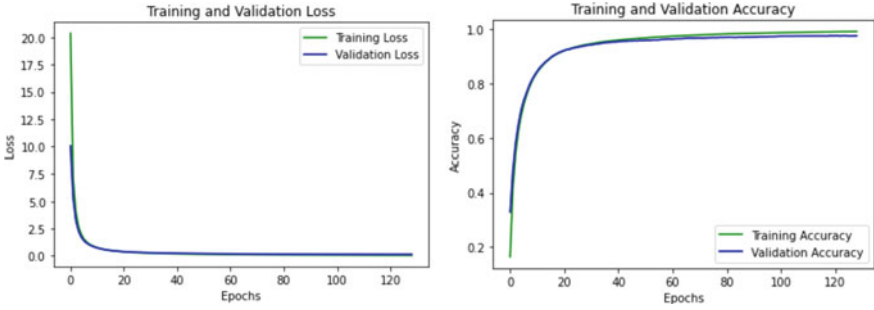
Fig. 5 VGG-16, left graph depicts training and validation loss, right graph depicts training and validation accuracy
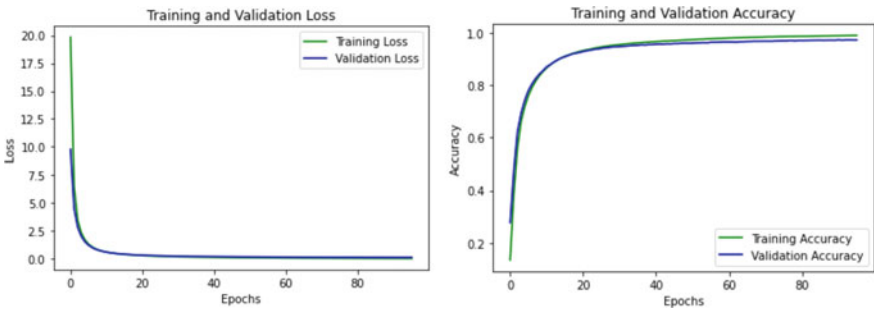
Fig. 6 VGG-19, left graph depicts training and validation loss, right graph depicts training and validation accuracy
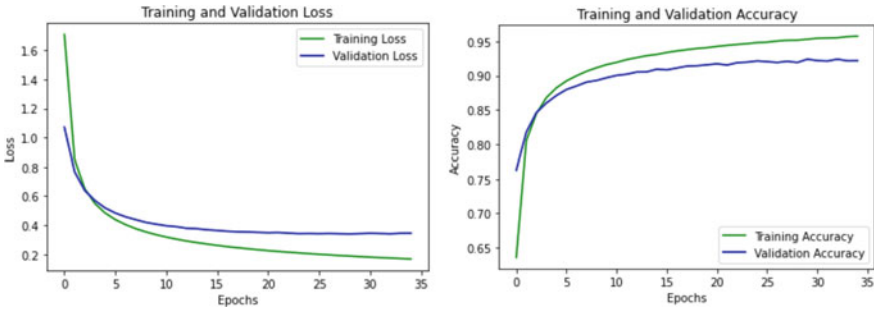
Fig. 7 ResNet152-V2, left graph depicts training and validation loss, right graph depicts training and validation accuracy

Table 1 Table with information of training accuracy, validation accuracy, testing accuracy, Top-1 & Top-5 accuracy for test dataset

Model name	Accuracy (%)	Validation accuracy (%)	Test top 1 accuracy (%)	Test top 5 accuracy (%)
VGG-16 [14]	99.27	97.69	85.45	94.54
VGG-19 [14]	98.91	97.16	83.63	94.54
ResNet152-V2 [15]	95.75	92.17	79.09	91.81

Figures 6, 7 and 8 are the graphs of VGG-16, VGG-19 and ResNet152-V2 respectively. Here, left graph of each fig represents training and validation loss, and right graph represents training and validation accuracy.

From the above table and graphs, we can conclude that among the above three models (VGG-16, VGG-19, and ResNet152-V2) VGG-16 has given good results with accuracy of 99.27%.

7   Conclusion and Future Work

Text recognition is a difficult task, and it is more difficult in a diversified country like India which has many different languages and scripts. The most frequently used languages in any part of India are mother tongue, Hindi, and English. So far, there are many tools that can extract text from major scripts like Hindi and English. But the extraction of text from the images of local languages is a problem. This paper has presented a printed Meetei-Mayek text/script recognition system that can recognize images from newspapers, books, etc., (where there are Meetei-Mayek text and nothing else than text) using VGG-16, VGG-19, and ResNet152-V2 which are well known for use in classification of images.

This work can be further developed in many ways; following are some ideas: It can be developed to recognize text from images containing text and non-text objects, and other popular models can be trained and checked for better recognition accuracy.

References

1. Rahman MM, Akhand M, Islam S, Shill PC, Rahman M et al (2015) Bangla handwritten character recognition using convolutional neural network. Int J Image Graph Signal Process (IJIGSP) 7(8):42–49
2. Devi CN, Das D, Devi HM (2020) A meetei mayek handwritten digit recognition system. J Comput Theor Nanosci 17(1):334–339
3. Thokchom T, Bansal P, Vig R, Bawa S (2010) Recognition of handwritten character of Manipuri script. J Comput 5(10):1570–1574

4. Inunganbi SC, Choudhary P (2018) Recognition of handwritten Meitei Mayek and English alphabets using combination of spatial features. In: International conference on intelligent systems design and applications. Springer, pp 1133–1142

5. Meetei LS, Singh TD, Bandyopadhyay S (2019) Extraction and identification of Manipuri and Mizo texts from scene and document images. In: International conference on pattern recognition and machine intelligence. Springer, pp 405–414

6. Garris MD, Blue JL, Candela GT, Dimmick DL, Geist J, Grother PJ, Janet SA, Wilson CL (1995) Off-line handwriting recognition from forms. In: 1995 IEEE international conference on systems, man and cybernetics. Intelligent systems for the 21st century. vol 3. IEEE, pp 2783-2788

7. Bunke H, Wang PS (1997) Handbook of character recognition and document image analysis. World Scientific

8. Nongmeikapam K, Kumar W, Singh MP (2017) Exploring an efficient handwritten Manipuri Meetei-Mayek character recognition using gradient feature extractor and cosine distance based multiclass k-nearest neighbor classifier. In: Proceedings of the 14th international conference on natural language processing (ICON-2017). pp 328-337

9. Blais A, Mertz D (2001) An introduction to neural networks-pattern learning with back propagation algorithm. Gnosis Software, Inc

10. Fukushima K (1988) Neocognitron: a hierarchical neural network capable of visual pattern recognition. Neural Netw 1(2):119–130

11. Ammina S (2019) Transfer learning using vgg-16 with deep convolutional neural network for classifying images. Int J Sci Res Publ (IJSRP) 9(10):143–150

12. Wen L, Li X, Li X, Gao L (2019) A new transfer learning based on vgg-19 network for fault diagnosis. In: 2019 IEEE 23rd international conference on computer supported cooperative work in design (CSCWD). IEEE, pp 205–209

13. Ferreira CA, Melo T, Sousa P, Meyer MI, Shakibapour E, Costa P, Campilho A (2018) Classification of breast cancer histology images through transfer learning using a pre-trained inception resnet v2. In: International conference image analysis and recognition. Springer, pp 763–770

14. Simonyan K, Zisserman A (2014) Very deep convolutional networks for large-scale image recognition. https://doi.org/10.48550/ARXIV.1409.1556

15. He K, Zhang X, Ren S, Sun J (2015) Deep residual learning for image recognition. https://doi.org/10.48550/ARXIV.1512.03385

16. Khuman YLK, Devi SD, Singh CP, Devi HM, Singh NA (2022) A benchmark dataset for printed Meitei/Meetei script character recognition. Data in Brief 45

An Efficient Intrusion Detection System Using Feature Selection and Long Short-Term Memory (LSTM)

Hidangmayum Satyajeet Sharma and Khundrakpam Johnson Singh

1 Introduction

Due to the increased reliance of various sectors such as governmental, medical, military, education, and industrial organizations on the Internet for daily operations, cyber threats are continually evolving and also becoming a serious problem in the last few years. One of the essential and crucial components aimed at enhancing computer system security is the intrusion detection system (IDS) [1]. The objective of an IDS is to monitor, evaluate, and classify network data as it moves throughout the network. Over the last few years, a wide range of models based on machine learning were created in order to assess and categorize the network structure. From the various machine learning models evaluated, deep learning models have experienced tremendous success in detecting and classifying the network attacks. In this work, we create an IDS model using longest short-term mean (LSTM) method which is a part of the deep learning algorithm [2]. To verify the efficacy of our proposed model, we implement it on a recent dataset: CICDDoS2019. Since the dataset has a very large number of features, we extract some important features using feature selection [3]. The detection accuracy is one of the main parameters used to measure the performance while classifying the attacks. In this work, the main aim is to test and enhanced the accuracy using hyperparameter optimization and to compare with other relative works.

The main objectives of this work are

- Building an IDS using deep learning algorithm.
- Training the proposed LSTM method using hyperparameter optimization.
- Evaluating the accuracy of the model.

H. S. Sharma (✉) · K. J. Singh

Department of Computer Science and Engineering, National Institute of Technology Manipur, Imphal West, Manipur 795004, India
e-mail: satya4hidang@gmail.com

- Comparison of the suggested model with relevant previous works.

In the sections that follow, we'll discuss about the following: Sect. 2 will discuss about the related works. Section 3 will give a brief description of the CIC DDoS 2019 dataset and steps of preprocessing the data. Section 4 will discuss about the proposed methodology and implementations of this work. Section 5 discussed the results discovered after execution. The study's conclusion is presented in Sect. 6.

2   Literature Review

Althubiti et al. [4] proposed a deep learning model using LSTM to implement the IDS. The model was tested using CICIDS-001 dataset. To increase the training and test accuracies, the hyperparameters were tuned. The final accuracy after testing the model was 84.83%.

Xiao et al. [5] proposed dimensionality reduction method using principal component analysis (PCA) and autoencoder (AE). The reduced set of features are implemented using convolution neural network (CNN) model. The implementation was carried out using KDDCup99 dataset, and the highest accuracy was obtained while using PCA with CNN with an accuracy of 94%. On comparing with DNN and RNN, the result shows that the CNN-IDS model has faster training speed than the other two models.

Cil et al. [6] proposed a deep learning model using feedforward deep neural network model to detect the DDoS attack types in CICDDoS2019 dataset. The DNN models contain 3 hidden layers with 50 neurons in each layer. The dataset was divided into two different types, one for detecting the normal traffics and the second dataset for classifying the attack types. The final accuracy of the proposed FFDNN model on the two datasets was 99.97 and 94.57%, respectively.

Jamadar [7] proposed an IDS using decision tree classifier and recursive feature elimination (RFE). The implementation was carried out using CIC IDS 2017 dataset. Since the dataset contains a large number of features (83 features), a subset of 13 features was selected using RFE. The model has a 99% accuracy rate in classifying the different attack types.

Sharma and Yadav [8] proposed an ensemble model consisting of support vector machine (SVM), random forest (RF), and decision tree (DT). The RFE technique was applied in order to remove the redundant features in the dataset. The RFE technique was applied on all the three machine learning model to identify the best result. The model was implemented on KDD Cup 99 dataset. The model could achieve an accuracy of 99.99%.

Bhati et al. [9] proposed an ensemble method based on discriminant classifiers using a Random Subspace algorithm to detect the attack class present in KDD Cup 99 dataset. A fivefold cross-validation is performed to train the model in each fold. After testing the model, the system takes a decision to check if the data is benign or attack class. The model could achieve an overall accuracy of 98.90%.

3 Data Preparation

The steps of analyzing and preprocessing the data are given in this section.

3.1 CIC DDoS 2019 Dataset

The CICDDoS2019 [10] dataset was created for the purpose of detecting and classifying the different types of DDoS attacks [11]. It contains total 86 features. For our experiment, we used 8 different attack types recorded from the first day: benign, LDAP, Portmap, NetBIOS, Syn, MSSQL, UDP Lag, UDP.

3.2 Preprocessing

The CICDDoS2019 dataset contains a very large amount of data sample. In order to make it easier for our experimental analysis, we extract 2% of the data from each attack classes. The dataset was decreased by randomly selecting rows in so as to ensure that the selection was random. The 'infinity' values present in all the rows are replaced with 'NaN' values. After that, all the 'NaN' values were deleted. The categorical data present in the sample like 'Flow ID', 'Source IP', 'Timestamp', 'Flow ID' are converted to numerical data using the 'Label Encoder'. Additionally, the 'SimilarHTTP' characteristic containing only value of '0' was eliminated from the dataset. The final dataset used for implementation of our model contains 488,882 rows and total 85 features.

3.3 Feature Selection

Feature selection is an important stage during the data processing stage where the irrelevant, as well as redundant features, are removed in order to reduce the training time and increase the accuracy [12]. Moreover, the feature selection process reduces the overfitting problem that can affect the accuracy of the model. In this work, an ensemble learning approach based on the aggregate of multiple decision trees, also known as extra tree classifier, is used for selecting features. Algorithm 1 gives the steps of the extra tree classifier algorithm.

Algorithm 1: Extra Tree Classifier

BEGIN

    Input: Training set, S with k samples where each sample contains f number
                of features, i number of features in each node and n number of trees
                in the ensemble and k_e as the minimum numbers of values to split.

  Output: Ensemble of Trees, $T = t_1, t_2, \ldots, t_n$

  Step 1.   Function Extra_Tree (S, i, n, k_e)

  Step 2.   If $|S| \leq k_e$, then

  Step 3.   return frequency of each class

  Step 4.   else

  Step 5.   Select i attributes randomly without replacing the attributes in
                    sample S

  Step 6.   Generate i splits, such that each split is in the range of attribute
                    in S.

  Step 7.   Select s such that Score(s , S) = $\max_{i \in \{1 \ldots k\}}$ Score(s_i , S), where
                    the score function gives a value indicating the split
                    performance.

  Step 8.   From s, split two sets S_l and S_m

  Step 9.   For the split sets, let t_l = Extra_Tree(S_l) and t_m = Extra_Tree(S_m)

  Step 10.   Create a node, P with split s and attach t_l and t_m as left and right
                    nodes.

  Step 11.   Return P

END

The information gain value determines the metrics of the importance of a feature. The information gain is calculated using the entropy value from each class. The formula used to calculate the entropy is given in Eq. (1).

$$\text{Entropy} = -\sum_{k=1}^{C} P_k \log_2 P_k \tag{1}$$

where P_k is the probability of the randomly selected class. Using the entropy value of all the class, the information gain can be calculated using Eq. (2).

$$\text{Gain}(G) = \text{Initial Entropy} - \text{Final Entropy} \tag{2}$$

25 features were selected based on the information gain as shown in Fig. 1. Moreover, for training the model on the 25 selected attributes, feature standardization was performed using 'StandardScalar' method. It is a feature scaling technique that uses mean and standard deviation as the starting points to locate specified feature values before rescaling data values to match the distribution between 0 and 1.
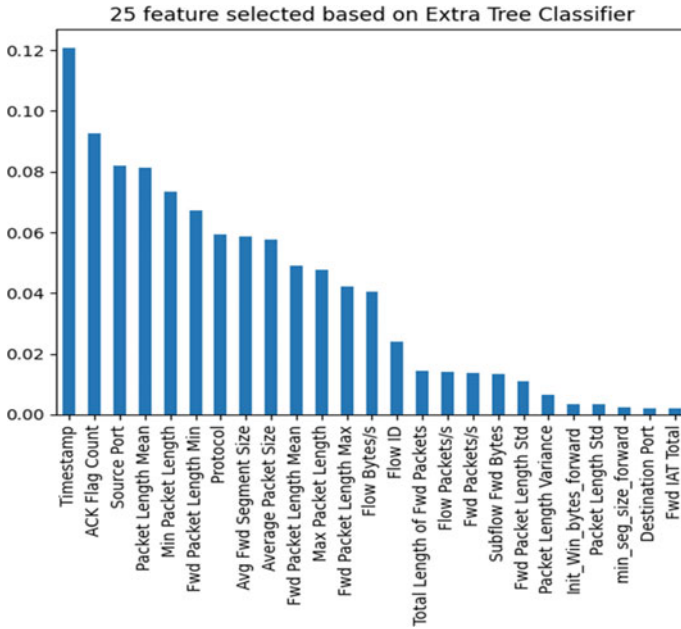
Fig. 1 25 features selected based on feature importance score

4 Proposed Model

The model is built using the LSTM method which is a part of the recurrent neural network (RNN) model with feedback connections. When compared to a typical RNN network, an LSTM network can significantly lessen gradient vanishing and explosion and improve the model's capacity for long-term memory. The basic architecture of an LSTM is given in Fig. 2. An LSTM cell has three distinct gates that regulate the flow of information: the input, output, and forget gates. The input gate selects the data using the sigmoid functions from the present state, and the activation function proposes a new vector that needs to be added. The output gate determines the input of the next hidden gate by passing through an activation function. The forget gate retains the essential information retrieved from the previous state. In the Fig. 2, C_t denotes the cell state, σ denotes the sigmoid function, tanh is the activation function, and H_t is the output.

The hyperparameter optimization process helps to achieve the optimum accuracy while training the deep learning model [13]. The LSTM model used for implementation in our work consists of three hidden layers, an input layer, and output layer. The 'ReLU' activation function is used for the input layer and three hidden layers. 'Softmax' activator is used for the output layer. The three hidden layers consist of 120, 100, and 60 neurons in each layer, respectively. Since the classification type is multiclass, 'categorical crossentropy' is used for the loss function, and 'adam'
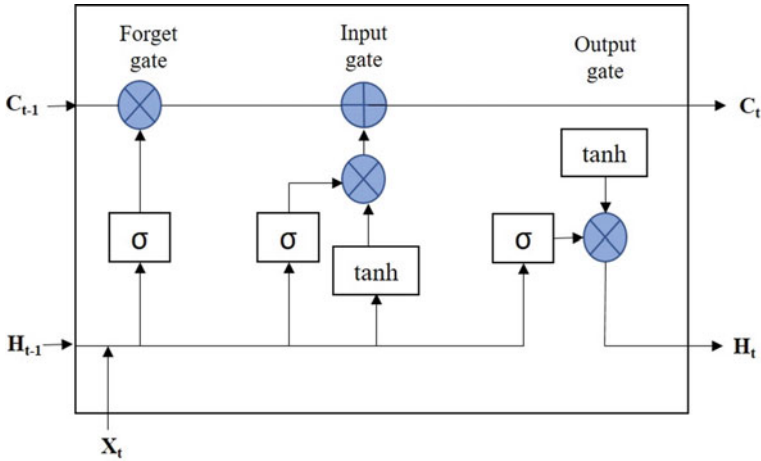
Fig. 2 Architecture of the LSTM model used in the proposed model

optimizer is used as the optimization algorithm for the proposed system. The batch size of the model is kept at 100 and dropout value as 0.02.

The resultant model has been applied using the CICDDoS2019 dataset, and results obtained are discussed in Sect. 5.

5 Experimental Results

In this section, the setup used for the experiment, performance metrics, and results is discussed.

5.1 Experiment Setup

The experiment was carried out on a system which includes Windows 10 (64-bit) operating system, Intel(R) Xeon(R) CPU E5-2687W v3 @ 3.10 GHz 3.10 GHz, and 64.0 GB RAM. The model was prepared using Python 3.7 programming language. To integrate the deep learning and machine learning models, tools like Keras-TesorFlow and Scikit-Learn are used.

5.2 Performance Metrics

In this section, the various performance measure such as accuracy, precision, F1-score, and recall is considered. The equation of the performance metrics is calculated based on true positive (T^+), true negative (T^-), false positive (F^+), and false negative (F^-). Accuracy (Acu): It is the proportion of all outcomes that were accurately predicted to all other outcomes, as stated in Eq. (3).

$$Acu = \frac{T^+ + T^-}{T^+ + T^- + F^+ + F^-} \tag{3}$$

Precision (Pre): It is the ratio of the actual normal attribute to the total normal and abnormal attributes detected correctly as given in Eq. (4).

$$Pre = \frac{T^+}{T^+ + F^+} \tag{4}$$

Recall (Re): It is the ratio of the normal attribute to the total number of attributes in the dataset as provided in Eq. (5).

$$Re = \frac{T^+}{T^+ + F^-} \tag{5}$$

F-Score (FS): It can be denoted as the harmonic mean of the precision and recall measure as given in Eq. (6).

$$FS = \frac{2\,Pre}{Pre + Re} \tag{6}$$

5.3 Results

The experiment was performed on the CICDDoS2019 dataset with 8 different attack classes using 25 features. The LSTM model can achieve an overall accuracy of 99.71% on the training set while using 50 epochs as depicted in Fig. 3. Whereas the testing set gives an accuracy of 99.68% which is very close to the training accuracy. The performance measure like precision, recall, and F-score of each attack types is shown in Table 1. Moreover, the overall value of the precision, recall, and F-score was 0.97, 0.88, and 0.88%, respectively.

The final loss percentage while training the model was very close to zero as shown in Fig. 4. Table 2 illustrates a comparison of the suggested model with other similar models.
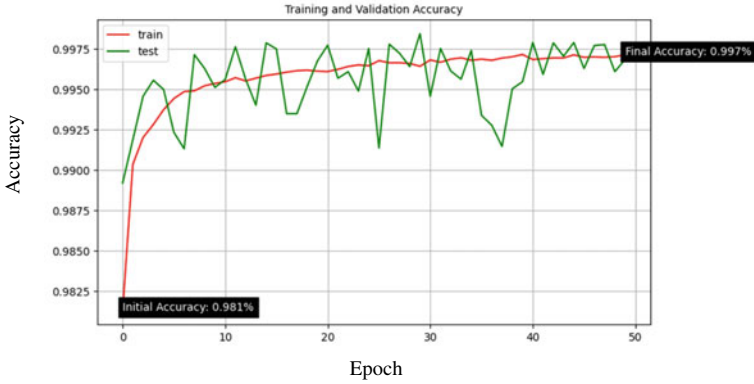
Fig. 3 Accuracy versus epoch for both the training and testing set

Table 1 Percentage measure of each attack class

Attack class	Pre	Re	Fs
Benign	0.99	0.98	0.98
LDAP	1.00	1.00	1.00
MSSQL	1.00	1.00	1.00
NetBIOS	1.00	0.99	0.99
Portmap	0.79	0.96	0.87
Syn	1.00	1.00	1.00
UDP	1.00	1.00	1.00
UDP Lag	1.00	0.09	0.17
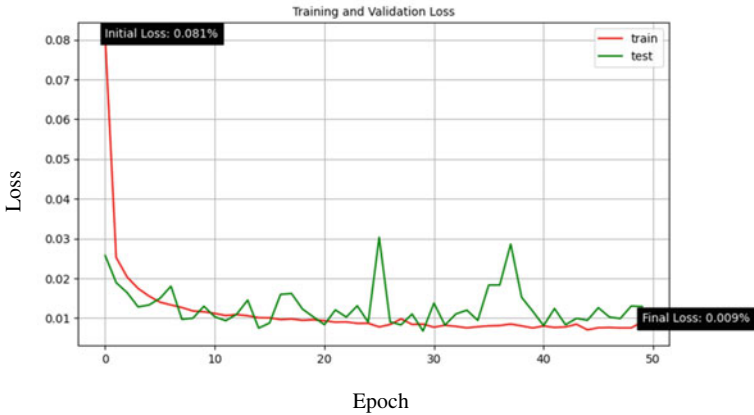
Fig. 4 Loss versus epoch for both the training and testing set

Table 2 Comparison of the proposed work with prior models

Author	Classifier	Dataset	Accuracy (%)
Althubiti et al. [4]	LSTM	CIC IDS-001	84.83
Cil et al. [6]	DNN	CICDDoS2019	94.57
Jamadar [7]	DT	CIC IDS 2017	99
Ferrag et al. [14]	LSTM	CIC IDS 2018	96
Proposed model	LSTM	CICDDoS2019	99.68

The experimental findings indicate that the suggested model exhibits the greatest overall performance in various parameters like accuracy, precision, F-score, and recall values while comparing with other almost related models. Also, the graphs depicted in Figs. 3 and 4 show that the training curve elevates uniformly with each epoch and loss curve descends uniformly with every epoch.

6 Conclusion

In this work, the LSTM model which is a part of the deep learning model is implemented using a more recent CICDDoS2019 dataset. Since the dataset is very huge, we use only a small fraction of the whole dataset. The experimental result shows an effective accuracy with very low value of loss percentage. The feature selection method using the extra tree classifier makes the model more effective while detecting the attacks. The 25 selected features are classified using the LSTM classifier. The detection accuracy of the classifier shows that the LSTM model can detect the attack types accurately, although a fractional amount of the dataset is used for the implementation.

In future studies, the similar model can be used to investigate using other more complex datasets in addition to other neural network models like DNN and CNN. Also, we will apply other feature selection method based on metaheuristic models in order to check their effectiveness of the IDS.

References

1. Ring M, Wunderlich S, Scheuring D, Landes D, Hotho A (2019) A survey of network-based intrusion detection data sets. Comput Secur 86:147–167
2. Laghrissi F, Douzi S, Douzi K, Hssina B (2021) Intrusion detection systems using long short-term memory (LSTM). J Big Data 8(65):1–16
3. Kasongo SM, Sun Y (2020) A deep learning method with wrapper based feature extraction for wireless intrusion detection system. Comput Secur 92:101752
4. Althubiti SA, Jones EM, Roy K (2018) LSTM for anomaly-based network intrusion detection. In: 2018 28th International telecommunication networks and applications conference (ITNAC). IEEE, pp 1–3

5. Xiao Y, Xing C, Zhang T, Zhao Z (2019) An intrusion detection model based on feature reduction and convolutional neural networks. IEEE Access 7:42210–42219
6. Cil AE, Yildiz K, Buldu A (2021) Detection of DDoS attacks with feed forward based deep neural network model. Expert Syst Appl 169:114520
7. Jamadar RA (2018) Network intrusion detection system using machine learning. Indian J Sci Technol 11(48):1–6
8. Sharma NV, Yadav NS (2021) An optimal intrusion detection system using recursive feature elimination and ensemble of classifiers. Microprocess Microsyst 85:104293
9. Bhati BS, Rai CS, Balamurugan B, Al-Turjman F (2020) An intrusion detection scheme based on the ensemble of discriminant classifiers. Comput Electr Eng 86:106742
10. Sharafaldin I, Lashkari AH, Hakak S, Ghorbani AA (2019) Developing realistic distributed denial of service (DDoS) attack dataset and taxonomy. In: 2019 International carnahan conference on security technology (ICCST). IEEE, pp 1–8
11. Singh KJ, Chanu US, De T (2017) Analysis of application layer DDoS attack detection parameters using statistical classifiers. Internetworking Indonesia J 9(2):23–31
12. Zhou Y, Cheng G, Jiang S, Dai M (2020) Building an efficient intrusion detection system based on feature selection and ensemble classifier. Comput Netw 174:107247
13. Kunang YN, Nurmaini S, Stiawan D, Suprapto BY (2021) Attack classification of an intrusion detection system using deep learning and hyperparameter optimization. J Inf Secur Appl 58:102804
14. Ferrag MA, Maglaras L, Moschoyiannis S, Janicke H (2020)Deep learning for cyber security intrusion detection: approaches, datasets, and comparative study. J Inf Secur Appl 50:102419

A Review on Speech Biomarkers for Obstructive Sleep Apnea(OSA)

Himanshu Sharma and Pradip K. Das

1 Introduction

Obstructive sleep apnea(OSA) is considered as a common sleep disorder designated by a decrease(hypopnea) or absolute stop(apnea) in the airflow, despite constant breathing efforts and this happens when pharyngeal muscles relaxes [1]. It is a serious sleep disorder affecting numerous people round the globe. Often multiple cases remain undiagnosed. Mulitple treatments have been developed for treating and addressing the disease where the diagnosis can be done either at home or in a clinic. A common symptom includes loud snoring many a times, while taking pauses in between as the airflow drops inside the pharyngeal path [2]. There is a tendency of breathlessness, insomnia, gasping sensation or frequent awakenings, etc. Elderly people have high chances of getting affected from the disease rather than middle-aged person with an estimation of around 28–67% in case of men and 20–54% in case of women [3]. There are majorly three types of OSA namely mild OSA, moderate OSA, and severe OSA. This review addresses the material and methods used to diagnose OSA. It is investigated that any change in the functional part of vocal tract unit affects speech acoustically [4]. It has been found that acoustic features of subjects with OSA differ from those of non-OSA subjects [5]. A frequent study [5] claimed that some people with OSA syndrome experiences non-intelligible speech.

H. Sharma (✉) · P. K. Das
Indian Institute of Technology Guwahati,Guwahati, Assam781039, India
e-mail: h.sharma@iitg.ac.in

P. K. Das
e-mail: pkdas@iitg.ac.in

2   OSA's Effect on Speech and Background Study

OSA affects more than 60% of adults, and its significant diagnostic marker is snoring [6]. The narrowing of the upper airway's cross-sectional area causes turbulent airflow, which in turn causes the upper airway's soft tissues to vibrate. In this paper, they developed a mathematical model [7] of snoring that measures airflow via an elastic tube. Based on the elasticity, flow, geometry and resistance, the tube can be made close and open repetitively, producing a sound resembling snoring. This model depicted that anatomical and other operational abnormalities in the patient's upper airway can lead to an OSA attack. If any structural change in the upper airway causes variability in an acoustic filter, then the acoustical features of snoring also vary.

In several studies, a review has been done on the relationship between snoring and OSA like by counting the snores of the patient [8], estimating the intensity of snoring sounds [9], and the analysis of jitter and pitch [10, 11]. In these studies, the methods mentioned don't justify the crucial aspects of acoustic features of the upper airway. Despite the carefully monitored sleep lab, the snore sounds recordings come out to be extremely corrupted from background acoustic noise and electromagnetic interference that, in turn, affects accuracy.

An established speech analysis technique known as Linear Predictive Coding (LPC) can be used to determine speech signal frequency models for the vocal tract. Formants are the identifiers for the upper airway's resonances, which are the implications for the maximal energy of resonant frequencies. The three formants f1, f2, and f3 mainly carry essential impulses for classifying snore signals as they relate to acoustic features of the upper airway. It has been revealed in the study that f1 and f2 relate to the constriction amount in the upper airway, while f3 refers to the length of the upper airway [25].

When diagnosing OSA, polysomnography(PSG) [12] is considered the gold standard. In general, a PSG includes electroencephalography (EEG), electromyography (EMG), electrooculography (EOG), electrocardiogram (ECG), movements of legs, upper airflow, thoracic movements, oxygen saturation of the patient and acoustic features recording of snoring signals. By exploiting the PSG approach, sleep apnea and hypopnea events have been calculated on per hour of sleep and an index (apnea-hypopnea index, AHI) is evaluated. OSA depicts if AHI>=5 events/hour. This process is very expensive and time consuming, hence so many patients remains undiagnosed from OSA.

It has been examined in the literature whether sustained vowel productions[1], read speech [13–18, 27] and yes/no questions and word lists [17, 18, 27] can be used as biomarkers for OSA. In this study [26], the automatic OSA detection has been performed on Portuguese speech and also from video blogs.
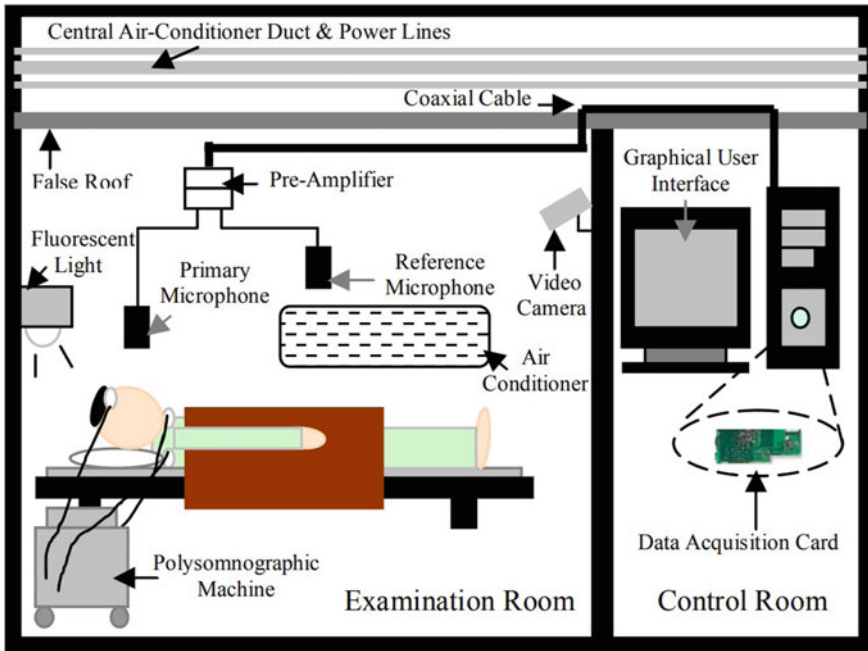
Fig. 1 Setup of snore acquisition system [19]

3 Materials and Methods

Several methods have been used for OSA detection. **Snore Acquisition System** in Fig. 1, [19] shows the Sleep Disorders Unit. Singapore General Hospital, built this high-precision snore acquisition system for sleep disorder detection. They used two non-invasive microphones with unidirectional condensers having 20-20000Hz, SM81 microphones. As snoring signals has an extensive range of dB, i.e., 90dB , hence snoring signals with low intensity can be easily masked by noise sources as there are many background acoustical noises possible like AC's sound, sounds of machines, equipment, etc., that majorly affect the signal-to-noise(SNR) ratio.

Normalized Least-Mean Square Adaptive filter: In this paper, they enhanced SNR by MNLMS adaptive filter design. SNR at the output has been increased by optimizing the adaptive filter to reduce output noise without distorting the signal of interest. In addition to all the advantages of conventional LMS and NLMS adaptive filters, MNLMS adaptive filters [20] may also have faster convergence and higher stability as shown in Fig. 2.

In another method [20], the acoustic characteristics of speech signals of awake subjects were compared with those of non-OSA (healthy) subjects. The vowels a,e,i,o,u were segmented and reviewed. Each segment of speech from each subject was framed in 30 ms intervals. A total of 51 acoustic features were extracted from each speech
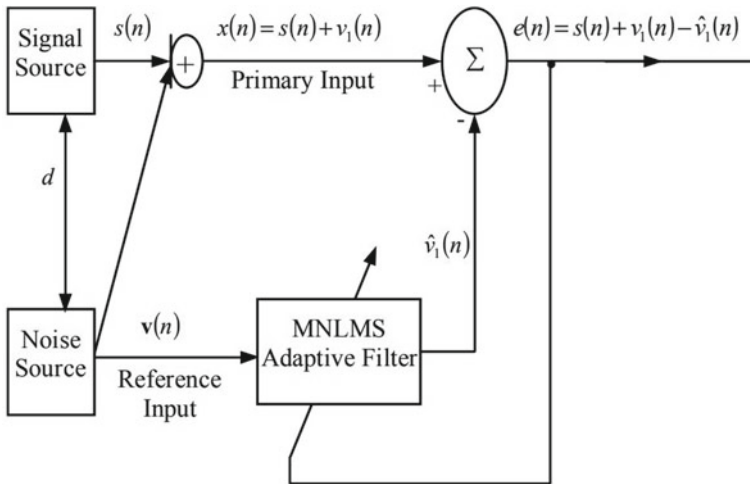
Fig. 2 Adaptive noise canceller [19]

frame, including spectral features such as formants, linear prediction coefficients (LPCs), cepstral features, Mel frequency cepstral coefficients (MFCCs), and their first derivatives and other prosodic features. On a subset of the twenty-dimension feature space, a Gaussian mixture model (GMM) classifier was implemented and selected based on a feature selection algorithm.

In another method [21], 131 male subjects (speakers) were analyzed for speech signals in order to create an automatic OSA severity estimation system. The signals were then preprocessed and features were extracted, as shown in Fig. 3. An apnea-hypopnea index(AHI) estimate from the regression model is computed, i.e., is the average of apnea and hypoapneic patient during a one-hour duration sleep. If the AHI value is less than five, that is considered as normal. In preprocessing they removed DC, performed normalization and signals of 30 msec frames were overlapped by 50% and removal of silence [22].

In another study [23], a total of 398 subjects (male and female) were taken for training purposes. The study included 305 patients who were referred to the Sleep-Wake Unit at Soroka University Medical Center for PSG and 93 Ben-Gurion University students who volunteered to participate using the WatchPAT 200 (Itamar Medical Ltd., Israel) [24].

3.1 Feature Extraction

Feature extraction is a crucial part of any study. In this review, we focus on the feature extraction of speech signals for OSA estimation. Several features have been used in the different studies like pitch frequency extracted from SIFT algorithm [25],
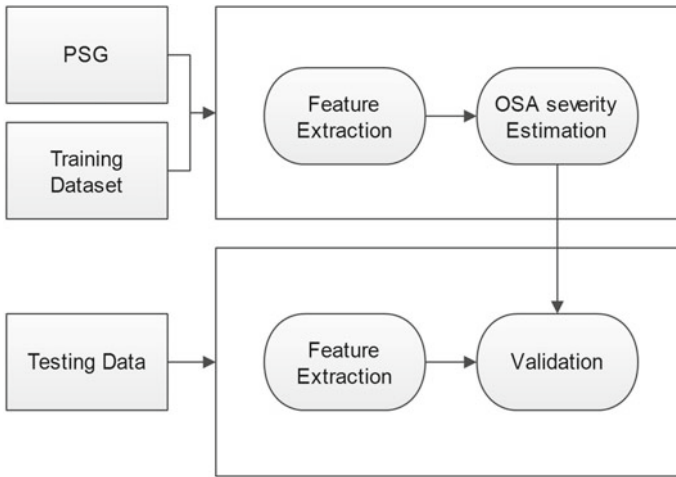
Fig. 3 Flow chart of PSG data analysis

LPC's evaluated via auto-correlation method [25]. As the human auditory system approximates MFCCs, they appear to have frequency bands that are positioned in logarithmic order (on the Mel scale) [25].

Knowledge based feature extraction Knowledge-based feature extraction is done in the study [26]. They have represented each audio file as a vector of 109 features. Out of these 25 features are common to eGeMAP [27] consists of the standard deviation(std) and mean of bandwidth and frequency of formant 1, 2, and 3: std and mean of jitter; also std and mean of Harmonics-to-noise ratio(HNR) [26]. The set has a mean of 12 MFCC with their first and second derivatives and 48 LPCC, resonance anomalies should be revealed by the formant frequencies, while phonation anomalies should be revealed by jitter and HNR. Soft tissue properties (pharyngeal mucosa and soft tissue) are expected to alter formant bandwidths.This feature set includes features related to spectral flux and F0-that have shown to be important for detecting sleepiness [26]. Based on their relationship with vocal tract shape, LPCC and MFCC were chosen.

Automatic feature extraction Two additional subsets were derived from the original feature set, referred to as OFS, With the corpus's reduced size, they tried to avoid high dimensionality problems.

– **Random Forest Feature selection(RF)**: To the RF ranking, here are the five most relevant features: the mean of $f'(MFCC)$ and $f''(MFCC)$. With mean, a percentile of 20.0 and a percentile of 50.0 for F0 are provided. In this technique, each component is ranked according to how much it decreases entropy.The reason for using it was that it is considered to be resistant to overfitting. However, the selection of features is clearly based on gender.

– **Mann-Whitney U test ranking (M-W)**: An analysis of Mann-Whitney U tests revealed 18 features with a lower p-value.
– **SVM**: with linear kernel enabled the best classification results with the number of features integrated into each feature subset in the Portuguese Sleep Disorders corpus. As a result, they could lead to overoptimistic results in that dataset. Hence, a normalization was done between 0 and 1 for the feature values.

4   Conclusion

In this paper, we have reviewed several techniques through which OSA can be detected. We acknowledged that there is already a gold standard for OSA detection, i.e., PSG. Various feature extraction techniques like knowledge-based and automatic feature extraction are discussed in detail. Knowledge-based extraction includes LPCC and MFCC as central feature extraction coefficients whereas automatic feature selection includes techniques like random forest, Mann-Whitney U test ranking, and SVM for feature extraction purposes. As a future scope, we can use a wearable device, i.e., smart watches to measure oxygen level during sleep as apnea pateints oxygen level goes down during sleep.

References

1. Arnold J, Sunilkumar M et al (2017). Obstructive sleep apnea. https://doi.org/10.4103/jpbs. JPBS_155_17
2. Sojot AJ, Meisami T, et al (2001) The epidemiology of mandibular fractures treated at the Toronto general hospital: a review of 246 cases, pp 640–645
3. Goodday RH (1997) Nasal respiration, nasal airway resistance, and obstructive sleep apnea syndrome, pp 167–77. https://doi.org/10.1016/S1042-3699(20)30989-4
4. Deller JR, Proakis JG et al (2000) Discrete-time processing of speech signals
5. Fox AW, Monoson PK et al (1989) Speech dysfunction of obstructive sleep apnea, 589–595. https://doi.org/10.1378/chest.96.3.589
6. Ohayon MM, Guilleminault C et al (1997) Snoring and breathing pauses during sleep: telephone interview survey of a United Kingdom population sample, 860–863. https://doi.org/10.1136/bmj.314.7084.860
7. Gavriely N, Jensen O (1993) Theory and measurement of snores, 2828–2837. https://doi.org/10.1152/jappl.1993.74.6.2828
8. Esnaola S, Duran J et al (1996) Diagnostic accuracy of a portable recording device (Mesam IV) in suspected obstructive sleep apnoea, 2597–2605. https://doi.org/10.1183/09031936.96.09122597
9. Brunt D, Lichstein KL et al (1997) Intensity pattern of snoring sounds as a predictor for sleep-disordered breathing, 1151–1156. https://doi.org/10.1093/sleep/20.12.1151
10. Sola-Soler J, Jane R et al (2000) Towards automatic pitch detection in snoring signals, 2974–2976. https://doi.org/10.1109/IEMBS.2000.901503
11. Abeyratne UR, Patabandi CKK et al (2001) Pitch-jitter analysis of snoring sounds for the diagnosis of sleep apnea, 2072–2075. https://doi.org/10.1109/IEMBS.2001.1020642

12. Sole-Casals J, Munteanu C et al (2014) Detection of severe obstructive sleep apnea through voice analysis, 346–354. https://doi.org/10.1016/j.asoc.2014.06.017
13. Pozo R, Murillo J et al (2009) Assessment of severe' apnoea through voice analysis, automatic speech, and speaker recognition techniques, 1687–6172 https://doi.org/10.1155/2009/982531
14. Kriboy M, Tarasiuk A et al (2014) A novel method for obstructive sleep apnea severity estimation using speech signals, 3606–3610. https://doi.org/10.1109/ICASSP.2014.6854273
15. Benavides A, Pozo R et al (2014) Analysis of voice features related to obstructive sleep apnea and their application in diagnosis support, 434–452. https://doi.org/10.1016/j.csl.2013.08.002
16. Espinoza-Cuadros F, Pozo R et al (2016) Reviewing the connection between speech and obstructive sleep apnea, 15–20. https://doi.org/10.1186/s12938-016-0138-5
17. Elisha O, Tarasiuk A et al (2012) Automatic detection of obstructive sleep apnea using speech signal analysis, 1373–1382. https://doi.org/10.1109/TBME.2010.2100096
18. Kriboy, Tarasiuk A et al (2014) Detection of obstructive sleep apnea in awake subjects by exploiting body posture effects on the speech signal, 4224–4227. https://doi.org/10.1109/EMBC.2014.6944556
19. Ng, Andrew K, Koh TS et al (2006) Speech-like analysis of snore signals for the detection of obstructive sleep apnea, 99–103
20. Zigel Y et al (2008) Analysis of speech signals among obstructive sleep apnea patients. 760–764. https://doi.org/10.1109/EEEI.2008.4736637
21. Kriboy M et al (2014) A novel method for obstructive sleep apnea severity estimation using speech signals, 3606–3610. https://doi.org/10.1109/ICASSP.2014.6854273
22. Sohn J, Kim NS et al (1999) A statistical model-based voice activity detection, 1–3. https://doi.org/10.1109/97.736233
23. Simply, Melody R et al (2019) Diagnosis of obstructive sleep apnea using speech signals from awake subjects, 251–260. https://doi.org/10.1109/JSTSP.2019.2955019
24. Ceylan T, Fırat H et al (2012) Quick diagnosis in obstructive sleep apnea syndrome: WatchPAT-200, 475–478
25. Deller JR, Proakis JG et al (1987) Discrete-time processing of speech signals
26. Botelho, Catarina M et al (2019) Speech as a biomarker for obstructive sleep apnea detection 5851–5855. https://doi.org/10.1109/ICASSP.2019.8682431
27. Goldshtein E, Tarasiuk A et al (2011) Automatic detection of obstructive sleep apnea using speech signals, 1373–1382. https://doi.org/10.1109/TBME.2010.2100096

Real-Time Object Detection for Unmanned Underwater Vehicles Using Movidius Neural Compute Stick

K. Amal Thomas, Soumyajit Poddar, Mourina Ghosh, and Amitava Nag

1 Introduction

Around 70% of the earth's surface is covered by the sea, while 97% of the earth's water is contained in the oceans. However, our understanding of the seas is limited, and fewer than 5 percent of the oceans have been investigated so far. Nevertheless, utilizing ships is costly and hence cannot be done on a bigger scale. In order to save costs, there is a growing trend toward the development and use of autonomous underwater vehicles [1, 2].

Counting objects is a hot area of study, and the underwater objects counting issue is difficult for several reasons. Underwater visibility, color suppression, and haze make detection particularly challenging. Objects within the field of view are frequently so obscure that neither humans nor deep models can identify them. Detection and counting on a low-powered autonomous underwater vehicle (AUV) impose considerable limits on the selection of the detection model and the detection process's frequency. Convolutional neural networks (CNNs) have been utilized in a wide variety of disciplines, including object detection, object identification [3], and object mapping. Specifically, in general purpose systems based on the employment of a microcontroller, the limited accessible memory restricts the network's complexity, which may affect the system's precision. Similarly, microcontroller-based systems have the poorest trade-off between energy usage and timing performance. Due to this, commercial hardware accelerators for CNNs such as Neural Compute Stick (NCS) and Google Coral have been developed [4]. These devices offer hardware designs

K. A. Thomas (✉) · S. Poddar · M. Ghosh
Indian Institute of Information Technology Guwahati, IIIT Guwahati, Bongora, Guwahati, Assam 781015, India
e-mail: amal.thomas@iiitg.ac.in

A. Nag
CSE Department, Central Institute of Technology Kokrajhar, Kokrajhar, Assam 783370, India

that are tuned to enable CNN model inferences with minimal latency and decreased power consumption.

Authors extending one of their patented works in this study [5]. The remainder of this particular work is organized in 4 sections. Section 2 discusses briefly about the state of the art. Section 3 describes the proposed model and Sect. 4 comprehensively portraits the experimental results. Section 5 gives the conclusion. This paper proposes an object tracking algorithm which is useful in an underwater vision scenario. Video stream which captured by the unmanned underwater vehicle areas are classified according on object population density. The areas are then adaptively dissected into an image patch tree. In addition, the surfaces are later efficiently saved and recovered as patches. As a result, the disclosed approach is an effective and adaptable method for generating a route tracking data structure of arbitrary surfaces with improved computability.

2   Related Works

The summary of related art consists of methodologies, experiments, surveillance datasets, baseline object recognition, and efforts to improve performance. Notably, the majority of prior research on landmark estimation use highly linked elastic networks that are challenging to optimize. Consequently, a great deal of work has been devoted to the development of optimization techniques for avoiding local minima. For instance, Zhu and Deva demonstrate that multi-view trees are an effective alternative because they can be globally optimized with dynamic programming and, surprisingly, they still capture a significant amount of globally elastic structure [6].

On the early 2000s Cai et al. introduced graph embedding methods. Which were primarily developed, and which turned into the development of graph neural networks in the upcoming year. Graph embedding method helps to minimize the high dimensionality of non-relational data [7]. This resulting in the development of a method that heavily relies on computer vision techniques and graph theory optimization. When it comes to unmanned underwater or surface vehicles, computer vision research started active in the field of object tracking and segmentation during the period of early years of last decade. After the integration of AI in this field, many developments happened including the uses of graph neural networks for various object tracking in the unstructured environments. Due to better computation speed and effective learning procedures, CNN-based trackers are incredibly popular [8, 9]. When it comes to computation in Edge platforms, a huge number of samples may be learned by CNN with a little computational resource usage and power consumption.

3 Methodology

An image from a video frame is recorded and divided into a plurality of areas depending on the density of faces of individuals or characteristic surfaces of target items inside each zone. Further, preferably, but optionally and also, adaptive segmentation is utilized to find and identify the surface locations of all items or the face positions of all individuals present in regions of changing surface or face density. In accordance with a further possible technique of the study, the granularity of segmentation adjusts over time to the numerical density of underwater objects.

Every subsequent frame and its associated graph are compared to identify position changes, and an object motion trail is created that may be saved locally on system server and regularly transferred to the cloud. According to the above mentioned block diagram, a frame is detected. The approach is depicted as a Block diagram/Schematic in Fig. 1. Underwater camera is aimed in a specific direction. The specifications include its field of view, focal length, height above ground, and viewing angle. From the camera, a single frame is extracted from the video stream. The frame's aspect ratio and area are fixed. Tracking of the surface is performed, and a set of graphs is generated for all surfaces visible in the frame. Before plotting the graph, created boundary box of the tracked object. The proposed embodiment detects each surface using a neural network-based classifier and constructs the center of each surface as a set of Cartesian coordinates. Later, horizontal and vertical tree diagrams constructed and labeled it with an object ID for seamless tracking. The route traversed by each surface is denoted as a series of places and consists of a set of provided frames. The aforementioned locations are determined using the UUV's preset specifications.

Fig. 1 Block diagram of the proposed method
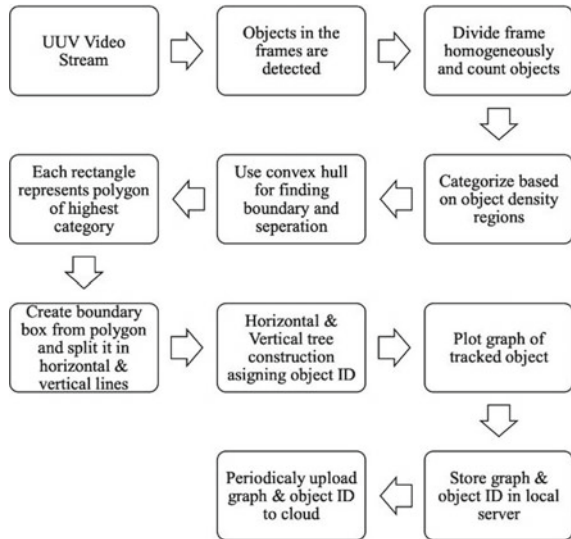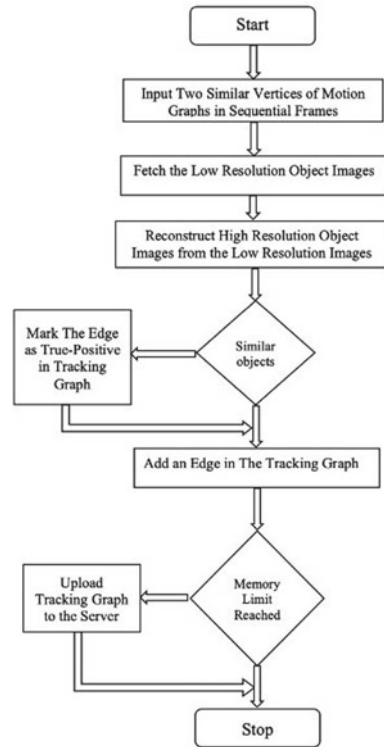
Figure 2 depicts the technique for constructing a tracking graph for camera-observed surfaces. The inputs of two comparable vertices of sequential motion graphs are generated using motion frame. Low resolution image of the object and pictures that correspond to the vertices are retrieved, and high-resolution object images are rebuilt from the low resolution photos. If comparable boundary boxes are found, images utilizing super resolution methods are produced. The edge is then shown as True-Positive on the tracking graph. If this is not the case, a new edge is added to the tracking graph depicted in the flowchart. Tracking graph is published to the server if the memory exceeded check-in returns are true. If not, then this iteration of the algorithm is complete.

4   Experimental Results and Discussion

This section presents some simulation results to demonstrate the effectiveness of the proposed tracking methods. The proposed model is elaborated using python and the efficiency of the proposed approach is validated by measuring F-measure (Fig. 3).

(a) (b) (c) (d)
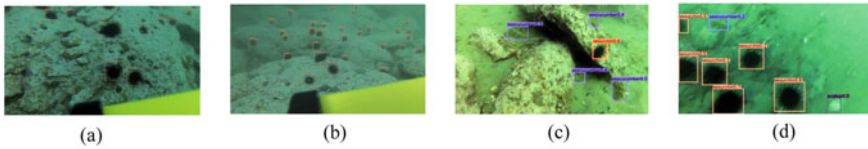
Fig. 3 Sematic object detection results of UDD dataset

The underwater open sea farm object detection dataset (UDD) is one of the first dataset which collected the real open sea farm for underwater robot picking, which is the 4K HD in quality [10]. UDD contains 3 different set of categories, which are sea cucumber, sea urchin, and scallop with 2,227 images. There are various other datasets related to underwater vision based. According to the Underwater Object Tracking Benchmark Dataset UOT100 and UOT32 also few among the other datasets which got feasible possibility among better object tracking demonstration underwater. To perform the inductive learning, we initially fix the bellow mentioned sets of hyper-parameters: $\alpha = 0.5$, $\lambda = 1.0$ and learning rate of 0.001. Due to the large amount of training materials, we decided to confirm the dropout rate to 0.3 as well as the weight decay to zero.

We compare the proposed model with the following methods: GraphSAGE [11], VR-GCN [12], GAT [13], GeniePath [14]. The metrics are consolidated in Fig. 4. We found out that larger predictive-power can be imposed with increasing network depth.

This particular study we set it up on the combined platform of PYNQ-Z2 board and Movidius NCS. Results are mentioned in Table 1. In concordance with our model expectations, proposed model achieved a much better performance while positioning NCS onto FPGA board. The proposed hardware system achieved an 84.37 ms runtime through utilizing NCS whereas the system alone obtains 49.37 ms without positioning NCS.
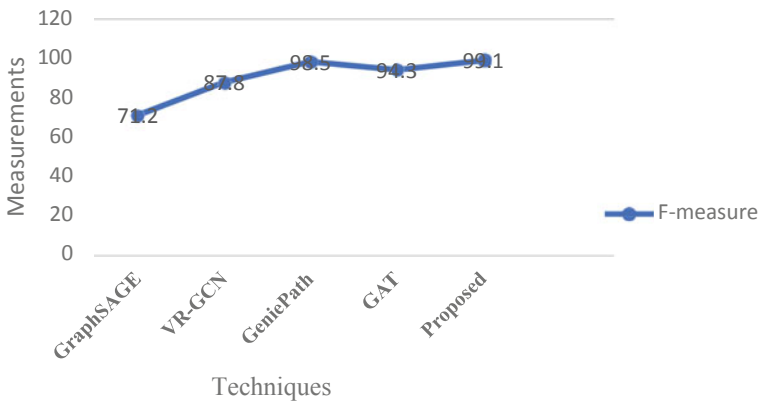
Fig. 4 Comparison of proposed model based on F-Measure

Table 1 Proposed model on NCS runtime analysis

Hardware platform	Runtime (ms)
Without-NCS	84.75
With-NCS	49.37

5   Conclusion

The unmanned underwater vehicle-based object tracking and image segmentation on GCNN architecture was modeled and adapted for UDD dataset. The proposed method of object tracking achieves better performances according to different assessment metrics. It was also observed that the object tracking performance result got improved with F-Measure values as 99.1 when runtime with NCS deployed at the rate of 49.37 ms.

We are planning to extend this graph-based object tracking mechanism in the various unstructured environments for better application permeability.

References

1. Langner F, Knauer C, Jans W, Ebert A (2009) Side scan sonar image resolution and automatic object detection, classification and identification. In: OCEANS 2009-EUROPE. IEEE, pp 1–8
2. Lee D, Kim G, Kim D, Myung H, Choi H-T (2012) Vision-based object detection and tracking for autonomous navigation of underwater robots. Ocean Eng 48:59–68
3. Ren S, He K, Girshick R, Sun J (2017) Faster R-CNN: towards real-time object detection with region proposal networks. IEEE Trans Pattern Anal Mach Intell 39(6):1137–1149
4. Antonini M, Vu TH, Min C, Montanari A, Mathur A, Kawsar F (2019) Resource characterisation of personal-scale sensing models on edge accelerators. In: Proceedings of the first international workshop on challenges in artificial intelligence and machine learning for internet of things, pp 49–55
5. Poddar S, Kakkassery AT, Naidu PCRJ (2022) Method for real time surface tracking in unstructured environments. In: U.S. Patent No.: US 11,315,257 B2
6. Zhu X, Ramanan D (2012) Face detection, pose estimation, and landmark localization in the wild. In: 2012 IEEE Conference on computer vision and pattern recognition, providence. IEEE, pp 2879–2886
7. Goyal P, Ferrara E (2018) Graph embedding techniques, applications, and performance: a survey. Knowl-Based Syst 151:78–94
8. Tijjani AS, Chemori A, Creuze V (2022) A survey on tracking control of unmanned underwater vehicles: experiments-based approach. Annu Rev Control 54:125–147
9. Lin C, Wang H, Fu M, Yuan J, Yu D (2019) A convolution neural network based obstacle avoiding method for unmanned underwater vehicle. ICIC Express Lett 13(11):1079–1086
10. Liu C et al (2022) A new dataset, Poisson GAN and AquaNet for underwater object grabbing. IEEE Trans Circuits Syst Video Technol 32(5):2831–2844
11. Hamilton WL, Ying R, Leskovee J (2017) Inductive representation learning on large graphs. In: 31st Conference on neural information processing systems (NIPS 2017), pp 1–11

12. Ye R, Li X, Fang Y, Zang H, Wang M (2019) A vectorized relational graph convolutional network for multi-relational network alignment. In: Proceedings of the twenty-eighth international joint conference on artificial intelligence (IJCAI-19), pp 4135–4141
13. Veličković P, Cucurull G, Casanova A, Romero A, Lio P, Bengio Y (2017) Graph attention networks. arXiv:1710.10903 [stat.ML], arXiv:1710.10903v3 [stat.ML], https://doi.org/10.48550/arXiv.1710.10903
14. Liu Z, Chen C, Li L, Zhou J, Li X, Song L, Qi Y (2019) Geniepath: graph neural networks with adaptive receptive paths. In: Proceedings of the AAAI conference on artificial intelligence, vol 33, issue no 1, pp 4424–4431

Dysarthric Speech Characterization and Classification Based on Affinity Propagation

Komal Bharti, Sandeep Agri, and Pradip K. Das

1 Introduction

In recent years automatic speech recognition (ASR) systems have been enhanced significantly with the help of Artificial Intelligence and Deep Learning Technologies but the state-of-the-art result keeps impaired speech far away. Since ASR systems operate hands-free, they can be used widely by disabled people. However, ASR systems are difficult to use by people who have speech disorders due to various reasons [1]. Cerebral palsy, which is a central nervous system disorder, is one of the causes of speech disorders. This paper is about dysarthric speech in which the speaker leads to slow movements of the muscles in the left hemisphere of the brain while speaking. They can't coordinate with their vocal cords and muscles' actions resulting in unclear and unstable utterances [2]. Dysarthric speakers can often access technology more efficiently and effectively using hands-free or speech-enabled interfaces rather than remote control, equipment switches and keywords. For those with dysarthria to be able to use ASR effectively, there is a serious need for a reliable system [3]. A particular condition that creates problems in the formation of speech sounds and great difficulties while communicating with others is termed as speech disorder. This may create a problem for others to understand their language. This condition affects a person's ability to speak fluently. There are several kinds of communication disorders broadly categorized into language disorders and speech disorders. People can have both speech and language disorders at a time [4, 15].

K. Bharti (✉) · S. Agri · P. K. Das
Indian Institute of Technology Guwahati, Guwahati, Assam 781039, India
e-mail: kbharti@iitg.ac.in

S. Agri
e-mail: sandeep.agri@iitg.ac.in

P. K. Das
e-mail: pkdas@iitg.ac.in

2 Literature Survey

The classification of vowels is one of the most important components of modeling spoken units since they are steady-state segments. Many studies have been conducted on the phonetic characteristics of vowels in different languages, as well as methods to characterize the spoken units, such as voice activity detection and keyword spotting. Assamese vowels were studied using formant frequencies in RNNs and KNNs by Sharma et al. [5]. In another approach, they also classify assamese fricatives using standard deviation, skewness, kurtosis, etc., [6]. Matthew and Yossi et al. [7] designed an algorithm that measures vowel duration using a forced aligner based on HMMs and Deep Neural Networks (DNNs).

Affinity Propagation was published by Frey and Dueck in 2007 [8]. In other clustering algorithms like k-means, an initial set of exemplars are taken, which should be close to a good solution in order to get the right clustering. On other hand, "affinity propagation" which takes as input measures of similarity matrix that consist of similarity values between pairs of data points. Wang et al. [9] used affinity propagation in Binaural sound localization to estimate the direction of sound source. They applied the clustering analysis to the similarity matrix based on affinity propagation. Affinity propagation is advantageous over common clustering algorithms such as hierarchical clustering and k-means. In affinity propagation, the number of the clusters does not need to be specified beforehand as done in k-means. It has been used in various fields including speech, such as speaker clustering [10], image processing [11], EEGs [12]. Bhagat et al. [13] used PLP coefficient as features for the classification of vowels and used a similar approach of pre-clustering through affinity propagation to select features. Smeulders and Nguyen [14] showed that pre-clustering is an effective way to improve classification by identifying a set of features that contribute to good classification.

3 Methodology

Pre-processing of raw data and feature selection is one of the most important keys for getting an accurate and fast trainable model. First, we take the raw speech signal and do some pre-processing followed by feature extraction using LPCCs, then from extracted features, we select the best features through affinity propagation by taking the best clustering score. After that, we classify the data and validate the selection of features with SVM. The proposed approach is shown in Fig. 1.

All steps are explained in the next section. The pre-processing, feature extraction, and affinity propagation implementation are done in Microsoft Visual Studio 2010 in C language. The pre-clustering analysis and SVM classification are done in Jupyter Notebook in python.
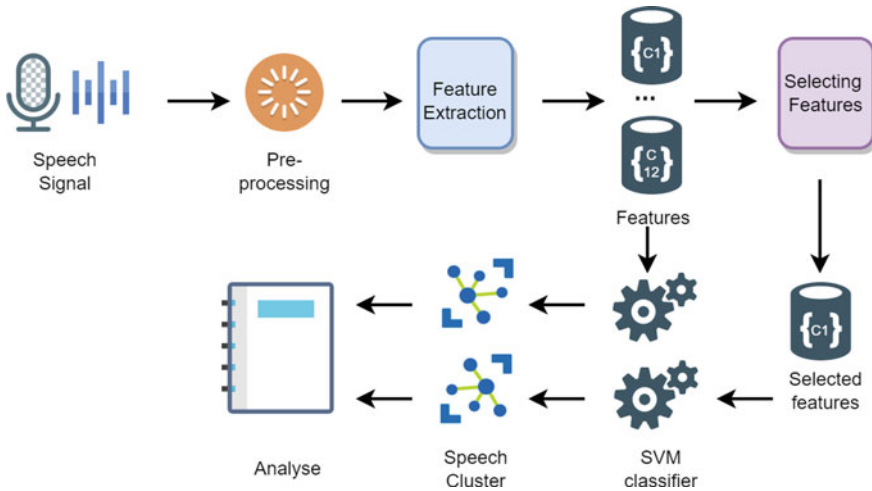
Fig. 1 Proposed approach

3.1 Pre-processing

We take vowel samples and digit samples from the UASpeech database [15] for all speakers with different severity levels. We correct the DC-shift of speech samples by taking the summation of silence part and averaging it by dividing it by the total samples, which will be the shift. Then we cut the stable frames, i.e., trimming the silence part from the beginning and end of required speech samples. The trimming is done by calculating the Short Term Energy(STE) of all frames of 320 samples each. when the STE increases by 400%, we mark it as the start of a stable frame. We do a similar process for marking the end point of vowel speech utterances. We normalize the speech samples in the range of +5000 to –5000, i.e., multiplying the samples with a factor such that maximum and minimum of sample becomes +5000 and –5000, respectively.

3.2 Feature Extraction

In first stage of the system after pre-processing, the audio signal is converted into some parametric values with distinctiveness which represent speech and speaker characteristics. The transformation of signal to parameters is called feature extraction and researchers have proposed many different methods for this like MFCC, LPCC, PLP, etc. In this approach, we have used LPCCs.

Framing of the sample is done on the pre-processed signal which is basically cutting of signals into small time slots which can be considered to have a constant

property. In this experiment, speech signal is divided into frames of size 320 samples. To avoid the loss of information, frames should be 25% overlapping. Now, we calculated the Ri values through auto-correlation method. It is the degree of similarity/agreement between a given time series(audio signal) and a lagged version of itself over successive intervals of time. It is calculated using Eq. 1

$$R_i = \sum_{m=0}^{N-i-1} S(m) * S(m+i)$$

(1)

In above equation value of i will vary from 0 to 12. So, for each frame of size 320 sample with the above equation we got 13 values naming R_0, R_1, R_2...R_{12}. Then we used Levinson and Durbin's algorithm to compute the linear prediction filter coefficients. It uses the auto-correlation method to estimate the linear prediction parameter for a segment of speech.

3.3 Affinity Propagation

Now we got 12 Ci values $C1$, $C2$, $C3$...$C12$ for each frame. We took all C1 from each frame together and treated it as one set of features and similarly, we took other Ci's and got the final 12 sets of different features Fig. 2. Now we select the features which get the closest optimal classification so as to reduce the modeling time. For selecting features we find the clustering score of each feature set using Affinity Propagation and select the feature sets which gave the highest score, as the better clustering score implies that the set of features differentiates the classes well.

The Affinity propagation algorithm finds the number of clusters and the mapping of data points to the cluster. It doesn't require the initial centroids or number of clusters. It models each data point as a node in a network.

First, the similarity matrix is generated with the data points by negating the sum of the squares of the differences between data points. The similarity matrix contains information of preferences and similarity. Preferences are the data points that are suitable to be an exemplar. It is stored in the diagonal and the non-diagonal elements represent the similarity values between data points.

As a starting point, we construct an availability matrix with zero elements. Then we use Eq. 2 to calculate the responsibility matrix.

$$r(i, k) \leftarrow s(i, k) - \max_{k', k' \neq k} \{a(i, k') + s(i, k')\}$$

(2)

where 'i' is rows and k' is column of responsibility matrix. "Responsibilities" $r(i, k)$ is sent from data points to candidate exemplars to indicate how strongly each data point favors the candidate data exemplar.
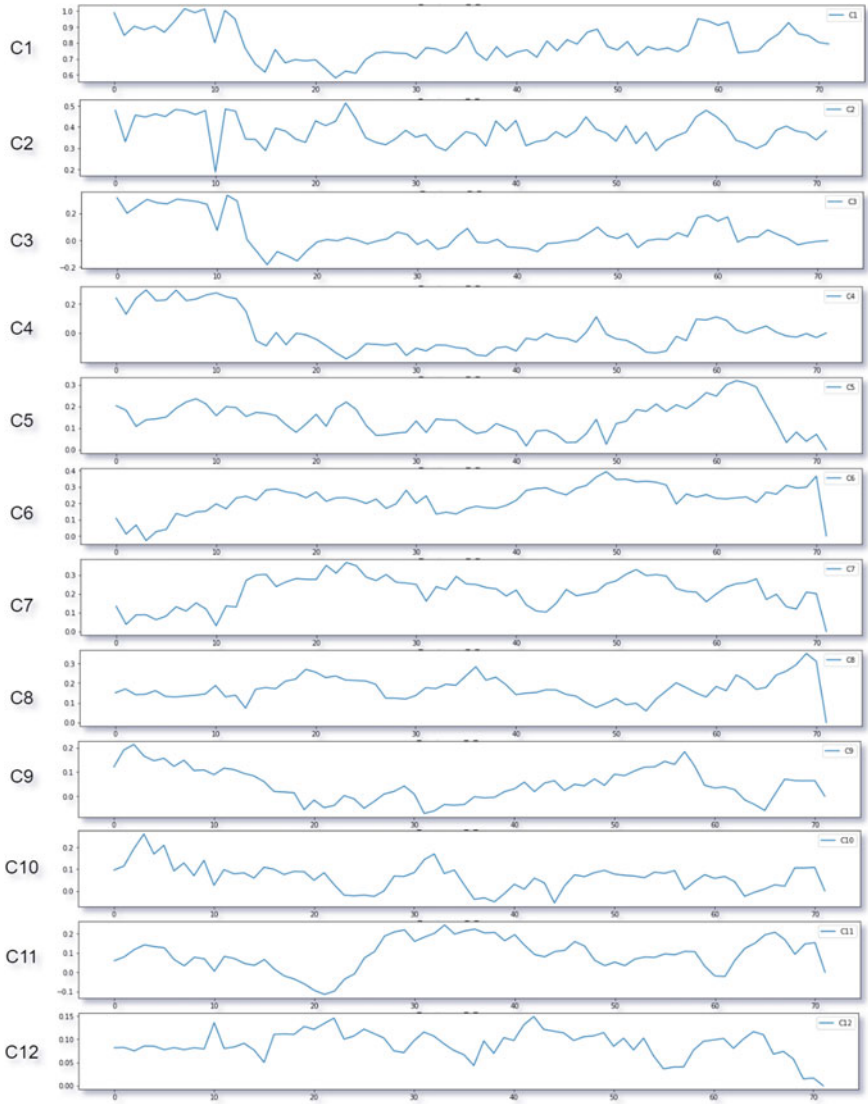
Fig. 2 Features of 1st utterance of vowel /a/

Table 1 Parameters of SVM

No.	Specification	Description
1.	Regularization	1.0
2.	Kernel	RBF
3.	gamma	scale
4.	Cache size	200 MB

3.4 Classification

Now that we have selected features for labeled data points. We model and classify the vowels and digit data. From the dataset, we took training samples and then tested the remaining testing samples. We also used SVM to validate the selection of features by classifying the data using all feature sets and getting the same accuracy as with selected features.

4 Experiments and Results

We took the vowel and digit data from UASpeech database. English vowels have 5 classes/a/, /e/, /i/, /o/, /u/, and 10 classes of digits 0–9 with 7 utterances of each class. The LPCCs are extracted through self-written C code in Microsoft Visual Studio 2010. Affinity Propagation is done using a C-written code from scratch for understanding the algorithm well and then shifted to python with sklearn library for ease of experimenting with different data and changing parameters quickly. SVM for classification and verification of pre-clustering is also done in Python with sklearn library. Parameters for SVM are shown in Table 1.

The set of features for vowel /a/ is shown in the Fig. 2. Each feature is the set Ci for each frame from a voice signal. Using pre-clustering we got the clustering score, i.e., random index shown in Table 2. So according to the results shown in table, $c1$ and $c3$ has the highest clustering score and we select them. The selected features were used to classify by SVM and got 75% accuracy.

We also verified the results of selection by taking all the features and applied SVM and got the same accuracy of 75%. In the digit dataset we again selected the best features shown in Table 3 on the basis of pre-clustering scores which were c4 and c10. We applied the SVM on the selected features and got 92% accuracy which was then verified using all the features and got the same accuracy of 92%. Figure 3 shows the original cluster v/s the clustering done by Affinity Propagation where blue cluster represents /a/, green represents /e/, red represents /i/, cyan represents /o/, and magenta represents /u/.
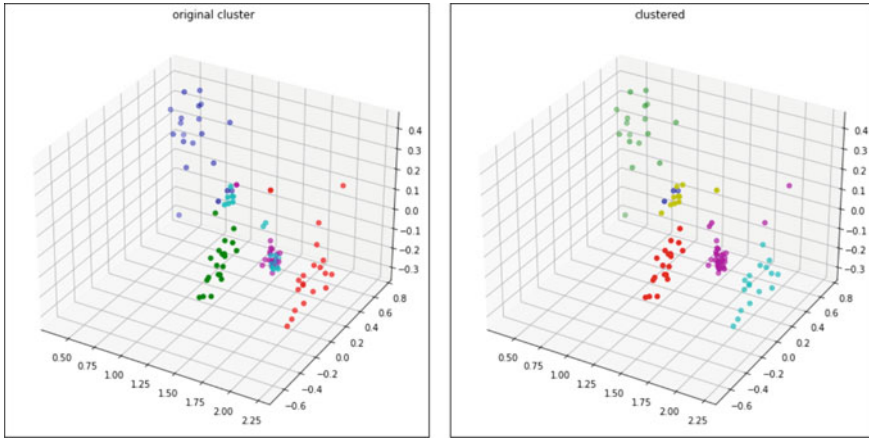
Fig. 3 Original clustering versus clustering by affinity propagation

Table 2 Clustering score(Random Index) of each feature set for vowel dataset

Feature	Score
C1	0.81
C2	0.78
C3	0.83
C4	0.72
C5	0.70
C6	0.73
C7	0.64
C8	0.70
C9	0.75
C10	0.68
C11	0.66
C12	0.69

The time taken by the model to train using all the features in vowel dataset was 3 ms, which got reduced to 0.9 ms after using the selected features. Similarly, for digit dataset the time for modeling was reduced from 18.9 ms to 1.9 ms.

5 Conclusion and Future Work

The experiments done in this system focused on using a pre-clustering algorithm technique over English-spoken vowels and digits spoken by a dysarthria patient. We used the LPCCs as feature extraction and did the classification using SVM. We

Table 3 Clustering score(Random Index) of each feature set for dysarthric digit dataset

Feature	Score
C1	0.87
C2	0.78
C3	0.87
C4	0.93
C5	0.08
C6	0.83
C7	0.85
C8	0.89
C9	0.84
C10	0.90
C11	0.86
C12	0.83

found that pre-clustering through AP works well with LPCCs for feature selection which reduced the modeling time and does not reduce the accuracy. In the future, this method needs to be verified for other datasets. This approach has been experimented for single speaker and can be combined with other feature extraction techniques. The given approach works well for a single speaker and can be used to build a better classification model for dysarthric speech.

References

1. Wu B et al (2017) An end-to-end deep learning approach to simultaneous speech dereverberation and acoustic modeling for robust speech recognition. IEEE J Select Topics Signal Process 11(8):1289–1300. https://doi.org/10.1109/JSTSP.2017.2756439. Dec.
2. Rudzicz F (2011) Articulatory knowledge in the recognition of Dysarthric speech. IEEE Trans Audio Speech Lang Process 19(4):947–960. https://doi.org/10.1109/TASL.2010.2072499. May
3. Biadsy F et al (2019) Parrotron: an end-to-end speech-to-speech conversion model and its applications to hearing-impaired speech and speech separation. arXiv preprint arXiv:1904.04169
4. Day M, Dey RK, Baucum M, Paek EJ, Park H, Khojandi A (2021) Predicting severity in people with aphasia: a natural language processing and machine learning approach. In: 2021 43rd annual international conference of the IEEE engineering in medicine and biology society (EMBC), pp 2299–2302, https://doi.org/10.1109/EMBC46164.2021.9630694
5. Sharma M, Sarma KK (2015) Dialectal Assamese vowel speech detection using acoustic phonetic features, KNN and RNN. In: 2015 2nd International conference on signal processing and integrated networks (SPIN), pp 674–678, https://doi.org/10.1109/SPIN.2015.7095270
6. Patgiri C, Sarma M, Sarma KK (2013) Recurrent neural network based approach to recognize assamese fricatives using experimentally derived acoustic-phonetic features. In: 2013 1st International conference on emerging trends and applications in computer science, pp 33–37, https://doi.org/10.1109/ICETACS.2013.6691390

7. Adi Y, Keshet J, Goldrick M (2015) Vowel duration measurement using deep neural networks. In: 2015 IEEE 25th international workshop on machine learning for signal processing (MLSP), pp 1–6, https://doi.org/10.1109/MLSP.2015.7324331

8. Frey B, Dueck D (2007) Clustering by passing messages between data points. Science, New York, N.Y., 315:972–6. https://doi.org/10.1126/science.1136800

9. Jing W, Jin W, Kai Q, Xiang X, Jingming K (2020) Binaural sound localization based on deep neural network and affinity propagation clustering in mismatched HRTF condition. EURASIP J Audio Speech Music Process. https://doi.org/10.1186/s13636-020-0171-y

10. Zhang X, Gao J, Lu P, Yan Y (2008) A novel speaker clustering algorithm via supervised affinity propagation. In: 2008 IEEE international conference on acoustics, speech and signal processing, pp 4369–4372. https://doi.org/10.1109/ICASSP.2008.4518623

11. Kang Z, Xingsheng G (2014) An affinity propagation clustering algorithm for mixed numeric and categorical datasets. Math Problems Eng. https://doi.org/10.1155/2014/486075

12. John T, Jin J, Justin D, Sydney C, Brandon WM (2017) Automated epileptiform spike detection via affinity propagation-based template matching. In: Conference proceedings: annual international conference of the IEEE engineering in medicine and biology society, 3057–3060. https://doi.org/10.1109/EMBC.2017.8037502

13. Bhagath P, Bharti K, Kotiya A, Das PK (2021) Feature selection using pre-clustering via affinity propagation for speech classification in low-resource languages. In: IEEE international conference on artificial intelligence in engineering and technology (IICAIET) 1–6. https://doi.org/10.1109/IICAIET51634.2021.9573696

14. Nguyen HT, Smeulders A (2004) Active learning using pre-clustering. In Proceedings of the twenty-first international conference on machine learning (ICML '04). Association for computing machinery, New York, NY, USA, 79. https://doi.org/10.1145/1015330.1015349

15. Kim H, Hasegawa-Johnson M, Perlman A, Gunderson J, Watkin K, Frame S (2008) Dysarthric speech database for universal access research. In: Proceedings of the annual conference of the international speech communication association, INTERSPEECH, 1741–1744

16. Latif S, Qadir J, Qayyum A, Usama M, Younis S (2021) Speech technology for healthcare: opportunities, challenges, and state of the art. IEEE Rev Biomed Eng 14:342–356. https://doi.org/10.1109/RBME.2020.3006860

Securing Secret Information

updates

Zeba Shamsi and Laiphrakpam Dolendro Singh

1 Introduction

Digital communication is an easy way to share information in this age of technology. Both cryptography and steganography can be used to protect valuable data from enemies. In cryptography, the communication is encrypted into a cipher that no human can decipher. Steganography is the technique of secretly encoding information in a media, such as a text file, image, audio file, video file, by utilizing certain mathematical formulae to mask the data's existence [1]. People think the QR code tag is the best example of image steganography [2]. Denso Wave, a Japanese company that is part of Toyota, came up with Quick Response codes (QR codes) in 1994 [3]. Today, they are very widely known in Asian countries like Japan, South Korea, China, and Taiwan. They are also growing in popularity in western countries every day. The Quick Response (QR) code is a kind of two-dimensional bar code that takes the form of a square made up of black square dots on a white backdrop. Information is encoded in both horizontal and vertical directions. It encodes the information both horizontally and vertically, which means it can store many times more information than bar codes [2, 4]. QR codes can store not just numbers and letters but also the Japanese kanji alphabet. Apart from steganography and cryptography methods, QR codes might potentially be utilized for encrypted data transmission. The generation of QR codes is accomplished by combining encryption and steganography. These QR codes are utilized for various purposes, including confidential communication, copyright protection, marketing, business, and education, among others. Significant growth in QR code use has been seen recently. They were first employed for marketing goals [5], delivering static information regarding advertising. Another everyday use is

Z. Shamsi (✉) · L. D. Singh

Department of Computer Science and Engineering, National Institute of Technology, Silchar, Assam 788010, India
e-mail: zeba_rs@cse.nits.ac.in

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2024 565
B. P. Swain and U. S. Dixit (eds.), *Recent Advances in Electrical and Electronic Engineering*, Lecture Notes in Electrical Engineering 1071,
https://doi.org/10.1007/978-981-99-4713-3_55

to provide details on a city's surroundings. Nonetheless, QR codes and other forms of mobile technology are becoming more popular in the classroom because of the ways in which they encourage more complex forms of student behavior and involvement. Sahu et al., [6] introduced two reversible data hiding (RDH)-based systems with better reversible data concealment. Their first system is an enhanced reversible dual image least significant bit system (LSB). The second method employs four cover images for concealing the payload with changed pixel value differences (CPVD) and n-rightmost bit replacement (n-RBR). Experimental evidence demonstrates that the presented system is very resistant to noise. Zhou et al. [7] introduced coverless, non-embedded images steganography-based system. In their approach, a group of selected images is used to create a database, and the database stores these images using hash sequences. The payload is compartmentalized into portions. If the image's hashing sequence matches the segment's, the image is sent. Their method is resistant to fluctuating brightness and noise. Barrera et al., [8] proposed a system employing optically encrypted QR codes as containers. Due to its tolerance for pollutant speckle noise, QR codes are utilized as containers in their approach. In addition, QR codes may be easily decoded using the cameras on mobile devices. The analysis shows that their technique is more susceptible to noise than conventional optical encryption. Ananth et al. [9] utilized both steganography and cryptography. Using a barcode encoder, the data or text to be concealed is first encoded into an analogous barcode image. After the encoding procedure, the original secret data is watermarked over a thermal image using enhanced LSB. On the steganography side, the least significant bit (LSB) replacement method is used to construct the stego image by hiding the barcode image over the thermal image. The experimental results of this proposed strategy have no effect on the image quality of stego.

A steganography approach must provide assurance of visual concealment. A QR code generator converts the input text into a QR code that cannot be read or comprehended by humans. However, any smartphone with a camera can readily interpret the text message buried in these QR codes. To secure communication from illegal access, a method is proposed that combines QR codes with steganography techniques. The proposed approach includes encoding and decoding processes at the sender and receiver, respectively.

1.1   Proposed Secure Stego Image Generation Process

The encoding procedure comprises converting text data into QR codes. Enciphering the QR code. Then ciphered QR code is embedded into the cover image at secret coordinates. The following steps are involved in the proposed secure stego image generation:

1. Input the secret text data and convert it into a QR code.
2. $\dim_x \leftarrow$ QR code x-dimension and $\dim_y \leftarrow$ QR code y-dimension.

3. Generate a 256 bits hash value (hash) using Secure Hash Algorithm 256. The first 128 bits (b_1) are used to generate x, and the later 128 bits (b_2) are used to generate y where x and y are given as:

$$x = \left| 0.5 - \text{BaseConvert}(b1, 2)/2^{128} \right|$$
$$y = \left| 0.5 - \text{BaseConvert}(b2, 2)/2 \quad /2^{128} \right| \tag{1}$$

Using X and Y as the initial parameters in the Henon map [11], random coordinate values and random sequence are generated as given below:

$$x' = -\left(1.4 \times x^2\right) + y + 1$$
$$y' = \underset{(\text{Cor}_x, \text{Cor}_y)=(x' \times 10^{15} \bmod \dim_x +1, y' \times 10^{15} \bmod \dim_y +1)}{\overset{1}{0.3}} \times x$$
$$r_{\text{bit}} = \text{Round}\left(\left|x'\right|\right)$$
$$X \leftarrow x'$$
$$Y \leftarrow y' \tag{2}$$

5. The barcode cipher (b_{cipher}) is generated as $b_{\text{cipher}} = r_{\text{bit}} \oplus \text{QR}_{\text{code}}$, where QR_{code} is the bit information of the QR code.
6. Embedding of the b_{cipher} information into random coordinates of the cover image (C_{img}) at least significant bit (LSB) or most significant bit (MSB) is performed to generate the stego image (S_{img}) as follows:

$$\text{IF}(\text{secret} == 0 \text{ AND MSB} == 0 \text{ at}(\text{Cor}_x, \text{Cor}_y))$$
$$\text{MSB}\left[(\text{Cor}_x, \text{Cor}_y)\right] \leftarrow 0$$
$$\text{LSB}\left[(\text{Cor}_x, \text{Cor}_y)\right] \leftarrow 0$$
$$\text{IF}(\text{secret} == 0 \text{ AND MSB} == 1 \text{ at } (\text{Cor}_x, \text{Cor}_y))$$
$$\text{MSB}\left[(\text{Cor}_x, \text{ Cory })\right] \leftarrow 1$$
$$\text{LSB}[(\text{Corx}, \text{Cory})] \leftarrow 0$$
$$\text{IF}(\text{secret} == 1 \text{ AND MSB} == 0 \text{ at } (\text{Cor}_x, \text{Cor}_y))$$
$$\text{MSB}[(\text{Cor}_x, \text{Cor}_y)] \leftarrow 0$$
$$\text{LSB}[(\text{Cor}_x, \text{Cor}_y)] \leftarrow 1$$
$$\text{IF } (\text{secret} == 1 \text{ AND MSB} == 1 \text{ at } (\text{Cor}_x, \text{Cor}_y))$$
$$\text{MSB}[(\text{Cor}_x, \text{Cor}_y)] \leftarrow 1$$
$$\text{LSB}[(\text{Cor}_x, \text{Cor}_y)] \leftarrow 1 \tag{3}$$

The stego image (S_{img}) can be shared to the communicating party. The initial keys x and y of the henon map can be securely shared with a standard public key encryption scheme with the communicating party.

1.2   Proposed Secure Stego Image Generation Process:

The decoding procedure comprises extracting the pixels at a random location from the stego image. Out of the extracted pixels, LSB or MSB information is extracted based on the condition given below in Step 5. The selected LSBs or MSBs are XORed with the random sequence generated from the Henon Map using the same initial keys as the sender, yielding the secret QR code. The deciphered QR code is converted to secret text using a Barcode recognizer.

The following steps are involved in the proposed secret data extraction:

1. Import stego image (S_{img}) and get the initial keys x and y of the Henon Map.
2. Using x and y as the initial parameters in the Henon Map, random coordinate values (cor_x, cor_y) and random sequence (r_{bit}) are generated using the same calculation given in Step 4 of Sect. 1.1.
3. Pixel values (pix) are extracted from the stego image from coordinates (cor_x, cor_y).
4. b_{cipher} data are extracted from each pixel's pix as follows:

$$\text{IF}\left(\text{MSB}\big[(cor_x, cor_y)\big] == 0 \underset{b_{cipher}\leftarrow 0}{\text{AND}} \text{LSB}\big[(cor_x, cor_y)\big] == 0\right)$$

$$\text{IF}\left(\text{MSB}\big[(cor_x, cor_y)\big] == 1 \underset{b_{cipher}\leftarrow 0}{\text{AND}} \text{LSB}\big[(cor_x, cor_y)\big] == 0\right)$$

$$\text{IF}\left(\text{MSB}\big[(cor_x, cor_y)\big] == 0 \underset{b_{cipher}\leftarrow 1}{\text{AND}} \text{LSB}\big[(cor_x, cor_y)\big] == 1\right)$$

$$\text{IF}\left(\text{MSB}\big[(cor_x, cor_y)\big] == 1 \underset{b_{cipher}\leftarrow 1}{\text{AND}} \text{LSB}\big[(cor_x, cor_y)\big] == 1\right) \tag{4}$$

5. The secret QR code is generated as $QR_{code} = b_{cipher} \oplus r_{bit}$.
6. The QR code is converted to the secret text data.

2   Simulation

The simulation of the proposed method is executed on a laptop with 11th Gen Intel(R) Core (TM) i7-1165G7 @ 2.80 GHz with 16 GB RAM using Wolfram Mathematica. The sample images are obtained from USC-SIPI [12] image database. Figure 1 b–f shows the secret text, the QR code of a secret text, the ciphered QR code, the cover image, the stego image and the recovered QR code.

The proposed approach has been evaluated utilizing QR codes for secret text messages. The recovered secret message is identical to the original secret text. The outcomes of these tests are discussed in the next section.
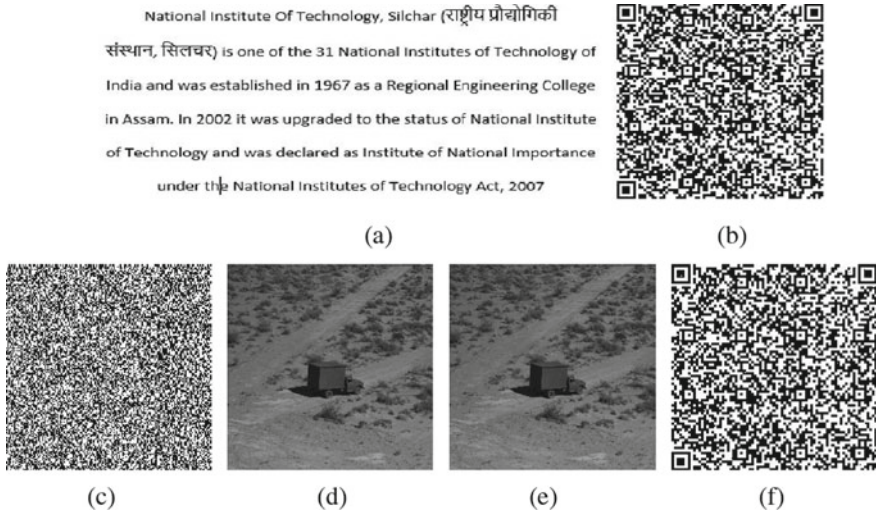
Fig. 1 **a** Text. **b**, **c** Plain and Cipher QR code. **d** Cover image. **E** Stego Image. **f** Extracted

3 Analysis

The most essential need for any data embedding approach is imperceptibility. Mean Square Error and Peak Signal to Noise Ratio are used as metrics for assessment. A high PSNR value indicates an almost imperceptible signal.

3.1 Mean Square Error:

The Mean Square Error (MSE) measures the mean square of the differences between the hidden message and the stego image. It provides us with a quantitative estimate of the cover image error caused by the data embedding procedure.

$$\text{MSE} = \frac{1}{M \times N} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} \left(C_{ji} - S_{ij} \right)^2 \tag{5}$$

where M and N are the numbers of rows and columns of the cover image. C_{ij} and S_{ij} are the pixel value from the cover image and the stego image. Higher values of MSE show that the cover image and stego image are not the same.

3.2 Peak Signal to Noise Ratio

Peak Signal to Noise Ratio (PSNR) is another way to measure the quality of the stego image compared with the cover image. It is the ratio between the signal's highest possible value and the noise's power (MSE). dBs are used to measure it. A better embedding is one with a higher PSNR value [10]

$$PSNR = 20 \log_{10}\left(\frac{MAX}{\sqrt{MSE}}\right) \qquad (6)$$

where $MAX = 255$ for grayscale image.

3.3 Structural Similarity Index

SSIM is way to compare the cover image and the stego image to see how similar they are. It measures how different the two images seem to the eye.

$$SSIM = \left(2\mu_x\mu_y + C_1\right)\left(2\sigma_{xy} + C_2\right)/\left((\mu_x)^2 + (\mu_y)^2 + C_1\right)\left((\sigma_x)^2 + (\sigma_x)^2 + C_2\right) \qquad (7)$$

Here,

μ_x and μ_y are two mean intensity values of images x and y.

$\sigma_x{}^2$ is the variance of x. $\sigma_y{}^2$ is the variance of y. σ_{xy} is the covariance of x and y. C_1 and C_2 are two stabilizing parameters.

3.4 Salt and Pepper Noise Attack

To test the proposed method against a salt and pepper noise attack, the stego image is induced with 10% salt and pepper noise. The proposed method can recover the secret text even with 10% salt and pepper noise attacks. Figure 2a, b, d, and e shows the cover image, stego image with 10% salt and pepper noise, original QR code and recovered QR code, respectively.

3.5 Occlusion Attack

To test the proposed method against occlusion attack, the stego image is induced with 6.25% occlusion attack. The proposed method can recover the secret text even with a 6.25% occlusion attack. Figure 2a, c, d, f shows the cover image, stego image
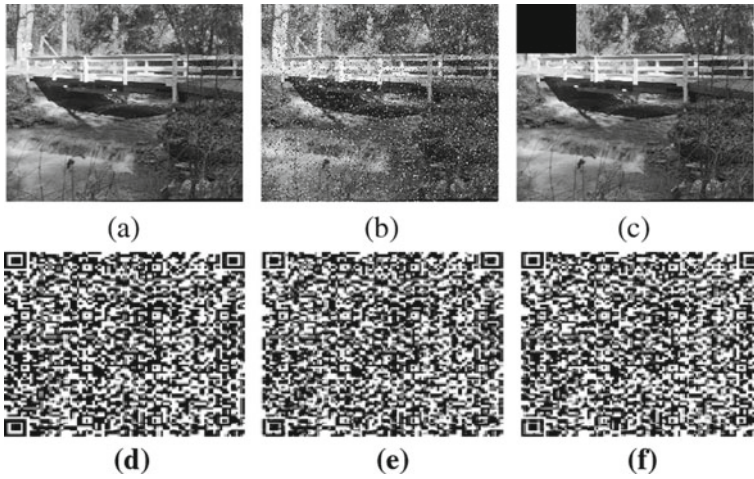
Fig. 2 a Cover image. **b, c** Stego image with 6.25% occlusion attack and 10% salt and pepper noise attack. **d** Original QR code. **e–f** Recovered QR code from (**b**) and (**c**)

with 6.25% occlusion attack, original QR code, and recovered QR code, respectively. Table 1 shows the MSE, PSNR, and SSIM of the proposed technique when the cover image is embedded with the QR code data. Table 2 shows the comparison of the proposed method with other state-of-the-art techniques with respect to MSE, PSNR, and SSIM of the stego-image when the cover image is embedded with secret data.

Table 1 MSE, PSNR, and SSIM of the cover image and the stego image

Cover image	Cover image	QR code	MSE	PSNR	SSIM
	Dimension	Dimension			
Plane	512×512	154×154	0.0451	61.5798	0.9999
Stream and bridge	512×512	162×162	0.0500	61.1362	0.9999
Aerial	512×512	177×177	0.0599	60.3555	0.9999

Table 2 Comparison

Method	Cover data size	Secret data size	PSNR	SSIM
Ref. [13]	19 KB	6.6 KB	57.12	0.99
Ref. [14]	256 KB	52.4 KB	55.69	0.91
Proposed	264 KB	24.9 KB	60.35	0.99

4 Conclusion

A method is proposed that securely hides the secret information both in the MSBs and LSBs of a cover image. A hash value is obtained by applying SHA256 on the QR code (generated from the secret text). The hash value is used to derive the initial key for processing the Henon Map. The coordinates where the secret data are hidden are determined by a random sequence generated using the Henon Map. The secret QR code is ciphered with a random sequence generated from Henon Map to provide additional security. From the analysis, it is seen that the proposed method maintains good imperceptibility, security, and robustness.

References

1. Morkel T, Jan HP Eloff, Olivier MS (2005) An overview of image steganography. Inf Secur South Africa 1(2):1–11
2. Rani MMS, Rosemary Euphrasia K (2016) Data security through qr code encryption and steganography. Adv Comput: Int J (ACIJ) 7(1/2):1–7
3. Iso/Iec 18004 (2006) Information technology—automatic identification and data capture techniques QR Code 2005 bar code symbology specification
4. Lin P-Y (2016) Distributed secret sharing approach with cheater prevention based on QR code. IEEE Trans Industr Inf 12(1):384–392
5. Hernando R, Mac´ıas JA (2022) Development of usable applications featuring QR codes for enhancing interaction and acceptance: a case study. Behav Inf Technol 1–19. https://doi.org/10.1080/0144929X.2021.2022209
6. Sahu AK, Swain G (2022) High fidelity based reversible data hiding using modified LSB matching and pixel difference. J King Saud Univ-Comput Inf Sci 34(4):1395–1409
7. Zhou Z, Sun H, Harit R, Chen X, Sun X (2015) Coverless image steganography without embedding. In: International conference on cloud computing and security. Springer, Cham, pp 123–132
8. Barrera JF, Mira A, Torroba R (2013) Optical encryption and QR codes: secure and noise-free information retrieval. Opt Express 21(5):5373–5378
9. Ananth SV, Sudhakar P (2016) Performance analysis of a combined cryptographic and steganographic method over thermal images using barcode encoder. Indian J Sci Technol 9(7):1–5
10. Obaid ZK, Saffar NFHA (2021) Image encryption based on elliptic curve cryptosystem. Int J Electric Comput Eng 11(2):1293
11. H´enon M (1976) A two-dimensional mapping with a strange attractor. Commun Math Phys 50:69–77. https://doi.org/10.1007/BF01608556.
12. The USC-SIPI Image Database. http://sipi.usc.edu/database/. Accessed 19 May 2020
13. Basu S, Debnath A, Basu A et al (2022) An image data hiding technique using differential evolution. Multimed Tools Appl 81:39995–40012. https://doi.org/10.1007/s11042-022-12557-0
14. Younus ZS, Hussain MK (2022) Image steganography using exploiting modification direction for com-pressed encrypted data. J King Saud Univ—Comput Inf Sci 34(6):2951–2963. https://doi.org/10.1016/j.jksuci.2019.04.008

A Comprehensive Study of DDoS Attack on Internet of Things Network

Nitin Anand and Khundrakpam Johnson Singh

1 Introduction

A network of gadgets with sensors that connect with one another online is preferred to be known as the Internet of Things (IoT). Without Internet access, it is impossible to function in the modern world. In many situations, the Internet is essential for communication. As we utilise the Internet for more daily tasks, it becomes less safe. Attacks are typically polymorphic in nature. DDoS attacks are used in some well-known attacks. DDoS attacks use multiple computers rather than just one system or host to carry out the attack. DDoS attacks have evolved from DoS attacks in recent years. Kumar and Selvakumar [1]. By flooding a website or online service with a lot of Internet traffic coming from different places, a Distributed Denial of Service (DDoS) attack tries to take down the target website or online service. A computer or other networked resource, such as an IoT device, is one example of exploited machinery. A DDoS attack uses multiple computers and numerous Internet connections, typically dispersed throughout the world in a botnet, as opposed to a DoS attack, which employs one computer and one Internet connection to flood an action plan to ensure with packets. According to Tuan et al. [2], a large-scale volumetric network can produce tens of gigabits of bandwidth per second (or possibly hundreds of gigabits). The main focus is how the attacker interacts with the Control Server to gain command as well as control of the system. Powerful servers with lots of memory, processing speed, and bandwidth make up a control server. Agents, also known as handlers, are in charge of keeping an eye on Botnets (BN01–BN05) and

N. Anand (✉) · K. J. Singh
Department of Computer Science and Engineering, National Institute of Technology, Imphal, Manipur, India
e-mail: nitin1036@nitmanipur.ac.in

K. J. Singh
e-mail: johnsonkh@nitmanipur.ac.in

receiving commands from an attacker. They provide the Botnets with configuration and upgrading instructions. A computer network's owner has no knowledge what malware is on their device or whether they are a part of a botnet if their network is vulnerable to attack. The attackers start attacking the targets after the agents.

In the following sections Challenges and Security issues are discussed in 1.1. Moving further, the Subsect. 1.2 deals with the working of DDoS attack on IoT networks. The next Subsect. 1.3 talks about the Research Methodology carried out for the completion of the comprehensive study Sect. 2 includes the Literature Review which is based on the highly cited articles in the field of our study. After this we will move to the Sect. 3 which is all about intrusion detection, types and its applications in the IoT-based Environment. We will end the paper with Conclusion and Future work which will be seen in the 4th and the Final Section.

1.1 Challenges and Security Issues in IoT

Bandwidth and Power Consumption: Devices for the Internet of Things (IoT) are made to be compact and have a weaker CPU and memory**.** Both bandwidth and power usage increase as a result. Therefore, it is not possible to employ modern encryption algorithms due to the IoT system's high memory and computing needs. While respecting security concerns, IoT uses numerous linked sensors to complete activities, which might consume a lot of bandwidth. IoT systems should therefore incorporate security measures that have the least amount of overhead possible as mentioned in Mohsin and Hamad [3].

Inadequate mechanisms for authentication and authorization: The most of Internet of Things (IoT) devices lack access control, have weak default passwords, and unsecure access. As a consequence, an attacker could use this to compromise the confidentiality and integrity of data.

Web interface insecurity: There are a number of Internet of Things (IoT) devices with web interfaces that don't demand users use difficult passwords. Some of them nevertheless permit those who have repeatedly failed to log in. As a result, these interfaces are vulnerable to numerous attacks, such as scripting, injections, as well as brute-force password attempts, and unsecured network services. All Internet of Things (IoT) devices require network connectivity, hence these networks must be secure. If not, buffer overflows, fuzzing, DDoS, and other attack types will jeopardise network services.

Poor physical security: If a hostile actor has physical access to an IoT device, he could exploit it to recover the stored.

Information by breaking or removing the storage card.

The Internet of Things (IoT) will gather and store a tonne of personal data as the number of connected devices rises. Hackers and other online criminals are therefore drawn to this information. An IoT system can be attacked using a variety

of methods, such as reducing network availability, introducing false data into the network, collecting personal information, etc. In order to secure consumer privacy and security, valid security, privacy, and trust models for IoT applications should be created, and practical technical solutions (such as intrusion prevention systems, firewalls), should be found as a result of which IoT technology uptake and acceptance will depend on how well users' privacy is protected. It keeps track of network activity at various levels and employs a set of rules and protocols to regulate it. As mentioned in Goel et al. [4], a firewall could be either software or hardware. Firewalls are grouped according to where they are located or how they operate. The firewall can be further divided into four groups based on its way of operation: Stateful Inspection Firewall, Circuit-Level Gateway, Packet Filtering Firewall, and Application Level Gateway are the order in which they are used. A list of certain firewalls may be found in Gupta et al. [5].

1.2 How DDoS Attack Work on IoT Network

The Internet of Things is a mechanism that Hadrami and Hussain [6] highlight as having a major uptick in both the industrial and consumer markets (IoT). This technology will change many areas of our daily lives, from the way we sip our coffee to the way smart objects makes interaction with industrial applications, which could have a significant impact on the customer experience. Such quick development and deployment are hampered by a number of issues, including the sheer amount of data collected, its size and variety of the network, as well as privacy and security concerns. One of the most recent issues that has emerged and needs attention is Distributed Denial of Service (DDoS) attacks in IoT networks. DDoS attacks exacerbate this issue in the IoT application by taking advantage of IoT devices' limited resources, such as storage space as well as network capacity.

Arguments are made for the importance of data security, especially for individuals in crucial positions, based on various IoT vulnerabilities as described in Table 1. These includes power utilities, industrial facilities, building automation, etc. Data storage and transmission with encryption ensures that only those with the proper authorization can access the information. Resources and known algorithms determine the security of cryptosystems. Many IoT devices leave ports open while executing vulnerable services [7], which enables an attacker to access and take advantage of a number of flaws. It is important to implement strong credential management to stop unauthorised access to IoT systems and networks. The majority of Internet of Things (IoT) devices and associated cloud administration platforms do not need adequately complicated passwords.

Table 1 Vulnerabilities in internet of things network

Vulnerabilities	Examples
Insufficient authentication	Lack of access control
Privacy concerns	Redundant user data, lack of confidentiality and other well-known CIA triads
Insecure interface	Default usernames and passwords cannot be changed. Poor password recovery techniques
Insecure software	Insecure server, lack of encryption and verification
Insecure network services	Vulnerability to Denial of service and buffer overflow attacks
Insufficient configurability	Cannot differentiate between administrator and users
Poor physical security	Easy to dissemble the devices
Lack of integrity and encryption verification	Unencrypted data are heavily transmitted

1.3 Research Methodology

To accomplish the security in IoT, a research methodology was applied. The authors read up on IoT security literature in order to assess the risks taxonomy and issues surrounding IoT security. Following the completion of these taxonomies, the writers evaluated the several approaches provided in those surveys and chose a number of major difficulties that they genuinely consider are crucial for network security. Based on the writers' personal security domain experience, these topics were chosen. There were other possibilities which were well-known in the field. Additionally, the authors searched the web for intriguing security-related publications using search engines.

1.3.1 Evaluation of Research Content

An impartial search technique was needed to locate initial reviews of intrusion detection systems as well as collect all pertinent literature. Search terms must be specified, and the search procedure must be as thorough and logical as it is feasible to be. We see that the word "intrusion" appears in the titles of various works on anomaly-based intrusion detection.

Therefore, we have mentioned the search phrase as "network intrusion detection" in order to include papers on anomaly-based intrusion detection.

1.3.2 Research Questions

To direct our research, we have created clear research questions (RQs) as well as in subquestions. These queries provide a thorough summary of the entire scenario and were raised during the duration of the course.

Application Domain

This deals with the aspects related to the domains that are covered by intrusion detection techniques. It also discusses about the studies based on the distribution among various areas with appropriate reasons. It also evaluates the variation of results from country by country.

Data Pre-processing Methods

It finds out the well-known data pre-processing methods used in network intrusion detection and discusses the implementation of pre-processing technologies with their technical features. We discuss the various pre-processing techniques.

(a) **Data Cleaning**:

Cleansing data fixes records that are incomplete or wrong. The following are potential quality standards.

- Reliability: Data may need to be of a specific type, such as numeric or Boolean.
- Precision: Data must match the circumstances. For instance, the recording procedure itself may have produced anomalies. Because validating actual data sources requires real data sources, accuracy cannot be guaranteed through data cleansing.
- Completion: There could be unknown or missing values in some data. Completeness problems are typically fixed by setting default values, zeroing out existing values, or removing them.

Uniformity: Conflicts within a dataset lead to inconsistency. A source IP, for instance, can be different between two receivers. Finding the most dependable datum is necessary for solving this kind of issue.

Errors: Here, the values can sometimes be attributed to a mean or other statistical metric, which can be found through data analysis and is reliant on different factors including means, standard deviations, and clustering methods.

(b) **Imbalanced Learning**

An uneven distribution of positive as well as negative cases in the sample will bias learning in favour of the higher proportion. For instance, the model won't be useful in the utmost scenario of a dataset with 95% positive as well as 5% negative results. Attacks are typically few, hence datasets for intrusion detection are frequently unbalanced.

(c) **Data Conversion**

It is frequently necessary to convert and map training data before feeding it to the model in order to satisfy the demands as well as requirements and increase the detection speed as well as accuracy. IDS datasets two categories of data are impacted by this.

(d) **Feature Selection**

A subset of the original dataset is chosen for the model's input through the feature selection process. This can help in generalisation and prevent dimensional catastrophes. Data must have redundant or irrelevant features in order to execute feature selection and prevent excessive information loss. There are numerous methods for selecting features.

- **Manual choice**. A feature's removal decision is made manually.
- **A thorough search**. It can take a lot of work to evaluate every conceivable feature subset and identify the subgroup with the lowest error rate.
- **Embedded approach**: While building the model, features are chosen. The L1 and L2 penalties of ridge regression are combined in the Bolasso technique to reduce multivariate regression coefficients towards zero.
- **Wrapper strategy**: One trains a prediction model that is tested on at withstanding set after each subgroup. The result of the subset is determined by the model test's error rate. This requires a lot of processing and is typically utilised simply to identify the best feature subset.
- **Filtering strategy**. While methods like mutual information, Pearson correlation coefficients, in addition to significance scores, like inter- or intra-class distances and rankings features, can be used to score a subset of attributes, they are unable to produce the optimal subset.

(e) **Feature Extraction**

Adding additional features to aid in learning is seen as a crucial component of model construction. The following algorithms are capable of carrying out this.

Principal Components Analysis (PCA). Principal Component Analysis (PCA), one of the most popular linear dimensionality reduction techniques, modifies the data's foundation in accordance with both the eigenvectors as well as eigen components of the data covariance matrix.

Linear Discriminant Analysis (LDA). LDA is a well-known dimensionality reduction algorithm that seeks linear combinations of attributes to define various object classes. As a supervised learning technique, it explores the low-dimensional space for vectors that most improved decision making data classes.

Auto encoder. Unsupervised learning is accomplished using hidden layers, which map high-dimensional characteristics using a nonlinear transformation to create a representation that is as close to the original input as is practical.

(f) **Data Visualisation**

Data visualisation, which uses graphics to depict data, is used to assist academics better comprehend features like data distribution. By enhancing data features and characteristics, visualisation in intrusion detection aids in our understanding of an attack's characteristics.

Detection Techniques

The mechanisms used in intrusion detection approaches are highlighted. The most crucial question is how machine learning as well as deep learning methods are used for intrusion detection. It also covers the fundamentals and traits of several intrusion detection techniques.

Evaluation Metrics

This analyses the evaluation of the performance of intrusion detection technologies and their application. The evaluation criteria that are frequently utilised in intrusion detection articles are being discussed. The most often utilised metrics are **precision, recall, F1 value, accuracy** as well as the **False-Alarm Rate (FAR)**. Throughout many articles, recall as well as accuracy are exploited. Recall, also referred to as detection rate or true positive rate, is the ratio of successfully identified attacks to all attacks (TPR).

Recall is a metric that can be used to determine how well the classifier identified attacks. The accuracy of a classifier is determined by dividing the number of samples that were properly predicted by the total number of samples that were forecasted.

$$F1 \text{ value} = 2[\text{Pr ecision} \times \text{Recall}]/(\text{Precision} + \text{Recall}) \tag{1}$$

$$\text{Where Precision} = \text{True Positives}/(\text{True Positives} + \text{False Positives}) \tag{2}$$

$$\text{And Recall} = \text{True Positives}/(\text{True Positives} + \text{False Negatives}) \tag{3}$$

Authors

It highlights the main contributors and their expertise in the relevant fields. By looking at the overall number of citations of the listed works through Scopus, we evaluated the key contributors to intrusion detection.

Datasets

It deals with public datasets available for research and development. The usage of those datasets specifically in our area of interest. Along with these, it also discusses the reason behind the wide use of these datasets. We have examined current network intrusion detection datasets. We have obtained specific datasets from the survey, as indicated in Table 2. We identified each dataset's year of creation, development technique, data volume, annotation state, number of tags, and links using data presented

by the dataset publishers as well as other searches. Fresh datasets [8] have been regularly entering the community since the DARPA 1998 dataset. The year 2009 served as the node for the examination of the growth in network intrusion detection datasets.

A dataset called **NSL-KDD** has been proposed to address some of the fundamental flaws with the KDD99 dataset. Due to the dearth of publicly available data sets for network-based IDSs, we presume this updated version of the KDD dataset can still be used as an efficient benchmark dataset to aid researchers in comparing various intrusion detection techniques.

The NSL-KDD train and test sets also contain a sizable amount of records. Due to this benefit, the full data set can be used for the tests rather than just a tiny sample that must be chosen at random. As a result, evaluation outcomes from various research projects will be comparable and consistent.

The Australian Cyber Security Centre's Cyber Range Laboratory developed UNSW-NB15. Due of its vast range of cutting-edge strikes, it is commonly used. Fuzzers, analysis, DoS, exploit, backdoors, generic, reconnaissance, worms and shellcode, are just a few of the numerous types of attacks. The testing set has 175,341 records, whereas the training set has 82,332 records. In addition to benign and common attacks with source data (PCAPs), the CICIDS2017 database also includes network traffic analysis findings (CSV files) based upon timestamps, source and destination IP addresses, source as well as destination ports, protocols, and token flows of attacks. The B-Profile technology was utilised by the researchers to provide secure background traffic and evaluate the non-local behaviour of individual interactions.

In place of ranking a particular feature selection algorithm, Singh et al. [9]'s ensemble approach for feature selection has been created. This methodology combines the actions of such seven feature selection algorithms. The average of

Table 2 Datasets for intrusion detection

Sl. No.	Datasets	Year	Authenticity	Count	Label	Number of labels
1	UNSW-NB15	2015	Emulated	2,540,044	Yes	9
2	CIDDS 001	2017	Emulated	31,959,267	Yes	6
3	CICIDS 2017	2017	Emulated	2,830,743	Yes	7
4	CICDoS 2017	2017	Emulated	32,925	Yes	8
5	SUEE 2017	2017	Emulated	19,301,217	Yes	3
6	PUF	2018	Emulated	6,000,000	Yes	4
7	ISOT CID	2018	Real	36,938,985	Yes	18
8	BoT-IoT	2019	Real	73,360,900	Yes	2
9	InSDN	2020	Emulated	343,139	Yes	7
10	CIRA-CIC-DoHBrw 2020	2020	Emulated	1,185,286	Yes	3

each feature selection technique is then determined, along with the threshold value, h, which is obtained by averaging the results of the seven feature selection algorithms.

As Saeys et al. [10] point out, feature selection is a crucial pre-processing step that is typically used to choose the lowest subset of features that will maximise the model's performance across a variety of machine learning applications. In addition to improving model performance, feature selection has additional benefits such as allowing the design of simpler and faster models by focusing on a subset of features and improving comprehension of the processes shown by the data.

In an anomaly IDS, the majority of network traffic features in a dataset are rarely useful for creating a machine learning model. When duplicate features are added, the machine learning model's generalizability is diminished, as well as the classifier's performance might suffer as well.

The usage of many attributes for anomaly detection raises the size of the model for machine learning significantly. As a result, when developing machine learning models, it is critical to carefully analyse the feature selection procedures utilised in anomaly IDS by Kshirsagar and Kumar [11]. The filter-based feature selection technique suite specifies the IG, GR, CR, and ReliefF as follows:

Information Gain (IG)
The filter-based method for feature selection aids in calculating the significance of a dataset property. The unpredictability of samples is represented by a feature's entropy. IG is calculated using the following formula:

$$IG(M, N) = H(M) - H(M|N) \tag{4}$$

$H(M \mid N)$ is the probability of a class M given a given characteristic N, whereas $H(M)$ is the entropy of a class M.

Gain Ratio (GR)
It rectifies the bias issue when features present the dataset with a collection of values, as represented by IG. GR is a modest enlargement of IG. The formula for calculating GR [4] is

$$GR(M, N) = [\text{Gain}(M, N)]/\text{Intrinsic Information}(N) \tag{5}$$

where Intrinsic Information $= -\sum[|S_i|/|S|]*[\log_2|S_i|/|S|]$.

$|S|$ is the number of possible values a feature N and $|S_i|$ is the number of real values of feature.

ReliefF
The ReliefF algorithm determines the feature vector that is closest to the same target class as the near-hit and the neighbouring alternate class that is referred to as the near-miss using Euclidean separation. The following equation gives the weight (W) of feature i as

$$W_i = W_i - (x_i - \text{nearHit}_I)^2 + (x_i - \text{nearMiss}_i)^2 \tag{6}$$

Correlation

Correlations between attributes pertaining to the class are measured using a method based on conditional entropy. The following equation is used to determine correlation.

$$\text{Correlation} = \frac{\sum(M - M')(N - N')}{\sqrt{\sum(M - M')2}\sqrt{\sum(N - N')2}} \tag{7}$$

where M' and N' are respective sample mean for M and N.

Chi Square (χ^2)

By calculating a score to represent the level of independence between these two features, the chi-square (2) assessment is used to examine the freedom of two features.

$$\chi^2 = \sum (F_{\text{observed}} - F_{\text{Expected}})^2 / F_{\text{Expected}} \tag{8}$$

where F_{observed} is the frequency of either an attack or a typical occurrence for each child, and F predicted is determined by multiplying the total rows and columns by both, and then dividing by the total frequency. The outcome of the calculation will demonstrate the degree to which each feature is related to forecasting the clients' status (either under attack or normal).

2   Literature Review

Since identifying DDoS attacks is fundamentally a distributed problem, a strong distributed solution is ideal. The activity of legitimate users shouldn't be impeded by a DDoS protection system. A defence system's security elements must guard against external and internal threats that could support DDoS attacks launched from within a network. Scalable as well as deployable defence systems should have compelling financial incentives. Instead of being a one-size-fits-all answer to all problems, defence should be planned such that it can be built into a larger solution over time. According to Singh and De [12], Distributed Denial of Service (DDOS) is a serious security threat at the moment. DDoS attacks are difficult to recognise and consume less bandwidth at the application layer. This study examines the many elements of such an attack and pinpoints the crucial components of the protocols used in each layer seven attack. The following parameters have been identified: HTTP count, IP density over a short time, constant mapping of random source port to regular destination port, and frame length. The dataset for the classification model was created by integrating all conceivable outcomes. Based on the supplied input attribute, they classified arriving packets in their article as either belonging to the attack category or the regular category.

Johnson et al. [13] looked into the effects of DDoS attacks and important variables that affect the attack. To identify the DDoS attack based on the quantity of HTTP GET requests, the entropy of the requests, and the variance of entropy, they trained an MLP using a GA learning approach. Entropy is found to be more valuable when dealing with regular clients and less valuable when dealing with attacks, they found. Additionally, they discovered that the variance of an attacking client is almost zero. It demonstrates that there were a consistent number of HTTP GET requests made during the attack. The proposed method has detection accuracy of 98.32%, sensitivity of 0.9962, and specificity of 0.9962 for DDoS attacks.

According to inter-arrival time, more than 100 distinct IP addresses from the CIADA 2007 dataset are grouped into several clusters by Johnson et al. [14]'s proposed DDoS detection procedure is based on SOM. Additionally, they offer a computer simulation of a DDoS attack and look at how it interacts to variables like the inter-arrival time of the attackers when hitting the server. They found that the likelihood of resource exhaustion is influenced by the inter-arrival rate of the attackers entering the victim server.

Sajith et al. [15] employed the Adaptive Neuro-Fuzzy Inference System (ANFIS) as a classifier for classification into harmful categories and analysed the outcomes with Fuzzy GNP and HHO with the use of the DARPA 1999 Database. The outcomes show that ANFIS is better in terms of accuracy.

These attacks that were primarily based on the UNSW-NB-15 as well as CICDoS2019 Datasets were discovered by Priya et al. [16] using Rankers as a search approach. Furthermore, it emphasises the categorization of multi-class attacks for machine learning-based anomaly detection systems.

An integrated method for identifying DDoS attacks using machine learning (with a 97.8% accuracy) and reducing DDoS attacks employing line-of-defences to prevent resource exhaustion is proposed by Chanu et al. [17]. The feature selection process reduces the dataset's size, which reduces the chance of an attack.

Feinstein et al. [18] highlighted the detection accuracy and performance after examining real traffic traces from diverse network environments and estimating entropy as well as frequency-sorted distributions of chosen packet properties. The findings demonstrate that these strategies can be efficient against current threats and provide suggestions for enhancing detection of future attacks that are more complex.

3   Intrusion Detection and Prevention

Software called the intrusion detection system (IDS) keeps an eye on and guards against hostile activity on a network. Intrusion detection is the technique of identifying illegal entry and intrusions into computer networks and information systems. While external intruders seek to access networks and/or information systems from outside the network, internal intruders are legitimate users who attempt to elevate privileges in order to obtain unauthorised data or services. IDS is made up of two parts: a reporting system and a sensor. Data collection is the main goal of the sensors.

According to the authors' claims, IDS ideas for IoT can be categorised depending on the classes of attacks that can be identified and the kinds of attacks that can be detected. Many authors have asserted that IoT systems may also be vulnerable to security weaknesses associated with traditional technologies and middleware used to build the IoT ecosystem, such as insecure HTTP connections as well as malicious code injection. IDS strategies for IoT can be separated into two groups: We focus on the techniques to detect DoS attacks and techniques to identify routing attacks. Their findings indicate that conventional attacks and man-in-the-middle attacks are both possible risks.

As noted in Smys et al. [19], due to a lack of security and an increase in intruders, intrusion detection systems are now a crucial processing component in modern wireless networks. IoT networks must create an intrusion detection system in order to prevent performance degradation because they are heterogeneous networks that face serious security threats comparable to those faced by wireless networks. Research work that was proposed analysed the various IoT attacks and suggested a hybrid convolutional neural network module that included a lengthy short term memory mechanism. The proposed model achieves a greater detection accuracy of 98% when compared to traditional recurrent neural networks through experimental verification, making the application appropriate for various IoT contexts.

IDSs can be several types, such as host-based IDS (HIDS) as well as network-based IDS (NIDS), depending upon the way they are used. While an NIDS analyses and monitors data for abnormal network traffic, a HIDS functions on a particular computing device and monitors doubtful or malicious software components or unidentified malicious programmes that affect its operating system. In addition, based on how IDS describe them, abuse or signature-based and anomaly-based network problems can be separated into two categories. During attacks, an IDS that is misuse- or signature-based searches for compromised systems utilising signatures and patterns such the network traffic byte sequence.

IDS is generally required at the communication level in to keep track of network activities as well as communication links and also to produce an alarm in the event of any anomaly, such as when a predefined policy is disregarded. WSNs or the conventional Internet are commonly taken into consideration in classic IDS techniques, as noted by Mosenia and Jha [20]. IDSs are able to identify malicious nodes that attempt to introduce false information into the system or go against the rules. IDS-based solutions to the injection problem have been offered in a number of recent research projects. For instance, Son et al. [21] have proposed a novel tool that accurately and effectively identifies code injection attacks on servers.

4   Conclusion and Future Work

IoT security is a current research area that is becoming more prevalent in governmental, industrial, and scientific research. The creation and implementation of IoT-based systems changing information is being carried out by numerous international

organisations and multinational enterprises. Designers must overcome a lot of obstacles to offer a wide range of reliable services, particularly in security-related research topics. Extensive research projects are now examining potential dangers and developing defences. This survey summarises numerous IoT security flaws and mitigation strategies on a level-by-level basis. The primary purpose of this paper is to give the reader the opportunity to study which attacks were successfully launched, and how they were prevented, and which dangers remain the most serious concerns.

References

1. Kumar PAR, Selvakumar S (2011) Distributed denial of service attack detection using an ensemble of neural classifier. Comput Commun 34(11):1328–1341
2. Tuan TA, Long HV, Son LH et al (2020) Performance evaluation of Botnet DDoS attack detection using machine learning. Evol Intel 13:283–294. https://doi.org/10.1007/s12065-019-00310-w
3. Mohsin MA, Hamad AH (2022) Implementation of entropy-based DDoS attack detection method in different SDN topologies. Am Acad Sci Res J Eng Technol Scis 86(1):63–76
4. Goel AK, Rose A, Gaur J, Bhushan B (2019) Attacks, countermeasures and security paradigms in IoT. In: 2019 2nd international conference on intelligent computing, instrumentation and control technologies (ICICICT), vol 1. IEEE, pp 875–880
5. Gupta N, Naik V, Sengupta S (2017) A firewall for internet of things. In: 2017 9th international conference on communication systems and networks (COMSNETS). IEEE, pp 411–412
6. Al-Hadhrami Y, Hussain FK (2021) DDoS attacks in IoT networks: a comprehensive systematic literature review. World Wide Web 24(3):971–1001. https://doi.org/10.1007/s11280-020-00855-2
7. Malhotra P, Singh Y, Anand P, Bangotra DK, Singh PK, Hong WC (2021) Internet of things: evolution, concerns and security challenges. Sensors 21(5):1809
8. Yang Z, Liu X, Li T, Wu D, Wang J, Zhao Y, Han H (2022) A systematic literature review of methods and datasets for anomaly-based network intrusion detection. Comput Secur 102675
9. Singh KJ, Thongam K, De T (2018) Detection and differentiation of application layer DDoS attack from flash events using fuzzy-GA computation. IET Inf Secur 12(6):502–512
10. Saeys Y, Abeel T, Peer YVD (2008) Robust feature selection using ensemble feature selection techniques. In: Joint European conference on machine learning and knowledge discovery in databases. Springer, Berlin, Heidelberg, pp 313–325
11. Kshirsagar D, Kumar S (2022) Towards an intrusion detection system for detecting web attacks based on an ensemble of filter feature selection techniques. Cyber-Phys Syst 1–16
12. Singh KJ, De T (2017) Analysis of application layer DDoS attack detection parameters using statistical classifiers. Internetworking Indonesia 9(2):23–31
13. Johnson Singh K, Thongam K, De T (2016) Entropy-based application layer DDoS attack detection using artificial neural networks. Entropy 18(10):350
14. Johnson Singh K, De T (2017) Mathematical modelling of DDoS attack and detection using correlation. J Cyber Secur Technol 1(3–4):175–186
15. Sajith PJ, Nagarajan G (2022) Network intrusion detection system using ANFIS classifier. Soft Comput. https://doi.org/10.1007/s00500-022-06854-x
16. Priya Devi A, Johnson Singh K (2021) A machine learning approach to intrusion detection system using UNSW-NB-15 and CICDDoS2019 datasets. In: Smart computing techniques and applications. Springer, Singapore, pp 195–205
17. Chanu US, Singh KJ, Chanu YJ, An ensemble method for feature selection and an integrated approach for mitigation of distributed denial of service attacks. Concurren Comput: Practice Experience e6919

18. Feinstein L, Schnackenberg D, Balupari R, Kindred D (2003) Statistical approaches to DDoS attack detection and response. In: Proceedings DARPA information survivability conference and exposition, vol 1. IEEE, pp 303–314
19. Smys S, Basar A, Wang H (2020) Hybrid intrusion detection system for internet of things (IoT). J ISMAC 2(04):190–199
20. Mosenia A, Jha NK (2017) A comprehensive study of security of internet-of-things. IEEE Trans Emerg Topics Comput 5(4):586–602. https://doi.org/10.1109/TETC.2016.2606384
21. Son S, McKinley KS, Shmatikov V (2013) Diglossia: detecting code injection attacks with precision and efficiency. In: Proceedings on ACM SIGSAC conference computer and communications security, pp 1181–1192

A Matrix Factorization Algorithm for Movie Recommendation

Disinlung Kamei and Khundrakpam Johnson Singh

1 Introduction

A recommender system (RS) is a system or technique that tells us the highest possible rating value that a user will likely give to an item and accordingly suggest an item that is relevant to the user [1]. Because of its potential to extract useful information from the vast amount of data available [2] many companies like Amazon [3], Movie Lens [4], Youtube, etc. have adopted RS of their own so as to enhance their user's experience and make things more personalized thereby achieving many customer's loyalty [5]. For such a company being able to recommend items that suit a user's personal taste makes a huge upliftment in maintaining the relationship between the company and the user.

Most of the RS adopts two strategies content filtering and collaborative filtering (CF) [5]. Simply speaking content filtering tries to find the similarity among items through their description which is explicitly created. For example, say items such as movies, we find the similarity among movies through its explicitly created description such as plot, genre or crew involved. And in CF, it tries to find the similarity among items or users through the interaction that the user or item has in the past with the RS. Example a rating that is given by the user on an item in the past. A RS improves its recommendation by looking at its own database in which it stores implicit or explicit feedback that a user has given to those items recommended by it.

The CF suffers from the cold start problem [5] despite its high accuracy rate as compared to content-based because it relies on the past or previous information. Example [6] a newly enrolled user in a movie RS will not rate any movies, therefore

D. Kamei (✉) · K. J. Singh
Department of CSE, National Institute of Technology, Manipur, India
e-mail: disin9856@gmail.com

K. J. Singh
e-mail: johnkh34@gmail.com

it will be tough to find similarities with current users to produce a recommendation for him. Thus the alternative is to use the content filtering as it has no cold start problem despite its low accuracy rate. Thus in this paper, we use both content and CF so that it solves the cold start and accuracy problem where we mainly use the particular Singular Value Decomposition (SVD) matrix factorization technique which was popularized by Simon funk in CF for predicting the rating of a given user.

2 Literature Survey

In 1990 CF technique was introduced since then the RS has become an important field of research [7–9]. Tapestry [10] is one of the earliest RS that was built successfully using the CF it was design to filter large volumes of mails and get the sets of document of interest to a particular user. It relies on the feedback of users. Lyle et al. [11] suggested a CF technique that uses clustering method, a variation of K-mean clustering and Gibbs sampling, claiming that reliable prediction can be made if people can be grouped into cluster with similar movie and putting movies together in groups that are likely to appeal to the same audience. They classify movies and user into various type of classes and these classes were found using their statistical method. GroupLens [12] system is another RS where CF was used. It helps people to find news article of their interest from the stream of available news. It makes use of a rating server to collect and distribute ratings. It forecasts ratings based on the heuristic that those who agreed previously will most likely agree again.

In movie RS it mainly adopts the CF method and it turns out to be very effective. In this paper, the CF method such as the user-based and item-based [13] techniques, here the system tries to find the similarity among the user by observing the user's past history of interaction with the item such as the rating given by the user to the item, two such users who have the same rating for the same item are likely to have the same preference in the upcoming future likewise it operates entirely on the user and item space to search for its nearest neighbours and the model-based CF in this paper is our RS that is build using K-nearest neighbour (KNN), KNN with mean SVD and SVD++ they predict the rating that user will give to an item it does so by first creating a model that stores rating designs in the light of client rating information base which can manage the versatility and scarcity issue.

3 Methodology

The dataset we use for our task is the MovieLens dataset. Four main files in.csv format namely movies_metadata.csv, links.csv, rating_small and keywords, were used the data set consists of 100 k ratings from 700 users over 9000 movies. The movie_medata.csv contains columns/features (fields) shown in Table 1.

Table 1 Features of Movies_Metadata.Csv

Field	Type	Description
Title	String	It tell us the name of the movie
Overview	String	Describe the plot of the movie
ImdbId	Integer	Gives us the Id of Imdb movie
Genre	String	It tell us the genre of the movie
Runtime	Integer	It tell us the length of the movie in minute
Year	Integer	It tell us the year of release
Vote_count	Integer	Total number of votes earned by a movie
Vote_average	Float	Average value of the total no. of votes earned by a movie

A. *Simple RS*

First, we build a simple recommender system using the Internet Movie Database (IMDb) weighted rating Eq. (1) as our rating matric

$$\text{Weighted Rating (\textbf{WR})} = \left(\frac{v}{v + m} \times R \right) + \left(\frac{m}{v + m} \times c \right) \tag{1}$$

where $v =$ Total no. of votes garnered/earned by the movie, $m =$ the minimum number of votes required for the movie to be recommended, $R =$ mean rating of the movie and $C =$ mean rating of all the movies in the dataset. Here v is given by the field 'vote_count' and R value by the field 'vote_average'. Here we made the value of m to be the number of votes earned by 70% of the movie hence a movie is suitable for ranking only when its total earned votes is more than or equal to m.

- *Algorithm for simple recommender*

    Step 1. We load in the movie dataset.
    Step 2. Amongst the vote_count of every movie, we find out the one where 70% of the vote_count falls below it and we set that vote_count as the value of m.
    Step 3. We calculate the mean of all the vote_average of each and every movie and set that value as C.
    Step 4. For i = 1 till total no. of movies

       If (movies[i].vote_count > = m)

          Take the movie[i].vote_count and make it v.
          Take the movie[i].vote_average and make it R.

       Using the Eq. (1) we calculated the weighted rating of the movie[i].

- **Step 5**. Sort the movie according to the weighted rating in descending order.

We choose the IMDb equation as our rating matric (score) for the case study because the advantage of using the IMDb formula is that movie A rated 9 out of 10 by 100,000

users will be considered a lower quality movie than movie B rated 9.5 by only 100 users because the average rating of B is higher than A, this problem is resolved if we use the IMDb formula. However, this very primitive RS does not take into account the user interest hence we go on to build a knowledge-based RS.

B. *Knowledge-based RS*

In this system, we take into account the user's input information, which are the genre, duration(runtime) and timeline(year), and use the IMDB formula to get a high weighted rating on the movie that satisfies the given input criteria from the user.

- *Algorithm for Knowledge-based RS*

  Step 1. Input user preferences, i.e. input the genre of the movie user want to watch, input the shortest duration of the movie, input the longest duration of the movie, input the earliest year and input the latest year of the movie that the user wants to watch.

  Step 2. Extract all the movies that match the condition set by the user in step 1.

  Step 3. We then rate those movies obtained from step2 using the IMDb Eq. (1) which we used in the simple RS and obtain the best-rated movie.

  Step. 4. We sort the best-rated movie obtained at step 3 in descending order and get the top 10 of them and recommend it to the user as shown in Fig. 1.

The problem with the knowledge-based RS is that it fails to capture the user's interest like what kind of movie a person will like, and fails to suggest a movie that is related or similar to the movie that the user has inputted. The quick solution to it is to ask the user for more metadata input. Say two users are asked to give the genre

```
preferred genre
action
give min duration
60
give max duration
150
year in between->starting at
2000
year in between ->ending at
2010
```

Out[33]:

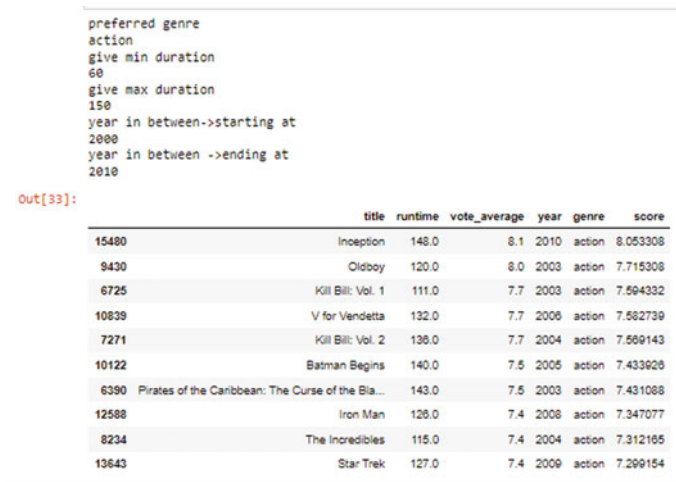|	title	runtime	vote_average	year	genre	score
15480	Inception	148.0	8.1	2010	action	8.053308
9430	Oldboy	120.0	8.0	2003	action	7.715308
6725	Kill Bill: Vol. 1	111.0	7.7	2003	action	7.594332
10839	V for Vendetta	132.0	7.7	2006	action	7.582739
7271	Kill Bill: Vol. 2	136.0	7.7	2004	action	7.569143
10122	Batman Begins	140.0	7.5	2005	action	7.433926
6390	Pirates of the Caribbean: The Curse of the Bla...	143.0	7.5	2003	action	7.431088
12588	Iron Man	126.0	7.4	2008	action	7.347077
8234	The Incredibles	115.0	7.4	2004	action	7.312165
13643	Star Trek	127.0	7.4	2009	action	7.299154

Fig. 1 Top ten movies recommended by knowledge-based recommender

of the movie and they inputted comedy as the genre, then they are further asked for a subgenre that is whether it is a family comedy or a romantic comedy, likewise, we can find the interest of the person's interest.

However, it becomes cumbersome and uninteresting for the user as each and every time a user is asked to give metadata resulting in annoyance and a bad user experience, hence solving the problem of the knowledge-based RS. We build the content-based RS. There are two types First content-based RS is the plot-based RS, which is built by using the corpus taken from the description of a given movie and the second one is the metadata-based RS, in this system, it takes into account the information such as the genre of the movies the keywords of the movie which highlight the subgenres of the movie and finally the information about the cast and crew involved.

C. *Content-based RS*

Here we build two types of content-based recommenders system plot description-based and metadata-based.

- *Plot description-based RS*

For building a plot-based RS, we use the field 'overview' Fig. 2. To create the Term frequency-Inverse document frequency (TF-IDF) matrix. Using the Eq. (2)

here we consider each movie's overview as a document, it consists of English stop words, punctuations, upper case letters, and in some cases 'overview' of the movie is not given thus we fill it with an empty string, we then remove all the English stop words and punctuations, convert the upper case letter to lower case latter,

$$w_{i,j} = tf_{i,j} \times \log\left(\frac{N}{df_i}\right) \tag{2}$$

where wi,j is the weight of word i in document j

where tf is term frequency it tells us the no. of times a words occur in a document (overview of a movie)

dfi (document frequency) is the number of documents (overview of a movie) that contain the term/words i

Fig. 2 Overview of the movie

	title	overview
0	Toy Story	Led by Woody, Andy's toys live happily in his ...
1	Jumanji	When siblings Judy and Peter discover an encha...
2	Grumpier Old Men	A family wedding reignites the ancient feud be...
3	Waiting to Exhale	Cheated on, mistreated and stepped on, the wom...
4	Father of the Bride Part II	Just when George Banks has recovered from his ...
5	Heat	Obsessive master thief, Neil McCauley leads a ...
6	Sabrina	An ugly duckling having undergone a remarkable...
7	Tom and Huck	A mischievous young boy, Tom Sawyer, witnesses...
8	Sudden Death	International action superstar Jean Claude Van...
9	GoldenEye	James Bond must unmask the mysterious head of ...

```
                  ancient andy discover encha family feud happily judy \
Toy Story              0    1        0     0      0    0       1    0
Jumanji                0    0        1     1      0    0       0    1
Grumpier Old Men       1    0        0     0      1    1       0    0

                  led live peter reignites siblings toys wedding woody
Toy Story           1    1     0         0        0    1       0     1
Jumanji             0    0     1         0        1    0       0     0
Grumpier Old Men    0    0     0         1        0    0       1     0
```

Fig. 3 *idf* of the first three movie 'overview'

N is the total number of documents (overview of a movie)

and $\log(N/df_i)$ gives us the *idf_i* (inverse document frequency), it tells us the weight to each of the word/term *i*

Hence the *tf-idf* matrix for the three movies '*toy story*', '*Jumanji*' and '*Grumpier*' Old men would be, to first create our *tf* (term frequency) matrix for the three movies as shown in Fig. 3.

Next we do our multiplication of TF matrix with IDF respectively and obtain the *Tfidf* matrix Fig. 4

Next, we calculate the movie-movie cosine similarity matrix using Eq. (3) between each and every movie's vector of TF-IDF matrix.

$$Cosine(\mathbf{x}, \mathbf{y}) = \frac{x \cdot y^T}{\|x\| \cdot \|y\|} \tag{3}$$

Here, in Fig. 4 'Toy story's review(document x) will be the entire 1st row and the Jumanji's review(document y) is the entire 2nd row and both are a 1×16 row vector, from the Eq. 3. We find the similarity value between each and every movie's review (document) and obtain the movie-movie similarity matrix in Fig. 5.

```
                  ancient      andy discover     encha   family      feud \
Toy Story        0.000000  0.477121 0.000000  0.000000 0.000000  0.000000
Jumanji          0.000000  0.000000 0.477121  0.477121 0.000000  0.000000
Grumpier Old Men 0.477121  0.000000 0.000000  0.000000 0.477121  0.477121

                  happily      judy      led      live    peter reignites \
Toy Story        0.477121  0.000000 0.477121  0.477121 0.000000  0.000000
Jumanji          0.000000  0.477121 0.000000  0.000000 0.477121  0.000000
Grumpier Old Men 0.000000  0.000000 0.000000  0.000000 0.000000  0.477121

                 siblings      toys  wedding woody
Toy Story        0.000000  0.477121 0.000000   0.0
Jumanji          0.477121  0.000000 0.000000   0.0
Grumpier Old Men 0.000000  0.000000 0.477121   0.0
```

Fig. 4 TF-IDF matrix

Fig. 5 Movie-movie similarity matrix

	Toy Story	Jumanji	Grumpier Old Men
Toy Story	1.0	0.0	0.0
Jumanji	0.0	1.0	0.0
Grumpier Old Men	0.0	0.0	1.0

D. *Collaborative Filtering method*

In this method given a user and a movie, we are to predict the rating that the user would give to that movie. It does so by learning from other similar users who have rated the movie.

Two type of CF are going to be build, user-based or item-based CF and model-based CF for our movie recommender system.

Here we are going to use the rating.csv files of the IMDB database of 100 k ratings with 700 users and 9000 movies (Table 2).

- *User-based collaborative filtering RS*

In this system we are given a user u and movie, m the system is to predict the rating that is highly likely to be given by the user (even if the user have not rated the movie). Before we build our system that can make such prediction we have to create a user-user similarity matrix using Table 3

From Table 4 we create the user-movie matrix A, (here each cell presence the rating given by the user on the movie). Then we derive the user-user similarity matrix by performing cosine similarity between the matrix A and itself as

$cosine(A,A) = (A.A^T)/\|A\|.\|A\|=$			
	u_1	u_5	u_{10}
u_1	$(u_1.u_1^T)/\|u_1\|.\|u_1\|$	$(u_1.u_5^T)/\|u_1\|.\|u_5\|$	$(u_1.u_{10}^T)/\|u_1\|.\|u_{10}\|$
U_5	$(u_5.u_1^T)/\|u_5\|.\|u_1\|$	$(u_5.u_5^T)/\|u_5\|.\|u_5\|$	$(u_5.u_{10}^T)/\|u_1\|.\|u_{10}\|$
U_{10}	$(u_{10}.u_1^T)/\|u_1\|.\|u_{10}\|$	$(u_{10}.u_5^T)/\|u_{10}\|.\|u_5\|$	$(u_{10}.u_{10}^T)/\|u_{10}\|.\|u_{10}\|$

Table 2 Features of rating.csv

Fields	Description
userId	The id of the users
movieId	The movie Id given by tmdb
rating	The rating that the user have given to a movie its value ranges from 1 to 5

Table 3 User–movie rating

UserId	MovieId	Rating
u_1	m_1	2.5
u_1	m_3	3
u_1	m_7	3
u_5	m_3	2
u_5	m_9	4
u_{10}	m_1	2
u_{10}	m_7	2
u_{10}	m_9	2.5

where user u1 is a 1×4 row vector with values [2.5,3,3,0] and user u10 is also a 1×4 row vector with values [2,0,2,3.5] calculating their cosine similarity we get as 0.496391 likewise we find the cosine similarity values between each and every users and obtain the user-user cosine similarity. Figure 6 is the cosine similarity value. Their values lie between -1 and 1, the higher the value between the two user means the more they are similar to each other.

- **_Item-based collaborative RS_**

We create the movie-user pivot matrix Table 5 from Table 4

In this movie-user matrix, B its row represents the movie and column represents the user and each cell represents the rating given by the user. We find the movie-movie similarity matrix by performing $cosine(B, B) = (B.B^T)/\|B\|.\|B\|$ just like how we did for the user-user similarity matrix in the user-user content-based RS and get the movie-movie similarity matrix as shown in Fig. 7.

- **_KNN model_**

Users are classified into classes/clusters in KNN, and only users from the same cluster are considered when rating predictions are made. It is almost the same as that of the above two collaborative RSs, i.e. item-based and user-based but the only difference here is that we take into consideration only the K-nearest neighbouring users (if we

Table 4 User-movie rating matrix A

	m_1	m_3	m_7	m_9
u_1	2.5	3	3	0
u_5	0	2	0	4
u_{10}	2	0	2	3.5

Fig. 6 User-user similarity matrix

	u1	u5	u10
u1	1.000000	0.272446	0.496391
u5	0.272446	1.000000	0.695666
u10	0.496391	0.695666	1.000000

Table 5 Movie-user matrix B

	u_1	u_5	u_{10}
m_1	2.5	0	2
m_3	3	2	0
m_7	3	0	2
m_9	0	4	3.5

Fig. 7 Movie-movie similarity

	m1	m3	m7	m9
m1	1.000000	0.649722	0.996241	0.411365
m3	0.649722	1.000000	0.692308	0.417454
m7	0.996241	0.692308	1.000000	0.365273
m9	0.411365	0.417454	0.365273	1.000000

use the user-user similarity matrix Fig. 6) or neighbouring item (if we use the movie-movie similarity matrix Fig. 7) here we consider the user-user similarity matrix for KNN-based CF, here we compute the rating given by individual user u on the movie m using the Eq. (4)

$$\hat{r}_{um} = \frac{\sum\limits_{v \in N_m^k(u)} \text{sim}(u, v) \cdot r_{vm}}{\sum\limits_{v \in N_m^k(u)} \text{sim}(u, v)} \tag{4}$$

where $Nm^k(u)$ is the set of k users that are neighbours to user u who have rated movie m, and sim(u,v) is the cosine similarity between user u and v as given by Fig. 6, rv,m is the rating that the user v gives on the movie m. For the 100 k rating dataset it gives the accuracy rate *RMSE* of 0.9918 and time taken is 4.7157 s.

- *SVD model*

It is a matrix factorization technique where the given user-movie rating matrix, Anxm is represented as the dot product of two matrices Pu and Qm, i.e. $An \times m = Pu.Q_i^T$ where Pu is a user-latent factor matrix of n row and Qi is a movie-latent factor matrix of m rows. In this particular version of the SVD matrix factorization technique which was popularized by Simon Funk [14] during the Netflix Prize the predicted rating of each user u over movie i is given by using Eq. (6)

$$\hat{r}_{u,m} = \mu + b_u + b_m + q_m^T p_u \tag{6}$$

where pu is the features of item m and qm denotes the features of movie i and $\mu + bu + bm$ is the baseline estimates
    μ = overall mean rating of movie.
    bu = (average rating of u)-μ.
    bm = (average rating of movie m)-μ.
    for the given 100 k rating dataset it gives the accuracy rate RMSE of 0.898 and the time taken is 2.914 s.

- *Hybrid RS*

It is combination of content-based RS and collaborative-based RS.

Algorithm for Hybrid RS

Given a user id *u*, and movie id *m* were commend list of top 10 movie that is similar to movie *m* and which will also be liked by user *u*.

> **Step 1**. We load in the movie-movie similarity matrix Fig. 5 that we derived at content-based recommender system.
> **Step 2**. Next we load in the predicted rating matrix that our SVD model-based CF have predicted for every user and movie.
> **Step 3**. Get the list of movies that is similar to our input movie *m* using the similarity matrix of step 1.
> **Step 4**. Sort the list of similar movie obtained at step 3 in descending order using similarity values.
> **Step 5**. We compute the predicted rating for the movies obtained at step 4 by the user *u* using the predicted rating matrix of step 2.
> **Step 6**. Sort the movies in descending order using the predicted rating values given to each movie and take the top 20 movies for a recommendation.

4 Experimental Result and Discussion

In this paper our experiment is performed to evaluate our method based on the MoveLens dataset in terms of time complexity and accuracy, our dataset contains 100 k ratings from 700 users over 9000 movies and it has the feature user ID, movie ID and the rating feature. Each user have rated at least 20 ratings to the movie with values ranging from 1 to 5, we have split our rating dataset into 75% as a training dataset and 25% as our test dataset, and in our experiment, the metrics for checking the accuracy of our model is root mean square error (RMSE) given by Eq. (8)

$$\text{RMSE} = \sqrt{\frac{1}{|R|} \sum_{r_{\text{um}}^{\wedge} \in R} \left(r_{\text{um}} - r_{\text{um}}^{\wedge}\right)^2} \tag{8}$$

where $\hat{R}$ is the testing dataset and the *ru,m* is the actual rating given by user u on movie *m* and the $\hat{r}_{u,m}\hat{r}u,m$ is the predicted rating of that movie *m* by the user *u*,the smaller the *RMSE* value means higher the accuracy. For the recommender that uses the KNN-based model, we have set our k = 40. We then obtain our time consumption and *RMSE* value for each of the four model-based CFs as shown in the table.

From the above Table 6, we can see that the svd++ has the most accurate prediction rate however in terms of time consumption, and it is the most time-consuming. KNN is the lowest in accuracy rate. We also see that the SVD is the least time-consuming model with a fairly accurate prediction rate. And we also saw by observing our RS that uses the user-based and item-based CF that item-based is better in explaining the reason behind the prediction made. Thus we move on to build the hybrid RS that uses SVD and item-item similarity matrix.

Table 6 Time and RMSE comparison

Model	KNN	KNN with mean	SVD	SVD + +
Time(sec) 4.7157	5.364		2.914	49.575
RMSE	0.9918	0.9246	0.898	0.890

5 Conclusion

We observed that SVD is the best in terms of accuracy and time complexity for making predictions of the rating, and the use of item-item similarity matrices turns out to be more reliable than the usage of user-user similarity matrices. So, through RMSE and time consumption analysis of our proposed method (of using SVD as a rating prediction model and use of movie-movie similarity matrix), the performance of our recommender has improved a lot and can solve the cold starting problem and reduces time and space complexity problem.

References

1. Ricci F, Rokach L, Shapira B (2011) Introduction to recommender systems handbook. Recommender systems handbook. Springer, Boston, MA, pp 1–35
2. Kumar M, Yadav D K, Singh A et al (2015) A movie recommender system: Movrec[J]. Int J Comput Appl 124(3)
3. Linden G, Smith B, York J (2003) Amazon.com recommendations: item-to-item collaborative filtering. IEEE Internet Comput 7(1):76–80
4. Miller, Bradley N et al (2003) Movielens unplugged: experiences with an occasionally connected recommender system. In: Proceedings of the 8th international conference on intelligent user interfaces.
5. Koren Y, Bell R, Volinsky C (2009) Matrix factorization techniques for recommender systems. Computer 42(8):30–37
6. Xiong L et al (2012) A novel nearest neighborhood algorithm for recommender systems. In: 2012 Third global congress on intelligent systems. IEEE
7. Hill W et al (1995) Recommending and evaluating choices in a virtual community of use. In: Proceedings of the SIGCHI conference on Human factors in computing systems
8. Resnick P et al (1994) Grouplens: An open architecture for collaborative filtering of netnews. In: Proceedings of the 1994 ACM conference on computer supported cooperative work
9. Shardanand U, Maes P (1995) Social information filtering: algorithms for automating "word of mouth. In: Proceedings of the SIGCHI conference on Human factors in computing systems
10. Herlocker JL et al (1999) An algorithmic framework for performing collaborative filtering. In: Proceedings of the 22nd annual international ACM SIGIR conference on Research and development in information retrieval
11. Sarwar B et al (2001) Item-based collaborative filtering recommendation algorithms. In: Proceedings of the 10th international conference on World Wide Web
12. MovieLensDataset. https://files.grouplens.org/datasets/movielens/ml-latest-small.zip (Accessed on 8 Jan 2022)
13. Takács G et al (2007) Major components of the gravity recommendation system. AcmSigkdd Explor Newsletter 9(2):80–83
14. Funk S (2006) Netflix update: try this at home. https://sifter.org/simon/journal/20061211.html (Accesssed on 12 Oct 2021)

15. Goldberg D et al (1992) Using collaborative filtering to weave an information tapestry. Commun ACM 35(12):61–70
16. Ungar, Lyle H, Foster DP (1998) Clustering methods for collaborative filtering. AAAI workshop on recommendation systems 1
17. Bell RM, Koren Y (2007) Scalable collaborative filtering with jointly derived neighborhood interpolation weights. In: Seventh IEEE international conference on data mining (ICDM 2007), IEEE