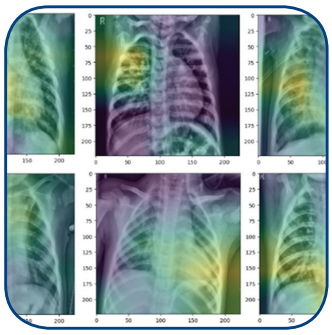
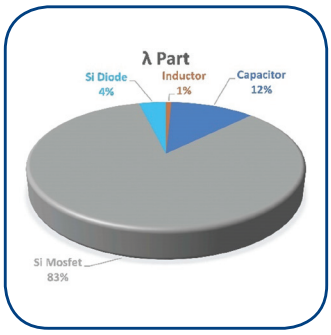
Volume 15, Number 4, 2024 ISSN 1847-6996

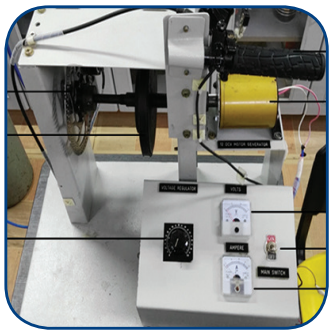
FERIT
FACULTY OF ELECTRICAL ENGINEERING, COMPUTER SCIENCE AND INFORMATION TECHNOLOGY **OSIJEK**

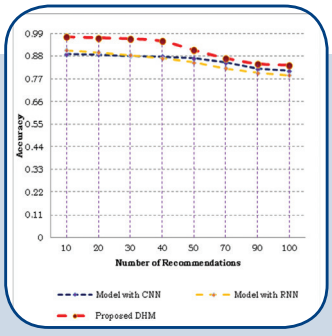
IJECES International Journal of Electrical and Computer Engineering Systems

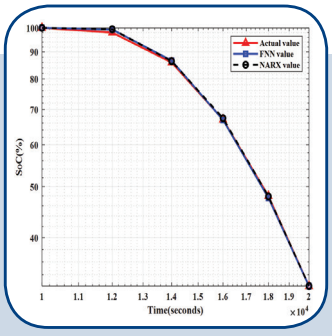
International Journal of Electrical and Computer Engineering Systems

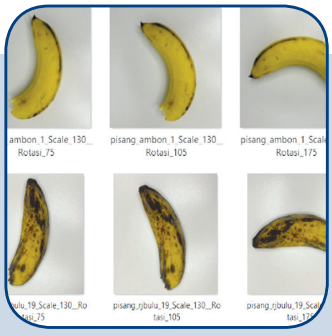












ISSN: 1847-6996

INTERNATIONAL JOURNAL OF ELECTRICAL AND COMPUTER ENGINEERING SYSTEMS

Published by Faculty of Electrical Engineering, Computer Science and Information Technology Osijek, Josip Juraj Strossmayer University of Osijek, Croatia

Osijek, Croatia | Volume 15, Number 4, 2024 | Pages 305 - 386

The International Journal of Electrical and Computer Engineering Systems is published with the financial support of the Ministry of Science and Education of the Republic of Croatia

CONTACT

International Journal of Electrical and Computer Engineering Systems (IJECES)

Faculty of Electrical Engineering, Computer Science and Information Technology Osijek, Josip Juraj Strossmayer University of Osijek, Croatia
Kneza Trpimira 2b, 31000 Osijek, Croatia
Phone: +38531224600, Fax: +38531224605
e-mail: ijeces@ferit.hr

Subscription Information

The annual subscription rate is 50€ for individuals, 25€ for students and 150€ for libraries.
Giro account: 2390001 - 1100016777, Croatian Postal Bank

Journal is referred in:

- Scopus
- Web of Science Core Collection (Emerging Sources Citation Index - ESCI)
- Google Scholar
- CiteFactor
- Genamics
- Hrčak
- Ulrichweb
- Reaxys
- Embase
- Engineering Village

Bibliographic Information

Commenced in 2010.
ISSN: 1847-6996
e-ISSN: 1847-7003
Published: quarterly
Circulation: 300

IJECES online

https://ijeces.ferit.hr

TABLE OF CONTENTS

Measurement of State of Charge of Lithium-Nickel Manganese Cobalt Battery using Artificial Neural Network and NARX Algorithm **305**
Original Scientific Paper
Divya. R | Karunanithi. K | Ramesh. S | Raja. S.P

Boosting Reliability: A Comparative Study of Silicon Carbide (Sic) and Silicon (Si) in Boost Converter Design Using MIL-HDBK-217 Standards **313**
Original Scientific Paper
Elaid Bouchetob | Bouchra Nadji

Design of Regenerative Braking System and Energy Storage with Supercapacitors as Energy Buffers **321**
Original Scientific Paper
Siluvai M. Michael | Bokani Mtengi | S.R.S Prabaharan | Adamu Murtala Zungeru | James Garba Ambafi

Optimizing Enhanced Extended Topological Active Net Model Using Parallel Processing **335**
Original Scientific Paper
Ranjita Akash Asati | M.M. Raghuwanshi | K.R. Singh

DHM-OCR: A Deep Hybrid Model for Online Course Recommendation and Sustainable Development of Education **345**
Original Scientific Paper
Sagar Mekala | Padma Tns | T Rama Rao

Comparative Analysis of Banana Detection Models: Deep Learning and Darknet Algorithm **367**
Original Scientific Paper
Abdul Haris Rangkuti | Varyl Hasbi Athala | Sian Lun Lau | Rudi Aryanto

Gray Level Co-occurrence Matrix based Fully Convolutional Neural Network Model for Pneumonia Detection **376**
Original Scientific Paper
Shubhra Prakash | B Ramamurthy

Data-driven Gait based Severity Classification for Parkinson's Disease using Duo Spatiotemporal Convoluted Kernel Boosted ResNet model **386**
Original Scientific Paper
Arogia Victor Paul M | Sharmila Shankar

About this Journal
IJECES Copyright Transfer Form

Measurement of State of Charge of Lithium-Nickel Manganese Cobalt Battery using Artificial Neural Network and NARX Algorithm

Original Scientific Paper

Divya. R
Research Scholar, School of Electrical and Communication Engineering,
Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology,
Chennai, Tamil Nadu, India.
divyarajendran.kr@gmail.com

Karunanithi. K
Professor, School of Electrical and Communication Engineering,
Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology,
Chennai, Tamil Nadu, India.
k.karunanithiklu@gmail.com

Ramesh. S
Professor, School of Electrical and Communication Engineering,
Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology,
Chennai, Tamil Nadu, India.
rameshsme@gmail.com

Raja. S.P
Associate Professor, School of Computer Science and Engineering,
Vellore Institute of Technology,
Vellore, Tamil Nadu, India.
avemariaraja@gmail.com

Abstract *– The battery's SoC is a crucial variable since it reflects its performance. An accurate estimation of SoC protects the battery, prevents overcharging or discharge, and extends its life time. Since most of the traditional methods use complex equations, ANN has been implemented to reduce the complications and provide better accuracy. In this research, Li-NMC with capacity rating of 2000mAh is used for the estimation of SoC. In this paper, Feedforward Neural Network (FNN) algorithm and Nonlinear Auto-Regressive network with exogenous inputs (NARX) have been used for designing a neural network model. Here, the performance matrixes of both neural network models have been compared and analyzed with the same dataset.*

Keywords: *ANN, SoC estimation, FNN algorithm, NARX algorithm, Li-NMC battery*

Received: November 9, 2023; Received in revised form: January 9, 2024; Accepted: January 19, 2024

1. INTRODUCTION

According to European Green Deal, Commission has boosted its rules by setting forth essential policies [1] to attain net-zero global warming emission by 2050. Due to the development of industries, there is a considerable increase of greenhouse effects and emission of carbon [2]. The transportation is biggest source of greenhouse gas emissions globally [3]. As a result, Electric Vehicles (EVs) must be introduced into the transportation industry [4, 5]. In Electric Vehicles (EVs), Lithium-ion battery is mainly used due to its longevity [6]. The life time, safety and charging capability need to be enhanced inorder to improve the performance of the Li-ion battery [7].

The SoC of the battery denotes the available capacity as the function of rated capacity. The value of the SoC varies from 0%-100% [8]. SoC is indirectly assessed using proxies like temperature, potential difference, and capacity [9]. Accurate prediction of SoC is a vital feature in a cell pack utilized in EV's [10]. Electric vehicles require accurate cell SoC prediction for safe and effective operation thereby extending battery life [11]. Estimating battery SoC poses challenges like non-linear battery behavior, model complexity, calibration needs, and limited observability. The research paper covers a literature survey in Section 2, data preparation and collection in Section 3, ANN architecture in Section 4, FNN model in Section 5, NARX model in Section 6, and concludes with results and discussion in Section 7. Section 8 provides the work's conclusion.

2. LITERATURE SURVEY

There is no clear and concise method for calculating the SoC accurately. A Li-ion cell at 100% SoC has all cyclable lithium ions in the negative electrode, while at 0% SoC, they are all in the positive electrode [12].

In [13] the SoC prediction techniques are classified into four groups such as model-based, ampere-hour, open circuit voltage (OCV) and data-driven prediction methods. According to the author [14], the Open Cir-

cuit Voltage (OCV) of Li-ion is a critical measure for analyzing changes and estimating the SoC. The charging state slope, measured offline at different temperatures and aging stages, is prone to errors in the OCV-SoC relationship due to operational condition variations [15].

The Ampere-hour is a simple and convenient method used to evaluate the SoC [16]. In [17] the author proposed improved coulomb counting approach with compensation coefficient in order to reduce the error. The author [18] suggested a novel capacity prediction technique on enhanced coulomb counting process and the error is about 1.7%. In [19], the drawback of this SoC calculation method is that a considerable estimation inaccuracy can result from an incorrect initial battery current.

According to the model-based estimation principle, the estimating process cannot manage the inaccuracies from the system model [20]. In [21], the author proposed SoC prediction using HIF and Extended Kalman Filter (EKF). SoC prediction can approximate to the precise value in 30 seconds while maintaining 0.5% efficiency. According to author [22], using Kalman filter incorrect parameters decrease the battery model's accuracy which results in an increase in SoC estimation error. In [23], the author proposed a sliding observer approach that relies on a variable adaptable system model, which had a precision of less than 2%. To provide a rapid prediction model for cell charging state and impedance, a multi-level PI observer is used [24]. In [25] the author suggested a GRU-RNN is used for an accurate SoC estimation. In [26], the author develops a neural network-based BMS (NN-BMS) for a through-the-road hybrid electric vehicle (TtR HEV), with an emphasis on the TtR HEV's recharging capacity. The machine learning approach, which includes Artificial Neural Networks (ANN), is also known as the data-driven approach [27]. An ANN is a mathematical framework made up of a series of independent processing units called neurons that are connected by weights [28].

In this paper, data driven approach (also known as black box model) is used for an accurate SoC prediction. Since this method requires minimum knowledge and time for modeling a system comparing with other methods. The nntool (Neural Network Toolbox) in MATLAB offers a combination of user-friendly interface, extensive functionality, flexibility, integration with other MATLAB tools, and a supportive community, making it advantageous for designing and implementing neural network models. The problem statement includes an accurate estimation of SoC since it is very important factor that should be measured accurately to protect the battery life. Comparison analysis of various SoC measurements is shown in Table 1.

Table 1. Comparison analysis of various SoC measurements

Slno	SoC measurement	Merits	Demerits
1	Circuit based model	Accurately representing physical systems, enabling simulations and versatility.	Accurate parameters are needed, potential simplifications, and limitation to specific types of systems.
2	Neural network model	Parallel processing leads to faster training, adaptability, flexibility, automatic feature learning.	Large data requirements, takes significan-t time for training and lack of transpare-ncy.
3	Pseudo-two-Dimensional(P2D) Model	Computa-tional accuracy, easier parameteridentification.	Assumption and approximtion, simplified geometry, applicatio-n, specific.
4	Single Particle model	Computa-tional efficient, quick sensitivity analysis, conceput-al clarity.	Lack of spatial information, limited applicability to fuel cell, sensitivity to particle size.

3. DATA COLLECTION AND PREPARATION

The input dataset has been gathered from the Battery Research Group of the Center for Advanced Life Cycle Engineering (CALCE) [29]. The feature extraction used in this paper consist of voltage, current, charge capacity, discharge capacity, test time, step time, change in voltage with respect to time (dv/dt), charge energy, discharge energy and internal resistance. Table2 denotes number of sample Data used in Training, Validation and Testing.

Table 2. Data used for training, Validation and Testing

Data	No of Samples
training	11814
validation	3938
testing	3938

4. FEEDFORWARD NEURAL NETWORK (FNN)

A FNN is made up of layers, each of which is made up of neurons. The input layer takes the data to be analyzed and feeds it to one or more hidden layers, that perform the categorization function, before sending it to the output layer [30]. Many neurons make up a FNN, which is also the fundamental unit of information processing [31]. Weights connect each neuron, resulting in probability-weighted correlations among source and result [32]. Fig. 1 shows the architecture of FNN.

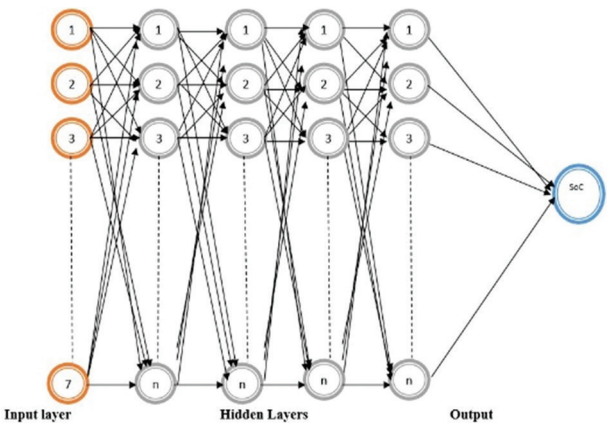


Fig. 1. FNN architecture

Each neuron has activation layer and pre-activation layer. Where, activation layer is denoted as 'h' and spre-activation layer is represented as 'a'. Equation (2) and (3) represents the matrix of the weight 'W_1'and Activation function 'h_1' for first layer respectively. Pre-activation function is 'a_{1n}'for first layer for n neurons.

$$W_1 = \begin{bmatrix} w_{111} & \cdots & w_{11k} \\ \vdots & \ddots & \vdots \\ w_{1j1} & \cdots & w_{1jk} \end{bmatrix} \quad (1)$$

$$h_1 = \begin{matrix} h_{11} \\ h_{12} \\ \vdots \end{matrix} \quad (2)$$

$$= w_1 j_1 x_1 + w_1 j_2 x_2 + w_1 j_3 x_3 + w_1 j_4 x_4 \ldots\ldots + w_1 j_k b_1 \quad (3)$$

At the output layer 'L,' the activation function is provided by,

$$\hat{y} = f(x) = h_L = o(a_L) \quad (4)$$

Where, o = output activation function.

In this paper, 4-layer FNN has been designed inorder to obtain less MSE. Levenberg-Marquardt learning function and GD transfer function has been used. The reason for choosing GD and Levenberg-Marquardt is fast convergence, adaptive learning rate, simplicity, parallelization and requires less memory compared to more complex optimization algorithms so it is used for large dataset. The computational time is 1000 epochs. The MSE at 1000 Epochs obtained is 1.231e^{-06}. This is the better performance obtained using the proposed FNN model. At 681 epochs, 0.00069379 MSEREG is obtained as better performance using the proposed FNN model. In this, 0.0074991 is obtained as SSE using 1000 Epochs. Comparing with MSE and MSEREG, this SSD error is high.

5. NARX MODEL

The NARX technique improves learning performance and computing efficiency in addressing battery nonlinearity. Its predicted output is consistently validated against the true value, enhancing accuracy in time series forecasts by storing both input and previous output values as feedback [33]. Fig. 2 depicts the NARX model architecture.

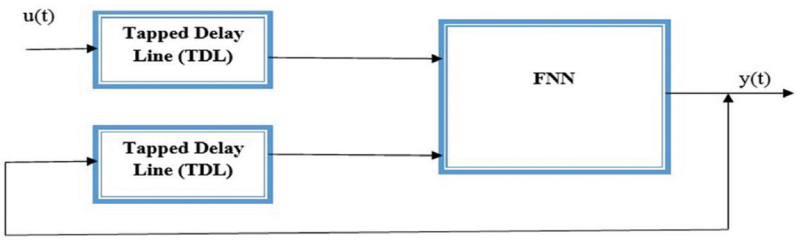


Fig. 2. NARX architecture

The mathematical formula for NARX model provided by the equation 5,

$$y(t) = f(y(t-1), y(t-2), \ldots, y(1-n_y), u(t-1), u(t-2), \ldots, u(t-n_u)) \quad (5)$$

Where, f= non-linear function, n_u= maximal lags input, n_y= output to the model respectively.

In proposed model, 7 inputs and 1 output are designed. In this model, the result of output layer is given as the feedback to the hidden layer. This feedback makes the difference between FNN model and NARX model. This is done to compare and analyze both model results. The best validation performance of MSE is 9.3021e^{-06}. The best validation performance using MSEREG is 8.2739e^{-05}. The SSE obtained is 0.01233 at 1000 Epochs. Comparing with MSE and MSEREG, SSE error is high. Table 3 shows the best validation performance of FNN and NARX Algorithm for MSE, MSEREG and SSE.

Table 3. Best validation performance of FNN and NARX Algorithm

Performance Matrices	FNN Algorithm	NARX Algorithm
MSE	1.231e^{-06}	9.3021e^{-06}
MSEREG	0.00069379	8.2739e^{-05}
SSE	0.0074991	0.01233

6. RESULTS AND DISCUSSION

To verify the trained models, MSE, MSEREG and SSE have been selected as model's performance indicator. In the fig. 3(a), predicted SoC value of FNN and NARX and actual SoC is plotted with the voltage(V). At initial voltage 3.44V, both the actual and predicted FNN SoC value is 100% while the predicted NARX SoC value is 99%. From 3.44V to 4.19V, the actual and predicted FNN value is 100%. The SoC value of both actual and predicted FNN is reduced from voltage 4.12V to 2.57V and SoC of 32.38% is reached for both actual value and FNN predicted value. The SoC value of NARX gets reduced from the voltage 4.19V to 2.57V and a SoC value of 32.39% is obtained.

In the fig. 3(b), actual, MSEREG predicted FNN and NARX SoC value is plotted with voltage. Initially, both actual and predicted FNN values are 100%, while the NARX prediction stands at 99.99% for an initial voltage of 3.44V. In the voltage range of 3.44V to 4.19V, FNN maintains a 100% SoC. Subsequently, the SoC decreases from 4.12V to 2.57V, reaching 32.38% for both actual and predicted FNN values. For the NARX model, the SoC prediction remains at 99% from 3.44V to 3.52V and reaches 100% from 3.52V to 4.19V. The NARX prediction decreases from 4.12V to 2.57V, resulting in a SoC of 32.40%.

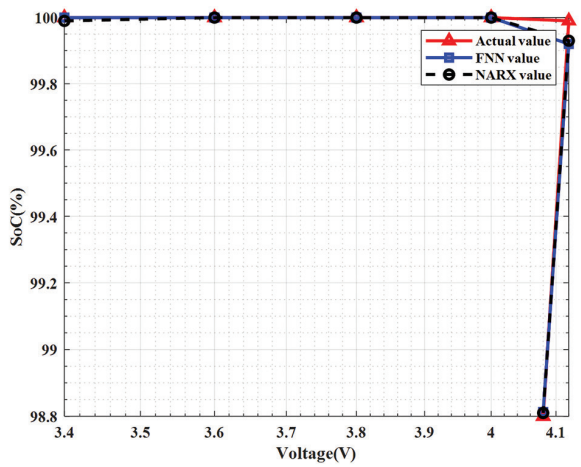


Fig. 3(a). MSE value of FNN, NARX, Actual value Vs Voltage

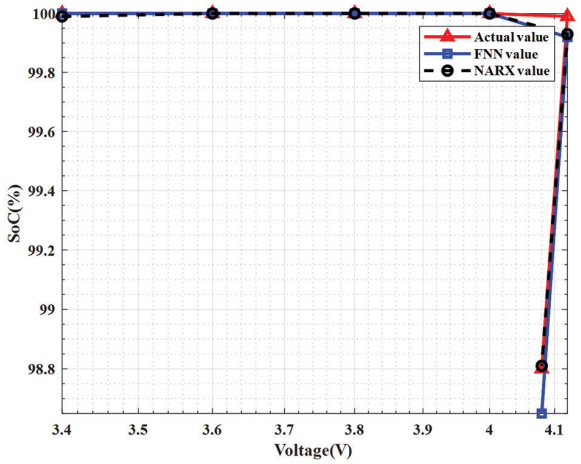


Fig. 3(b). MSEREG value of FNN, NARX, Actual value Vs Voltage

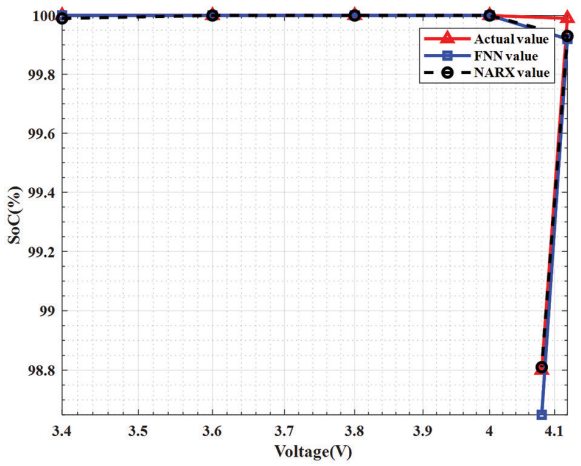


Fig. 3(c). SSE value of FNN, NARX and Actual value Vs Voltage

In the above fig. 3(c), SSE predicted value for FNN & NARX and actual value is plotted with the voltage. At initial voltage 3.44V, both the actual and predicted FNN SoC value is 100% whereas the predicted NARX SoC value is 99%. From the voltage range 3.44V to 4.19V, the actual SoC value is 100%. Whereas from 344V to 4.11V, the predicted FNN value is 100%. NARX SoC of 99% is obtained from the voltage 3.44V to 3.44V. During this period, the current and charge capacity is increased. Finally, SoC of 32.38% is obtained. From 3.44V to 4.19V, the actual SoC value is 100%.

In Fig. 4(a), the actual value, predicted FNN and NARX SoC value is plotted with respect to test time. At initial test time 1.12 seconds, the voltage is 3.44V. From 1.002seconds to 1.25seconds, the actual SoC and FNN value is 100%. However, from time period 1.002 seconds to 64.03 seconds, the NARX value is 99%. At 65.04 seconds, the NARX value is 100%and it is continued till time period 1.22 seconds. The NARX SoC value begin to reduce from 12293.32 seconds to 1.98 seconds and SoC value of 32.39% is obtained. From 1.12 seconds to 1.25 seconds, the actual & FNN value is of 100% while from 1.25 seconds to 1.98 seconds, the actual & FNN value is reduced and 32.38% of SoC is obtained.

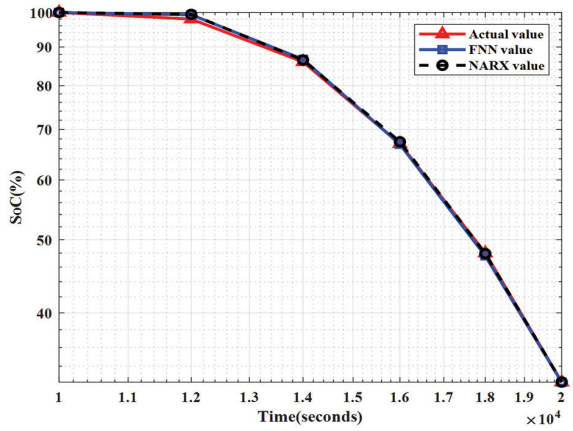


Fig. 4(a). MSE value of FNN, NARX and Actual SoC value Vs Test time

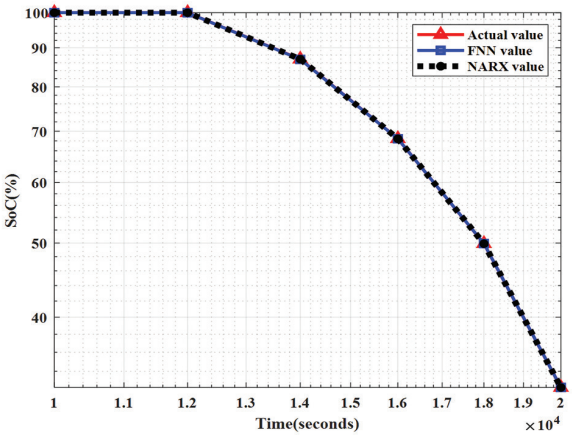


Fig. 4(b). MSEREG value of FNN, NARX and Actual SoC value Vs Test time

In Fig. 4(b), the actual, predicted FNN and NARX SoC value is plotted with respect to test time. At initial test time 1.002 seconds, the actual & predicted FNN SoC is 100% and NARX value is 99%. From 1.002 seconds to 1.25 seconds, the actual & FNN value is 100%. In NARX,

the SoC value is 99% from 1.002 seconds to 60.00 seconds. From 61.0017seconds to 1.25.15seconds, the NARX SoC value is 100%. During this time period, the current is 0A. From 1.25seconds to 1.98seconds, the actual & FNN SoC value is reduced and 32.38% is obtained. From the time period 1.25seconds to 1.98seconds, the NARX value is reduced and SoC of 32.40% is obtained. From 1.07seconds to 1.25seconds, the current reaches 0A whereas the current value gets reversed from the time period 1.25seconds to 1.98seconds.

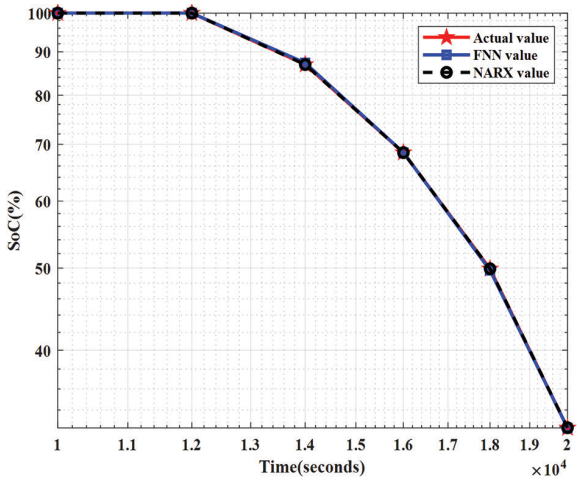


Fig. 4(c). SSE value of FNN, NARX and Actual SoC value Vs Test time

In the above Fig. 4(c), the actual value, predicted FNN and NARX SoC value is plotted with the test time. At initial time period 1.002 sec, both actual and predicted FNN SoC value is 100%. Whereas, NARX SoC value is 99.9%. From the time period 1.002166seconds to 12593.15seconds, the actual SoC value is 100%. This actual value decreases from 1.25 to 1.98seconds and SoC of 32.38% is obtained. From 1.002166seconds to 1.26seconds, the predicted FNN value is 100%. This value starts decreasing from the time period 1.26seconds to 1.98seconds and SoC of 32.38% is obtained. Similarly, from the time period 1.002 seconds to 64.03 seconds, the NARX value is 99%. From 65.04 seconds to 1.25seconds, the SoC value is 100%. This SoC value of NARX begins to reduce from 1.25seconds to 1.98seconds and 32.38% is obtained.

Table 4 shows comparison table based on different neural network algorithm with the existing model. In this table, most of the algorithms are compared with respect to performance matrices and input parameters. Considering MSE value, FNN model has better performance accuracy while NARX is less comparing with FNN model. In MSEREG and SSE, NARX model has better accuracy than FNN model. Even though nntool provide accurate estimation it has some limitations such as dependency on MATLAB (i.e it works on MATLAB platform alone), slow code execution, limited scalability, lack of advanced deep learning features.

Table 4. Error comparison of different Neural Network Algorithms.

Algorithm	Training function/learning Function/Transfer function	Input parameter	Error Rate
Multilayer Perceptron [34]	Levenberg Marquardt/ Gradient Descent/Hyperbolic Tangent Sigmoid	Time, Current, Voltage	$3.11x10^{-6}$
Feed forward [35]	Gradient Descent/ sigmoid activation function.	Current, voltage	>1%
Feed forward [36]	Gradient Descent	Current, voltage, time	>2%
feed-forward [37]	Levenberg–Marquardt,	Current, voltage	0.025
LSTM [38]	Levenberg–Marquardt	Current, voltage, temperature	>2%
LSTM-RNN [39]	-	Voltage, current	2.088%MSE 2.44% RMSE
FNN and NARX	Levenberg–Marquardt/GD	Current, voltage, charge capacity, discharge capacity, dv/dt, DOD, test time	FNN=$1.231e^{-06}$ MSE, 0.00069379 MSEREG, 0.0074991 SSE NARX=$9.3021e^{-06}$ MSE, $8.2739\ e^{-05}$ MSEREG, 0.01233 SSE

7. CONCLUSION

In this paper, a 4-layer FNN and recurrent NARX neural network have been designed and sigmoid is used as activation function. The performance matrix MSE of FNN is found to be $1.231e^{-06}$ and NARX is $9.3021e^{-06}$. Similarly, MSEREG and SSE of NARX model has better accuracy than FNN model. Finally, an accurate performance indicator for SOC prediction of Li-NMC battery has been found employing nntool MATLAB2021b. The future work is to design a neural network with many numbers of neurons and hidden layers. Different type of batteries can be trained and a comparative study can be done with respect to performance matrices.

8. REFERENCES:

[1] S. Barja-Martinez, F. Rücker, M. Aragüés-Peñalba, R. Villafafila-Robles, Í. Munné-Collado, P. Lloret-Gallego, "A novel hybrid home energy management system considering electricity cost and greenhouse gas emissions minimization", IEEE Transactions on Industry Applications, Vol. 57, No. 3, 2021, pp. 2782-2790.

[2] M. Mohsin, H. W. Kamran, M. A. Nawaz, M. S. Hussain, A. S. Dahri, "Assessing the impact of transition

from nonrenewable to renewable energy consumption on economic growth-environmental nexus from developing Asian economies", Journal of Environmental Management, Vol. 284, 2021, p. 111999.

[3] M. Liu, X. Zhang, M. Zhang, Y. Feng, Y. Liu, J. Wen, L. Liu, "Influencing factors of carbon emissions in transportation industry based on CD function and LMDI decomposition model: China as an example", Environmental Impact Assessment Review, Vol. 90, 2021, p. 106623.

[4] T. Teoh, O. Kunze, C. C. Teo, Y. D. Wong, "Decarbonisation of urban freight transport using electric vehicles and opportunity charging", Sustainability, Vol. 10, No. 9, 2018, p. 3258.

[5] B. V. Ayodele, S. I. Mustapa, "Life cycle cost assessment of electric vehicles: A review and bibliometric analysis", Sustainability, Vol. 12, No. 6, 2020, p. 2387.

[6] K. Ogura, M. L. Kolhe, "Battery technologies for electric vehicles", Electric Vehicles: Prospects and Challenges, Elsevier, 2017, pp. 139-167.

[7] R. Schmuch, R. Wagner, G. Hörpel, T. Placke, M. Winter, "Performance and cost of materials for lithium-based rechargeable automotive batteries", Nature Energy, Vol. 3, No. 4, 2018, pp. 267-278.

[8] H. Abdi, B. Mohammadi-ivatloo, S. Javadi, A. R. Khodaei, E. Dehnavi, "Energy storage systems", Distributed Generation Systems, Vol. 7, 2017, pp. 333-368.

[9] M. H. Lipu, A. Hussain, M. H. M. Saad, A. Ayob, M. A. Hannan, "Improved recurrent NARX neural network model for state of charge estimation of lithium-ion battery using pso algorithm", Proceedings of the IEEE Symposium on Computer Applications & Industrial Electronics, Penang, Malaysia, 28-29 April 2018, pp. 354-359.

[10] R. Zhang, B. Xia, B. Li, L. Cao, Y. Lai, W. Zheng, W. Wang, "State of the art of lithium-ion battery SOC estimation for electrical vehicles", Energies, Vol. 11, No. 7, 2018, p. 1820.

[11] M. A. Hannan, M. H. Lipu, A. Hussain, P. J. Ker, T. I. Mahlia, M. Mansor, Z. Y. Dong, "Toward enhanced state of charge estimation of lithium-ion batteries using optimized machine learning techniques", Scientific Reports, Vol. 10, No. 1, 2020, p. 4687.

[12] B. Sundén, "Hydrogen, batteries and fuel cells", Academic Press, 2019.

[13] X. Wu, M. Li, J. Du, F. Hu, "SOC prediction method based on battery pack aging and consistency deviation of thermoelectric characteristics", Energy Reports, Vol. 8, 2022, pp. 2262-2272.

[14] R. Zhang, B. Xia, B. Li, L. Cao, Y. Lai, W. Zheng, M. Wang, "A study on the open circuit voltage and state of charge characterization of high-capacity lithium-ion battery under different temperature", Energies, Vol. 11, No. 9, 2018, p. 2408.

[15] R. Xiong, Q. Yu, "Open circuit voltage and state of charge online estimation for lithium-ion batteries", Energy Procedia, Vol. 142, 2017, pp. 1902-1907.

[16] L. Zhao, M. Lin, Y. Chen, "Least-squares based coulomb counting method and its application for state-of-charge (SOC) estimation in electric vehicles", International Journal of Energy Research, Vol. 40, No. 10, 2016, pp. 1389-1399.

[17] L. He, D. Guo, "An improved coulomb counting approach based on numerical iteration for SOC estimation with real-time error correction ability", IEEE Access, Vol. 7, 2019, pp. 74274-74282.

[18] C. Youssef, D. Omar, G. Ahmed, E. Fatima, E. S. Najia, "Design and simulation of an accurate neural network state-of-charge estimator for lithium-ion battery pack", International Review of Automatic Control, Vol. 10, No. 2, 2017, pp. 186-192.

[19] M. U. Ali, A. Zafar, S. H. Nengroo, S. Hussain, M. J. Alvi, H. J. Kim, "Towards a smarter battery management system for electric vehicle applications: A critical review of lithium-ion battery state of charge estimation", Energies, Vol. 12, No. 3, 2019, p. 446.

[20] J. Meng, D.I. Stroe, M. Ricco, G. Luo, R. Teodorescu, "A simplified model-based state-of-charge estimation approach for lithium-ion battery with dynamic linear model", IEEE Transactions on Industrial Electronics, Vol. 66, No. 10, 2018, pp. 7717-7727.

[21] C. Chen, R. Xiong, W. Shen, "A lithium-ion battery-in-the-loop approach to test and validate multi-scale dual H infinity filters for state-of-charge and

capacity estimation", IEEE Transactions on Power Electronics, Vol. 33, No. 1, 2017, pp. 332-342.

[22] J. Xing, P. Wu, "State of charge estimation of lithium-ion battery based on improved adaptive unscented Kalman filter", Sustainability, Vol. 13, No. 9, 2021, p. 5046.

[23] B. Ning, J. Xu, B. Cao, B. Wang, G. Xu, "A sliding mode observer SOC estimation method based on parameter adaptive battery model", Energy Procedia, Vol. 88, 2016, pp. 619-626.

[24] T. Chen, M. Huo, X. Yang, R. Wen, "A fast lithium-ion battery impedance and SOC estimation method based on two-stage PI observer", World Electric Vehicle Journal, Vol. 12, No. 3, 2021, p. 108.

[25] S. El Fallah, J. Kharbach, Z. Hammouch, A. Rezzouk, M. O. Jamil, "State of charge estimation of an electric vehicle's battery using Deep Neural Networks: Simulation and experimental results", Journal of Energy Storage, Vol. 62, 2023, p. 106904.

[26] Y. Ko, K. Cho, M. Kim, W. Choi, "A novel capacity estimation method for the lithium batteries using the enhanced coulomb counting method with kalman filtering", IEEE Access, Vol. 10, 2022, pp. 38793-38801.

[27] M. Charkhgard, M. Farrokhi, "State-of-charge estimation for lithium-ion batteries using neural networks and EKF", IEEE Transactions on Industrial Electronics, Vol. 57, No. 12, 2010, pp. 4178-4187.

[28] E. J. Miriam, S. Sekar, S. Ambalavanan, "Artificial Neural Network technique for predicting the lifetime and performance of lead-acid battery", International Journal ESTIJ, Vol. 3, 2013, pp. 393-401.

[29] Battery Data, https://calce.umd.edu/battery-data (accessed: 2023)

[30] X. Li, J. Sun, "Facial emotion recognition via stationary wavelet entropy and particle swarm optimization", Cognitive Systems and Signal Processing in Image Processing, Academic Press, 2022, pp. 145-162.

[31] A. Das, P. K. Kumawat, N. D. Chaturvedi, "A Study to Target Energy Consumption in Wastewater Treatment Plant using Machine Learning Algorithms", Computer Aided Chemical Engineering, Vol. 50, 2021, pp. 1511-1516.

[32] D. Darbar, I. Bhattacharya, "Application of machine learning in battery: state of charge estimation using feed forward neural network for sodium-ion battery", Electrochem, Vol. 3, No. 1, 2022, pp. 42-57.

[33] M. G. Sani, N. Abdul Wahab, Y. M. Sam, S. I. Samsudin, I. W. Jamaludin, "Comparison of NARX neural network and classical modelling approaches", Applied Mechanics and Materials, Vol. 554, 2014, pp. 360-365.

[34] C. Youssef, D. Omar, G. Ahmed, E. Fatima, E. S. Najia, "Design and simulation of an accurate neural network state-of-charge estimator for lithium-ion battery pack", International Review of Automatic Control, Vol. 10, No. 2, 2017, pp. 186-192.

[35] C. Chen, R. Xiong, R. Yang, W. Shen, F. Sun, "State-of-charge estimation of lithium-ion battery using an improved neural network model and extended Kalman filter", Journal of Cleaner Production, Vol. 234, 2019, pp. 1153-1164.

[36] S. Tong, J. H. Lacap, J. W. Park, "Battery state of charge estimation using a load-classifying neural network", Journal of Energy Storage, Vol. 7, 2016, pp. 236-243.

[37] F. Yang, X. Song, F. Xu, K. L. Tsui, "State-of-charge estimation of lithium-ion batteries via long short-term memory network", IEEE Access, Vol. 7, 2019, pp. 53792-53799.

[38] E. Chemali, P. J. Kollmeyer, M. Preindl, R. Ahmed, A. Emadi, "Long short-term memory networks for accurate state-of-charge estimation of Li-ion batteries", IEEE Transactions on Industrial Electronics, Vol. 65, No 8, 2017, pp. 6730-6739.

[39] M. A. Hannan, M. S. H. Lipu, A. Hussain, M. H. Saad, A. Ayob, "Neural network approach for estimating state of charge of lithium-ion battery using backtracking search algorithm", IEEE Access, Vol. 6, 2018, pp. 10069-10079.

Boosting Reliability: A Comparative Study of Silicon Carbide (Sic) and Silicon (Si) in Boost Converter Design Using MIL-HDBK-217 Standards

Original Scientific Paper

Elaid Bouchetob
Faculté des hydrocarbures et de la chimie, Laboratoire d'électrification des entreprise industrielles, LREEI
Université de M'hamed Bougara
Boumerdes, Algeria
e.bouchetob@univ-boumerdes.dz

Bouchra Nadji
Faculté des hydrocarbures et de la chimie, Laboratoire d'électrification des entreprise industrielles, LREEI
Université de M'hamed Bougara
Boumerdes, Algeria
b.nadji@univ-boumerdes.dz

Abstract *– Reliability is very important in the world of electronic device design and production, particularly in applications where continuous and flawless performance is a necessity. This directs our attention to the boost converter, which forms the foundation of power electronics, renewable energy systems, and electric vehicles. However, as technology progresses, the choice of materials for these converters is a big challenge. For that, in this paper, the impact of using Silicon Carbide (SiC) devices, with their promising material properties, on the reliability of boost converters is presented. Because the results showed that more than 80% of boost converter failures are caused by semiconductors, the use of SiC materials is assessed by determining its reliability using MIL-HDBK-217 standard. In addition, a comparative study with the use of traditional Silicon (Si) is conducted. The results showed that the failure rate of boost converters based on SiC devices reduced from 8.335 failure/10-6h to 6.243 failure/10-6h. This notable shift in failure rates establishes SiC as a pivotal material in the evolution of boost converter technology, offering a compelling solution to address the persistent challenges associated with semiconductor-related failures.*

Keywords: *Boost converter, SiC devices, Schottky diode, Reliability prediction, MIL-HDBK 217*

Received: September 10, 2023; Received in revised form: February 29, 2024; Accepted: February 29, 2024

1. INTRODUCTION

In the dynamic landscape of electronic device design and manufacturing, the pursuit of reliability stands as an indispensable quest, especially in applications where unwavering performance is paramount. This pursuit leads us to boost converters, a very important element in modern technology which elevates modest input voltages to remarkable heights. This type of static converters form the backbone of power electronics, renewable energy systems, and the electric vehicles propelling us toward a sustainable future [1].

At its essence, the boost converter epitomizes the ingenuity of electronic engineering, orchestrating precise voltage transformations integral to powering a spectrum of electronic devices, ranging from portable gadgets to renewable energy sources. However, as technological progress persists, the material selection for these boost converters assumes heightened significance [2].

Silicon Carbide (SiC), characterized by its promising attributes, emerges as a substantial contender challenging the conventional dominance of traditional Silicon (Si). SiC presents a suite of superior material properties, encompassing a wide bandgap, outstanding thermal conductivity, and notable resistance to elevated temperatures [2]. These characteristics lay the groundwork for a prospective revolution in the design of boost converters, offering the prospect of improved efficiency and performance.

Researchers are actively pursuing investigations into the reliability of wide band gap semiconductors to assess their performance and stability under diverse operating conditions. This involves studying factors such as temperature fluctuations [2], electrical stress [3], and other environmental variables that can impact the long-term functionality of electronic devices based on these materials [4]. However, as the industry moves toward the use of SiC devices, it becomes imperative to compre-

hensively assess and predict their reliability in real-world applications. Reliability prediction is a critical aspect for designing and deploying power electronics systems, as it aids in understanding the potential failure modes, estimating component lifetimes, and ultimately ensuring the longevity of PV systems [5]. At 650V power conversion, various technologies, including SiC MOSFETs, GaN HEMTs, and silicon devices, offer distinct advantages, from high-speed GaN to SiC's temperature resilience and silicon's cost-effectiveness. In the paper [1] authors contrast MIL-HDBK-217 and FIDES methods, contributing to the collective understanding of electronic system dependability. In [6] authors examined their performance, body diodes, and reliability, shedding light on this competitive field. In Algeria, despite government support for PV system development, there is a notable absence of research labs focusing on PV system reliability. Authors in [4] studied the reliability of PV system MPPT and simulated the integration of a photovoltaic panel with a Disrupts & Regards-controlled DC/DC converter, considering various choppers and power level. The paper [2] investigated the reliability of driven design in power electronic circuits, with a specific emphasis on power loss in switches and diodes. It presents a reliability evaluation method for a buck converter using thermal analysis of an IGBT and a diode by switching frequency and duty cycle. The mean time to failure for the buck converter was calculated.

Yet, as with any technological leap, questions arise. What implications does the adoption of SiC in boost converters have on their reliability? How do we navigate this pivotal juncture in materials science and engineering to ensure that SiC-fortified boost converters stand as unwavering pillars of dependability?

To answer these questions, in this paper, we delve into the ramifications of integrating Silicon Carbide (SiC) devices, renowned for their favorable material properties, on the reliability of boost converters. An in-depth analysis of the reliability of SiC materials is conducted, employing the MIL-HDBK-217 standard as the benchmark. Moreover, a comprehensive comparative study is executed, drawing distinctions between the application of SiC and the conventional Silicon (Si) within the context of boost converters. This investigation aims to provide insights into how SiC's distinctive material attributes impact the reliability landscape, particularly when juxtaposed with the established Silicon technology.

2. MATERIAL AND METHOD

2.1. BOOST CONVERTERS AND THEIR IMPORTANCE

Boost converters, although inconspicuous in their physical stature, occupy a paramount position in the world of modern electronics. Their significance lies in their ability to efficiently transform and regulate voltage levels, making them indispensable in a myriad of applications. These unassuming devices act as voltage elevators, taking a lower input voltage and elevating it to a higher output voltage. This fundamental function is pivotal in ensuring the smooth operation of numerous electronic systems and applications. From extending the lifespan of battery-powered devices by providing the required voltage boost to facilitating the integration of renewable energy sources into power grids, boost converters are the unsung heroes behind the scenes, ensuring that our electronic world operates with precision and reliability. In the context of power electronics, renewable energy systems, and electric vehicles, boost converters take center stage, making their reliability and performance critical factors that warrant meticulous examination [7].

Following bellow equations:

$$Vin = Vout(1 - D) \tag{1}$$

V_{in}= input voltage ; V_{out}= output voltage.

And the other component (Capacitor, Inductor) value we follow the equation [8]

$$L \geq V_{in} \frac{D}{\Delta I_L} f_S \tag{2}$$

While:

$$\Delta L \geq \frac{D \cdot V_{in-min}}{f_S \cdot L} \tag{3}$$

ΔI_L= estimate ripple current, $V_{in\text{-}min}$= minimum input voltage, f_s= switch frequency, D= duty cycle, L= inductor.

$$C = \frac{I_{out} \cdot D}{f_S \cdot \Delta V_{out}} \tag{4}$$

C= capacitor, I_{out}= output current, ΔV_{out} = estimate ripple voltage.

We have calculated parameters of the boost converter for 2kW and boosted the voltage from 120 V to 400 V, and switching frequency 25 kHz, the results are presented in the Table 1. For a ripple current of 30% and ripple voltage of 2%.

Using equations (1) to (4) for designing the parameters of boost converter which are presented in Table 1.

Table 1. Parameters of DC-DC boost converter

Parameter	Device specification
Inductor [mH]	0.67
Capacitor [μF]	17.42
Resistor [Ω]	160
Input voltage[V]	120
Output voltage[V]	400
Estimate ripple current [A]	5
Switch frequency[kHz]	25
Duty cycle	0.697
Estimate ripple voltage[V]	8

2.2. SILICON CARBIDE (SiC) VS. SILICON (Si): RESHAPING RELIABILITY IN ELECTRONIC DEVICES

The ongoing comparison between Silicon Carbide (SiC) and Silicon (Si) as semiconductor materials represents a compelling juxtaposition of established incum-

bency versus emerging potential. Silicon, a longstanding cornerstone in the electronics industry, carries a legacy of reliability and widespread adoption. In contrast, Silicon Carbide, positioned as a nascent alternative, possesses distinctive properties and promising attributes that pose a challenge to the existing technological paradigm.

On the other hand, Silicon Carbide, with its wide bandgap and remarkable material properties, presents a tantalizing alternative. SiC exhibits superior thermal conductivity, enabling it to dissipate heat more efficiently, a quality that greatly enhances the reliability of electronic components operating at high temperatures. Additionally, SiC devices can handle higher voltages and power levels, reducing the stress on components and promising extended operational lifespans. These characteristics have propelled Silicon Carbide (SiC) into prominence as a contender challenging the dominance of Silicon.

In the realm of boost converters and beyond, the SiC versus Si comparison stands as a testament to the continual pursuit of enhancement and reliability in electronic systems. As we navigate the ever-evolving landscape of technology, the choice between these two materials has the potential to reshape industries, redefine efficiency, and ultimately dictate the reliability of the electronic devices that underpin our modern world.

Table 2. Mosfets and diodes parameters [9-12]

Manufacture	ROHM	IXYS	CREE	APT
Part number	SCH2080KEC	IXFN 32N120	C4D20120D	APT60D120B
Breakdown rated voltage [V]	1200	1200	1200	1200
Rated current [Amp]	28A @ 100 °C	32A @ 25°C	33A @ 135 °C	60A @ 25 °C
Maximum junction temperature [°C]	175	150	175	150
Gate-source voltage [V]	-10/+20	-30/+30	/	/
Thermal resistance junction case θ_{jc} [°C]	0.44	0.16	0.43	0.66
Diode forward voltage Vf [V]	/	/	1.5	12
Static drain-source on-resistance RDS [Ω]	0.08 Ω	0.35 Ω	/	/

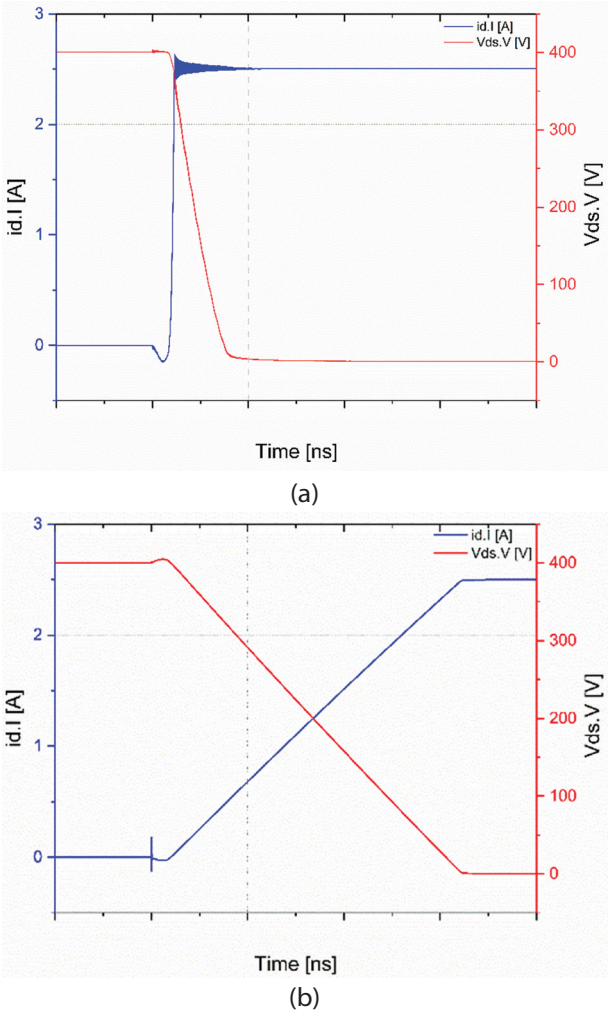


Fig. 1. Turn-ON of Mosfet – SiC (a) vs Si (b)

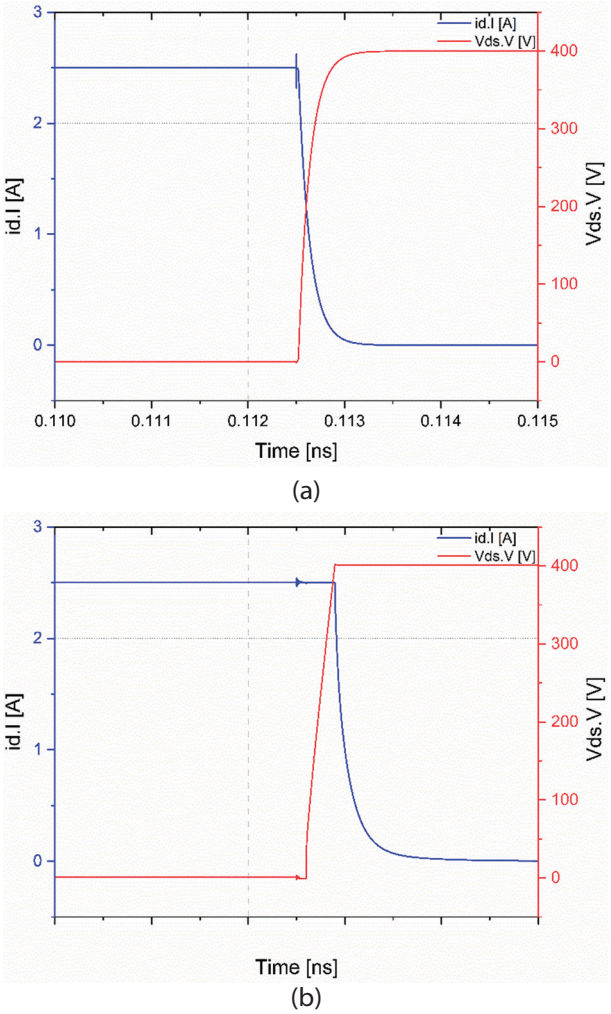


Fig. 2. Turn-OFF of Mosfet – SiC (a) vs Si (b)

The simulation of the switches is made using Ansys simplorer to define the power losses and response time. Silicon Carbide (SiC) Mosfets offer advantages such as quicker turn-on switching, reduced switching losses, and enhanced efficiency due to their rapid switching properties (Fig. 1.a). However, they necessitate gate drivers capable of supplying higher voltage levels. In contrast, Silicon (Si) Mosfets (Fig. 1.b) have slower turn-on speeds, increased switching losses, and lower threshold voltages, rendering them better suited for specific low- to medium-power applications. Choosing between SiC and Si Mosfets hinges on precise application requirements, considering factors like speed, efficiency, and gate driver capabilities.

Regarding turn-off (Fig. 2.a) characteristics, SiC Mosfets exhibit swift switching speeds, enabling rapid deactivation, and they experience lower turn-off losses, which enhances overall efficiency during turn-off transitions. Nonetheless, SiC Mosfets demand gate drivers with the capability for quick switching and precise control. Conversely, Si Mosfets (Fig. 2.b) feature slower turn-off speeds and tend to incur higher turn-off losses, potentially affecting overall efficiency. They typically have lower gate drive requirements in terms of voltage and current. The decision between Silicon Carbide (SiC) and Silicon (Si) MOSFETs for turn-off is a pivotal choice that hinges on the specific requirements of the application. Several critical factors must be carefully considered to optimize performance, efficiency, and overall reliability.

2.3. RELIABILITY PREDICTION USING MIL-HDBK-217 STANDARDS

Reliability prediction using the MIL-HDBK-217 standards represents a cornerstone in the field of electronic engineering. This venerable set of guidelines and methodologies provides a structured approach to assessing and forecasting the reliability of electronic components and systems, ensuring they meet the rigorous demands of various applications.

In the context of boost converters and the comparison between Silicon Carbide (SiC) and Silicon (Si) as semiconductor materials, MIL-HDBK-217 offers a robust framework to navigate this critical juncture. It takes into account an array of factors, including temperature, voltage, and environmental conditions, to estimate the failure rate of electronic components [13].

For boost converters, these standards become a compass, guiding engineers and designers through the labyrinth of material choices. By leveraging MIL-HDBK-217, one can scrutinize how SiC and Si-based boost converters fare under the relentless forces of time and environmental stressors. It enables us to make data-driven decisions, shedding light on which material SiC or Si ultimately yields the highest reliability.

By adopting this time-tested methodology, we initiate an exploration into the intricate synergy among material science, electronic engineering, and reliability. This endeavor holds the potential not only to yield insights but also to provide a systematic framework for fortifying our boost converters against the rigors of the most demanding operational challenges. Continuing to drive innovation and progress across the electronic landscape [14].

Temperature and Stress Levels

Operating temperature is a pivotal factor. SiC's superior thermal conductivity allows it to handle higher temperatures without degradation, potentially extending the operational life of the converter. Si-based devices may experience more significant stress due to elevated temperatures.

Voltage Ratings

The voltage ratings of semiconductor materials impact their reliability. SiC devices are known for their ability to handle higher voltage levels without breakdown, making them more suitable for high-voltage applications.

Environmental Conditions

MIL-HDBK-217 takes into account environmental factors such as humidity, vibration, and thermal cycling. These conditions can affect the reliability of electronic components and SiC's resistance to harsh environments may confer an advantage.

Redundancy and Mitigation Strategies

Reliability predictions should consider the deployment of redundancy and mitigation strategies. This can involve redundancy in critical components or systems and the use of protective features like overcurrent or overvoltage protection.

Testing and Validation

Real-world testing and validation data should complement reliability predictions. Actual performance in the field can provide valuable insights into the reliability of SiC and Si-based boost converters.

Application-Specific Considerations:

The intended application of the boost converter plays a significant role. Different applications may have varying demands and stress factors that influence reliability requirements.

In the subsequent sections, we delve into the results of these reliability predictions, conducting a comprehensive analysis of the SiC-based and Si-based converters' performance and reliability under various conditions and scenarios.

On the other hand, the parts count approach is more conservative and can be used in the early stages of design when limited information is available. The reliability expression R is

$$R(t) = e^{-\lambda .t} R(t) = e^{-\lambda t} \quad (5)$$

λ presents the failure rate and its value is constant, at t the average of $R(t)$ is determined while:

$$t = \frac{1}{\lambda} \quad (6)$$

In this context, λ represents the unchanging rate at which failures occur, and 't' denotes the time at which the initial failure happens. The Mean Time to Failure (MTTF) can be found by calculating the average of R(t), while the Mean Time Between Failure (MTBF) can be computed using the MTTF in conjunction with the Mean Time to Repair (MTTR), according to the following formula:

$$MTBF = MTTR + MTTF = \frac{1}{\lambda} \quad (7)$$

When it comes to particular components like Mosfets, diodes, inductors, and capacitors, determining their failure rates involves considering a range of factors. These factors encompass the construction factor, application factor, environment factor, quality factor, stress factor, and temperature factor, all of which contribute to the overall reliability assessment of these components. (λ_{system}) for the part stress method is:

$$\lambda_{system} = \sum_{i=1}^{N} (\lambda_{part})i \quad (8)$$

Failure rate of Mosfet is calculated by:

$$\lambda_{P(S)} = \lambda_{(b(S)} \cdot \pi_Q . \pi_A . \pi_E . \pi_T) \quad (9)$$

For the diode failure rate:

$$\lambda_{P(D)} = \lambda_{b(D)} \cdot \pi_Q . \pi_C . \pi_E . \pi_T . \pi_S \quad (10)$$

The Inductor failure rate is

$$\lambda_{P(L)} = \lambda_{b(L)} \cdot \pi_Q . \pi_C . \pi_E \quad (11)$$

For the capacitor failure rate:

$$\lambda_{P(D)} = \lambda_{b(C)} \cdot \pi_Q . \pi_{CV} . \pi_E \quad (12)$$

In summary, comprehending the reliability of electronic components holds immense importance in the design and operation of electronic systems. To assess this reliability, we employ methods such as part stress and parts count, which allow us to calculate failure rates accurately. These calculations involve various factors, including the contact construction factor (π_C), application factor (π_A), environment factor (π_E), quality factor (π_Q), stress factor (π_S), and temperature factor (π_T).

Equations (13) and (15) are instrumental in determining the thermal conductivity values of the inductor (λb) and capacitor ($\lambda_{P(C)}$), respectively. Equation (14) aids in finding the hot spot temperature of the inductor (T_{HS}), a critical parameter needed for equation (13). This equation relies on two essential inputs: the ambient operating temperature (T_A) and the temperature rise above ambient (ΔT). Furthermore, equation (15) hinges on the ratio of operating voltage to rated voltage (S) and the capacitance value (C). By leveraging these equations and factors, we can pinpoint the reliability of electronic components, identify key factors impacting system reliability, and propose strategies to enhance reliability while minimizing maintenance costs.

$$\lambda_b = 0.00035 \cdot e^{\left(\frac{T_{HS}+273}{329}\right)^{15.6}} \quad (13)$$

$$T_{HS} = T_A + 1.1 \times \Delta T \quad (14)$$

$$\lambda_b = 0.00254 \left(\left(\frac{S}{0.5}\right)^3 + 1\right) \exp\left(5.09 \times \left(\frac{(TA+273)}{378}\right)^5\right) \quad (15)$$

Equations (16) and (17) come into play when calculating the temperature factor (π_T) for both the switch and diode components. To apply these equations, we first need to ascertain the junction temperature (T_j), a parameter essential for these calculations. Equation (11) assists in determining the junction temperature, requiring inputs such as the heat sink temperature (T_C), the thermal resistance of the switch or diode (θ_{jc}), and the total power loss experienced by the switch or diode (P_{loss}). It's worth noting that the junction temperature is typically established through simulation results, allowing for an accurate assessment of component temperature and, by extension, the temperature factor crucial for reliability calculations.

$$\pi_{T(S)} = \exp\left(-1925 \times \left(\frac{1}{(T_j + 273)} - \frac{1}{298}\right)\right) \quad (16)$$

$$\pi_{T(D)} = \exp\left(-1925 \times \left(\frac{1}{(T_j + 273)} - \frac{1}{293}\right)\right) \quad (17)$$

$$T_j = T_C + \theta_{jc} \times P_{loss} \quad (18)$$

The Equation (19) and (20) present, respectively, the ratio of operating voltage to rated voltage, VS, and the factor π_{CV} for a Capacitance C.

$$\pi_S = V^{2.43} \quad (19)$$

$$\pi_{CV} = 0.34 \times C^{0.18} \quad (20)$$

3. RESULTS AND DISCUSSION

The MIL HDBK-217F standard, in its traditional form, lays the foundation for conducting reliability analysis and predictions. This process entails the incorporation of adjustment factors, with a primary focus on the base failure rate, which acts as a fundamental parameter for modifying the component's failure rate.

Critical data concerning both average and effective power values hold pivotal importance as they serve as the building blocks for calculating losses and subsequently determining the junction temperature of the MOSFET component. Leveraging the Arrhenius model, the temperature factor is fine-tuned, leading to the ultimate calculation of the failure rate for the MOSFET.

For the purpose of this analysis, it is assumed that the ambient temperature remains constant at 27°C.

Caculation of failure rate for main Switch

For SiC Mosfet

$P_{losss} = P_{M\text{-}Con} + P_{M\text{-}SW}$

$P_{M\text{-}Con} = R_{DS}R \cdot I^2_{RMS} = 0.08 \,.\, (16.66)^2 = 22.2$ W

$P_{M\text{-}SW}=(E_{off}+E_{on})\,f_{sw}$
$=(87.2+31.7)10^{-6}\,.\,25.103=2.97$
$P_{loss\text{-}s}=25.17W$
$T_c=27°c$
$T_j=T_c+\theta_{jc}\cdot P_{loss}$
$=38.4$

$$\pi_{T(S)}=\exp\left(-1925\times\left(\frac{1}{(T_j+273)}-\frac{1}{298}\right)\right)$$

$\pi_T=1.311$
$\pi_Q=5.5$, $\pi_E=6$, $\pi_A=10$, $\lambda_b=0.012$
λ_P-(S) $=5.191$

For Si Mosfet

$P_{loss\text{-}s}=P_{M\text{-}Con}+P_{M\text{-}SW}$
$P_{M\text{-}Con}\ Rds*I^2RMS=0.35*(16.66)^2=97.14$ W
$P_{M\text{-}SW}=(E_{off}+E_{on})f_{sw}=(55.2+168.4)\ 10^{-6}*25.103=5.59$ W
$P_{loss\text{-}s}=102.73$ W
$T_c=27°c$
$T_j=T_c+\theta_{jc}*P_{loss\text{-}s}=53.52$

$$\pi_{T(S)}=\exp\left(-1925\times\left(\frac{1}{(T_j+273)}-\frac{1}{298}\right)\right)$$

$\pi_T=1.758$
$\pi_Q=5.5$, $\pi_E=6$, $\pi_A=10$, $\lambda_b=0.012$
λ_P-$(S)$$=6.962$

Calculation of failure rate for output Diode:

Schottky Diode

$P_{loss\text{-}D}=P_{con\text{-}d}+P_{sw\text{-}d}$
$P_{con\text{-}d}=V_{fd}*I_{RMS}=1.8*16.66=29.98$ W
$P_{sw\text{-}d}=Q_c.V_{fsw}=8.10^{-8}.400.25.10^3=0.8$ W
$T_c=27°c$
$T_j=T_c+\theta_{jc}*P_{loss}=40.23°C$

$$\pi_{T(S)}=\exp\left(-1925\times\left(\frac{1}{(T_j+273)}-\frac{1}{298}\right)\right)$$

$\pi_T=1.4511$
$\pi_Q=5.5$, $\pi_E=6$, $\pi_S=0.07$, $\pi_C=1$, $\lambda_b=0.003$
λ_P-(D)$=0.01005$

For Si Diode

$P_{losss\text{-}D}=P_{Con\text{-}D}+P_{SW\text{-}D}$
$P_{con\text{-}d}=V_{fd}+I_{out\text{-}RMS}=2.15\ .16.66=35.81$ W
$P_{sw\text{-}d}=Q_c\cdot V_{out}\cdot f_{sw}=1.88.10^{-6}.400.25.103$
$=18.8$ W
$T_c=27°c$
$T_j=T_c+\theta_{jc}\cdot P_{loss}=63.04$

$$\pi_{T(S)}=\exp\left(-1925\times\left(\frac{1}{(T_j+273)}-\frac{1}{298}\right)\right)$$

$\pi_T=2.077$
$\pi_T=2.077$
$\pi_Q=5.5$, $\pi_E=6$, $\pi_S=0.07$, $\pi_C=1$, $\lambda_b=0.069$
$\lambda_{P\text{-}(D)}=0.3310$

Calculation of failure rate for each of output capacitor

$\pi_{CV}=0.34\cdot C^{0.18}=0.34\ (17.42)^{0.18}=0.5687$
$S=400/600=0.667$, $T_A=27°C$

$$\lambda_b=0.00254\left(\left(\frac{S}{0.5}\right)^3+1\right)\exp\left(5.09\times\left(\frac{(TA+273)}{378}\right)^5\right)$$

$\lambda_b=0.043$
$\pi_Q=10$, $\pi_E=4$
$\lambda_{P\text{-}(C)}=0.043.4.10.0.5687=0.9781$

Calculation of failure rate for each of input inductor

$T_{HS}=T_A+1.1\times\Delta T$

$$\lambda_b=0.00035\times\exp\left(\frac{(T_{HS}+273)}{329}\right)^{15.6}=0.00162$$

$\pi_Q=20$, $\pi_E=2$. $\pi_C=1$.
$\lambda_{P\text{-}(L)}==0.00162.2.20.1=0.0648$

Table 3. Adjustment factors (MIL HDBK-217F standard)

	π_Q	π_E	π_S	π_T	π_A	π_C	π_{CV}	λ_b
SiC Transistor FET	5.5	6	/	2	10		/	0.012
Schottky Diode	5.5	6	0.07	3	10	1	/	0.003
Si Transistor	5.5	6	/	2	10		/	0.012
Diode	5.5	6	0.07	3	10	1	/	0.069
Inductor	20	2	/	/		1	/	0.00162
Capacitor	10	4	/	/	/	/	0.56	0.043

Table 4. Reliability of the boost converter based on different semiconductors generation

	Element	Failure rate λpart [failure/10^{-6}]	System failure rate
Boost converter based on SiC device	Inductor	0.0648	$\lambda_{system}=\sum_{i=1}^{N}\lambda_{i-part}$ 6.243 failure/10-6h
	Capacitor	0.9781	
	SiC Mosfet	5.191	
	Schottky diode	0.01005	
Boost converter based on Si device	Iductor	0.0648	$\lambda_{system}=\sum_{i=1}^{N}\lambda_{i-part}$ 8.335 failure/10-6h
	Capacitor	0.9781	
	Si Mosfet	6.962	
	Diode	0.3310	

Table.4 provides insights into the expected reliability results, rigorously evaluated in accordance with the well-established Military Handbook (MIL-HDBK-217) standard. This analysis takes various factors into ac-

count, including thermal stress, voltage fluctuations, and current variations, delivering valuable insights into the Mean Time Between Failures (MTBF) of the converter. The forecasted reliability once again confirms the resilience of the SiC-based system in enduring demanding conditions, assuring steadfast performance across the converter's entire operational spectrum.

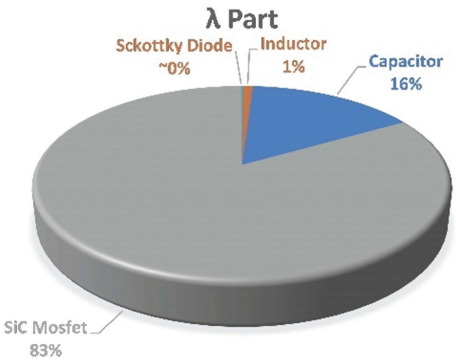


Fig. 3. λ part of boost converter based on SiC device

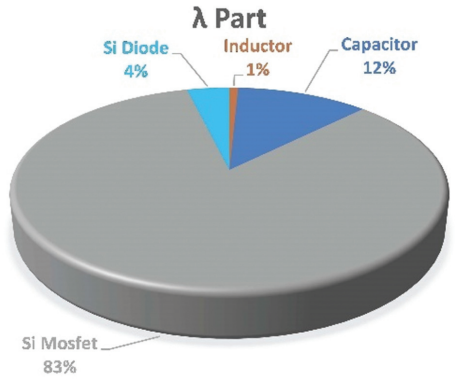


Fig. 4. λ part of boost converter based on Si device

Figs. 3 and 4 present the λ part of boost converter based on SiC and Si devices, respectively. The results show that more than 83% of boost converter failures are caused by semiconductors. This justifies the importance of improving the reliability of these devices.

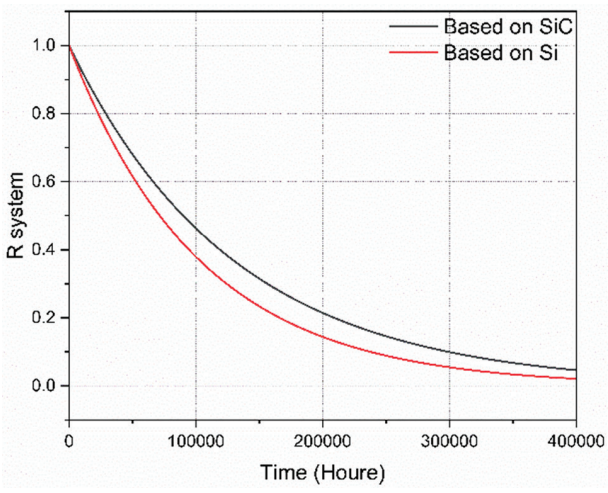


Fig. 5. Comparative Reliability Analysis - Si vs SiC

The comparison of reliability predictions between SiC-based and Si-based boost converters, as shown in Fig. 5, tells a compelling story of transformation. In this narrative, the SiC-based system emerges as a symbol of robustness and trustworthiness. Its superior thermal stability, coupled with the intrinsic durability of SiC devices, positions it for a prolonged operational life, surpassing its Si-based counterpart in terms of sustained reliability.

In this intersection of data, our analysis resonates with the potential of SiC-based boost converters to redefine the reliability landscape. They not only exhibit qualities such as higher efficiency, superior voltage and current handling but also emphasize their capacity for enduring and dependable performance. This reliability prediction further bolsters our confidence in the SiC-based boost converter's ability to gracefully navigate the challenges posed by real-world applications, seamlessly aligning with the exacting demands of contemporary power electronics systems.

4. CONCLUSION

In conclusion, our study delved into the intricate interplay between Silicon Carbide (SiC) and traditional Silicon (Si) in the realm of boost converters' reliability, guided by the MIL-HDBK-217 standard. The primary focus was on assessing the impact of SiC devices, with their promising material properties, on the failure rates of these essential electronic components.

The results of our analysis revealed a substantial decrease in the failure rate of boost converters when utilizing SiC devices, dropping from one failures on 13.7 year to one failures 18.3 year. This significant reduction underscores the heightened reliability achieved with SiC technology, aligning with the expectations outlined in the MIL-HDBK-217 standard. The standard's comprehensive guidelines facilitated a meticulous evaluation, ensuring the robustness and consistency of our reliability assessment.

Our findings not only corroborate the efficacy of SiC in enhancing reliability but also provide valuable insights for practitioners and researchers in the field of electronic engineering. The MIL-HDBK-217 standard played a pivotal role in guiding our study, offering a reliable framework for comparison and assessment.

In the broader context of boost converter technology, our results suggest that Silicon Carbide emerges as a material that holds great promise in elevating the reliability standards of electronic systems. The visual representation of our results, presented graphically, will further enhance the accessibility and clarity of our findings.

As we move forward, this study contributes to the ongoing discourse on material selection in electronic design, providing a foundation for future advancements in boost converter technology. The promising outcomes with SiC materials underscore their potential role as key

enablers for achieving robust and dependable electronic systems in the pursuit of technological progress.

This study can be improved using others recent and updated benchmarks, as well as it can be extended by including other methods of reliability evaluation.

5. REFERENCES

[1] P. I. Prodanov, D. D. Dankov, N. D. Madzharov, "Research of reliability of power thyristors using methods MIL-HDBK-217F and FIDES", Proceedings of the 21st International Symposium on Electrical Apparatus & Technologies, Bourgas, Bulgaria, 3-6 June 2020.

[2] M. Mojibi, M. Radmehr, "Reliability evaluation of buck converter based on thermal analysis", Journal of Microelectronics, Electronic Components and Materials, Vol. 48, No. 4, 2018, pp. 217-227.

[3] B. Abdi, A. H. Ranjbar, G. B. Gharehpetian, J. Milimonfared, "Reliability considerations for parallel performance of semiconductor switches in high-power switching power supplies", IEEE Transactions on Industrial Electronics, Vol. 56, No. 6, 2009, pp. 2133-2139.

[4] F. Khelifi, B. Nadji, Y. Chelali, "Study of the Reliability of Static Converter for Photovoltaic Application", Energy Procedia, Vol. 74, 2015, pp. 564-574.

[5] S. Ghosh, V. K. Yadav, V. Mukherjee, "A Novel Hot Spot Mitigation Circuit for Improved Reliability of PV Module", IEEE Transactions on Device and Materials Reliability, Vol. 20, No. 1, 2020, pp. 191-198.

[6] J. O. Gonzalez, R. Wu, S. Jahdi, O. Alatise, "Performance and Reliability Review of 650 v and 900 v Silicon and SiC Devices: MOSFETs, Cascode JFETs and IGBTs", IEEE Transactions on Industrial Electronics, Vol. 67, No. 9, 2020, pp. 7375-7385.

[7] E. Bouchetob, B. Nadji, "Using ANN based MPPT controller to increase PV central performance", Proceedings of the 2nd International Conference on Advanced Electrical Engineering, Constantine, Algeria, 29-31 October 2022.

[8] A. M. Noman, K. E. Addoweesh, H. M. Mashaly, S. Energy, "A Fuzzy Logic Control Method for MPPT of PV Systems", Proceedings of the 38th Annual Conference on IEEE Industrial Electronics Society, Montreal, QC, Canada, 25-28 October 2012, pp. 874-880.

[9] "SiC Power Devices and Modules", https://fscdn.rohm.com/en/products/databook/datasheet/discrete/sic/mosfet/sch2080ke-e.pdf. (accessed: 2023)

[10] "Advanced Technical Data HiPerFETTM Power MOSFETs N-Channel Enhancement Mode Avalanche Rated", https://datasheetspdf.com/pdf-file/697470/IXYSCorporation/IXFN32N120/1. (accessed: 2023)

[11] Advanced Power Technology, "APT60D120B DIODE Datasheet", https://datasheetspdf.com/pdf-file/379649/AdvancedPowerTechnology/APT60D120B/1. (accessed: 2023)

[12] "C4D20120D Wolfspeed Schottky Diodes & Rectifiers SIC SCHOTTKY DIODE 1200V, 2x10A", https://assets.wolfspeed.com/uploads/2020/12/C4D20120D.pdf. (accessed: 2023)

[13] I. Villanueva, N. Vázquez, J. Vaquero, C. Hernández, H. López-Tapia, R. Osorio-Sánchez, "Photovoltaic Inverter Reliability Study through SiC Switches Redundant Structures", Technologies, Vol. 11, No. 2, 2023, p. 59.

[14] "MIL-HDBK-217F (NOTICE 2), MILITARY HANDBOOK: RELIABILITY PREDICTION OF ELECTRONIC EQUIPMENT (28 FEB 1995) A", https://www.quanterion.com/wp-content/uploads/2014/09/MIL-HDBK-217F.pdf (accessed: 2023)

Design of Regenerative Braking System and Energy Storage with Supercapacitors as Energy Buffers

Original Scientific Paper

Siluvai M. Michael
Sri Sivasubramaniya Nadar College of Engineering
Chemical Sciences Research Centre, Department of Chemistry.
Tamil Nadu 603110, India
michaelm@ssn.edu.in

Bokani Mtengi
Botswana International University of Science and Technology
Department of Electrical, Computer and Telecommunications Engineering,
Plot 10071 Boseja, Palapye, Botswana
mtengib@biuts.ac.bw

S.R.S Prabaharan
Inventus Battery Energy Technologies Pvt Ltd
Tamil Nadu 600090, India
srsprabaharan1611@gmail.com

Adamu Murtala Zungeru
Botswana International University of Science and Technology
Department of Electrical, Computer and Telecommunications Engineering,
Plot 10071 Boseja, Palapye, Botswana
zungerum@biust.ac.bw

James Garba Ambafi
Federal University of Technology
Department of Electrical and Electronics Engineering,
P. M.B. 65, 920101, Minna, Niger, Nigeria
ambafi@futminna.edu.ng

Abstract *– Vehicles are part of urban area transport and are subjected to variable loads as they traverse the city with varying slopes and stop-and-go traffic. Electric Vehicles (EVs) can be a good option because of their high efficiency under stop-and-go conditions and ability to gain energy from braking. However, limited battery energy makes EVs less efficient and degrades their lifetime. In contrast to a Li-Ion battery, supercapacitors work well under high power charge and discharge cycles. However, their high cost and low energy density prevent them from being viable replacements for batteries. Due to the slow charging and discharging process of batteries, they have a low power density, but a high energy density compared to the supercapacitor. In this paper, we discussed our system design consisting of both a battery and a supercapacitor. The main aim is to design and develop a scheduling algorithm to optimize energy flow between the battery, supercapacitor, and motor. We further described an analogue-based control methodology and algorithm for the supercapacitor, augmented battery-powered motoring process. This is in addition to a charge controller designed to optimize the supercapacitor bank's current-based charge-discharge profile. The system design and tests are developed on PSPICE and a hardware platform.*

Keywords: *Energy storage, Supercapacitors, Energy buffers, Regenerative braking systems, Electric vehicles, scheduling algorithm, DC-DC converter*

Received: September 17, 2023; Received in revised form: February 14, 2024; Accepted: February 15, 2024

1. INTRODUCTION

An increase in the percentage of the emissions of hazardous greenhouse gases and other pollutants comes from high road traffic transportation systems. Carbon emissions could be reduced using electric vehicles. However, there have been some limitations over EVs, like energy backup, which depends on battery power. Power management systems involving supercapacitors have been of great interest in the power engineering arena, and various control approaches and topologies have been proposed for numerous supercapacitor-augmented applications [1-15]. With the advancement of Energy Storage Systems (ESSs), several proposals have emerged for managing and designing such systems. Some work have detailed the comparison of the ESSs, which can be found in Ref. [16]. Recently, the concept of Hybrid ESSs has emerged. This concept combines two or more ESSs to benefit from their complementary characteristics [17-

19]. Sang Young Park et al. [20] proposed an electronic-based power interface for supercapacitor-based hybrid energy storage systems and battery management for electric vehicles. The effectiveness of the supercapacitor-based HESSs has been demonstrated in a large number of research works, showing improved performance in terms of acceleration of the EV [17], energy efficiency during regenerative braking [9], driving range of the EV [2,6], and the lifetime of battery [21-24]. In addition, other noticeable research is outlined in [25-28], which has made many contributions to this field.

To this end, we developed a new control strategy for a supercapacitor-battery vehicle. It is anticipated that the hybrid power source with a combination of a supercapacitor bank and Lithium-Ion battery will improve the range and efficiency of the vehicle by charging the battery through the supercapacitor. The energy generated from regenerative braking and the advantageous features of supercapacitors are defined as essential metrics for propulsion.

Supercapacitors have a very high specific power (W/kg). Its augmentation into battery-powered motoring systems has gained popularity recently, particularly in the electric vehicle (EV) and hybrid vehicle (HEV) industries. The contribution of supercapacitor modules as a power buffer during motor transients reduces the stress induced onto the supply batteries, as observed during motor acceleration. Similarly, for regenerative motor braking instances, injection of high power into the battery cells in a non-buffered manner would indubitably induce an electrochemical strain onto the battery chemistry, reducing its lifespan.

In recent years, numerous power-flow management and control strategies have been proposed for HEV-based applications [29]. Zhang et al. [30, 31] researched active hybrid system control strategies that propose charge-sharing mechanisms in storage elements involving supercapacitors. At the same time, Musat et al. [32] presented dynamic adapting topologies for hybrid storage systems. Though the numerous approaches and topologies are investigated, they are generally process-based strategies that enable the supercapacitor to cater to the application-specific designs. Hence, there is no substantial emphasis on the supercapacitor charge controller design and optimization for current fast-charge profiles. This article provided a detailed approach to designing a regenerative braking system and energy storage with supercapacitors as energy buffers through modelling and simulation in PSpice. We also undertake a laboratory prototype as shown in Figs. 5, 13 and 14. The major contributions of this paper are:

- A simplified Regenerative Braking model is designed with Supercapacitors as an electrical energy storage buffer.
- A new scheduling algorithm was designed to optimize energy flow between the battery, supercapacitor and motor.
- Designed a battery charging system based on supercapacitors expending energy onto Li-Ion batteries.

Hence, this paper uses an analogue-based charge controller to optimize the supercapacitor's charging and discharge profile. The charge controller is designed to be used in an experimental platform for an electric vehicle powertrain. With respect to the application, a control algorithm is proposed to manipulate the storage elements in the system effectively for effective power management. Our energy-efficient prototype system consists of batteries and supercapacitors connected through a DC/DC converter to achieve the expected performance, as shown in Fig. 1.

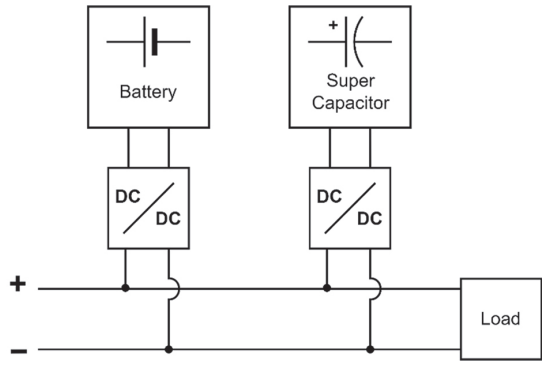


Fig. 1. A simplified view of architecture for a battery and supercapacitor energy source

Design verification is performed via software simulations in the PSPICE environment, and hardware realization using various components is done successfully. This system can be helpful for the further development of Electrical Vehicles.

2. SUPERCAPACITOR-AUGMENTED REGENERATIVE BRAKING SYSTEM

The proposed complete system is used to demonstrate the regenerative braking mechanism shown in Fig. 2. The kinetic energy from braking is stored in the supercapacitor, which charges the Lithium-Ion battery.

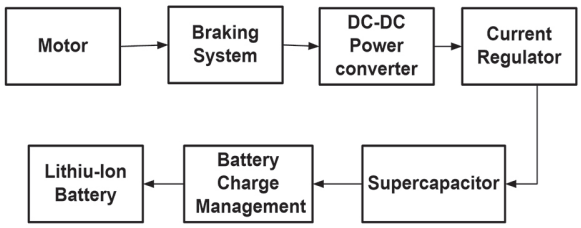


Fig. 2. Block diagram of the overall structure of the system

Regenerative braking technology captures energy from moving vehicles as they stop when braking is applied [33]. This energy is stored in the supercapacitor to charge the battery. This novel technology makes EV and HEV vehicles more advantageous in the automo-

bile sector. Supercapacitors are attractive due to their fast charging-discharging profile and long life use. Although the supercapacitor works the same way as the conventional capacitor, it can store much more power than the conventional capacitor in the quickest possible time due to its double-layer architecture in a small space and its ESR [34-36].

The regenerative braking system is augmented with a supercapacitor and a corresponding SCC through Cadence OrCAD PSPICE. The SCC is designed based on a controlled DC-DC converter [37] and exhibits mirrored topology for non-side-specific buck-boost operations and bi-directional power flow.

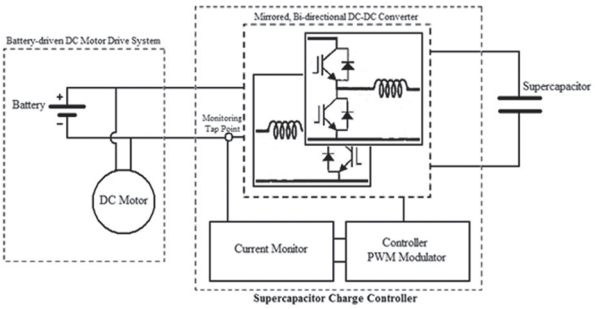


Fig. 3. Implementation and components of the SCC

Fig. 3 demonstrates integrating the proposed SCC and supercapacitor module into a battery-powered DC motor drive circuit devised for regenerative motor braking. The proposed SCC design incorporates a current monitor block and a subsequent controller that produces a corresponding PWM signal stream based on the current tracker. Therefore, the concurrent monitor and control cycles will produce an optimal current charge-discharge profile for the supercapacitor, which helps in charging the battery.

3. PRIMARY SYSTEM MODEL

The primary system model incorporated with DC-DC converter and motor supercapacitor module with a protection circuit. CAP-XX Ltd. only provided the SPICE model library for the supercapacitor, and the low-ESR GS206 supercapacitor is used in the proposed design for simulation. The specifications of the GS206 supercapacitor are displayed in Table 1. This supercapacitor is only used for simulation in PSPICE.

Table 1. CAP-XX GS206 supercapacitor model specifications

Nominal capacitance	600 mF
Rated Terminal Voltage	4.5 V
Nominal ESR	40 mΩ
Test temperature	20 °C

The designed primary system is defined herein as the circuit structure coupled directly to the high-power bus, which includes the battery for the supply to the motor, the supercapacitor buffer to achieve the desired prototype, the DC motor, and the DC-DC power converter to maintain the range of the voltage, as shown in Fig. 4. In order to use a supercapacitor as a peak power supply for the system till the end of the life span of the supercapacitor in a real-time application, there is a need to check the electric characteristics and energy capacity so the circuit protection unit is designed to avoid any overvoltage condition [37, 38].

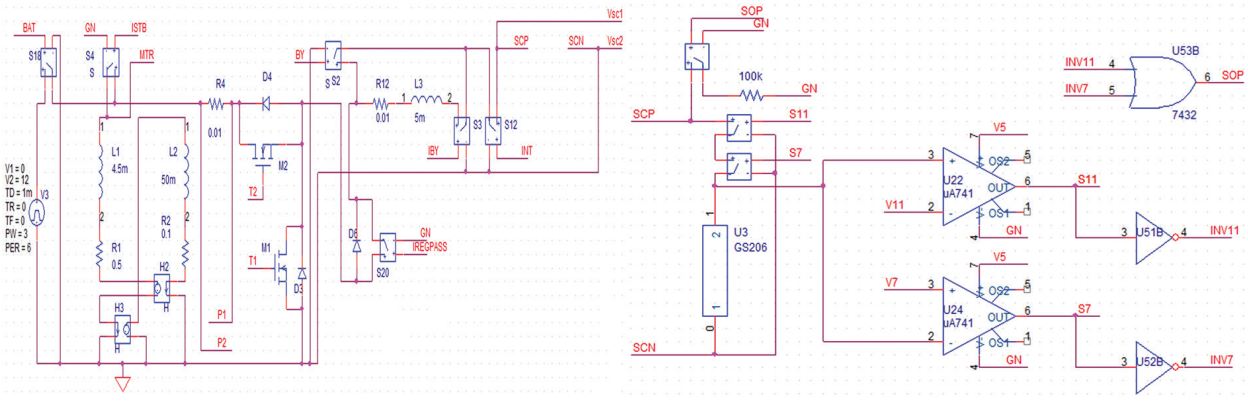


Fig. 4. SPICE model of the primary system

The DC motor is modelled based on its electrical and mechanical equivalents.

$$V_{in} = I_a R_a + L_a \frac{dI_a}{dt} + V_{emf} \qquad (1)$$
$$T_{dev} = K_t I_a$$

$$T_{dev} = J \frac{d\omega}{dt} + B\omega \qquad (2)$$

where V_{in} is the DC motor terminal input voltage, Ia is the motor armature current, R_a is the motor armature resistance, L_a is the motor armature inductance, V_{emf} is the motor back-emf, T_{dev} is the shaft developed torque, K_t is the current-torque relational coefficient, J is the induced shaft inertia, B is the shaft frictional coefficient, and ω is the shaft angular velocity.

The motor build specifications are shown in Table 2.

Table 2. DC motor experimental specifications

Armature inductance	4.5 mH
Armature resistance	0.5 Ω
Torque coefficient	1.21 Nm A-1
Load Inertia	0.05 Nm rad-1 s2
Friction	0.1 Nm rad-1 s
Speed coefficient	0.7 V rad-1 s

4. MOTORING AND BRAKING MECHANISM

A flywheel is used to retrieve the energy from the motor. The flywheel stores the mechanical energy from the rotation of the motor. The motor is considered the first stage of the model, and braking is considered a consecutive part of the model after the motor. The motor is running with 12 Volts power supply, and as soon as the brake is applied, the energy conversion is done by disconnecting the power supply to the motor. The momentum in the flywheel causes it to continue to rotate at a certain reducing velocity in free rotation. The circuit is now connected to the supercapacitors while isolated from the supply voltage of 12 Volts that runs the motor. When the brake is released, it reconnects the flywheel to the motor. This isolation is provided by the connection of the relay and is explained in other sections elsewhere in this article. The regenerative braking system shown in Fig. 5 is designed to incorporate a motor connected to a flywheel and a disc brake by a shaft running across.

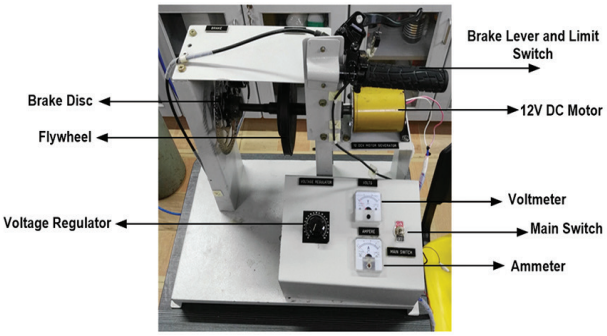


Fig. 5. Regenerative braking motor and braking system

The circuit diagram of the electric model of regenerative braking is shown in Fig. 6(a). The circuit limit switch is originally in a normally closed condition, which means if the main switch is closed, the current will flow through the relay coil, which is magnetized. An energized relay coil will make the normally open switch connection a closed circuit, and the normally closed circuit becomes an open circuit.

The motor's motoring cycle will rotate the flywheel by transferring mechanical energy through the shaft connected to it. The relay coil gets energized, bypassing current, so all normally closed switches become normally opened and vice-versa. The current flow is shown in Figs. 6(b) and 6(c) using red color in the motoring and braking cycle, respectively. In a braking situation, a limit switch is opened, which will cause the current to flow into the supercapacitor bank since the motor is in the regenerative cycle.

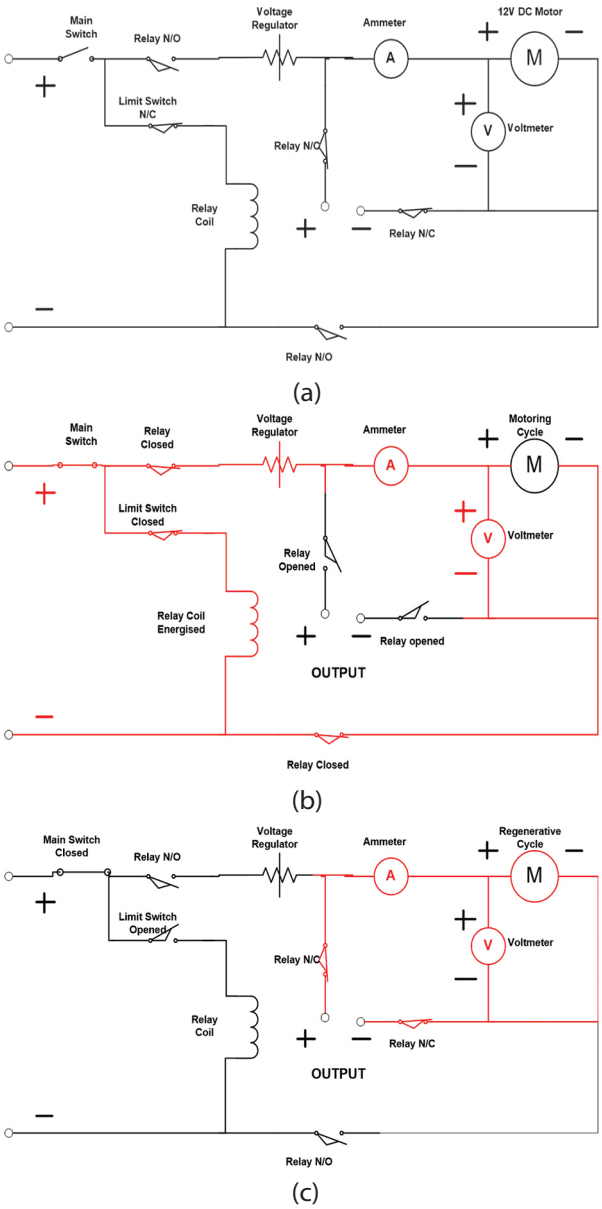


Fig. 6. (a) Electrical circuit schematic of Regenerative braking system, (b) Motoring cycle, (c) Regenerative cycle

5. SCC MONITOR AND CONTROL CIRCUIT DESIGN

The SCC monitor and control (MnC) circuit is an integrated part of the SCC design, developed to feed precise gating signals to the buck-boost MOSFETs in the primary system SCC. The supercapacitor's state of charge (SOC) is essential for the smooth running of the system. The SCC MnC is developed to decode and modulate based entirely on analogue schemes; hence, it is an extremely sensitive controller compared to DSP-based encoders and modulation techniques exhibiting finite sampling resolution.

5.1. SCC MNC ALGORITHM

The SCC MnC control sequencing and algorithm are shown in Fig. 7. During system startup, the EDLC bank state of charge (SOC) will be gauged, battery-charged, and maintained at a specified nominal SOC. The main purpose of the SOC nominalization is to prepare the EDLC for motor transients, where high power is drawn during acceleration and stored during regeneration. In a scaled-down system of 12 V, the nominal voltage is approximately 50 % of the effective EDLC energy spectrum, which is 9 V nominal. Before motor acceleration, SOC will be below nominal due to power buffering, and similarly, SOC will be above nominal prior to regenerative braking. During constant velocity cruising, no transients are observed, and the SCC will re-level the EDLC SOC to its nominal state.

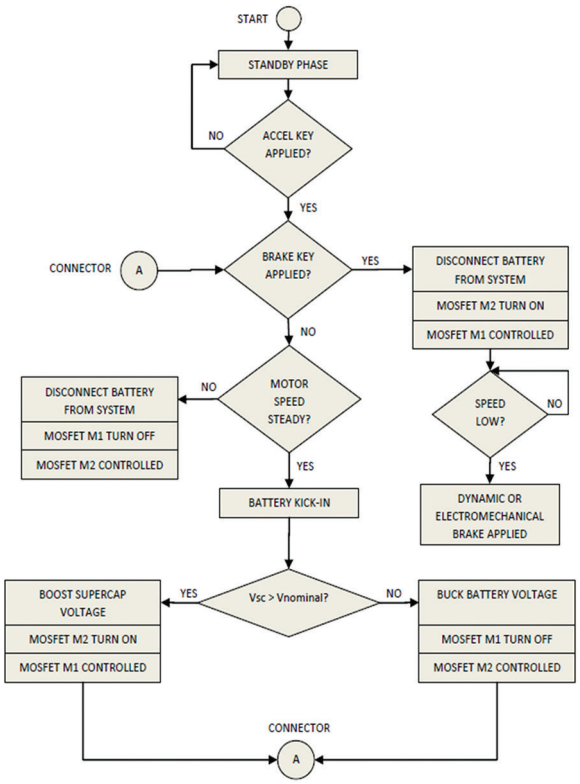


Fig. 7. SCC MnC flowchart

The SCC will disconnect the system battery during motor acceleration from zero or low speeds and reconfigure the DC converter for forwarding power transfer in boost configuration. The SCC MnC package will perform concurrent gating control onto the boost converter based on the current tracker, allowing optimized acceleration. The MnC will reconnect the battery into the system at near steady speeds while reconfiguring the SCC to restore the supercapacitor nominal SOC. Similarly, when regenerative braking is performed, the SCC is reconFigd to allow backward power flow in boost configuration and is gated appropriately through the MnC scheme. As regenerative power is insignificant at low speeds, shaft speed is not effectively reduced, and hence electromechanical braking is invoked as an effective, low-speed braking mechanism through the MnC.

5.2. CHARGE CONTROLLER MNC CIRCUIT DESIGN

The power flow in the primary system has a rather deterministic implication on the states of the primary system blocks. Therefore in a simple motor driver design, the supercapacitor charge controller MnC is designed based purely on system power trackers, such that it suffices the switching between modes of operation.

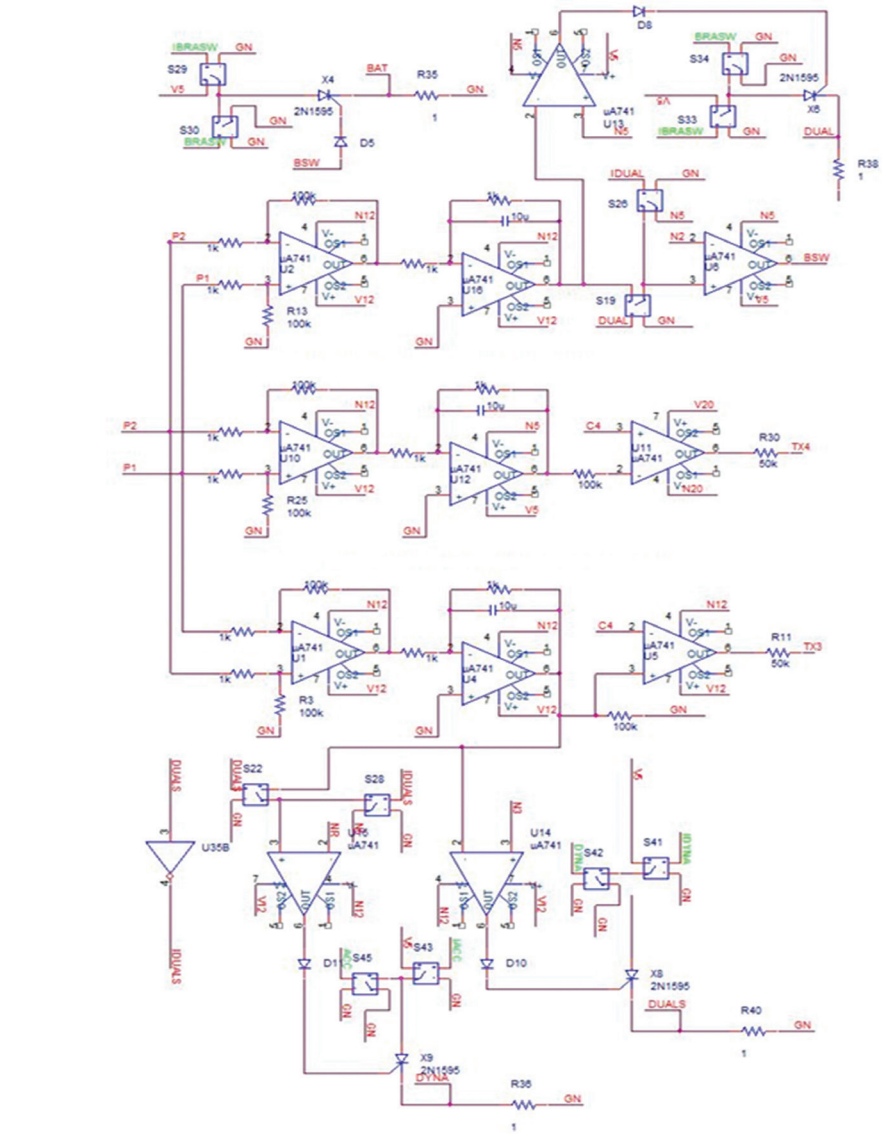
The charge controller MnC circuit was designed in SPICE and then implemented on hardware. A total of 2 trackers are implemented into the MnC design, which monitors the primary system line power and the supercapacitor SOC respectively. The first tracker is a 10 mΩ sense resistor placed on the power line between the DC motor and the SCC, as shown in Fig. 4, which provides concurrent monitoring of the system power flow. A subsequent tracker monitors the EDLC's SOC.

Dedicated PWM modulators and controllers are designed for each MOSFET transistor at the primary SCC DC-DC converts based on uA741 op-amp topological configurations. Power information is tracked at the sense resistor, which is performed with differentiation and integration-based analogue decodes. Tracked potential at the sense resistor is differentiated to obtain a magnitudal representation of the primary system line current, which is later fed into a low RC op-amp integrator for a reference signal responsive to current variations.

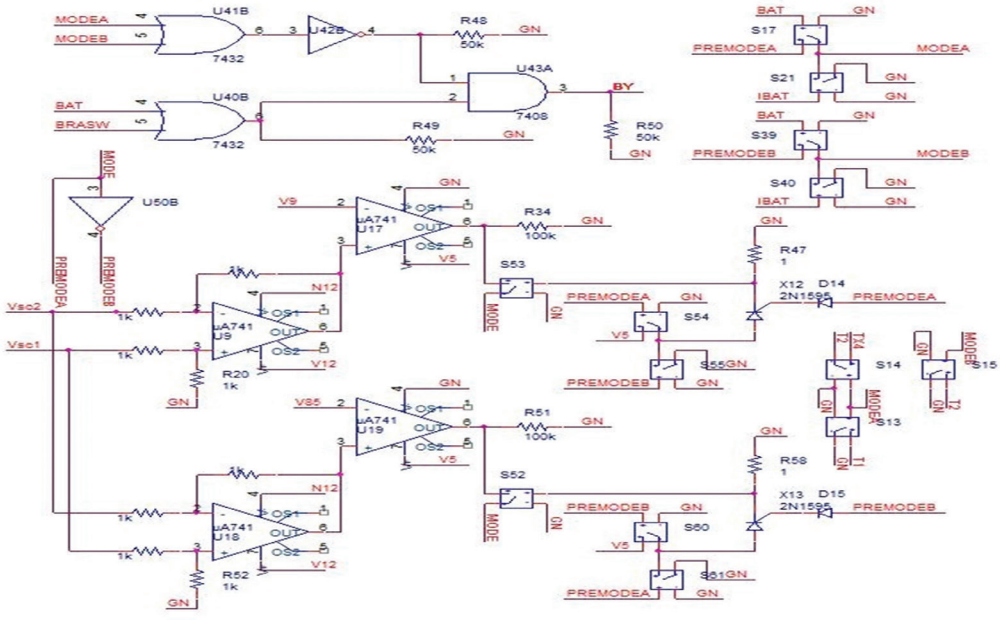
The processed output is referenced to a pulse train via natural sampling to produce a pulse width modulated (PWM) MOSFET gating signal fed into the primary SCC for an optimized supercapacitor charge-discharge profile. Fig. 8 shows the MnC design.

In a solely current tracked MnC scheme, accelerative and regenerative operations require mode switches that trigger based on an impending steady-state at motor transients, as seen in steady-speed motoring Dual-point (DP) controllers are developed as an event-triggered controller that responds to 2 sequenced current reference levels to perform an operational mode switch. Fig. 9 demonstrates the DP triggering at an accelerative transient. In order to switch modes appropriately, the flag point must be triggered prior to the real reference point to bypass the false point. Fig. 8(a) shows the DP triggers and control circuits.

At all non-transient states, the SCC nominal SOC controller must initiate to level the EDLC SOC at nominal SOC before an impending transient state. Fig 8(b) shows the nominal SOC controller circuit design. After a successful acceleration buffer, the depleted EDLC SOC is replenished at an optimized charge-rate preset based on an acknowledged battery discharge strain and system size. Similarly, during regeneration, the excess charge stored in the EDLC is fed back into the battery cell.



(a)



(b)

Fig. 8. SCC MnC SPICE design for (a) PWM modulators with DP triggers and (b) Nominal voltage controller and operation mode-switching logic

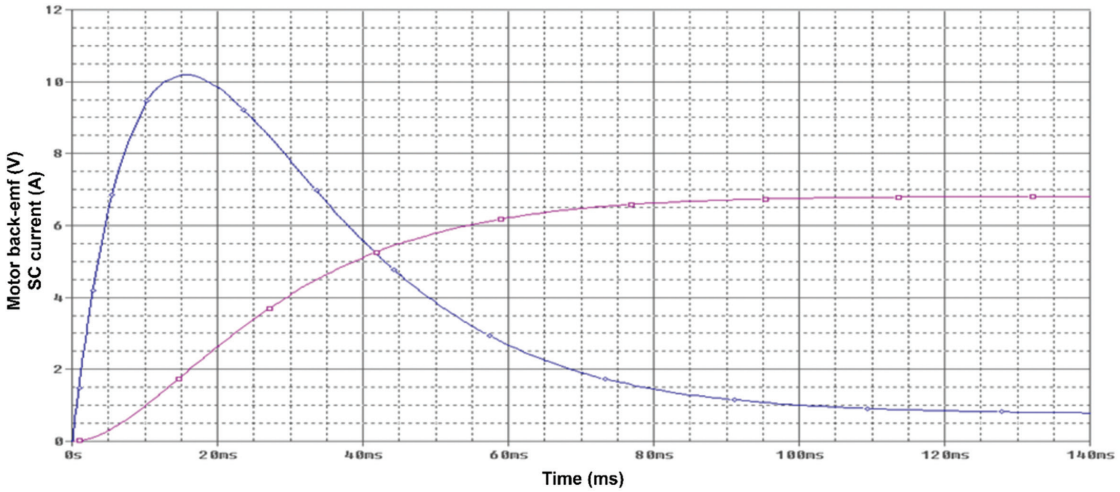
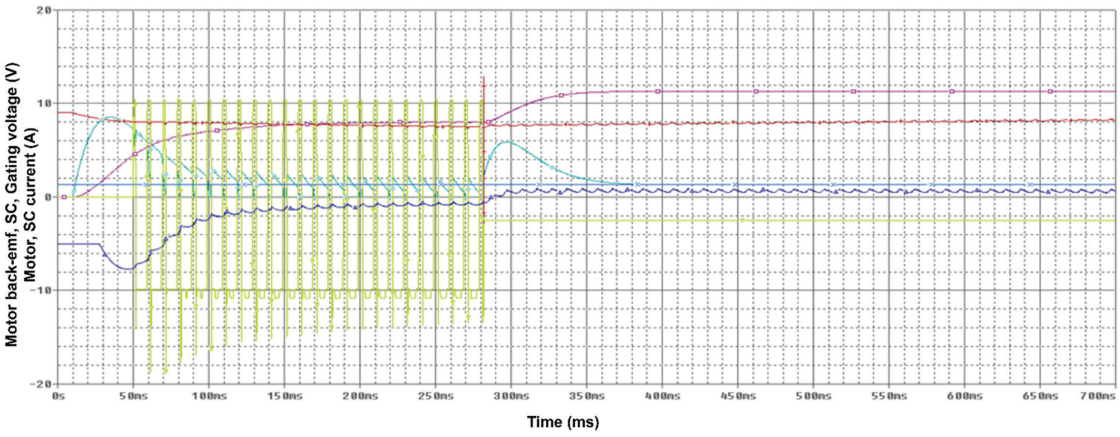


Fig 9. DP control approach at transients

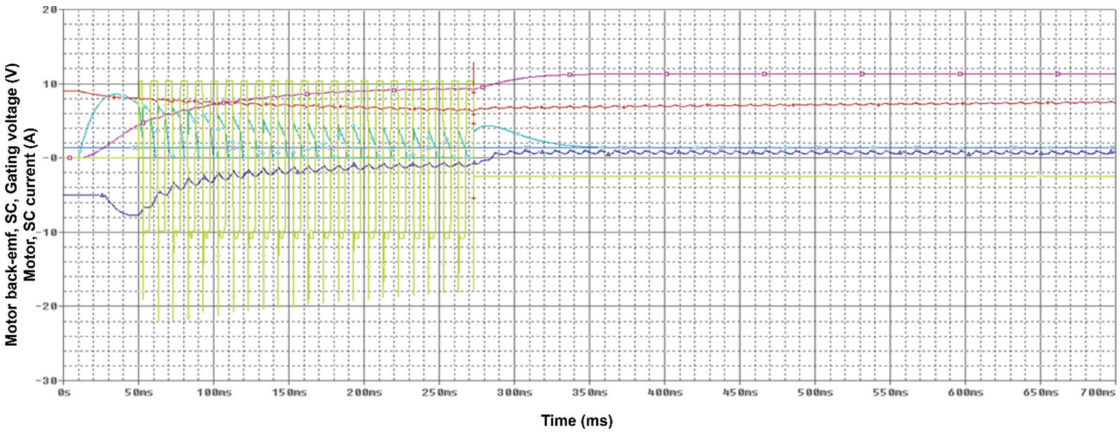
6. EXPERIMENTAL RESULTS USING SIMULATION

6.1. A MOTOR ACCELERATION POWER BUFFER

The supercapacitor buffers the high power demand of the DC motor during accelerative transients at experimental values of the duty cycle. Fig. 9 shows a test result for the accelerative action. In Fig. 10(a), a 600 mF EDLC pre-charged at a nominal 9 V SOC can pull the motor back-emf up by 8 V from a motor standstill. A similar EDLC module with equal pre-charge can pull the motor back-emf up by 9.2 V when fed with a higher valued fixed duty cycle, at the expense of substantially higher EDLC SOC depletion, as simulated in Fig. 10(b). At a predetermined supercapacitor current cutoff value, which is an optimistic determinant of a near steady-state operation, the MnC will have the battery kick-in into the primary system with low current operation thereon.



(a) $D = 0.25$, fconverter = 0.1 kHz



(b) $D = 0.50$, fconverter = 0.1 kHz

Fig. 10. Acceleration mode simulation result for (a) duty cycle, $D = 0.25$ and (b) $D = 0.5$

6.2 SUPERCAPACITOR SOC BALANCING

Supercapacitor SOC is balanced at nominal value for all system non-transient states. Fig. 11 demonstrates a test of the SCC nominal SOC controller. In this case, the system is worked into the steady-speed motoring region from a standstill, in which the aforementioned depleted EDLC SOC during the accelerative power buffer is replenished to a nominal state. The supercapacitor rate-of-charge in this mode is independent of the supercapacitor size but is highly dependent on the battery size and mean system steady-speed power. Considerations of the above accomplish optimization of the charging-discharging profile in this mode.

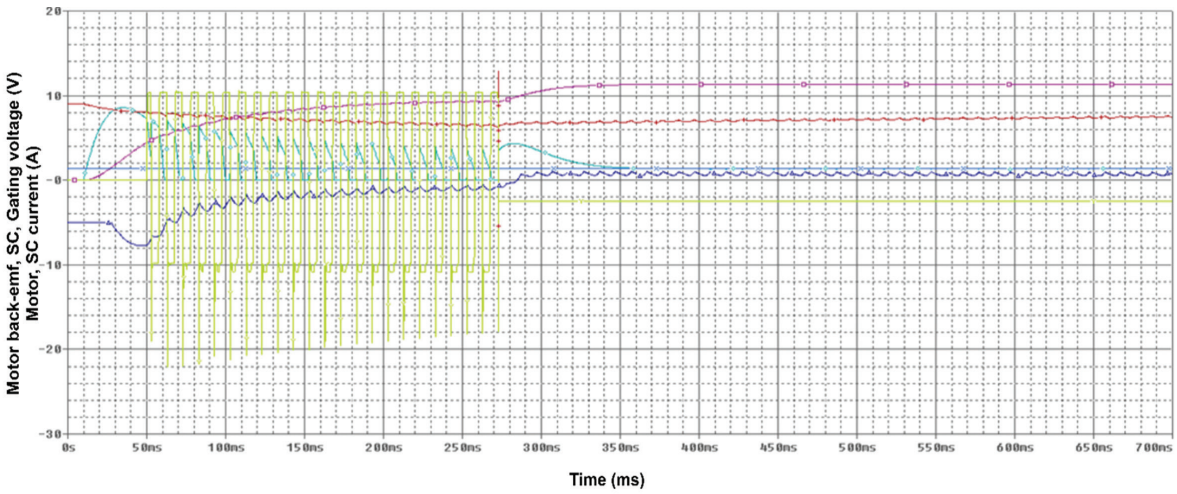


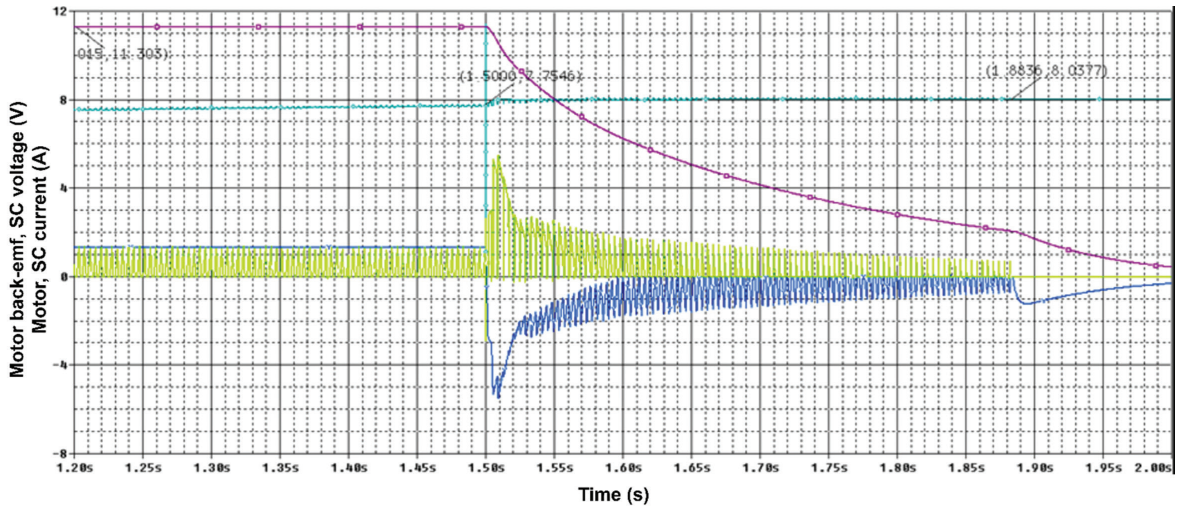
Fig. 11. Nominal SOC forced test

6.3. REGENERATIVE BRAKING

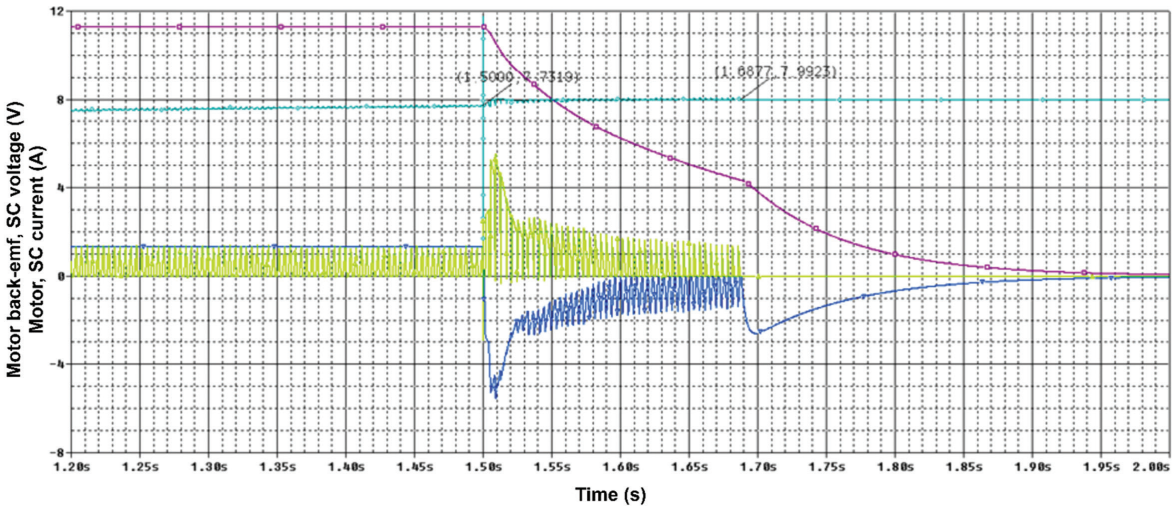
Regenerative braking is executed via a forced power-flow reversal in the primary system through the SCC controls. Fig. 12 shows the regenerative braking profile of the primary system. Control of the reference current has projected a substantial change in the regenerative cutoff, where the reference current sets a minimum quantity at which the tracked system current is required to reach before electromechanical brakes are effectively applied.

The adequate braking time, defined herein as the average time required to bring the motor to a standstill from steady-speed cruising, is an essential design factor in most motor systems since some degree of control over physical braking must be enabled.

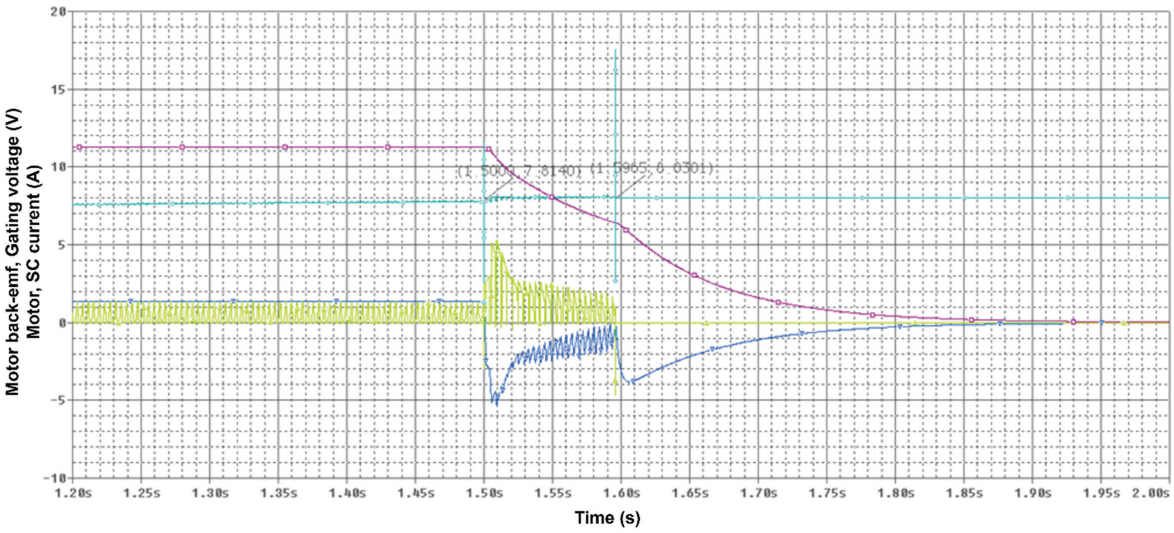
Assuming a boundary condition where the minimum applicable reference current is set as Fig. 12(a), it has been found that the regenerative efficiency will drop by an approximated 9 per cent when the braking action is performed with a significantly higher reference, as such in Fig. 12(b). However, in this situation, the effective braking time has been reduced by 50 percent, denoting a non-linear proportionality between regenerative efficiency and the effective braking time in the low energy spectrum. Therefore, the system design must consider a level of compromise between the parameters.



(a) *Iref* = 200 mA, *fconverter* = 0.25 kHz

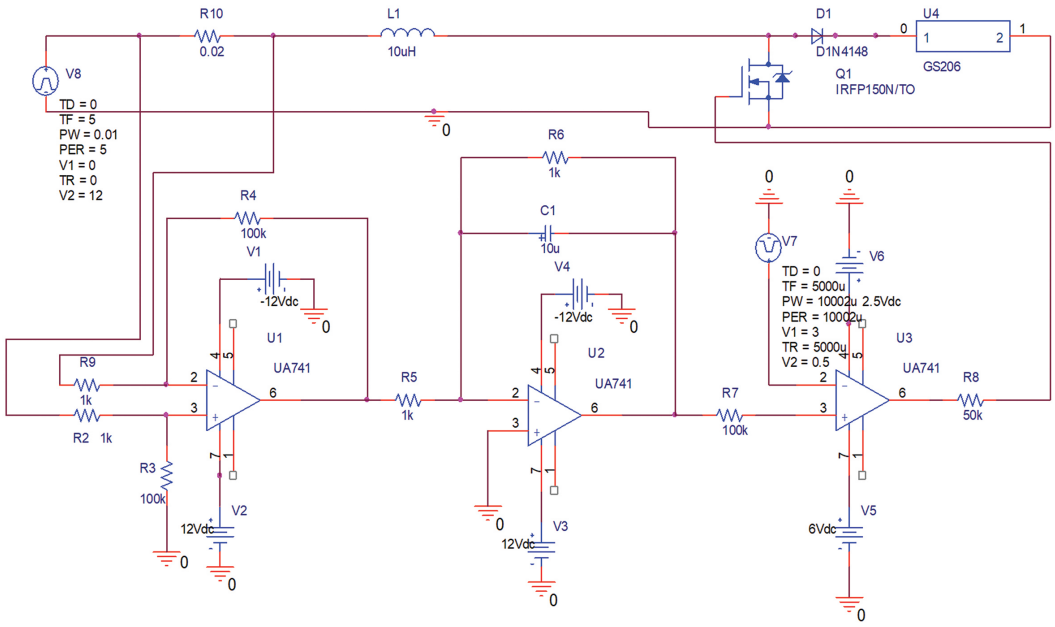


(b) *Iref* = 500 mA, *fconverter* = 0.25 kHz

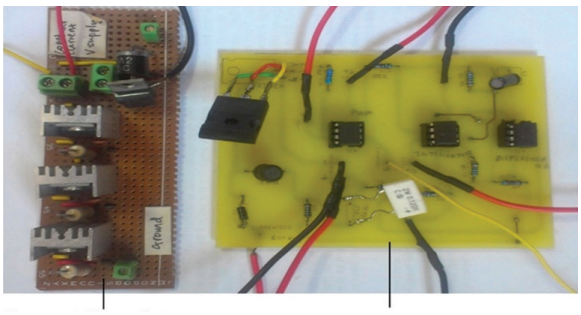


(c) *Iref* = 1000 mA, *fconverter* = 0.25 kHz

Fig 12. Regenerative mode simulation result for (a) reference current of 0.2 A, (b) 0.5 A and (c) 1 A



(a)



(b)

Fig. 13. (a) PSPICE Schematic of DC-DC converter and controller, (b) Hardware circuitry of the current regulator and DC-DC converter and controller.

DC-DC converter maintains a continuous flow of charge from the flywheel to the supercapacitor bank during regenerative braking. Fig. 13(a) shows the PSPICE model of the DC-DC converter and controller. The sensing resistor works as a reference for the controller, and the inductor serves as a tool to boost the output voltage at the OFF state of the converter. In this simulation, the constant current regulator is not shown, but the diode is supposed to be connected to the constant current regulator. The diode's purpose is to forbid the flow of charge into the supercapacitor back to the circuit. The DC-DC converter can produce a different level of boosting since the controller's output, which is a square wave in varying duty cycles due to the inputs from the sensing resistor, is taken as a duty cycle to operate the switch of the DC-DC converter. The experimental setup is shown in Fig. 13(b) in these three parallel configurations of the current regulator to successfully produce the desired 3 Amps of constant current to charge the bank.

When the supercapacitor is incorporated into the system, it will serve two important purposes. The first is to act as a buffer before the regenerative energy can be stored in the lithium-ion battery. The second role is an energy storage device that holds excess charge from the braking energy. The supercapacitor used in this has a capacitance of 10 Farad and a rated voltage of 2.5 Volts manufactured by Maxwell; specifications are shown in Table 3. Supercapacitors are connected in series to meet high voltage demands, so the capacitance is 2 Farad, and a 12.5 Volts rating voltage is shown in Fig. 14.

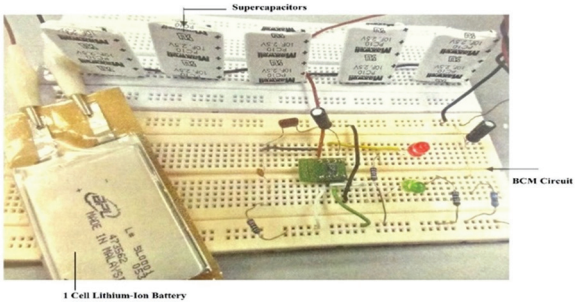
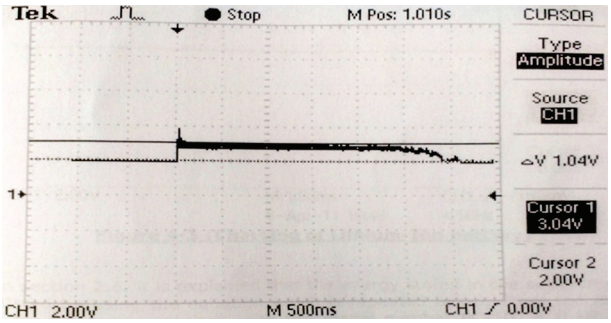


Fig 14. Combined circuitry of charging Lithium-Ion battery using supercapacitor banks incorporating BCM

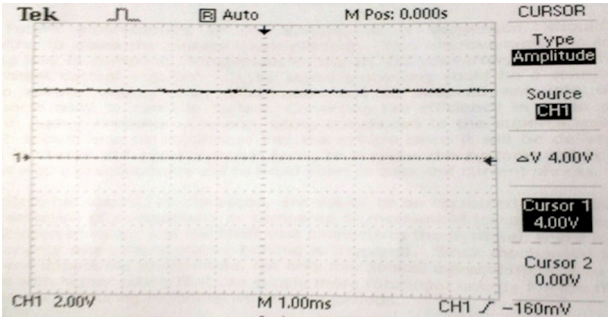
Table 3. Maxwell technologies pc-10 series model specifications

Nominal capacitance	10 F
Rated Terminal Voltage	2.5 V
Nominal ESR	0.18 Ω
Operating temperature Range	-40 °C to 70 °C

Hitherto, the supercapacitor charging process was going on now; the next part was to charge a battery. Before connecting a battery directly to the supercapacitor, a battery charge management circuit is required for a safety mechanism. Internal MOSFET of LTC4052 works to prevent the reverse battery current if the input voltage is shorted to the ground. 1- Cell lithium-ion battery is used. The supercapacitors are connected in series having 12.5 Volts, which creates a higher potential difference on the supercapacitor side to charge the battery up to the level of 4.2 Volts managed by the BCM system. The supercapacitor banks can be increased at any time for more energy, but this configuration is sufficient and successfully tested for testing and prototype.



(a)



(b)

Fig 15. Output voltage after (a) constant current regulator, (b) Charging of Lithium-Ion battery

DC voltage recorded by oscilloscope is shown in Fig. 15(a), which depicts that the voltage rises as soon as the regenerative braking scheme is activated and is constant until all energy from the flywheel reduces when it comes to halting voltage becomes zero. This energy is stored as a charge in a supercapacitor, proving that the system can use a supercapacitor as a buffer in regenerative braking. The system is then completed by connecting the supercapacitor to BCM to charge the battery. Fig. 15(b) shows the charging of a Lithium-Ion

battery. The voltage of the battery continues to rise to 4.2 Volts. This measurement indicates that the system is working as proposed.

7. CONCLUSIONS

We proposed and designed an analogue-based supercapacitor charge controller design and algorithm, incorporated and analyzed in a simple regenerative braking-based application. The EDLC module substantially contributed to the power buffering of a battery-powered motor system and has demonstrated a decent current charge-discharge profile for a fast charging scheme optimized for performance and endurance. In line with using an EDLC module and analogue SCC in the proposed system, a modest control algorithm has also been developed and verified on hardware support. Supercapacitors effectively reduce peak battery current, but the range is limited. The developed system mainly concentrates on proof of concept, which is proved by results, and therefore, aspects such as optimization and efficiency were not much considered a high priority. Although the designed system can be used as the foundation for future work, an increment in the system's functionality will add more value and quality to the system. To begin with, a program is needed to check any controller's non-functionality of the supercapacitor bank This can be managed using the timer function by providing a predefined time (say in ms) to check the supercapacitor's charging duration. With this, if any bank or supercapacitor does not meet the time requirement, which should have a fault in the bank. Work is in progress to add more design features by interfacing the current system status with other modules in the vehicle. Using processors to customize implementation can be possible. With the above additional features, the commercial prototype can be developed.

AUTHORS CONTRIBUTIONS

All co-authors of this manuscript have equally contributed to the content of this manuscript. S. R. S. Prabaharan – A Senior author that conceived the idea, designed and performed the simulation work, and wrote the manuscript. M. S. Michael– conceived the idea, designed and performed the simulation work, and co-wrote the manuscript. Adamu Murtala Zungeru – Co-Senior author who conceived the idea alongside S. Prabaharan, contributed to the design analysis and co-wrote the manuscript. J. G. Ambafi – provided critical feedback, helped shape the research analysis, and co-wrote the manuscript. B. Mtengi provided critical feedback, helped shape the research analysis, and co-wrote the manuscript.

COMPETING INTERESTS

The authors declare no competing interests.

DATA AVAILABILITY

There is no external data used for this research

ADDITIONAL INFORMATION

Correspondence and requests for materials should be addressed to Adamu Murtala Zungeru and S. R. S. Prabaharan.

ACKNOWLEDGEMENTS

Authors are indebted to Inventus Battery Energy Technologies Pvt Ltd, Tamil Nadu 600090, India, and Botswana International University of Science and Technology. The authors are thankful to the management of the organizations for their support.

8. REFERENCES

[1] F. K. Abo-Elyousr, R. M. Kamel, H. Abo-zeid, F. Abdelbar, “Characterization and Effectiveness of an Ultracapacitor Bank to Enhance a Battery Electric Vehicle Energy Storage System Dynamics”, Proceedings of the International Conference on Electrical Machines and Systems, Busan, Korea, 26-29 October 2013, pp. 235-241.

[2] X. Shen et al. "ConFig Methodology of Onboard Supercapacitor Array for Recycling Regenerative Braking Energy of URT Vehicles", IEEE Transactions on Industry Applications, Vol. 49, No. 4, 2013, pp. 1678-1685.

[3] Y. Xing, E. W. M. Ma, K. L. Tsui, M. Pecht, “Battery Management Systems in Electric and Hybrid Vehicles”, Energies Vol. 4, 2011, pp. 1840-1857.

[4] B. P. Divakar et al. "Battery management system and control strategy for hybrid and electric vehicle”, Proceedings of the 3rd International Conference on Power Electronics Systems and Applications, Hong Kong, China, 20-22 May 2009, pp. 1-6.

[5] P. Aishwarya, O. M. Hari, "A Review of Optimal Energy Management Strategies for Hybrid Electric Vehicle", International Journal of Vehicular Technology, Vol. 2014, 2014.

[6] Y. Zou, H. Shi-jie, L. Dong-Ge, G. Wei, X. Hu, "Optimal Energy Control Strategy Design for a Hybrid Electric Vehicle", Discrete Dynamics in Nature and Society, Vol. 2013, 2013.

[7] J. Wu, X. Wang, L. Li, C. Qin, Y. Du, Hierarchical control strategy with battery aging consideration for hybrid electric vehicle regenerative braking control, Energy, Vol. 145, 2018, pp. 301-312.

[8] A. Adib, R. Dhaouadi, "Modeling and analysis of a regenerative braking system with a battery-su-

percapacitor energy storage", Proceedings of the 7th International Conference on Modeling, Simulation, and Applied Optimization, Sharjah, United Arab Emirates, 4-6 April 2017, pp. 1-6.

[9] B. Noh, "Supercapacitor Assisted Hybrid Electric Vehicle Powertrain and Power Selection using Fuzzy Rule-Based Algorithm", Proceedings of the AEIT International Conference of Electrical and Electronic Technologies for Automotive, Turin, Italy, 18-20 November 2020, pp. 1-5.

[10] W. Zhao, G. Wu, C. Wang, L. Yu, Y. Li, "Energy transfer and utilization efficiency of regenerative braking with hybrid energy storage system", Journal of Power Sources, Vol. 427, 2019, pp. 174-183.

[11] J. Partridge, D. I. Abouelamaimen, "The Role of Supercapacitors in Regenerative Braking Systems", Energies, Vol. 12, 2019, p. 2683.

[12] Q. Li, W. Huang, W. Chen, Y. Yan, W. Shang, M. Li, "Regenerative braking energy recovery strategy based on Pontryagin's minimum principle for fell cell/supercapacitor hybrid locomotive", International Journal of Hydrogen Energy, Vol. 44, No. 11, 2019, pp. 5454-5461.

[13] F. Naseri, E. Farjah, T. Ghanbari, "An efficient regenerative braking system based on battery/supercapacitor for electric, hybrid, and plug-in hybrid electric vehicles with BLDC motor", IEEE Transactions on Vehicular Technology, Vol. 66, No. 5, 2016, pp. 3724-3738.

[14] D. A. Permana, U. Sholahuddin, D. Hamdani, T. D. Rachmildha, "Study of Supercapacitor Utilization on Regenerative Braking System: Design and Simulation", Proceedings of the 5th International Conference on Electric Vehicular Technology, Surakarta, Indonesia, 30-31 October 2018, pp. 88-94.

[15] J. Hamid, R. Sheeba, S. Sofiya, "Energy Harvesting through Regenerative Braking using Hybrid Storage System in Electric Vehicles", Proceedings of the IEEE International Conference on Intelligent Techniques in Control, Optimization and Signal Processing, Tamilnadu, India, 11-13 April 2019, pp. 1-6.

[16] E. Pipitone, G. Vitale, "A regenerative braking system for internal combustion engine vehicles using supercapacitors as energy storage elements - Part 1: System analysis and modelling", Journal of Power Sources, Vol. 448, 2020, p. 227368.

[17] H. Beltran, S. Harrison, A. Egea-Alvarez, L. Xu, "Techno-economic assessment of energy storage technologies for inertia response and frequency support from wind farms", Energies, Vol. 13, No. 13, 2020, p. 3421

[18] F. Naseri, E. Farjah, T. Ghanbari, "An efficient regenerative braking system based on battery/supercapacitor for electric, hybrid, and plug-in hybrid electric vehicles with BLDC motor", IEEE Transactions on Vehicular Technology, Vol. 66, No. 5, 2017, pp. 3724-38.

[19] T. Ibrahim, D. Stroe, T. Kerekes, D. Sera, S. Spataru, "An overview of supercapacitors for integrated PV–energy storage panels", Proceedings of the IEEE 19th international power electronics and motion control conference, Gliwice, Poland, 25-29 April 2021, p. 828-35.

[20] C. Pinto, J. V. Barreras, R. de Castro, E. Schaltz, S. J. Andreasen, R. E. Araújo, "Influence of Li-ion battery models in the sizing of hybrid storage systems with supercapacitors", Proceedings of the IEEE Vehicle Power and Propulsion Conference, Coimbra, Portugal, 27-30 October 2014, pp. 1-6.

[21] P. Sangyoung, K. Younghyun, C. Naehyuck, "Hybrid Energy Storage and Battery Management for Electric Vehicles", Proceedings of the 50th Annual Design Automation Conference, Austin, TX, USA, 29 May - 7 June 2013.

[22] Z. Zou, J. Cao, B. Cao, W. Chen, "Evaluation strategy of regenerative braking energy for supercapacitor vehicle", ISA Transactions, Vol. 55, 2015, pp. 234-40.

[23] M. Conte, A. Genovese, F. Ortenzi, F. Vellucci, "Hybrid battery-supercapacitor storage for an electric forklift: a life-cycle cost assessment", Journal of Applied Electrochemistry, Vol. 44, No. 4, 2014, pp. 523-532.

[24] J. Park, B. Raju, A. Emadi, "Effects of an ultra-capacitor and battery energy storage system in a hybrid electric vehicle", SAE Technical Paper, 0148-7191, 2005.

[25] J. Cao, A. Emadi, "A new battery/ultracapacitor hybrid energy storage system for electric, hybrid,

and plug-in hybrid electric vehicles", IEEE Transactions on Power Electronics, Vol. 27, No. 1, 2012, pp. 122-32.

[26] W. Jing, C. H. Lai, W. S. H. Wong, M. L. Dennis Wong, "A comprehensive study of battery-supercapacitor hybrid energy storage system for standalone PV power system in rural electrification", Applied Energy, Vol. 224, 2018, pp. 340-356.

[27] W. Jing, C. H. Lai, W. S. H. Wong, M. L. Dennis Wong, "Battery-supercapacitor hybrid energy storage system in standalone DC microgrids: a review", IET Renewable Power Generation, Vol. 11, No. 4, 2017, pp. 461-469.

[28] K. Jayasawal, A. K. Karna, K. B. Thapa, "Topologies for Interfacing Supercapacitor and Battery in Hybrid Electric Vehicle Applications: An Overview", Proceedings of the International Conference on Sustainable Energy and Future Electric Transportation, Hyderabad, India, 2021, pp. 1-6.

[29] Z. Cabrane, M. Ouassaid, M. Maaroufi, "Analysis and evaluation of battery-supercapacitor hybrid energy storage system for photovoltaic installation", International Journal of Hydrogen Energy, Vol. 41, No. 45, 2016, pp. 20897-20907.

[30] L. Brabetz, M. Ayeb, D. Tellmann, "Efficient Vehicle Power Supply by Adaptive Energy, Charge and Heat Management of an Alternator - Super Capacitor System", SAE International Journal of Passenger Cars - Electronic and Electrical Systems, Vol. 2, No. 1, pp. 359-366.

[31] Y. Zhang, Z. Jiang, X. Yu, "Control Strategies for Battery/Supercapacitor Hybrid Energy Storage Systems", Proceedings of the IEEE Energy 2030 Conference, Atlanta, GA, USA, 17-18 November 2008, pp. 1-6.

[32] R. Carter, A. Cruden, "Strategies for control of a battery/ supercapacitor system in an electric vehicle", Proceedings of the International Symposium on Power Electronics, Electrical Drives, Automation and Motion, Ischia, Italy, 11-13 June 2008, pp. 727-732.

[33] A. Musat, M. Carp, P. Borza, R. Musat, D. Sojref, "Hybrid storage systems and dynamic adapting topologies for vehicle applications", Proceedings of the 13th International Conference on Optimization of Electrical and Electronic Equipment, 24-26 May 2012 , pp. 1559-1566.

[34] M. G. S. P. Paredes, J. A. Pomilio, A. Santos, "Combined regenerative and mechanical braking in Electric Vehicle", Proceedings of the Brazilian Power Electronics Conference, Gramado, Brazil, 27-31 October 2013, pp. 935-941.

[35] C. Farcas, D. Petreus, I. Ciocan, N. Palaghita, "Modeling and simulation of supercapacitors", Proceedings of the 15th International Symposium On Design and Technology of Electronics Packages, Gyula, Hungary, 17-20 September 2009, pp.195-200.

[36] M. Choi, S. Kim, S. Seo, "Energy management optimization in a battery/supercapacitor hybrid energy storage system", IEEE Transactions on Smart Grid, Vol. 3, No. 1, 2012, pp. 463-472.

[37] A. Wangsupphapholl, N. R. N. Idris, A. Jusoh, N. D. Muhamad, L. W. Yao, "The Energy Management Control Strategy for Electric Vehicle Applications", Proceedings of the International Conference and Utility Exhibition 2014 on Green Energy for Sustainable Development, Pattaya, Thailand, 19-21 March 2014.

[38] D. Montesinos-Miracle, J. Bergas-Jane, S. Galceran-Arellano, A. Rufer, "Design and control of a Modular Multilevel DC/DC Converter for Regenerative Applications", IEEE Transactions on Power Electronics, Vol. 28, No. 8, 2013, pp 3970-3979.

Optimizing Enhanced Extended Topological Active Net Model Using Parallel Processing

Original Scientific Paper

Ranjita Akash Asati
Department of Computer Technology, Yeshwantrao Chavan College of Engineering
Nagpur, India
ranjita.asati@gmail.com

Dr. M.M. Raghuwanshi

Dr. K.R. Singh
Department of Computer Technology, Yeshwantrao Chavan College of Engineering
Nagpur, India

Abstract *– In numerous clinical applications that support the diagnosis and treatment planning of a broad variety of disorders, medical image segmentation is essential. Medical picture segmentation using the Enhanced Extended Topological Active Net (EETAN) model has proven to be successful in correctly identifying structures. This study suggests a novel way to combine the best clustering techniques and parallel processing approaches to maximize the segmentation performance of the EETAN model. The Probabilistic Depth Search Optimization (PDSO) Algorithm, which makes the parallel searching technique to find the ideal contour set, is responsible for this. This work implements parallel processing and ideal clustering to improve the EETAN model's performance in medical image segmentation. Performance metrics like accuracy, precision, recall, dice similarity, and computational time are used for a comparison study. The results demonstrate the notable enhancements attained by employing parallel processing and effective clustering.*

Keywords: *Image Segmentation, Parallel Computing, Probabilistic Depth Search Optimization (PDSO), Enhanced Extended Topological Active Net (EETAN)*

Received: December 26, 2023; Received in revised form: February 3, 2024; Accepted: February 20, 2024

1. INTRODUCTION

Medical picture segmentation is an essential part of diagnostic imaging that helps medical professionals identify exact anatomical structures and make clinical decisions based on those findings. The increase in volume and resolution of medical image datasets due to better medical imaging technology has made robust segmentation models necessary to handle the complexity of these datasets. Because it can capture complex structures in medical images, the Extended Topological Active Net (ETAN) model has emerged as a possible option [1]. However, it becomes more and more clear that more effective and scalable segmentation techniques are required as the size and complexity of medical datasets increase. Medical image data volume and resolution have increased exponentially as a result of the quick development of medical imaging technologies. For significant information to be extracted from these massive databases, effective segmentation techniques are crucial. A promising solution to the problems associated with medical image segmentation is the ETAN model, which is an extension of the Topological Active Net.

In contemporary healthcare, medical picture segmentation is an essential task that is critical to diagnosis, therapy planning, and image-guided therapies [2]. To extract relevant information from medical images and support healthcare practitioners in making decisions, accurate delineation of anatomical features is essential. The Extended Topological Active Net (ETAN) model is a useful tool for medical image segmentation because it is recognized for handling intricate anatomical components.

The amount and complexity of medical image data have significantly increased as a result of the growing use of high-resolution imaging modalities like computed tomography (CT) and magnetic resonance imaging (MRI). Although the ETAN model performs well in segmentation tasks, applying it to large-scale datasets may provide a challenge to its computing efficiency. The need for segmentation in real-time or almost real-time, especially in clinical contexts, makes it necessary to investigate novel

approaches to improve the ETAN model's speed and scalability. The necessity to close the gap between the ETAN model's promise and the changing needs of modern medical imaging drives this research. We seek to overcome the computational issues with the ETAN model and advance it toward a more effective and scalable solution for medical picture segmentation by utilizing parallel processing capabilities and implementing optimal clustering algorithms. Through this investigation, we aim to provide a valuable contribution to the continuous endeavors aimed at refining segmentation techniques, which will promote progress in medical imaging technology and ultimately enhance patient care.

The idea for this study came from the realization that when dealing with large-scale medical picture datasets, the traditional sequential implementation of the ETAN model can run into processing difficulties. To overcome this difficulty, the incorporation of parallel processing methods is investigated as a way to improve the ETAN model's computational effectiveness. The potential to improve segmentation accuracy within the parallelized framework is another reason for including optimal clustering techniques. The combination of best clustering and parallel processing is meant to offer a fast and precise way to segment medical images, which could lead to better clinical results.

We explore the complexities of the ETAN model in this work, recognizing both its advantages and disadvantages. The next step of our trip is to investigate how parallel processing and optimum clustering strategies might work together to overcome computational bottlenecks and improve the EETAN model's performance. The particle weight search parallel model improves performance by cutting down on time without compromising system accuracy. The PDSO optimization algorithm handled this. The approaches used, the integration of parallel processing and optimal clustering, and the thorough assessment of the suggested strategy using pertinent performance measures are all covered in detail in the sections that follow. The research's conclusions have potential ramifications for improving medical picture segmentation in the larger healthcare context in addition to aiding in the ETAN model's optimization.

One significant difficulty is the exponential increase in volume and complexity of medical image databases brought about by advances in imaging technologies. When used on these larger datasets, conventional segmentation models—such as the ETAN model—may encounter scalability and processing speed issues. The need for segmentation techniques that can successfully manage the inherent complexity of contemporary medical images is growing as the need for more thorough and detailed medical imaging increases. The ETAN model's traditional sequential processing may make it more difficult for it to deliver findings quickly, particularly in situations where making decisions quickly is essential, like in clinical settings. The goal of enabling quicker and more effective segmentation capabilities within the EETAN framework is what drives the investigation of parallel processing techniques as a means of overcoming these constraints. Because the PDSO uses many optimization objectives to simplify contour searching and detect particles simultaneously, it also optimizes time consumption. Moreover, the realization that improving segmentation accuracy is just as important is what propels the incorporation of the best clustering approaches. Contour searching is the depiction of the boundaries or outlines of objects or shapes within an image. Contours are used to represent the structural information of objects within an image and are valuable in image segmentation. These are formed by connecting adjacent points with similar pixel intensity. So set of coordinates that outline the boundary of an object helps in segmenting objects from the background in an image.

In the ETAN algorithm contours are represented as a series of connected points in Cartesian coordinates. In the ETAN algorithm, Chan-Vese segmentation is applied to the image. Chan-Vese segmentation is a level set-based image segmentation method that partitions an image into regions based on intensity homogeneity. The Chan-Vese segmentation algorithm minimizes an energy function that consists of two terms: an internal term promoting smoothness within regions and an external term penalizing deviations from a given intensity or gray level. This is employed to evolve a contour or boundary that separates different regions in the image. We seek to improve the accuracy and consistency of the segmentation outputs generated by the EETAN model by integrating sophisticated clustering techniques, guaranteeing that the segmented structures closely match the ground truth.

2. ENHANCED EXTENDED TOPOLOGICAL ACTIVE NET MODEL

Building on the fundamental ideas of the classic Topological Active Net, the Enhanced Extended Topological Active Net (EETAN) model offers a comprehensive and flexible method for medical image segmentation. The EETAN model was created to address the difficulties presented by complex anatomical features and intensity variations in medical pictures. It accomplishes accurate and thorough segmentation results by combining topological geometry with dynamic contour evolution. The core idea of the EETAN model is the use of deformable contour representations that dynamically change across repetitions while successfully respecting anatomical structure boundaries. The EETAN model is unique in that it incorporates topological flexibility, which makes it possible to depict several interrelated components and makes it easier to delineate complicated systems subtly. The EETAN model is well-suited to the challenges presented by medical imaging datasets because it incorporates topological information into the segmentation process, which enhances its ability to capture fine-grained anatomical characteristics. The EETAN model's principal strength is its high degree of realism while handling complex anatomy. Multiple connected component scenarios may be difficult for typical active contour models to handle, but the topological adaptabil-

ity of the EETAN model allows it to navigate and define such systems with accuracy. Because of its versatility, the EETAN model can be used for a variety of medical imaging tasks, such as organ, tissue, and lesion segmentation. The EETAN model has limits, especially concerning computational efficiency. The following sections of this paper explore these limitations and suggest creative solutions that make use of parallel processing and efficient clustering approaches. With these improvements, the EETAN model should be able to meet the changing needs of modern healthcare by being more computationally efficient, scalable, and adaptive for medical image segmentation. The EETAN model is an advanced framework in the field of medical image segmentation that was created to tackle the difficulties caused by intricate anatomical structures and different levels of intensity in medical images. By utilizing active contour evolution and topological geometry, the EETAN model achieves reliable and precise segmentation outcomes. Although the EETAN model performs admirably in segmentation, it is not without flaws. Its computational efficiency is one of its main limitations, particularly when dealing with large-scale medical imaging datasets. Traditional implementations' sequential design may cause extended processing times, which would make the model less applicable in situations where outcomes are crucial, including in healthcare settings. Furthermore, as medical datasets get more complex, the model may encounter difficulties with scalability and adaptability. To improve the performance of the EETAN model, this research attempts to investigate fresh methodologies, particularly parallel processing approaches and optimal clustering algorithms. Through the resolution of these issues, we want to fully realize the promise of the EETAN model and further the development of sophisticated segmentation techniques for medical imaging. The approaches used to include parallel processing and optimal clustering into the EETAN model are described in depth in the following sections, which aim to improve its usability in modern healthcare applications while reducing its drawbacks. The EETAN model's intrinsic dependency on sequential processing is one of its main drawbacks. Processing duration may be prolonged due to the computing demands resulting from the expansion of medical imaging collections with higher quality and complexity. This presents difficulties, especially in situations involving patients when outcomes are crucial for making well-informed decisions. Concerns about the EETAN model's scalability arise as medical datasets grow larger. The accuracy of segmentation can be affected by the EETAN model's sensitivity to initialization factors. In circumstances of unclear or difficult anatomy, suboptimal initialization might result in contour convergence problems that affect the model's capacity to precisely define structures. Several user-defined parameters in the EETAN model may need to be fine-tuned depending on the particulars of the medical imaging task at hand. Due to the model's sensitivity to these factors, precise calibration is required, which adds a level of subjectivity and may make it difficult to achieve the best results in various applications. Improving the EETAN model's performance requires addressing these constraints. To address these limitations, the following sections of this study investigate how parallel processing and appropriate clustering strategies might be combined to improve the overall effectiveness and suitability of the EETAN model for medical image segmentation.

3. PARALLEL PROCESSING TECHNIQUES

In response to the computational challenges posed by the EETAN model, parallel processing techniques are explored to harness the power of concurrent computation, accelerate segmentation tasks, and address the growing demands of large-scale medical image datasets. The parallelization of the EETAN model is essential to overcome computational bottlenecks and enhance its efficiency in handling large-scale medical image datasets. The parallelization of the EETAN model through multi-threading, GPU acceleration, task parallelism, and data parallelism collectively address the computational challenges. The subsequent integration with optimal clustering techniques further refines segmentation accuracy. In the multithreading approach, the segmentation algorithm is decomposed into concurrent threads allowing for the simultaneous execution of independent tasks. This approach enhances the utilization of multi-core processors resulting in faster iterations and reduced overall processing time.

The segmentation process within the EETAN model involves iterative tasks providing opportunities for task parallelism. Each iteration can be treated as an independent task enabling concurrent execution and reducing the overall processing time. Efficient load-balancing mechanisms are implemented to ensure optimal resource utilization and performance. Dynamic task scheduling mechanisms are employed to adaptively distribute computational tasks based on workload variations. This ensures that processing units remain engaged and productive throughout the segmentation process. The dynamic load balancing mitigates the risk of idle resources and maximizes the utilization of available computational power.

The data parallelism is achieved by partitioning the medical image dataset into smaller subsets and each subset is then processed independently by different processing units enabling parallel execution of segmentation tasks. The results from each subset are aggregated to produce the final segmented output. This approach enhances scalability and facilitates the efficient processing of large and high-resolution medical image datasets. The deformable contour evolution, a core component of the EETAN model is parallelized by distributing contour evolution tasks across processing units which accelerates the convergence of contours to anatomical boundaries contributing to faster and more efficient segmentation [3].

In conjunction with parallel processing, the integration of optimal clustering techniques into the EETAN model plays a crucial role in refining segmentation accuracy and addressing challenges associated with complex anatomical structures. Various clustering algorithms are explored

to enhance the robustness of the segmentation process [4]. The k-means clustering is applied to group pixels based on intensity similarities. By partitioning the image into clusters with similar intensity levels the EETAN model benefits from improved discrimination between different tissue types. To mitigate sensitivity to initialization robust initialization strategies are employed for K-Means clustering. Smart initialization methods such as K-Means++ are implemented to improve convergence speed and ensure representative cluster assignments [5]. Hierarchical clustering is integrated to capture structural relationships within the image data. This method enhances the adaptability of the EETAN model to complex anatomies by incorporating information at different hierarchical levels contributing to more nuanced segmentation results.

The Fuzzy C-Means clustering is introduced to handle pixel memberships with degrees of uncertainty [5, 6]. This is particularly beneficial in regions where anatomical boundaries are ambiguous. The fuzzy clustering approach allows the model to represent partial memberships, enabling a more nuanced and accurate representation of anatomical structures. By combining optimal clustering with parallel processing, the segmentation algorithm gains the advantages of both enhanced accuracy from clustering and accelerated computation from parallelization. The resulting synergy creates a comprehensive solution for image segmentation demonstrating improved efficiency and adaptability across diverse datasets [7]. Clustering algorithms such as K-Means [8], hierarchical clustering, and fuzzy C-Means [9] are integrated into the parallelized segments of the EETAN model. A hybrid parallelization approach is implemented combining the strengths of both CPU and GPU architectures. CPU-based parallelization handles complex control logic, task scheduling, and communication, while GPU acceleration is employed for computationally intensive tasks within both clustering and segmentation components. Load balancing mechanisms dynamically adjust the workload distribution based on the capabilities of CPU and GPU units preventing resource underutilization and optimizing overall processing speed [10]. The performance metrics including segmentation accuracy, precision, recall, and computational time are employed to quantitatively assess the effectiveness of the integrated parallel processing and clustering approach.

4. METHODOLOGY

The methodology, which uses the Probabilistic Depth Search Optimization (PDSO) algorithm in Figure 1, describes the methodical approach used to include parallel processing techniques and effective clustering approaches in the EETAN model. PDSO algorithm is based on Particle swarm optimization (PSO) algorithm, where particles represent potential solutions that evolve over iterations. Particles for the optimization model are initialized with random coordinate positions within a specified range. The particles are split into random coordinates and start searching for updates. The completion time for each direction is estimated based on starting and finishing times. The difference between the traditional ETAN and the proposed EETAN is the implementation of contour updates by using PDSO. In the PDSO optimization, the Virtual Member set is initialized and processed as serial and parallel methods. The EETAN was implemented as a serial and parallel type of PDSO optimization to evaluate the time consumption. As per the implementation of PDSO in the EETAN model, this will enhance the performance level of segmentation. The actions made to improve segmentation accuracy and computing efficiency to create a comprehensive framework for advanced medical image analysis are covered in this part.

The main function of the PDSO optimization algorithm is to find the best particles for contour update. The estimation of contour update will be based on the update of particle weight value as calculated from the objective function. This will make the large searching process which leads to time consumption. In that, some of the parameters are independent to find the best contour. This we can compute by the separate thread which refers to parallel computing. This type of parallelization leads to reducing the time consumption for searching the best. This will update the weight and refer to the global parameters to validate the update of particles.

For the optimal searching process, the input of the function refers to the feature set, 'T' that was initialized as in equation (1).

$$T = \{T_1, T_2, \dots T_n\} \tag{1}$$

To achieve the parallel process of optimization, this initializes the virtual member set 'VM' which is represented in (2). The virtual member set (VM) represents the entities that interact with the feature set during the optimization process. In this algorithm, the virtual members are associated with different parameters for the processing of endoscopic images. These parameters include:

1. Contour Update Parameters:

- Parameters governing the update of contours in the images.

2. Optimization Parameters:

- Parameters controlling the behavior of the algorithm, such as swarm size, maximum iterations, and inertia weight.

$$VM = \{VM_1, VM_2, \dots VM_n\} \tag{2}$$

From these parameters, the particles for the optimization model can be initialized by the random coordinate positions as 'X_0'. This can be evaluated by the equation (3).

$$X_0 = X_{min} + (X_{max} - X_{min}) \times rand \tag{3}$$

From the particle update, the particles are now split in the random coordinates and start searching for the particle updates. For that, the particles are now updated in the right way by estimating the completion

time for each direction to search for the update which means the particles take less time to update the parameters. This is referred to in (4).

$$T_{CT(ij)} = FT\big(T(j)\big) - ST\big(T(j)\big) \tag{4}$$

Where,

ST – Starting time

FT – Finishing Time

i=1,2,…n // Index of the feature list

j=1,2,…m // Index of VM list.

From the estimation of completion time in (4), the maximum value is estimated, and this is the reference for other particles to update the parameters as in (5).

$$M_T = max\{T_{CT(ij)}\} \tag{5}$$

The average utilization of particles for each 'VM' is calculated to find the amount of usage that is to perform the update. Equation (6) and (7) represents the average utilization and response time for iteration count respectively.

$$Average\ Utilization_{VM} = \frac{\sum_{i=1}^{n} T_{CT(ij)}}{M_T \times m} \tag{6}$$

$$Response\ Time\ (RT) = n \times T_{CT(0)} \tag{7}$$

Then calculate the value of process in each particle as from (8). This is to refer to the weight value of each particle to find the best update of contour.

$$L\big(VM_j, t\big) = \frac{NT(T,t)}{SR\big(VM_j, t\big)} \tag{8}$$

Where,

NT – Number of features in each VM at a time instant.

SR – Service Rate running in a VM.

The particles in the VM need to estimate the over process and under process to estimate the limit of particle movement for each iteration count. This can be evaluated from (9) and (10).

$$OL_{VM(j)} = \begin{cases} VM(j), & if\ (L(j) > S(j) \\ \phi, & otherwise \end{cases} \tag{9}$$

$$UL_{VM(j)} = \begin{cases} VM(j), & if\ (L(j) < (S(j) - B) \\ \phi, & otherwise \end{cases} \tag{10}$$

Where,

$S(j)$ – Maximum size of each VM.

B – Boundary limit of VM.

OL – OverProcess of VM.

UL – UnderProcess of VM.

Based on this evaluation, the amount of resources that are scheduled and the available resources to apply search processes are estimated from (11) and (12).

$$T_{R_{Used}} = \sum_{i=1}^{n} T_i \tag{11}$$

$$T_{R_{Avail}} = \big(T_R - T_{R_{Used}}\big) \tag{12}$$

Arrange the logical Confirm List (CL) from the estimation of OL and UL. This is to eliminate the irrelevant features that are not updated in the contour validation. The list consists of logical values to represent the confirmed '1' and not allowed '0' respectively. This was represented in the equation (13).

$$CL = \begin{cases} 1, & if\ \big(OL_{VM(j)}(or)UL_{VM(j)} \neq \phi\big) \\ 0, & otherwise \end{cases} \tag{13}$$

Based on the evaluation factors, the particles are validated to perform the update of contour matrix that is to represent the relevant pixels that are matched with the neighboring. For that, the similarity between the patterns is referred to form the connected components that represent the contour for the current update. For the similarity measures, the angle of the findings 'θ' is estimated, and the similarity of particles. This was represented in (14) and (15).

$$\theta = \alpha \times angle + (1 - \alpha) \times \frac{T_{CT(ij)}}{M_T} \tag{14}$$

Where α – Similarity constant is 0.5.

$$angle = \cos^{-1}\left(T_{CT(ij)} \times \frac{T_{R_{avail}}}{\big(\big|T_{CT(ij)}\big|\big|T_{R_{avail}}\big|\big)}\right) \tag{15}$$

From the update of parameters, the fitness value is calculated as in (16).

$$Fitness = \frac{1}{M_T} \times Average\ Utilization_{VM} \tag{16}$$

The resource sequence are updated as in the format of elements in the array set by (17).

$$R_i^t = (R_i^1, R_i^2, R_i^3, \ldots, R_i^n) \tag{17}$$

Where,

$$R_i^1 = S_i^1 mod\ m \tag{18}$$

S_i^t – Discrete Permutation sequences for the feature-length 'n'.

The standard deviation of updated particles is to represent the evaluated points to update the particles for the next iteration position which is represented in (19).

$$\sigma = \sqrt{\frac{1}{m}\sum_{i=1}^{m}\big(PT_j - PT\big)^2} \tag{19}$$

$$\text{Where, } PT_j = \frac{L(VM_i, t)}{Average\ system\ process} \tag{20}$$

Equation (21) represents the normalized particle parameters for updating the coordinates of particles for the next iteration period.

$$PT = \sum_{i=1}^{m} PT_j \tag{21}$$

The detailed flowchart is presented in Fig. 1. The Kvasir-SEG dataset (size 46.2 MB) containing 1000 endoscopic gastrointestinal images and their corresponding ground truth from the Kvasir Dataset v2 is used. The resolution of the images contained in Kvasir-SEG varies from 332x487 to 1920x1072 pixels. Implementing the EETAN model incorporates topological processing on an image, including smoothening, adding borders, computing persistence diagrams, and segmenting

components based on topological features. The basis for such advancements is a comprehensive description of the sequential EETAN model. To parallelize the EETAN model, the particle searches are done in batches. The parameters such as the VM weights, Average particle utilization for each VM, and time estimation to update the particles for iteration count. This makes the PDSO optimization parameters independent of estimating the prediction of contour update. Utilizing resources as efficiently as possible is the goal of load-balancing systems.

The ETAN and the PDSO optimization techniques were combined to form the parallelized EETAN model. To work concurrently on segmented regions, the clustering approaches leverage task and data parallelism. If the robust initialization and parameter adjustment techniques are applied, the clustering results are more trustworthy. Between the segments and clustering sections, a feedback loop is established. Findings from clustering techniques have an iterative impact on the segmentation process. Flexible responses to evolving anatomical configurations and imaging qualities are ensured by mechanisms for dynamic parameter adjustments.

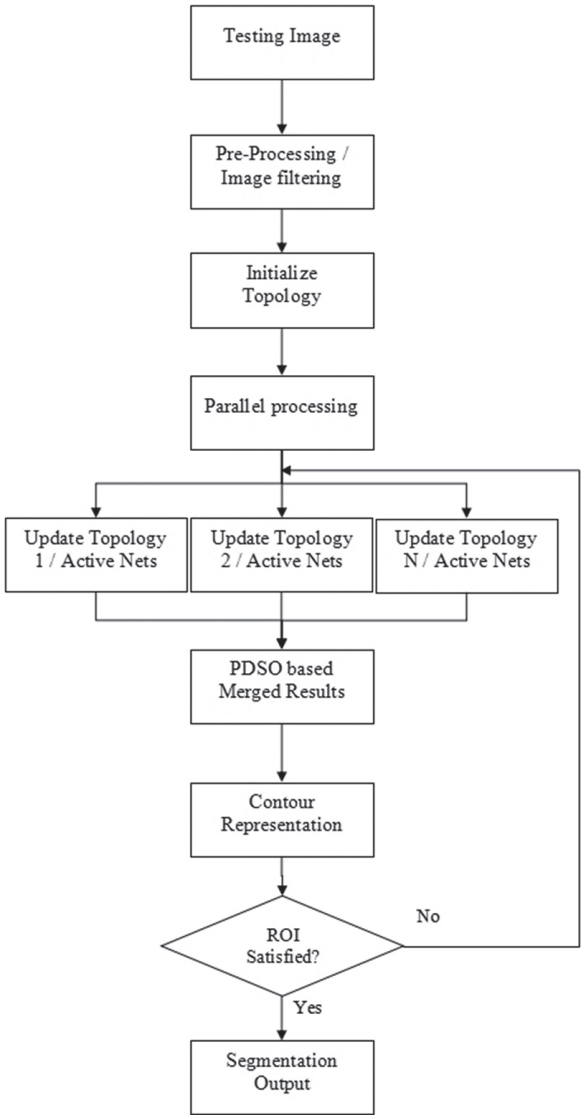


Fig 1. Flowchart for EETAN model

5. EXPERIMENTAL RESULTS

The experimental results provide a thorough assessment of the suggested methodology that integrates the best clustering algorithms and parallel processing techniques into the EETAN model for medical image segmentation. By showcasing increases in segmentation accuracy and processing efficiency, the trials hope to show how successful the integrated method is. The algorithm's output is compared to ground truth annotations to determine how accurate the segmentation results are. To measure segmentation accuracy, the Dice coefficient, and pixel-wise accuracy are calculated. We track and compare the segmentation computational time required with conventional ETAN models. Precision and Recall parameters are calculated by the pre-defined formula estimated from the confusion matrix which is framed by the difference estimation between the actual and predicted result from the algorithms. The actual result is represented as the ground truth and the predicted result is referred to as the clustered/segmented result from the result of simulation output. To assess the efficiency advantages attained through parallel processing, speedup ratios are computed. Consistency and reproducibility are ensured by the use of appropriate programming languages and frameworks in the implementation of the methodology. A thorough evaluation of the advancements made is possible through comparisons with alternative segmentation techniques and conventional EETAN models. A qualitative comprehension of the visual quality and correctness of the segmented structures is facilitated by visual inspection of the segmented findings, expert reviews, and comparisons against ground truth annotations. Python is used to implement the methodology for the EETAN model and parallel processing components.

Table 1 shows the time comparison of ETAN and EETAN for different methods on average time taken. For all these comparisons the EETAN statistics are best with the average time taken for EETAN_serial as 26.73 seconds and EETAN_parallel as 3.67 seconds so a reduction in time of image segmentation is nearly eight times whereas in Table 2 comparison based on percentage average accuracy, Kappa coefficients, MCC, TPR, and F1-macro is discussed and the various metrics in percentage resulted in accuracy as 95, Kappa coefficient as 73.96, MCC as 76.86, TPR as 96.87 and F1-macro as 86.87. These evaluation metrics are related to the quality of segmentation and found that in the proposed work they are improved so the quality of segmentation is improved. Values of parameters are same in serial and parallel because whether we run it in serial or parallel, it is not affecting the quality of segmentation. Table 3 compares the performance measures using the Kvasir-SEG dataset in precision, dice similarity, and recall with 98.24, 96.63, and 96.75 respectively. In Table 4 the comparison of segmentation results with the existing methods on the Kvasir-SEG dataset is presented in precision, dice similarity, and recall with 98.24, 96.19, and 96.75 respectively.

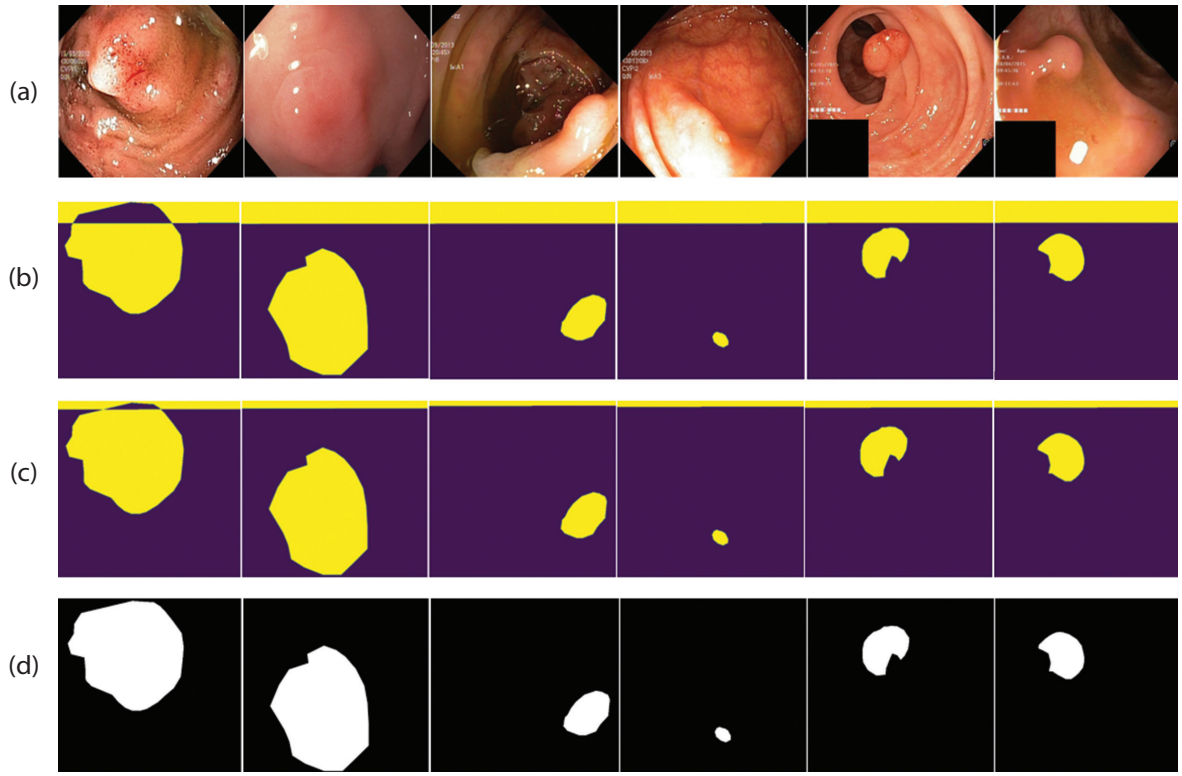


Fig. 2. Sample segmentation Results. (a) Input images, (b) Segmentation result from ETAN, (c) Segmentation result from EETAN, and (d) Ground-Truth

Table 1. Time Comparison of ETAN and EETAN

Methods	Average Time taken (Sec)
ETAN_Serial	26.73
ETAN_Parallel	5.73
EETAN_Serial	24.56
EETAN_Parallel	3.67

Table 2. Comparison of Parameters in Average

Methods	Average Accuracy (%)	Average Kappa Coefficient (%)	Average MCC (%)	Average TPR (%)	Average F1-Macro (%)
ETAN_ Serial	86.95	52.02	58.61	91.38	75.19
ETAN_ Parallel	86.95	52.02	58.61	91.38	75.19
EETAN_ Serial	95	73.96	76.86	96.87	86.87
EETAN_ Parallel	95	73.96	76.86	96.87	86.87

Table 3. Comparison of Performance Measures in Kvasir-Seg Dataset

Methods	Precision	Dice Similarity	Recall
CCS-Net [11]	92.47	90.89	91.41
CCS-Net with HFP [11]	96.81	92.93	90.01
MFRA-Net [12]	93.12	94.19	95.71
EETAN (Proposed)	98.24	96.63	96.75

Table 4. Comparing Segmentation Results with Existing Methods on The Kvasir-Seg Database

Methods	Precision	Dice Similarity	Recall
A-DenseUNet [13]	97.66	90.85	94.48
MSRF-Net [14]	96.66	92.17	91.98
ResUnet++ [11]	81.33	79.27	87.74
CCS-Net [11]	92.47	90.89	91.41
MFRA-Net [12]	93.12	94.19	95.71
EETAN (Proposed)	98.24	96.19	96.75

Visual inspection of segmented results supported by expert evaluations provides qualitative insights into the visual quality and accuracy of the segmentation outcomes. Comparative visualizations against traditional ETAN models and ground truth annotations help validate the improvements achieved through clustering and parallel processing. The experimental results provide a comprehensive evaluation of the proposed methodology, showcasing improvements in segmentation accuracy and computational efficiency. The subsequent sections of the research will discuss the implications of these findings, potential limitations, and avenues for future research in the context of medical image segmentation using the integrated approach. Comparative analyses in fig. 3-5 demonstrate improvements in segmentation accuracy achieved through the integration of optimal clustering techniques. PDSO optimization algorithm contributes to better discrimination of anatomical structures, reflected in higher Dice coefficient.

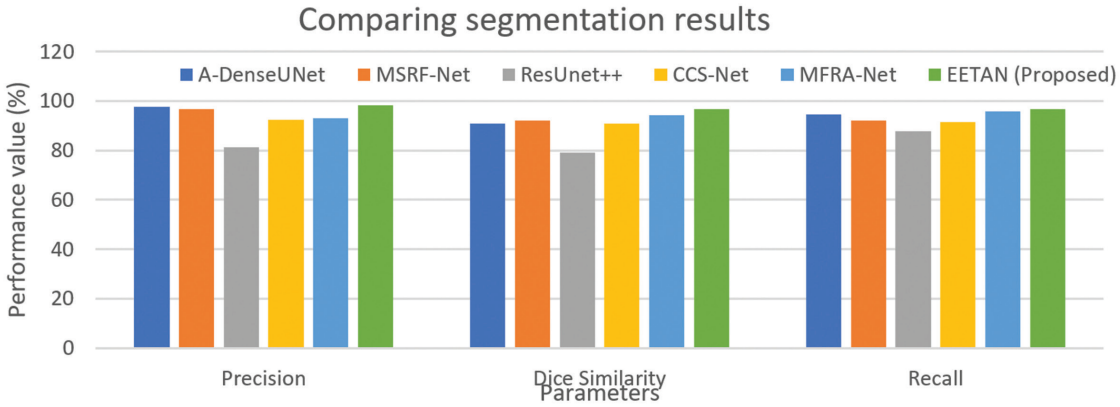


Fig. 3. Comparing segmentation results chart

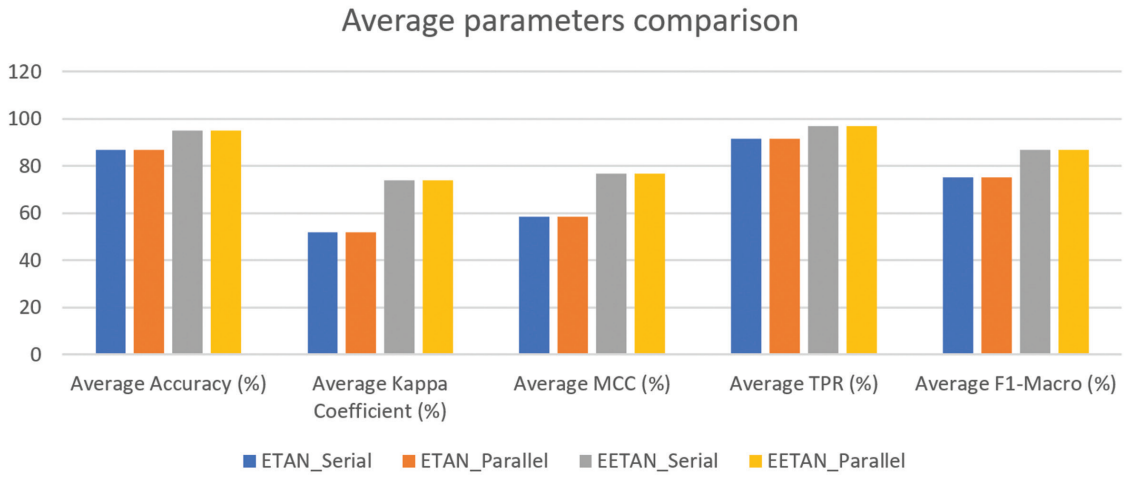


Fig. 4. Average parameters comparison chart

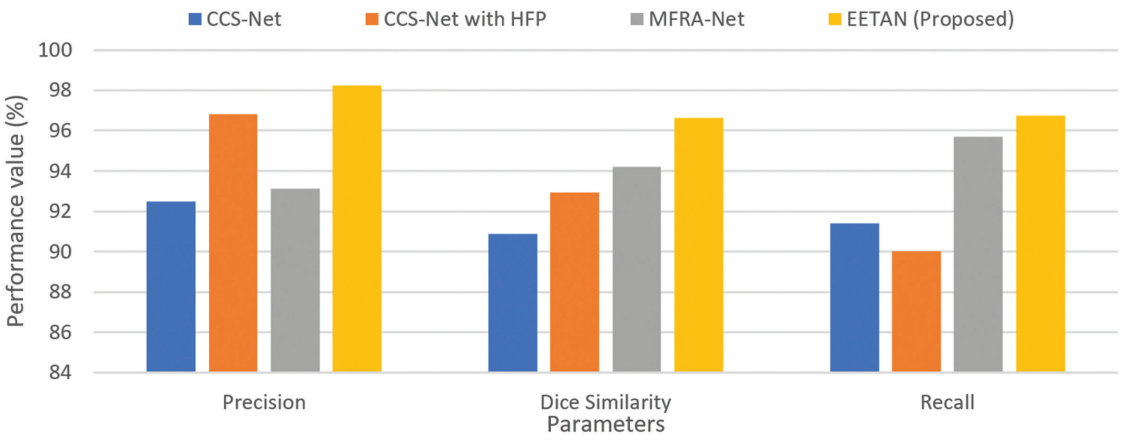


Fig. 5. Comparison of the performance chart

6. FUTURE CONSIDERATIONS

Even if the research to date shows encouraging developments, there is a need for more investigation and improvement. To achieve even greater efficiency and scalability, clustering algorithms and parallel processing techniques may be further optimized. Finally, the integration of deep learning approaches, such as convolutional neural networks, may be explored to improve the segmentation model's learning capacity. The study shows that the EETAN model's incorporation of optimal clustering and parallel processing provides a potent remedy for dealing with the difficulties presented by huge and intricate medical image collections. With this, the area of diagnostic imaging will benefit from more precise and efficient applications of image segmentation techniques.

7. CONCLUSION

This proposed work has investigated the incorporation of optimal clustering techniques and parallel processing techniques into the Enhanced Extended

Topological Active Nets model for medical image segmentation. The extensive studies and experimental results show notable improvements in segmentation accuracy and computing efficiency. The conclusions reached from this investigation are summed up in the following important elements. To increase the accuracy of segmentation, the PDSO optimization method is integrated and a concurrent process of particle searching for contour updates is carried out. A higher Dice coefficient is the outcome of improved anatomical structure discrimination made possible by these clustering techniques. The feedback loop that exists between segmentation and clustering makes it easier to develop the model iteratively, which improves its capacity to adapt to different anatomies. By using an integrated approach, segmentation accuracy is improved and a comprehensive framework that tackles the EETAN model's computational issues is created. When combined with appropriate clustering, the parallelized EETAN model has a synergistic impact that offers a comprehensive solution for advanced medical image segmentation.

8. REFERENCES

[1] N. Bova, Ó. Ibáñez, Ó. Cordón, "Extended Topological Active Nets", Image and Vision Computing, Vol. 31, No. 12, 2013, pp. 905-920.

[2] H. Hooda, O. P. Verma, "Fuzzy clustering using gravitational search algorithm for brain image segmentation", Multimedia Tools and Applications, Vol. 81, 2022, pp. 29633-29652.

[3] L. He et al. "A comparative study of deformable contour methods on medical image segmentation", Image and Vision Computing, Vol. 26, No. 2, 2008, pp. 141-163.

[4] H. Ahmad, S. Kim, S. J. Hong, "Development of two-phase flow regime map for thermally stimulated flows using deep learning and image segmentation technique", International Journal of Multiphase Flow, Vol. 146, 2022, p. 103869.

[5] H. Zhang, H. Li, N. Chen, S. Chen, J. Liu, "Novel fuzzy clustering algorithm with variable multi-pixel fitting spatial information for image segmentation", Pattern Recognition, Vol. 121, 2022, p. 108201.

[6] C. Kumah, N. Zhing, R. K. Raji, Z. Li, R. Pan, "Unsupervised segmentation of printed fabric patterns based on mean shift algorithm", The Journal of The Textile Institute, Vol. 113, No. 1, 2022, pp. 1-9.

[7] A. Haider, M. Arsalan, S. H. Nam, J. S. Hong, H. Sultan, K. R. Park, "Multi-scale feature retention and aggregation for colorectal cancer diagnosis using gastrointestinal images", Engineering Applications of Artificial Intelligence, Vol. 125, 2023, p. 106749.

[8] K. P. Sinaga, M.-S. Yang, "Unsupervised K-Means Clustering Algorithm", IEEE Access, Vol. 8, 2020, pp. 80716-80727.

[9] X.-Q. Tang, P. Zhu, "Hierarchical Clustering Problems and Analysis of Fuzzy Proximity Relation on Granular Space", IEEE Transactions on Fuzzy Systems, Vol. 21, No. 5, 2013, pp. 814-824.

[10] Z. Tang, L. Du, X. Zhang, L. Yang, K. Li, "AEML: An Acceleration Engine for Multi-GPU Load-Balancing in Distributed Heterogeneous Environment", IEEE Transactions on Computers, Vol. 71, No. 6, 2022, pp. 1344-1357.

[11] D. Jha et al. "ResUNet++: An Advanced Architecture for Medical Image Segmentation", Proceedings of the 21st IEEE International Symposium on Multimedia, San Diego, CA, USA, 9-11 December 2019.

[12] A. Haider et al. "Multi-scale feature retention and aggregation for colorectal cancer diagnosis using gastrointestinal images", Journal of Engineering Applications of Artificial Intelligence Vol. 125, 2023, p. 106749.

[13] S. Safarov, T. K. Whangbo, "A-DenseUNet: Adaptive Densely Connected UNet for Polyp Segmentation in Colonoscopy Images with Atrous Convolution", Sensors, Vol. 21, No. 4, 2021, p. 1441.

[14] A. Srivastava et al. "MSRF-Net: A Multi-Scale Residual Fusion Network for Biomedical Image Segmentation", IEEE Journal of Biomedical and Health Informatics, Vol. 26, No. 5, 2022, pp. 2252-2263.

DHM-OCR: A Deep Hybrid Model for Online Course Recommendation and Sustainable Development of Education

Original Scientific Paper

Sagar Mekala
Department of Computer Science and Engineering,
CVR College of Engineering, Telangana, Hyderabad, India
msagarphd@cvr.ac

Padma TNS
Assistant Professor, Department of CSE- Data Science,
Sreenidhi Institute of Science and Technology, Hyderabad, India

Rama Rao Tandu
Department of Computer Science and Engineering,
CVR College of Engineering, Telangana, Hyderabad, India

Abstract *– In recent years, there has been an increase in online education resources to help learners improve their skills. However, it is difficult to select the right course from available online education resources due to the demands and needs of learners with different knowledge domains. To solve this problem, an online course recommendation model has the important factor of enhancing learner's knowledge. Many existing recommendation systems (RS) use collaborative filtering (CF) to recommend courses to learners. The major problems with the Collaborative Filtering Recommendation System (CFRS) are the sparse preferences and the scalability of the data. According to the similarity of items, many recommendation models are proposed and developed, but none of these provide suggestions to users without their associations or preferences. We propose a deep hybrid model-online course recommendation (DHM-OCR) that uses high-level learner behavior and course objective features. We demonstrate the improvements and efficiency of the model for suggesting online e-learning courses. According to the analysis and evaluation results, it seems that our DHM-OCR outperforms the parallel research recommendation system. Experimental findings from online course data reveal that the suggested model and approach significantly improve classification accuracy and training efficiency, particularly limited available data.*

Keywords: *Recommendation System, Convolutional Neural Network, Content Based Recommendation System, Collaborative Filtering Recommendation System, Deep Hybrid Model, Ranking, Recurrent Neural Networks, Similarity, Preferences*

Received: November 1, 2023; Received in revised form: January 2, 2024; Accepted: January 5, 2024

1. INTRODUCTION

During and after the COVID-19 epidemic, the rapid development of online course providers focused on online course learners. As per parallel research on the online education system, those who have completed courses through the online learning system have benefited and enhanced their skill set in their domain area. Online education resources such as Udamy, MIT, Coursera, EdX, and Swayam are widespread and increasing in number. Courses learned through these sources, learners can develop academically and professionally. Many course modules and sub-modules are available through these online education platforms. These platforms offer hundreds of courses to learners. Parallel research results on online education platforms indicate that learners face great difficulty selecting an appropriate course due to the many online course providers and the large set of course data [1].

The recommendation system is a process of suggesting and/or predicting a priority set of items based on the user's interest and is constructed for many applications such as e-learning, games, movies, and e-commerce to simplify human daily life. As per the previous study, recommendation systems have received greater attention in academic and research circles. The recommendation system (RS) is a model that can obtain and produce a higher indexed and/or ranked collection of items in which users might be interested, according to their existing and present choices of items. The goal of RS is to make daily life easier for people by proposing

and/or forecasting a prioritized list of goods based on the user's interests. RS makes user's daily activities on the internet and social media easy. RS is used in many daily applications, such as news article suggestions, movie and music recommendations, e-learning, e-commerce, and many more. Online course recommendation models are divided into four basic approaches: content-based recommendation systems (CBRS), collaborative filtering recommendation systems (CFRS), hybrid recommendation systems (HRS), and sequence-based recommendation systems (SBRS). With recent advancements in information technology, many algorithms and schemes have been adapted to enhance the performance of RS.

The Content-Based Recommendation System (CBRS) approach is built based on users' preferences. In the CBRS model, the items are suggested based on the user's previous preference information, data, or user profile. According to the CBRS, suggestions can be made mainly from the data files and past preferences. The CBRS is useful to solve the problem of cold starts. In the Collaborative Filtering Recommendation System (CFRS) [2, 3] technique, items are suggested or predicted for a particular user when other similar kinds of users also prefer those items. The CFRS focuses mainly on finding similarities between learners and courses. The outcomes of CFRS conclude that learners with similar behaviors would take similar online courses [4]. It can recommend courses to target learners based on similar behaviors, but it is difficult to solve the cold start problems. A mixed approach combining CBRS and CFRS enhanced the performance of the model and reduced the cold-start problem. A specialized recommendation model proposed for course recommendation is called the Sequence Based Recommendation System (SBRS). It is mainly based on updating data through sequences, learning behavior sequences, and the time series approach of the learner. A tripartite graph, known as a location-query-browse graph (LQB), is proposed to provide sophisticated contextual suggestions [5]. REQUEST a query-based RS [6] proposed to utilize a multidimensional recommender system paradigm, including contextual dimensions, user and item dimensions, and OLAP-type aggregation and filtering capabilities. Social psychology-based CF for recommendations proposed in [7]. Authors [8], proposed a social recommender system that utilizes a temporal clustering technique.

Parallel research has discussed the applications of deep learning, collaborative filtering, and content-based recommendation systems in online course recommendation systems. For example, academics or industries have used CBRS to model learners' browsing histories and construct their preferences for courses. Shan et al. [9] proposed a model for e-learning recommendation based on CNN. CNN is used to extract course features from a text dataset of e-learning resources, such as introductions and content. The primary goal of the online course recommendation model is to predict learner preferences for personalized and preferable courses. For learners, they can study online through learning resources to gain knowledge and skills in their domain of interest. And for service providers, they can prepare and upload all kinds of course materials with different objectives.

The study on the integrated DHM framework with an intelligent recommendation system for online course assistance is strongly justified for several compelling reasons. The traditional online learning experience often lacks personalized guidance and tailored recommendations for individual learners. This study addresses the need to enhance the overall learning experience by providing intelligent recommendations that align with the unique preferences, capabilities, and learning styles of each student. With the proliferation of online courses, students may feel overwhelmed by the abundance of available resources. An intelligent recommendation system helps in efficient resource allocation, directing learners to the most relevant and beneficial courses and materials. The Deep Hybrid Model (DHM) has become a major domain of research because of the increasing amount of sequential data available in online learning education. DHM uses this to construct a learner's course preferences and interests based on their past course objectives and behavior. It is evident that CBRS and CFRS are accurate to some extent, but they fail to extract the complex features and dependencies in sequential data. All these online learning recommendation models have high content, but most of them use recommendation techniques without considering the distinction of behavior, preferences, objectives, and time of learners, which are crucial parts of learning in online learning. Most of the recommendation models for online learning don't consider the course objectives to be accurate and necessary to reach learner goals. For every course, there are certain objectives. If the similarity of two or more course objectives is high, such courses will be suggested to target and present learners. We proposed a deep hybrid model (DBM) that combines a Convolutional Neural Network (CNN) and a Recurrent Neural Network (RNN) to learn semantic features of learners and courses. And build a sequence model for the courses. In our point of view, users with similar learning behaviors have similar preferences, so our proposed DHM consists of CNN and RAN to learn the learner's behavior and focuses on time series. The primary contributions in this article are summarized as follows:

A framework for modeling and analyzing the e-learning learner-course induced data.

A deep hybrid model consisting of CNN and RNN is used together to construct an efficient learning process and enhance classification precision.

A DHM-OCR with CNN and RNN is designed to recognize and extract the combination of semantic and time series characteristics.

A recommendation model to predict learners with course guidance and course objective advice based on various training mechanisms for intelligent online course support

Successive sections in this paper are organized as follows. A brief introduction related to the work is presented in Section 2. We present the intelligent online course recommendation system framework with DHM in Section 3. In Section 4, we introduce a deep hybrid model experimental framework and evaluation results based on the online learner course. We conclude this article in Section 5.

2. RELATED WORK

A Recommendation System with a Deep Hybrid Model utilizes more than one deep learning approach. The advantage of DHM makes it possible to merge several neural building blocks to complement one another and form a more efficient hybrid model.

Garivaldis et al. [10], presented how existing learning theories might improve online education. In recent years huge surveys done and published papers on recommendation systems (RS), for instance, Adomavicius et al [11] presented a comparative survey on content-based, collaborative filtering, and hybrid recommendation models. The authors highlighted some of the important limitations of various recommendation systems. The possible merits and demerits of the hybrid recommender system model were studied by Burke et al [12]. Park et al [13], presented basic and real-world application areas of recommendation system approaches. Many algorithms and approaches have been published, but these papers impact a particular area of recommender system development. Zhang Q et al [14] analyzes commonly used E-learning recommendation algorithms and suggests new research possibilities. This study presents content-based (CBR), collaborative filtering-based (CFBR), and knowledge-based recommendations (KBR). These techniques and how they meet E-learning needs are explained.

To enhance the accuracy of RS a trust model collaborative filtering scheme is introduced by Jiang et al. [15]. Trust awareness-based CFRS implemented [16] for user behavior filtering. Social psychology-based CF for recommendations developed by Liu et al. [17]. Yang et al. [18], proposed a model to adapt SRL techniques to design a hybrid job recommender by combining content-based and collaborative filtering. Chadem et al. [19] implemented a recommendation system based on user-based and item-based collaborative filtering by introducing time decay, and item–to–item similarity models based on time intervals for suggesting top N-items.

Deshpande et al. [20], specifies the user-item matrix, item-rating matrix, and their binary values. To construct similarity matrices Sarwar et al. [21] use the cosine and correlation-based approaches. Hui Chen et al. [22], introduced a recommendation model to specify the behavior of e-learners and proposed an adaptive model based on learner behavior. Parvatikar et al. [23] have adapted different approaches for suggesting books to purchase. They primarily according to the book-based collaborative filtering apply data mining schemes. Implementation results of their approach simplify the scalability and data sparsity problems in Badarerah et al. [24] proposed a model for course selection and recommendation according to the similarities among learners and to construct the cluster with similar learners they used K-mean and KNN algorithms to get the most closed learner clusters. Initially, generates the learner clusters as per the e-learning style and performs collaborative filtering and association mining to retrieve the user behavior and interests.

Dien et al. [25], propose a deep matrix decomposition approach that recommends learning courses depending on learners' skills and needs. The authors provide to researcher with a simpler explanation of conventional matrix factorization and deep matrix factorization. The authors performed a comparison of the deep and conventional matrix factorizations. The presented model may be used for course recommendation and mapped to a recommendation system issue. Altaf et al. [1] presented various distinct models for course performance prediction. The methodology developed in this research is broad enough to produce accurate predictions based on many inputs. The proposed system uses 1D Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks to analyze online learning data, which is complicated and diverse. This research's hybrid deep learning approach has increased accuracy in visualizing advanced study specialists and aided weak pupils in achieving precision instruction. Li [26] a deep learning-based course recommender system (DECOR) is proposed to record user behaviors and course attributes. DECOR reduces information overload, solves high-dimensional data sparsity, and increases feature extraction performance. Biletskiy et al. [27], proposed ontological models of learners and learning objects, as well as methods for determining and adjusting preferences using these models. Based on a student profile and learning objectives, adaptive e-learning systems may automatically create individualized learning experiences [28]. Modern model-based hybrid recommendation systems need substantial feature engineering to create a user profile that combines these two filtering methods. Authors in [18, 29], suggested using SRL state-of-the-art techniques to build a hybrid course and job recommendation system. Helping recommender system interfaces grow is the goal of this TOCHI special area [30]. The Alternating Least Squares (ALS) algorithm-based collaborative filtering recommendation system prototype was proposed [31]. Multiple applications benefit from hybrid recommender systems. A hybrid fuzzy linguistic recommender system to assist the Technology Transfer Office personnel in distributing user-friendly research materials is proposed [32].

We advocate scaling educational best practices to additional students and institutions and developing sustainable future practices. Using enduring education theory and practice to develop learning methodology makes the past relevant to the present and future and allows innovation to be scaled and expanded from one impact to many.

3. DEEP HYBRID BASED MODEL FOR ONLINE COURSE RECOMMENDATION (DHM – OCR)

An introduction to recommender systems explains the current generation of content-based, collaborative, and hybrid recommendation techniques [11]. This study also analyzes present recommendation technique limitations and suggests additions to increase recommendation capabilities and make recommender systems more versatile.

In this section, we propose to design an intelligent approach to provide learners with automatic e-learning course resource guidance based on the Deep Hybrid Model-Online Course Recommendation System (DHM-OCRS). In many cases, learners may find it difficult to identify appropriate e-learning resources and to get the course updated on time. Here we try to identify an effective way to suggest a more efficient recommendation based on the learner's course learning process, and personalized course inquiry. The proposed DHM model consists of CNN, which is used to identify and extract features from data sources. RNN can model the feature dynamics and time series of content data. More than one deep learning algorithm is combined to construct effective recommendation models. These deep hybrid models consist of deep architectures that contain both discriminative and generative components.

3.1. ANALYSIS OF LEARNER – COURSE GENERATED DATA

The online course recommendation system has two main modules: an intelligent learner behavior learning module and an intelligent course suggestion module. The former is used to guide a learner to online learning resources and courses based on his/her inquiries, simulating the role of the guiding teacher, while the latter predicts medicine names based on learner behavior, simulating the role of the instructor. For text feature extraction, DCNN (deep convolutional neural network) uses two parallel CNNs to predict learner behavior and course objectives from the dataset. This model addresses the sparsity problem and enhances personalized recommendations by extracting similar features from learner feedback texts. The extracted features are passed through a single convolutional layer (SCL) with various kernels, a max-pooling layer (MPL), and a fully connected layer (FCL) consecutively. The outputs of the learner network (Li) and course network (Ci) are finally annotated. It is input to the prediction layer (RNN), where the matrix factorization technique is applied to combine learner and course interaction for rating prediction. In the following, we discuss in detail the main functional components of our DHM.

We utilize a convolutional neural network (CNN) to map learner preferences to other course vectors. Convolutional operations construct the considerable local similarities of a learner. Formally, CNN can learn the optimal learner behavior and preference specifications for course recommendation. The CNN contains an SCL, an MPL, a FCL, and a CL. The primary role of the convolutional layer is to obtain local features from the learner and course data. The CNN for learner-course local feature extraction can be defined in equation 1. RNN is another important approach of deep learning neural networks for text extraction, generally used on the data with a sequential representation. An RNN model consists of the RNN layer, FCL, and a CL. The primary job of the RNN layer is to obtain sequential dependencies. $X = F(Y, \Theta)$ of RNN can be specified in equation 2.

$$L_c = classifier(FC(Max\_Pool(CoNN(X,\theta_{CNN}),\theta_{MP}), \theta_{FC})) \quad (1)$$

$$L_c = classifier\ (FC(RNN(X, \theta_{RNN}), \theta_{FC})) \quad (2)$$

CNN models effectively use spatial data characteristics, such as pictures. Traditional CNN cannot handle sequential data. On the other hand, RNN-based models excel in modeling sequential data, such as conversations. To create a new model, DHM, a mixture of CNN and RNN is suggested in equation 3. Since the input data contains short learner questions and course datasets, we created the DHM model based on CNN and RNN. In our DHM, a max pooling layer precedes the RNN layer, generating two max pooling layers. The deep hybrid model is in equation 4.

$$Lc = classifier\ (FC(RNN(max\_pool\ (CNN(X, \theta_{CNN}), \theta_{pool}), \theta_{RNN}), \theta_{FC})) \quad (3)$$

$$Lc=classifier(FC(max\_pool\ (RNN(max\_pool\ (CNN(X, \theta_{CNN}), \theta_{pool}), \theta_{RNN}), \theta_{pool}),\theta_{FC})) \quad (4)$$

Where the classifier describes the classification layer, θ_{CNN} denotes collection of attributes of the function CNN, FC specifies the fully connected layer, θ_{FC} is collection of attributes of the function Fully Connected Layer (FC), Max_Pool denotes the pooling layer, and θ_{MP} denotes collection of attributes of the procedure max pooling layer., function RNN specifies RNN Layer, θ_{RNN} denotes collection of attributes of the procedure RNN, FC represents the fully connected layer, θ_{FC} collection of attributes of the method Fully Connected Layer. Semantic features are retrieved via the max-pool layer after convolution, followed by sequential RNN feature extraction and another max-pooling layer. The procedure emphasizes the combination of semantic and sequential elements to be compressed. The DHM effectively handles short input data sequences by placing the max-pooling layer with salient characteristics after the RNN layer. As an example, the training dataset may consist of labeled online learners' course questions. Set CNN as the convolutional layer and max-pool as the two pooling layers. A basic RNN is utilized for short consecutive words. Input data is sent over many layers

during training, including CNN, max_pool, RNN, and complete connection. Dropout is used to avoid overfitting in classification models. The suggested DHM algorithm for online course recommendation is defined in subsequent sections.

3.2. INTELLIGENT ONLINE COURSE RECOMMENDATION FRAMEWORK

The intelligent support module assigns learners to online learning courses based on their course history. The module receives learner course history and online learning materials (course feedback, grade, learning resource}) as input. The intelligent behavior module is responsible for pre-suggesting courses based on student inquiries and the teacher replies in the database. The module receives student course history and grade (course feedback, grade, learning resource). The suggested deep hybrid model is applied in both modules for the classification of short learner history, highlighting semantic and sequential aspects. The first module classification model assigns one learning resource per input history, whereas the second module classification model assigns one or more courses per learner.

The deep hybrid model in all modules uses a single-label classification model by assigning one learning resource to input data (i.e., courses). The multi-label classification methodology gives many categories (courses) to input data that course feedback, grades, and learning resources. Thus, we use Softmax in the first module of DHM's classifier layer and Sigmoid in the second module. In the training process, models are built using three steps: 1) Training the first deep hybrid model for inquiry-based classification using the dataset {Course - C; Learning Resource-L}. 2) Clustering data in learning resources into many groupings. 3) Train the second deep hybrid model for feedback-grade categorization using the dataset {Grade – G; Feedback - F; Course-C} in each group. For intelligent online course assistance, the first inquiry-based classifier will recommend the course recourse with the greatest likelihood of using Softmax for a given learner with a new query. Using the proposed deep hybrid model; the K-means algorithm clusters data into groups, selecting the most relevant one for the query q. Finally, the second feedback-grade classifier uses Sigmoid to propose courses with probabilities over a threshold. In the mathematical analysis for the proposed deep hybrid recommender, we consider the non-functional similarity to obtain the learner's behavior. The given dataset consists of 'm' course learners and 'n' courses offered in various platforms, this can be represented as m x n a matrix we call a learner-course matrix. The following equations from 5 to 8 demonstrate the learner-course matrix, the behavior of the learner is split into two types such as learner-based similarity (LSIM) and course-based similarity (CSIM) measures. Person Correlation Coefficient (PCC) was utilized in most of the RS for finding similarity, learner-based behavior measures the PCC can be computed to measure the behavior between two learners' u, and v by using.

$$LSIM(u,v) = \Sigma_{k\epsilon C}(P_{uk} - Q_u)(P_{vk} - Q_v)/\sqrt{A}\sqrt{B} \quad (5)$$

Where $\sqrt{A}$ is $\Sigma_{k\epsilon C}(P_{uk} - Q_u)^2$ and $\sqrt{B}$ is $\Sigma_{k\epsilon C}(P_{vk} - Q_v)^2$, and $C = Cu \cap Cv$ is the set of courses offered by learners u, v and P_{uk} is the satisfaction rating of the learner on the course C_k, and Q_u specifies the mean value of learner 'u' over the course in 'C_k'. Rates between -1 to1 are calculated using the behavior equation. In course-based similarity measure, the PCC can be evaluated to measure the learner behavior within any two courses by using.

$$CSIM(m,n)=\Sigma_{k\epsilon C}(P_{mk} - Q_m)(P_{nk} - Q_n)/\sqrt{A}\sqrt{B} \quad (6)$$

Where $\sqrt{A}$ is $\Sigma_{k\epsilon C}(P_{mk} - Q_m)^2$ and $\sqrt{B}$ is $\Sigma_{k\epsilon C}(P_{nk} - Q_n)2$, and $C = Cm \cap Cn$ is the set of courses both courses taken by learners u, v and Pmk is the satisfaction rating (feedback) of the learner on the course (Cm) taken by learner 'u' denoted as P_{mk}, Q_m and specifies the mean value of course 'm'. Similarly, the satisfaction rate of course (Cn) taken by learner 'v', is denoted as P_{nk} and Q_n specifies the mean evaluated value of the course (Cn) taken by learner 'v'. Re-ranking has been introduced to obtain similarity between target learners and courses based on the time of the satisfaction rating in the learner-course matrix. Then we perform computations of learner-based and course-based similarity measures. In learner-based similarity measures, the PCC is utilized to compute the similarity within the learners' 'u' and 'v' based on other courses.

$$LSIM(u,v) = 1 - \sqrt{\Sigma_{k\epsilon C}\{A\text{-}B\}} / \sqrt{|C|} \quad (7)$$

Where A is $(f_{uk} - f_{umin}) / (f_{umax} - f_{umin})$, B is $(f_{uk} - f_{umin}) / (f_{umax} - f_{umin})$, and $C = C_u \cap C_v$ is the set of courses taken by learners 'u' and 'v', the number of courses are denoted by $|C|$, and fuk represents the course 'k' with learner 'u' value in the modified learner-course matrix. The highest and lowest values for 'u' and 'v' are specified f_{umin}, f_{umax}, f_{umin}, and fumax respectively. The learner similarity (LSIM) is shown in equation 8.

$$LSIM(u,v) = \begin{cases} 0 & iff\ u_k \neq v_k \\ 1 & otherwise \end{cases} \quad (8)$$

Where 0 indicates that learners are dissimilar while considering courses (Ck), whereas 1 indicates both learners are similar by considering the course (Ck). The course matrix is constructed with rows consisting of courses and course objectives are separated by columns. These are important for constructing a course objective matrix. The course rating matrix for each learner according to the course ID is interpreted in a binary format. Every learner has rated one or more feedback rating each course. According to the course objective matrix and course rating matrix, we constructed the dot product matrix the result is again interpreted in binary format. An algorithm designed for building the hybrid recommendation model is as follows.

Input: 1. Identified Online Learning Resources with course objectives, course title, course Instructor, platform, and Course Identification (CID) as a

dataset. 2. Learner Objectives, history, goals, and Feedbacks as a dataset

Output: Top-N courses for target learner.

1. *Construct a Course Data Frame (CDF) from the various Learning Resources.*
2. *Construct a list that contains all the columns in the course.*
3. *For every Course Objective (CO), Course Feedbacks (CF), and Course Rating (CR) in course dataset.*
4. *if CO in CDF:*
5. *if CF is positive then:*
6. *CFeatures=Classifier(FC(MaxPool(RNN(MaxPool(CNN(CO, CF))))))*
7. *For every Learner Objective (LO), Learner History (LH) from Learner dataset.*
8. *if LO in LDF:*
9. *if LH is positive then:*
10. *LFeatures = Classifier(FC(MaxPool(RNN(MaxPool(CNN(LH, LO))))))*
11. *Step 5. Predicting_Learning_ Resource = Classifier(FC(RNN(MaxPool(CNN(LF,CF)))))*
12. *For every CO in CDF and CR.*
13. *13. for every CID:*
14. *if CR gt_Eql 3:*
15. *CR_matrix=1*
16. *else*
17. *17. CR_matrix = -1*
18. *Find the dot product between course matrix and CR_ matrix.*
19. *Convert the course dot matrix to binary format.*
20. *Find the Euclidian distance between the present learner and target learners.*
21. *Consider the rows with min(Euclidian distance): they are Top-N recommended courses for the target learner.*

This model based on the sequence of learners/users preferences and interests, and courses rated by the learners. Let us consider C as the set of courses, L, R, and P as the set of learners, rating and preferences in the system. Then $Sl = < x_1, x_2, x_3, \ldots, x_n >$ represents the sequence of rating of learner of l_i. $Slp = < p_1, p_2, p_3, \ldots, p_n >$ represents the sequence of preferences and interests of learner of l_i, and x_i in the above sequence indicates that $(C_i, r|p)$ specifies preference or rating of course given by learner l_i.

4. RESULTS

We undertake experiments to assess the performance of our suggested deep hybrid model and mechanism, comparing it to parallel research techniques. The most commonly performed operations are data preprocessing which is used for eliminating and avoiding incomplete data, and to find the order of course pattern events for each learner. For each course there will be a grade value between 0 - 1, we have converted grades into new grade scales. Evaluation trials acquired data from around 50 web pages, a realistic online learning service. There were 24,423 student feedback and grades. After screening data, including removing duplicates, incorrect formats, and misclassifications, 17,432 grades, and feedback were accessible for experiments and data descriptions from various online course resources. Out of 17,432 learner data, the average feedback length is 41, with 14,301 feedbacks ranging from 10 to 70. We utilize data from the "learning resources" to evaluate the performance of grade-feedback-based categorization in the intelligent course module. Based on 130 courses mentioned in learner feedback, we suggest the top 20 most commonly used courses based on 2654 samples. These courses include Python, C, C++, Java, Data Analysis, Big Data, Business Analysis, Deep Learning, Cloud Computing, Machine Learning, Cloud Analysis, Cyber Security, PHP, Pearl, R, Network Analysis, XML, SQL, Data Visualization, and Neural Network.

In our training system, input data is transformed to a 300 x 18 matrix for parameter setup. A 32-length vector describes each character in a sentence, with a maximum of 500 characters per phrase. The CNN convolutional layer uses 256 kernels (5 x 32) to generate 256 feature maps and extract textual information in a 496 x 256 matrix. Set the first max_pool layer to get 24 x 256 feature maps, and the basic RNN layer (RNN) to extract sequential features in a 24x32 matrix. A further max_pool layer compresses the output into a 32x1 matrix, integrating semantic and sequential data effectively. In inquiry-based classification, extracted characteristics are fed via a completely connected layer, forming a 256x1 vector, and categorized by the Softmax layer based on the six departments. The inquiry-answer-based classification classifies aspects using the sigmoid layer for the 20 courses mentioned. As previously mentioned, the initial intelligent learner module uses a single-label classification algorithm to propose courses based on learner feedback and grades in previous courses. We evaluate the recommendation performance of CNN and RNN neural network models with our proposed deep hybrid model on four categories (Table 2). Accuracy, Precision, Recall, and F1 are commonly used metrics to assess the performance of three classification models. Results may be measured using following equations from 9 to 12:

NTP: associated with learning resources and suggested

NFP: suggested but not part of the available learning resource;

NFN: not recommended, but associated with a learning resource;

NTN: not recommended and not related to learning resources.

$$Accuracy = NTP / TOTAL \quad (9)$$

$$Precision = NTP / (NTP + NFP) \quad (10)$$

$$Recall = NTP / (NTP + NFN) \quad (11)$$

$$F1=2 * Precision * Recall / (Precision + Recall) \quad (12)$$

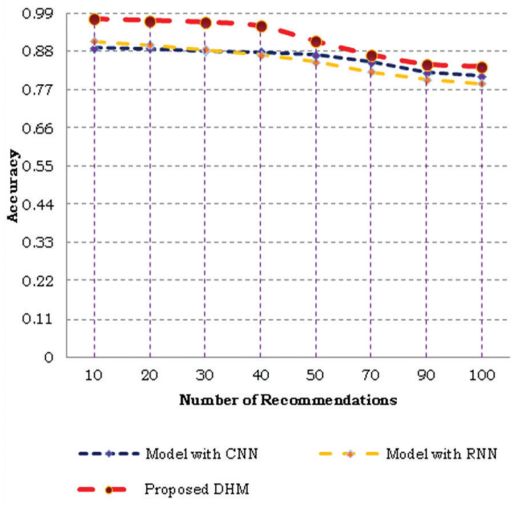
Table 2. Top 10 Recommendations Accuracy, Precision, Recall, and F1 Comparison among CNN, and RNN with proposed DHM-OCR

Course Name	Accuracy			Precision			Recall			F1		
	CNN	RNN	DHM	CNN	RNN	DHM	CNN	RNN	DHM	CNN	RNN	DHM
Data Science	0.879	0.911	0.974	0.891	0.901	0.972	0.881	0.934	0.965	0.904	0.918	0.964
Business Analysis	0.882	0.883	0.926	0.712	0.812	0.941	0.823	0.885	0.931	0.842	0.877	0.941
Machine Learning	0.929	0.932	0.944	0.909	0.949	0.967	0.909	0.898	0.946	0.871	0.895	0.946
Artificial Intelligence	0.911	0.887	0.948	0.891	0.911	0.934	0.889	0.810	0.946	0.871	0.888	0.937
Deep Learning	0.883	0.899	0.911	0.803	0.866	0.900	0.871	0.840	0.97	0.899	0.910	0.941
Web Development	0.871	0.881	0.954	0.851	0.887	0.912	0.836	0.810	0.837	0.830	0.810	0.868
Devops	0.886	0.898	0.921	0.896	0.866	0.895	0.867	0.816	0.816	0.816	0.799	0.833
Python Programming	0.885	0.895	0.949	0.885	0.918	0.932	0.885	0.811	0.821	0.905	0.905	0.912
Cyber Security	0.893	0.923	0.957	0.813	0.813	0.875	0.851	0.810	0.822	0.811	0.810	0.866
Data Visualization	0.893	0.890	0.933	0.898	0.754	0.801	0.888	0.821	0.833	0.877	0.867	0.919
Network Analysis	0.887	0.915	0.969	0.870	0.887	0.933	0.798	0.809	0.821	0.899	0.888	0.923
R	0.893	0.898	0.943	0.799	0.803	0.824	0.803	0.816	0.842	0.821	0.811	0.859
MAT Lab	0.865	0.915	0.960	0.790	0.821	0.868	0.845	0.837	0.878	0.910	0.899	0.922
Psychology	0.881	0.888	0.954	0.901	0.893	0.912	0.783	0.819	0.839	0.888	0.899	0.936
Marketing	0.892	0.879	0.945	0.892	0.910	0.926	0.881	0.801	0.866	0.865	0.851	0.895
Biology	0.890	0.876	0.955	0.755	0.846	0.875	0.789	0.827	0.850	0.805	0.833	0.878
C Programming	0.895	0.895	0.911	0.789	0.863	0.896	0.855	0.798	0.836	0.820	0.844	0.899

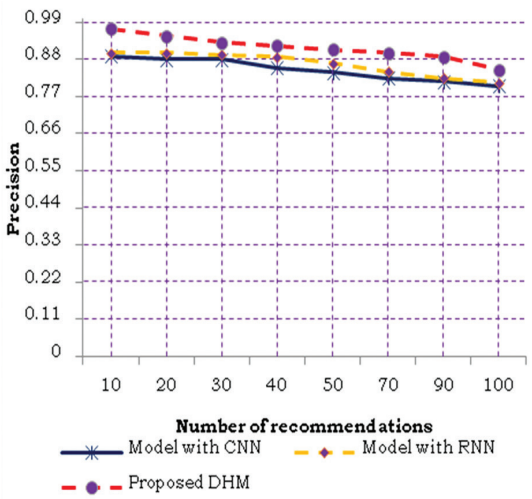
The testing data includes 6000 "feedback, grade, and learning resource" data from all learning resources, ensuring an average of 1000 samples per resource. In particular, the Stochastic Gradient Descent method and batch training are used during training. To address the imbalance in sample numbers, we guarantee that each batch has samples in each category throughout training. Accuracy, Precision, Recall, and F1 comparative data are in Table 1. The deep hybrid model performs best in Accuracy, Precision, Recall, and F1. The accuracy of the two baseline models across the above ten learning resources is above 0.96, whereas our suggested model averages 0.91. The suggested model achieves an average F1 value of 0.94, whereas the two baseline models among six learning resources have an F1 value above 0.92. This is because the deep hybrid model uses two max-pooling layers to highlight semantic and sequential features, which can better handle short input data lengths. In particular, our model performs better in "machine learning" and "data analysis" than in learning resources. Compared to other learning materials, learners in "machine learning" and "data analysis" may have a better understanding of their characteristics, which are easier to summarize.

The characteristics are complicated to convey in a few words. These findings suggest that providing intelligent assistance for diverse learners in online courses is challenging. Learners need time to improve their abilities. One group improves abilities throughout academics, while the other learns specialized skills for a short period as per projects, personal interests, and organizational desires. Learning goals are predicted over time using learner objectives, needs, and interests in this article. We are efficient in the first 10 recommendations, but as the number of ideas rises, the model's precision, accuracy, and recall diminish. CNN and RNN examine learner behavior over time. Based on this, each learner's group or cluster based on aims and interests might alter over time. To assess uniqueness, a 150-m dataset analysis was shown. CNN and RNN algorithms complete all tests 200 times.

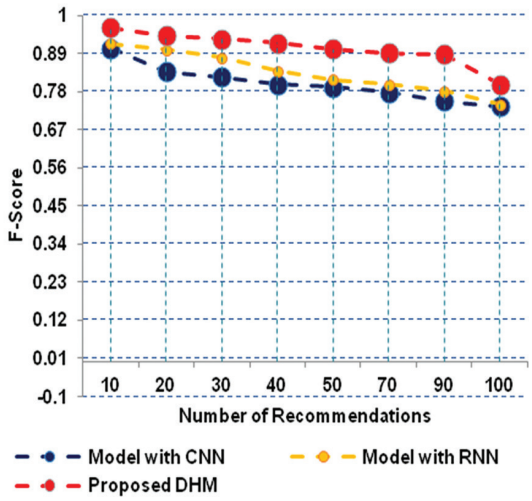
The DHM in Figures 3 (a) – 3 (d) and Table 2 underwent evaluation for precision, recall, F1-score, and accuracy. Compared to CNN and RNN models, the suggested DHM for different courses has the greatest precision, recall, accuracy, and F1 score. The first 10 recommendations based on DHM have an accuracy exceeding 0.964, but as the number of suggestions rises, all measures diminish. According to experiment findings, the suggested DHM outperforms CNN and RNN models in precision, recall, accuracy, and F1-score for different courses. According to Figures 3 (a) – 3 (d), the DHM system is extremely excellent in the top 10 recommendations.



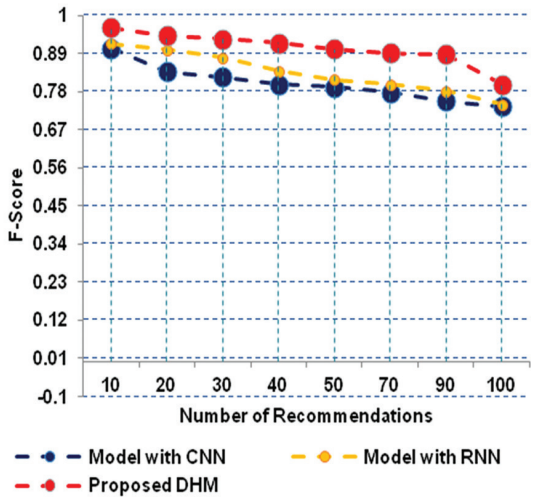
(a)



(b)



(c)



(d)

Fig. 1. (a) Accuracy (b) Precision (c) Recall and (d) F1 values of DHM compared with other approaches

5. CONCLUSION

This research introduces an integrated DHM framework with an intelligent recommendation system for online course assistance. CNN-RNN classification and clustering models modeled and analyzed course-resource-generated data. We created a novel CNN-RNN framework structure and devised a deep hybrid method to extract and highlight semantic and sequential information. The architecture for intelligent online course assistance includes a learning resource module that guides learners to the right course and a course module that predicts courses. A clustering model was created to optimize the learning process and reduce the course scope using more representative characteristics. A recommendation system was created to automatically offer learning materials and courses to learners. Experimental findings from online course data show that our suggested model and approach enhance classification accuracy and training efficiency, particularly for short input data. We want to address high-dimensional course-related data in future investigations. We will conduct evaluation tests to address challenging circumstances in online course learning data contexts.

6. REFERENCES

[1] S. Altaf, R. Asad, S. Ahmad, I. Ahmed, M. Abdollahian, M. Zaindin, "A Hybrid Framework of Deep Learning Techniques to Predict Online Performance of Learners during COVID-19 Pandemic", Sustainability, Vol. 15, 2023, p. 11731.

[2] Z. Tan, L. He, "An efficient similarity measure for user-based collaborative filtering recommender systems inspired by the physical resonance principle", IEEE Access, Vol. 5, 2017, pp. 27211-27228.

[3] Z. Liu, L. Wang, X. Li, S. Pang, "A multi-attribute personalized recommendation method for manufacturing service composition with combining collaborative filtering and genetic algorithm", Journal of Manufacturing Systems, Vol. 58, 2021, pp. 348-364.

[4] L. Baltrunas, F. Ricci, "Experimental evaluation of context-dependent collaborative filtering using item splitting", User Modeling and User-Adapted Interaction, Vol. 24, No. 1-2, 2014, pp. 7-34.

[5] Y. Ren, M. Tomko, F. D. Salim, J. Chan, C. L. A. Clarke, M. Sanderson, "A location-query-browse graph for a contextual recommendation," IEEE Transactions on Knowledge and Data Engineering, Vol. 30, No. 2, 2018, pp. 204-218.

[6] G. Adomavicius, A. Tuzhilin, R. Zheng, "REQUEST: A query language for customizing recommendations", Information Systems Research, Vol. 22, No. 1, 2011, pp. 99-117.

[7] N. Joorabloo, M. Jalili, Y. Ren, "A probabilistic graph-based method to improve recommender system accuracy", Proceedings of the International Conference on Engineering Applications of Neural Networks, Xersonisos, Greece, 24-26 May 2019, pp. 151-163.

[8] S. Ahmadian, N. Joorabloo, M. Jalili, M. Meghdadi, M. Afsharchi, Y. Ren, "A temporal clustering approach for social recommender systems", Proceedings of the IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, Barcelona, Spain, 28-31 August 2018, pp. 1139-1144.

[9] O. C. Santos, J. G. Boticario, D. Pérez-Marín, "Extending web-based educational systems with personalized support through User Centred Designed recommendations along the e-learning life cycle", Science of Computer Programming, Vol. 88, 2014, pp. 92-109.

[10] F. Garivaldis, S. McKenzie, D. Henriksen, S. Studente, "Achieving lasting education in the new digital learning world" Australasian Journal of Educational Technology, Vol. 38, No. 4, 2022.

[11] G. Adomavicius, A. Tuzhilin, "Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions", IEEE Transactions on Knowledge and Data Engineering, Vol. 17, 2005, pp. 734-749.

[12] R. Burke, "Hybrid recommender systems: survey and experiments", User Modeling and User-Adapted Interaction, Vol. 12, 2002, pp. 331-370.

[13] D. H. Park, H. K. Kim, I. Y. Choi, J. K. Kim, "A literature review and classification of recommender systems research", Expert Systems with Applications, Vol. 39, 2012, pp. 10059-10072.

[14] Q. Zhang, J. Lu, G. Q. Zhang, "Recommender Systems in E-learning", Journal of Smart Environments and Green Computing, Vol. 1, 2021, pp. 76-89.

[15] L. Jiang, Y. Cheng, L. Yang, J. Li, H. Yan, X. Wang, "A trust-based collaborative filtering algorithm for E-commerce recommendation system", Journal of Ambient Intelligence and Humanized Computing, Vol. 10, No. 8, 2019, pp. 3023-3034.

[16] L. Tang, Z. Duan, Y. Zhao, "Toward using social media to support ridesharing services: challenges and opportunities", Transportation Planning and Technology, Vol. 42, No. 4, 2019, pp. 355-379.

[17] F. Y. Liu, X. Q. Gao, Z. Zhang, "Improved Bayesian probabilistic model based recommender system", Computer Science, Vol. 44, No. 5, 2017, pp. 285-289.

[18] S. Yang, M. Korayem, K. AlJadda, T. Grainger, S. Natarajan, "Combining content-based and collaborative filtering for job recommendation system: A cost-sensitive Statistical Relational Learning approach", Knowledge-Based Systems, Vol. 136, 2017, pp. 37-45.

[19] C. Xia, X. Jiang, S. Liu, Z. Luo, Z. Yu, "Dynamic item-based recommendation algorithm with time decay", Proceedings of the Sixth International Conference on Natural Computation, Yantai, 10-12 August 2010, pp. 242-247.

[20] M. Deshpande, G. Karypis, "Item-Based Top-N Recommendation Algorithms", ACM Transactions on Information Systems, Vol. 22, No. 1, 2004, pp. 143-177.

[21] B. Sarwar, G. Karypis, J. Konstan, J. Riedl. "Item-Based Collaborative Filtering Recommendation Algorithms", Proceedings of the 10th international conference on World Wide Web, May 2001, pp. 285-295.

[22] Y. Ma, S. Ji, Y. Liang, J. Zhao, and Y. Cui, "A hybrid recommendation list aggregation algorithm for group recommendation", Proceedings of the IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology, Singapore, 6-9 December 2015, pp. 405-408.

[23] S. Parvatikar, B. Joshi, "Online book recommendation system by using collaborative filtering and association mining", Proceeding of the IEEE International Conference on Computational Intelligence and Computing Research, 10-12 December 2015, pp. 1-4.

[24] A. Al-Badarenah, J. Alsakran, "An Automated Recommender System for Course Selection", International Journal of Advanced Computer Science and Applications, Vol. 7, No. 3, 2016.

[25] T. T. Dien, N. Thanh-Hai, N. T. Nghe, "An approach for learning resource recommendation using

deep matrix factorization", Journal of Information and Telecommunication, Vol. 6, No. 4, 2022, pp. 381-398.

[26] Q. Li, J. Kim, "A Deep Learning-Based Course Recommender System for Sustainable Development in Education", Applied Sciences, Vol. 11, 2021, p. 8993.

[27] Y. Biletskiy, H. Baghi, I. Keleberda, M. Fleming, "An adjustable personalization of search and delivery of learning objects to learners", Expert Systems with Applications, Vol. 36, 2009, pp. 9113-9120.

[28] N. Capuano, M. Gaeta, P. Ritrovato, S. Salerno, "Elicitation of latent learning needs through learning goals recommendation", Computers in Human Behavior, Vol. 30, 2014, pp. 663-673.

[29] S. Yanga, M. Korayemb, K. AlJaddab, T. Grainger, S. Natarajan "Combining content-based and collaborative filtering for job recommendation system: A cost-sensitive Statistical Relational Learning approach", Knowledge-Based Systems, Vol. 136, 2017, pp. 37-45.

[30] P. Resnick, H. R. Varian, "Recommender systems – introduction to the special section", Communications of the ACM, Vol. 40, No. 3, 1997, pp. 56-58.

[31] J. B. Li, S. Y. Lin, Y. H. Hsu, Y. C. Huang, "An empirical study of alternating least squares collaborative filtering recommendation for movie lens on Apache Hadoop and spark," International Journal of Grid and Utility Computing, Vol. 11, No. 5, 2020, pp. 674-682.

[32] C. Porcel, A. Tejeda-Lorente, M. A. Martínez, E. Herrera-Viedma, "A hybrid recommender system for the selective dissemination of research resources in a technology transfer", Information Sciences, Vol. 184, 2012, No. 1, pp. 1-19.

Comparative Analysis of Banana Detection Models: Deep Learning and Darknet Algorithm

Original Scientific Paper

Abdul Haris Rangkuti
Bina Nusantara University
School of Computer Science, Computer Science
Jl Paskal, Bandung, Indonesia
rangku2000@binus.ac.id

Sian Lun Lau
Sunway University
School of Engineering and Technology
Kuala Lumpur, Malaysia
sianlunl@sunway.edu.my

Varyl Hasbi Athala
Bina Nusantara University
School of Computer Science, Computer Science
Jl Paskal, Bandung, Indonesia
varyl.athala@binus.ac.id

Rudi Aryanto
Bina Nusantara University
School of Computer Science, Computer Science
Jl Paskal, Bandung, Indonesia
rudiaryanto@binus.ac.id

Abstract *– This study aims to compare and evaluate the performance of banana detection models utilizing deep learning techniques and the Darknet algorithm. The objective is to identify the most effective approach for accurately detecting bananas in various real-world scenarios. The analysis involves training and testing multiple models using different datasets and evaluating their performance based on precision, recall, and overall accuracy. The results provide valuable insights into the strengths and weaknesses of each approach, enabling researchers and practitioners to make informed decisions when implementing banana detection systems. To detect banana objects, several convolutional neural network (CNN) models were used, including MobileNetV2, YOLOv3-Nano, YOLO Fastest 1.1, YOLOv3-tiny-PRN, YOLOv4-tiny, YOLOv7, and DenseNet121-YOLOv3. The training process utilizes the Darknet algorithm to facilitate the identification of banana types/classes captured by a camera, resulting in an MP4 film file. In this research, various experiments were carried out using different CNN models. However, these six models achieve optimal accuracy above 80%. Among them, the YOLOv7 model shows the highest average accuracy (MAP) at 100%, followed by the small model YOLOv4 at 92%. Meanwhile, for performance measurements, the accuracy of the YOLOv4-tiny model was 87%, followed by the YOLOv7 model at 84%. In the banana fruit experiment, several models showed very good performance, such as recognition of the Ambon, Kepok, and Emas banana classes up to 100% using the YOLOv7 and YOLOv4-tiny models. The YOLOv7 model itself can recognize other banana classes up to 100% in the Barangan, Rjbulu, Uli, and Tanduk classes. Furthermore, theYOLOv4-tiny model can recognize other banana classes, up to 90% of the Barangan, Rjbulu, Rjsereh, and Uli banana types. Thus, this experiment provides very good average accuracy results on 2 CNN models, namely YOLOv7 and YOLOv4-tiny. Future research will involve grouping pictures of bananas, which produces different image shapes, so it requires a different way to recognize them. It is hoped that this research can become a basis for further research in this field.*

Keywords*: information security, information system, security awareness, user behavior*

Received: November 16, 2023; Received in revised form: February 26, 2024; Accepted: February 26, 2024

1. INTRODUCTION

Bananas are the most popular type of fruit, in demand by all levels of society worldwide. Additionally, they offer numerous benefits as they are rich in essential nutrients for the human body. In general, bananas stand out as one of the most renowned fruits in Indonesia. The characteristic sweet taste of bananas makes them highly favored among Indonesian citizens, despite the fruit having various types with distinct shapes and colors. This support enables the automation of diverse tasks in agriculture and plantations with maximum accuracy, making smart farming concepts a reality. An overview of computer vision, coupled with the acquisition of high-quality images using remote cameras or drones, facilitates efficient non-contact and technology-based solutions in agriculture, plantations, and forestry [1].

An intelligent fruit ranking system is imperative due to the slow, labor-intensive, error-prone, and tedious nature of grading and sorting performed by humans in fruit categorization. Deep learning has proven to be highly effective in this area, with significant advancements in accurately and efficiently categorizing fruits. This plays a crucial role in agriculture, quality control, and automated fruit sorting systems, impacting fruit quality evaluation and the export market for producers [2-4]. A broad review of fruits vand vegetables among various horti-

cultural products in agriculture was carried out. The review covered specific models, data pre-processing, data analysis methods, and the overall value of performance accuracy using clear performance metrics [5]. Another research informs a convolutional neural network-based predictive model in identifying bruised apples based on shape information (in the form of three-dimensional [3D] surface meshes) obtained from a 3D infrared imaging system. Often there are irregularities on the surface of bruised apples, which can be used to distinguish bruised from unbruised apples [6]. The absence of an automatic system for classifying dates poses a challenge in the fruit industry, leading to the reliance on manual expertise, which is labor-intensive, costly, and prone to bias [7]. Based on the Machine Learning techniques development, it is possible to create automation in the agricultural sector, including especially fruit farming. An automatic system based on ML would effectively handle the process of classifying and sorting fruit which is often carried out by human experts [8]. This research is based on performance analysis of several deep learning models including the impact of various parameters on accuracy and efficiency in fruit categorization systems. However, Convolutional Neural Networks are used through different approaches in evaluating effectiveness [9].

This research identifies various types of bananas found in Indonesia, which is known as the producer of the most banana varieties in the world. The research is focused on detecting and classifying 9 types of bananas that are widely consumed by Indonesian people. The research aims to analyze the performance of deep learning models and the influence of various parameters on the accuracy and efficiency of fruit categorization systems. By understanding these factors, researchers can further improve the automation of fruit classification in industry.

2. RELATED STUDY

Bananas are widely known in Indonesia as a delicious fruit that can be enjoyed in various ways. This means that manual expertise is required, which involves hard work, costs, and potential bias [10]. The agricultural sector, especially fruit farming, has greatly benefited from the development of ML techniques. This progress has resulted in the emergence of automated systems that utilize ML to effectively carry out fruit classification and sorting tasks, which were previously reliant on human expertise [11]. The objective of this study is to analyse the performance of deep learning models and assess the impact of various parameters on the accuracy and efficiency of fruit categorization systems. Specifically, Convolutional Neural Networks with different approaches will be employed for evaluation to support the image recognition [12].

This application of AI in the food industry, specifically in maximizing resource utilization and reducing human error. By leveraging artificial intelligence and data science, the quality of restaurants, cafes, online food delivery aims, hotels, and food outlets can be improved through increased production using different pairing algorithms for sales prediction [13]. The important aspect of research is the training and processing part. A study found that by using a GPU as the main processing power, they achieved a 177x acceleration on training data and a 175x acceleration on test data [14]. To ensure the reliability of the dataset used, the researchers utilized the fruis-360 dataset, which had already been successfully used in previous research [15]. This dataset includes photos of 30 different fruit classes. The researchers employed prominent deep learning architectures such as VGG16 and ResNet50 to build their classification system. The models achieved 86% and 85% accuracy on the public dataset and 99% and 98% accuracy on their custom dataset [16]. Another study, a wider variety of fruit types was used. The experiment involved 24 classes of fruit, consisting of 3924 pictures. The authors pre-processed the data by applying augmentation techniques and trained a convolutional neural network (CNN) with a batch size of 16 and 100 epochs. The model achieved a test accuracy of 95.5% [17]. The Study used a similar fruit dataset and focused on various types of apples. The average accuracy values for training and test datasets were 100% and 73% respectively. There would be a difference between the normal image and the kernel image. The Kernel Process achieved the desired effects such as various blurring or sharpening effects [18].

In terms of image equalization, CLAHE (Contrast Limited Adaptive Histogram Equalization) differs from ordinary adaptive histogram equalization by using clipping histograms to equalize the image [19]. When it comes to computer vision-based technology for fruit identification and classification, one study focused on bananas. They aimed to build a computer vision-based model identified bananas images, to determine the variety of bananas, and assess the quality of bananas [20]. In the fruit detection for automatic harvesting, deep learning had been widely used due to its ability to extract high-dimensional features from fruit images. DL is employed to detect fruits in scenarios [21].

This research aims to determine the performance of detecting optimal banana image objects. Examples of banana images can be seen in Fig. 1.

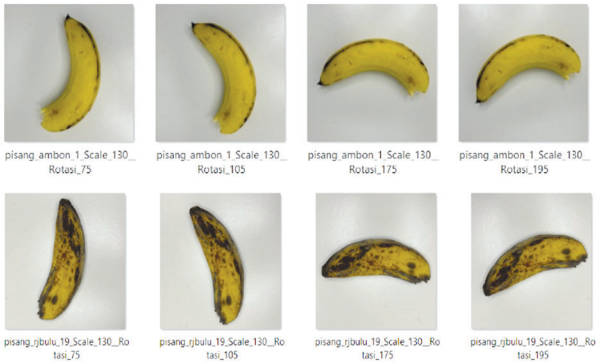


Fig. 1. The banana dataset consists of Ambon, Rjbulu and Rjsereh bananas

Fig. 1 provides examples of several images of several types of bananas that have undergone changes in rotation and scale from the original image. The rotations are 75, 105, 175 and 195 degrees and the scales are 130, 133, 137 and 140 percent of the original image. In this experiment there were 9 classes of banana datasets, each of which had different characteristics and orientation. From this dataset, 7504 images of bananas were obtained, consisting of 4506 training datasets, 1122 validation datasets and 1876 testing datasets. Thus, the system can detect 9 classes of banana objects with image conditions of different scales and rotations. An example of banana object detection can be seen in Fig. 2.

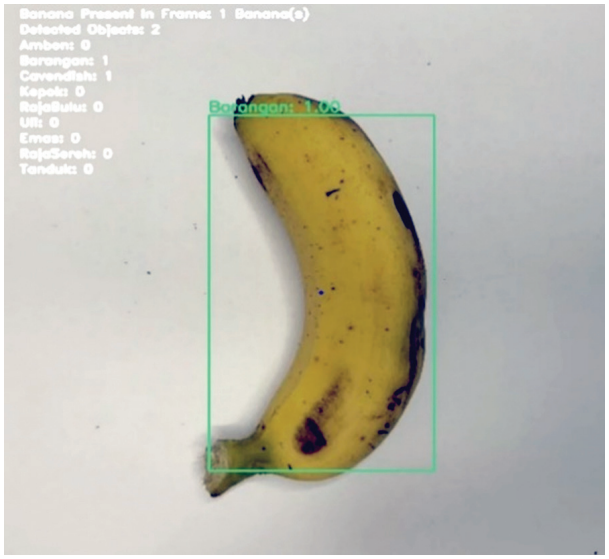


Fig. 2. Examples of Banana image object detection in Banana datasets

In Fig. 2 present object detection involves identifying different objects within an image and outlining them with a boundary to provide localization coordinates. Researchers in the field have taken an interest in object detection specifically in banana images, as aerial views from cameras mounted on drones offer stereo perspectives [22]. The purpose of this paper is to provide accurate banana image detection using deep learning-based object detection methods, particularly on Banana datasets. It aims to serve as a comprehensive resource for current advancements in deep learning-based Banana object detection in Banana datasets.

Our research focuses on identifying suitable convolutional network models for detecting banana fruit image objects. The banana dataset can be classified into nine different classes. The goal of this research is to determine the accuracy of object classification based on the characteristics depicted in the fruit dataset. We aim to determine the extent to which the average accuracy improves in recognizing detected objects. The main objectives of this research are as follows:

1. Evaluate the performance of different convolutional network models in detecting banana fruit image objects.
2. Determine the accuracy of object classification based on the characteristics depicted in the fruit dataset.
3. Assess the improvement in average accuracy in recognizing detected objects.

The detection system described the stages of banana object detection using deep learning using several CNN models. Where the dataset was collected using a digital camera positioned correctly between 30 - 40 cm from the banana object.

In conducting research on detecting banana objects using a camera positioned between 30 - 40 cm from the banana object, the data collection process for several banana classes was carried out. The objective of this process was to facilitate the detection of image objects using deep learning algorithms. Two main factors need to be considered in this detection process: the accuracy of object detection and the calculation of banana similarity across different classes. The detection results for two types of bananas using one of the CNN models can be seen in Fig. 2 and Fig. 3.

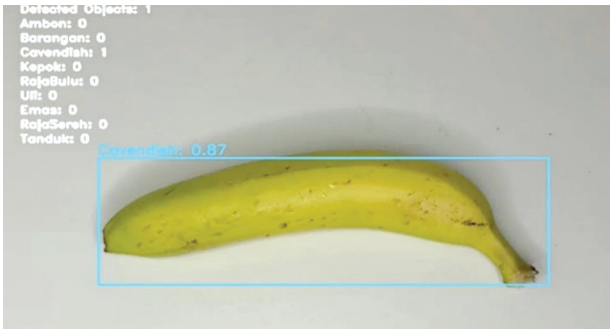


Fig. 3. The result of Banana Detection using MobilenetV2 - YoloV3-Nano

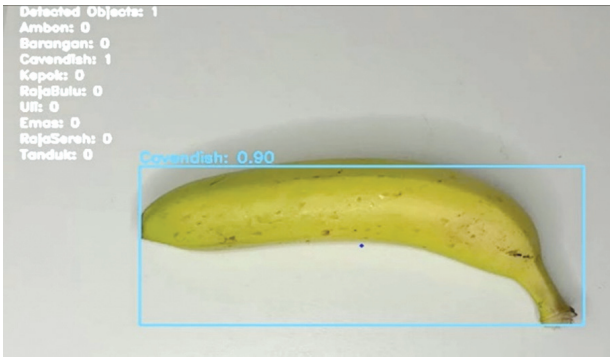


Fig. 4. The result of Banana Detection using Yolo V7

Fig. 3 and Fig. 4 illustrate the results of the accuracy in detecting Banana image class objects from processing the Banana Dataset. The class of Banana objects is determined using two different CNN models. The training process utilizes the Darknet algorithm, which characterizes the class of Banana objects detected within a bounding box. The concept of object class detection from Banana Datasets in video form is a novel approach that holds promise for further development in MAP.

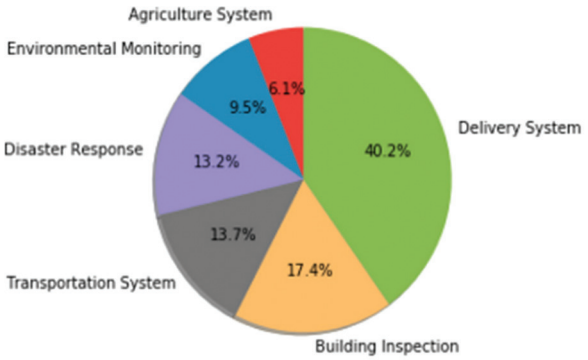


Fig. 5. Several examples of the use of object detection in various fields can be seen in Figure 4.0. [22]

Fig. 5 clearly informed the surge in research publications in recent years, driven by the advent of deep learning-based object detection. However, the accuracy achievable in detecting bananas or other fruits directly from images remains limited. While the object detection domain is vast, encompassing various object types and development stages, the focus remains on algorithms that perform well on images of objects captured directly from a predetermined distance [23]. The literature on object detection in images can be broadly categorized into two approaches: classical and modern.

3. PROPOSED METHODS

This research on detecting banana objects in several classes is divided into three main parts: the preparation stage, the training stage, and the testing stage. The preparation stage involves collecting or creating banana video datasets and analyzing banana class datasets. In the training stage, all collected banana data is processed, and the weight of each data is calculated to enable recognition in the testing process During the testing stage, the banana test data is recognized based on the training data that has been collected. For a more thorough explanation refer to the figure below:

A. Collecting Materials and Preparation

The initial process carried out in detecting the 9 classes of bananas that must be carried out is the preprocessing stage or also referred to as the preparation stage where the steps carried out can be seen in Fig. 6.

Fig. 6 provides information about the preparation stage. Description of the fields of image object detection can be broadly categorized into two approaches: classical and modern. Classical categorization includes conventional techniques consisting of vision-based approaches based on machine fruit classifiers. Modern categorization refers to deep learning-based algorithms, which is the focus in this research. Classical approaches to object detection include all major advances in the field of imaging using human-generated feature-based machine learning approaches [24]. The collected data will be used for training and testing. Therefore, video obtained from video dataset needs to be processed first to make it understandable for machines. Using FFMPEG, each banana video sample is processed to extract video screenshots at regular intervals. This process is done to prepare the images for the image labeling process. Image labeling is essential for the machine to understand and identify the objects it needs to detect. Fig. 6 shows an example of the drone video being converted into a new image, along with a text file for each image containing object data for the machine to learn from. The labeled data then undergoes a training process using pretrained convolutional weights to determine which model produces the highest accuracy.

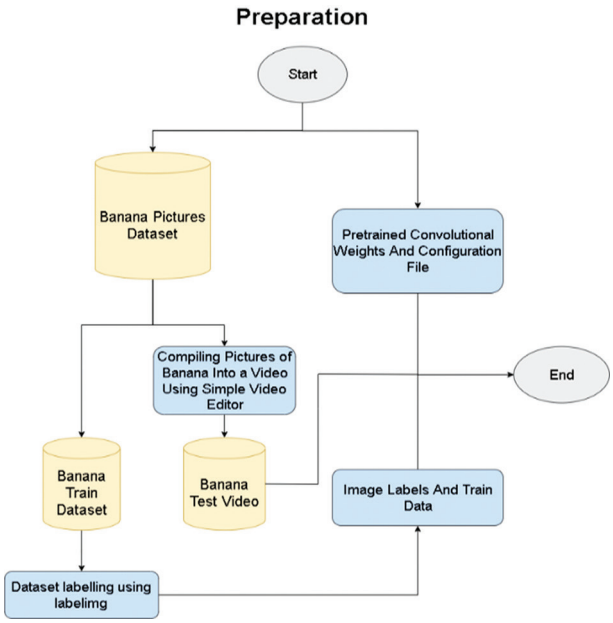


Fig. 6. The preparation stages in Banana fruit image object detection

After converting and labeling the images, another file called "trainer. Data" needs to be prepared. This file contains labeled images of banana fruits, the total number of classes, and the training output folder. Additionally, at this stage, pre-trained convolutional weights and the configuration of the CNN model are obtained from the internet. The configuration file must be adjusted based on the total number of classes. These configurations are known as training hyperparameters.

B. Training Dataset

The next stage is the training process which can be seen in Fig. 7.

Fig. 7 After completing the initial processing of the required banana image data, the next step is to begin the training process. During this training stage, an iterative process will be performed until either the iteration modulus is complete or the 1000th iteration module is completed. The objective of this stage is to train the banana image data, along with the supporting files, using the Darknet algorithm at each iteration. The outcome of the training stage will be weight files for each banana image. In general, the choice of weights is based

on the object space of the target image, which heavily relies on the properties of the objects in the training set and the predicted properties. To ensure the most accurate data, the training process will be conducted using the same training set and image data set. The resulting weight file will strike a balance between being neither too ambiguous nor too clear. This object detection process was executed to support experiments involving various classes of banana fruit images. The Darknet framework is leveraged solely to assist in the training process carried out by several pretrained CNN models.

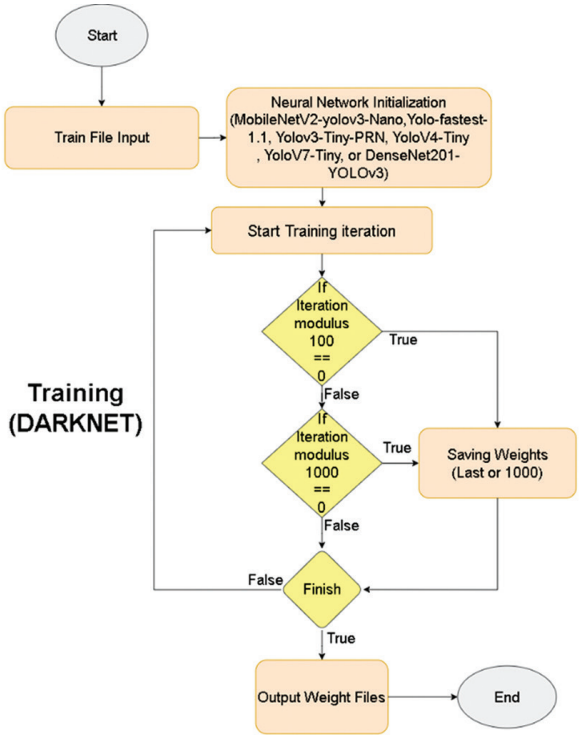


Fig. 7. The training stage in Banana fruit image object detection

C. Testing Process

After completing the training process, the testing process will be carried out on the banana fruit object data taken from the Banana Dataset. The goal is to detect all test banana data optimally based on the training data. In this experiment, 6 CNN models were used to determine which model provided optimal performance results for detecting banana fruit objects. The stages of the testing process in Fig. 8.

Fig. 8 explains and describes the process after the training stage is completed. The next step is to begin the testing stage. During this stage, the first step is to import the necessary files for testing, which include the trained weight files, the configuration files for training results, and the data files for the trainer. The testing process starts by extracting frames from the video dataset of banana fruit images. Then, the prediction is calculated using the non-max Suppression (NMS) function. The subsequent step involves drawing a bounding box around the banana object and determining its type/class.

This also includes calculating the prediction accuracy of the banana object class as the target object. The bounding box also includes other crucial information such as class and coordinates, which play a vital role in detecting banana class objects. The system also draws a line to indicate the direction of the object if it has moved within the frame. Finally, the banana image bounding box must respond to the accuracy percentage as per the specified binding.

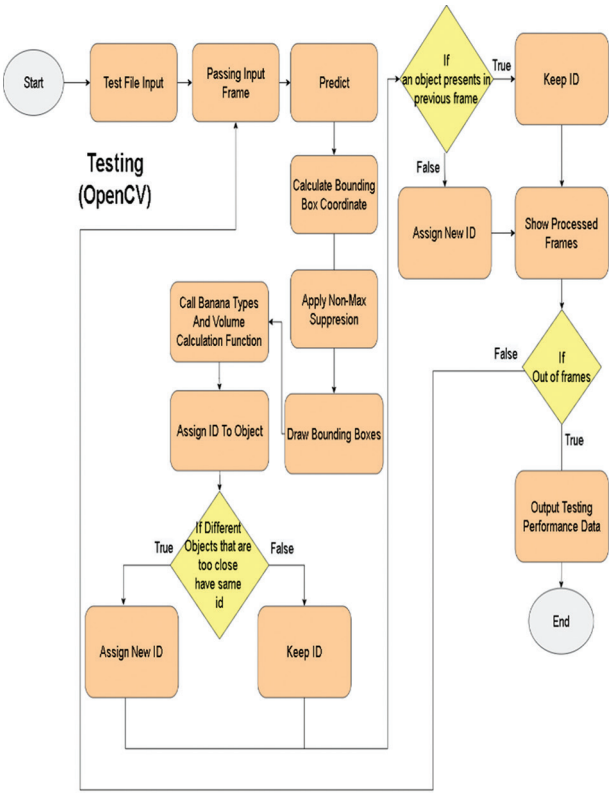


Fig. 8. The testing stage in Banana fruit image object detection

During the testing stage, the output produced is in the form of accuracy, precision, and F1 Score. These metrics are used to measure the performance of the Convolutional Neural Network (CNN) models. These metrics help determine which approach is effective for detecting objects in several classes of bananas based on the banana fruit dataset in video form. The testing process for validation data for the stages can be seen in Fig. 9.

Fig. 9 explains the validation stages using the darknet algorithm. The input data uses three parameters: the output weight, the configuration model used, and the training data. All files input with these three parameters will be processed with the validation function. The resulting accuracy of the validation banana fruit data will be predicted. The performance of each model is assessed using standard evaluation metrics, including precision (the ratio of accurately detected bananas to all detected objects), recall and overall accuracy. Additionally, computational efficiency and speed are compared to determine the trade-offs between accuracy and real-time performance.

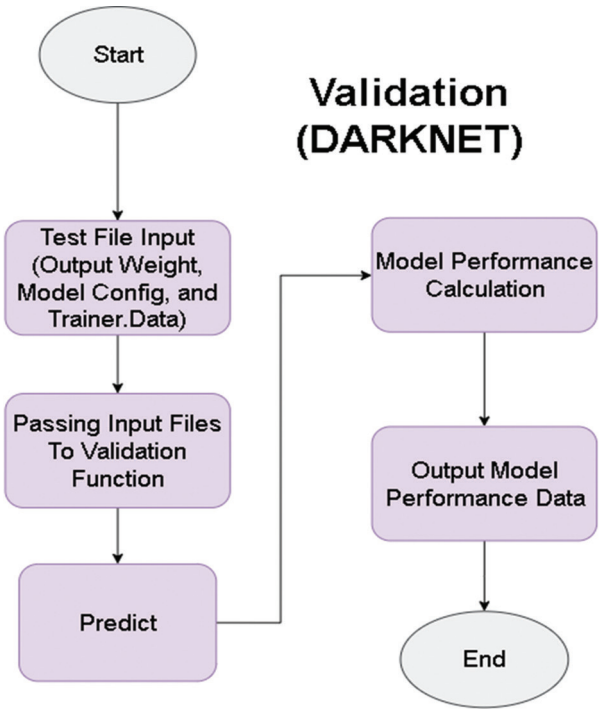


Fig. 9. The Validation stage in Banana image object detection

1) Precision and Recall

To determine the performance in detecting banana objects, the process of calculating the precision value must be supported by 2 important parameters, namely the first to get the right number of samples which are classified correctly by the model, while the second parameter must get the total samples which are classified correctly, regardless of whether the sample can be classified correctly or not. Precision calculations are between 0 and 1, with the reference being that 0 is the lowest value and 1 is the highest value. Conversely, recall is calculated by dividing the number of positive samples correctly classified by the model by the total number of positive samples. Both precision and recall formulas are illustrated below:

$$Precision = \frac{true\ positive}{true\ positi\ ve + false\ positive} \quad (1)$$

$$Recall = \frac{True\ Positive}{True\ Positive + False\ Negative} \quad (2)$$

Where:

1. True Positive: number of correct examples after a correct classification process.
2. False Positive: the number of negative samples from the wrong model but the correct samples.
3. False Negative: the number of positive samples but cannot be classified by the model.

2) F1-Score and mAP (mean Average Precision)

The F1 score is a measure of the harmonic mean of precision and recall. It is commonly used as an evaluation metric in binary and multi-class classification. The F1 score integrates precision and recall into a single metric to gain a better understanding of model performance [25]. The resulting F1-score value ranges from 0 to 1, where 1 represents the highest accuracy value.

$$F1\ Score = \frac{2\ (Precision\ x\ Recall)}{(Precision + Recall)} \quad (3)$$

In addition to the F1-score, which encapsulates the two-preceding metrics, the mean average precision (mAP) is a metric that represents the average value of average precision across the detection process for all previously identified classes. An object detection performance accuracy test was conducted on the Banana image dataset. The mean average precision (mAP) was employed as the performance metric [26], with the formula as follows:

$$\text{MAP} = \frac{\sum_{q=1}^{Q} AveP(q)}{Q} \quad (4)$$

In the formula, the variable Q is the total query in a data set, and the parameter AveP(q) describes the average precision (AP) value in a particular query process for parameter q. Next, for each query, q, the corresponding AP is calculated, then averaging all the scores from the AP parameters to produce a number called MAP, which is a measure of the effectiveness of the model's query processing capabilities. This performance metric can essentially be one of the metrics used to determine which model shows the best overall performance.

4. EXPERIMENTAL RESULTS

A. Training Results

In the training process on the banana dataset, PC technology is used which is equipped with the latest Central Processing Unit and Graphic Processing Unit specifications or as needed for banana image processing. An approach with the right PC technology will provide advantages in speeding up the duration of image training with CNN models supported by the Darknet framework in making GPU acceleration during the training phase process. The Specification of Hardware needed in the experiment in Table 1.

Table 1. Hardware Specification on the Experiment

No.	Description	Specification
1	CPU	Intel R7 RTX 4060 B760
2	RAM	32 GB DDR4
3	VGA	NVIDIA RTX 3 060
4	VRAM	6 GB

Precise and larger model architectures result in slower training times. The size of the CNN model also reaches the size of the model output.

All tested models produced good average loss results, which can be visualized in the form of a CNN model training graph, as shown in Figure below.

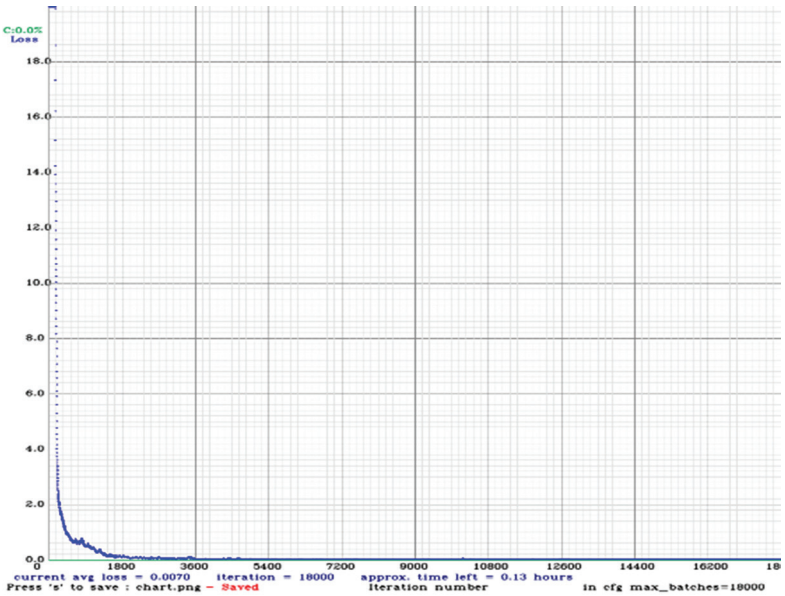


Fig. 10. Banana image Training using DenseNet201-YOLOv3

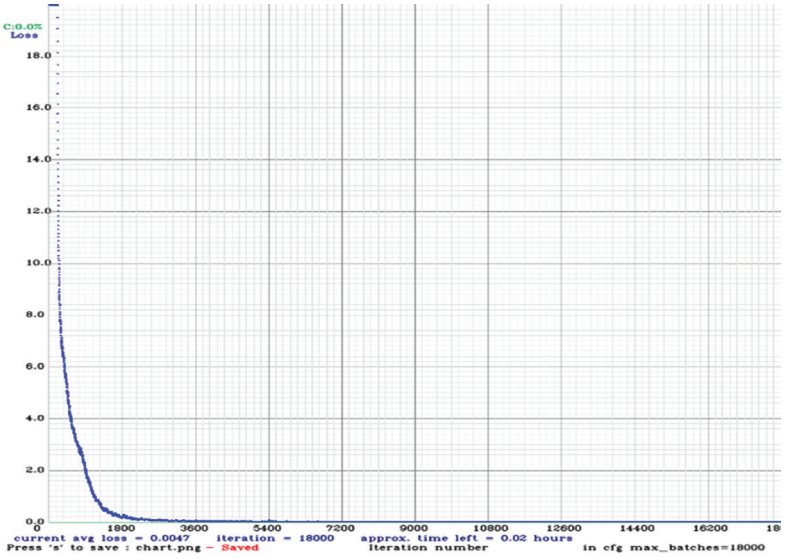


Fig. 11. Banana image Training using MobilenetV2-YoloV3-Nano

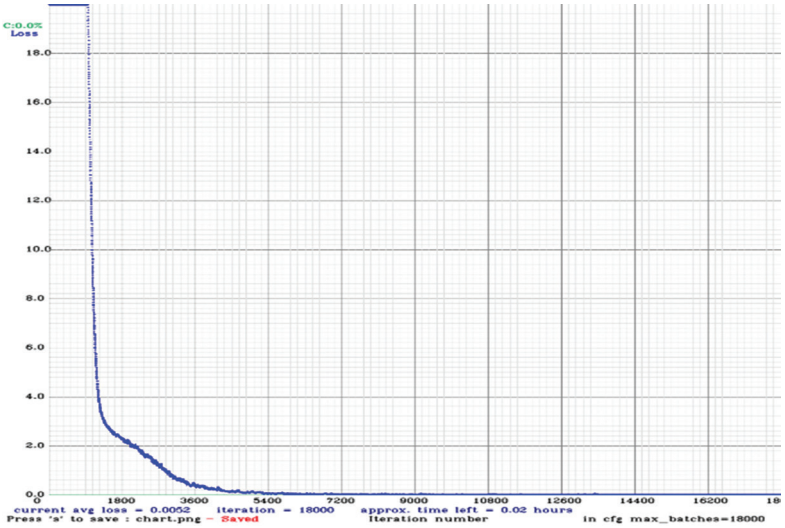


Fig. 12. Banana image Training using Yolo-Fastest 1.1

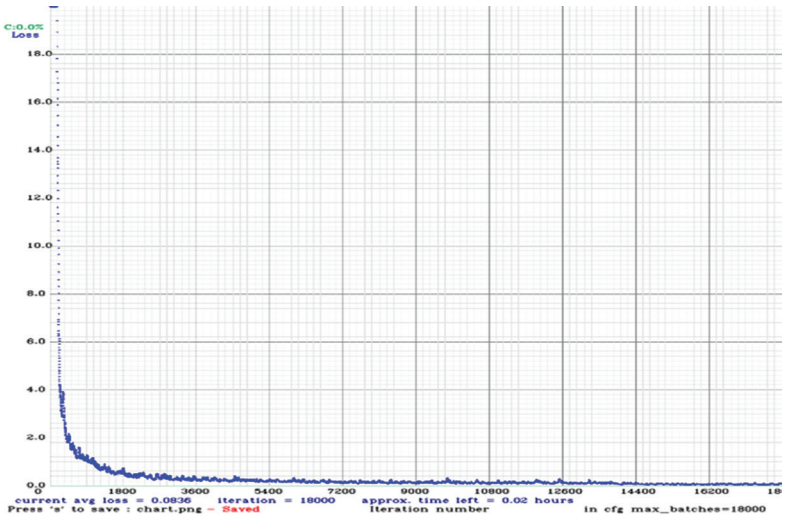


Fig. 13. Banana image Training using YOLOv3-tiny-prn

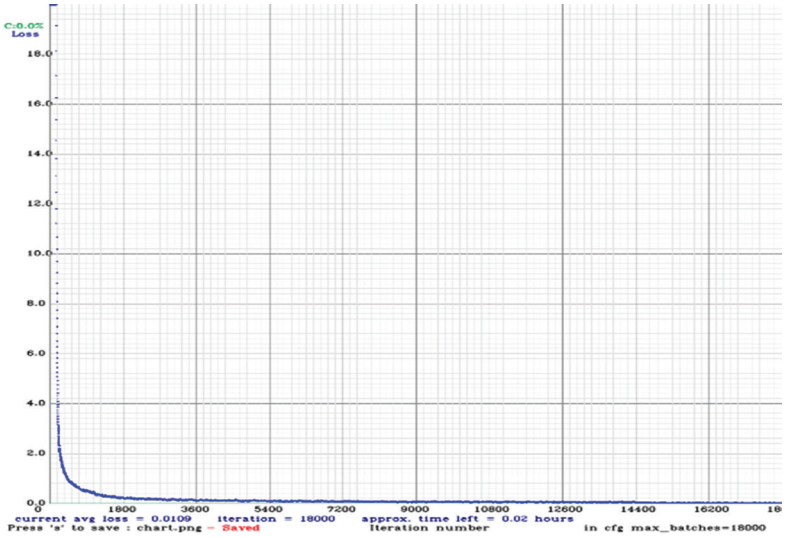


Fig. 14. Banana image Training using Yolov4-tiny

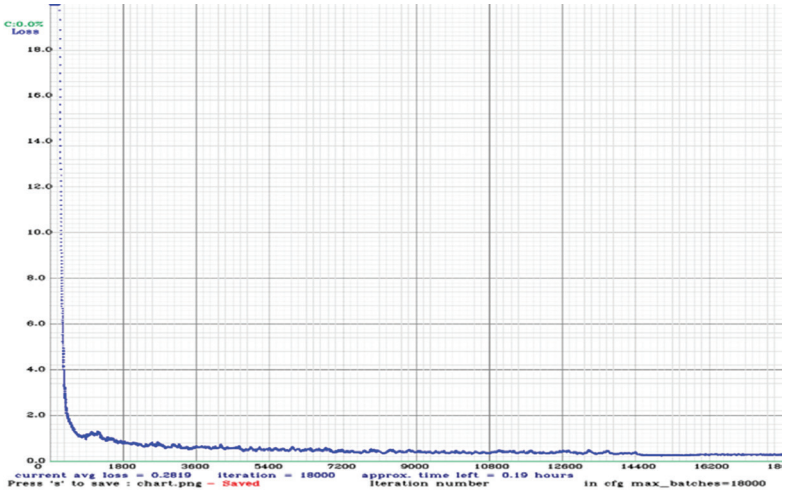


Fig. 15. Banana image Training using Yolov7

Figs. 10, 11, 12, 13, 14, 15 depicts the curve graph of the training stage results for the banana image, which can vary depending on the CNN model. In contrast, YoloV4-tiny and YOLOv7 exhibit faster degradation than the other models. However, the training process is not affected if the loss drops to a certain level. After a sharp initial decline, the losses stabilize into a gentle curve. This indicates that the model is beginning to under-

stand the dataset and is adapting to it gradually. The graph illustrates that the loss remains stable until the end of the iteration, which is 18,000 iterations for all models. This result implies that the trained model has successfully learned the specific banana image dataset without issues. In training result in Table 2.

Table 2. Training result of detection Banana Image using CNN model or Yolo Methods

Model	Average Loss(%)	Approx time
YOLOv3-tiny-prn	0.0836	0.02 hours
YOLO-fastest 1.1	0.0052	0.02 hours
YOLOv7	0.2819	0.19 hours
DenseNet201-YOLOv3	0.0070	0.13 hours
MobilenetV2-YoloV3-Nano	0.0047	0.02 hours
Yolov4-tiny	0.0109	0.02 hours

In Table 2, the average loss of the CNN model in determining performance in detecting bananas is provided. Table 2 shows that all models have an average loss below 0.1, which is considered quite low and acceptable for experiments in detecting fruit images. The MobilenetV2-YoloV3-Nano model has the lowest average loss of 0.0047, followed by YOLO-fastest 1.1 with 0.0052, and DenseNet201-YOLOv3 with 0.0070. When detecting banana objects, the CNN YOLOv3-tiny-prn model has the second highest average loss of 0.0836 with an estimated training time of 5.1 hours. The final model tested, YOLOv7, had the highest average loss of 0.2819 with an estimated training time of 5.2 hours. However, the difference in training time between the two models is only 0.1 hour. A solely based on the average loss value. An overview of the training configuration can be seen in Table 3.

Table 3. Configuration information parameter's

CNN Model Name	Batch	Subdivision	MaxBatch	Classes
Densent201-YoloV3	28	24	18000	9
MobilenetV2-YoloV3-nano	32	24	18000	9
Yolo-Fastest 1.1	32	24	18000	9
Yolo3-tiny-pm	32	24	18000	9
Yolov4-tiny	32	24	18000	9
Yolov7	32	24	18000	9

In Table 3 describes the configuration information from the processing results in detecting banana images. In general, this process describes the batch starting from 32 to a maximum of 18,000 to recognize 9 classes of banana images in each CNN model.

B. Simulation and Results

Before testing the weights trained on the system, the banana object needs to import the necessary supporting files. This file includes training labels for banana images, image positions, model configurations, and a .data file called trainer. Data. The trainer. data file is necessary to run the banana data testing process using the OpenCV library for inference. Once the supporting files are imported, the self-service of accuracy and banana image class detection will automatically be initiated. Then, an experiment of banana image object detection ensures the processed the trained object detection works properly. An overview detect banana fruit class objects in the banana dataset can be seen in Fig. 16.

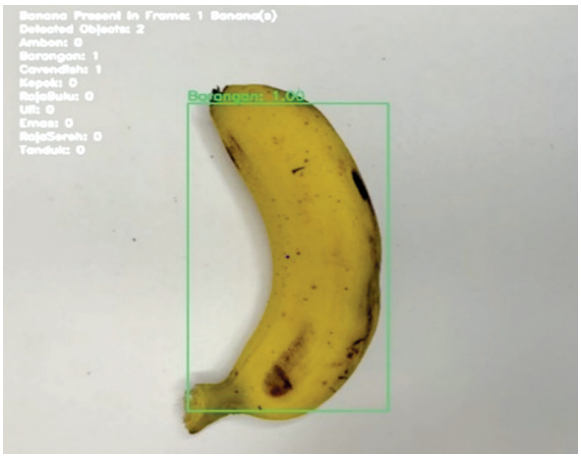


Fig.16. Banana fruit image class object detection simulation on Banana fruit dataset

Fig. 16 showcases the system's capability to accurately detect multiple classes of bananas in banana videos, while maintaining acceptable performance and excellent response times. However, achieving real-time banana object detection requires the implementation of appropriate methods and concepts to ensure seamless processing between frames and accurate interactions within the boundary box.

This approach reduces the processing load, which can potentially increase processing time. The results of the bounding box correctness calculation are shown in Fig. 17.

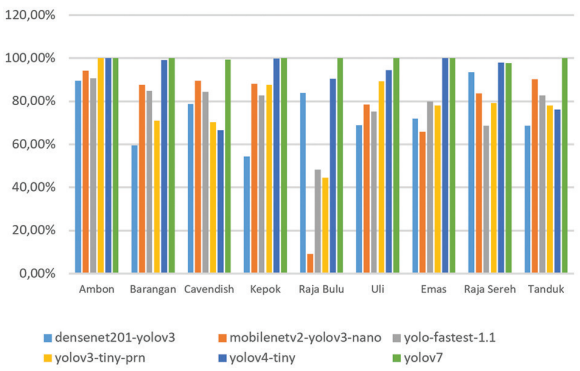


Fig.17. Banana image of average precesion after optimization CNN Model Comparison

Fig.17 shows the average precision results for banana image detection for 9 classes using 6 CNN models. Based on the experiments carried out, the Yolov 7 model was able to detect all banana classes very well

with an average precision of 98%, followed by Yolo v4 - tiny that delivers an average accuracy of 89%. Based on experiment this model only the banana Tanduk class has a detection precision of less than 78%. Another model whose precision is quite good is Pisang Rjbulu 43% only. the Yolo v3-tiny-pm whose precision reaches 70.5%, while there is 1 class of bananas whose precision is low namely Rjbulu.

The results of this experiment can be used as a guide for further research. The mean of Procession (mAP) and time processing for each Banana object is shown in Fig. 18.

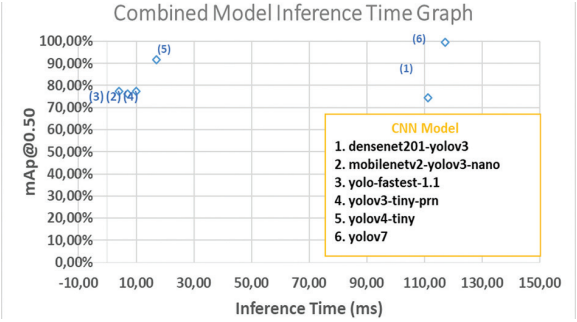


Fig. 18. Combined inference time model graph before and after optimization

Fig. 18 illustrates the inference time of 6 CNN models in relation to their mAP@0.50 percentage or their accuracy in detecting the banana class. Inference time measures the processing time between capturing an image and obtaining the detection results in the form of predictions. Based on this figure, the CNN model exhibits an additional 1% accuracy with approximately half the original inference time. In Banana object detection informed the average inference time significantly higher than the average accuracy of the images predicted by the system. For example, the YOLOv7 model achieves 100% accuracy, but its inference time is less than 50%.

Instead, the model's inference time is over 100%, reaching 110 ms. Other models, such as YOLOv4-Tiny, demonstrate results above 90%. The YOLOv4-Tiny model achieves an accuracy of 92.2%, while its inference time exceeds 50%, reaching 14.2 ms. The YOLO-fastest 1.1 model exhibits an average accuracy of 80.3%, but its inference time is also above 50%, reaching 8.2 ms. MobileNetV2-YOLOv3-Nano and YOLOv3-Tiny have inference times of 9-10 ms and average accuracies of 78% and 80%, respectively. DenseNet201-YOLOv3 possesses the highest inference time at 90.8 ms but also has the lowest mAP@0.50 percentage at 72.3%. Conversely, YOLO-fastest 1.1 exhibits a low inference time (8.2 ms) while maintaining good compatibility (80.3% mAP@0.50). The results are shown in Fig. 19.

Fig. 19 presents an experiment conducted using a Banana dataset to detect nine object classes. The experiment reveals that the CNN Yolo4-Tiny model achieves the highest accuracy, with precision, recall, and F1 Score reaching 84%. The YOLOv7 model follows closely behind with an average accuracy of 80%. On the other hand, the YOLOv3-Tiny-prn and Yolov4-Tiny models exhibit the lowest accuracy, with an average value of 53-54%. The minimum average accuracy exceeds 50%, indicating that this approach has performed well and can detect objects almost perfectly. To illustrate the combined average Precision for every CNN model can be in Fig. 20.

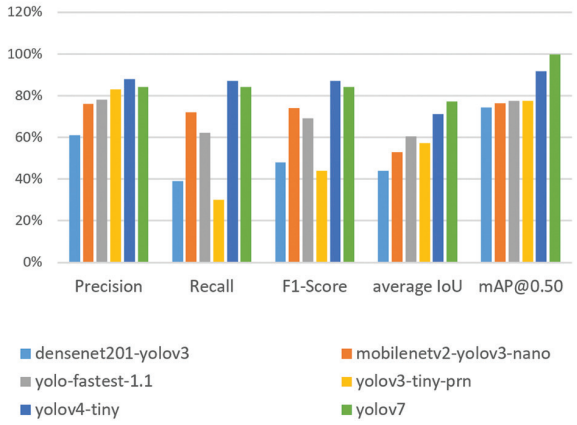


Fig. 19. The information of Precisions, recall, F1-score, average IoU and mAP@0.50 of each model CNN

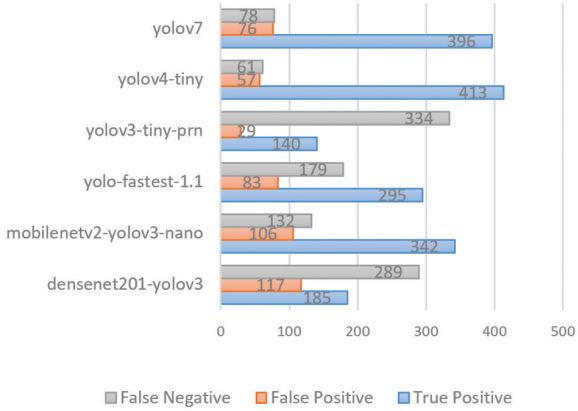


Fig. 20. Banana Image object detection performance using confusion matric

Fig. 20 presents the Yolov4-tiny can recognize more banana images than other models, specifically 413 images, which have true positive. However, there are 61 Banana images that are falsely identified as negative. The Yolov7 model recognizes 396 Banana images as true positive, but there are 78 images that are falsely identified as negative, which is more than the previous model. Another model that recognizes a significant number of banana images is MobileNetV2-Yolov3-Nano, with 342 images identified correctly. However, there are 132 images that cannot be recognized or are falsely identified as negative. On the other hand, the Yolo-fastest-1.1 model can recognize 295 banana images, but it fails to identify 179 images. this figure can effectively recognize banana images, with a total of 400 images recognized accurately.

In this figure, it can be concluded that only two models, yolov4-tiny and yolov7, effectively recognize banana images. In the next following tables are the ac-

curacy results for detecting each class of banana using each CNN model.

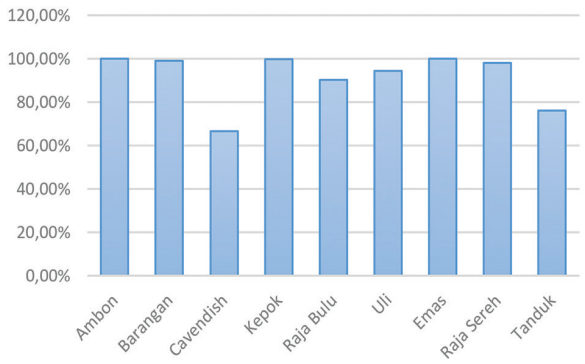


Fig. 21. Accuracy results in detecting banana images using the Yolov4- tiny model

Fig. 21 presents comparison of accuracy results for each Banana Class Detected Using the YOLOv4-Tiny Model Based on the experimental results obtained using the YOLOv4-tiny model, the highest accuracy was achieved for the Ambon, Barangan, Kepok, and Emas banana classes, with 100% accuracy. The Rjsereh banana class followed with an accuracy of 98%, while Uli and Rjbulu achieved accuracies of 96% and 94%, respectively. Additionally, the model demonstrates good performance in recognizing other banana classes.

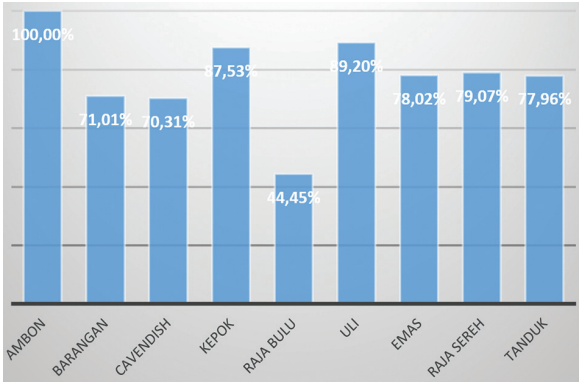


Fig. 22. Accuracy results in detecting banana images using the Yolov3- tiny-prn model

Fig. 22 shows a comparison of the accuracy results for each class of banana objects detected using the yolov3-tiny-prn model. Based on the results of experiments using this model, the highest accuracy results were found in the Ambon banana classes only, namely 100%. But the others have accuracy lest than 80% such as Uli, Emas, Rjsereh and Tanduk Banana Class. All the class banana has accuracy used this model less than 80% but more than 70%. However, based on the experiment this model has a good enough to detect and recognize the 9 class of Banana.

Fig. 23 presents comparison of the accuracy results for each class of banana objects detected using the YOLO-fastest-1.1model. Based on the results of experiments using this model, the highest accuracy results were found in the Ambon, Barangan, Cavendish Kepok, Emas, Tanduk classes is more than 80%. followed by the RjSereh banana class with an accuracy of less than 70% and Rjbulu less than 50%. . Based on this image, the accuracy produced by this model can detect several classes of bananas perfectly. However, for other banana classes, it is also good at recognizing the detected banana class.

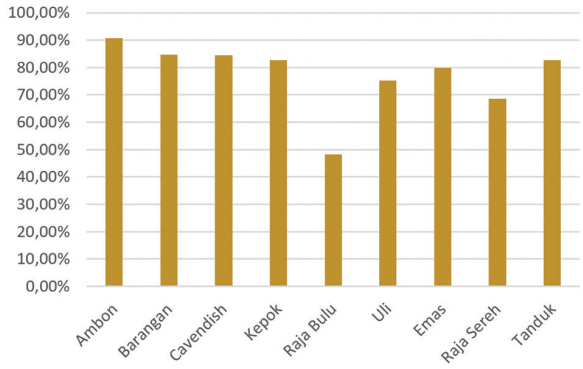


Fig. 23. Accuracy results in detecting banana images using the Yolov4- fastest – 1.1 model

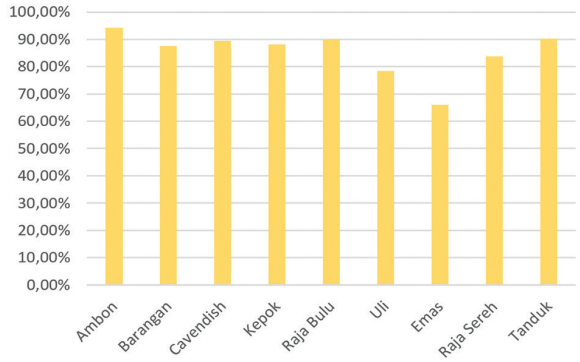


Fig. 24. Accuracy results in detecting banana images using the mobilenetV2-yolov3-nano

Fig. 24 presents comparison of the accuracy results for each class of banana objects detected using the yolov3-nano model. Based on the results of experiments using this model, the highest accuracy results were found in the Ambon, Cavendish, Rjbulu and Tanduk banana classes, namely 90% or more, followed by the Barangan, Kepok banana class with an accuracy of less than 90%. But Uli and Emas have accuracy at 78% and 63%. Based on this experiment, the accuracy produced lower than 90%. However, from other banana classes, it is also good at recognizing the detected banana class.

Fig. 25 presents a comparison of the accuracy results for each class of banana objects detected using the YOLO-Fastest-1.1 model. Based on the experimental results, the Ambon class achieved the highest accuracy of 90%, followed by the Barangan, Cavendish, Kepok, and Tanduk classes with accuracies above 80%. However, the Uli and Rjsereh classes had accuracies below 80%, at 73% and 68%, respectively. The Rjbulu class had the lowest accuracy of 48%. These results suggest

that the YOLO-Fastest-1.1 model can effectively detect several banana classes.

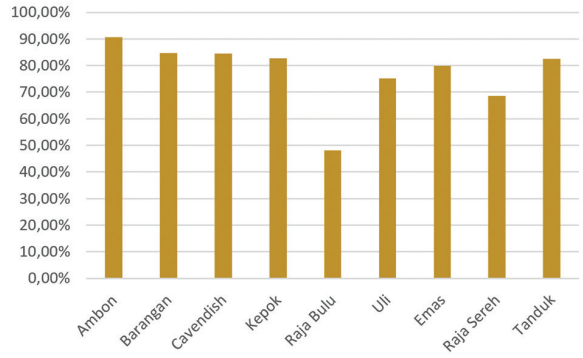


Fig. 25. Accuracy results in detecting banana images using the Yolo-fastest- 1.1

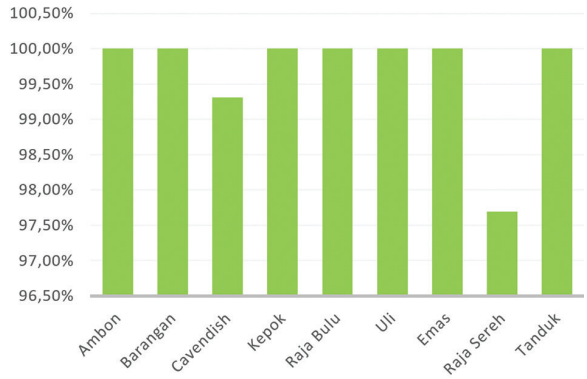


Fig. 26. Accuracy results in detecting banana images using the Yolov7

Fig. 26 presents a comparison of the accuracy results for each class of banana objects detected using the yolov4-tiny model. Based on the results of experiments using this model, the highest accuracy results were found in the Ambon, Barangan, Kepok, and Emas banana classes, namely 100%, followed by the Rjsereh banana class with an accuracy of 98%, Pisang Uli at 96%, and Raja Uli at 94%. Based on this figure can be seen that the accuracy produced by this model can detect several classes of bananas perfectly.

5. EVALUATION

In the experiment conducted to detect banana fruit classes, six CNN models were used. These models had been tested in object detection technology but had not been previously tested on the banana fruit dataset. While most of the CNN models used can detect the Ambon banana class with an average accuracy of over 90%, the average accuracy for Kepok, Barangan, and Emas bananas reaches over 80%. The experiments revealed significant variations in the number of detections for each banana class across the different CNN models. One of the challenges encountered is the issue of camera movement and the size of the objects being detected. These factors make it challenging for the system to accurately track the targeted objects. Despite these challenges, the number of objects detected can serve as an indicator of the model's performance.

6. CONCLUSION

1. Based on the experimental results on the banana dataset, there are 2 CNN models that have optimal accuracy. The YOLOv4-Tiny model produces an mAP@50 performance percentage of 92%, an average precision of 88%, while recall and F1 Score reach 86% and 84%. Meanwhile, the YOLOv7 model produces an mAP@50 performance percentage of 10%, while precision, recall and F1 Score produce the same performance of 84%. However, the two models have different Intersection over Union (IoU) accuracy, where YOLOv7 is greater at 77% compared to YOLOv4-Tiny at 71%.
2. Furthermore, the YOLOv4 Tiny model has the best potential when using the Banana fruit dataset because of its low inference time, only 13 ms, and a high mAP@0.50 accuracy percentage, namely 90.2%. The YOLOv7 model has an inference time of 96 ms and mAP@0.50 accuracy percentage reaching 100%.
3. Based on the low inference time of 12 MS and high accuracy reaching 90%, YOLOv4-Tiny has the potential to be the best option to be applied in the case of real-time banana object detection using video streams of various bananas on processor devices with power limitations. Another option is to use YOLOv7 because of its high accuracy capability reaching 100%, but the inference time is also very high, reaching 110 ms. However, due to its high inference time, it is only recommended when using high-power devices.
4. In further research recommended to using more complex and clearer datasets and better object tracking systems. These improvements have the potential to enhance the performance of real-time object detection systems.

7. REFERENCES

[1] S. Elayabalan, S. Subramaniam, V. Shobana, K. A. Kumar, "An overview on phytochemical Composition of Banana (Musa spp.)", Indian Journal of Natural Sciences, Vol. 7, No. 42, 2017, pp. 12408-12419.

[2] A. Koirala, K. B. Walsh, Z. Wang, C. McCarthy, "Deep learning–Method overview and review of use for fruit detection and yield estimation", Computers and Electronics in Agriculture, Vol. 162, 2019, pp. 219-234.

[3] S. Anatya, V. C. Mawardi, J. Hendryli, "Fruit maturity classification using convolutional neural networks method", IOP Conference Series: Materials Science and Engineering, Vol. 1007, No. 1, 2020, p. 012149.

[4] F. M. A. Mazen, A. A. Nashat, "Ripeness classification of bananas using an artificial neural network", Arabian Journal for Science and Engineering, Vol. 44, No. 8, 2019, pp. 6901- 6910.

[5] R. A. D. Putri, A. N. F. Rahman, A. Syarifuddin, "Proximate composition and sensory evaluation of Indonesia traditional food", AIP Conference Proceedings, Vol. 2596, No. 1, 2023.

[6] J. Wäldchen, M. Rzanny, M. Seeland, P. Mäder, "Automated plant species identification—Trends and future directions", PLoS Computational Biology, Vol. 14, No. 4, 2018, p. e1005993.

[7] K. Albarrak, Y. Gulzar, Y. Hamid, A. Mehmood, A. B. Soomro, "A deep learning-based model for date fruit classification", Sustainability, Vol. 14, No. 10, 2022, p. 6339.

[8] L. Zhou, C. Zhang, F. Liu, Z. Qiu, Y. He, "Application of deep learning in food: a review", Comprehensive Reviews in Food Science and Food Safety, Vol. 18, No. 6, 2019, pp. 1793-1811.

[9] A. Gupta, R. Kushwaha, K. Patidar, "A Survey: Fruit Classification Using Machine Learning", International Journal of Engineering Technology and Applied Science, Vol. 5, No. 7, 2019.

[10] S. Naik, B. Patel, "Machine vision-based fruit classification and grading a review", International Journal of Computer Applications, Vol. 170, No. 9, 2017, pp. 22-34.

[11] A. Bhargava, A. Bansal, "Fruits and vegetables quality evaluation using computer vision: A review", Journal of King Saud University-Computer and Information Sciences, Vol. 33, No. 3, 2021, pp. 243-257.

[12] B. Dhiman, Y. Kumar, M. Kumar, "Fruit quality evaluation using machine learning techniques: review, motivation and future perspectives", Multimedia Tools and Applications, Vol. 81, 2022, pp. 16255-16277.

[13] J. Ye, Z. Shen, P. Behrani, F. Ding, Q. Shi, "Detecting USM image sharpening by using CNN", Signal Processing: Image Communication, Vol. 68, 2018, pp. 258-264.

[14] A. H. Rangkuti, A. Harjoko, A. E. Putra, "Improvement of Accuracy in Batik Image Retrieval Due to Scale and Rotation Changes Using M2ECS-LBP Algorithm", Journal of Theoretical and Applied Information Technology, Vol. 97, No. 14, 2019, pp. 97-14.

[15] A. H. Rangkuti, V. H. Athala, N. F. Luthfi, S. V. Aditama, J. M. Kerta, "Reliable of traditional cloth pattern Retrieval Using Convolutional Neural Network", Proceedings of the 2nd International Conference on Artificial Intelligence and Data Sciences, 2021, pp. 1-6.

[16] Y. Hamid, S. Wani, A. B. Soomro, A. A. Alwan, Y. Gulzar, "Smart Seed Classification System based on MobileNetV2 Architecture", Proceedings of the 2nd International Conference on Computing and Information Technology, Tabuk, Saudi Arabia, 25-27 January 2022, pp. 217-222.

[17] H. Liu, R. Wang, S. Shan, X. Chen, "Deep supervised hashing for fast image retrieval", Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27-30 June 2016, pp. 2064-2072.

[18] D. C. Lepcha, B. Goyal, A. Dogra, S. H. Wang, J. S. Chohan, "Medical image enhancement strategy based on morphologically processing of residuals using a special kernel", Expert Systems, 2022, p. e13207.

[19] A. Bilal, G. Sun, S. Mazhar, "Finger-vein recognition using a novel enhancement method with convolutional neural network", Journal of the Chinese Institute of Engineers, Vol. 44, No. 5, 2020, pp. 407-417.

[20] N. Darapaneni, A. Tanndalam, M. Gupta, N. Taneja, P. Purushothaman, S. Eswar, T. Arichandrapandian, "Banana Sub-Family Classification and Quality Prediction using Computer Vision", arXiv:2204.02581, 2020.

[21] F. Xiao, H. Wang, Y. Xu, R. Zhang, "Fruit Detection and Recognition Based on Deep Learning for Automatic Harvesting: An Overview and Review", Agronomy, Vol. 13, No. 6, 2023, p. 1625.

[22] P. Mittal, R. Singh, A. Sharma, "Deep learning-based object detection in low-altitude UAV datasets: A survey", Image and Vision computing, Vol. 104, 2020, p. 104046.

[23] Y. Amit, P. Felzenszwalb, R. Girshick, "Object detection", Computer Vision: A Reference Guide, Springer International Publishing, 2021, pp. 875-883.

[24] Z. Li, W. Shi, P. Lu, L. Yan, Q. Wang, Z. Miao, "Landslide mapping from aerial photographs using change detection-based Markov random field", Remote Sensing of Environment, Vol. 187, 2016, p. 76-90.

[25] Esri, "How the Compute Accuracy For Object Detection tool works," ArcGIS Pro, https://pro.arcgis.com/en/pro-app/latest/tool-reference/image-analyst/how-compute-accuracy-for-object-detection-works.htm (accessed: 2023)

[26] A. F. Gad, "Accuracy, Precision, and Recall in Deep Learning", https://blog.paperspace.com/deep-learning-metrics-precision-recall-accuracy/ (accessed: 2023)

Gray Level Co-occurrence Matrix based Fully Convolutional Neural Network Model for Pneumonia Detection

Original Scientific Paper

Shubhra Prakash
Computer Science Department,
CHRIST (Deemed to be University)
Bengaluru, India
shubhra.prakash@res.christuniversity.in

B Ramamurthy
Computer Science Department,
CHRIST (Deemed to be University)
Bengaluru, India
ramamurthy.b@christuniversity

Abstract *– This study presents a new method to improve the detection ability of a convolutional neural network (CNN) in pneumonia detection using chest X-ray images. Using Gray-Level Co-occurrence Matrix (GLCM) analysis, additional channels are added to the original image data provided by Guangzhou Children's Hospital in Guangzhou, China. The main goal is to design a lightweight, fully convolution network and increase its available information using GLCM. Performance analysis is performed on the new CNN model and GLCM-enhanced CNN model, and results are compared with Transfer Learning approaches. Various evaluation metrics, including accuracy, precision, recall, F1 score, and AUC-ROC, are used to evaluate the improved analysis performance of CNN. The results showed a significant increase in the ability of the model to detect pneumonia, with an accuracy of 99.57%. In addition, the study evaluates the descriptive properties of the CNN model by analyzing its decision process using Grad-CAM.*

Keywords: *CNN, Pneumonia, Chest X-ray, Diagnostic, Explainability, GLCM*

Received: September 27, 2023; Received in revised form: February 20, 2024; Accepted: February 20, 2024

1. INTRODUCTION

Pneumonia is a leading cause of infant death globally. Hence, a precise diagnosis is essential for successful treatment. The current standard of care for diagnosing pneumonia consists of imaging techniques and physical examination, with chest X-rays and computed tomography (CT) scans being the most common imaging procedures. Radiologists are tasked with interpreting these images, using their training to spot signs of pneumonia and inform further decisions about patient care. However, radiologists' views of this procedure vary, making it subjective. Furthermore, there are disadvantages to relying solely on human interpretation, and the healthcare system faces difficulties due to the growing need for diagnostic imaging. More objective, scalable, and practical techniques for pneumonia detection are required because of the limitations associated with subjectivity, human factors, and technological limits [1].

Convolutional Neural Networks (CNNs) are a powerful method of deep learning techniques that can identify complex patterns in images [2-4]. A typical CNN has many layers of artificial neurons, frequently with different kinds of layers for various purposes. They are extensively used in the medical imaging diagnosis and classification of a wide range of diseases, from brain tumors to skin lesions and, more recently, respiratory illnesses, including pneumonia. Even with their efficacy, there are still many obstacles to overcome. One major obstacle is the large amount of computational power required to train reliable CNN models. Large-scale datasets or complex models may require complex and time-consuming training methods, which call for sophisticated systems that aren't always readily available.

Models for transfer learning, like VGGNet, AlexNet, MobileNet, and ResNet, have been widely used for various applications, including the classification of medical images. Although these models have shown to be successful, their original architectures were in-

tended for multiclass classification on datasets such as Imagenet, which has more than 100 classes. These models are considered excessive for applications such as pneumonia detection, where input images are frequently grayscale with a single channel, and there are just two classification classes. Their larger number of channels and intricate network architecture also lead to bulkier model sizes, which require significant processing power. As a result, there is an increasing need for low-complexity, lightweight models, especially for resource-constrained setups. This paper explores a novel strategy for creating a lightweight X-ray image classification model that uses point-wise convolution and max-pooling to achieve progressive dimension reduction and channel count reduction. This research also examines the combination of the Gray-Level Co-occurrence Matrix (GLCM) to improve the detection power of CNN models. By incorporating additional data channels from the GLCM into the original image.

The main idea of this work is to show that a small and robust CNN model may produce predictions that are comparable to or better than those produced by current techniques, even though it is much smaller in size. Additionally, adding texture feature information from images can further improve this performance.

The key contributions of this manuscript include:

- Creation of a lightweight fully convolutional CNN model for pneumonia classification
- Performing necessary preprocessing on the Xray images
- Extracting GLCM-based channels from the given images and appending them to original images for the classification task
- Experimentation with channel dimensions and counts while keeping crucial spatial information for precise predictions
- Comparison of the proposed model's classification performance to transfer learning models (trained on the same dataset), such as VGG and ResNet

The sections that follow are arranged as follows: The findings of the literature review are shown in Section 2. In Section 3, the dataset used and methods are described in detail. The results are shown and discussed in Section 4. The manuscript is concluded in Section 6, after Section 5 examines potential future developments.

2. LITERATURE REVIEW

Souid et al. [5] used a modified MobileNet V2 model to make predictions from radiographic images. They applied transfer learning and metadata integration, extracting data from the NIH Chest-Xray-14 database. The performance of their method is evaluated using the AUC statistic, giving an average AUC of 0.558 and achieving an accuracy of more than 90%.

In their study, Oh et al. [6] introduced a deep neural network model based on patch-based analysis. They trained the model on a small data set and based their decisions on broad observations from random lung patches. In addition, they introduced the Grad-CAM saliency map for detailed information.

Khan et al. [7] developed a CoroNet model using the Xception architecture to train it on a dataset containing X-ray images of COVID-19 and pneumonia. Their model achieved an accuracy of 89.6% and a precision and recall rate of 90% and 89.9%, respectively, for COVID-19.

Ozturk et al. [7] proposed a new model in their study to improve the diagnosis of COVID-19. This model achieved a good accuracy of 87% in two variable classifications.

Wijaya et al. [8] used K-Nearest Neighbor and Gray Level Co-Occurrence methods, achieving the highest accuracy of 66.20% for K=5.

Polsinelli et al. [9] designed a lightweight CNN based on SqueezeNet, which correctly detected COVID-19 in CT chest scans with 85.03% accuracy,

Joshi et al. [10] conducted a study where a CNN model called LiMS-Net was proposed to solve the problem of overfitting in training samples and the detection of COVID-19 by CT scans. With 2.53 million parameters, the model outperformed the transfer learning approaches and achieved 92.11% accuracy and 92.59% F1 score, demonstrating its effectiveness even with small CT data.

Several other studies [11-13] were also performed to identify pneumonia using machine learning and X-ray images of the chest, thus reinforcing the critical role of machine learning and deep learning in automatic detection.

3. MATERIALS AND METHODS

This study uses X-ray images showing different pneumonia patterns as the dataset to build and validate our model. We designed a simple CNN model to handle the complexity of pneumonia diagnosis, carefully considering the balance between model complexity and diagnosis accuracy within the limits of a real clinical setting. Our dataset is divided into training, validation, and testing sections to evaluate the model's performance. In addition to creating and evaluating the model, we use Grad-CAM visualization to see the areas in the image that contribute more to pneumonia classification.

The proposed framework and workflow are shown in Fig. 1. The first step in the process is loading the images, which are then resized to 224 by 224 pixels and normalized. These images are then processed using a feature extractor to extract different texture-based GLCM feature channels. After that, the extracted channels are added back to the original image. Now that this enhanced image has 1+9 channels, we train our

lightweight CNN model. A loss computation is done, and the network weights are modified following the resultant computed loss.

3.1. DATASET

The dataset used for this study consisted of chest X-ray images of two kinds: one representing pneumonia and the other showing normal conditions. These radiographic images were provided by Guangzhou Children's Hospital, China [14]. The dataset is publically available and can be accessed using this link: https://data.mendeley.com/datasets/rscbjbr9sj/3. A pneumonia case includes images that show different levels and degrees of the disease, while a normal case includes images without any signs of pneumonia. This diverse dataset serves as a reliable basis for developing and evaluating our simple CNN model for diagnosing pneumonia. The image has different dimensions in the study, although the majority is 255 x 255 pixels with a depth of 24 bits, and the images are in PNG format.

Table 1. Train and test split for dataset

Class label	Training	Testing
Normal	1340	3874
Pneumonia	224	241

3.2. PREPROCESSING

In the study, all images were processed by resizing them to a size of 224 x 224 pixels. The main goals of resizing were to enhance model performance and computational efficiency. First, resizing ensures consistency in the data provided to the network by helping to standardize input dimensions throughout the dataset. This uniformity is essential for the CNN to efficiently learn and generalize patterns. Additionally, resizing lessens the computational effort. This is particularly useful when training big models on hardware that is constrained. Apart from this, resizing also allows effective comparison of results.

Next, we normalize the pixel values in the image; Normalizing improves the training stability and convergence of the model. Normalizing guarantees that features in various channels are on a similar scale and helps reduce pixel intensity variances. By avoiding problems like vanishing or exploding gradients. Additionally, it strengthens the network's resistance to variations in lighting and enhances its capacity to recognize patterns in various images. The calculation of normalized pixel values is performed using the formula [15]:

$$normalized\_value(z) = (pixel\_value(x) - average\_value(\mu)) \,/\, standard\_deviation(\sigma) \quad (1)$$

3.2.1. Gray-Level Co-occurrence Matrix (GLCM)

Apart from resizing the images, we also apply the GLCM method to extract and create nine additional channels, including Mean, Standard Deviation, Contrast, Dissimilarity, Homogeneity, Angular Second Moment, Energy, Maximum, and Entropy values. These additional channels are appended back to the original image (refer to Fig. 2), and the image is then converted into a tensor object for model training.

Gray-Level-Co-occurrence Matrix is an image texture analysis technique commonly used in image processing and computer vision. It defines the spatial relationship between pixel intensity values in a digital image. A GLCM is constructed by counting the occurrences of two pixels with strong values at different distances and directions in the image.

3.3. DEEP NEURAL NETWORK CLASSIFIER

The proposed CNN model is suitable for pneumonia classification, distinguishing between normal and pneumonia conditions. Its purpose is to diagnose pneumonia in patients by analyzing chest x-ray images. The following sections will provide a detailed description of this CNN model and go into the details of its development and analysis of its results.

3.3.1. Components of Model

The proposed CNN model is designed to classify pneumonia, with the primary goal of distinguishing between normal and pneumonia-related conditions from chest X-ray images. This model uses layers with filters of different sizes (kernels), such as 3x3 or 5x5, to extract essential features from the input image. These filters combine, creating feature maps representing different parts of the original image [16]. Batch Normalization is used after every layer. It involves normalizing the intermediate feature maps within a batch of training samples to have zero mean and unit variance. By applying BatchNorm, the network becomes less sensitive to variations in the distribution of inputs, leading to improved training stability and faster convergence [17].

We can define the normalization formula of Batch Norm as:

$$Z^n = \left(\frac{Z - m_z}{s_z}\right) \quad (2)$$

where m_z and s_z represent the mean and standard deviation

Non-linearity is introduced by using an activation function, namely ELU, which is given by the following formula [18]:

$$f(x) = \{ x, if\ x > 0\ and\ -\alpha * (exp(-x) - 1),\ if\ x < 0 \} \quad (3)$$

In addition, the model uses a reduction method, namely Max pooling, to reduce space dimensions while maintaining essential features. It divides the feature map into non-overlapping areas and selects the maximum value in each area, preserving the essential features and removing the less important details. Global Average Pooling (GAP) is used at the end to reduce an entire channel to one value.

Finally, during training, the model uses Softmax activation and negative log loss as a function to transform the existing values into a probability distribution with a wrong prediction penalty, thus making predictions more accurate [19].

3.3.2. Architecture of the Proposed Model

The proposed CNN model comprises convolutional layers, pooling layers, and GAP layers. The PyTorch library is used to design the model in Python. This model follows the process of compression and expansion to add channels and consists of two types of blocks: convolutional and transition blocks.

In convolutional block, the input or feature maps are convolved over by a 3x3 filter with a stride of 1 and padding of 0, the number of channels in the filter progressively increases twice (e.g., 16,32,64), and then the transition block is applied to bring the channels back to the starting value (e.g., 16). In the study, we have trained the same CNN model twice, once with the original image of 3 channels and again with GLCM images of 3+9 channels. Thus, the number of input channels varies for both cases.

In a convolutional block, we have used a 3x3 kernel to create feature maps of channel sizes 16, 32, and 64; after every convolution, a batch normalization and ELU are added. Once the model reaches 64 channels, the dimension of the feature map is reduced by using max-pooling. Additionally, the size of the channels is also reduced using point-wise convolution.

This transition block can be understood with the help of Fig. 3. Suppose we have an input of Channels (C) x Width (W) x Height (H). On this input, we apply max-pooling, which reduces the dimension of the image by a factor of r (generally 2). Thus, the resulting output will become C x W/r x H/r. Note that the number of channels are still the same, just the dimensions of the image are reduced. Next, we use point-wise convolution to reduce the number of channels.

This is done by using a filter of C x 1 x 1. Now, this filter can be used multiple times depending on the channels we need as output. Thus, the output after every transition block will become C/p x W/r x H/r where r is the reduction factor for dimension and p is the reduction factor for a number of channels.

This transition prepares the model for another cycle of increasing channels using a convolutional block. This cycle is repeated four times until the image size is down to 11x11 with 16 channels. After this, we convolve using 3x3 filters twice with channel size increasing to 32 and 64, and then Global Average Pooling is applied, which gives an output of 64x1x1. We apply point-wise convolution to reduce this to 2x1x1, our final output before softmax.

The step-by-step operation is shown in the model summary table below.

Table 2. Model Summary

Layer	Kernel Size	Input Shape	Output Shape
Input	-	-	[3, 224, 224]
Conv2d	3x3	[3, 224, 224]	[16, 222, 222]
Conv2d	3x3	[16, 222, 222]	[32, 220, 220]
Conv2d	3x3	[32, 220, 220]	[64, 218, 218]
MaxPool2d	2x2	[64, 218, 218]	[64, 109, 109]
Conv2d	1x1	[64, 109, 109]	[16, 109, 109]
Conv2d	3x3	[16, 109, 109]	[32, 107, 107]
Conv2d	3x3	[32, 107, 107]	[64, 105, 105]
MaxPool2d	2x2	[64, 105, 105]	[64, 52, 52]
Conv2d	1x1	[64, 52, 52]	[16, 52, 52]
Conv2d	3x3	[16, 52, 52]	[32, 50, 50]
Conv2d	3x3	[32, 50, 50]	[64, 48, 48]
MaxPool2d	2x2	[64, 48, 48]	[64, 24, 24]
Conv2d	1x1	[64, 24, 24]	[16, 24, 24]
Conv2d (with Padding)	3x3	[16, 24, 24]	[32, 24, 24]
Conv2d	3x3	[32, 24, 24]	[64, 22, 22]
MaxPool2d	2x2	[64, 22, 22]	[64, 11, 11]
Conv2d	1x1	[64, 11, 11]	[16, 11, 11]
Conv2d	3x3	[16, 11, 11]	[32, 7, 7]
Conv2d	3x3	[32, 7, 7]	[64, 5, 5]
AvgPool2d	5x5	[64, 5, 5]	[64, 1, 1]
Conv2d	1x1	[64, 1, 1]	[2, 1, 1]

3.4. MODEL TRAINING AND TESTING

For training and testing, the dataset is divided into two parts: training and testing data. During the training phase, the CNN model is trained using the training data. Stochastic Gradient Descent has been used as an optimizer which updates the model parameters to minimize the loss function. The learning rate is set to 0.01, allowing the optimizer to control the step size during parameter updates. In addition, a momentum value of 0.9 is specified, which helps speed up the convergence during training. The epochs and batch size are set to 25 and 32, respectively. The model is trained by feeding batches through the network, calculating losses, and fitting gradients to improve model parameters. This process continues until all the epochs are over. Finally, the trained model is evaluated using test data.

Table 3. Training Hyperparameters

Hyperparameter	Value
Optimizer	SGD
Learning Rate	0.01
Momentum	0.9
Epochs	20
Batch Size	32

3.5. EXPERIMENTAL SETUP

In this study we have performed all the experiments on a machine with an NVIDIA GeForce GTX 1050 Ti GPU, 24 GB of RAM, and an Intel Core i5 8th Gen processor.

All the scripts were written and executed using Python 3.9.16 and the major libraries include PyTorch 2.0.0 with CUDA 11.8, and torchvision 0.15.1.

It is important to emphasize that a model's training time depends on the system specification hence for comparison, we have chosen to use the total number of trainable parameters in the model. Smaller parameter counts lead to faster training, especially when using transfer learning on a range of machines, as the number of parameters is directly proportional to the training time.

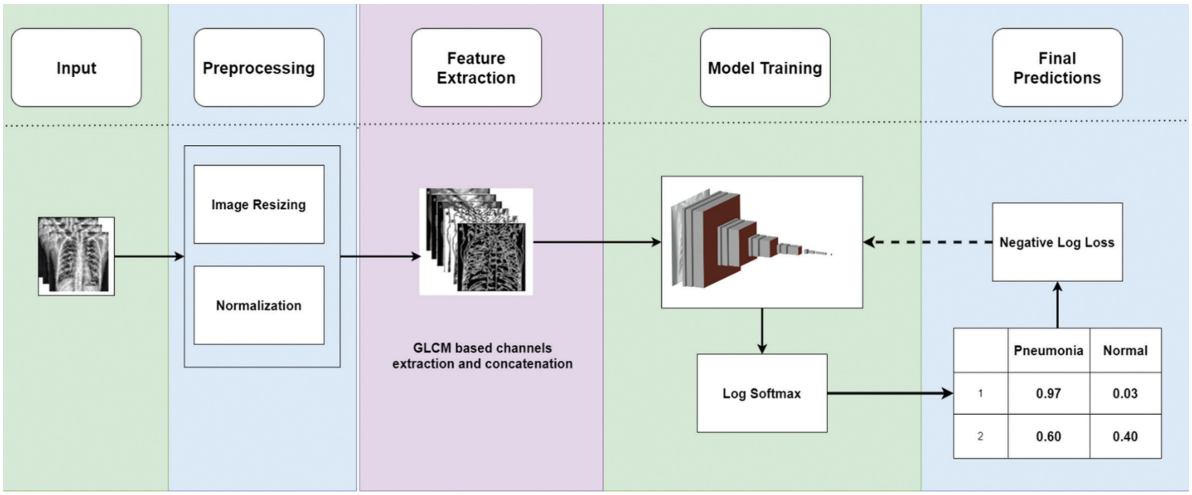


Fig. 1. Proposed framework and workflow

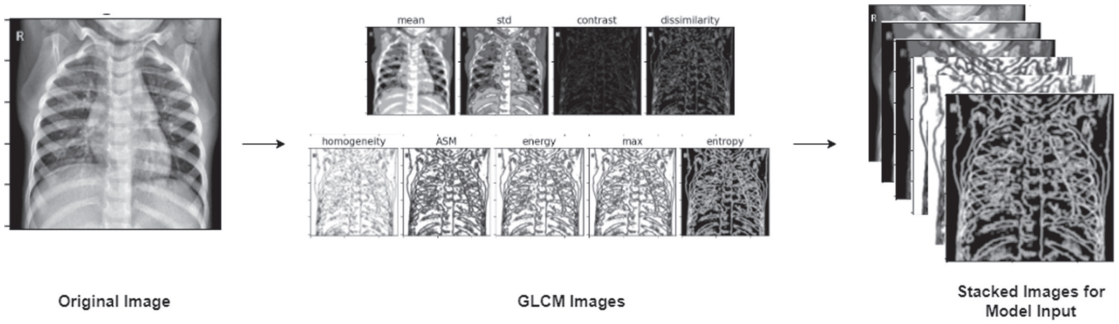


Fig. 2. GLCM-based channel concatenation

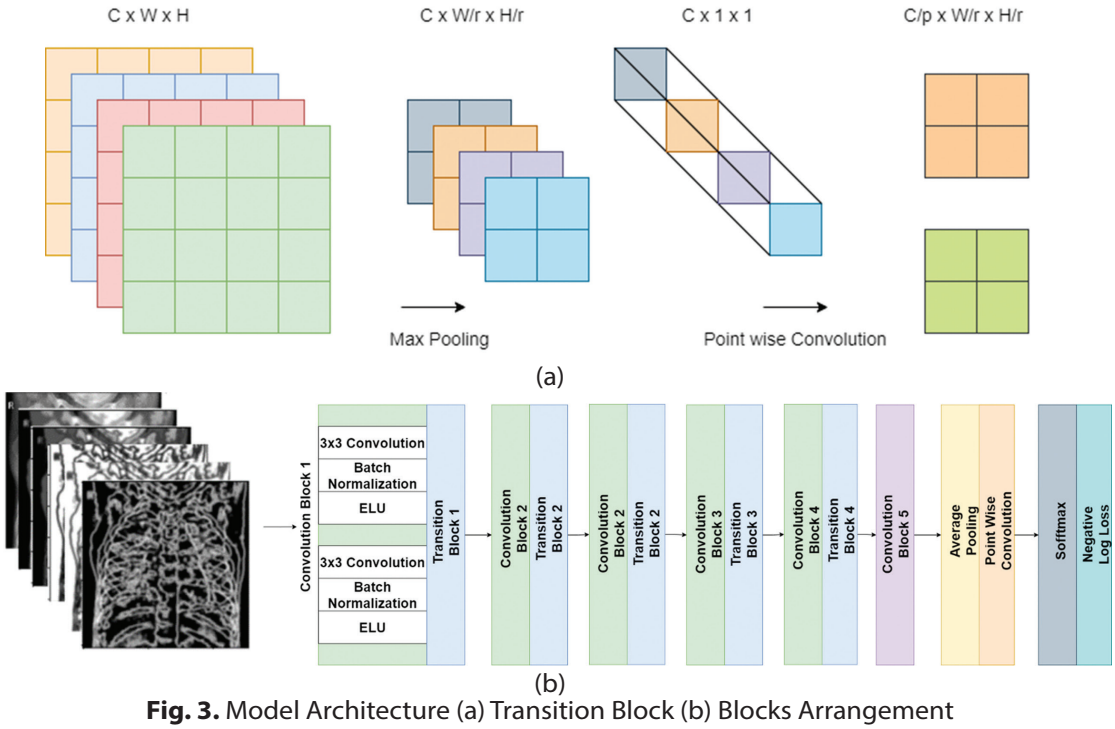


Fig. 3. Model Architecture (a) Transition Block (b) Blocks Arrangement

4. RESULTS AND DISCUSSIONS

Results from the proposed models are reported and discussed in this section and a comparative analysis is presented, highlighting the differences between our proposed models and existing models. Shown next in Table 3 and Table 4 we can see the training and test evaluation for both the CNN model.

Table 3. Training evaluation for CNN and GLCM CNN models

Model	Accuracy	Precision	Recall	F1
GLCM CNN	0.999	0.999	0.999	0.999
Base CNN	0.999	0.999	0.999	0.99

Table 4. Test evaluation for CNN and GLCM CNN models

Model	Accuracy	Precision	Recall	F1
GLCM CNN	0.9784	0.9785	0.9784	0.978
Base CNN	0.974	0.974	0.974	0.974

Based on the tables presented above, the GLCM-based CNN model performs slightly better than the base CNN model. Throughout the training process for both models, we consistently monitored accuracy for each epoch to ensure no over fitting. The results for the same can be seen below in Fig. 4 and Fig. 5.

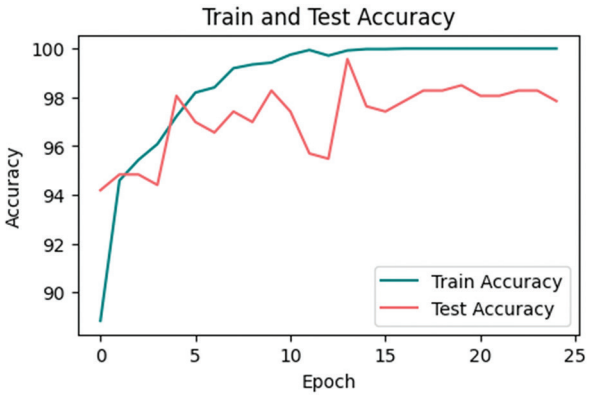


Fig. 4. Train and test accuracy for GLCM-based CNN model

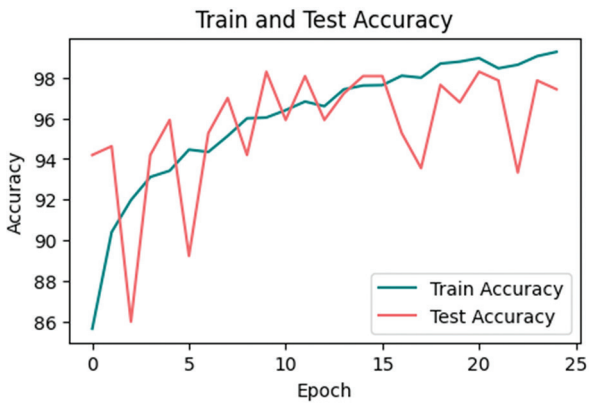


Fig. 5. Train and test accuracy for base CNN model

It can be seen in the above image that fluctuations in test accuracy is lot lesser in GLCM CNN model in comparison.

Below is the comparison of test results of both the models with transfer learning approaches. The size of the models is also shown (K denotes 1000, and M denotes 100000).

Table 5. Comparison with transfer learning approaches

Model	Accuracy	Precision	Recall	F1	Size
GLCM CNN	0.9784	0.9785	0.9784	0.978	129 K
Base CNN	0.974	0.974	0.974	0.974	129 K
MobileNet V2	0.96	0.96	0.96	0.96	22.26 M
Resnet34	0.95	0.95	0.95	0.95	21.28 M
AlexNet	0.97	0.97	0.97	0.97	57.01 M
VGG-16	0.97	0.97	0.97	0.97	134.26 M

It can be seen from the table above that the proposed models perform better than transfer learning approaches, although there is a huge difference in model size. Hence, it can be concluded that the proposed models are more compact and comparable in diagnostic capabilities.

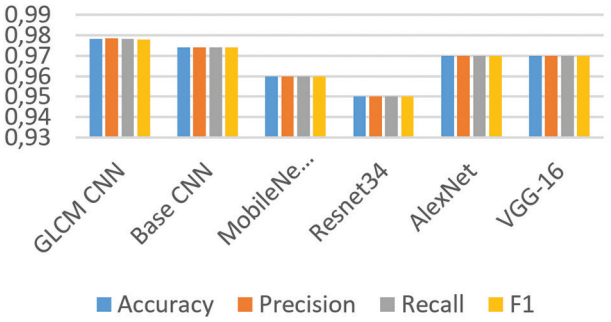


Fig. 6. Comparison with transfer learning approaches

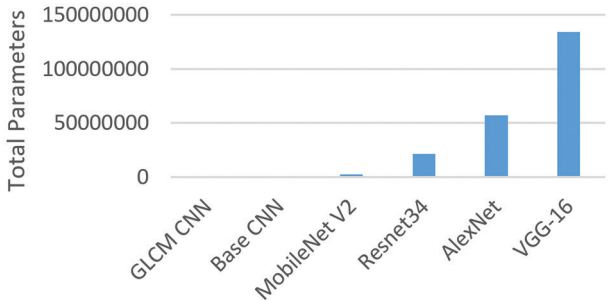


Fig. 7. Size comparison of different models

Fig. 8 shows a heat map generated using the Gradient-weighted Class Activation Mapping (GradCAM) technique, which is applied to predict pneumonia in a given input image. These heat maps serve as a critical medium, providing vital information about some regions of the input image that are critical to influencing the predictions of our Deep-learning model. By identifying these essential areas of the input image, Grad-

CAM allows decision-makers to better understand how the model approaches its predictions.

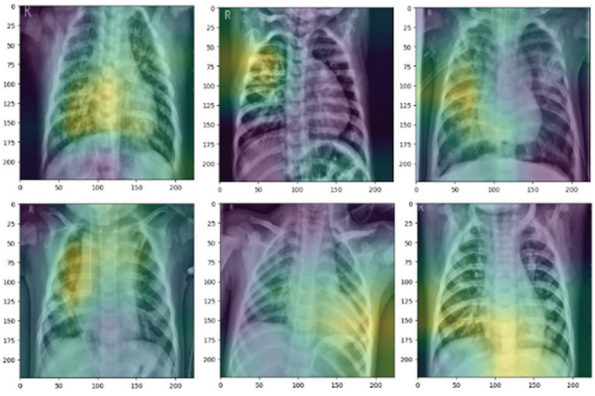


Fig. 8. GradCAM output

In Table 4, we assess the performance of our proposed model using the test dataset and compare it to previous studies. We have only shown the studies that have used the same dataset. It can be seen from the table below that the GLCM CNN model performs better in terms of all the evaluation metrics compared to work done by Moujahid et al. [20] and Zhang et al. [21]. Study conducted by Singh et al. [22] and Srivastav et al. [23] did not report all the metric however our model still shows better Accuracy and F1 Score.

Table 4. Model comparison

	Accuracy	Precision	Recall	F1Score
GLCM CNN Model	**0.9784**	**0.9785**	**0.9784**	**0.978**
Moujahid et al. [20]	0.9681	0.91	0.97	0.94
Zhang et al. [21]	0.9607	0.9441	0.9082	0.9258
Singh et al. [22]	0.9375			0.9405
Srivastav et al. [23]	0.945			

In summary, better model performance scores and less test accuracy fluctuations show that the GLCM-based CNN model performs marginally better in pneumonia identification than the base CNN model. Furthermore, the proposed models perform better than transfer learning strategies, demonstrating their effectiveness and diagnostic potential. Notably, even if the GLCM-based CNN produces better results, the base CNN might make sense when resources are limited because of its comparable performance.

5. LIMITATIONS AND FUTURE SCOPE

Although the current GLCM-based CNN model gives better accuracy than the base CNN model. The preprocessing and Model training time for the GLCM-based CNN model is huge. Further work is required to reduce time and make the process more efficient. It is suggested to adapt our base CNN model if the decision time required is less. Additionally, while the current approach only uses the GLCM method, further improvement of performance is possible by incorporating different texture-based features.

6. CONCLUSION

Our research has led to the creation of a novel model for pneumonia detection. We have developed a simple, Fully Convolutional Neural Network (CNN) algorithm that handles the difficult task of detecting pneumonia. The CNN model showed exceptional performance in pneumonia detection. The evaluation metrics demonstrated the model's ability to distinguish between pneumonia and non-pneumonia. The CNN model based on the Gray-Level Co-occurrence Matrix (GLCM) performed the best, achieving training and testing accuracy of 99.99% and 97%, respectively. Although the base CNN model achieved less accuracy, it significantly takes less time to train and predict. From a broader point of view, this study not only shows an advanced approach but also a bright future for AI-based healthcare. Although we focused on pneumonia, the models and principles applied in this research have great potential to address similar challenges in healthcare.

7. REFERENCES

[1] G. Labhane, R. Pansare, S. Maheshwari, R. Tiwari, A. Shukla, "Detection of Pediatric Pneumonia from Chest X-Ray Images using CNN and Transfer Learning", Proceedings of the 3rd International Conference on Emerging Technologies in Computer Engineering: Machine Learning and Internet of Things, Jaipur, India, 7-8 February 2020, pp. 85-92.

[2] D. Saida, K. K. Vardhan, P. Premchand, "Effective Brain Tumor Classification Using Deep Residual Network-Based Transfer Learning", International Journal of Electrical and Computer Engineering Systems, Vol. 14, No. 6, 2023, pp. 625-634.

[3] S. Tounsi, K. Boukari, A. Souahi, "The impact of collarette region-based convolutional neural network for iris recognition", International Journal of Electrical and Computer Engineering Systems, Vol. 13, No. 1, 2022, pp. 37-47.

[4] Y. Dahdouh, A. A. Boudhir, M. B. Ahmed, "A New Approach using Deep Learning and Reinforcement Learning in HealthCare: Skin Cancer Classification", International Journal of Electrical and Computer Engineering Systems, Vol. 14, No. 5, 2023, pp. 557-564.

[5] A. Souid, N. Sakli, H. Sakli, "Classification and Predictions of Lung Diseases from Chest X-rays Using MobileNet V2", Applied Sciences, Vol. 11, No. 6, 2021, p. 2751.

[6] Y. Oh, S. Park, J. C. Ye, "Deep Learning COVID-19 Features on CXR Using Limited Training Data Sets", IEEE Transactions on Medical Imaging, Vol. 39, No. 8, 2020, pp. 2688-2700.

[7] T. Ozturk, M. Talo, E. A. Yildrim, U. B. Baloglu, O. Yildrim, U. R. Acharya, "Automated detection of COVID-19 cases using deep neural networks with X-ray images", Computers in Biology and Medicine, Vol. 121, 2020, p. 103792.

[8] C. Wijaya, H. Irsyad, W. Widhiarso, "KLASIFIKASI PNEUMONIA MENGGUNAKAN METODE K-NEAREST NEIGHBOR DENGAN EKSTRAKSI GLCM", Jurnal Algoritme, Vol. 1, No. 1, 2020, pp. 33-44.

[9] M. Polsinelli, L. Cinque, G. Placidi, "A light CNN for detecting COVID-19 from CT scans of the chest", Pattern Recognition Letters, Vol. 140, 2020, pp. 95-100.

[10] A. M. Joshi, D. R. Nayak, D. Das, Y. Zhang, "LiMS-Net: A Lightweight Multi-Scale CNN for COVID-19 Detection from Chest CT Scans", ACM Transactions on Management Information Systems, Vol. 14, No. 1, 2023, pp. 1-17.

[11] S. D. Thepade, H. Jha, "COVID-19 Identification using Machine Learning Classifiers with GLCM Features of Chest X-ray Images", Trends in Sciences, Vol. 18, No. 23, 2021, p. 46.

[12] Y. J. Kim, "Machine Learning Model Based on Radiomic Features for Differentiation between COVID-19 and Pneumonia on Chest X-ray", Sensors, Vol. 22, No. 17, 2022, p. 6709.

[13] N. Dey, Y.-D. Zhang, V. Rajinikanth, R. Pugalenthi, N. Sri Madhava Raja, "Customized VGG19 Architecture for Pneumonia Detection in Chest X-Rays", Pattern Recognition Letters, Vol. 143, 2021, pp. 67-74.

[14] D. Kermany, K. Zhang, M. Goldbaum, "Large Dataset of Labeled Optical Coherence Tomography (OCT) and Chest X-Ray Images", Cell, Vol. 172, No. 5, 2018, pp. 1122-1131.

[15] S.-C. Pei, C.-N. Lin, "Image normalization for pattern recognition", Image and Vision Computing, Vol. 13, No. 10, 1995, pp. 711-723.

[16] Y. Lecun, L. Battou, Y. Bengio, P. Haffner, "Gradient-based learning applied to document recognition", Proceedings of the IEEE, Vol. 86, No. 11, 1998, pp. 2278-2324.

[17] S. Ioffe, C. Szegedy, "Batch normalization: accelerating deep network training by reducing internal covariate shift", Proceedings of the 32nd International Conference on International Conference on Machine Learning, Vol. 37, Lille, France, July 2015. pp. 448-456.

[18] D.-A. Clevert, T. Unterthiner, S. Hochreiter, "Fast and Accurate Deep Network Learning by Exponential Linear Units (ELUs)", Proceedings of the 4th International Conference on Learning Representations, San Juan, Puerto Rico, 2-4 May 2016.

[19] A. Krizhevsky, I. Sutskever, G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks", Advances in Neural Information Processing Systems, Vol. 25, 2012, pp. 1-9.

[20] H. Moujahid, B. Cherradi, O. E. Gannour, L. Bahatti, O. Terrada, S. Hamida, "Convolutional Neural Network Based Classification of Patients with Pneumonia using X-ray Lung Images", Advances in Science, Technology and Engineering Systems Journal, Vol. 5, 2020, pp. 167-175.

[21] D. Zhang, F. Ren, Y. Li, L. Na, Y. Ma, "Pneumonia Detection from Chest X-ray Images Based on Convolutional Neural Network", Electronics, Vol. 10, No. 13, 2021, p. 1512.

[22] S. Singh, B. K. Tripathi, "Pneumonia classification using quaternion deep learning", Multimedia Tools and Applications, Vol. 81, No. 2, 2022, pp. 1743-1764.

[23] D. Srivastav, A. Bajpai, P. Srivastava, "Improved Classification for Pneumonia Detection using Transfer Learning with GAN based Synthetic Image Augmentation", Proceedings of the 11th International Conference on Cloud Computing, Data Science & Engineering, Noida, India, 28-29 January 2021, pp. 433-437.

Data-driven Gait based Severity Classification for Parkinson's Disease using Duo Spatiotemporal Convoluted Kernel Boosted ResNet model

Original Scientific Paper

Arogia Victor Paul M
Research Scholar, Department of Computer Science and Engineering,
B.S. Abdur Rahman Crescent Institute of Science and Technology,
Chennai, India
victorpaul_cse@crescent.education

Sharmila Sankar
Professor & Dean, Department of Computer Science and Engineering,
B.S. Abdur Rahman Crescent Institute of Science and Technology,
Chennai, India
sharmilasankar@crescent.education

Abstract *– Parkinson's disease (PD) is one of the reformed brain syndromes that results in unintended stiffness and difficulty with balance and dexterity. To detect PD in medical scenery, physicians commonly use experimental indicators like motorized and non-motor symptoms and the severity rating depends on the unified PD Rating Scale (UPDRS). However, these medical assessments highly rely on expertized clinicians and lead to inter-variability discrepancies. Nowadays, gait sensor data assists doctors in diagnosing PD and estimates the severity level of gait abnormalities in patients. However, the gait sensor data increases the dimensionality issues and is subjected to high non-linear complexity. Hence, this study suggests an innovative deep learning (DL) technique for accurate PD analysis using gait patterns. Initially, the gait sensor data is preprocessed by performing data cleaning, and decimal scaling normalization (DS-Norm) to enhance the data quality. The Hoehn and Yahr (H&Y) scale is a commonly used rating scale for measuring the progression of Parkinson's disease symptoms. It's typically used to assess motor symptoms like tremors, rigidity, and bradykinesia. The scale ranges from 0 to 5, with higher numbers indicating more severe symptoms and disability. The preprocessed data are then fed into the proposed Duo spatiotemporal convoluted kernel boosted ResNet (DSCK-RNet) model for classifying the PD severity rating by learning the gait spatiotemporal features. The developed method is processed and scrutinized via the Python platform and a publicly available Physio-Net dataset is utilized for the simulation process. Various assessment measures like accuracy, precision, sensitivity, specificity, PPV, FPR, and MCC are examined and compared with traditional studies. In the experimental section, the developed DSCK-RNet model achieved an accuracy of 100%, 99.6%, 99%, and 99.64% for different classes like healthy, severity-2, severity-2.5, and severity-3 respectively. Compared to the conventional techniques, our suggested approach performs better. The experimental findings demonstrate the clinical significance of the suggested approach for the impartial evaluation of gait motor impairment in PD patients.*

Keywords: *Parkinson's disease classification, severity estimation, Hoehn and Yahr scale, gait patterns, decimal scaling normalization, spatiotemporal features, Duo Spatiotemporal Convoluted Kernel Boosted ResNet model*

Received: December 22, 2023; Received in revised form: February 10, 2024; Accepted: February 25, 2024

1. INTRODUCTION

PD (Parkinson's disease) is a distressing condition that causes serious illness to both non-motorized and motorized functions of the human body [1]. The insufficient nerve cells in the brain result in the production of a chemical named dopamine, which is the main reason behind PD illness [2]. Dopamine assists in sending messages for controlling coordination and movement. However, the shortage of dopamine leads to slowness in movement, uncontrolled shaking, problems with balance, etc. These kinds of problems are termed motor symptoms that can be easily identified [3]. The PD can also produce some non-motor symptoms that can be determined using gait patterns effectively. The integration of force-sensitive devices and machine learning algorithms for gait analysis holds promise for improving the early diagnosis and management of neurodegenerative diseases [4]. The most common traditional

techniques like high-speed cameras are utilized for analyzing the gait patterns that measure the pressure sensors and motion trails to obtain leg movements and muscle activities.

However, neuroimaging techniques are cost-effective in capturing motion illuminations and other force estimations [5]. Hence, the gait analysis via placing sensors on the human foot becomes an integral role in analyzing leg movements which is a highly accurate and affordable technique [6]. Also, the spatiotemporal (ST) features provide the quantitative measure of stroke, pace length, and strike time that can manipulate the motor deficiencies and predict the PD severity level with minimal complexities [7]. Therefore, determining the gait patterns not only identifies PD but also analyses the severity level of PD effectively. Despite this, the gait analysis based on sensor data faces numerous challenges such as increased data dimensionality and non-linear correlations among the spatiotemporal features [8]. To overcome this issue, highly effective data processing methods are required to assist physicians in detecting PD accurately.

Currently, artificial intelligence (AI) established machine learning (ML) and deep learning (DL) techniques have played an important role in the medical field, especially in the disease identification process [4, 7]. Several existing studies pointed out that ML-based techniques like support vector machine (SVM), decision tree (DT), random forest (RF), multilayer perceptron (MLP), artificial neural network (ANN) and naïve Bayes (NB) assist in detecting the abnormal PD level based on the changes in gait patterns from the motor symptoms [9]. The ML techniques can overcome complex nonlinear features by considering the essential gait patterns [10]. However, the ML models have low learning capability while processing the data with a larger population. Moreover, it fails to detect invisible gait features (non-motor symptoms) due to increased processing time and high amenability issues [11]. To overwhelm this issue, an effective and automatic severity rating classification technique is required for investigating invisible gait patterns. Nowadays, DL-based techniques aid doctors in assessing quantitative gait spatiotemporal features using an enhanced feature learning process [12]. Subsequently, it can train larger data for a huge population and can generate quality treatment with proven clinical outcomes. The DL models have various neurological-related applications such as diagnosis, severity assessments, and disease identification [13]. Henceforth, this study presents an innovative DL approach in PD decision making and severity assessment using gait ST analysis.

PD has been portrayed for the past few centuries as a reformed neurodegenerative illness that causes movement, cognitive abnormalities, and other non-motor symptoms. Even though PD cannot be cured, early detection and treatment with the right drugs and surgery can manage the symptoms. Even sleep issues, depression, tremors, uncontrollably shaking, and cognitive difficulties can be brought on by Parkinson's disease. Nevertheless, some of the concealed symptoms are imperceptible and necessitate precise methods to identify PD. Although there are other methods available, gait-based sensor data is becoming more and more common since it is an inexpensive and effective way to analyze the hidden symptoms of Parkinson's disease. Unfortunately, the gait data lead to low accuracy, complicated interpretation, and high dimensionality, particularly when identifying the severity level of Parkinson's disease. Numerous previous studies have described using gait spatiotemporal characteristics to determine the PD scores through machine learning. Nevertheless, the machine learning model requires more time to train on a larger population. Furthermore, it performed poorly in categorization since it was unable to precisely learn the non-motor symptoms. Therefore, to detect and categorize PD levels with minimal complexity, an efficient and highly accurate technique is needed. This served as inspiration for the created study, which presented a novel DL study that used gait patterns to analyze the PD. Below is a detailed illustration of the developed study's principal contributions:

- To introduce a novel DL model (DSCK-RNet) that uses sensor data based on gait to determine Parkinson's disease.
- To include crucial preprocessing steps to improve the quality of the gait data, such as data cleaning and decimal scaling normalization (DS-Norm).
- To present a novel Duo spatiotemporal convoluted kernel boosted ResNet (DSCK-RNet) model that learns the gait spatiotemporal features in order to classify Parkinson's disease.

The forthcoming sections are: Section 2 interprets the literature studies associated to PD classification using DL techniques. Section 3 describes the developed methodology. Section 4 determines the results and Section 5 represents the Concluded portion of the developed study.

2. RELATED WORKS

The medical industry has seen a significant surge in the use of AI and deep learning for disease diagnosis and prognosis due to their rapid expansion and research. Using a range of datasets, numerous research has been carried out to diagnose Parkinson's disease.

Balaji et al. [14] defined the DL method for perceiving the severity rating of PD accurately. Here, the LSTM model was presented to analyze the severity level of PD using gait data. For solving the overfitting issues, L2 regularization and dropout were utilized. To reduce the cost-utility, SGO and ADAM optimizers were utilized. Finally, the rating of PD was estimated via the H&Y and the unified PD Rating (UPDR) scales. The overall accuracy obtained by this method was about 98.6%. However,

this method faces high gradient sufficiency problems and high training time.

Aşuroğlu and Hasan [15] introduced the multiclass PD analysis using the hybridized DL technique. A hybridized CNN with a Locally Weighted Random Forest (LWRF) technique was introduced to classify the PD effectively. Here, essential gait features like frequency and time components were extracted. The overall Correlation Coefficient (CC), MAE, and RMSE obtained values of 0.89, 3.0, and 4.5 respectively. However, this method was a high cost for training with larger data.

Setiawan and Che-Wei [16] established the DL model for analyzing PD severity using time and frequency features. Three major stages were performed namely preprocessing, feature extraction, and classification. Here, the PD rating was analyzed based on VGRF signals. For the preprocessing, the signals were separated into different time-varying windows. Then, the continuous wavelet transform (CWT) was exploited to extract the phase-frequency features. In addition to this, principal component analysis (PCA) was introduced to enhance the extracted features accurately. Finally, 5 types of CNNs were utilized to detect and classify the PD severity levels. The overall accuracy achieved by this study was about 96.5%. However, this method increases the error due to a lack of noise filtering techniques.

Sai et al. [17] introduced the DL technique for predicting PD using gait spatiotemporal features. In this study, CNN-LSTM was introduced to forecast the PD rating on the basis of H&Y scale study. Various spatiotemporal features like swing and stance phases were extracted to identify the PD disease accurately. The overall accuracy, precision, and recall obtained the value of 88%, 86%, 94%, and 90% respectively. However, this method failed to maintain generalizability while appliedto wider datasets.

Vidya and Sasikumar [18] defined the PD severity analysis using a hybridized DL technique based on gait signals. Initially, the useful VGRF signals were obtained by the EMD technique that accurately extracts the fundamental functions. Then, power spectral analysis was utilized to select the essential intrinsic functions of the VGRF signals. Then, the CNN LSTM technique was introduced to analyze the severity of the PD accurately. Then, the overfitting issues were solved using L2 regularization and dropout mechanisms. The overall accuracy achieved by this method was about 98.3%. However, this method was highly time-consuming and subjected to increased error.

A residual network with 50 layers called ResNet50 was suggested by Omar El Ariss et al. [19] as a tool for Parkinson's disease diagnosis. The patient's speech recordings were subjected to spectral analysis techniques, which resulted in the gathering of frequency features that were used as data. Next, a 2-D heat map was created using the frequency features. ResNet50 receives this heat map and uses it to forecast whether or not the patient has Parkinson's disease.

A novel methodology for the precise identification of Parkinson's disease using handwritten records from a standard NewHandPD dataset was proposed by Sura Mahmood Abdullah et al. [20]. To lessen the strain of training time, the suggested framework is built on transfer learning models like ResNet, VGG19, and InceptionV3. To create an optimized feature vector for improved classification outcomes, the combined features from the TL models are fed into the genetic algorithm optimization process.

3. PROBLEM STATEMENT

From the deep scrutiny of the existing approaches, many issues have been reported such as increased dimensionality issues, high training time and error, etc. Several existing studies pointed out electroencephalogram (EEG) signals and vocal and handwritten features for detecting PD accurately. However, these features failed to analyze the severity level of PD due to low capability in processing with complex nonlinear features. Moreover, the gait signals are also utilized in various recent studies to analyze the PD conditions effectively. However, the gait signals are highly complex and increase the diagnosing time, and are highly cost-effective. To overcome this issue, gait sensor data is utilized in present studies in which the sensors are placed on the foot of a human leg to obtain useful spatiotemporal features. Various scaling methods like VGRF and H&Y scale are employed to analyze the rating of PD. However, the accuracy of using sensor data is not very effective due to high nonlinear correlation, poor data quality, complex interpretation, etc. Hence, this study presents a novel DL technique to learn the gait spatiotemporal features and to achieve better performance in classifying the PD levels accurately. To our knowledge, the proposed study overcomes all the issues faced in the existing studies and provides outstanding performance with minimal complexity.

4. PROPOSED METHODOLOGY

PD is one of the progressive brain disorders that results in inadvertent stiffness and difficulty with balance and agility. To detect PD in medical scenery, physicians commonly use irrefutable indicators like motorized and non-motorized symptoms and the rating based on the UPDRS. However, these medical assessments highly rely on expertized clinicians and lead to inter-variability discrepancies. Fig. 1 illustrates the workflow of the developed model.

Nowadays, gait sensor data assists doctors in diagnosing PD and estimates the severity level of gait abnormalities in patients. However, the gait sensor data increases the dimensionality issues and is subjected to high nonlinear complexity. Hence, this study suggests an innovative model for accurate PD analysis using gait patterns. Initially, the gait sensor data is preprocessed by perform-

ing data cleaning, and decimal scaling normalization (DS-Norm) to enhance the data quality. The preprocessed data are then fed into the proposed Duo spatiotemporal convoluted kernel boosted ResNet (DSCK-RNet) model for classifying the PD severity rating by learning the gait spatiotemporal features. Finally, different classes like healthy, severity-2, severity-2.5, and severity-3 are classified by the developed model accurately.

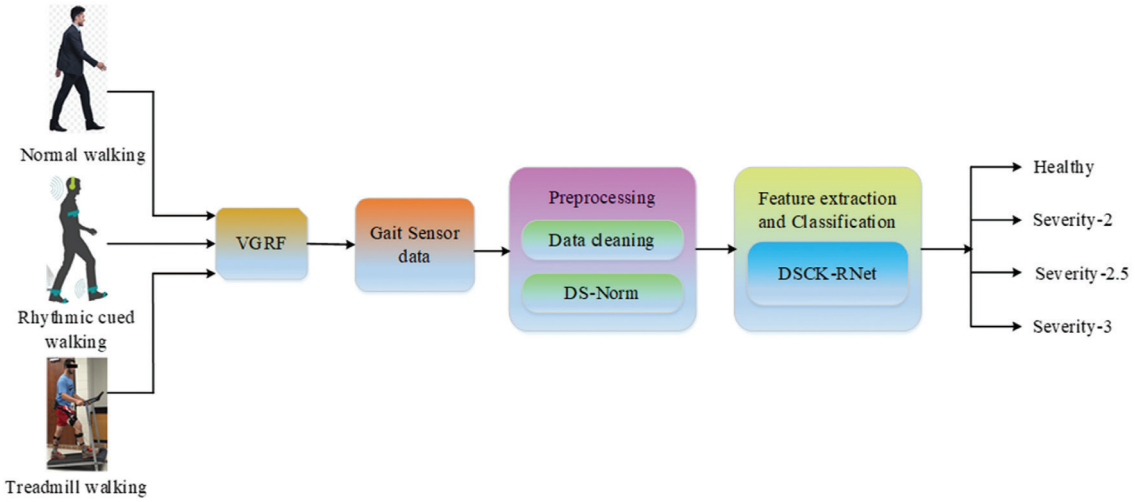


Fig. 1. Workflow of the developed model

4.1. PREPROCESSING

The raw sensor data collected from the human foot contains high noises, missing values, and data redundancy problems. To overcome this issue, data cleaning, and decimal scaling normalization (DS-Norm) processes are introduced in the initial stage to enhance the data quality. Initially, the data cleaning process is done in which the corrupted or missing values are identified and removed from the database. This process improves the accuracy performance and prevents data redundancy problems. Moreover, the overfitting issues and unwanted training complexities are eliminated using a data cleansing process.

After cleaning the data, a decimal scaling normalization technique is used that can minimize the non-linear complexities and speed up the training process. In the DS-Norm technique, the decimal points of the gait parameters are moved to obtain a normalized value. Here, the decimal points are altered using the extreme absolute value of the gait patterns. It can be mathematically formulated as,

$$\widetilde{D}_x = \frac{D_x}{10^n} \quad (1)$$

Here, $\tilde{D}_x$ indicates the normalized data, D_x represents the actual gait sensor data, and nrepresents the smallest integer in which $Max([d])<1$.

4.2. FEATURE EXTRACTION AND PD CLASSIFICATION USING THE DSCK-RNET MODEL

The preprocessed data are then fed into the proposed Duo spatiotemporal convoluted kernel-boosted ResNet (DSCK-RNet) model to classify the presence of PD by learning the gait spatiotemporal features.

The developed network model consists of a spatiotemporal network for extracting the gait spatiotemporal patterns and it is finally fed into a convoluted kernel-boosted ResNet model to select the appropriate features and to classify the PD severity levels accurately. The developed model decreases the processing time and prevents dimensionality issues with the aid of an effective kernel mechanism.

Initially, the preprocessed data are given into two different streams for extracting the gait spatiotemporal features. The initial stream considered the temporal features based on convoluted residual blocks and temporal attention (AT) blocks. Here, the narrow multi-layer feature extractor, max-pooling (MP) layer, AT blocks, and multi-scale residual blocks (RBs) are deployed. At first, the narrow gait features are extracted through multi-scale convolution blocks (CBs). Then, the AT blocks are emphasized to evaluate the inter-channel relationship between the gait parameters. Then, the MP layer is implemented to overcome the dimensionality issues during the training process. The multi-scale RBs help the model to understand the extracted features for enhancing the accuracy performance. In the same way, the spatial attention (SA) block and MP layer were determined effectively. At the final stage, the extracted features are then fed into the multi-scale RBs and SA blocks to generate high-level gait features. The second stream evaluates the extracted gait patterns and concentrates mostly on spatial features while processing with SA blocks. Similar to The A block, the SA contains single multi-scale CB and triple multi-scale RBs. After connecting the CBs, the TA blocks and MP layers are determined. Fig. 2 illustrates the architecture of the DSCK-RNet model.

After extracting the features, the convoluted kernel boosted the ResNet (CK-RNet) model to choose the

best features thereby minimizing the time complication and data idleness problems. In the RBs, convoluted kernel widths are utilized to select the most relevant gait parameters from the entire feature. The resultant Y_a of the a^{th} kernel is determined using a convolutional process and it is depicted below:

$$Y_a(i) = f(\sum \beta_a(i) * p + Bias_a(i)) \quad (2)$$

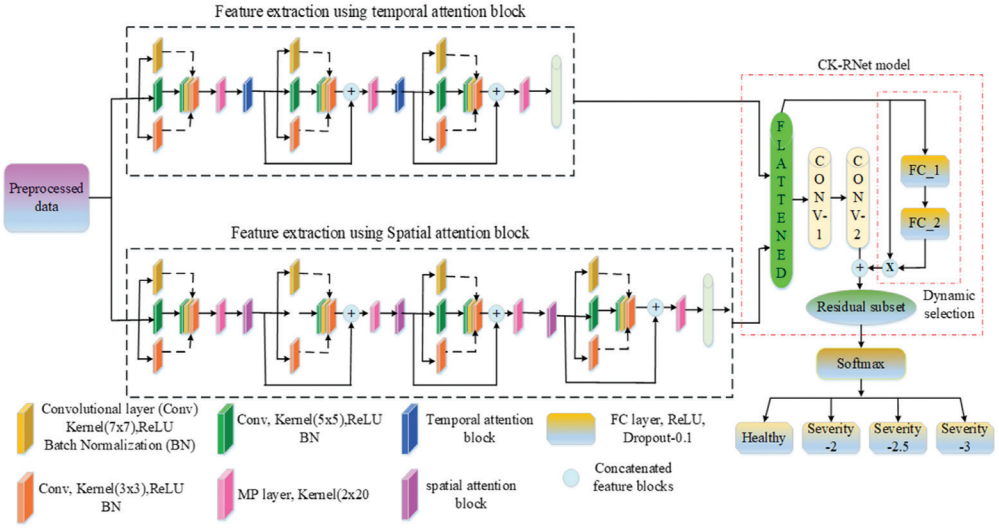


Fig. 2. Architecture of DSCK-RNet model

Here, findicates the activation function (AF), βrepresents the convoluted kernel, *represents the convolutional operation, and iindicates the number of channels. To highlight the useful features and to prevent irrelevant ones, an active selection of gait feature subsets is introduced behind the MP layers of each stream. The active selection comprises global average pooling, (GAP), and dual fully connected (FC) layers in which the GAP indicates the squeeze operation, and the FC layer indicates the excitation operation. For generating the relevant gait patterns, the obtained weights and channel-wise multiplication between incoming feature maps are trained carefully. The mathematical operations for selecting the appropriate features are depicted as follows:

$$x_c = G_{sq}(y_c) = \frac{1}{N}\sum_{n=1}^{N} y_c(n) \quad (3)$$

$$u = G_{ex}(x, U) = \sigma\big(\sigma(x, U)\big) = \sigma\big(U_2\sigma(x, U_1)\big) \quad (4)$$

$$\hat{y}_c = u_c \times y_c \quad (5)$$

Here, $U_1 \in R^{c/d}$and $U_1 \in R^{c/d}$ indicates the weights of dual FC layers, drepresents the reduction ratio and its value is 2, σ indicates the rectified linear unit (ReLU), u_c represents the scalar, and y_c indicates the actual extracted features.The feature subsets obtained by varying kernel widths are then combined using channel-wise concatenation operation and it is mathematically formulated as,

$$\tilde{Y} = con(Y_a), a = 1,2,\ldots..M \quad (6)$$

Here, M indicates the total number of convoluted kernels. Finally, the outcome from the CK-RNet model can be mathematically defined as,

$$z = F(p, [U]) + u \bullet p \quad (7)$$

Here, p indicates the input of RB, $F(p,[U])$indicates the residual subsets, and Uindicates the weights of active feature selection. The outcome of a convoluted kernel block contains the features obtained under varying kernels. The outcome of MP in the a^{th} channel can be mathematically expressed as,

$$MP_a(m) = max[z_a(mU, (m+1)U], 0 \le m \le \frac{n}{r} \quad (8)$$

Here, z represents the input, U indicates the width of the pooling window, r indicates the stride windows, and n indicates the feature-length. After the MP layers, the learned features are given into the dual FC layers to reduce the feature dimensions and severity classification. A total 9 different features like phase time, pace time, strike time, posture time (s), rhythm (phase/sec), phase length (cm), pace length (cm), and speed (m/s) are considered for learning the severity levels of PD. The consequence of a a^{th} neuron can be mathematically formulated as,

$$F_a = f\big(u_a^{fc} MP_a + Bias_a\big) \quad (9)$$

Here, findicates the AF (ReLU) and the softmax are utilized to predict the outcome under each class. Also, u_a^{fc} indicates the weights of FC layers in a^{th} neurons.

5. RESULTS AND DISCUSSION

The developed study is processed and investigated via the Python platform and a publicly available Physio-Net dataset [21] is utilized in this study. This dataset consists of both the H&Y scale and UPDRS score in which three walking styles were collected from the movement disorder unit of the Tel-Aviv Sourasaky Medical Centre, Israel, and Laboratory for Gait & Neuro-dynamics. Totally 8 resistive force sensors are placed at the insole of indi-

vidual foot and asked to walk at a regular interval of 2 minutes. Finally, 16 VGRF signals are sampled to 100Hz. The utilized gait dataset was contributed by three researchers from [22]. For investigating the cause of outer cueing especially in pace length, and gait speed, [22] is utilized that consists of 29 PD patients with the severity-2.5 on the H&Y measure. In addition to this, 18 healthy samples for three mobile scenarios such as rambling on the ground without assistance, rambling on the ground with a rambler, and rambling on the drudgery.

In, gait patterns from 29 PD samples and 26 normal samples are considered for analyzing the pulsating acoustic stimulus (PAS) on pace-to-pace changeability. The cognitive and gait functions can be analyzed via which consist of 29 healthy samples and 35 PD patients with an average age of 71. It is noted that the pressure level of foot sensors in PD patients is half while compared with the healthy samples. In this dataset, a total of 18K samples are recorded for 2 minutes under rhythmic auditory stimulation (RAS). Totally 173 VGRF sensor data are available with a total number of gait samples represented in18×13K×173.

5.1. ASSESSMENT METRICS

In this section, the several metrics like accuracy (Acc), sensitivity (Sen), specificity (Spe), PPV (positive predictive value (PPV), FPR (false positive rate), MCC (Mathew's correlation coefficient (MCC) and precision (Pre) are analyzed and its mathematical formula is provided below:

$$Acc(\%) = \frac{w+x}{w+x+y+z} \times 100\% \quad (10)$$

$$Sen(\%) = \frac{x}{x+y} \times 100\% \quad (11)$$

$$Spe(\%) = \frac{w}{w+z} \times 100\% \quad (12)$$

$$PPV(\%) = \frac{x}{x+z} \times 100\% \quad (13)$$

$$F-score(\%) = 2 \times \frac{Pre \times Sen}{Pre+Sen} \times 100\% \quad (14)$$

$$MCC = \frac{x \times w - z \times y}{\sqrt{(x+z)(x+y)(w+z)(w+y)}} \quad (15)$$

$$FPR(\%) = \frac{z}{w+z} \times 100\% \quad (16)$$

Here, *w, x, y*, zindicates the true negative (TN), true positive (TP), false negative (FN), and false positive (FP) respectively.

5.2. PERFORMANCE STUDY OF DEVELOPED METHOD OVER CONVENTIONAL METHODS

In Fig. 3, 0 represents healthy, 1 represents severity-2, 2 represents severity-2.5, and 3 represents severity-3. For class 0 (healthy), a total of 3704 trials are deliberated and in these 3445 trials are correctly classified as healthy. The remaining 259 trials are wrongly classified as class 1 (84 trials), class 2 (158 trials), and class 3 (17 trials) respectively. For class 1 (severity-2), a total of 2090 trials are considered and in this 1886 trials are correctly classified as class 1. The residual 204 trials are misclassified as class 0 (121 trials), class 2 (64 trials) respectively, and class 3 (19 trials). For class 2 (severity-2.5), a total of 2843 trials are deliberated and in this 2245 trials are correctly classified as class 2. The remaining 598 trials are misclassified as class 0 (487 trials), class 1 (101), and class 3 (10 trials) respectively. For class 3 (severity-3), a total of 323 trials are considered and in these 307 trials are correctly classified as class 3. The residual 16 samples are wrongly classified as class 0 (15 trials) and class 1 (1 trial) respectively. The performance obtained by the developed method over the existing methods is analyzed via graphical illustration.

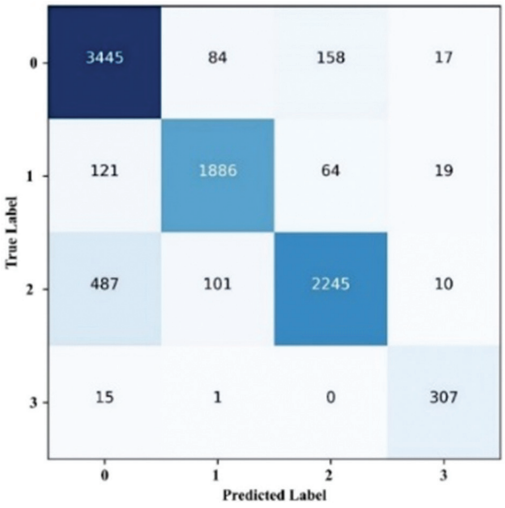


Fig. 3. Confusion matrix to describe the performance of the classification model

Numerous performance measures like accuracy, F-measure, sensitivity, specificity, PPV, FPR, and MCC are analyzed and compared with other conventional techniques like SVM, NB, KNN, and EC. A detailed analysis of the graphical illustrations depicted below: Fig. 4 indicates the accuracy analysis for different techniques under varying classes. In the graphical elucidation, it is clear that the suggested DSCK-RNet technique shows better outcomes compared to existing studies. Further, the proposed and existing techniques show better performance for identifying healthy samples by learning the gait patterns. However, the existing studies reduce its performance while analyzing the gait patterns for PD severity levels. This is due to that the existing techniques are trained with complex data and failed to consider non-motor symptoms effectively. The developed method considers the most relevant spatiotemporal patterns and can determine the invisible symptoms with minimal complexity.

Fig. 5 indicates the sensitivity analysis for different techniques under varying classes. In the graphical interpretation, it is clear that the suggested DSCK-RNet technique shows better outcomes compared to current studies. It is also illustrious that the proposed and existing techniques show better sensitivity performance for identifying healthy samples by learning the gait pat-

terns. However, the existing studies reduce its performance while analyzing the gait patterns for PD severity ratings. The existing techniques increase the overfitting issues due to high data redundancy problems. The developed method considers the most relevant spatiotemporal patterns and can determine the hidden symptoms without the aid of experienced clinicians.

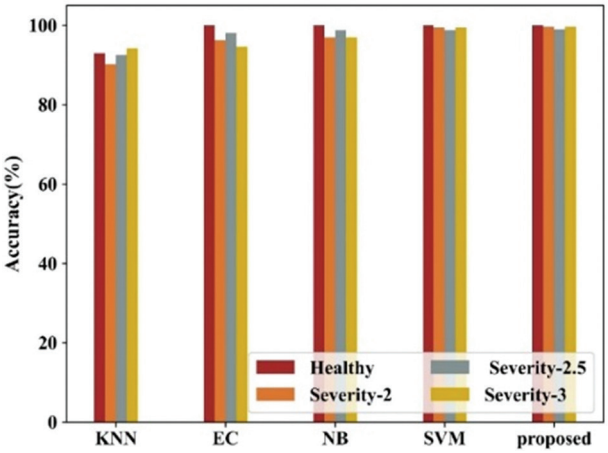


Fig. 4. Accuracy analysis for different techniques under varying classes

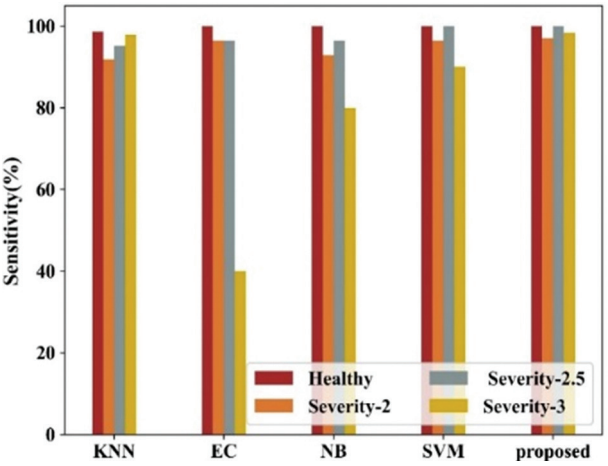


Fig. 5. Sensitivity analysis for different techniques under varying classes

Fig. 6 indicates the specificity analysis for different techniques under varying classes. In the graphical interpretation, it is clear that the developed DSCK-RNet technique demonstrates better outcomes compared to existing studies. It is also illustrating that the proposed and existing techniques show better specificity for identifying healthy samples by learning the gait patterns. However, the existing studies reduce its performance in learning the gait patterns, especially for PD severity levels. This is due to that the existing techniques lack effective preliminary stages to overcome the non-linear problems and also fail to consider non-motor symptoms effectively. The developed method considers the most relevant spatiotemporal gait patterns and can determine the motor and non-motor symptoms with minimal complexity.

Fig. 7 indicates the PPV analysis for different techniques under varying classes. In the graphical interpretation, it is clear that the developed DSCK-RNet technique demonstrates better outcomes compared to existing studies. It is also illustrating that the proposed and existing techniques show better performance for identifying healthy samples by learning the gait patterns. However, the existing studies reduce its performance while analyzing the gait patterns for PD severity levels. Moreover, the existing studies have very low performance while the complex data are under advanced PD levels. The proposed model tackles the discrepancies by considering the larger gait patterns into small ranges to minimize the network complexity.

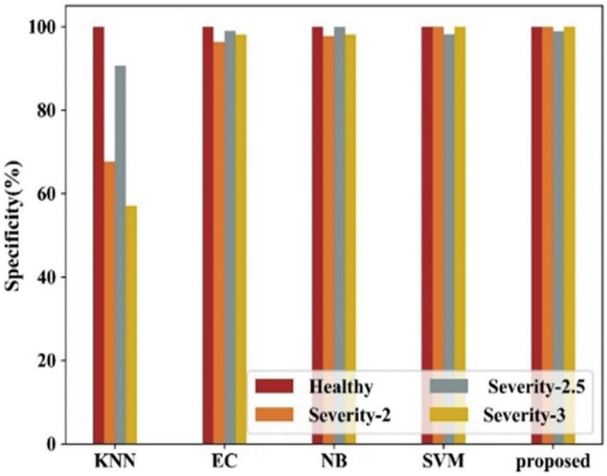


Fig. 6. Specificity analysis for different techniques under varying classes

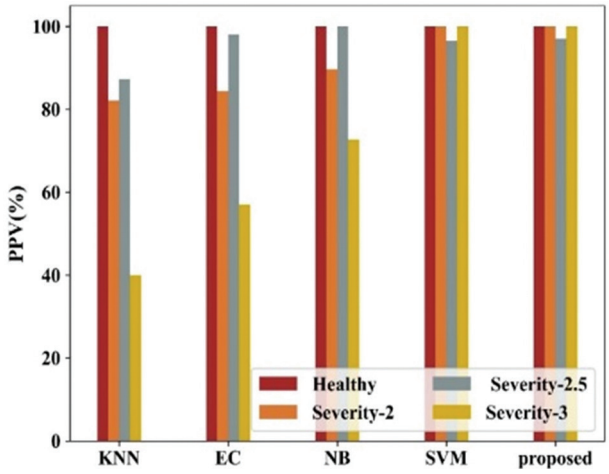


Fig. 7. PPV analysis for different techniques under varying classes

Fig. 8 indicates the FPR analysis for different techniques under varying classes. In the graphical elucidation, it is clear that the developed DSCK-RNet technique illustrates less performance compared to existing studies. It is also noted that the proposed and existing techniques show better performance for identifying healthy samples by learning the gait patterns. However, the present techniques reduce its performance while

analyzing the gait patterns for PD severity levels. This is due to that the existing techniques are trained with complex data and failed to consider non-motor symptoms for the training process. The developed method considers the most relevant spatiotemporal patterns and minimizes the training error. Fig. 9 indicates the precision analysis for different techniques under varying classes. In the graphical explication, it is clear that the developed DSCK-RNet technique contemplates better outcomes compared to traditional studies. It is also emaciating that the proposed and existing techniques show better precision for determining healthy samples by learning the gait patterns. However, the existing studies reduce its performance while analyzing the gait patterns for PD severity levels. This is due to that the existing techniques failed to consider the hidden gait patterns and trained with unwanted or corrupted samples. The developed method considers the most relevant spatiotemporal patterns and can determine the invisible symptoms with minimal complexity.

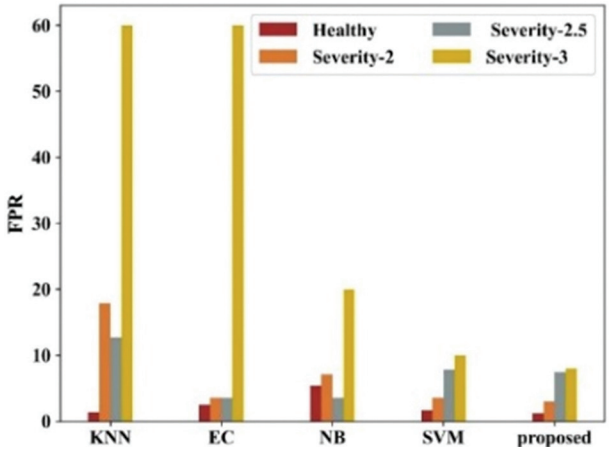


Fig. 8. FPR analysis for different techniques under varying classes

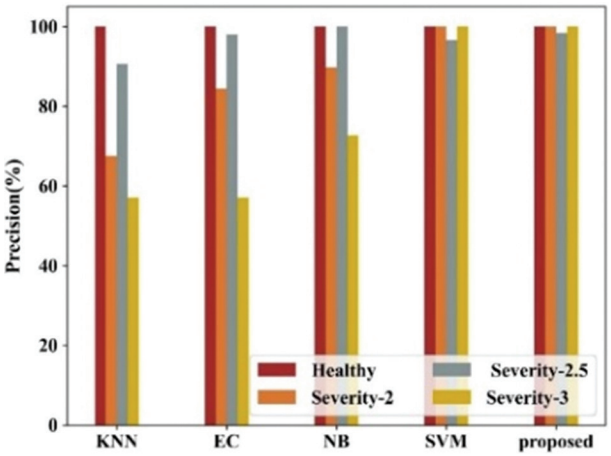


Fig. 9. Precision analysis for different techniques under varying classes

Fig. 10 indicates the F-measure analysis for different techniques under varying classes. In the graphical illustration, it is clear that the developed DSCK-RNet technique provides better outcomes compared to traditional approaches. It is also defined that the proposed and existing techniques show better performance for identifying healthy samples by learning the gait patterns. However, the existing techniques showed less score while it is trained with complex data and failed to consider non-motor symptoms effectively. The developed method considers the most relevant spatiotemporal patterns and can determine the invisible symptoms with minimal complexity.

Fig. 11 indicates the MCC analysis for different techniques under varying classes. In the graphical elucidation, it is clear that the developed DSCK-RNet technique shows better outcomes compared to traditional techniques. It is also illustrating that the proposed and existing techniques show better performance for identifying healthy samples by learning the gait patterns.

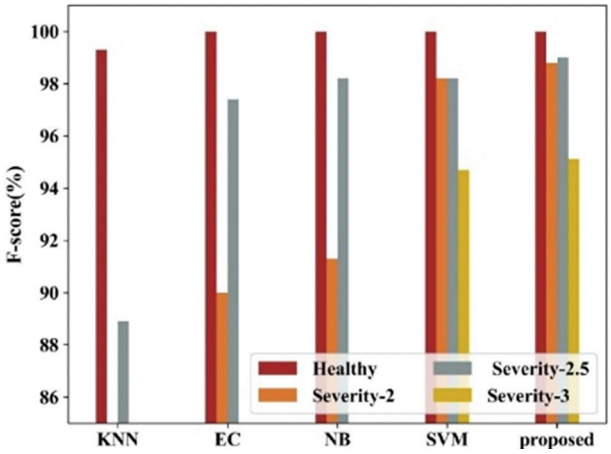


Fig. 10. F-measure analysis for different techniques under varying classes

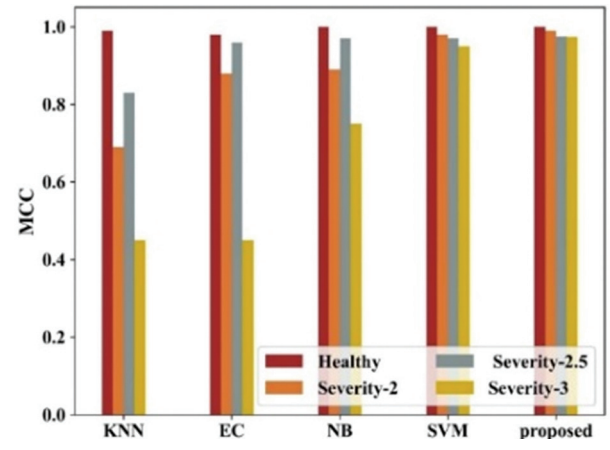


Fig. 11. MCC analysis for different techniques under varying classes

Table 1 shows the comparative analysis of different techniques for PD severity levels. From the table, it is elucidating that the developed approach achieved better outcomes compared to conventional studies. Some mL models like NB, SVM, EC, and KNN are compared with the developed DL model.

Table 1. Comparative analysis of different techniques for PD severity levels

Methods	Severity levels	Acc	Spe	FPR	Sen	F-score	PPV	Pre	MCC
Proposed (DSCK-RNet)	**0**	**100**	**100**	**1.2**	**100**	**100**	**100**	**100**	**1**
	1	**99.6**	**100**	**3**	**97**	**98.8**	**100**	**100**	**0.99**
	2	**99**	**98.9**	**7.45**	**100**	**99**	**97**	**98.4**	**0.975**
	3	**99.64**	**100**	**8**	**98.3**	**95.12**	**100**	**100**	**0.974**
EC [9]	0	100	100	-	100	100	100	100	-
	1	96.3	96.3	3.64	96.4	90	84.4	84.4	0.88
	2	98.1	99	3.62	96.4	97.4	98.1	98.1	0.96
	3	94.6	98.1	60	40	47	57.1	57.1	0.45
NB [9]	0	100	100	-	100	100	100	100	1.0
	1	97	97.8	7.13	92.9	91.3	89.7	89.7	0.89
	2	98.8	100	3.62	96.4	98.2	100	100	0.97
	3	97	98.1	20	80	76.2	72.7	72.7	0.75
SVM [9]	0	100	100	-	100	100	100	100	1.0
	1	99.4	100	3.62	96.4	98.2	100	100	0.98
	2	98.79	98.19	-	100	98.19	96.49	96.49	0.97
	3	99.39	100	10	90	94.69	100	100	0.95
KNN [9]	0	93	100	1.41	98.6	99.3	100	100	0.99
	1	90.2	67.6	17.94	91.9	74.1	82.1	67.6	0.69
	2	92.5	90.6	12.72	95.2	88.9	87.3	90.6	0.83
	3	94.2	57.1	60	97.9	47	40	57.1	0.45

6. CONCLUSION

The developed approach presented and investigated an innovative DL technique for classifying the PD severity level using gait spatiotemporal data. For learning the complex spatiotemporal features accurately, the developed study conquered an effective DSCK-RNet model that proved an outstanding performance in determining PD levels. Moreover, the developed DSCK-RNet technique deliberates the most appropriate features and minimizes the training complexity efficiently. In addition to this, some preliminary stages like data cleaning, missing value imputation, and DS-norm mechanisms helped to achieve better data quality and minimize the nonlinear correlations among the gait patterns. The developed study is processed and analyzed via the Python simulation and a publicly available Physio-Net dataset is utilized for the simulation process. Various assessment measures like accuracy, precision, sensitivity, specificity, PPV, FPR, and MCC are examined and compared with traditional studies like NB, SVM, EC, and KNN. In the experimental section, the developed DSCK-RNet model achieved an accuracy of 100%, 99.6%, 99%, and 99.64% for different classes like healthy, severity-2, severity-2.5, and severity-3 respectively. Even though the developed study achieved a better accuracy, there are some limitations and further extension of the effective PD analysis is required. The gait patterns considered in the present study consist of only healthy and moderate PD levels and failed to consider advanced PD levels due to a lack of data availability. Moreover, several unidentified motor symptoms, especially tremor signals will be integrated with the novel DL frameworks to improve the medical assessments and PD prognosis. Finally, the developed study failed to determine the time and frequency features that can be assessed. While data-driven gait-based severity classification using advanced machine learning models offers significant potential benefits for Parkinson's disease management, addressing the associated challenges is crucial for maximizing the model's clinical utility and ethical appropriateness. Collaboration between clinicians, researchers, and technology experts is essential to overcome these challenges and harness the full potential of this approach in improving patient care. A data-driven strategy utilizing machine learning approaches may, with additional validation, offer a more effective diagnostic and prognostic tool that can support doctors in their decision-making.

7. REFERENCES:

[1] H. Abujrida, E. Agu, K. Pahlavan, "DeepaMed: Deep learning-based medication adherence of Parkinson's disease using smartphone gait analysis", Smart Health, Vol. 30, 2023, p. 100430.

[2] A. Balakrishnan, J. Medikonda, P. K. Namboothiri, M. Natarajan, "Mahalanobis Metric-based Oversampling Technique for Parkinson's Disease Severity Assessment using Spatiotemporal Gait Parameters", Biomedical Signal Processing and Control, Vol. 86, 2023, p. 105057.

[3] M. T. García-Ordás, J. A. Benítez-Andrades, J. Aveleira-Mata, J. M. Alija-Pérez, C. Benavides, "Determining the severity of Parkinson's disease in patients using a multi task neural network", Multimedia Tools and Applications, Vol. 83, 2023, pp. 6077–6092.

[4] L. Borzì, L. Sigcha, D., Rodríguez-Martín, G. Olmo, "Real-time detection of freezing of gait in Parkinson's disease using multi-head convolutional neural networks and a single inertial sensor", Artificial Intelligence in Medicine, Vol. 135, 2023, p. 102459.

[5] Ç. B. Erdaş, E. Sümer, S. Kibaroğlu, "Neurodegenerative diseases detection and grading using gait dynamics", Multimedia Tools and Applications, Vol. 82, 2023, pp. 22925-22942.

[6] Y. Zhou, L. Xiao, X. Su, X. Xu, "Assessment of Parkinson's Motor Severity by Multi-feature and Multi-scale Motion Convolutional Neural Network with Positional Encoding", Proceedings of the 4th International Seminar on Artificial Intelligence, Networking and Information Technology, Nanjing, China, 16-18 June 2023, pp. 179-183.

[7] B. Guan, L. Yu, Y. Li, Z. Jia, Z. Jin, "Assessment of patients with Parkinson's disease based on federated learning", International Journal of Machine Learning and Cybernetics, Vol. 15, 2024, pp. 1621-1632.

[8] G. Cai et al. "Specific Distribution of Digital Gait Biomarkers in Parkinson's Disease Using Body-Worn Sensors and Machine Learning", The Journals of Gerontology: Series A, Vol. 78, No. 8, 2023, p. 101.

[9] J. F. Pedrero-Sánchez, J. M. Belda-Lois, P. Serra-Añó, S. Mollà-Casanova, J. López-Pascual, "Classification of Parkinson's disease stages with a two-stage deep neural network", Frontiers in Aging Neuroscience, Vol. 15, 2023, p. 1152917.

[10] P. L. Jadhwani, P. Harjpal, "A Review of Artificial Intelligence-Based Gait Evaluation and Rehabilitation in Parkinson's Disease", Cureus, Vol. 15, No. 10, 2023.

[11] M. Ullrich, "Fall Risk Prediction in Parkinson's Disease Using Machine Learning and Real-World Inertial Sensor-Based Gait Analysis", Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany, Ph.D. thesis, 2023.

[12] C. Dong, Y. Chen, Z. Huan, Z. Li, B. Zhou, Y. Liu, "Static-Dynamic Temporal Networks for Parkinson's Disease Detection and Severity Prediction", IEEE Transactions on Neural Systems and Rehabilitation Engineering, Vol. 31, 2023, pp. 2205-2213.

[13] U. K. Lilhore et al. "Hybrid CNN-LSTM model with efficient hyperparameter tuning for prediction of Parkinson's disease", Scientific Reports, Vol. 13, No. 1, 2023, p. 14605.

[14] E. Balaji, D. Brindha, V. K. Elumalai, R. Vikrama, "Automatic and non-invasive Parkinson's disease diagnosis and severity rating using LSTM network", Applied Soft Computing, Vol. 108, 2021, p. 107463.

[15] T. Aşuroğlu, H. Oğul, "A deep learning approach for Parkinson's disease severity assessment", Health and Technology, Vol. 12, No. 5, 2022, pp. 943-953.

[16] F. Setiawan, C. W. Lin, "Implementation of a deep learning algorithm based on vertical ground reaction force time–frequency features for the detection and severity classification of Parkinson's disease", Sensors, Vol. 21, No. 15, 2021, p. 5207.

[17] K. V. S. Sai Kumar, I. Sirisha, K. Vathsalya, K. K. V. V. Vamsi, "Parkinson Disease Diagnosis and Severity Rating Prediction Based on Gait analysis using Deep Learning", International Research Journal on Advanced Science Hub, Vol. 5, No. 5S, pp. 418-425.

[18] B. Vidya, P. Sasikumar, "Parkinson's disease diagnosis and stage prediction based on gait signal analysis using EMD and CNN–LSTM network", Engineering Applications of Artificial Intelligence, Vol. 114, 2022, p. 105099.

[19] O. El Ariss, K. Hu, "ResNet-Based Parkinson's Disease Classification", IEEE Transactions on Artificial Intelligence, Vol. 4, No. 5, 2023, pp. 1258-1268.

[20] S. M. Abdullah et al. "Deep Transfer Learning Based Parkinson's Disease Detection Using Optimized Feature Selection", IEEE Access, Vol. 11, 2023, pp. 3511-3524.

[21] J. Hausdorff, "Gait in Parkinson's disease", https://www.physionet.org/content/gaitpdb/1.0.0/ (accessed: 2023)

[22] S. Frenkel-Toledo, N. Giladi, C. Peretz, T. Herman, L. Gruendlinger, J. M. Hausdorff, "Treadmill walking as an external pacemaker to improve gait rhythm and stability in Parkinson's disease", Movement disorders: Official Journal of the Movement Disorder Society, Vol. 20, No. 9, 2005, pp. 1109-1114.

INTERNATIONAL JOURNAL OF ELECTRICAL AND COMPUTER ENGINEERING SYSTEMS

Published by Faculty of Electrical Engineering, Computer Science and Information Technology Osijek,
Josip Juraj Strossmayer University of Osijek, Croatia.

About this Journal

The International Journal of Electrical and Computer Engineering Systems publishes original research in the form of full papers, case studies, reviews and surveys. It covers theory and application of electrical and computer engineering, synergy of computer systems and computational methods with electrical and electronic systems, as well as interdisciplinary research.

Topics of interest include, but are not limited to:

- Power systems
- Renewable electricity production
- Power electronics
- Electrical drives
- Industrial electronics
- Communication systems
- Advanced modulation techniques
- RFID devices and systems
- Signal and data processing
- Image processing
- Multimedia systems
- Microelectronics
- Instrumentation and measurement
- Control systems
- Robotics
- Modeling and simulation
- Modern computer architectures
- Computer networks
- Embedded systems
- High-performance computing
- Parallel and distributed computer systems
- Human-computer systems
- Intelligent systems
- Multi-agent and holonic systems
- Real-time systems
- Software engineering
- Internet and web applications and systems
- Applications of computer systems in engineering and related disciplines
- Mathematical models of engineering systems
- Engineering management
- Engineering education

Paper Submission

Authors are invited to submit original, unpublished research papers that are not being considered by another journal or any other publisher. Manuscripts must be submitted in doc, docx, rtf or pdf format, and limited to 30 one-column double-spaced pages. All figures and tables must be cited and placed in the body of the paper. Provide contact information of all authors and designate the corresponding author who should submit the manuscript to https://ijeces.ferit.hr. The corresponding author is responsible for ensuring that the article's publication has been approved by all coauthors and by the institutions of the authors if required. All enquiries concerning the publication of accepted papers should be sent to ijeces@ferit.hr.

The following information should be included in the submission:

- paper title;
- full name of each author;
- full institutional mailing addresses;
- e-mail addresses of each author;
- abstract (should be self-contained and not exceed 150 words). Introduction should have no subheadings;
- manuscript should contain one to five alphabetically ordered keywords;
- all abbreviations used in the manuscript should be explained by first appearance;
- all acknowledgments should be included at the end of the paper:
- authors are responsible for ensuring that the information in each reference is complete and accurate. All references must be numbered consecutively and citations of references in text should be identified using numbers in square brackets. All references should be cited within the text;
- each figure should be integrated in the text and cited in a consecutive order. Upon acceptance of the paper, each figure should be of high quality in one of the following formats: EPS, WMF, BMP and TIFF;
- corrected proofs must be returned to the publisher within 7 days of receipt.

Peer Review

All manuscripts are subject to peer review and must meet academic standards. Submissions will be first considered by an editor-in-chief and if not rejected right away, then they will be reviewed by anonymous reviewers. The submitting author will be asked to provide the names of 5 proposed reviewers including their e-mail addresses. The proposed reviewers should be in the research field of the manuscript. They should not be affiliated to the same institution of the manuscript author(s) and should not have had any collaboration with any of the authors during the last 3 years.

Author Benefits

The corresponding author will be provided with a .pdf file of the article or alternatively one hardcopy of the journal free of charge.

Units of Measurement

Units of measurement should be presented simply and concisely using System International (SI) units.

Bibliographic Information

Commenced in 2010.
ISSN: 1847-6996
e-ISSN: 1847-7003

Published: semiannually

Copyright

Subscription Information

The annual subscription rate is 50€ for individuals, 25€ for students and 150€ for libraries.

Postal Address

Faculty of Electrical Engineering,
Computer Science and Information Technology Osijek,
Josip Juraj Strossmayer University of Osijek, Croatia
Kneza Trpimira 2b
31000 Osijek, Croatia

IJECES Copyright Transfer Form

(Please, read this carefully)

This form is intended for all accepted material submitted to the IJECES journal and must accompany any such material before publication.

TITLE OF ARTICLE (hereinafter referred to as "the Work"):

COMPLETE LIST OF AUTHORS:

The undersigned hereby assigns to the IJECES all rights under copyright that may exist in and to the above Work, and any revised or expanded works submitted to the IJECES by the undersigned based on the Work. The undersigned hereby warrants that the Work is original and that he/she is the author of the complete Work and all incorporated parts of the Work. Otherwise he/she warrants that necessary permissions have been obtained for those parts of works originating from other authors or publishers.

Authors retain all proprietary rights in any process or procedure described in the Work. Authors may reproduce or authorize others to reproduce the Work or derivative works for the author's personal use or for company use, provided that the source and the IJECES copyright notice are indicated, the copies are not used in any way that implies IJECES endorsement of a product or service of any author, and the copies themselves are not offered for sale. In the case of a Work performed under a special government contract or grant, the IJECES recognizes that the government has royalty-free permission to reproduce all or portions of the Work, and to authorize others to do so, for official government purposes only, if the contract/grant so requires. For all uses not covered previously, authors must ask for permission from the IJECES to reproduce or authorize the reproduction of the Work or material extracted from the Work. Although authors are permitted to re-use all or portions of the Work in other works, this excludes granting third-party requests for reprinting, republishing, or other types of re-use. The IJECES must handle all such third-party requests. The IJECES distributes its publication by various means and media. It also abstracts and may translate its publications, and articles contained therein, for inclusion in various collections, databases and other publications. The IJECES publisher requires that the consent of the first-named author be sought as a condition to granting reprint or republication rights to others or for permitting use of a Work for promotion or marketing purposes. If you are employed and prepared the Work on a subject within the scope of your employment, the copyright in the Work belongs to your employer as a work-for-hire. In that case, the IJECES publisher assumes that when you sign this Form, you are authorized to do so by your employer and that your employer has consented to the transfer of copyright, to the representation and warranty of publication rights, and to all other terms and conditions of this Form. If such authorization and consent has not been given to you, an authorized representative of your employer should sign this Form as the Author.

Authors of IJECES journal articles and other material must ensure that their Work meets originality, authorship, author responsibilities and author misconduct requirements. It is the responsibility of the authors, not the IJECES publisher, to determine whether disclosure of their material requires the prior consent of other parties and, if so, to obtain it.

- The undersigned represents that he/she has the authority to make and execute this assignment.
- For jointly authored Works, all joint authors should sign, or one of the authors should sign as authorized agent for the others.
- The undersigned agrees to indemnify and hold harmless the IJECES publisher from any damage or expense that may arise in the event of a breach of any of the warranties set forth above.

_____ _____

Author/Authorized Agent **Date**

CONTACT
International Journal of Electrical and Computer Engineering Systems (IJECES)
Faculty of Electrical Engineering, Computer Science and Information Technology Osijek
Josip Juraj Strossmayer University of Osijek
Kneza Trpimira 2b
31000 Osijek, Croatia
Phone: +38531224600,
Fax: +38531224605,
e-mail: ijeces@ferit.hr