



Statistical Machine Learning Methods for Condition Monitoring of Electric Drive Trains

Ms.V. Priyadharshini, Ms.C. Visali

PG Scholar, Department of Information Technology, Vivekanandha College of Engineering for Women, Tiruchengode, Tamil Nadu, India

Assistant Professor, Department of Information Technology, Vivekanandha College of Engineering for Women, Tiruchengode, Tamil Nadu, India

Publication History: Received: 25.02.2026; Revised: 20.03.2026; Accepted: 25.03.2026; Published: 28.03.2026.

ABSTRACT: This study presents the design and implementation of a non-contact vibration picker for capturing data from spinning machinery in order to detect bearing faults early. The Hilbert transform is used to denoise the collected vibration signals, and the dataset is then analysed by Principal Component Analysis (PCA) and Sequential Floating Forward Selection (SFFS) for dimensionality reduction and feature selection, respectively. The most essential attributes are then used to identify and categorize various bearing issues using Support Vector Machines (SVM) and Multi-Layer Perceptron (MLP) algorithms. The entire methodology provides an effective and proactive way for bearing health monitoring and maintenance, emphasizing fast defect identification and leading to significant savings in time, effort, and equipment maintenance costs.

KEYWORDS: Machine Learning, Electric Drive Trains, Condition Monitoring

I. INTRODUCTION

The increasing reliance on electric drive trains in a variety of sectors highlights the vital need for robust condition monitoring to maintain optimal performance and dependability. As these systems get more sophisticated, there is a greater need for advanced statistical machine learning algorithms to improve condition monitoring accuracy and efficiency. The objective of this comparative study is to investigate and assess several statistical machine learning algorithms in the context of

electric drive train status monitoring. This study aims to give useful insights into the creation of resilient and adaptable monitoring systems by analysing the strengths and limits of various methodologies.

1.1 MACHINE LEARNING

Machine learning, a type of artificial intelligence, has transformed how we tackle complicated issues in a variety of disciplines. It is a data-driven method that allows computers to learn and forecast or make judgments without being explicitly programmed. This game-changing technology has implications ranging from healthcare and banking to driverless cars and natural language processing. Machine learning is based on algorithms and models that can analyse and extract patterns from massive information, providing answers that were previously considered to be science fiction. Machine learning has become a vital tool for making sense of the massive and complicated information accessible to us in this era of data abundance, frequently outperforming human skills in tasks such as picture identification, language translation, and even game-playing. A number of reasons have contributed to the expansion of machine learning, including the availability of large datasets, advancements in processing power, and breakthroughs in algorithm development.

1.2 ELECTRIC DRIVE TRAINS

Electric drive trains represent a paradigm shift in transportation and industrial machinery, ushering in a new era of better efficiency, lower environmental impact, and improved performance. Electric drive trains, as opposed to traditional mechanical drive systems, use electrical power to propel cars and run machines. The worldwide goal to reduce the environmental footprint of transportation and industrial activities is driving this trend toward electrification. Electric drive trains have a wide range of applications, including electric automobiles and hybrid vehicles, as well as



industrial robots and renewable energy systems. As the world pursues more sustainable technologies, the need for dependable, high-performance electric drive trains has increased.

1.3 CONDITION MONITORING

Condition monitoring is an important aspect of contemporary engineering and maintenance procedures, since it provides a proactive method to assessing the health of machinery and systems. It entails the constant and real-time evaluation of numerous parameters to assure optimal equipment operation, forecast probable defects, and prevent unexpected failures. This strategy contrasts sharply with typical reactive maintenance, moving the emphasis from correcting problems after they occur to avoiding them before they impair performance. Condition monitoring is a strategic technique of improving dependability, reducing downtime, and optimizing operating efficiency in the context of machinery and industrial systems. Condition monitoring gives a full picture of equipment health by utilizing modern sensors, data analytics, and machine learning approaches. This enables early interventions and cost-effective maintenance plans.

II. LITERATURE SURVEY

PAPER NAME	PARAMETERS / TECHNICAL ASPECTS	ADVANTAGES	DISADVANTAGES
1. Deep Learning for Detection: A Survey – Li Liu et al.	Generic Object Detection frameworks (RCNN, Fast RCNN), feature extraction, proposal generation, datasets & metrics	Comprehensive view of object detection progress, highlights taxonomy, discusses training strategies	Too broad to cover all works, no experimental results, excludes video-based detection
2. Survey on Safety & Trustworthiness of DNNs – Xiaowei Huang et al.	Adversarial testing, robustness checks, interpretability, coverage-guided testing	Clear classification of safety methods, strong focus on certification and reliability	Scalability issues in verification, lack of unified standard across techniques
3. Increasing Trustworthiness via Accuracy Monitoring – Zhihui Shao et al.	Softmax-based accuracy monitor, transfer learning, Monte Carlo dropout ensemble	Estimates true accuracy without labels, adaptable to new datasets	Dependent on pre-trained models, limited to cost (1%), overhead in deployment
4. Online Monitoring of Object Detection Performance – Quazi M. Rahman et al.	Cascaded neural network, sliding window map prediction, feature-map extraction	Real-time monitoring, predicts mAP drop, improves safety in deployment	Requires feature access to detector backbone, additional complexity and computation
5. Per-Frame mAP Prediction – Quazi M. Rahman et al.	Frame-wise mAP thresholding, introspective perception, internal feature usage	Detects silent failures, need for ground-truth labels during deployment	Threshold sensitivity, risk of false alarms or missed failures
6. MetaDetect: Uncertainty & Quality Estimation – Marius Schubert et al.	Meta-classification, regression, IoU gradient boosting, meta-metrics	Accurate false positive detection, works with any handcrafted detector, interpretable metrics	Extra post-processing step, requires training structured meta-model
7. Introspective Vision for SLAM – Rabiee et al.	Context-aware noise modeling, autonomously supervised learning, feature prediction	Reduces SLAM tracking failures, improves robustness by 70%, adapts feature selection	Needs large training data, static image dependence, no temporal info used yet
8. Monitoring & Diagnosability of Perception Systems – Pasquale Antonante et al.	Graph theory-based diagnosability model, fault detection, system-level monitoring	Formal method for fault detection, scalable to multi-module systems, useful for certification	Theoretical model, minimal experimental validation, complex to implement in real-time
9. Automated Evaluation of Semantic Lidar –	Lidar-assisted ground truth generation, environmental	No manual labeling required, handles weather & lighting class, lidar dependency,	Focused mainly on road



PAPER NAME	PARAMETERS / TECHNIQUES USED	ADVANTAGES	DISADVANTAGES
Segmentation Robustness – Wei Zhou et al.	condition testing, robustness metrics	variations, validation pipeline	scalable limited to segmentation tasks
10. Online Monitoring for Neural Network-Based Pedestrian Pose Estimation Arjun Gupta et al.	ATOM (Adversarial.Trained Monitor) ORB feature extraction + Canny edge-based contour matching Keypoint distance threshold (≤ 26 px) & descriptor ratio test Loss prediction for 3D human pose outputs	Detects incorrect pose estimations in real-time ,Improves average error by 12.5% and worst-case error by 126.5% traditional model-based monitors ,First online monitoring system for 3D pose estimation	Requires adversarial training (computationally expensive) Performance drops under heavy occlusion or poor lighting ,Outperforms traditional model-based segmentation or mask alignment ,Additional latency added to inference pipeline

III. EXISITING SYSTEM

Fault detection and diagnosis are crucial for precise industrial machinery maintenance and management. In this context, data-driven condition monitoring models play a significant role in the diagnosis and management of equipment faults. This research looks on the use of various statistical machine learning techniques in modeling vast amounts of data in the condition monitoring of electric drive trains in supply chains. To discriminate between eleven failure states, large data sets are employed to train linear discriminant analysis, K-nearest neighbor method, naive Bayes, kernel naive Bayes, decision trees, and support vector machine. The decision trees achieved 93.8% accuracy in the testing data set, followed by kernel naive Bayes (91.9%), radial basis function (Gaussian) support vector machine (89.3%), linear discriminant analysis (84.5%), k-NN algorithm (80.5%), and Gaussian naive Bayes (71.3%). As a result, the statistical machine learning technique used effects classification accuracy in electric drive problem diagnostics. Furthermore, decision trees learn and categorize new examples from huge data in a matter of seconds. This simplifies the selection of decision trees for electric drive train status monitoring and management.

IV. PROPOSED SYSTEM

The proposed model acquires data from spinning machinery using a non-contact vibration pickup, allowing for early failure diagnosis in bearings. The Hilbert transform is used to denoise vibration signals, and the dataset is then analysed using Principal Component Analysis (PCA) and Sequential Floating Forward Selection (SFFS) for dimensionality reduction and feature selection, respectively. The most essential attributes are then used to identify and categorize various bearing issues using Support Vector Machines (SVM) and Multi-Layer Perceptron (MLP) algorithms. This complete methodology provides an effective and proactive way for bearing health monitoring and maintenance, emphasizing fast defect identification and leading to significant savings in time, effort, and equipment maintenance costs.

4.1 LOAD BEARING FAULT DATASET

The project focuses on obtaining and curating a comprehensive dataset especially specialized to load-bearing problems in rotating machinery in this module. This entails gathering vibration data under various load situations. The dataset serves as the foundation for further analysis, ensuring that the system is trained and tested on a wide variety of load-induced bearing defects. The thorough selection and compilation of this dataset are critical to the overall system's accuracy and dependability in recognizing and diagnosing errors connected to fluctuating loads.

4.2 FEATURE REDUCTION USING PCA BASED ON FEATURE EXTRACTION AND NORMALIZATION

This module addresses the need for fast feature extraction and normalization following the capture of the load-bearing fault dataset using Principal Component Analysis (PCA). PCA is used to minimize the dimensionality of a dataset while maintaining crucial information, improving computing performance and reducing the danger of overfitting. This phase is critical in preparing the data for following stages of analysis, as it ensures that the most significant properties are kept, adding to the system's capacity to properly distinguish between different fault scenarios.

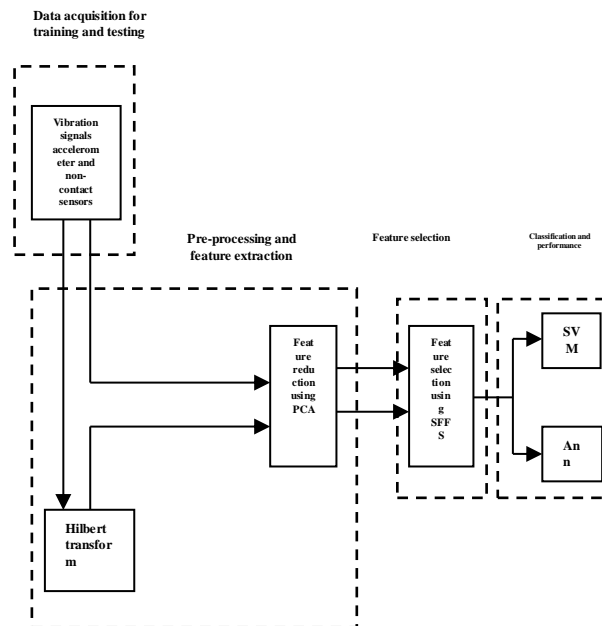


Figure 1 block diagram

4.3 SVM CLASSIFICATION BASED ON FEATURE SELECTION USING SFFS

The SVM classification module uses Support Vector Machines (SVM) to accurately classify faults based on the pre-processed information. Sequential Floating Forward Selection (SFFS) is used to enhance feature selection, a strategy that iteratively discovers and adds the most discriminative features to improve the model's performance. This guarantees that the SVM classifier is trained on the most relevant data, improving its capacity to detect and categorize load-bearing problems in the rotating gear under examination. The SVM classifier constructs an optimal hyperplane that separates fault categories with maximum margin, ensuring reliable decision boundaries even in high-dimensional spaces. To prevent overfitting, kernel functions such as Radial Basis Function (RBF), linear, or polynomial kernels are employed based on the nature of the feature distribution. Furthermore, cross-validation is implemented to fine-tune hyperparameters like the penalty parameter (C) and kernel coefficient (γ), ensuring optimal generalization on unseen data. The inclusion of Sequential Floating Forward Selection (SFFS) significantly reduces computational redundancy by eliminating irrelevant and redundant features. This dynamic feature selection approach not only adds the most informative features but also removes previously selected ones if they no longer contribute positively to classification accuracy. As a result, the feature subset remains compact, efficient, and highly discriminative. In addition, the decision-making process of the SVM module is enhanced by analyzing support vectors, which represent the most critical data points influencing the classifier's output. These support vectors help improve the interpretability of the model, allowing a deeper understanding of the underlying fault characteristics in rotating gear systems. The combination of SFFS and SVM thus offers a robust and precise diagnostic framework capable of handling complex vibration signals and noise interference. To further increase reliability, the system can be integrated with probabilistic outputs, providing confidence levels for each classification decision. This is particularly beneficial in real-time industrial applications, where timely and accurate fault detection is crucial to prevent equipment failure and costly downtime. Ultimately, the method enhances the efficiency and safety of rotating machinery by enabling early detection of abnormalities such as imbalance, misalignment, gear wear, or bearing defects.

4.4 MLP CLASSIFICATION BASED ON FEATURE SELECTION USING SFFS

This part, like the SVM classification module, incorporates Multi-Layer Perceptron (MLP) for fault classification. The SFFS feature selection procedure is used to discover and include the most useful characteristics, hence maximizing the MLP model's ability to distinguish between distinct load-bearing faults. The use of MLP adds a neural network-based method, providing flexibility in collecting nuanced patterns within data, thus improving the system's overall accuracy and resilience in detecting and diagnosing load-induced bearing failures. The MLP consists of multiple hidden layers with nonlinear activation functions such as ReLU or sigmoid, enabling it to capture complex, non-linear relationships present in vibration and acoustic signals. During training, backpropagation is employed in conjunction with optimization algorithms like Adam or stochastic gradient descent (SGD) to minimize classification error and enhance



convergence speed. To avoid overfitting, techniques such as dropout, L2 regularization, and early stopping are integrated into the training process, ensuring robust performance on unseen data. By combining SFFS with MLP, the system ensures that only the most discriminative and relevant features are used as inputs, reducing computational load while increasing classification precision. This selective learning process enhances the neural network's ability to focus on fault-sensitive signal attributes, such as harmonics, kurtosis, spectral energy, and time–frequency domain features. Furthermore, the adaptive learning capability of the MLP enables real-time adjustment to varying operating conditions, such as fluctuating speeds or load variations, which are common in industrial machinery. The output layer of the MLP uses a softmax activation function for multi-class classification, allowing the model to assign probability scores to different fault categories. This probabilistic output aids in decision-making, especially in cases where fault signatures overlap or appear ambiguous. Additionally, performance metrics such as accuracy, precision, recall, F1-score, and confusion matrices are used to evaluate and validate the effectiveness of the MLP-based classifier. To further improve diagnostic reliability, the system can be enhanced through ensemble techniques, where multiple MLP models are trained and their outputs combined to reduce variance and increase robustness. The integration of MLP with SFFS also promotes interpretability, as the selected features highlight the key signal components responsible for fault detection. Overall, this hybrid approach significantly boosts the system's reliability and ensures early detection of bearing defects, minimizing machine downtime and maintenance costs.

V. ALGORITHM DETAILS

A. Multilayer Perceptron (MLP)

This Model The pre-processed dataset was visualized using Matplotlib. The realization led to the insights of a classification problem upon the data pre-set with class labels (human presence or absence). These insights led to the decision to train an MLP model over other different options available. The data collected in the first phase of the experimentation was used to train an MLP over Tensor flow library in python. A sequential neural network was modelled with five hidden layers of varied neuron. The input dimensions of the data were 1429×16 . Model training was divided into training, validation, and testing. The training matrix was 1143×16 ; a 30% validation set was derived from the training set itself, whereas a matrix of 230×16 was utilized for the testing purpose. A 30% validation split was given to the dataset. Several network topologies with various layers and units in each layer were tested, and the bestperforming model was worked upon for further computations. A five-layer highly optimized MLP was built with the first four layers having the activation function as Relu, whereas the fifth layer uses the Sigmoid activation function to get a binary classification output

B. Support Vector Machine

Support Vector Machine (SVM) is a popular supervised machine learning model that is used for classification and prediction of unknown data. It is asserted by several researchers that SVM is a very accurate technique for text classification. It is also widely used in sentiment classification. For instance, if we have a dataset in which data is pre-labeled into two categories: positive and negative reviews, then we can train a model to classify new data into these two categories. This is exactly how SVM works. It is the model that we train on a dataset, so it can analyse and classify unknown data into the categories that were present in the training set. SVM is a linear learning method. It finds an optimal hyper-plane to differentiate two classes. Being a supervised classification model, it tries to maximize the distance between the closest training point and either class so as to achieve better classification performance on test data. The process for classification functions is as follows: It takes the labeled sample of data, and draws a line separating the two classes. This line is called the decision boundary. The solution is based only on those training data points which are really close to the decision boundary. The data points are called Support Vectors. For example, if we are categorizing movie reviews (in our case), one side of the boundary will have positive reviews while the other side has negative reviews. Now when new data needs to be classified, it goes either into the left or right side of the decision boundary.

VI. RESULT ANALYSIS

Two machine learning techniques, Support Vector Machines (SVM) and Multi-Layer Perceptron (MLP), were used to evaluate the suggested bearing failure detection system. The SVM method was 84% accurate overall, with precision, recall, and F-measure values of 0.79, 0.75, and 0.84, respectively. The MLP algorithm, on the other hand, outperformed the others, obtaining an accuracy of 92%. Notably, MLP demonstrated good accuracy (0.99) and recall (0.92) metrics, but with a somewhat lower F-measure of 0.89. These findings support the suggested methodology's efficacy, with the MLP algorithm beating SVM in terms of accuracy and precision, making it a potential choice for bearing defect recognition and classification in spinning equipment.



VII. CONCLUSION

In conclusion, the created non-contact vibration pickup, along with modern data processing and machine learning approaches, has shown to be a reliable and practical method for monitoring bearing health in rotating machinery. The Hilbert transform was used for denoising, PCA was used for dimensionality reduction, and SFFS was used for feature selection, which simplified the dataset and allowed for the exact detection and categorization of various bearing faults. The effective deployment of SVM and MLP algorithms illustrates the system's capacity to detect faults in real time. This complete technique not only improves the proactive nature of maintenance procedures, but it also offers significant savings in time, resources, and equipment upkeep expenses.

VIII. FUTURE WORK

It is critical for future work to investigate and enhance the suggested system to fit a greater range of industrial environments and machinery types. Further research into the non-contact vibration pickup's flexibility across diverse operational circumstances and settings will increase the system's versatility. Incorporating real-time monitoring capabilities and investigating the incorporation of future technologies like as edge computing or the Internet of Things might also give a more dynamic and responsive approach to bearing health monitoring. Continuous research efforts should be directed toward optimizing machine learning algorithms, with the possibility of adding deep learning models for enhanced pattern identification and fault detection.

REFERENCES

1. L. Liu, W. Ouyang, X. Wang, P. Fieguth, J. Chen, X. Liu, and M. Pietikäinen, "Deep learning for generic object detection: A survey," *International Journal of Computer Vision*, vol. 128, no. 2, pp. 261-318, Feb. 2020.
2. X. Huang, D. Kroening, W. Ruan, J. Sharp, Y. Sun, E. Thamo, M. Wu, and X. Yi, "A survey of the safety and trustworthiness of deep neural networks: verification, testing, adversarial attack and defense, and interpretability," *Comput. Sci. Rev.*, vol. 37, Aug. 2020, Art. no. 100270.
3. Z. Shao, J. Yang, and S. Ren, "Increasing the trustworthiness of deep neural networks through accuracy monitoring," in *Proc. Workshop Artif. Intell. Saf.*, 2020, pp. 1-8
4. Q. M. Rahman, N. Sünderhauf, and F. Dayoub, "Online monitoring of object detection performance post-deployment 2020; arXiv:2011.07750. [Online]. Available at <http://arxiv.org/abs/2011.07750>.
5. Q. M. Rahman, N. Sünderhauf, and F. Dayoub, "Per-frame mAP prediction for continuous performance monitoring of object detection during deployment," in *Proc. IEEE/CVF Winter Conf. Appl. Comput. Vis. (WACV) Workshops*, Jan. 2021, pp. 152-160.
6. M. Schubert, K. Kahl, and M. Rottmann, "MetaDetect: Uncertainty quantification and prediction quality estimates for object detection," 2020. arXiv:2010.01695. [Online], <http://arxiv.org/abs/2010.01695>.
7. S. Rabiee and J. Biswas, "IV-SLAM: Introspective vision for simultaneous localization and mapping," 2020, arXiv:08.02760. [Online], <http://arxiv.org/abs/2008.02760>.
8. C.Nagarajan and M.Madheswaran - 'Stability Analysis of Series Parallel Resonant Converter with Fuzzy Logic Controller Using State Space Techniques'- Taylor & Francis, *Electric Power Components and Systems*, Vol.39 (8), pp.780-793, May 2011. DOI: 10.1080/15325008.2010.541746
9. C.Nagarajan and M.Madheswaran - 'Experimental verification and stability state space analysis of CLL-T Series Parallel Resonant Converter' - *Journal of Electrical Engineering*, Vol.63 (6), pp.365-372, Dec.2012. DOI: 10.2478/v10187-012-0054-2
10. C.Nagarajan and M.Madheswaran - 'Performance Analysis of LCL-T Resonant Converter with Fuzzy/PID Using State Space Analysis'- Springer, *Electrical Engineering*, Vol.93 (3), pp.167-178, September 2011. DOI 10.1007/s00202-011-0203-9
11. S.Tamilselvi, R.Prakash, C.Nagarajan, "Solar System Integrated Smart Grid Utilizing Hybrid Coot-Genetic Algorithm Optimized ANN Controller" *Iranian Journal Of Science And Technology-Transactions Of Electrical Engineering*, DOI10.1007/s40998-025-00917-z,2025
12. S.Tamilselvi, R.Prakash, C.Nagarajan, " Adaptive sliding mode control of multilevel grid-connected inverters using reinforcement learning for enhanced LVRT performance" *Electric Power Systems Research* 253 (2026) 112428, doi.org/10.1016/j.epr.2025.112428
13. S.Thirunavukkarasu, C. Nagarajan, 2024, "Performance Investigation on OCF and SCF study in BLDC machine using FTANN Controller," *Journal of Electrical Engineering And Technology*, Volume 20, pages 2675–2688, (2025), doi.org/10.1007/s42835-024-02126-w



14. C. Nagarajan, M.Madheswaran and D.Ramasubramanian- 'Development of DSP based Robust Control Method for General Resonant Converter Topologies using Transfer Function Model'- Acta Electrotechnica et Informatica Journal , Vol.13 (2), pp.18-31, April-June.2013, DOI: 10.2478/aeii-2013-0025.
15. C.Nagarajan and M.Madheswaran - 'DSP Based Fuzzy Controller for Series Parallel Resonant converter'- Springer, Frontiers of Electrical and Electronic Engineering, Vol. 7(4), pp. 438-446, Dec.12. DOI 10.1007/s11460-012-0212-0.
16. C.Nagarajan and M.Madheswaran - 'Experimental Study and steady state stability analysis of CLL-T Series Parallel Resonant Converter with Fuzzy controller using State Space Analysis'- Iranian Journal of Electrical & Electronic Engineering, Vol.8 (3), pp.259-267, September 2012.
17. C.Nagarajan and M.Madheswaran, "Analysis and Simulation of LCL Series Resonant Full Bridge Converter Using PWM Technique with Load Independent Operation" has been presented in ICTES'08, a IEEE / IET International Conference organized by M.G.R.University, Chennai.Vol.no.1, pp.190-195, Dec.2007
18. Suganthi Mullainathan, Ramesh Natarajan, "An SPSS and CNN modelling based quality assessment using ceramic materials and membrane filtration techniques", Revista Materia (Rio J.) Vol. 30, 2025, DOI: <https://doi.org/10.1590/1517-7076-RMAT-2024-0721>
19. M Suganthi, N Ramesh, "Treatment of water using natural zeolite as membrane filter", Journal of Environmental Protection and Ecology, Volume 23, Issue 2, pp: 520-530,2022
20. P. Antonante, D. I. Spivak, and L. Carlone, "Monitoring and diagnosability of perceptual systems," 2020; arXiv:2005.11816. [Online], <http://arxiv.org/abs/2005.11816>.
21. W. Zhou, J. S. Berrio, S. Worrall, and E. Nebot, "Automated evaluation of semantic segmentation robustness for autonomous driving," IEEE Transactions on Intelligence and Transactional Systems, vol. 21, no. 5, pp. 1951-1963, May 2020.
22. Gupta and L. Carlone, "Online monitoring for neural network based monocular pedestrian stance estimation 2020; arXiv:2005.05451. [Online]: <http://arxiv.org/abs/2005.05451>.
23. Anand, L., Tyagi, R., & Mehta, V. (2024, January). Food recognition using deep learning for recipe and restaurant recommendation. In Proceedings of Eighth International Conference on Information System Design and Intelligent Applications (pp. 269-279). Singapore: Springer Nature Singapore.
24. Kumar, K. A., Vimal, V. R., & Karthick, V. (2024, June). Error Rate Increment of Quantum Computer using Quantity Computational Algorithm Comparing with Deutsch Jozsa Algorithm. In 2024 OPJU International Technology Conference (OTCON) on Smart Computing for Innovation and Advancement in Industry 4.0 (pp. 1-5). IEEE.
25. Poornima, G., & Anand, L. (2025). Medical image fusion model using CT and MRI images based on dual scale weighted fusion based residual attention network with encoder-decoder architecture. Biomedical Signal Processing and Control, 108, 107932.