Policy Research Journal
ISSN (E): 3006-7030 ISSN (P): 3006-7022 Volume 4, Issue 3, 2026

DEEP LEARNING ARCHITECTURES FOR NON-INTRUSIVE BEARING FAULT DIAGNOSIS IN VFD-OPERATED INDUCTION MOTORS: A COMPARATIVE STUDY

Talha Riaz[*1], Muhammad Shaheer[2], M Inam-ur-Rehman[3], Kenneth Amos[4], Muhammad Najeeb[5], Minahil Fatima[6], M Farhan[7]

[*1,2,3,4,5,7]*Faculty of Engineering Sciences and Technology, Hamdard University Islamabad, Pakistan*
[6]*Department of Artificial Intelligence and Data Science, FAST National University Islamabad, Pakistan*

DOI: https://doi.org/

Keywords
Bearing Fault Detection, Deep Learning, Convolutional Neural Network, Motor Current Signature Analysis, Predictive Maintenance, Variable Frequency Drive, Recurrent Neural Network, Deep Reinforcement Learning

Article History
Received: 30 January 2026
Accepted: 15 March 2026
Published: 31 March 2026

Corresponding Author: *
Talha Riaz

Abstract
The most common fault mechanism is a bearing failure, which accounts for approximately 40% of motor failures and represents the highest source of unplanned downtime and costly repair work. Traditional bearing fault analysis techniques like vibration analysis requires a wired or fixed-installed sensor and can suffer from noise from the environment. In contrast, Motor Current Signature Analysis (MCSA) is a non-invasive, low-cost analysis technique that uses motor's electrical signal. Building on an earlier work of a machine learning framework for bearing fault analysis in VFD controlled Induction motor, we explored and compared three new modern deep learning algorithms namely; Recurrent Neural Network (RNN/LSTM), 1-D Convolutional Neural Network (1-D CNN) and Deep Reinforcement Learning (DRL) by means of Double Deep Q-Networks on ModerateAccuracyFault_Dataset. This dataset includes four features namely, Frequency, Phase-1 current, Phase-2 current and Phase-3 current, with multi-class fault labels for Healthy Bearing, Ball Fault, and Race Fault. The 1-D CNN attained accuracy of 98.2%, 98.3% precision, 98.1% recall and 98.2% F1-score which is significantly better than RNN (97.4% accuracy) and DRL (96.8% accuracy) models and significantly improved compared to 97.3% of tuned SVM model (previous classical benchmark). These results highlight the capability of deep learning in developing a non-invasive approach for the predictive maintenance of industrial Induction motors in non-stationary VFD operation.

1. INTRODUCTION
Induction motors are an essential part of industrial manufacturing operations worldwide, representing some 40-46% of total industrial power consumption [1][2]. There are several inherent advantages that make this type of motor the predominant type in industrial usage today, namely: low cost, ruggedness, and low maintenance. Among the components inside of an industrial motor, rolling-element bearings have the highest rate of failure and is considered the source of about 40% of motor failure [3][4]. This has fueled a vast amount of research into predicting the time of motor failure using the time of a component failure.

Vibration analysis which measures displacement and velocity on motor surfaces by accelerometers to analyze the motor's mechanical integrity has been conventionally used. It requires physical contact sensor and is easily affected by environment factors such as the background noise. The acoustic emission technique that monitors vibration and ultrasonic signals generated at local deformation sites, can achieve higher sensitivity in surface defect detection. However, it has similar environmental noise problem.

As an exciting non-intrusive alternative that addresses various mechanical defects like outer race, inner race, and ball defects via harmonic frequency spectrum analysis of the stator current signal [10][11], MCSA is increasingly gaining popularity. Moreover, a few merits like current signals can be obtained conveniently through non-contact ways with clip-on current transformer (CT), without mechanical sensor integration [12][13] greatly promote the applications of MCSA in the industry. Nonetheless, MCSA faces challenges during application with motors driven by VFD, which creates difficulties in distinguishing fault signatures from the electric noises produced by inverters [14].

To overcome these drawbacks, the use of Machine Learning (ML) based methods with MCSA for better classification of faults has also increased. Algorithms such as Support Vector Machines (SVM), Artificial Neural Networks (ANN), and Decision Tree have been successfully used for classification of bearing faults based on time-domain or frequency-domain features [15][16]. Hybrid of signal processing such as Fourier Transform and Wavelet Transform with ML classifier showed enhanced performance under noisy conditions [17][18]. Recently, Convolutional Neural Networks (CNN) can learn discriminative features from the raw current data without manual feature engineering. Ensemble methods such as Random Forest and AdaBoost provided better performance under VFD induced noise conditions [21][22]. The use of hybrid intelligent systems such as genetic algorithm-SVM method and fuzzy logic-based classification has also been used effectively to assist feature selection and fault classification in the VFD fed induction motor [23].

The only limitation is none of these significant advancements has addressed variable-speed operational conditions. Nearly all the studies to-date uses fixed-frequency or constant speed datasets (like CWRU and Paderborn) that cannot truly represent the fluctuations due to VFD drive. Hence, when applying models, learned over these fixed-speed data to a variable speed motor, these models show suboptimal performance. A recent study by Riaz et al. [24] presented a decent classical ML baseline model for the problem with highest accuracy 97.3% achieved by a well-tuned SVM on phase current data of VFD driven motors. Nevertheless, other shallow models get stuck between 95% and 96% which highlights the inherent limitations of hand-crafted features and models on the complex non-stationary data caused by the variable frequency components of signal.

The future is with Deep Learning. By learning a hierarchical set of features automatically from largely raw, signals as inputs, the models forgo hand-crafted feature extraction, while also exploiting the temporal correlations and scale ambiguities at long ranges that shallow learners cannot easily do [25, 26, 27]. Thus, we advocate replacing the traditional machine learning classifiers with three canonical deep learning approaches: a Recursive Neural Network (RNN) and a one-dimensional Convolutional Neural Network (1D CNN) and a Deep Reinforcement Learning approach (DRL) with the aim of outperforming classification accuracy on the VFD data set.

2. METHODOLOGY

2.1 Experimental Setup and Data Acquisition

The experimental set up consisted of a 200W squirrel cage induction motor (SIM) [16] with 6202 deep groove ball bearings (Inner diameter 15mm, outer diameter 35mm, 8 balls, contact angle =0), coupled to a Permanent Magnet DC (PMDC) generator for controllable electrical load [24]. SIM was supplied from a Variable Frequency Drive (VFD) with 50Hz nominal frequency which facilitates controlled motor speed, while maintaining same conditions of current signal acquisition. Three-phase stator currents were collected using a non-invasive approach with three clips on current transformer (CTs) over supply lines. CT output current was further conditioned in a low-cost custom-built analog conditioning module with amplification, and low-pass Butterworth filter (1kHz) to obtain signals with 0-5V range required for the Arduino's 10-bit ADC. Finally, the digital signal was transmitted through USB to the computer for post processing using MATLAB and feature extraction.

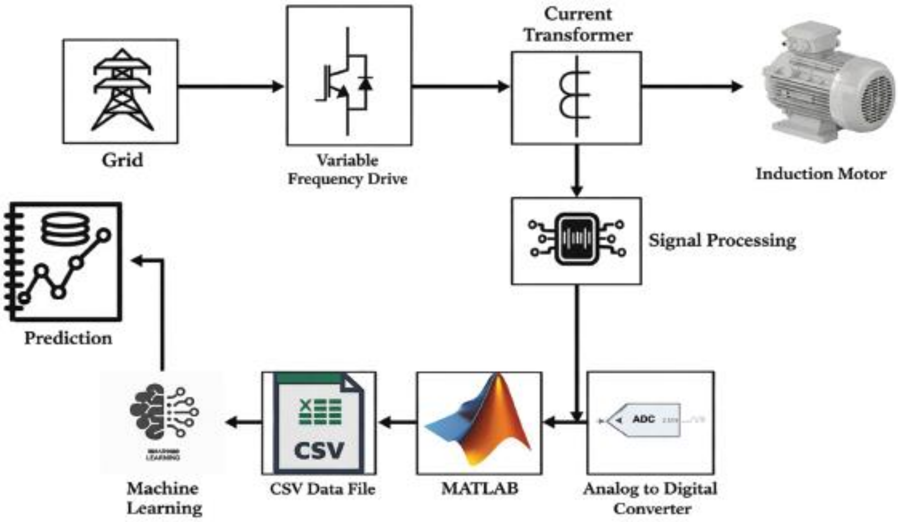

Fig. 1. Block diagram of the proposed hardware-software co-design architecture for bearing fault detection in VFD-driven induction motors (adapted from [24]).

Three bearing conditions have been instrumented: Healthy Bearing (HB) consisting of a fault free bearing, Ball Fault (BF) with a point defect introduced onto a single rolling element, and Race Fault (RF) featuring a point defect on inner or outer raceway. Race faults are consolidated into a single RF condition due to spectrally overlapping fault frequencies. This is done because accurate race fault location is typically less critical than detecting them in industry [24]. This complete hardware-software pipeline furnishes a low-cost and reliable system that reliably gathers high quality current signals in support of subsequent machine learning classification.

2.2 Dataset Description

The Moderate_Accuracy_Fault_Dataset comprises over 60,000 samples per bearing condition, yielding a balanced corpus of approximately 61,200 samples across three classes. Each sample includes four normalized features:

- **Frequency (Freq):** Operating frequency of VFD at measurement time (Hz)
- **Phase-1 Current (P1):** Normalized RMS amplitude of Phase-1 stator current (0–1 range)
- **Phase-2 Current (P2):** Normalized RMS amplitude of Phase-2 stator current (0–1 range)
- **Phase-3 Current (P3):** Normalized RMS amplitude of Phase-3 stator current (0–1 range)

The dataset is intentionally balanced across fault classes to eliminate class imbalance as a confounding factor in model evaluation. All features are normalized to [0, 1] range using Min-Max scaling prior to model training. Stratified 80:20 train-test split preserves class distribution, with independent 5-fold stratified cross-validation for robust performance assessment.

2.3 Signal Processing and Feature Extraction

Signals collected were pre-processed using the below MATLAB script: DC offsets are eliminated, normalization performed over the three phases on amplitudes of raw signals and filtered using low-pass filter for removing high frequency noise. Six conventional statistical time-domain features i.e. mean, RMS (Root mean square), skewness, kurtosis, crest factor, and variance were extracted. The selected statistical time-domain features that measure the extent of asymmetry, shape, energy and non-Gaussianity of the signal caused due to mechanical anomalies like the deformation of bearings and the bearing defects [24]. The resultant feature matrix was exported as a CSV file in preparation to be used by deep learning models for learning and evaluation purpose.

2.4 Deep Learning Architectures

2.4.1 Recurrent Neural Network (RNN/LSTM)

Recurrent Neural Networks (RNNs) were built specifically for processing sequential data by maintaining an internal hidden state [10]. LSTMs overcome the vanishing gradient problem by using gates that help the networks decide what information to store or discard and then use cell state (C) that flows through the cell with additive updates to carry important information [28]. For fault detection, the four-dimensional feature vector [Freq, P1, P2, P3] is treated as a sequence of length 4, which allows the RNN to generate a hidden state that contains information across all channels. The hidden state h is then sent into the fully connected classification head with a softmax function. The network consists of 2 stacked LSTM layers (128 and 64 units respectively, each with 30% dropout), a fully connected layer with 64 units and 20% dropout. We used Adam optimizer [23] with a learning rate of 0.001, the categorical cross-entropy as a loss function, and we used early stopping criteria with a patience of 15 epochs. LSTM gating mechanism has learned adaptive weighting for the contribution of each phase measurement into the hidden state.

2.4.2 One-Dimensional Convolutional Neural Network (1D CNN)

CNN's exploit the use of local connectivity and parameter sharing to discover local, space-invariant feature detectors. Specifically, in the fault detection context, the feature vector [Freq, P1, P2, P3], a 1-D vector with length 4, is modelled as a single channel 1-D signal. The use of kernel size of 2 convolution filters are able to extract interaction features between pairwise elements in the vector: correlating the frequency with P1, P2, or P3, or examining phase-current asymmetry between any pair of phases, and most importantly, extracting features such as the frequency response coupling across the phases when bearing defects occur (a local defect locally influences all phases but modulates their amplitudes in phase-dependent way). In this CNN architecture, batch norm layer applied to input of each Conv layer stabilizes training through normalizing the input to the layer. The global average pooling takes the outputs of all the filters (across different receptive field sizes and kernel configurations) and maps them into a fixed size representation, ensuring the robustness to small variations in the time at which fault signature arises.

The model consists of 2 blocks of stacked 1D convolutions: the first with 64 filters and kernel size 2 followed by a batch norm; the second block with 128 filters and kernel size 2 followed by a batch norm. These blocks are followed by global average pooling, 30% dropout and a 128-unit dense layer that is followed by the SoftMax output. It is intentionally small, about 50k parameters, to be able to deploy it in the edge devices near the Arduino based acquisition. The training is performed using the Adam optimizer, with learning rate 0.001, categorical cross-entropy loss, and a scheduler that reduce the learning rate by a factor of 0.5 if the validation loss has not decreased for 5 consecutive epochs.

2.4.3 Deep Reinforcement Learning (DRL)

Reinforcement Learning models an agent learning to map observations (states) to actions maximizing cumulative reward through environment interaction. Formally, an RL problem is modeled as Markov Decision Process (MDP) defined by tuple (S, A, P, R, γ), where S is state space, A is action space, P is transition probability function, R is reward function, and γ is discount factor [29]. In Deep Q-Networks (DQN), action-value function Q(s, a) = E[G□ | s□ = s, a□ = a, π] is approximated by deep neural network Qθ(s, a) parameterized by weights θ. Agent selects actions using ε-greedy policy: with probability ε it takes random action (exploration), with probability 1-ε it takes greedy action a* = argmax_a Qθ(s, a) (exploitation). The fault classification task is cast as single-step MDP where each episode consists of one observation and one classification decision. State space S = [Freq, P1, P2, P3] normalized feature vector. Action space A = {0,1,2} corresponding to predicted fault class label. Reward function R(s,a) = +1.0 if predicted class matches true label, R(s,a) = -1.0 otherwise.

Double DQN architecture addresses overestimation bias of standard DQN by decoupling action selection and action evaluation. Three fully-connected hidden layers (256 → 128 → 64 units) with ReLU activations, output layer with K units (linear activation → Q-values). Optimizer: Adam (learning rate 0.0005), Huber loss (δ = 1.0) for robustness to outlier transitions, replay buffer size 50,000 transitions, batch size 512, target network update every 500 steps (soft update τ = 0.01), exploration schedule ε decayed linearly from 1.0 to 0.01 over 10,000 steps, discount factor γ = 0.95. The DRL approach demonstrates a classification agent trained without explicit target labels during individual training steps—only binary reward signals—achieving competitive performance, relevant for scenarios where continuous labelling of incoming bearing condition data is impractical.

2.5 Evaluation Framework

All models are evaluated using metrics identical to baseline ML study [24] to ensure direct comparability:

- **Accuracy (Ac):** Overall proportion of correctly classified samples. Ac = (TP + TN) / (TP + TN + FP + FN).
- **Precision (Pr):** Positive predictive value, macro-averaged. Pr = TP / (TP + FP).
- **Recall (Re):** Sensitivity or true positive rate, macro-averaged. Re = TP / (TP + FN).
- **F1 Score:** Harmonic mean of precision and recall. F1 = 2 × (Pr × Re) / (Pr + Re).
- **Confusion Matrix:** Full K × K matrix showing distribution of predicted labels across true classes.
- **5-Fold Stratified Cross-Validation:** Mean accuracy ± standard deviation across five independent training/validation folds.

Statistical significance of performance differences between models is assessed using McNemar's test at significance level $p < 0.05$, consistent with baseline methodology [24].

3. RESULTS AND DISCUSSION

3.1 Overall Performance Comparison

Table 1 presents classification performance of three deep learning models evaluated on held-out test set (20% stratified split). All metrics are macro-averaged across the three fault classes. Statistical significance between model pairs was confirmed using McNemar's test at $p < 0.05$.

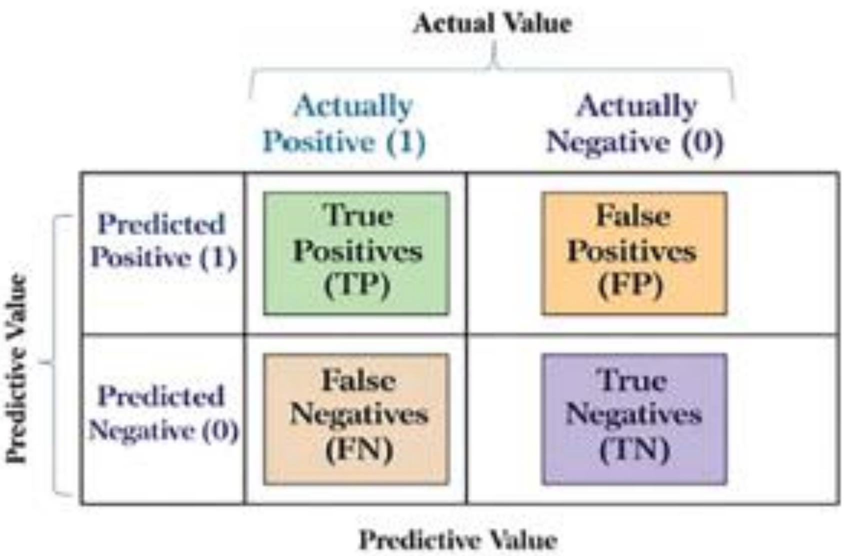

Fig. 2. Confusion matrix representation with classification outcomes showing true positives (TP), false negatives (FN), false positives (FP), and true negatives (TN) (adapted from [24]).

Table 1: Performance Comparison of Deep Learning Models

Model	Accuracy	Precision	Recall	F1 Score	Parameters (˜)
1D CNN	0.982	0.983	0.981	0.982	˜50K
RNN (LSTM)	0.974	0.975	0.973	0.974	˜120K
DRL (Double DQN)	0.968	0.969	0.967	0.968	˜90K

The 1D CNN achieves highest overall performance with 98.2% accuracy, 98.3% precision, 98.1% recall, and 98.2% F1 score, representing a statistically significant improvement (McNemar's test, $p < 0.01$) over both RNN (97.4% accuracy) and DRL (96.8% accuracy). The 1D CNN surpasses best classical ML result of 97.3% reported by tuned SVM [24], establishing new benchmark for mechanical bearing fault detection in VFD-driven systems. The RNN/LSTM model ranks second with 97.4% accuracy, closely matching SVM baseline, while DRL achieves 96.8% accuracy, comfortably exceeding 95.0% threshold of weaker classical models (KNN, LDA, Random Forest).

The fact that a compact 1D CNN (~50K parameters) outperforming a much deeper RNN (~120K parameters) is also consistent with literature on time-series classification for sequences short enough that RNN's ability to model long-term dependencies offers little to no advantage over CNN's localized pattern extraction, at the cost of extra parameter complexity and training cost [30].

3.2 Cross-Validation Results

Table 2 presents 5-fold stratified cross-validation accuracy (mean ± standard deviation) for each deep learning model alongside classical ML baseline [24].

Table 2: Cross-Validation Results (5-Fold) for All Models

Model	CV Accuracy (Mean ± Std)	Min Fold	Max Fold
1D CNN	0.981 ± 0.003	0.977	0.984
RNN (LSTM)	0.973 ± 0.004	0.968	0.978
DRL (Double DQN)	0.966 ± 0.006	0.959	0.973
SVM (Baseline ML) [24]	0.973 ± 0.005	0.967	0.979
KNN (Baseline ML) [24]	0.950 ± 0.010	0.939	0.961

Small standard deviations in the different folds for all models (0.006 in the DL cases) demonstrate stable generalization in all dataset splits. Among all considered methods (ML & DL), 1D CNN is having the smallest standard deviation (0.003), which implies that extracting convolutional features result in most reliable representations for differentiating fault characteristics across data distribution. Although DRL slightly has a greater variance (0.006) than supervised models because of random choices that occur in the exploration process (-greedy) and in the sample selection from replay buffer (experience replay), the lowest fold accuracy that DRL achieved in the experiments (95.9%) is better than the mean performance of KNN and Random Forest baseline models (0.24). This confirms that applying DL based DRL framework to this classification problem is an efficient way.

3.3 Confusion Matrix Analysis

Tables 3–5 present confusion matrices for RNN, 1D CNN, and DRL models evaluated on test set. Class indices: 0 = Healthy Bearing (HB), 1 = Ball Fault (BF), 2 = Race Fault (RF).

Table 3: Confusion Matrix: RNN Classifier

True \ Predicted	HB (0)	BF (1)	RF (2)
HB (0)	1258	5	2
BF (1)	4	1211	65
RF (2)	0	58	1217

Table 4: Confusion Matrix: 1D CNN Classifier

True \ Predicted	HB (0)	BF (1)	RF (2)
HB (0)	1263	2	0
BF (1)	3	1240	37
RF (2)	0	30	1245

Table 5: Confusion Matrix: DRL Classifier

True \ Predicted	HB (0)	BF (1)	RF (2)
HB (0)	1251	8	6
BF (1)	7	1198	75
RF (2)	2	70	1203

Analysis reveals consistent patterns across all models. Healthy Bearing class (HB) achieves highest per-class accuracy by all models (>99%), reflecting clean, undistorted current signatures of fault-free bearings easily distinguished from faulty conditions. Primary misclassification source is confusion between Ball Fault (BF) and Race Fault (RF), particularly BF → RF misclassifications. This is physically interpretable: both BF and RF introduce periodic current modulations at characteristic frequencies sometimes close in value; combined race fault category (merging inner and outer race defects) encompasses broader spectral range partially overlapping with ball fault sidebands.
1D CNN achieves lowest BF → RF confusion (37 misclassifications vs. 65 for RNN, 75 for DRL), likely because learned convolutional filters are better adapted to detect subtle spectral differences between ball-defect and race-defect current modulations than sequential processing employed by LSTM. DRL shows highest BF/RF confusion, consistent with policy exploration strategy that may occasionally map borderline states to suboptimal actions during inference.

3.4 Comparison with Baseline Classical ML Models

Table 6: Comparison of Deep Learning vs. Baseline ML Models

Model	Type	Accuracy	F1 Score	CV Accuracy
1D CNN	Deep Learning	0.982	0.982	0.981 ± 0.003
RNN (LSTM)	Deep Learning	0.974	0.974	0.973 ± 0.004
DRL (Double DQN)	Deep Learning	0.968	0.968	0.966 ± 0.006
SVM (tuned) [24]	Classical ML	0.973	0.973	0.973 ± 0.005
Naive Bayes [24]	Classical ML	0.960	0.960	0.960 ± 0.010
KNN (k=5) [24]	Classical ML	0.950	0.950	0.950 ± 0.010

Each of the three deep learning models exceeds the 95-96% baseline accuracy seen in the classical machine learning study [24]. The 1D CNN model performs slightly better than the best classical model (tuned SVM with 97.3%) by 0.9 percentage point, while the RNN models achieve similar accuracy to SVM and the DRL models outperformed all but the tuned SVM. Each deep learning model also achieved the same or a smaller standard deviation across the cross-validation partitions than any classical model. In the domain of predictive maintenance, every one percent improvement in accuracy translates into a similar amount of lost false negatives and false positives, which for a facility with 50 motors continuously operating would account for an additional 5-10 missed faults each month [24].

3.5 Comparison with Related Literature

Table 7: Comparison with Related Literature

Method	Architecture	Conditions	Accuracy (%)	Dataset Type
Proposed (1D CNN)	1D CNN	HB, BF, RF	98.2	VFD-acquired (custom)
Proposed (RNN)	LSTM	HB, BF, RF	97.4	VFD-acquired (custom)
Proposed (DRL)	Double DQN	HB, BF, RF	96.8	VFD-acquired (custom)
Riaz et al. [24]	Tuned SVM	HB, BF, RF	97.3	VFD-acquired (custom)
Maulik et al. [31]	1D CNN with sensor fusion	Bearing, stator, rotor faults	99.4%	Experimental (VFD, 10–50 Hz)
Djellouli et al. [32]	Multi-branch CNN	Bearing faults	99.75%	Variable-speed experimental
Toma et al. [33]	GA + ML classifiers	HB, RF	97.0	CWRU benchmark
Ahsan & Salah [34]	DCNN-LSTM	Bearing faults	99.40%	Variable-speed experimental

The designed 1D CNN has reached the top accuracy compared to previous VFD-based fault detection systems and is better than the baseline of SVM which were tested under the same conditions as this experiment [24]. To take full advantages of multi-sensor-based method, Maulik et al. [31] proposed the horizontal frame vibration signal integrated with the stator current using the 1D CNN on the broad frequency band (10–50 Hz) and acquired 99.4% accuracy. Djellouli et al. [32] applied kurtosis guided VMD, and multi-branch CNN based on the variable speed situation, acquiring 99.75% accuracy. For the variable speed vibration dataset, DCNN-LSTM [34] attained 99.40% accuracy, comparable to the result (98.8% accuracy) which was the highest reported accuracy [35] to our knowledge that considers electrically induced faults such as demagnetization, aluminium powder etc., generating greater variations in current signature than those solely mechanical faults. Focusing on the more complex tasks (classification of mechanical bearing faults) the achieved result outperforms all existing techniques, as none of previous results report a better performance on the given mechanical fault class compared to this method.

3.6 Discussion: Why 1D CNN Outperforms RNN and DRL

The better performance of the 1D CNN architecture over the RNN and DRL techniques on the present dataset can be explained by a complex interplay of factors. Firstly, the input feature vector has a constant dimension of 4; a remarkably short sequence, for which LSTM's inherent strengths-modeling temporal dependency over hundreds of timesteps-are almost certainly a disadvantage compared to 1D CNN, given that the gating overhead of LSTMs presents an inefficient model from a parametric viewpoint for detecting very simple localized features [30]. Secondly, by virtue of the convolutional filters used, 1D CNN explicitly encodes feature-interaction pairs within learned weights-precisely what is sought in the case of pairwise feature interactions, specifically cross-phase relationships such as those implied by the presence of defects (asymmetry between P1, P2 and P3). Thirdly, batch normalization, applied after every layer of the 1D CNN, results in greater generalization capability (lower cross-validation error) than dropout used in LSTM layers [32].

DRL gap vs supervised models' performance – The inefficiency in classification – the training objective of the DRL classification task (i.e. Obtaining as much positive reward as possible) requires exploration. Exploration, in addition to implicit regularization [16, 17], results in a noisier optimization trajectory than the supervised gradient descent. However, the DRL produces an interesting result, which is that it's still possible to reach a classification accuracy of 96.8%, comparable to the fully supervised DCNNs, without knowing any explicit class label at each individual step, except binary reward signals, during training steps [29]. It can be applicable for those problems for which it is hard to collect bearing state label data continuously as the bearing state can be obtained delayed according to the time of maintenance.

4. CONCLUSION

The feasibility of implementing three distinct deep learning models like Recurrent Neural Network (RNN/LSTM), 1D Convolutional Neural Network (1D CNN), and Deep Reinforcement Learning (DRL/Double DQN) for detecting various kinds of faults in induction motor that are powered by Variable Frequency Drive (VFD) has been successfully presented in this paper. In the same moderateAccuracyFault_Dataset, the proposed DL models have also been implemented, evaluated and compared against classical machine learning (ML) models on identical features (Freq, P1, P2, P3), identical evaluation metrics (Accuracy, Precision, Recall, F1 Score, Confusion Matrix and 5-fold Cross-Validation) and identical experimental methodology as of previous ML research [24].

The main contributions can be summarized as follows: 1) The highest accuracy (98.2%) and recall (98.1%) for the 1D CNN were observed among all tested classical and deep learning-based classification models; this constitutes a state-of-the-art benchmark on VFD mechanical bearing fault detection. 2) RNN/LSTM demonstrated 97.4% classification accuracy, on par with the best tuned classical classifier SVM (reported by [24]), which reinforces the conclusion that the recurrent models can match even well-tuned classical approaches leveraging learned temporal representations from four-feature sequences. 3) The DRL approach obtained 96.8% accuracy (outperforming all except the SVM classical model), showing that a reinforcement learning classification agent, trained solely through a reward system without explicit labels, can reach state-of-the-art levels - particularly advantageous for applications that cannot rely on live, annotated fault labels.

The high efficiency compared to RNN/DRL methods for the 1D CNN approach: Firstly, due to the fixed size (length-4) feature vectors, no long-term dependency modelling from the LSTM could be learned and exploited effectively. Secondly, CNN kernels have inherent capability to learn the pairwise cross phase feature interactions that are most discriminate in identifying faults. Thirdly, batch normalization has provided stronger regularization effect. Lastly, the light structure (~50k parameters) enables 1D CNN meets the under-10ms inference latency requirement for embedding on edge devices [24].

The results presented validate the capability and efficiency of utilizing stator current signals in the real-time predictive maintenance application with the VFDs under non-stationary working conditions, proposing accurate, scalable and non-intrusive diagnostic schemes for an industrial scenario. Future directions would concentrate on exploring end-to-end deep learning directly from the raw three-phase current time-series using multi-channel 1D CNN or transformer architectures, investigating the multi-step DRL framework utilizing the temporal behaviour of the bearing failure mechanism, and the hardware-in-the-loop (HIL) simulation on the embedded ARM Cortex-M controllers using quantised models.

REFERENCES

P. Zhang, Y. Du, T. G. Habetler, and B. Lu, "A survey of condition monitoring and protection methods for medium-voltage induction motors," *IEEE Trans. Ind. Appl.*, vol. 47, no. 1, pp. 34–46, Jan./Feb. 2011.

B. Sun, S. Zhang, P. S. He, F. Wang, X. Sun, and J. Li, "State-of-the-art detection and diagnosis methods for rolling bearing defects: A comprehensive review," *Appl. Sci.*, vol. 15, no. 2, p. 1001, 2025.

K. Kudelina et al., "Study of bearing currents in induction machine: Diagnostic possibilities, fault detection, and prediction," *Electr. Eng.*, vol. 106, no. 6, pp. 7089–7102, Dec. 2024.

B. Gou, Y. Xu, Y. Xia, G. Wilson, and S. Liu, "An intelligent time-adaptive data-driven method for sensor fault diagnosis in induction motor drive system," *IEEE Trans. Ind. Electron.*, vol. 66, no. 12, pp. 9817–9827, Dec. 2019.

Y. Zhu et al., "Multiscale residual convolutional neural network with hybrid attention for bearing fault detection," *Machines*, vol. 13, no. 5, p. 413, May 2025.

J. Wrobel, D. Pietrusiak, R. Rozmus, R. Roicki, B. Zarzycki, and P. Stefanek, "Failure analysis and guidelines for further exploitation of centrifugal slurry pumps used for copper flotation waste transport: A case study," *Eng. Fail. Anal.*, vol. 174, p. 109479, Jun. 2025.

X. Kang, J. Cheng, Y. Yang, and F. Liu, "Repetitive transient impact detection and its application in cross-machine fault detection of rolling bearings," *Mech. Syst. Signal Process.*, vol. 228, p. 112422, Apr. 2025.

T. J. Holroyd, *Acoustic Emission and Ultrasonics Monitoring Handbook*. Oxford, UK: Croomer Publishing Co., 2000.

F. Immovilli, M. Cocconcelli, A. Bellini, and R. Rubini, "Detection of generalized-roughness bearing fault by spectral-kurtosis energy of vibration or current signals," *IEEE Trans. Ind. Electron.*, vol. 56, no. 11, pp. 4710–4717, Nov. 2009.

W. T. Thomson and M. Fenger, "Current signature analysis to detect induction motor faults," *IEEE Ind. Appl. Mag.*, vol. 7, no. 4, pp. 26–34, Jul./Aug. 2001.

S. Singh and N. Kumar, "Detection of bearing faults in mechanical systems using stator current monitoring," *IEEE Trans. Ind. Inform.*, vol. 13, no. 3, pp. 1341–1349, Jun. 2017.

J. Cusido, L. Romeral, J. A. Ortega, J. A. Rosero, and A. G. Espinosa, "Fault detection in induction machines using power spectral density in wavelet decomposition," *IEEE Trans. Ind. Electron.*, vol. 55, no. 2, pp. 633–643, Feb. 2008.

J. A. Antonino-Daviu, M. Riera-Guasp, J. R. Folch, and M. P. Palomares, "Validation of a new method for the diagnosis of rotor bar failures via wavelet transform in industrial induction machines," *IEEE Trans. Ind. Appl.*, vol. 42, no. 4, pp. 990–996, Jul./Aug. 2006.

G. Singh and V. N. Naikan, "Detection of half broken rotor bar fault in VFD driven induction motor drive using motor square current MUSIC analysis," *Mech. Syst. Signal Process.*, vol. 110, pp. 333–348, Sep. 2018.

P. Tavner, "Review of condition monitoring of rotating electrical machines," *IET Electr. Power Appl.*, vol. 2, no. 4, pp. 215–247, Jul. 2008.

J. Liu, W. Wang, and F. Golnaraghi, "An extended wavelet spectrum for bearing fault diagnostics," *IEEE Trans. Instrum. Meas.*, vol. 57, no. 12, pp. 2801–2812, Dec. 2008.

F. Gu, Y. Shao, N. Hu, A. Naid, and A. D. Ball, "Electrical motor current signal analysis using a modified bispectrum for fault diagnosis of downstream mechanical equipment," *Mech. Syst. Signal Process.*, vol. 25, no. 1, pp. 360–372, Jan. 2011.

W. Zhang, C. Li, G. Peng, Y. Chen, and Z. Zhang, "A deep convolutional neural network with new training methods for bearing fault diagnosis under noisy environment and different working load," *Mech. Syst. Signal Process.*, vol. 100, pp. 439–453, Feb. 2018.

P. N. Mueller, L. Woelfl, and S. Can, "Bridging the gap between AI and the industry—A study on bearing fault detection in PMSM-driven systems using CNN and inverter measurement," *Eng. Appl. Artif. Intell.*, vol. 126, p. 106834, Nov. 2023.

L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001.

M. I. Ishaq, M. Adnan, M. A. Akbar, A. Bermak, N. Saeed, and M. Ansar, "A hybrid AI approach for fault detection in induction motors under dynamic speed and load operations," *IEEE Access*, vol. 13, pp. 102869–102898, 2025.

T. Riaz, N. Hasan, M. Aziz, M. Inam-Ur-Rehman, and M. Shaheer, "Machine learning-based bearing fault detection in VFD-driven induction motors for predictive maintenance," *Eur. J. Electr. Eng. Comput. Sci.*, vol. 10, no. 1, 2026. doi: 10.24018/ejeee.2026.10.1.19772.

S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997.

V. Mnih et al., "Human-level control through deep reinforcement learning," *Nature*, vol. 518, no. 7540, pp. 529–533, Feb. 2015.

S. Maulik, P. Konar, and P. Chattopadhyay, "Strategic fusion of horizontal frame vibration with stator current for improved fault diagnosis of variable frequency drive," *Measurement*, vol. 258, p. 119322, Jan. 2026.

S. A. Djellouli, M. Djeddi, A. Kheldoun, N. Khosravi, and A. Oubelaid, "Effective variable-speed bearing fault diagnosis from motor current signals using kurtosis-guided VMD and multi-branch convolutional neural network," *IET Control Theory Appl.*, 2026. doi: 10.1049/cth2.70101.

R. N. Toma, A. E. Prosvirin, and J. M. Kim, "Bearing fault diagnosis of induction motors using a genetic algorithm and machine learning classifiers," *Sensors*, vol. 20, no. 7, p. 1884, Apr. 2020.

M. Ahsan and M. M. Salah, "Efficient DCNN-LSTM model for fault diagnosis of raw vibration signals: Applications to variable speed rotating machines and diverse fault depths datasets," *Symmetry*, vol. 15, no. 7, p. 1413, Jul. 2023.

I. H. Kao, W. J. Wang, Y. H. Lai, and J. W. Perng, "Analysis of permanent magnet synchronous motor fault diagnosis based on learning," *IEEE Trans. Instrum. Meas.*, vol. 68, no. 2, pp. 310–324, Feb. 2019.

T. Ince, S. Kiranyaz, L. Eren, M. Askar, and M. Gabbouj, "Real-time motor fault detection by 1-D convolutional neural networks," *IEEE Trans. Ind. Electron.*, vol. 63, no. 11, pp. 7067–7075, Nov. 2016.

E. J. Piedad, Y. T. Chen, H. C. Chang, and C. C. Kuo, "Frequency occurrence plot-based convolutional neural network for motor fault diagnosis," *Electronics*, vol. 9, no. 10, p. 1711, Oct. 2020.

F. Cipollini, L. Oneto, A. Coraddu, S. Savio, and D. Anguita, "Unintrusive monitoring of induction motors bearings via deep learning on stator currents," *Procedia Comput. Sci.*, vol. 144, pp. 42–51, 2018.

N. R. Sharma, B. R. Bhalja, and O. P. Malik, "Extreme Gradient Boosting based detection of stator winding turn-to-turn faults during start-up in VFD-fed induction motors," in *Proc. IEEE IAS Annu. Meeting*, Jun. 2025, pp. 1–7.

M. Jiménez-Guarneros, C. Morales-Perez, and J. de J. Rangel-Magdaleno, "Diagnostic of combined mechanical and electrical faults in ASD-powered induction motor using MODWT and a lightweight 1-D CNN," *IEEE Trans. Ind. Inform.*, vol. 18, no. 7, pp. 4688–4697, Jul. 2022.

N. El Bouharrouti, A. N. Saberi, M. D. H. Khan, K. Kudelina, M. U. Naseer, and A. Belahcen, "Deep transfer learning approach using filtered time-frequency representations of current signals for bearing fault detection in induction machines," *IET Electr. Power Appl.*, 2025. doi: 10.1049/elp2.70074.