Proceedings of the Estonian Academy of Sciences
2026, **75**, 4, 368–382

https://doi.org/10.3176/proc.2026.4.10

www.eap.ee/proceedings
Estonian Academy Publishers

MACHINE FAULT DIAGNOSIS

RESEARCH ARTICLE

Received 30 June 2026
Accepted 23 September 2026
Available online 16 November 2026

Keywords:
fault diagnosis, simulation-to-reality gap, broken rotor bar, deep learning, transfer learning, CNN-LSTM, induction motors

Corresponding author:
Bishal Silwal
Bishal.silwal@pcampus.edu.np

Citation:
Aryal, A., Bhandari, A., Gnawali, A., Fatima, M. and Silwal, B. 2026. CNN-LSTM transfer learning for sim-to-real broken rotor bar fault diagnosis in induction motors. *Proceedings of the Estonian Academy of Sciences*, **75**(4), 368–382. https://doi.org/10.3176/proc.2026.4.10

© 2026 Authors. This is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0).

CNN-LSTM transfer learning for sim-to-real broken rotor bar fault diagnosis in induction motors

Aalok Aryal[a], Aashirbad Bhandari[a], Arun Gnawali[a], Ankeet Gadtaula[a], Mehroz Fatima[b] and Bishal Silwal[a]

[a] Department of Electrical Engineering, Pulchowk Campus, Institute of Engineering, Tribhuvan University, Lalitpur, Nepal

[b] Department of Electrical Power Engineering and Mechatronics, Tallinn University of Technology, Tallinn, Estonia

ABSTRACT

Broken rotor bar (BRB) faults in three-phase induction motors remain challenging to diagnose reliably, particularly when models trained solely on simulated data are applied to real-world measurements. Conventional methods, such as motor current signature analysis (MCSA) and classical machine learning (ML), are sensitive to load variations, rely on manual feature extraction, and exhibit poor performance under simulation-to-reality domain shift.

This paper proposes a convolutional neural network-long short-term memory (CNN-LSTM) transfer learning framework for BRB fault diagnosis. A dq0-based mathematical model is used to generate a comprehensive synthetic dataset covering healthy and faulty conditions across multiple fault severities and load levels. The simulated currents are subjected to empirical noise augmentation to better approximate experimental measurements. Classical ML classifiers trained on time-domain features derived from the Park vector modulus (PVM) achieve approximately 99% accuracy on synthetic data but deteriorate to nearly 55% when tested on experimental data, revealing a significant domain gap.

The proposed CNN-LSTM model is pre-trained on the synthetic data to learn the spatial and temporal fault signatures and subsequently fine-tuned using limited experimental data. This two-stage approach achieves accuracy and an F1-score exceeding 99% when tested on the experimental dataset from IEEE DataPort. The findings demonstrate that synthetic data pretraining combined with transfer learning can substantially narrow the simulation-to-reality gap under the tested conditions, while reducing the reliance on large volumes of labelled experimental data.

1. Introduction

Three-phase squirrel-cage induction motors underpin modern industrial systems, driving pumps, compressors, conveyors, and manufacturing machinery, while consuming a substantial share of global industrial electricity. Their robustness and low maintenance requirements make them economically critical, yet they remain susceptible to progressive electromechanical degradation that, if undetected, culminates in costly unplanned downtime [1].

Motor faults are broadly classified as bearing failures (~42%), stator winding faults (~28%), rotor faults (~8%), and other fault types (~22%) [2]. Broken rotor bar (BRB) faults are particularly critical because a single broken bar disrupts the symmetric current distribution in the rotor cage, induces a backward-rotating magnetic field, causes torque pulsations and localised overheating, and may propagate to adjacent bars if left undetected, ultimately leading to catastrophic motor failure.

A broken bar produces characteristic sideband harmonics in the stator current at frequencies given by

$$f_{\mathrm{BRB}} = (1 \pm 2ks)\, f_1, \quad k = 1, 2, 3, \ldots, \tag{1}$$

where s is the per-unit slip and f_1 is the supply frequency [3]. Motor current signature analysis (MCSA) relies on these sidebands but is fundamentally limited by load dependence. As slip approaches zero at light load, the sidebands merge into the dominant fundamental component and become indistinguishable, making MCSA unreliable [4].

Signal processing techniques have been proposed in the literature to overcome the stationarity assumption of the fast Fourier transform (FFT). The short-time Fourier transform (STFT) provides a two-dimensional time-frequency representation suitable for transient startup conditions but is constrained by an inherent tradeoff between time and frequency resolution [4]. The discrete wavelet transform (DWT) addresses this through a multi-resolution, variable-window decomposition; sub-band energy coefficients provide quantitative fault severity indicators, though their accuracy depends on the choice of mother wavelet, for which no universally optimal criterion exists [4]. The Hilbert–Huang transform (HHT), combining empirical mode decomposition (EMD) with Hilbert spectral analysis, adaptively decomposes the signal into intrinsic mode functions (IMFs) derived entirely from the data, making it well suited to the nonlinear transient dynamics of fault-induced current modulations [5].

Despite their individual strengths, all of these signal processing methods share a dependency on expert parameter tuning and degrade under variable load conditions or supply-side disturbances, motivating the transition to data-driven approaches. Classical machine learning (ML) classifiers, such as support vector machines (SVMs), k-nearest neighbours (k-NN), and decision trees, as well as ensemble methods such as random forest and bagged trees, have all been applied to BRB classification: Bouchiba and Kaddouri [6] report 87% SVM accuracy, while Quiroz et al. [7] achieve 98.8% accuracy with random forest on startup current features. However, these techniques fail to generalise across domains because manually engineered features capture domain-specific rather than domain-invariant fault signatures.

Deep learning architectures bypass manual feature engineering by learning hierarchical representations directly from raw signals [8]. These models are often prone to overfitting due to limited training data. Real-world fault data are expensive and difficult to collect, as inducing fault conditions risks permanent machine damage. Moreover, fault signatures vary across motor types, load levels, and supply conditions. Ali et al. [9] address this by training on a large dataset of 57 500 spectral images for BRB diagnosis, applying a transfer learning-based ShuffleNet V2 architecture to achieve 98.856% accuracy with low computational overhead, while also demonstrating the value of fusing current and vibration signals for robust, lightweight diagnostic systems.

Transfer learning addresses the data scarcity issue by leveraging knowledge from a data-rich source domain (simulation) to improve performance in a data-scarce target domain (experiment) [10]. Yosinski et al. [11] established the theoretical basis for transfer learning by showing that early network layers learn general, transferable features (local pattern detection, frequency filtering), while later layers encode task-specific decision boundaries that require adaptation to the target domain. Building on this, Han et al. [12] demonstrated deep transfer networks for industrial fault diagnosis, and Xiang et al. [13] showed that a classifier-constrained domain adaptation network using simulation-derived signatures as the source domain successfully transfers to experimental BRB data.

A key obstacle for data-driven approaches in industrial fault diagnosis is the simulation-to-reality domain gap. Recent work has attempted to bridge this gap by combining limited real-world measurements with simulated signals during training for eccentricity faults [14]. However, most published studies address only one domain, either synthetic or experimental, leaving this gap insufficiently explored for BRB faults.

This paper addresses the gap through the following contributions:

- A direct-quadrature-zero (dq0)-based MATLAB/Simulink induction motor model that generates labelled synthetic BRB data at varying severity (1–5 broken bars) and load conditions, with empirical noise augmentation.
- A hybrid convolutional neural network-long short-term memory (CNN-LSTM) architecture for fault severity classification from raw stator current segments.
- A transfer learning strategy in which convolutional layers are pre-trained on augmented synthetic data and LSTM layers are fine-tuned on experimental data, achieving an F1-score of over 99% on a public experimental benchmark while converging ~50% faster than training from scratch.
- A systematic comparison with classical ML classifiers under both same-domain and cross-domain evaluation conditions.

2. Mathematical modelling and fault simulation

The proposed fault diagnosis framework consists of three main stages, namely synthetic data generation from the dq0-based mathematical model, classical ML classification, and CNN-LSTM-based transfer learning. The overall workflow of the proposed approach is illustrated in Fig. 1.

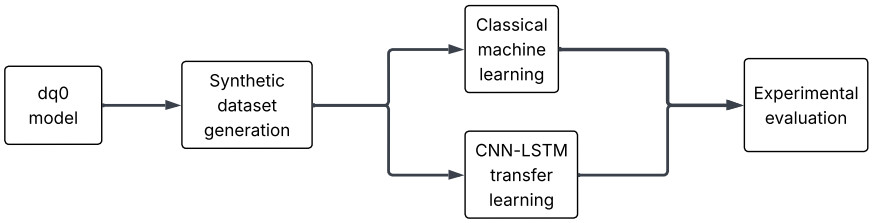


Fig. 1. Proposed workflow.

2.1. dq0-based mathematical model

The dq0 transformation projects the three-phase stator and rotor quantities onto a synchronously rotating reference frame, converting sinusoidal AC signals into constant DC quantities under steady-state operation. This eliminates the rotor-position-dependent inductances inherent in the abc-frame model. The dq-frame representation also forms the basis for the Park vector modulus (PVM), defined as the magnitude of the vector sum of d- and q-axis components, which is later used for feature extraction in ML. While the dq0 model is a simplification of the machine's behaviour, it retains the essential fault signatures necessary for diagnosis [15], making it suitable for generating large synthetic datasets for training ML models.

When the dq0 transformation is applied to the voltage and flux linkage equations of the stator and rotor in the abc frame, the following equations are obtained in the dq0 frame:

$$\mathbf{v}_{qd0s} = \mathbf{r}_s \mathbf{i}_{qd0s} + \omega \boldsymbol{\lambda}_{dqs} + p\boldsymbol{\lambda}_{qd0s}, \tag{2}$$

$$\mathbf{v}_{qd0r} = \mathbf{r}_r \mathbf{i}_{qd0r} + (\omega - \omega_r)\boldsymbol{\lambda}_{dqr} + p\boldsymbol{\lambda}_{qd0r}, \tag{3}$$

where

$$(\boldsymbol{\lambda}_{dqs})^T = \begin{bmatrix} \lambda_{ds} & -\lambda_{qs} & 0 \end{bmatrix}, \tag{4}$$

$$(\boldsymbol{\lambda}_{dqr})^T = \begin{bmatrix} \lambda_{dr} & -\lambda_{qr} & 0 \end{bmatrix}, \tag{5}$$

where the symbols $\mathbf{v}$, $\mathbf{i}$, $\boldsymbol{\lambda}$, $\mathbf{r}$, ω, and p represent the voltage, current, flux linkage, resistance, angular frequency of the stator supply, and the differential operator with respect to time, respectively. The subscripts s and r denote the stator and rotor quantities. This system of differential equations is the basis for the modelling of the electromagnetic system of a healthy induction motor. Further details on the derivation of the dq0 model and its application to induction motor analysis can be found in [16,17].

The equivalent circuit parameters of the simulated motor were obtained from a 1 hp, three-phase, four-pole induction machine through standard laboratory tests, including no-load, blocked-rotor, and DC resistance measurements, and are listed in Table 1. These parameters are representative of the 1 hp motor for which the experimental dataset was acquired.

2.2. Modelling of faulty induction motor

BRB faults are introduced by adding an asymmetric incremental resistance to the rotor circuit [18]. For n_{bb} broken bars out of N_b total rotor bars, the incremental resistance per rotor phase is

$$\Delta R = \frac{3\, n_{bb}}{N_b - 3\, n_{bb}}\, R_r. \tag{6}$$

In the three-phase abc frame, the rotor resistance matrix $\mathbf{r}_{r,abc}$ is diagonal but when transformed to the dq0 frame, it yields a position-dependent, non-diagonal rotor resistance matrix $\mathbf{r}_r$ in the dq0 reference frame, whose off-diagonal terms contain trigonometric functions of the rotor angle θ_r.

Table 1. Induction motor nameplate and equivalent circuit parameters

Nameplate parameters

Parameter	Value
Rated power	1 hp (746 W)
Rated voltage	220 V / 380 V
Rated current	3.02 A / 1.75 A
Nominal torque	4.1 N·m
Rated full load speed	1715 rpm
Supply frequency	60 Hz
Poles	4
No. of rotor bars	34

Simulation parameters

Parameter	Symbol	Value
Stator resistance	R_s	9.4236 Ω
Rotor resistance	R_r	9.4236 Ω
Stator leakage inductance	L_s	0.0553 H
Rotor leakage inductance	L_r	0.0553 H
Magnetising inductance	L_m	0.6050 H
Moment of inertia	J	0.01 kg·m²

These terms produce oscillatory components at the frequency of $2sf_1$ in the d- and q-axis currents, generating the characteristic PVM ripple and MCSA sidebands used as fault indicators. The matrix $\mathbf{r}_r$ is substituted into Eq. (3) for different fault severities and solved numerically to determine the corresponding d- and q-axis components of the stator current.

The modelling of the BRB fault in this work follows the approach described by Chen et al. [18], which has been validated against experimental data and is widely adopted in the literature for simulating BRB faults in induction motors. The system of equations in the dq0 reference frame, described above, is solved numerically in MATLAB/Simulink to generate synthetic current signals for healthy and faulty conditions across a range of load levels and fault severities.

2.3. Simulation and experimental datasets

The simulation runs are conducted for each combination of load level and fault severity. The load varies from no load to full load in increments of 20% of the rated load torque, while fault severity ranges from healthy (0 BRBs) to 5 BRBs. Each run lasts 60 seconds at a sampling frequency of 50 kHz. Startup transients are discarded and only steady-state signals are retained. The synthetic dataset totals 30 minutes of the three-phase stator current data and serves as the source domain for model training.

The experimental dataset [19] is acquired from a 1 hp, four-pole induction motor. Fault conditions range from healthy to four BRBs. Measurements are taken at eight load levels from 12.5% to 100% of full load, with 10 repeated acquisitions per condition. Each acquisition lasts 18 seconds at a sampling frequency of 50 kHz, covering both transient and steady-state operation. Startup transients are discarded for consistency with the simulation dataset. This dataset serves as the target domain for evaluating cross-domain generalisation.

For the classical ML experiments, PVM features extracted from the synthetic and experimental datasets are each partitioned using a window-level 90:10 stratified split into training and test sets. Same-domain evaluation is performed separately for both domains: synthetic data are used for training and testing within the source domain, and experimental data are used for training and testing within the target domain. Cross-domain evaluation trains exclusively on synthetic data and tests on the experimental test set. For the CNN-LSTM experiments, the synthetic dataset is used exclusively for pre-training. The experimental dataset is divided into a fine-tuning partition (training and validation) and a held-out test set, with temporal ordering preserved throughout to prevent leakage between adjacent windows.

2.4. Empirical noise augmentation

Simulation-generated data are clean and deterministic, while real-world measurements are contaminated with various sources of noise and interference. This domain shift poses a significant challenge for models trained solely on synthetic data when applied to experimental measurements. To mitigate this issue, a two-pronged preprocessing strategy is employed: empirical noise injection into the synthetic signals to approximate real-world sensor characteristics, and low-pass filtering of the experimental signals to suppress high-frequency components absent from the simulation model.

The synthetic data were augmented with noise extracted directly from the experimental stator current measurements under healthy conditions, preserving the true spectral characteristics of real-world sensor noise and yielding a comparable signal-to-noise ratio (SNR). The FFT was applied

to the healthy experimental current signal, and the fundamental frequency component at 60 Hz was removed. The remaining spectral components, representing measurement noise and harmonics, were then transformed back to the time domain via the inverse fast Fourier transform (IFFT). The extracted noise was superimposed onto the synthetic current signals at a controlled SNR of 31.4 dB, empirically determined to approximate the noise level observed in the experimental dataset.

This approach attempts to bring both distributions into closer alignment, reducing the effective domain gap and improving the transferability of models trained on synthetic data to real-world operating conditions.

3. Classical machine learning

ML-based fault diagnosis is reliant on the selection of features that faithfully represent the underlying fault characteristics. This work adopts the Park vector approach, wherein time-domain statistical features are extracted exclusively from the PVM defined as

$$\text{PVM} = \sqrt{i_d^2 + i_q^2}. \tag{7}$$

Under healthy operating conditions, a three-phase induction motor draws balanced stator currents which, upon application of the Park transformation in the synchronously rotating reference frame, yield nominally constant d- and q-axis components. Consequently, the PVM remains essentially constant, exhibiting only minimal ripple attributable to measurement noise. In the presence of BRB faults, the rotor asymmetry induces an oscillatory component in the dq-axis currents at a frequency equal to twice the slip frequency ($2sf_1$), which manifests as a characteristic periodic ripple in the PVM. The amplitude of this ripple serves as a direct and quantitative indicator of fault severity. This behaviour is illustrated in Fig. 2, which presents the PVM ripple of the experimental current data obtained from IEEE DataPort [19]. As observed, the ripple amplitude is negligible under healthy conditions and increases monotonically with the number of BRBs, confirming the suitability of PVM-derived features as discriminative inputs for fault severity classification.

Prior to feature extraction, the continuous time-domain signal is segmented into fixed-length windows of one second duration, corresponding to 50 000 data points at the adopted sampling rate of 50 kHz. The windows are non-overlapping to ensure independence of samples. The defined statistical features, which include mean, standard deviation, skewness, kurtosis, root mean square (RMS), peak-to-peak amplitude, energy, impulse factor, shape factor, crest factor, and coefficient of variation, are subsequently computed over each window independently, condensing the temporal information contained within 50 000 samples into a single feature vector per window.

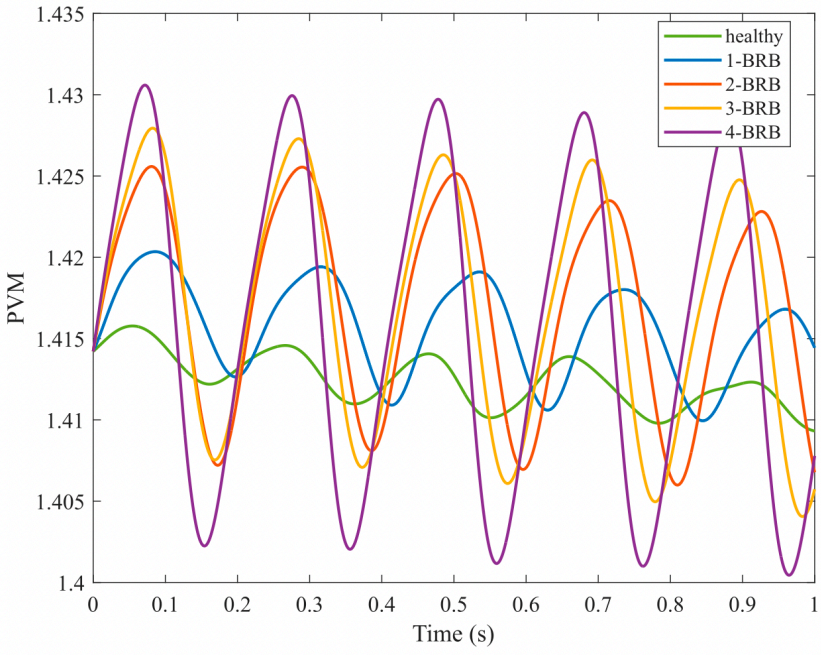


Fig. 2. PVM ripple for different fault severities at full load.

These features are selected based on their ability to capture the amplitude, variability, and shape characteristics of the PVM ripple associated with BRB faults. The magnitude of the PVM ripple increases with fault severity for a given load level, as additional broken bars intensify the rotor asymmetry and the associated current modulation. However, ripple magnitude also increases with load for a given fault severity, since higher loading increases the rotor current and amplifies the fault-induced sideband components. Consequently, amplitude-based features such as RMS and peak-to-peak value reflect both fault-related and load-dependent signal variations. Higher-order statistics such as skewness and kurtosis are less sensitive to amplitude scaling and provide insight into the asymmetry and peakedness of the ripple waveform, which are more directly indicative of fault-induced distortions. The impulse factor, shape factor, and crest factor characterise the waveform morphology of the PVM signal, capturing the sharpness and impulsiveness of the oscillatory components introduced by BRBs. The coefficient of variation normalises the standard deviation by the mean, providing a scale-invariant measure of variability useful for comparing signals across different load conditions.

The resulting feature vectors form the input dataset used for training and evaluation of the ML classification models. The feature vectors obtained across all loading conditions and fault severity levels (healthy, 1-BRB, 2-BRB, 3-BRB, 4-BRB, and 5-BRB) were partitioned into training and test sets in a 90:10 ratio. Stratified splitting was employed to ensure proportional representation of all fault categories in both subsets, preventing any class from being absent or underrepresented in the test set. Similar preprocessing and feature extraction steps were performed on the experimental dataset [19], which was used for cross-domain evaluation of the trained models. Model training and hyperparameter configuration were performed using the Classification Learner application in MATLAB. Three-fold cross-validation was applied during training to mitigate overfitting and obtain a reliable estimate of generalisation performance on unseen data.

The following classifiers were trained and evaluated: decision tree, SVM, *k*-NN, artificial neural network (ANN), as well as ensemble methods, including random forest, bagged trees, and boosted trees. All models were trained on the simulation-generated feature vectors and evaluated under two distinct testing conditions: same-domain validation, wherein the models were tested on held-out simulated data, and cross-domain evaluation, wherein the trained models were tested directly on unseen experimental data without any retraining or fine-tuning. The latter condition constitutes the primary measure of sim-to-real transfer capability and exposes the degree to which simulation-trained classifiers generalise to real-world motor measurements. The classification performance of all models under both evaluation conditions is presented and analysed in Section 5.

4. Deep learning framework

The proposed deep learning approach combines convolutional neural networks (CNNs) and long short-term memory (LSTM) networks in a hybrid architecture to automatically extract fault signatures from raw stator current signals and classify different severity levels of BRB faults. The proposed workflow is illustrated in Fig. 3.

4.1. CNN-LSTM model

The proposed framework is a hybrid deep learning architecture combining CNN and LSTM networks with transfer learning represented in Fig. 4 to automatically extract fault signatures from motor current signals and classify different severity levels of BRB faults. While CNNs excel at learning local spatial structures through convolutional operations with shared weights, LSTMs are specifically designed to capture long-range temporal dependencies through gating mechanisms that control information flow through time. This combination addresses the fundamental nature of motor fault diagnosis: faults manifest as time-evolving patterns in current signals that span multiple motor cycles. By combining these architectures, the model learns both the frequency domain characteristics of faults (through convolutions) and their temporal evolution (through LSTM recurrence). Additionally, transfer learning enables knowledge transfer from a well-labelled source domain to target domains with limited data, significantly reducing training requirements for new motor systems.

The classification problem is treated as supervised sequence learning under source–target domain shift. The inability of classical ML models to generalise to real-world measurements indicates a distribution mismatch between the source and target domains, although both domains describe the

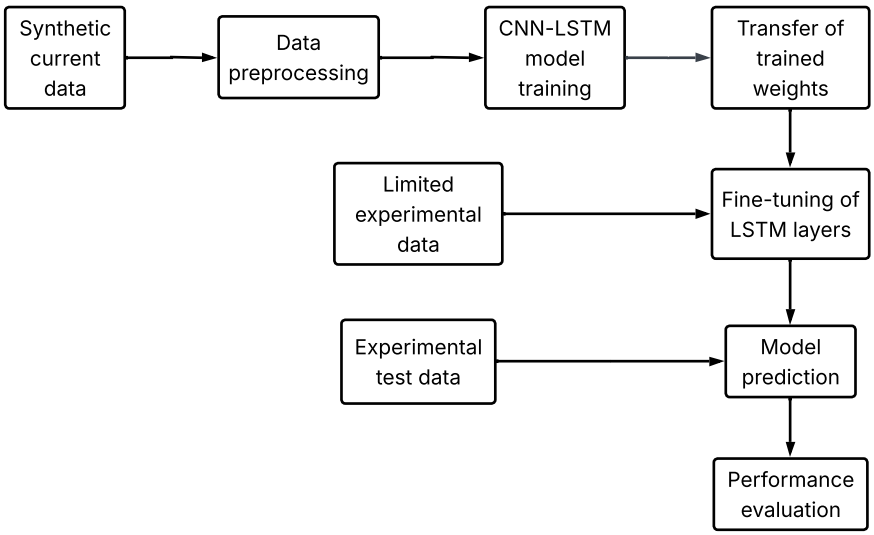


Fig. 3. Proposed deep learning workflow.

same BRB mechanism. Since conventional ML classifiers are strongly dependent on hand-crafted feature distributions, they are not sufficiently invariant to the mismatch between dq0-based simulation signals and measured stator currents. The proposed method therefore learns a discriminative representation directly from the current waveform, reducing dependence on manually selected statistical features.

Let $\mathbf{x} = [x_1, x_2, \ldots, x_T]^\top$ denote a phase current window with $T = 4000$. CNN layers extract local temporal and frequency-related patterns from $\mathbf{x}$, whereas LSTM layers model the sequential evolution of these features over multiple motor cycles [20]. Transfer learning adapts the source-domain representation to the experimental domain by preserving the low-level convolutional filters and updating the recurrent and classification layers [21]. The transfer learning model is illustrated in Fig. 4.

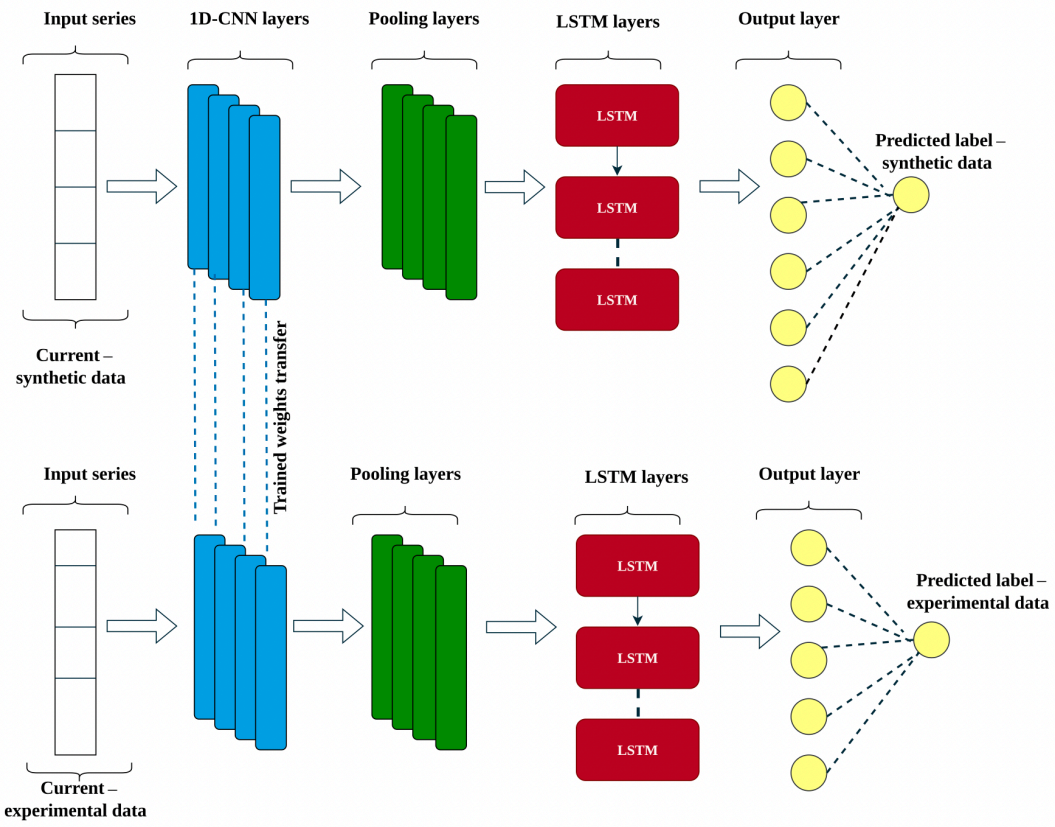


Fig. 4. CNN-LSTM transfer learning model.

4.2. Dataset preparation and preprocessing

The source-domain dataset consists of simulated three-phase stator current measurements for six classes: healthy, 1-BRB, 2-BRB, 3-BRB, 4-BRB, and 5-BRB. The signals are sampled at 50 kHz under no-load, 40%, 60%, 80%, and 100% load conditions. Phase-A current is used as the input channel because the broken-bar disturbance is reflected in the stator current waveform, while the temporal window preserves the modulation information needed by the CNN-LSTM model. The target domain is the experimental dataset, containing five classes: healthy, 1-BRB, 2-BRB, 3-BRB, and 4-BRB under 50%, 62.5%, 75%, and 87.5% load conditions. Source data are split into 82.5% training and 17.5% testing samples; target data are split into 88% fine-tuning and 12% validation samples. All splits preserve temporal ordering to avoid leakage between adjacent windows.

The current signal is segmented into fixed-length windows of 4000 samples (80 ms at 50 kHz). This window length balances three requirements: compatibility with the network's downsampling schedule ($4000 \rightarrow 800 \rightarrow 160 \rightarrow 80$ samples across stride and pooling layers), maximising the number of training sequences from limited experimental data, and exploiting the CNN-LSTM's ability to learn directly from raw time-domain signals without the frequency resolution constraints that necessitate longer windows in FFT-based methods. Fault-related temporal evolution is captured by the LSTM layers through sequential processing of consecutive windows. Training windows use 50% overlap to increase sample count and retain boundary information, while test windows use no overlap to avoid redundant evaluation [22]. For raw current $i_a[n]$, the kth window is $\mathbf{x}_k = [i_a[kH], i_a[kH+1], \ldots, i_a[kH+T-1]]^\top$, where $T = 4000$ and H is the hop size. Labels are encoded as 0–5 in simulation and 0–4 in experiment, where 0 represents the healthy condition and 1–4 represent the corresponding BRB fault classes. Z-score standardisation is applied using only training-set statistics:

$$\hat{x}_{k,t} = \frac{x_{k,t} - \mu_{\text{tr}}}{\sqrt{\sigma_{\text{tr}}^2 + \epsilon}}, \tag{8}$$

where μ_{tr} and σ_{tr}^2 are computed from the training current channel. This prevents test-set information from entering the training pipeline and improves convergence [23].

To reduce the gap between idealised simulation and real-world measurement data, noise augmentation is applied exclusively to the simulated training set. The SNR is set to 31.4 dB, determined empirically by measuring the SNR of the healthy experimental stator current signal at no-load conditions. A supply voltage unbalance of ±2% is adopted from the maximum permissible voltage unbalance specified in EN 50160 [24]. A gain mismatch of ±3% and a DC offset of ±0.114 A are adopted as representative values consistent with the analogue-to-digital converter (ADC) and sensor calibration errors reported in prior noise augmentation literature [8,25]. Together, these perturbations represent sensor noise, calibration error, offset drift, and non-ideal supply conditions, encouraging the model to learn features less sensitive to idealised simulation assumptions.

4.3. Architecture, training, and transfer learning

For each standardised input window, the first 1D convolutional layer uses 128 filters with kernel size 15, stride 5, and same padding, reducing the input sequence from 4000 to 800 samples. The second convolutional block uses 256 filters with kernel size 7 and stride 5, further reducing the sequence from 800 to 160 feature samples. With zero-padded input feature map $\tilde{a}_c[n]$, a 1D convolution output is computed as

$$y_j[n] = \phi\left(b_j + \sum_{c=1}^{C_{\text{in}}} \sum_{p=0}^{K-1} w_{j,c}[p]\, \tilde{a}_c[ns+p] \right), \tag{9}$$

where $w_{j,c}$ is the kernel connecting input channel c to the output map j, K is the kernel size, s is the stride, b_j is the bias, and $\phi(\cdot)$ is the nonlinear activation. Batch normalisation stabilises the feature distribution after convolution [26], max pooling of size 2 reduces the sequence from 160 to 80 samples, and layer normalisation stabilises recurrent inputs [27].

The recurrent section contains LSTM layers with 128, 64, and 32 units. The first two LSTM layers return full sequences, and the final LSTM layer returns a 32-dimensional hidden representation for classification. For training sample m, the dense output layer produces logits $\boldsymbol{z}^{(m)} = [z_1^{(m)}, \ldots, z_C^{(m)}]$.

The softmax function converts these logits into class probabilities:

$$p\left(y^{(m)} = c \mid \mathbf{x}^{(m)}\right) = \frac{\exp\left(z_c^{(m)}\right)}{\sum_{r=1}^{C} \exp\left(z_r^{(m)}\right)}, \quad c = 1, \ldots, C. \tag{10}$$

Here $\mathbf{x}^{(m)}$ is the mth input window, $y^{(m)}$ is its class label, c and r are the index classes, and C is the number of classes ($C = 6$ for source training and $C = 5$ after transfer). The model minimises the sparse categorical cross-entropy,

$$\mathcal{L} = -\frac{1}{N} \sum_{m=1}^{N} \log p\left(y^{(m)} \mid \mathbf{x}^{(m)}\right), \tag{11}$$

where N is the number of training samples. Optimisation uses the Adam optimiser with a learning rate of 0.00025 and a global gradient-clipping norm of 1.0 [28,29]. Training is performed for up to 70 epochs using ReduceLROnPlateau, which reduces the learning rate by 0.5 after five validation-loss plateau epochs.

The synthetic source dataset contains six classes (healthy and 1–5 BRBs), while the experimental target dataset contains five classes (healthy and 1–4 BRBs), reflecting the fault severity levels physically present in each dataset. To accommodate this discrepancy, the pre-trained six-class output layer is discarded and replaced with a randomly initialised dense layer sized for five target classes prior to fine-tuning.

For transfer learning, the convolutional and batch normalisation layers are frozen because they encode domain-invariant fault-related feature representations expected to remain useful across domains. The LSTM layers and the new output layer are fine-tuned on the experimental dataset, allowing the model to relearn task-specific temporal decision boundaries appropriate to the five-class target problem while retaining the general feature representations learned from simulation. This selective fine-tuning reduces trainable parameters, limits overfitting on the smaller target dataset, and preserves source-domain knowledge while correcting the simulation-to-reality mismatch [11,12]. The same pipeline supports the three evaluations reported later: source-domain testing, direct experimental transfer before fine-tuning, and experimental evaluation after fine-tuning.

5. Results and discussion

5.1. Classical machine learning

Classical ML models trained on PVM-derived statistical features achieved high classification accuracy of over 96% when trained and tested on the same experimental dataset, as shown in Table 2, confirming that the extracted features carry sufficient discriminative information for fault severity classification within a single domain. However, when these models were trained on simulated data and evaluated directly on the experimental dataset, accuracy degraded severely to the 45–55% range, demonstrating the critical sim-to-real domain gap. Real-world experimental signals contain measurement noise, sensor nonlinearity, and supply-side harmonics absent from idealised simulation, causing the statistical feature distributions to shift substantially. This indicates that the PVM statistical distributions also differ significantly between the clean simulation domain and the noisy experimental domain, rendering simulation-trained classical ML models unsuitable for deployment on real motors without extensive retraining.

The high same-domain classification accuracy of classical ML models requires some interpretation. The train/test split was performed after the windowing process, meaning adjacent non-overlapping windows from the same acquisition could appear in both the training and test sets. Since only steady-state signals are retained, windows from the same loading and fault condition are spectrally similar, which likely contributes to the high same-domain accuracy and may partly reflect run-specific signal characteristics.

Additionally, as discussed in Section 3, the extracted PVM features are sensitive to both fault severity and load level, and the two effects cannot be fully decoupled from time-domain statistics alone. The high same-domain accuracy therefore likely reflects a combination of fault-related, load-dependent, and run-specific factors rather than the purely fault-discriminative learning. This is acknowledged as a limitation of the same-domain evaluation.

Table 2. Performance of classical ML models under same-domain and cross-domain evaluation

Model	Trained on experimental data		Trained on simulation data	
	Validation[1] (%)	Test[2] (%)	Validation[1] (%)	Test[2] (%)
Bilayered neural network	97.59	99.34	97.02	52.80
Wide neural network	98.10	99.01	98.24	50.06
Ensemble (bagged trees)	98.57	98.68	98.77	55.46
Medium neural network	97.95	98.68	97.72	50.39
Fine *k*-NN	96.78	98.02	96.14	48.98
Fine decision tree	97.29	97.70	97.54	55.59
Trilayered neural network	97.15	95.72	98.77	55.59

[1] Mean three-fold cross-validation accuracy on the training domain
[2] Accuracy on the held-out experimental test set in both training conditions

These considerations do not affect the validity of the cross-domain results, which constitute the primary evaluation in this study. In the cross-domain setting, the exact load levels of the experimental dataset (12.5% increments) were not represented in the simulated training data (20% increments), as described in Section 2.3, meaning that the models were evaluated under genuinely unseen operating conditions.

The substantial degradation in cross-domain accuracy therefore reflects a true generalisation failure and highlights a fundamental limitation: hand-crafted PVM features capture domain-specific signal characteristics rather than the domain-invariant fault physics, rendering simulation-trained classical ML models undeployable on real motors without extensive retraining. The poor generalisation of classical ML models across domains motivates the adoption of deep learning architectures that can learn hierarchical representations directly from raw signals, potentially capturing more robust and transferable features for sim-to-real fault diagnosis.

5.2. CNN-LSTM transfer learning

The proposed CNN-LSTM transfer learning framework was evaluated in three stages: source-domain training using simulated current signals, direct cross-domain testing on experimental data, and fine-tuning on the experimental dataset. The simulated dataset contained six operating states, while the experimental dataset contained five states. This class count discrepancy intentionally tests the model's ability to adapt learned features to new motor systems without extensive retraining.

The CNN-LSTM model trained on synthetically augmented simulation data achieved high classification performance across 70 training epochs, as shown by the source-domain accuracy curve in Fig. 5a. On the simulation training set, the model exhibited excellent performance, obtaining accuracy, precision, recall, and an F1-score exceeding 99% , as shown in Table 3.

The source-domain training confusion matrix is presented in Fig. 6a. Misclassifications in the simulation domain were concentrated between adjacent rotor bar fault severity levels, indicating an adjacency-dominant error pattern that is physically meaningful because consecutive fault stages exhibit overlapping stator current signatures due to progressive degradation. This behaviour is verified by the source-domain confusion matrix in Fig. 6a, where the dominant off-diagonal errors occurred between neighbouring severity levels (e.g. healthy/1-BRB, 1-BRB/2-BRB, and so on). Figure 6b presents the confusion matrix for validation on simulated data, which does not show such adjacency-dominant misclassifications, indicating that the model generalises well to unseen simulated samples. The validation accuracy of 99.9439% confirms that the model has learned a robust representation of the fault signatures present in the simulation domain. The precision, recall, and F1-score metrics also exceed 99%, indicating that the model is both accurate and reliable in its predictions across all fault classes, as documented in Table 3.

The transfer-learning accuracy curve in Fig. 5b shows rapid convergence on the experimental dataset. The transferred model reached a stable region after approximately 25–30 epochs, whereas source-domain training on simulated data required approximately 60–70 epochs (Fig. 5a). This reduction in training effort is important for practical fault diagnosis, since collecting large labelled datasets for every new motor is time-consuming and expensive. The result shows that simulation can be used to learn general fault representations, while a smaller experimental dataset can be used to adapt the model to a specific machine and measurement setup.

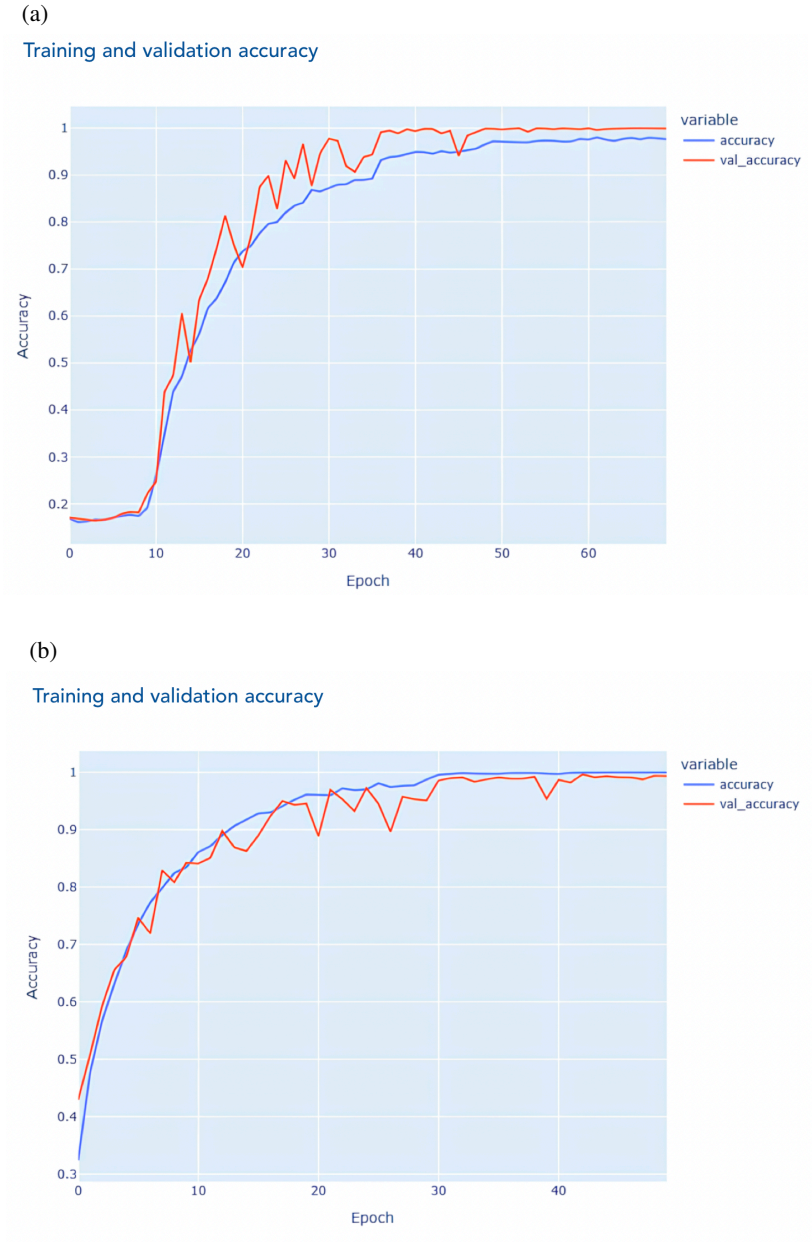


Fig. 5. Accuracy convergence for source-domain training and experimental fine-tuning: (a) source-domain training accuracy, (b) transfer-learning accuracy.

Table 3. CNN-LSTM classification performance

Evaluation condition	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
Simulation training data	99.1579	99.1591	99.1579	99.1579
Simulation validation data	99.9439	99.9441	99.9439	99.9439
Experimental test before fine-tuning	18.18	9.78	18.48	9.65
Experimental test after fine-tuning	99.3478	99.3489	99.3478	99.3475

When the simulation-trained model was evaluated on the experimental dataset without fine-tuning, accuracy collapsed to 18.18%, with an F1-score of 9.65%. This severe degradation is attributable to the sim-to-real domain gap, inter-bar currents, manufacturing tolerances, and real-world noise characteristics present in experimental data, which could not be replicated in simulation, rendering the learned decision boundaries ineffective. Figure 7a shows that before transfer learning, the model collapses to predicting 4-BRB almost exclusively, indicating that the simulated feature distribution maps onto the 4-BRB region of the experimental feature space.

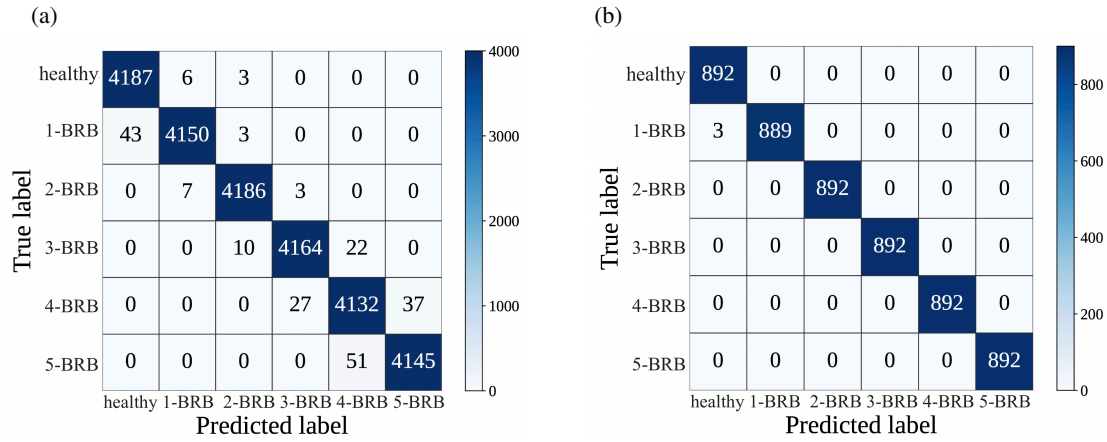


Fig. 6. Confusion matrices for CNN-LSTM training and validation on simulated data: (a) training on simulated data, (b) validation on held-out simulated data.

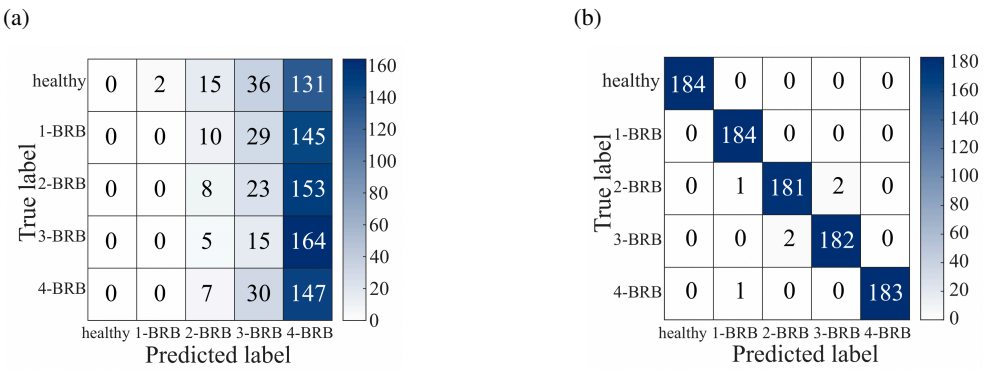


Fig. 7. Confusion matrices before and after transfer learning, evaluated on experimental test data: (a) before transfer learning, (b) after transfer learning.

After transfer learning, the experimental test performance increased sharply to 99.3478% accuracy, 99.3489% precision, 99.3478% recall, and 99.3475% F1-score, as listed in Table 3. The class-wise performance metrics, including precision, recall, F1-score, and support, are listed in Table 4. The convolutional and batch normalisation layers were kept frozen, while the LSTM layers and final classification layer were fine-tuned for the experimental classes. This selective fine-tuning strategy retained source-domain knowledge in frozen layers, while adaptation occurred only in necessary components. The confusion matrix in Fig. 7b shows that the fine-tuned model preserved high class-wise discrimination on unseen experimental samples.

The t-SNE (t-distributed stochastic neighbour embedding) visualisation in Fig. 8 illustrates the learned feature space for the experimental validation data after transfer learning. The projected clusters suggest that the fine-tuned CNN-LSTM model organises features according to fault severity, with adjacent clusters corresponding to neighbouring BRB severity levels, which is consistent with the adjacency-dominant misclassification pattern observed in the confusion matrix. While t-SNE provides a qualitative indication of feature space structure, the classification accuracy and F1-score reported in Table 3 serve as the primary quantitative evidence of successful domain transfer.

Table 4. Class-wise classification metrics – CNN-LSTM after fine-tuning on experimental dataset

Class	Precision (%)	Recall (%)	F1-score (%)	Support
Healthy	100.00	100.00	100.00	184
1-BRB	98.92	100.00	99.46	184
2-BRB	98.91	98.37	98.64	184
3-BRB	98.91	98.91	98.91	184
4-BRB	100.00	99.46	99.73	184

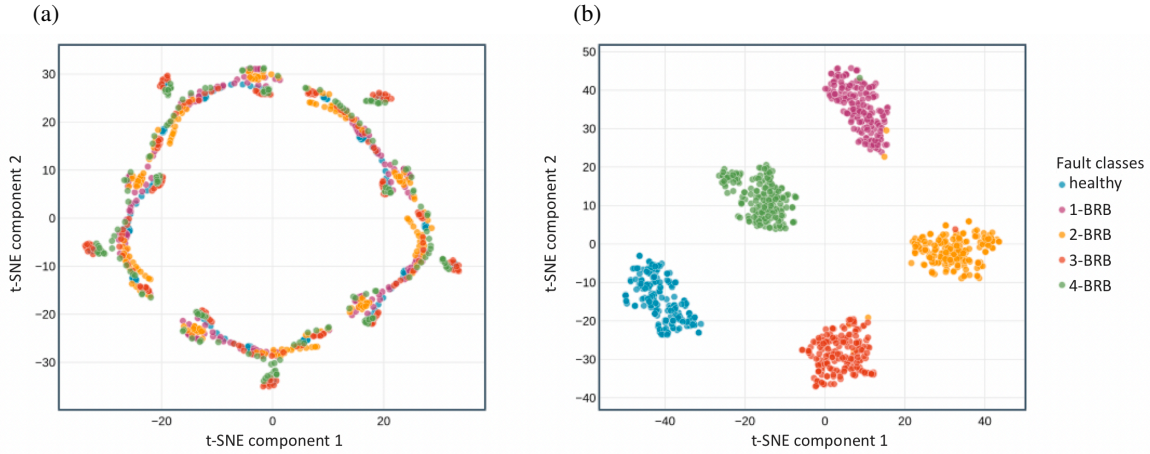


Fig. 8. t-SNE feature space visualisation for experimental validation data after transfer learning: (a) raw input features, (b) CNN-LSTM learned features.

Overall, the two-stage transfer learning approach demonstrated strong performance for BRB detection under the evaluated conditions, as summarised in Table 3. Simulation pre-training produced a discriminative representation of fault severity, direct transfer exposed the spectral mismatch between simulated and experimental current signals, and fine-tuning corrected this mismatch with high classification performance. The final experimental accuracy of 99.35% and an F1-score of 99.35% confirm that the proposed approach can classify rotor bar fault severity under real measurement conditions while reducing the amount of target-domain data required.

These results support the effectiveness of the proposed CNN-LSTM architecture and transfer learning strategy in narrowing the simulation-to-reality gap for the evaluated dataset and operating conditions. The findings suggest that synthetic pre-training followed by limited fine-tuning is a promising strategy for induction motor fault diagnosis in settings where labelled experimental data are scarce, though broader validation across different motor ratings and operating environments remains a natural direction for future work.

6. Conclusion

This paper presented a CNN-LSTM-based transfer learning framework for broken rotor bar fault diagnosis in three-phase induction motors, addressing the simulation-to-reality gap that limits the practical deployment of simulation-trained diagnostic models. Classical machine learning classifiers trained on Park vector modulus features achieved high accuracy on synthetic data but degraded substantially when evaluated on experimental measurements, confirming the domain gap as the primary obstacle to real-world generalisation. The proposed CNN-LSTM model, pre-trained on the augmented synthetic dataset and fine-tuned on experimental data, achieved classification accuracy and an F1-score exceeding 99%, demonstrating that two-stage transfer learning narrows this gap for the evaluated fault conditions without requiring large volumes of labelled experimental data.

These results suggest that combining physics-based simulation with targeted domain adaptation offers a viable pathway for data-driven fault diagnosis in settings where experimental fault data are costly or difficult to obtain. It should be noted that these findings are specific to the evaluated benchmark dataset and operating conditions, and generalisation to other motor ratings, supply conditions, or fault types remains a subject for future investigation. Future work can extend this framework to additional fault types, including bearing and stator winding defects, and evaluate its robustness across broader operating conditions and machine ratings. Further investigation into lighter-weight model architectures may also support deployment on embedded diagnostic hardware for real-time condition monitoring.

Data availability statement

The experimental dataset used in this study is publicly available via IEEE DataPort [19]. The synthetic dataset was generated using the dq0-based motor and fault model described in Section 2 and can be reproduced following by the methodology detailed therein.

Acknowledgement

The publication costs of this article were covered by the Estonian Academy of Sciences.

References

1. Bellini, A., Filippetti, F., Tassoni, C. and Capolino, G.-A. Advances in diagnostic techniques for induction machines. *IEEE Transactions on Industrial Electronics*, 2008, **55**(12), 4109–4126. https://doi.org/10.1109/TIE.2008.2007527
2. Hamani, K., Kuchar, M., Kubatko, M. and Kirschner, S. Advancements in induction motor fault diagnosis and condition monitoring: a comprehensive review. *Sensors*, 2025, **25**(19), 5942. https://doi.org/10.3390/s25195942
3. Thomson, W. T. and Fenger, M. Current signature analysis to detect induction motor faults. *IEEE Industry Applications Magazine*, 2001, **7**(4), 26–34. https://doi.org/10.1109/2943.930988
4. Hassan, O. E., Amer, M., Abdelsalam, A. K. and Williams, B. W. Induction motor broken rotor bar fault detection techniques based on fault signature analysis – a review. *IET Electric Power Applications*, 2018, **12**(7), 895–907. https://doi.org/10.1049/iet-epa.2018.0054
5. Antonino-Daviu, J., Riera-Guasp, M., Pineda-Sanchez, M. and Perez, R. B. A critical comparison between DWT and Hilbert–Huang-based methods for the diagnosis of rotor bar failures in induction machines. *IEEE Transactions on Industry Applications*, 2009, **45**(5), 1794–1803. https://doi.org/10.1109/TIA.2009.2027558
6. Bouchiba, N. and Kaddouri, A. Fault detection and localization based on decision tree and support vector machine algorithms in electrical power transmission network. In *2022 2nd International Conference on Advanced Electrical Engineering (ICAEE), Constantine, Algeria*. IEEE, 2022, 1–6. https://doi.org/10.1109/ICAEE53772.2022.9961970
7. Quiroz, J. C., Mariun, N., Mehrjou, M. R., Izadi, M., Misron, N. and Mohd Radzi, M. A. Fault detection of broken rotor bar in LS-PMSM using random forests. *Measurement*, 2018, **116**, 273–280. https://doi.org/10.1016/j.measurement.2017.11.004
8. Zhang, W., Peng, G., Li, C., Chen, Y. and Zhang, Z. A new deep learning model for fault diagnosis with good anti-noise and domain adaptation ability on raw vibration signals. *Sensors*, 2017, **17**(2), 425. https://doi.org/10.3390/s17020425
9. Ali, U. A multimodal lightweight approach to fault diagnosis of induction motors in high-dimensional dataset. arXiv preprint, 2025. https://doi.org/10.48550/arXiv.2501.03746
10. Kumar, P. Transfer learning for induction motor health monitoring: a brief review. *Energies*, 2025, **18**(14), 3823. https://doi.org/10.3390/en18143823
11. Yosinski, J., Clune, J., Bengio, Y. and Lipson, H. How transferable are features in deep neural networks? arXiv preprint, 2014. https://doi.org/10.48550/arXiv.1411.1792
12. Han, T., Liu, C., Yang, W. and Jiang, D. Deep transfer network with joint distribution adaptation: a new intelligent fault diagnosis framework for industry application. *ISA Transactions*, 2020, **97**, 269–281. https://doi.org/10.1016/j.isatra.2019.08.012
13. Xiang, L., Zhang, X., Zhang, Y., Hu, A. and Bing, H. A novel method for rotor fault diagnosis based on deep transfer learning with simulated samples. *Measurement*, 2023, **207**, 112350. https://doi.org/10.1016/j.measurement.2022.112350
14. Ji, D.-Y., Wang, B., Inoue, H. and Kanemaru, M. Simulation-to-reality domain adaptation for motor fault detection. In *2025 IEEE International Symposium on Diagnostics for Electric Machines, Power Electronics and Drives (SDEMPED), Dallas, TX, USA*. IEEE, 2025, 1–7. https://doi.org/10.1109/SDEMPED53223.2025.11154252
15. Konuhova, M. Induction motor dynamics regimes: a comprehensive study of mathematical models and validation. *Applied Sciences*, 2025, **15**(3), 1527. https://doi.org/10.3390/app15031527
16. Krause, P. C. Method of multiple reference frames applied to the analysis of symmetrical induction machinery. *IEEE Transactions on Power Apparatus and Systems*, 1968, **PAS-87**(1), 218–227. https://doi.org/10.1109/TPAS.1968.291992
17. Krause, P. C., Wasynczuk, O., Sudhoff, S. D. and Pekarek, S. D. *Analysis of Electric Machinery and Drive Systems*. 4th ed. Wiley-IEEE Press, Hoboken, NJ, 2025. https://doi.org/10.1002/9781394293896
18. Chen, S. and Živanović, R. Modelling and simulation of stator and rotor fault conditions in induction machines for testing fault diagnostic techniques. *European Transactions on Electrical Power*, 2010, **20**(5), 611–629. https://doi.org/10.1002/etep.342
19. Treml, A. E., Flauzino, R. A., Suetake, M. and Maciejewski, N. A. R. Experimental database for detecting and diagnosing rotor broken bar in a three-phase induction motor. IEEE DataPort, 2020. https://doi.org/10.21227/fmnm-bn95
20. Hochreiter, S. and Schmidhuber, J. Long short-term memory. *Neural Computation*, 1997, **9**(8), 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735
21. Pan, S. J. and Yang, Q. A survey on transfer learning. *IEEE Transactions on Knowledge and Data Engineering*, 2010, **22**(10), 1345–1359. https://doi.org/10.1109/TKDE.2009.191
22. Squicciarini, A., Valero Toranzo, E. and Zarzo, A. A time-series feature-extraction methodology based on multiscale overlapping windows, adaptive KDE, and continuous entropic and information functionals. *Mathematics*, 2024, **12**(15), 2396. https://doi.org/10.3390/math12152396
23. Bin James, H. H., Nurfadilah Binti Jaini, S., Bin Ismail, M. A. and Wai Heng, C. Influence of data pre-processing on CNN-LSTM model for regression-based analysis. In *2025 IEEE International Conference on Artificial Intelligence in Engineering and Technology (IICAIET), Kota Kinabalu, Malaysia*. IEEE, 2025, 621–625. https://doi.org/10.1109/IICAIET67254.2025.11265123
24. CENELEC - EN 50160:2022. *Voltage characteristics of electricity supplied by public distribution systems*. European Committee for Electrotechnical Standardization, Brussels, 2022.
25. Ince, T., Kiranyaz, S., Eren, L., Askar, M. and Gabbouj, M. Real-time motor fault detection by 1-D convolutional neural networks. *IEEE Transactions on Industrial Electronics*, 2016, **63**(11), 7067–7075. https://doi.org/10.1109/TIE.2016.2582729
26. Ioffe, S. and Szegedy, C. Batch normalization: accelerating deep network training by reducing internal covariate shift. *Proceedings of Machine Learning Research*, 2015, **37**, 448–456. https://proceedings.mlr.press/v37/ioffe15.html
27. Ba, J. L., Kiros, J. R. and Hinton, G. E. Layer normalization. arXiv preprint, 2016. https://doi.org/10.48550/arXiv.1607.06450
28. Kingma, D. P. and Ba, J. Adam: a method for stochastic optimization. arXiv preprint, 2014. https://doi.org/10.48550/arXiv.1412.6980
29. Pascanu, R., Mikolov, T. and Bengio, Y. On the difficulty of training recurrent neural networks. *Proceedings of Machine Learning Research*, 2013, **28**(3), 1310–1318. https://proceedings.mlr.press/v28/pascanu13.html

Asünkroonmootori rootorivarraste rikete diagnoosimine CNN-LSTM-siirdeõppe abil

Aalok Aryal, Aashirbad Bhandari, Arun Gnawali, Ankeet Gadtaula, Mehroz Fatima ja Bishal Silwal

Kolmefaasilisi asünkroonmootoreid kasutatakse tööstuses väga laialdaselt ja nende töökindlus mõjutab otseselt tootmisprotsesside toimimist. Üks võimalik rike on rootorivarraste purunemine, mis võib põhjustada vibratsiooni, ülekuumenemist ja täiendavaid kahjustusi. Seetõttu on oluline avastada sellised rikked võimalikult vara. Artiklis uuritakse, kuidas kasutada selleks tehisintellekti ja kuidas kohandada simulatsioonide põhjal treenitud diagnostikamudelit tegeliku elektrimootori mõõteandmetega. Uuringus loodi esmalt asünkroonmootori matemaatiline mudel, mille abil simuleeriti nii korras mootori tööd kui ka selle tööd ühe kuni viie purunenud rootorivardaga erinevatel koormustel. Simuleeritud voolusignaalidele lisati tegelikest mõõtmistest saadud müra, et muuta need tegeliku mootori signaalidega sarnasemaks. Seejärel võrreldi klassikalisi masinõppemeetodeid süvaõppel põhineva CNN-LSTM (*convolutional neural network-long short-term memory*)-mudeliga. Klassikalised meetodid töötasid hästi siis, kui mudelit treeniti ja katsetati sama tüüpi andmetega, kuid simulatsiooniandmetega treenitud mudeli täpsus langes tegelike mõõteandmete puhul ligikaudu 45–55%ni. See näitab selgelt erinevust ideaalse simuleeritud ja tegeliku mootori mõõtesignaalide vahel. Probleemi lahendamiseks kasutati siirdeõpet. CNN-LSTM-mudelit treeniti esmalt suure hulga simuleeritud andmetega ja seejärel kohandati seda väiksema hulga eksperimentaalsete andmete abil. Pärast kohandamist saavutas mudel tegelike mõõteandmete põhjal rootorivarraste rikete ja nende raskusastme määramisel umbes 99,35% täpsuse. Tulemused näitavad, et füüsikalistel mudelitel põhinevate simulatsioonide ja tehisintellekti kombineerimine võimaldab vähendada vajadust mahukate ja kallite eksperimentaalsete rikkeandmete järele. Selline lähenemine võib tulevikus lihtsustada elektrimasinate seisundiseire ja ennetava hoolduse süsteemide arendamist. Edasistes uuringutes tuleb meetodit katsetada erineva võimsusega mootoritel, mitmekesisemates töötingimustes ja teiste rikkeliikide korral.