IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 29 April 2026, accepted 3 June 2026, date of publication 17 June 2026, date of current version 24 June 2026.

Digital Object Identifier 10.1109/ACCESS.2026.3704536

RESEARCH ARTICLE

Machine Learning-Based System for Diagnosing Broken Bars in Three-Phase Induction Motors With Reduced Electrical Sensing

LEONARDO BATISTA DE SANTANA[1], NARCO AFONSO RAVAZZOLI MACIEJEWSKI[2], HEITOR SILVÉRIO LOPES[3], AND JOSÉ JAIR ALVES MENDES JUNIOR[3]

[1]Graduate Program in Electrical and Computer Engineering (CPGEI), Universidade Tecnológica Federal do Paraná (UTFPR), Curitiba, Paraná 80230-901, Brazil
[2]Academic Department of Electrotechnics (DAELT), Graduate Program in Energy Systems (PPGSE), Universidade Tecnológica Federal do Paraná (UTFPR), Curitiba, Paraná 80230-901, Brazil
[3]Academic Department of Electronics (DAELN), Graduate Program in Electrical and Computer Engineering (CPGEI), Universidade Tecnológica Federal do Paraná (UTFPR), Curitiba, Paraná 80230-901, Brazil

Corresponding author: Leonardo Batista de Santana (leonardosantana@alunos.utfpr.edu.br)

This work was supported by the Coordination for the Improvement of Higher Education Personnel (CAPES) for the Article Processing Charge under Grant ROR identifier: 00x0ma614.

ABSTRACT Integrating Artificial Intelligence (AI) into digitalization initiatives is vital for optimizing complex processes such as industrial maintenance. This paper develops a robust cost-effective methodology for diagnosing Broken Rotor Bar (BRB) defects in three-phase induction motors (IM). We address known limitations of traditional methods, such as Motor Current Signature Analysis (MCSA), in low-load and noisy environments. We introduce a data-driven pipeline based on reduced electrical sensing. The process begins with the acquisition of experimental data (current and voltage signals under various defect severities) and signal pre-processing. Next, a comprehensive feature extraction is performed using the Time Series Feature Extraction Library (TSFEL) library, followed by a feature selection stage using the Fast Correlation-Based Filter (FCBF) algorithm. This streamlined input feeds a supervised multiclass classification framework that includes Decision Trees (DT), Random Forest (RF), Support Vector Machine (SVM), and Multilayer Perceptron (MLP). Our findings demonstrate the practicality of monitoring only the three-phase stator currents, as the RF classifier achieved 95% accuracy and 94.92% F1-score, outperforming the MLP (87.5% accuracy and 97.38% F1-score) and DT (90% accuracy and 89.88% F1-score) models, with a feature selection relevance criterion of 110%. Furthermore, the proposed approach proved to be computationally efficient, with the RF model requiring only 0.029 MB of memory and exhibiting an inference latency of less than 4 ms. Crucially, this AI-driven method significantly outperforms MCSA, which failed to detect defects under low-load conditions (0.5 Nm and 1 Nm). This work validates an economically viable, highly efficient solution for predictive maintenance that aligns with Industry 4.0 by achieving high accuracy while minimizing sensory input and computational overhead.

INDEX TERMS Condition monitoring, defect detection, defect classification, machine learning, maintenance engineering, intelligent systems, three-phase induction motor.

I. INTRODUCTION

The accelerating adoption of Industry 4.0 paradigms is driving industrial digitalization and, consequently, increasing the demand for advanced solutions that optimize processes and enable predictive maintenance [1], [2]. Electric motors are crucial for the industrial sector, accounting for up to 80% of total energy consumption [3]. Therefore, improving the efficiency, reliability, and operational availability of these machines offers substantial economic advantages.

The associate editor coordinating the review of this manuscript and approving it for publication was Shih-Wei Lin.

Three-phase induction motors (IM), in particular, are widely employed. IMs are essential because they combine low cost with efficient conversion of electrical energy to mechanical energy [4]. However, IMs are susceptible to various faults, with Broken Rotor Bar (BRB) defects being among the most critical. If not detected, these defects can rapidly escalate into catastrophic failures. Such malfunctions, including voltage imbalances, increased torque pulsations, decreased efficiency, and excessive heat, can severely impair motor operation and lead to unexpected downtime [5], [6].

When a rotor bar breaks, it stops conducting current, and the motor loses magnetic flux in the area around the damaged rotor bar. This creates an asymmetry in the rotor's magnetic field, which induces a reverse rotating field at the rotor slip frequency. These physical effects superimpose fault-specific harmonic components in the stator current, which form the basis for diagnostic techniques such as Motor Current Signature Analysis (MCSA) [7], [8]. However, factors such as external noise, low motor shaft torque, varying defect severities, multiple simultaneous defects, or insufficient data-acquisition sampling rates reduce the efficacy of MCSA. These conditions cause diagnostic sidebands to merge with higher energy spectral components, which leads to unreliable diagnoses [5], [9]. Similarly, noise and external interference often challenge vibration analysis, another common IM monitoring method, requiring rigorous data pre-processing and filtering techniques [5], [9]. Most studies in this domain, including the research of [10] and [11], focus on analyzing steady-state signals, but researchers should explore more resource-efficient monitoring strategies.

Detecting these defects before they evolve into faults is critical, as unplanned shutdowns lead to significant financial losses, reduced productivity, and increased safety risks [1]. Thus, early identification enables proactive interventions, lowers maintenance costs, and supports sustainable industrial practices. As industrial systems evolve towards smart connected infrastructures enabled by Cyber-Physical Systems (CPS) and Digital Twins, researchers focus on real-time monitoring of electrical parameters — specifically current and voltage — for fault diagnosis [12]. These strategies support intelligent, data-driven decision-making and reduce reliance on complex, costly diagnostic tools, such as vibration analysis or infrared thermography. However, a critical challenge remains: the high cost and complexity associated with multi-channel Data Acquisition Systems (DAQ), which often require synchronized sensors for all electrical phases, hindering large-scale industrial deployment.

Consistent with the principles of Industry 4.0 and the growing digitalization, there is a clear shift toward Artificial Intelligence (AI)-driven methods for real-time defect detection and diagnosis. For example, advanced deep learning techniques, such as Convolutional Neural Networks (CNN)s combined with Long Short-Term Memory (LSTM) layers, have demonstrated improved precision in both defect prediction and estimating the remaining operational life of electrical machinery [13], [14], [15]. Furthermore, various Machine Learning (ML) algorithms, including neural networks and k-Nearest Neighbors (kNN), have proven to be highly effective in accurately categorizing bearing and other mechanical defects [16], [17], [18], [19]. While these methods achieve high accuracy, they often operate as "black boxes" with high computational overhead, lacking a clear justification for the necessity of each input sensor.

To further enhance the accuracy and efficiency of defect detection in complex systems, the integration of optimization strategies, such as GA and PSO, has also been proposed [20]. Currently, the use of IoT-enabled monitoring systems is being actively investigated for their ability to collect and analyze data in real time, thereby significantly improving responsiveness to defects and overall predictive maintenance [21].

Although there are numerous techniques for the detection and diagnosis of IM defects [9], [22], [23], current research often presents two key limitations: either a lack of comprehensive performance evaluations in various operating conditions and severity of defects, or focusing on a restricted set of diagnostic methods. Furthermore, conventional monitoring approaches that rely on electrical signals (such as phase currents and voltages) typically require data acquisition across all phases of the IM. This comprehensive sensing scheme inherently increases the costs associated with detection, installation, and signal processing. To address these limitations, this paper investigates the feasibility of reduced-phase monitoring. The core novelty of this work lies in the systematic identification of the minimum sensory threshold required for reliable diagnosis. Specifically, we explore reducing the required electrical sensing inputs to only phase currents or to a minimal subset of currents and voltages. By utilizing the Fast Correlation-Based Filter (FCBF) to iteratively evaluate feature relevance, we establish a formal boundary where voltage sensors become redundant, an aspect seldom quantified in prior literature. The objective is to simultaneously reduce the complexity and expense of the system while maintaining high diagnostic accuracy. This minimal monitoring hypothesis is critical for enhancing the cost-effectiveness and scalability of predictive maintenance, thereby accelerating its industrial adoption. Based on this premise, the present work focuses exclusively on diagnosing the severity of BRB defects in three-phase IMs, under varying shaft torques. Our methodological process integrates signal processing and ML methods to analyze the predictive capacity of different signal configurations for BRB diagnosis. The research uses steady-state signals acquired from an experimental test bench, enabling the controlled variation of key operational parameters, namely motor shaft load and BRB defect severity.

The main research gaps addressed in this paper and the distinctive contributions are centered on the systematic validation of a minimal sensing paradigm. While recent literature has explored complex architectures such as 1D CNNs and

EfficientNet to achieve high accuracy on standard datasets, these approaches often rely on high-dimensional features and significant computational resources, limiting their adoption in cost-sensitive edge devices. This paper advances the state of the art by achieving comparable diagnostic accuracy using exclusively stator current signals, even under critical low-load conditions (0.5 N.m and 4.0 N.m) where traditional MCSA is historically ineffective due to spectral overlap at low slip [8]. Furthermore, this study provides an engineering-oriented validation of the pipeline's efficiency, quantifying a minimal memory footprint (0.029 MB) and ultra-low latency (4 ms), which distinguishes this work as a viable solution for real-time Edge Computing deployment. By iteratively varying the relevance criterion of selected features, this study provides a formal quantification of the predictive power of each electrical variable, establishing the three-phase current monitoring as the optimal balance between diagnostic reliability and installation cost. In contrast to the non-interpretability of deep learning models, the proposed pipeline offers higher interpretability and lower inference latency, providing an empirically validated and resource-efficient configuration for practical industrial deployment.

This paper is structured as follows: Section II presents a literature review, emphasizing defect analysis in electrical machines and the application of machine learning methods. Section III details the materials and methods employed. Section IV discusses the results, focusing on each component of the proposed methodology. Section V concludes the paper and outlines future research directions.

II. RELATED WORKS

Integrating AI into information systems is crucial for industrial digitalization, significantly increasing competitiveness through its adaptability and capacity to manage complex processes [24]. In this context, maintenance plays a vital role in optimizing production workflows and minimizing operating costs [1]. Rotary actuators, particularly three-phase IM, are often the focus due to their widespread industrial use and inherent susceptibility to faults [25].

IM defects are broadly categorized as electrical or mechanical. Among electrical defects, BRBs stand out because of their progressive nature, which can rapidly lead to catastrophic failures if not detected early. Although bearing defects are quantitatively more frequent, accounting for over 40% of all IM defects [9], [22], research on BRB is highly prioritized given its severe and increasing impact.

Historically, condition monitoring has been based on methods such as vibration analysis and temperature monitoring. However, vibration analysis has limitations in detecting electrical defects, particularly in high-powered machines, where it may be less sensitive or economically unfeasible [26]. While recent studies attempt to fuse vibration and acoustic data, the high cost of specialized sensors and the requirement for complex mounting remain significant barriers for cost-sensitive industrial plants. With the advance of digital signal processing methods, current monitoring via MCSA has emerged as a non-invasive, low-cost, and effective diagnostic method. This is because internal defects manifest as unique, environment-dependent patterns within the current signals.

Although widely used, MCSA faces some difficulty in detecting BRBs under low-load or no-load conditions. This problem arises because the characteristic BRB frequencies are close to the fundamental frequency and have low amplitudes. Although Hilbert Transform (HT) can offer improvements by enhancing these low-amplitude components, other challenges still remain [27], [28]. More specifically, Fast Fourier Transform (FFT), frequently used in MCSA, is susceptible to spectral leakage caused by finite-time windows, potentially obscuring characteristic fault frequencies. The critical limitation of these traditional spectral techniques is their reliance on fixed thresholds, which lack the adaptability required to differentiate between incipient defects and operational fluctuations in dynamic environments. Consequently, a more robust, integrated, and comparative analytical framework is required [28].

In the frequency domain, MCSA is the most widely used method for detecting BRB in IMs. This method is characterized by identifying the sideband frequencies $f_b = f_s(1 + 2ks)$, which indicate the BRB defect incidence, where f_b is the characteristic frequency of the broken bar, f_s is the motor supply frequency, s is the motor slip, and k represents the order of the sideband around the fundamental frequency.

An analysis performed using the same experimental database employed in this work is summarized in Table 1, which shows similar values for a standard detectable range, i.e., with a minimum load torque of 1.5 N.m. It is observed that increasing the load also increases the characteristic frequency for the same degree of severity, since this modulation depends on the motor slip. This behavior was observed in both techniques investigated [29].

Table 1 corroborates the known sensitivity characteristics of conventional MCSA's, which isolate the spectral components associated with BRB, and modulate the stator current supply frequencies. The inherent dependence of the characteristic fault frequency on the motor slip and, consequently, on the shaft load, is an intrinsic property of BRBs. This understanding is fundamental for an accurate interpretation of diagnostic signals under various operating conditions.

Traditional MCSA remains the industrial standard due to its simplicity, but its limitations in automated monitoring are well-documented. Spectral leakage, noise interference from Variable Frequency Drive (VFD)s, and, most critically, the overlap of fault sidebands with the fundamental frequency under low-load conditions (0.5 N.m to 1.0 N.m) often lead to unreliable diagnoses [29]. Furthermore, while state-of-the-art Deep Learning (DL) architectures have been proposed to overcome these spectral limitations, they often introduce a

TABLE 1. Comparison of defect characteristic frequency obtained by MCSA using phase electric current signal.

Rotor cond.	Load (N·m)	MCSA	
		Defect Freq. $f_s - 2sf_s$ (Hz)	*Slip Freq.* $2sf_s$ (Hz)
HR[a]	0.5	-	-
1BRB[b]	0.5	-	-
1BRB	1.0	-	-
1BRB	1.5	58.6	1.4
1BRB	2.0	57.9	2.1
1BRB	2.5	57.3	2.7
1BRB	3.0	56.8	3.2
1BRB	3.5	56.2	3.8
1BRB	4.0	55.4	4.6

[a]HR=healthy, [b]1BRB=1 broken rotor bar.

"black-box" nature that hinders engineering interpretability and requires significant computational resources, often unavailable in Edge Computing gateways.

While MCSA is essential for providing the theoretical basis and initial validation of defect signatures, it must be integrated with advanced techniques. ML-based approaches overcome MCSA limitations through several key mechanisms. First, they employ advanced feature extraction, moving beyond simple spectral frequencies to refine discriminative features from the statistical, temporal, and time-frequency domains. Second, they apply optimization and filtering techniques to select optimal feature subsets, eliminating noise and redundancy to maximize predictive capability. However, a notable research gap in current ML-based studies is the lack of a systematic evaluation regarding the necessity of voltage-based features, with most authors opting for full sensory redundancy without justifying the associated hardware costs.

Integrating ML enables the development of accurate, intrinsically robust diagnostic systems that maintain reliable performance even in noisy environments and under varying operational conditions. This robustness is crucial for automating Predictive Maintenance in Industry 4.0, where minimizing human intervention is very relevant. In this context, several advanced diagnostic techniques have been employed in the literature to detect defects during the transient operating state of electrical machines, including: Wavelet Transform (WT) [31], Weibull analysis [32], deep learning approaches [33], ensemble learning methods [34], fuzzy logic methodologies [35], and hybrid feature selection techniques [36]. Despite the high accuracy reported in these studies (often exceeding 98%), their reliance on high-dimensional feature spaces poses a challenge for real-time inference on low-cost microcontrollers.

Computational intelligence methods surpass traditional techniques in accuracy and flexibility, enabling the analysis of complex variables under diverse operating conditions. Pre-processing and feature extraction from raw sensor data, such as stator currents, are critical for the performance of subsequent ML models. While techniques like FFT analyze the frequency domain, they are most effective in steady-state operation. However, oscillations or instantaneous frequency changes cause harmonic overlap, seriously limiting defect detection. Time-frequency methods, such as the Short-Time Fourier Transform (STFT), WT, Discrete Wavelet Transform (DWT), and HT, enable the analysis of transient signals and low-frequency spectral regions. Yet, no single technique is universally effective. These methods often fail to accurately classify defect type or severity due to their limited generalization capacity in nonlinear systems. This highlights the need for more adaptive or hybrid approaches to achieve robustness in dynamic industrial environments. Specifically, the current state-of-the-art lacks a quantitative "cost-vs-accuracy" analysis, where the minimal sensory configuration is identified to support scalable predictive maintenance.

The advent of AI and ML has revolutionized defect diagnosis, opening new possibilities that surpass the constraints of traditional techniques. ML algorithms learn complex patterns from large data volumes, identify defect signatures, and deliver accurate diagnostic decisions without the need for manual programming tailored for specific defect conditions [37]. Researchers increasingly rely on Artificial Neural Networks (ANN), hybrid models, and standard algorithms such as Support Vector Machine (SVM), renowned for their robustness in high-dimensional spaces, as well as Decision Trees (DT) and kNN [38].

In real-world applications, diagnostic techniques can be hindered by electromagnetic noise, such as that from frequency inverters, as well as electrical contact noise and environmental noise. Therefore, robust diagnostic systems are essential. In this context, HT is a common technique for enhancing the signal-to-noise ratio by emphasizing fault features by eliminating fundamental components. Combining HT-derived features with ML models has proven effective in hostile environments. For instance, studies have shown that configurations that use the FFT and HT features, along with statistical and complexity measures (such as the Kernel Density Estimator, entropy, and peak values), are critical for diagnostic resilience. Specifically, the most robust configurations achieved high accuracy and sustained stability when subjected to 40% added white noise in steady-state BRB diagnosis [29]. Furthermore, noise robustness has been demonstrated even under transient conditions, where advanced intelligent systems achieved satisfactory performance with 80% added Gaussian noise by considering feature sets from multiple domains, including statistical, information-based, and complexity analysis [30].

Signal processing outputs can be highly correlated, making feature selection methods essential for reducing dimensionality, optimizing learning, and preventing overfitting in online applications [39]. The feature selection methods are categorized into three main types: filter, wrapper, and embedded methods [9]. Filter methods assess features based on their intrinsic properties, such as correlation, independently of the ML model. They are computationally efficient but may overlook complex interactions between features and classifier-specific performance, and fail to address multicollinearity. On the other hand, wrapper methods rely on ML algorithms to evaluate the quality of feature subsets, thereby optimizing model performance and capturing complex feature dependencies, albeit at the cost of repeated training and increased computational complexity. Also, embedded methods integrate feature selection directly into the model training process, offering a balance between efficiency and the ability to capture model-specific interactions. While these selection methods are common, they are rarely used to investigate the complete elimination of a sensor type (e.g., voltage sensors), which is the focus of the present work.

This paper uses the Best-First method for subset generation [40], using the Fast Correlation-Based Filter (FCBF) algorithm, where feature selection occurs before the classification. As a filtering method, FCBF may struggle to capture strongly interacting features because it evaluates them independently. A key contribution of this work is to filter representative features from signals with low correlation with the class, thereby supporting the feasibility of building models with a reduced set of signals.

A systematic review of the literature by [22] reinforces the issues mentioned above. Their results indicate that 40.83% of the reviewed studies focused on bearing defects, 8.85% on stator defects, 20.71% on broken rotor bars, and 17.75% on issues related to eccentricity. Furthermore, 11.83% of the studies addressed multiple defects, with or without discretization. Regarding the operating regime, 79.88% of the analyzed studies focused on steady-state defects, 2.96% on transient conditions, and 17.16% employed a hybrid approach.

Although most studies focus on the steady state, there is a significant gap in the comprehensive evaluation of techniques for identifying defect severity under varying load conditions and, crucially, in exploring reduced-electrical-sensing strategies for cost optimization and scalability. Many current studies, such as those employing 1D-CNNs on raw data, implicitly assume that more data from more sensors always leads to better models, neglecting the law of diminishing returns in sensor acquisition and the practical constraints of industrial DAQ systems.

The main research gaps addressed in this work and the contributions are the following:

- Predominant focus on defect detection without comprehensive severity identification, and under different operational load conditions [4], [9], [10], [11], [31], [32], [34], [36], [37], [41], [42], [43], [44], [45], [46], [48], [49], [50], [51], [52]. Many studies focus on defect detection or steady-state analysis, but neglect the crucial ability to classify defect severity under varying loads. This is essential for effective predictive maintenance planning and identifying anomalies in the element where the IMs trigger. Specifically, BRB detection is challenging at low loads due to the masking of fault frequencies and low amplitudes, necessitating explicit performance evaluation across diverse load conditions and defect severities. Additionally, identifying optimal electrical signal configurations for diagnosing the severity of broken bars under varying loads, and analyzing the predictive capacity of integrating signal processing, feature selection, and machine learning classifiers.
- Reliance on monitoring methods with full electrical sensing (three phases), resulting in high costs and complexity [9], [10], [44], [45], [46]. Conventional monitoring approaches require data acquisition from all phases, which increases installation and processing costs, thereby limiting the scalability and widespread adoption of predictive maintenance systems in industry. Analysis of the feasibility of reduced phase monitoring (e.g., using less than three phases) to maintain diagnostic accuracy, to substantially reduce system complexity and costs.
- Limited holistic evaluation of integrated defect diagnosis pipelines (signal processing + feature selection + machine learning) for diverse operational conditions [1], [8], [12], [22], [24], [53]. Prior studies often assess isolated components, neglecting their synergistic interaction for optimal performance. Comprehensive comparative evaluation of integrated techniques across all diagnostic stages.
- Demand for empirical validation and optimization of integrated pipelines for industrial deployment [23], [53], [54]. A clear understanding of 'what works best, and why' under challenging conditions is needed for deployable solutions. Identification of optimized, empirically validated integrated configurations for reliable and practical diagnostic systems.

The literature review reveals a strong trend in studying defect detection and diagnosis in three-phase IMs, often focusing on BRB defects under steady-state and offline analysis. However, a critical gap remains in the exploration of the predictive capability of reduced sets of electrical measurements, which is essential for real-world, cost-effective, and scalable industrial applications. Studies in the recent literature commonly rely on economically unviable sensor sets. This work addresses these gaps directly by proposing and validating an innovative approach that minimizes sensory input while maintaining high diagnostic accuracy, thereby paving the way for more efficient and widespread adoption of predictive maintenance systems.

III. MATERIALS AND METHODS

Figure 1 shows the proposed framework for detecting and diagnosing BRB in IM, as well as evaluating the reduced set of predictive signals using stator electric current and motor phase voltage. The proposed method includes six distinct steps, as follows.

A. EXPERIMENTAL SETUP

The database used in this work was constructed using an experimental workbench. The workbench, illustrated in

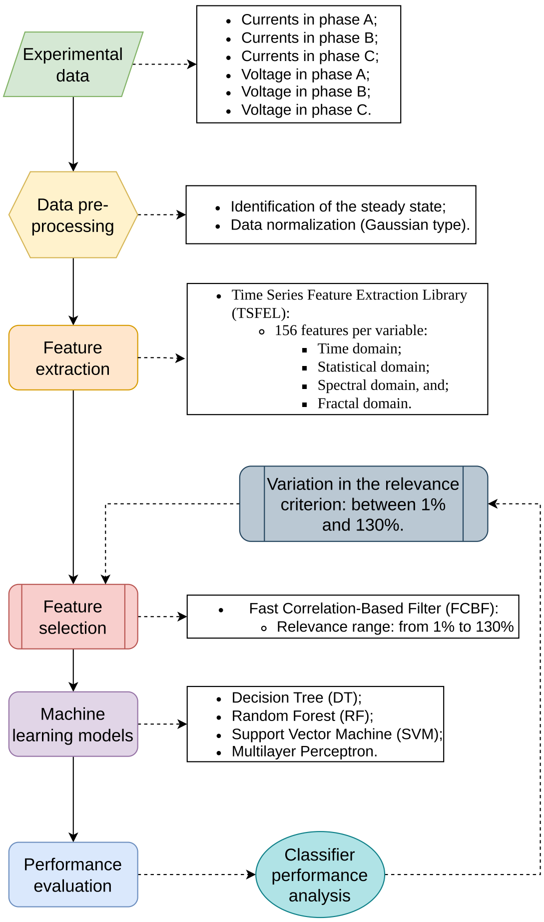

FIGURE 1. Methodological framework for diagnosis of BRB in three-phase IMs and reduced sensing analysis.

Figure 2, comprises a three-phase electric motor mechanically coupled to a torque transducer and a Direct Current (DC) machine operating as a generator. The DC machine simulates a resistive load, imposing a braking torque on the motor shaft. This configuration enables controlled test conditions for simulating defects and faults, as well as various operating conditions and loading on the motor shaft.

The IM used in this study has the following specifications: 1 HP, 220 V/380 V, 3.02 A/1.75 A, four poles, 60 Hz, with a nominal torque of 4.1 N.m and a nominal speed of 1,715 RPM. The rotor is a squirrel-cage type with 34 bars.

In this work, all data were acquired using a compact, modular system with signal filtering and amplification capabilities. A National Instruments data acquisition card (model NI DAQmx PCIe-6259) was used to acquire the electrical signals, configured with six input channels and a sampling rate of 50 kHz [43]. This sampling frequency ensures that the Nyquist-Shannon theorem is satisfied for harmonics up to 25 kHz, capturing high-frequency transient components that might be relevant for fractal analysis. The DAQ resolution was set to 16 bits, providing a dynamic

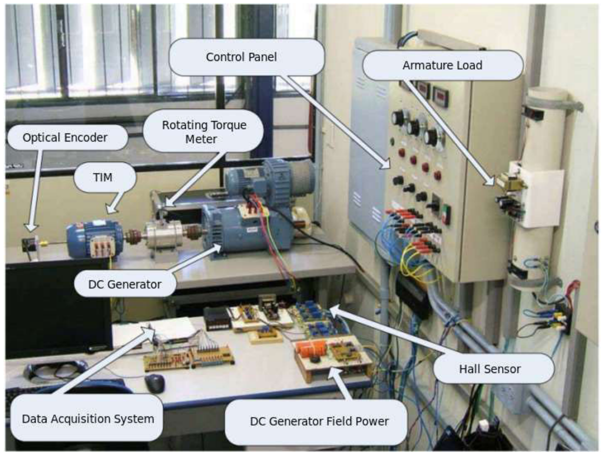

FIGURE 2. Experimental workbench [55].

range sufficient to distinguish low-amplitude fault sidebands from the 60 Hz fundamental component. The electrical signals are:

- Va: phase's A voltage;
- Vb: phase's B voltage;
- Vc: phase's C voltage;
- Ia: phase's A current;
- Ib: phase's B current;
- Ic: phase's C current.

Each experiment was run for 20 seconds. However, to ensure steady-state data, only the last 10 seconds were used. To simulate rotor defects, defects were introduced by drilling the squirrel cage using a 6 mm diameter drill, ensuring the hole exceeded the width of a rotor bar. These holes were drilled longitudinally in the center of the rotor, as shown in Figure 3(a).

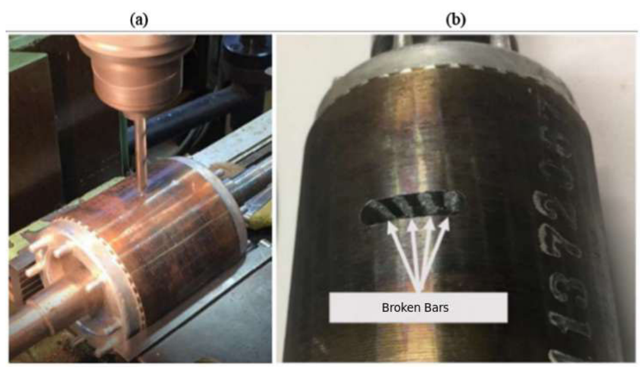

FIGURE 3. Procedure for inserting the broken bar defects into the rotor: (a) positioning the bench drill in the center of the rotor, (b) detail of the rotor with four adjacent broken bars [55].

Five rotor configurations were tested in this study to simulate progressive and common defects in real applications, where broken bars tend to occur adjacent to the rotor. The first rotor was defect-free and served as the baseline; we referred to this as the Healthy Rotor (HR). The second had One Broken Rotor Bar (1BRB), the third had Two adjacent Broken Rotor Bar (2BRB), the fourth had Three adjacent Broken Rotor Bar (3BRB), and the fifth rotor had Four adjacent Broken Rotor Bar (4BRB), as illustrated in Figure 3(b).

For each rotor configuration, eight distinct torque load conditions were applied, ranging from 0.5 N.m to 4.0 N.m, in increments of 0.5 N.m. Under each torque condition, data acquisition was repeated 10 times to ensure statistical reliability. Consequently, the experiment produced three-phase current and voltage signals for each combination of defect level and torque load, yielding a total of 2,400 time series.

All data collected in this experiment were published and are publicly available through the IEEE DataPort repository [55].

B. DATA PRE-PROCESSING

The time series collected presented different magnitudes due to the different torque conditions applied to the motor shaft. They were normalized using the Gaussian normalization method to avoid bias in the classification process. Furthermore, the database contained no missing values or systematic measurement errors. Signals collected at high torque (e.g., 2.5 N.m) will naturally have larger magnitudes than those collected at low torque (e.g., 0.5 N.m). An ML algorithm (such as SVM or Multilayer Perceptron (MLP)) commonly treats larger magnitudes as more important or assigns them greater weight, which can lead to bias in larger-scale features. The Z-score standardizes all samples, ensuring that the model treats the fluctuations and defect patterns (the features) extracted from 0.5 N.m with the same weight as those from 2.5 N.m. By centering the data at zero and scaling it by variance, the Z-score allows the ML model to focus on the relative variations within the signal, which are the valid signatures of the defects (rather than the absolute magnitude of the signal, which is only a function of the load). Standardization is applied to each feature x as follows (1):

$$z = \frac{x - \mu}{\sigma}, \tag{1}$$

where μ and σ are, respectively, the mean and standard deviation of each time series or feature in the training dataset. Furthermore, all subsequent methodological steps were based on the steady-state operating signals of the induction motor.

C. FEATURE EXTRACTION

To build effective machine learning models, we extracted predictive features from the normalized time series. We used the Time Series Feature Extraction Library (TSFEL) library, which generates 156 features across four domains: time, statistical, spectral, and fractal [56] The choice of these specific domains is grounded in the physics of BRB defects: spectral features capture frequency-domain sidebands, while fractal and statistical features quantify the non-linear irregularity and entropy changes in the current manifold caused by rotor asymmetry.

D. FEATURE SELECTION

The feature extraction step generated a high-dimensional dataset comprising 936 total features (156 features $\times$ 6 variables). To reduce dimensionality and eliminate redundant or irrelevant features, we implemented the FCBF algorithm. This method selects features based on relevance and redundancy using the Symmetrical Uncertainty (SU), an entropy-based metric that captures non-linear relationships. The FCBF operates in two main stages. First, SU between each feature and the class label (SU(Feature, Class)) was computed. Features with a SU value below a predefined relevance threshold were discarded as irrelevant. Second, the algorithm addresses redundancy by iterating through the remaining relevant features. It removes any feature F_i if it finds a more appropriate feature F_j (i.e., $SU(F_j, \text{Class}) > SU(F_i, \text{Class})$) that is also highly correlated with F_i (i.e., $SU(F_i, F_j) \geq SU(F_i, \text{Class})$). This ensures that the final subset contains only features highly correlated with the class, but poorly correlated with each other [57].

The FCBF relevance criterion functions as the algorithm's initial filter, designed to eliminate features with insufficient predictive power regarding the class label. This stage relies on SU, an information-theoretic metric that quantifies the correlation (linear and nonlinear) between an individual feature (F_i) and the target class (C). SU normalizes the Information Gain (IG) using the entropy (H) of both variables, defined as (2):

$$SU(F_i, C) = 2 \times \frac{IG(F_i|C)}{H(F_i) + H(C)}. \tag{2}$$

This equation scales the relevance value to a [0, 1] range, where 0 indicates no correlation, and 1 indicates perfect correlation. Unlike linear correlation coefficients, SU is robust against differences in the underlying distribution of features, making it ideal for a dataset containing both linear spectral peaks and non-linear fractal dimensions. In practice, the algorithm first calculates $SU(F_i, C)$ for every feature in the dataset. This value is then compared to a user-defined minimum relevance threshold (δ). Any feature F_i failing to meet this minimum threshold (i.e., $SU(F_i, C) < \delta$) is considered irrelevant, and the algorithm discards it immediately. Only the features that pass this relevance filter proceed to the second stage for redundancy analysis [57].

To assess the predictive ability of approaches that use exclusively electrical signals, the relevance criterion of the FCBF algorithm was varied. For each 10% increase in relevance, the number of selected features was reduced, allowing the predictive ability of each variable (Va, Vb, Vc, Ia, Ib, and Ic) to be analyzed in conjunction with the next step, since feature selection was based on their relevance.

E. MACHINE LEARNING MODELS

The ML techniques employed in this research are framed within a supervised multiclass classification context. The classification task aims to differentiate among five distinct rotor conditions: healthy rotor (no broken bars), and rotors with 1, 2, 3, or 4 broken bars.

Based on a review of related literature, this study adopts models from two complementary paradigms: symbolic learning methods offer higher interpretability and model transparency, which are essential for practical deployment and engineering validation. Linear learning models, on the other hand, are well-suited for capturing complex relationships within the extracted time-domain features.

After the feature selection step, the selected features were subjected to the machine learning models to construct the pattern classifier.

We split the dataset into 80% for training and 20% for testing, using stratified sampling to ensure balanced class representation. To enhance model reliability and robustness, we implemented a 5-fold stratified cross-validation on the training set. We paired this procedure with GridSearchCV for hyperparameter optimization.

To ensure the integrity of the evaluation and prevent data leakage, the feature selection process using FCBF was strictly coupled within the cross-validation folds. This ensures that the feature space is re-evaluated for each training subset, validating that the selected features are not artifacts of a specific data split but are truly representative of the physical fault signatures. Specifically, the relevance and redundancy criteria were calculated using only the training portion of each fold, and the resulting feature mask was then applied to the validation subset. This protocol ensures that no information from the validation or test sets influenced the selection of the most discriminative features, providing an unbiased estimation of the models' generalization capability.

The classification models implemented in this study were: Decision Trees (DT); Random Forest (RF); Support Vector Machine (SVM); and Multilayer Perceptron (MLP).

We selected the classification models based on their interpretability (e.g., DT) and suitability for our normalized, balanced, and highly class-related data (e.g., SVM). To evaluate the impact of reduced sensing, we iteratively tested this classifier set. In each iteration, we varied the feature selection relevance criterion, retrained the models using only the remaining features, and analyzed the resulting performance. This process allowed us to quantify the predictive capacity achievable under progressively minimized sensory input.

This selection specifically excludes probabilistic and linear models, such as Naive Bayes Classifiers (NBC), Bayesian Linear Discriminant (BLD), and Gaussian Mixture Models (GMM), due to their theoretical limitations regarding our dataset. For instance, NBC's assumption of feature independence is incompatible with the highly correlated multi-domain features extracted from current signals. Similarly, Bayesian linear methods and GMM often struggle with the non-linear boundaries and high-dimensional manifolds typical of electromechanical faults under varying loads. By prioritizing discriminative models (SVM and RF) and universal approximators (MLP), we employ architectures proven to be more robust for the complex nature of BRB diagnosis.

F. ASSESSMENT OF THE PREDICTIVE CAPACITY

After training the models, we expanded the performance assessment beyond classification accuracy. To provide a comprehensive evaluation of the models' sensitivity and specificity across the five rotor conditions, we included the F1-score (macro-averaged). Furthermore, all performance metrics are reported as the mean value $\pm$ the standard deviation across the 5-fold cross-validation, ensuring the statistical significance of the results. These indicators enable a detailed analysis of how the system handles ambiguities between adjacent defect severities, a common challenge in broken bar diagnosis.

We used this performance measure to compare the classifiers across the training scenarios defined by the varying feature selection relevance criterion. Additionally, we used the models to evaluate the predictive power of each experimental variable. This determined if we could reduce the required experimental signals (i.e., reduced sensing) while maintaining satisfactory performance.

G. CONFIGURATIONS AND PARAMETERS

One of the main components of this study is feature selection, as it significantly affects the classification stage's overall performance, which in turn requires an optimal configuration to achieve good performance for the project as a whole.

Based on preliminary tests, we defined the FCBF relevance range as 1% to 130%, with 10% increments. We set 130% as the upper bound because relevance criteria above this value yielded no features for training.

With the selected features, the ML models were evaluated for defect detection and diagnosis. Two classification approaches can be considered: binary and multiclass. Binary classification involves distinguishing between defective and non-defective conditions. In this paper, multiclass classification was employed, categorizing rotors into two conditions: those without defects and those with defects, indicating the severity (1, 2, 3, or 4 broken bars).

Accuracy was used as the sole performance indicator because our database was balanced. To ensure reproducibility, all models were implemented using the Scikit-Learn framework in Python 3.9. We optimized all machine learning hyperparameters using GridSearchCV to find the optimal configuration for each model. The resulting parameters were: for the DT, we set 'min_samples_split' as 10 to control overfitting by requiring at least 10 samples to split a node. The RF employed 50 trees ('n_estimators') and a 'min_samples_split' of 5. For the SVM, we chose the 'rbf' kernel to capture non-linearities and set the regularization parameter $C = 10$. This relatively high

C value imposes a significant penalty for misclassification, prioritizing a precise fit to the training data. Finally, we configured the MLP with two hidden layers (100 and 50 neurons), using the Adam optimizer and the ReLU activation function, a 'constant' learning rate, and L2 regularization ($\alpha = 0.001$).

IV. RESULTS AND DISCUSSION

The experimental results were analyzed to validate the hypothesis that a reduced sensory set, optimized through relevance-based feature selection, can provide high-fidelity diagnosis for BRB defects.Feature generation started with the normalized dataset using the TSFEL tool, which generates 156 features per variable. The dataset comprised 400 total tests, originating from 5 rotor conditions of the machine, each tested under eight load conditions (0.5 N.m, 1.0 N.m, 1.5 N.m, 2.0 N.m, 2.5 N.m, 3.0 N.m, 3.5 N.m, and 4.0 N.m) with 10 repetitions per load. Applying TSFEL to the six variables produced 936 total features (156 features $\times$ 6 variables). This process resulted in the final DataFrame for selection and classification, which had dimensions of 400 rows $\times$ 938 columns. The 938 columns include the 936 features plus columns for the test identifier and the rotor class label.

In this stage, a feature selection procedure was developed using the FCBF method. Unlike standard filter methods, the choice of FCBF is justified by its ability to handle feature redundancy through symmetrical uncertainty, which is crucial when dealing with highly correlated electrical signals from three-phase systems. However, to assess the predictive capability of each variable (currents and voltages), the procedure allowed the relevance criterion to be adjusted, thereby also allowing the number of relevant features to be changed. Table 2 reveals a critical transition point at the 110% relevance threshold, where all voltage-based variables (V_a, V_b, V_c) are eliminated, leaving only current signals (I_a, I_b, I_c). As shown in Table 3, at this 110% threshold, the feature space is composed exclusively of fractal descriptors. This suggests that while spectral features are sufficient in multi-sensory setups, the non-linear complexity of the current manifold is the only reliable indicator when voltage references are absent, especially at low-load conditions where spectral sidebands are obscured. This reduction confirms that for BRB diagnosis, the variance captured by stator currents contains the most significant discriminatory information, rendering the additional cost of voltage monitoring unnecessary for high-level classification.

We tested relevance levels between 1% and 120% with a 10% step size. No features met the selection criterion at thresholds of 130% or higher. These results provide the first evidence of the differing predictive capabilities among the electrical variables, a trend subsequently validated during the classification stage.

Throughout this study, we observed different behaviors across models, depending on the number of features provided to the model for classification. A total of 13 rounds of

TABLE 2. Number of selected features and corresponding variables as a function of the FCBF relevance criterion.

Relevance	*Features (Qty)*	*Variables*
1%	712	Ia /Ib /Ic /Va /Vb /Vc
10%	511	Ia /Ib /Ic /Va /Vb /Vc
20%	461	Ia /Ib /Ic /Va /Vb /Vc
30%	417	Ia /Ib /Ic /Va /Vb /Vc
40%	352	Ia /Ib /Ic /Va /Vb /Vc
50%	302	Ia /Ib /Ic /Va /Vb /Vc
60%	236	Ia /Ib /Ic /Va /Vb /Vc
70%	169	Ia /Ib /Ic /Va /Vb /Vc
80%	117	Ia /Ib /Ic /Va /Vb
90%	86	Ia /Ib /Ic /Va /Vb
100%	42	Ia /Ib /Ic /Va
110%	33	Ia /Ib /Ic
120%	3	Ib

classification were carried out, with the relevance criterion of the feature selection method varying between them. The selections resulted in limits from 712 to 3 selected features, based on their relevance level and domain, as shown in Table 3.

TABLE 3. Variation by domain of the selected features.

Dom. / *Rel.*	*Stat.*	*Temp.*	*Spec.*	*Frac.*	*Total*
1%	168	45	302	197	712
10%	104	88	177	142	511
20%	94	27	192	148	461
30%	79	77	142	119	417
40%	65	74	112	101	352
50%	57	69	89	87	302
60%	43	19	89	85	236
70%	33	15	50	71	169
80%	21	11	20	65	117
90%	17	7	5	57	86
100%	3	0	0	39	42
110%	0	0	0	33	33
120%	0	0	0	3	3

Rel.= Relevance, Dom.= Domain, Stat.= Statistical,
Temp.= Temporal, Spec.= Spectral, Frac.= Fractal

Analyzing the type of feature selected in each round, it is possible to conclude that fractal and spectral features are predominant during the selections. However, across the 60% to 70% of the relevance metric, we observe that fractal characteristics become more relevant than spectral characteristics. This remains the case until 100% of relevance is achieved, at which point only characteristics specific to this domain are selected. The isolation of these specific fractal components confirms that the proposed pipeline effectively targets the non-linear electromechanical signatures of the fault, rather than relying on linear correlations. Analyzing the last relevance percentage (120%), we can see which characteristics are most relevant for characterizing electrical signals from electrical machines:

- Wavelet energy at 4.17 Hz;
- Wavelet standard deviation at 4.17 Hz;
- Wavelet variance at 4.17 Hz.

The selection of Wavelet features at 4.17 Hz as the most discriminative indicators has a profound electromechanical justification. In a four-pole motor operating at 60 Hz, this frequency range is closely aligned with the pole-pass frequency ($f_{pp} = 2sf_s$), which is the physical rate at which rotor magnetic asymmetries interact with the stator field. When a bar breaks, the rotor's magnetic balance is disrupted, creating a backward-rotating magnetic field at slip frequency. This phenomenon induces torque pulsations and speed oscillations that modulate the stator current envelope specifically at $2sf_s$ [8], [58]. Consequently, the Wavelet transform effectively isolates these low-frequency modulations, which are often submerged by the fundamental 60 Hz component in traditional spectral analysis, especially under low-load conditions. This specific frequency capture demonstrates the framework's ability to act as a digital filter, focusing on the slip-dependent sidebands that characterize the physical loss of rotor symmetry.

Furthermore, the predominance of fractal domain features in the selection process underscores the non-linear nature of the BRB defect. Fractal metrics, such as complexity and irregularity measures, quantify the increase in the signal's electromagnetic entropy caused by the fault. In a healthy motor, the current signal exhibits a high degree of periodicity and symmetry. However, the progression of rotor bar breakage introduces localized distortions and non-stationary fluctuations in the magnetic flux. These variations increase the chaotic behavior of the current waveform, a characteristic that fractal descriptors are uniquely equipped to capture. By selecting these features, the FCBF algorithm identifies the mathematical signatures of the loss of physical symmetry within the machine, providing a robust basis for diagnosing defect severity across different operating torques.

From this point, it can be concluded that fractal and spectral features are fundamental for the classification of electrical signals, with particular emphasis on the fractal domain, which alone can provide sufficient information for classifier models to have satisfactory performance. Wavelets at 4.17 Hz contain the most discriminative information, enabling an efficient process with a significant reduction in the number of features generated. Wavelet analysis, unlike the Fourier transform, can provide resolution in both time and frequency, allowing for precise identification of the timing and occurrence of an event. To correlate these discriminative spectral components with the physical manifestation of the fault, the theoretical fault frequency for BRB can be used. The fault frequency in rotors with broken bars can be calculated using the following equation [59] (3):

$$f_{BRB} = (1 \pm 2s)f_s, \tag{3}$$

where f_{BRB} is the frequency of the broken rotor bar, f_s is the frequency of the power supply, and s is the motor slip.

However, 4.17 Hz is likely the sideband frequency caused by modulation of the central frequency [8], [58]. Wavelet analysis showed adequate behavior for isolating this failure frequency, even at such a low amplitude [59], [60].

Another critical analysis is the trend distribution curve generated by the variation in relevance, based on the quantity and type of features selected. The distribution of these reductions can be observed in Figure 4.

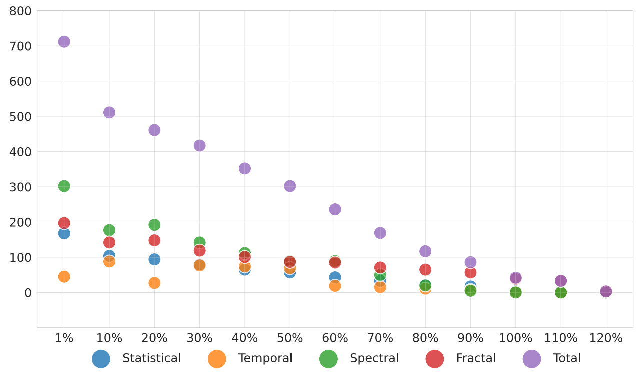

FIGURE 4. Feature selection trend.

We analyzed the performance of each classification model based on two main factors: the quantity of selected features and their source variables (i.e., the electrical signals from which they originated). To provide a comprehensive assessment, the models were evaluated using both accuracy and F1-score (macro) across the training and testing phases. Figures 5 and 6 present the testing performance, while Figures 7 and 8 illustrate the mean results obtained during the cross-validation (training) phase.

The results demonstrate how each classifier responds to the progressive reduction of sensory input. It is possible to state that the performance varies significantly depending on the selected variables, where: 6 corresponds to all three-phase currents and voltages (I_a, I_b, I_c, V_a, V_b, and V_c), 5 to I_a, I_b, I_c, V_a, and V_b, 4 to I_a, I_b, I_c, and V_a, 3 to I_a, I_b, and I_c, and 1 to I_b as the sole selected variable. The comparison between training and testing metrics confirms the generalization capability of the models, particularly for the RF and MLP architectures, which maintained high stability even under significant feature suppression.

The SVM demonstrated the most significant performance drop when one variable (Vc) was removed (relevance as 80%), indicating greater sensitivity to reduced feature sets and potentially less robustness to decreasing data dimensionality than other models. This behavior suggests a loss of linear separability in the reduced feature space. In industrial environments, where signal-to-noise ratios may vary, this sensitivity makes the SVM a less reliable candidate for minimal sensing compared to ensemble methods. One possible explanation is that removing voltage variables may have impacted the SVM's ability to find an optimal separating hyperplane in lower dimensions, especially if the voltages contributed nonlinear information crucial to class separability.

Furthermore, the DT model consistently performed poorly overall. This is likely due to its inherent susceptibility to

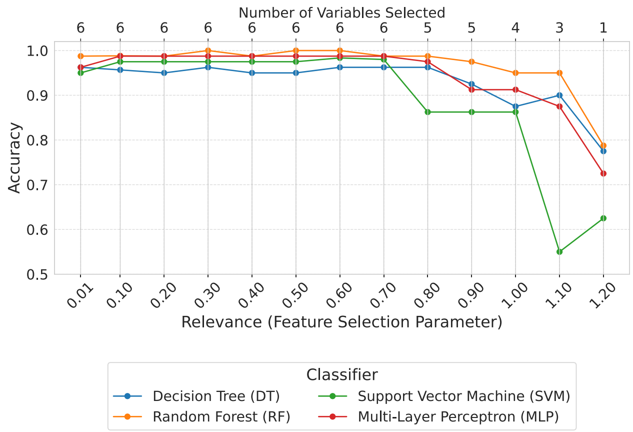

FIGURE 5. Comparison of classifier accuracy testing across different feature selection relevance thresholds, illustrating the performance variation as electrical variables are iteratively removed.

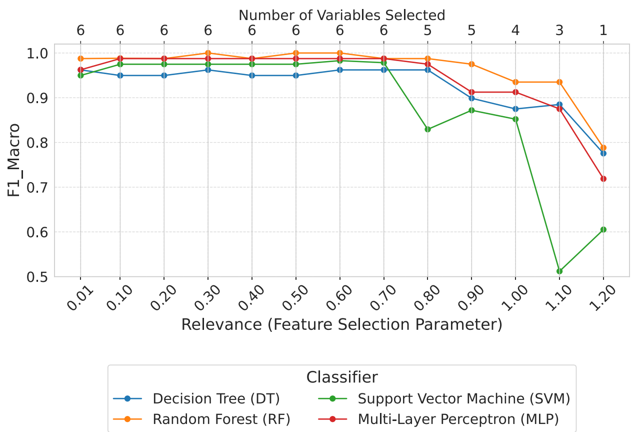

FIGURE 6. Comparison of classifier F1-score testing across different feature selection relevance thresholds, illustrating the performance variation as electrical variables are iteratively removed.

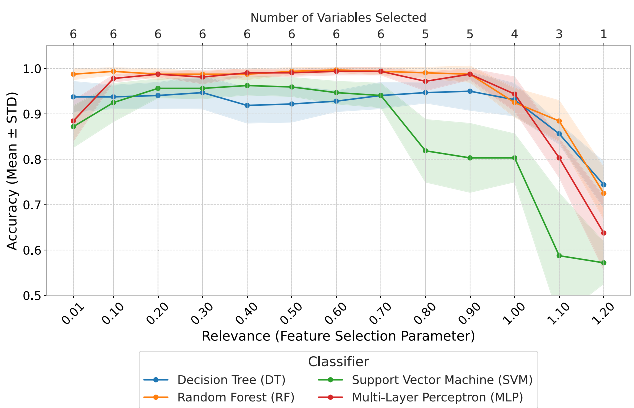

FIGURE 7. Comparison of classifier accuracy average training across different feature selection relevance thresholds, illustrating the performance variation as electrical variables are iteratively removed.

overfitting and its limitations in capturing complex, nonlinear relationships with fewer features. The high variance and

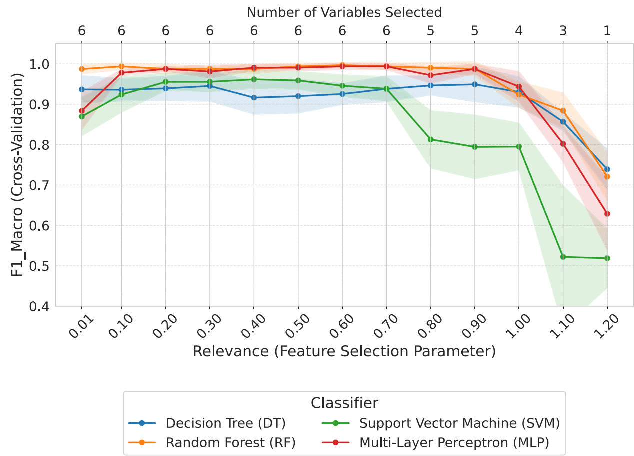

FIGURE 8. Comparison of classifier F1-score average training across different feature selection relevance thresholds, illustrating the performance variation as electrical variables are iteratively removed.

lack of regularization in a single tree architecture fail to generalize the subtle non-linearities of the current signals, highlighting the necessity of the ensemble approach adopted by the RF.

Thus, the best models were the RF and the MLP, with RF generally demonstrating superior performance throughout the experiment. A particularly significant result is the performance of RF as a feature selection relevance criterion of 110%. At this point, the selected features were exclusively related to electrical currents (Ia, Ib, and Ic), completely excluding voltage-type variables (Va, Vb, and Vc); and about the kind of features, just fractal domain features are selected. Despite this substantial reduction in input data, RF achieved 95% of accuracy, while MLP obtained 87.5%, and DT reached 90%. This result is of paramount practical importance, as it robustly validates the central hypothesis of this work: that reliable defect diagnosis can be achieved by monitoring only stator currents (Ia, Ib, and Ic). Current acquisition is typically easier and less costly in industrial settings than voltage acquisition, making this approach highly attractive for implementing more affordable and scalable predictive maintenance systems. The results, presented in Figure 9, are based on the confusion matrix of the four methods.

Figure 9 details the classifiers' performance specifically for the 110% relevance criterion, where the selected features are exclusively from the three-phase currents (I_a, I_b, and I_c). In this training scenario, RF maintained a superior performance with a mean accuracy of 95.21% $\pm$ 0.4% and an F1-score of 94.92% $\pm$ 4.4%, while DT achieved 85.62% $\pm$ 2.3% accuracy. The stability of the RF across the 5-fold cross-validation indicates that the model's performance is statistically consistent and not a result of localized data artifacts. The low standard deviation across the 5-fold cross-validation (σ < 0.5% for accuracy) provides strong evidence of the model's generalizability, confirming that the diagnostic performance is not an artifact of specific data splits but a robust reflection of the underlying fault physics. In contrast, models like SVM and MLP showed greater sensitivity to this reduction. In the case of SVM, showed a catastrophic performance drop at the 110% threshold, with F1-scores falling to $51.12\% \pm 17.6\%$. This precipitous decline, coupled with a sharp increase in the standard deviation, indicates that the current-only feature space is no longer linearly separable for the SVM. This critical failure highlights the necessity of the ensemble approach provided by RF, which remains stable by capturing non-linear feature interactions that simpler models fail to resolve without voltage-based normalization. Furthermore, the analysis of the standard deviation across the folds reveals that while the RF and DT models remain relatively stable, the sharp increase in the SVM error margin at this threshold indicates a loss of linear separability, validating the choice of ensemble and non-linear learners for minimal sensing industrial deployments.

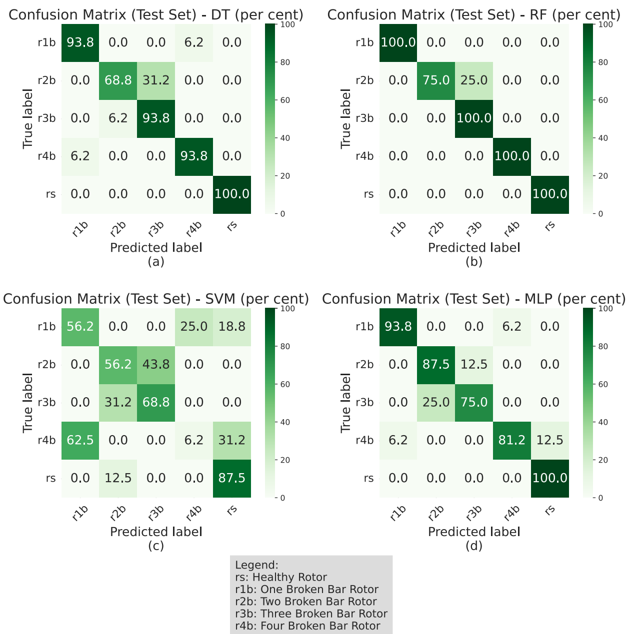

FIGURE 9. **Performance of pattern classifiers, with relevance of 110%: (a) DT, (b) RF, (c) SVM, (d) MLP.**

The results presented show that RF consistently achieves the highest accuracy values for most relevance levels (i.e., the number of variables). It maintains an accuracy close to 100% when the number of variables is high (relevance from 0.01 to 0.7), and its performance drop is more gradual and controlled than that of the other classifiers as relevance decreases (fewer variables are used).

The superior performance of RF can be attributed to several key characteristics inherent to its ensemble learning approach. In summary, the combination of RF's ensemble nature, ability to reduce overfitting, robustness, and effectiveness in high-dimensional environments makes it a very powerful and often high-performing classifier for various classification problems, such as rotor defect detection [5], [9].

However, it is worth noting that all the work developed envisioned a controlled environment as close as possible to the ideal, something very different from what a real plant would be like, with undesirable disturbances and noise. However, when analyzing the framework: Normalization + TSFEL + FCBF + RF, it can be estimated that:

- Gaussian normalization mitigates potential values generated by noise so that they do not significantly impact subsequent processing steps;
- The use of the TSFEL library for feature extraction enables time series analysis from the perspective of four distinct domains, while also offering techniques for detecting contouring, noise, and disturbances;
- The FCBF algorithm seeks to remove irrelevant features from the sample submitted to the classifier. By selecting the most relevant features to the target variable, FCBF eliminates information that might be affected by noise or disturbances, which could otherwise impair the classifier's effectiveness;
- Models such as RF have high robustness to outliers, and they are capable of operating with high-dimensional data.

Therefore, the proposed method could be applied in real-world environments to achieve satisfactory defect detection during real-time monitoring, even in the presence of noise disturbances characteristic of such environments. This study demonstrated the feasibility of an efficient and minimally invasive approach for monitoring conditions in three-phase induction motors. The results clearly indicate that it is possible to achieve reliable diagnoses using only the sensing of electrical currents, thereby eliminating the need for additional voltage sensors, which represents a significant reduction in the complexity and cost of implementing monitoring systems [9].

The proposed methodology, based on the integration of the TSFEL library for comprehensive extraction of temporal, fractal, spectral, and statistical features [56], followed by the application of the FCBF algorithm [57], for selecting more discriminative attributes (especially fractal-type attributes, based on Wavelets at 4.17 Hz) [58], [59], [60], proved highly effective in preparing the data for the classification process. The use of RF as the primary classifier proved particularly suitable for this application, demonstrating robustness against operational variations and generalization capability [38].

The experimental results validated the effectiveness of the proposed model, which achieved high accuracy and reliability in the detection and classification of different motor operating conditions [26], [34], [55]. Feature selection via FCBF allowed for the identification of a reduced set of the most relevant features, optimizing the computational process without compromising model performance [57].

This approach opens the possibility of investigating deep learning techniques for automatic feature extraction [13], [18], [33], [34], [61]. It expands the dataset to include more failure conditions and operating regimes, and implements the system in a real industrial environment for practical validation. Thus, this work presents an accessible and efficient approach for the predictive monitoring of IM, offering a practical solution for industry that combines modern ML techniques with simplified sensing, thereby promoting operational reliability and preventive maintenance of motor systems [1], [62].

A. COMPUTATIONAL EFFICIENCY AND COMPLEXITY ANALYSIS

The practical feasibility of the proposed diagnostic pipeline for Industry 4.0 applications depends on its computational footprint and near real-time response capability. Quantifying these complexities is essential for validating the scalability of the diagnostic system. Feature extraction via TSFEL involves multiple descriptors across four domains with varying complexities: statistical features generally follow $O(n)$, while spectral features based on the Fast Fourier Transform (FFT) are $O(n \log n)$, and more complex fractal descriptors can reach $O(n^2)$, where n is the number of data points per window [56]. To optimize this, our pipeline utilizes the FCBF algorithm for feature selection, which exhibits a computational complexity of $O(M \cdot N \log N)$, where M is the number of samples and N is the initial number of features. This multivariate filter approach is significantly more efficient than wrapper methods that require repeated classifier training, making it suitable for high-dimensional sensor data [57].

In comparison, classical MCSA methodologies rely primarily on the FFT, which maintains a complexity of $O(n \log n)$ [8]. While efficient, MCSA requires expert-driven visual inspection or fixed-threshold logic, which lacks the automated adaptability of ML. Furthermore, more advanced high-resolution spectral techniques such as Multiple Signal Classification (MUSIC) or Estimation of Signal Parameters via Rotational Invariance Technique (ESPRIT), often proposed to resolve sidebands at low slip, possess a significantly higher complexity of $O(n^3)$ due to the requirement of Singular Value Decomposition (SVD) on large autocorrelation matrices [63]. Our proposed pipeline strikes a balance between these extremes: it maintains lower asymptotic complexity than high-resolution subspace methods while providing more robust, automated classification than traditional frequency-domain analysis.

To quantify these metrics, the system was implemented in Python using a standard benchmark machine equipped with an Intel Core i5-12450HX processor (12 cores, 2.40 GHz) and 16 GB of RAM. The training phase, which includes hyperparameter optimization and 5-fold cross-validation, exhibited significant computational variance between architectures; at the 110% relevance threshold, the total training time for the RF model was approximately 0.52 s, using 0.155 GB of RAM, whereas the MLP required 33.55 s and 0.662 GB of RAM, due to its iterative gradient-based optimization. These offline training durations are considered negligible for industrial maintenance schedules, where model updates occur periodically.

The relationship between the feature selection threshold and the resulting computational overhead is detailed in Figures 10 and 11, which illustrate the total training and testing times, respectively. It is noteworthy that while the training duration reflects the offline computational demand for model optimization, the reduction in input variables directly minimizes inference latency during the online phase. This dual efficiency reinforces the viability of the proposed selection method, ensuring both rapid model updates and near-instantaneous diagnostic feedback for time-critical industrial tasks.

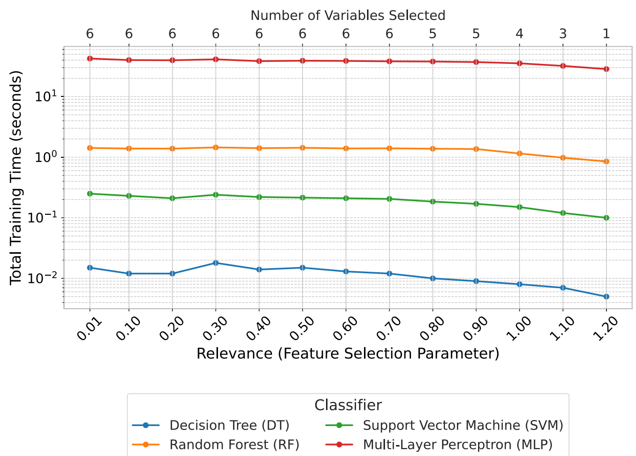

FIGURE 10. Variation in total training time for each classifier as a function of the feature relevance threshold, illustrating the reduction in offline computational demand via feature suppression.

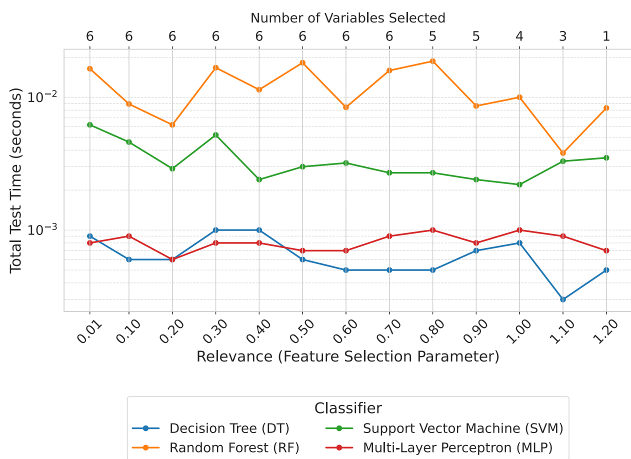

FIGURE 11. Variation in testing latency for each classifier as a function of the feature relevance threshold, highlighting the real-time operational efficiency gained through feature suppression.

For the online monitoring phase, the inference latency is the critical factor. The complete feature extraction and selection process is followed by the model's prediction, which for the RF at the 110% threshold required a total testing time of only 0.0038 s for the entire test set. On a per-sample basis, the inference latency for the MLP was even lower, at approximately 0.01092 ms, confirming that both architectures are well-suited for real-world industrial deployment where motor operating conditions require near-instantaneous diagnostic feedback.

Consequently, the total diagnostic cycle latency, including feature processing and model inference at the 110% relevance threshold, is approximately 0.004 s (4.0 ms), using 0.029 GB fo RAM, for the RF model. This memory footprint is approximately 22 times smaller than the requirements for the MLP at lower relevance levels, representing a significant breakthrough for deployment in low-cost ARM Cortex-M microcontrollers. This performance is highly compatible with real-time condition-monitoring requirements, where decision windows of < 0.5 s are typically acceptable for detecting incipient faults in induction motors. Furthermore, the low inference time of the RF model, which follows $O(n_{trees} \cdot \text{depth})$ during prediction, confirms that the pipeline is lightweight enough to be deployed on resource-constrained edge devices, such as industrial gateways or Micro-controller Unit (MCU)-based nodes. This efficiency eliminates the need for high-performance Graphics Processing Unit (GPU) resources, validating the feasibility of the proposed minimal sensing approach for decentralized industrial applications [29], [30].

B. INDUSTRIAL PERSPECTIVES AND LIMITATIONS

While the proposed methodology achieved high accuracy under controlled conditions, its transition to real industrial environments involves addressing specific challenges such as Electromagnetic Interference (EMI) and harmonic noise from VFDs. In practice, VFD switching frequencies can introduce spectral components that mimic or mask BRB signatures. However, integrating FCBF for feature selection provides a robust defense by identifying and retaining only the most dominant defect-related features, effectively filtering out non-stationary noise. This robustness is primarily due to the fractal and wavelet features selected, which capture the non-linear essence of the fault—a physical characteristic that remains distinct from the high-frequency stochastic nature of EMI. Furthermore, previous studies using this experimental dataset have shown that intelligent systems combining fractal features and ensemble classifiers can maintain reliable performance even under significant levels of added Gaussian noise [29]. Despite these strengths, a primary limitation of this study is the reliance on steady-state data. In industrial applications where motors frequently undergo transient states or operate under rapidly varying loads, the current-only features may require dynamic windowing to maintain reliability. Furthermore, the framework's performance under heavy harmonic distortion from aging power electronics remains a venue for future validation. Thus, the current pipeline is not only computationally efficient but also theoretically equipped to handle the spectral complexities of practical industrial deployment.

For field deployment, a practical adaptation strategy involves Transfer Learning or domain adaptation techniques. By pre-training the model on the comprehensive laboratory dataset and fine-tuning it with a small subset of labeled data from the specific industrial motor in operation, the system can account for unique machine characteristics and background noise profiles. Additionally, the low computational footprint of the Random Forest and TSFEL pipeline facilitates real-time execution on cost-effective edge devices, such as industrial microcontrollers, ensuring low latency and reduced data transmission costs in Industry 4.0 monitoring architectures. From a practical perspective, the ability to eliminate voltage sensors is not merely a cost-saving measure; it simplifies the non-invasive nature of the monitoring system. Current monitoring can be performed using split-core CTs without interrupting the power circuit or requiring high-voltage insulation for the DAQ inputs. This significantly reduces the risks associated with installation and maintenance, facilitating the transition from pilot projects to factory-wide predictive maintenance.

V. CONCLUSION

This study establishes that the diagnosis of broken rotor bars in three-phase induction motors can be reliably performed using a reduced electrical sensing scheme limited to stator currents (Ia, Ib, and Ic). By systematically identifying a relevance threshold of 110% via FCBF, we demonstrated that voltage-based features become redundant, allowing for a 50% reduction in the required sensory input channels without significant performance degradation. The practical implications of this finding are significant for the industrial sector: by eliminating the need for voltage and vibration sensors, the complexity and cost of the monitoring hardware are substantially reduced, enabling the cost-effective scaling of predictive maintenance across large industrial plants.

Our results demonstrate that an optimized pipeline (TSFEL + FCBF + RF) achieves a cross-validation accuracy of 88.4% and a test accuracy of 95% using only current signals. This performance provides a robust alternative to non-interpretable deep learning models, maintaining diagnostic stability (F1-score of 94.92% $\pm$ 4.4%) even under low-load conditions (0.5 N.m) where traditional MCSA fails due to spectral overlap. Furthermore, the validated computational overhead—characterized by a 0.029 MB memory footprint and 4 ms latency—proves that this computational efficiency and sensor parsimony directly support the transition toward Edge Computing in Industry 4.0, where localized, low-latency diagnostic decisions are paramount. Compared to state-of-the-art deep learning architectures, the proposed framework offers a superior balance between diagnostic accuracy and hardware feasibility on resource-constrained devices.

Despite the high performance achieved, we identify several concrete directions for future research to facilitate full-scale industrial deployment:

- Porting the proposed pipeline to resource-constrained hardware, such as industrial microcontrollers, to evaluate inference latency and energy consumption in a real-time monitoring context;
- Validating the methodology in environments powered by VFDs, where non-stationary harmonics and switching noise present additional challenges to fault signature isolation;
- Investigating the use of laboratory-trained models as a baseline for quick fine-tuning on field motors, accounting for unique machine signatures and background disturbances;
- Further exploring the correlation between fractal features and electromagnetic entropy to improve the transparency and engineering trust in automated diagnostic systems;
- Expanding the dataset to include concurrent faults (e.g., BRB combined with bearing wear) and testing the scalability of the reduced sensing approach in complex failure scenarios.

In conclusion, this research provides an empirically validated and economically viable framework for induction motor condition monitoring, quantifying the limits of sensor reduction and filling a critical gap between complex theoretical models and the practical requirements of modern industrial maintenance.

ACKNOWLEDGMENT

The authors would also like to thank Universidade Tecnológica Federal do Paraná (UTFPR) for the technical support and infrastructure provided during the experimental phase of this research.

REFERENCES

[1] N. Zhang, K. Cai, Y. Deng, and J. Zhang, "Joint optimization of condition-based maintenance and condition-based production of a single equipment considering random yield and maintenance delay," *Rel. Eng. Syst. Saf.*, vol. 241, Jan. 2024, Art. no. 109694.

[2] W. Yan, X. Wang, Q. Gong, X. Zhang, H. Zhang, and Z. Jiang, "A service-oriented energy assessment system based on BPMN and machine learning," *Auto. Intell. Syst.*, vol. 2, no. 1, p. 18, Aug. 2022, doi: 10.1007/s43684-022-00036-0.

[3] M. Hasanuzzaman, N. A. Rahim, R. Saidur, and S. N. Kazi, "Energy savings and emissions reductions for rewinding and replacement of industrial motor," *Energy*, vol. 36, no. 1, pp. 233–240, Jan. 2011, doi: 10.1016/j.energy.2010.10.046.

[4] Z. Ye, A. Sadeghian, and B. Wu, "Mechanical fault diagnostics for induction motor with variable speed drives using adaptive neuro-fuzzy inference system," *Electric Power Syst. Res.*, vol. 76, nos. 9–10, pp. 742–752, Jun. 2006.

[5] Y. Liu and A. M. Bazzi, "A review and comparison of fault detection and diagnosis methods for squirrel-cage induction motors: State of the art," *ISA Trans.*, vol. 70, pp. 400–409, Sep. 2017, doi: 10.1016/j.isatra.2017.06.001.

[6] S. Nandi, H. A. Toliyat, and X. Li, "Condition monitoring and fault diagnosis of electrical motors—A review," *IEEE Trans. Energy Convers.*, vol. 20, no. 4, pp. 719–729, Dec. 2005, doi: 10.1109/TEC.2005.847955.

[7] M. Katona, M. Kuczmann, and T. Orosz, "Accuracy of the robust design analysis for the flux barrier modelling of an interior permanent magnet synchronous motor," *J. Comput. Appl. Math.*, vol. 429, Sep. 2023, Art. no. 115228.

[8] N. A. R. Maciejewski, C. P. do Nascimento, L. Gremonini, B. A. T. Iamamura, Â. P. B. Da Silva Ferreira, J. P. A. Bastos, U. J. Dreyer, and T. De Paula Machado Bazzo, "Decision support system for diagnosing broken rotor bar severity in induction motors," *IEEE Access*, vol. 14, pp. 16751–16773, 2026.

[9] N. A. R. Maciejewski, R. Z. Freire, A. L. Szejka, T. P. M. Bazzo, S. M. A. Lopes, and R. A. Flauzino, "Artificial intelligence-based recommendation system for detecting and diagnosing broken bars in induction motors under transient operation," *IFAC-PapersOnLine*, vol. 58, no. 19, pp. 1156–1161, 2024.

[10] T.-J. Kang, C. Yang, Y. Park, D. Hyun, S. B. Lee, and M. Teska, "Electrical monitoring of mechanical defects in induction motor-driven V-belt–pulley speed reduction couplings," *IEEE Trans. Ind. Appl.*, vol. 54, no. 3, pp. 2255–2264, May 2018.

[11] W. A. Elhaija and Q. A. Al-Haija, "A novel dataset and lightweight detection system for broken bars induction motors using optimizable neural networks," *Intell. Syst. Appl.*, vol. 17, Feb. 2023, Art. no. 200167.

[12] D. Zhao, F. Fu, D. Wang, and Y. Shi, "Fault diagnosability evaluation for interconnected large-scale cyber–physical systems," *Automatica*, vol. 173, Mar. 2025, Art. no. 112090.

[13] H. Fu and Y. Liu, "A deep learning-based approach for electrical equipment remaining useful life prediction," *Auto. Intell. Syst.*, vol. 2, no. 1, p. 16, Dec. 2022, doi: 10.1007/s43684-022-00034-2.

[14] Y. Liu, C. Chen, T. Wang, and L. Cheng, "An attention enhanced dilated CNN approach for cross-axis industrial robotics fault diagnosis," *Auto. Intell. Syst.*, vol. 2, no. 1, p. 11, Dec. 2022, doi: 10.1007/s43684-022-00030-6.

[15] K. Lu, C. Chen, T. Wang, L. Cheng, and J. Qin, "Fault diagnosis of industrial robot based on dual-module attention convolutional neural network," *Auto. Intell. Syst.*, vol. 2, no. 1, p. 12, Dec. 2022, doi: 10.1007/s43684-022-00031-5.

[16] Z. Jin, C. Chen, A. Syntetos, and Y. Liu, "Enhanced bearing RUL prediction based on dynamic temporal attention and mixed MLP," *Auto. Intell. Syst.*, vol. 5, no. 1, p. 2, Jan. 2025, doi: 10.1007/s43684-024-00088-4.

[17] T. N. Babu, J. Saraya, K. Singh, and D. R. Prabha, "Rolling element bearing fault diagnosis using discrete Mayer wavelet and fault classification using machine learning algorithms," *J. Vibrat. Eng. Technol.*, vol. 13, no. 1, p. 87, Jan. 2025, doi: 10.1007/s42417-024-01549-z.

[18] M. Iqbal, A. K. Madan, and N. Ahmad, "Vibration and acoustic signal-based bearing fault diagnosis in CNC machine using an improved deep learning," *Iran J. Comput. Sci.*, vol. 7, no. 4, pp. 723–733, Dec. 2024, doi: 10.1007/s42044-024-00205-9.

[19] M. Iqbal and A. K. Madan, "Bearing fault diagnosis in CNC machine using hybrid signal decomposition and gentle AdaBoost learning," *J. Vib. Eng. Technol.*, vol. 12, no. 2, pp. 1621–1634, Feb. 2024, doi: 10.1007/s42417-023-00930-8.

[20] J. J. Guedes, A. Goedtel, M. F. Castoldi, D. S. Sanches, P. J. A. Serni, A. F. F. Rezende, G. H. Bazan, and W. A. de Souza, "Three-phase induction motor fault identification using optimization algorithms and intelligent systems," *Soft Comput.*, vol. 28, nos. 9–10, pp. 6709–6724, May 2024, doi: 10.1007/s00500-023-09519-5.

[21] B. H. Bahgat, E. A. Elhay, and M. M. Elkholy, "Advanced fault detection technique of three phase induction motor: Comprehensive review," *Discover Electron.*, vol. 1, no. 1, p. 9, Jun. 2024, doi: 10.1007/s44291-024-00012-3.

[22] N. A. R. Maciejewski, A. E. Treml, and R. A. Flauzino, "A systematic review of fault detection and diagnosis methods for induction motors," in *Proc. 40th Int. Conf. Electr. Eng. (FORTEI-ICEE)*, Sep. 2020, pp. 86–90.

[23] J. Aldrini, I. Chihi, and L. Sidhom, "Fault diagnosis and self-healing for smart manufacturing: A review," *J. Intell. Manuf.*, vol. 35, no. 6, pp. 2441–2473, Aug. 2024, doi: 10.1007/s10845-023-02165-6.

[24] M. Zdravković, H. Panetto, and G. Weichhart, "AI-enabled enterprise information systems for manufacturing," *Enterprise Inf. Syst.*, vol. 16, no. 4, pp. 668–720, Apr. 2022.

[25] F. V. Santos, L. Frosini, and D. M. Sotelo, "Diagnosis of rotor asymmetries in induction motors: A steady state approach using rotating frame references," *E-Prime-Adv. Electr. Eng., Electron. Energy*, vol. 12, Jun. 2025, Art. no. 100957.

[26] B. Corne, B. Vervisch, C. Debruyne, J. Knockaert, and J. Desmet, "Comparing MCSA with vibration analysis in order to detect bearing faults—A case study," in *Proc. IEEE Int. Elect. Mach. Drives Conf. (IEMDC)*, 2015, pp. 1366–1372, doi: 10.1109/IEMDC.2015.7409240.

[27] A. Giantomassi, F. Ferracuti, S. Iarlori, G. Ippoliti, and S. Longhi, "Electric motor fault detection and diagnosis by kernel density estimation and Kullback–Leibler divergence based on stator current measurements," *IEEE Trans. Ind. Electron.*, vol. 62, no. 3, pp. 1770–1780, Mar. 2015, doi: 10.1109/TIE.2014.2370936.

[28] M. B. Abd-el-Malek, A. K. Abdelsalam, and O. E. Hassan, "Novel approach using Hilbert transform for multiple broken rotor bars fault location detection for three phase induction motor," *ISA Trans.*, vol. 80, pp. 439–457, Sep. 2018, doi: 10.1016/j.isatra.2018.07.020.

[29] N. A. R. Maciejewski, R. Z. Freire, A. L. Szejka, T. P. M. Bazzo, V. B. Frencl, and A. E. Treml, "Comparative analysis of feature selection and classification techniques for robust broken rotor bar diagnosis in induction motors using current and vibration signals," *Auto. Intell. Syst.*, vol. 5, no. 1, p. 26, Oct. 2025.

[30] N. A. R. Maciejewski, R. Z. Freire, A. L. Szejka, T. D. P. M. Bazzo, S. M. de Andrade Lopes, and R. A. Flauzino, "Cognitive-based framework for detecting and diagnosing broken bars in induction motors for industry maintenance," *J. Ind. Inf. Integr.*, vol. 50, p. 101022, Nov. 2025.

[31] Q. Li, "New sparse regularization approach for extracting transient impulses from fault vibration signal of rotating machinery," *Mech. Syst. Signal Process.*, vol. 209, Mar. 2024, Art. no. 111101.

[32] J. Ben Ali, B. Chebel-Morello, L. Saidi, S. Malinowski, and F. Fnaiech, "Accurate bearing remaining useful life prediction based on Weibull distribution and artificial neural network," *Mech. Syst. Signal Process.*, vols. 56–57, pp. 150–172, May 2015.

[33] M. Sedghi, M. Zolfaghari, A. Mohseni, and J. Nosratian-Ahour, "Real-time transient stability estimation of power system considering nonlinear limiters of excitation system using deep machine learning: An actual case study in Iran," *Eng. Appl. Artif. Intell.*, vol. 127, Jan. 2024, Art. no. 107254.

[34] O. Attallah, R. A. Ibrahim, and N. E. Zakzouk, "Fault diagnosis for induction generator-based wind turbine using ensemble deep learning techniques," *Energy Rep.*, vol. 8, pp. 12787–12798, Nov. 2022.

[35] S. Raja and M. Rathinakumar, "Transient analysis of the fuzzy logic-based speed control of a three-phase BLDC motor," *Eng., Technol. Appl. Sci. Res.*, vol. 13, no. 1, pp. 9855–9860, Feb. 2023.

[36] H. Ehya, T. N. Skreien, and A. Nysveen, "Intelligent data-driven diagnosis of incipient interturn short circuit fault in field winding of salient pole synchronous generators," *IEEE Trans. Ind. Informat.*, vol. 18, no. 5, pp. 3286–3294, May 2022.

[37] S. Shin, Y.-Y. Choi, Y.-J. Sohn, M. Kim, I. S. Lim, and H. Oh, "Machine learning-based fault diagnosis for various steady conditions of proton exchange membrane fuel cell systems," *Int. J. Hydrogen Energy*, vol. 89, pp. 507–517, Nov. 2024.

[38] N. H. A. Malek, W. F. W. Yaacob, Y. B. Wah, S. A. M. Nasir, N. Shaadan, and S. W. Indratno, "Comparison of ensemble hybrid sampling with bagging and boosting machine learning approach for imbalanced data," *Indonesian J. Electr. Eng. Comput. Sci.*, vol. 29, no. 1, p. 598, Jan. 2022.

[39] P. Dhal and C. Azad, "A comprehensive survey on feature selection in the various fields of machine learning," *Appl. Intell.*, vol. 52, no. 4, pp. 4543–4581, Mar. 2022, doi: 10.1007/s10489-021-02550-9.

[40] I. H. Witten, E. Frank, and M. A. Hall, *Data Mining: Practical Machine Learning Tools and Techniques*, 3rd ed., San Mateo, CA, USA: Morgan Kaufmann, 2018.

[41] Z. Ye and B. Wu, "Online rotor bar breakage detection of three phase induction motors by wavelet packet decomposition and artificial neural network," in *Proc. IEEE 32nd Annu. Power Electron. Spec. Conf.*, vol. 4, Jul. 2001, pp. 2209–2216.

[42] Z. Chen, J. Meng, B. Liang, and D. Guo, "Fault diagnosis of AC squirrel-cage asynchronous motors based on wavelet packet-neural network," in *Proc. EUROCON*, 2007, pp. 720–724.

[43] M. Suetake, "Sistemas inteligentes para monitoramento e diagnósticos de falhas em motores de indução trifásicos," Ph.D. dissertation, Dept. Escola de Engenharia de São Carlos, Univ. Fed. Uberlândia, Uberlândia, Brazil, 2012.

[44] L. H. B. Liboni, "Diagnóstico de falhas em motores de indução trifásicos baseado em decomposição em componentes ortogonais e aprendizagem de máquinas," Ph.D. dissertation, Dept. Escola de Engenharia de São Carlos, Univ. São Paulo, São Paulo, Brazil, 2017.

[45] Y. Xie, P. Chen, F. Li, and H. Liu, "Electromagnetic forces signature and vibration characteristic for diagnosis broken bars in squirrel cage induction motors," *Mech. Syst. Signal Process.*, vol. 123, pp. 554–572, May 2019.

[46] M. F. Yakhni, S. Cauet, A. Sakout, H. Assoum, E. Etien, L. Rambault, and M. El-Gohary, "Variable speed induction motors' fault detection based on transient motor current signatures analysis: A review," *Mech. Syst. Signal Process.*, vol. 184, Feb. 2023, Art. no. 109737.

[47] R. M. Chagas, "Aplicação da técnica motor current signature analysis (MCSA) em motores de uma mesa de resfriamento de rolos," M.S. thesis, Dept. Centro Tecnológico, Univ. Fed. Itajub, Itajubá, Brazil, 2024.

[48] S. Tsuruta and D. Liu, "Harmonic signature extraction in motor fault detection via a weighted sparsity-driven and graph-based model," *IFAC-PapersOnLine*, vol. 58, no. 4, pp. 734–739, 2024.

[49] P. Kumar and A. S. Hati, "Dilated convolutional neural network based model for bearing faults and broken rotor bar detection in squirrel cage induction motors," *Expert Syst. Appl.*, vol. 191, Apr. 2022, Art. no. 116290.

[50] F. Fontoura, "Sistemas de detecção de falhas em motores elétricos baseados em análises de vibração: Estudo de um motor de indução trifásico," M.S. thesis, Dept. Centro Tecnológico, Univ. Fed. Santa Maria, Santa Maria, Brazil, 2023.

[51] R. Raja, J. Raja, S. Gogula, A. Katiyar, T. Thilagam, and V. S. Bhagavan, "Neural networks for fault diagnosis in electrical machine," in *Proc. Asian Conf. Intell. Technol. (ACOIT)*, 2024, pp. 1–4.

[52] P. V. Shinde, R. G. Desavale, P. M. Jadhav, and S. H. Sawant, "A multi fault classification in a rotor-bearing system using machine learning approach," *J. Brazilian Soc. Mech. Sci. Eng.*, vol. 45, no. 2, p. 121, Feb. 2023, doi: 10.1007/s40430-023-04015-1.

[53] M. Tiboni, C. Remino, R. Bussola, and C. Amici, "A review on vibration-based condition monitoring of rotating machinery," *Appl. Sci.*, vol. 12, no. 3, p. 972, Jan. 2022, doi: 10.3390/app12030972.

[54] F. Alyoussef, A. Akrad, R. Sehab, C. Morel, and I. Kaya, "Velocity sensor fault-tolerant controller for induction machine using intelligent voting algorithm," *Energies*, vol. 15, no. 9, p. 3084, Apr. 2022, doi: 10.3390/en15093084.

[55] A. E. Treml, R. A. Flauzino, M. Suetake, and N. A. R. Maciejewski, "Experimental database for detecting and diagnosing rotor broken bar in a three-phase induction motor," School Eng. of São Carlos Univ. of São Paulo (USP), São Carlos, Brazil, Tech. Rep., 2020, doi: 10.21227/fmnm-bn95.

[56] M. Barandas, D. Folgado, L. Fernandes, S. Santos, M. Abreu, P. Bota, H. Liu, T. Schultz, and H. Gamboa, "TSFEL: Time series feature extraction library," *SoftwareX*, vol. 11, Jan. 2020, Art. no. 100456, doi: 10.1016/j.softx.2020.100456.

[57] N. Gopika and A. M. Kowshalaya, "Correlation based feature selection algorithm for machine learning," in *Proc. 3rd Int. Conf. Commun. Electron. Syst. (ICCES)*, Oct. 2018, pp. 692–695, doi: 10.1109/CESYS.2018.8723980.

[58] M. Mehrjou, N. Mariun, M. Karami, S. Noor, S. Zolfaghari, M. Norhisam, Z. Kadir, M. A. M. Radzi, and M. H. Marhaban, "Wavelet-based analysis of MCSA for fault detection in electrical machine," in *Fault Detection Diagnosis*. Serdang, MY: IntechOpen, 2015, doi: 10.5772/61532.

[59] K. C. D. Kompella, M. V. G. Rao, and R. S. Rao, "Bearing fault detection in a 3 phase induction motor using stator current frequency spectral subtraction with various wavelet decomposition techniques," *Ain Shams Eng. J.*, vol. 9, no. 4, pp. 2427–2439, Dec. 2018, doi: 10.1016/j.asej.2017.06.002.

[60] C. Kar and A. R. Mohanty, "Monitoring gear vibrations through motor current signature analysis and wavelet transform," *Mech. Syst. Signal Process.*, vol. 20, no. 1, pp. 158–187, Jan. 2006, doi: 10.1016/j.ymssp.2004.07.006.

[61] E. L. de Aguiar, L. da Silva Nolasco, A. E. Lazzaretti, D. R. Pipa, and H. S. Lopes, "ST-NILM: A wavelet scattering-based architecture for feature extraction and multilabel classification in NILM signals," *IEEE Sensors J.*, vol. 24, no. 7, pp. 10540–10550, Apr. 2024, doi: 10.1109/JSEN.2024.3360188.

[62] O. Ayeni, "Integration of artificial intelligence in predictive maintenance for mechanical and industrial engineering," *Int. Res. J. Mod. Eng. Technol. Sci.*, vol. 7, no. 3, pp. 1–23, 2025.

[63] M. Kang and J.-M. Kim, "Singular value decomposition based feature extraction approaches for classifying faults of induction motors," *Mech. Syst. Signal Process.*, vol. 41, nos. 1–2, pp. 348–356, Dec. 2013.

LEONARDO BATISTA DE SANTANA received the B.S. degree in electrical engineering, the B.S. degree in control and automation engineering, and the M.Sc. degree in electrical engineering from the Federal University of Technology–Paraná (UTFPR), Curitiba, Brazil, in 2019, 2023, and 2023, respectively, where he is currently pursuing the Ph.D. degree with the Graduate Program in Electrical and Computer Engineering (CPGEI). He is also an Electrical Engineer with the Federal University of Paraná (UFPR), Curitiba, specializing in electrical installations and energy system planning. Previously, he was a Military Engineer with the Brazilian Air Force, where he developed projects for electrical infrastructure and communication networks. His research interests include intelligent systems applied to industrial processes, electrical machines, data analysis, and AI-based fault detection and predictive maintenance.

NARCO AFONSO RAVAZZOLI MACIEJEWSKI received the B.S. degree in mechanical engineering and the M.S. degree in electrical and computing engineering from the State University of Western Paraná (UNIOESTE), Foz do Iguaçu, Paraná, Brazil, in 2017 and 2018, respectively, and the Ph.D. degree in electrical engineering from the University of São Paulo (USP), São Carlos, São Paulo, Brazil, in 2021, with research focused on intelligent systems for fault detection in electrical machines. He is currently a Professor with the Federal University of Technology–Paraná (UTFPR), Curitiba, Paraná. His major work experience includes his current position and his previous research activities focused on intelligent systems applied to industrial processes. His research interests include electrical machines, biomechanics, data analysis, and the application of artificial intelligence to fault detection and predictive maintenance.

HEITOR SILVÉRIO LOPES received the B.Sc. and M.Sc. degrees in electrical engineering from the Federal University of Technology–Paraná (UTFPR), Curitiba, in 1984 and 1990, respectively, and the Ph.D. degree from the Federal University of Santa Catarina, in 1996. Currently, he is a tenured Full Professor with the Department of Electronics and the Graduate Program in Electrical Engineering and Applied Computer Science (CPGEI), UTFPR. His major research interests include computer vision, deep learning, evolutionary computation, and data mining.

JOSÉ JAIR ALVES MENDES JUNIOR received the Automation Technologist, M.Sc., and D.Sc. degrees in electrical engineering from the Federal University of Technology–Paraná, Curitiba, in 2015, 2016, and 2020, respectively, and the B.Sc. degree in electrical engineering from the Higher Education Center, Campos Gerais, in 2020. He is currently a Professor with DAELN, Federal University of Technology–Paraná. His research interests include biomedical instrumentation, signal processing, machine learning, and industrial automation.