Article

Improved Diagnostic Approach for BRB Detection and Classification in Inverter-Driven Induction Motors Employing Sparse Stacked Autoencoder (SSAE) and LightGBM

Muhammad Amir Khan [1], Bilal Asad [1,2,*], Toomas Vaimann [2] and Ants Kallaste [2]

[1] Department of Electrical (Power) Engineering, The Islamia University of Bahawalpur, Bahawalpur 6300, Pakistan; amirblouch41@gmail.com

[2] Department of Electrical Power Engineering and Mechatronics, Tallinn University of Technology, 12616 Tallinn, Estonia; toomas.vaimann@taltech.ee (T.V.); ants.kallaste@taltech.ee (A.K.)

[*] Correspondence: bilal.asad@taltech.ee

Abstract: This study introduces an innovative approach to diagnostics, employing a unique combination of techniques including a stratified group K-fold cross-validation method and a sparse stacked autoencoder (SSAE) alongside LightGBM. By examining signatures derived from motor current, voltage, speed, and torque, the framework aims to effectively detect and classify broken rotor bars (BRBs) within inverter-fed induction machines. In this kind of cross-validation method, class labels and grouping factors are spread out across folds by distributing motor operational data attributes equally over target label stratification and extra grouping information. By integrating SSAE and LightGBM, a gradient-boosting framework, we elevate the precision and efficacy of defect diagnosis. The SSAE feature extraction algorithm proves to be particularly effective in identifying small BRB signatures within motor operational data. Our approach relies on comprehensive datasets collected from motor systems operating under diverse loading conditions, ranging from 0% to 100%. Using a sparse stacked autoencoder, the model lowers the dimensionality and noise of the motor fault data. It then sends the cleaned data to the LightGBM network for fault diagnosis. LightGBM leverages the attributes of the sparse stacked autoencoder to showcase the distinctive qualities associated with BRBs. This integration offers the potential to improve defect identification by furnishing input representations that are both more precise and more concise. The proposed model (SSAE with LightGBM) was trained using 80% of the data, while the remaining 20% was used for testing. To validate the proposed architecture, we evaluate the accuracy, precision, recall, and F1-scores of the results using motor global signals, with the help of confusion matrices with receiver operating characteristic (ROC) curves. Following the training of a new LightGBM model with refined hyperparameters through Bayesian optimization, we proceed to conduct the final classification utilizing the optimal feature subset. Evaluation of the test dataset indicates that the BRBs diagnostic framework facilitates the detection and classification of issues with induction motor BRBs, achieving accuracy rates of up to 99% across all loading conditions.

Keywords: electrical machines; condition monitoring; induction motors; rotor faults; feature importance; broken rotor bars diagnosis; gradient boosting; LightGBM; artificial intelligence; predictive maintenance; fault classification; motor current signature analysis

1. Introduction

Industries heavily rely on induction motors due to their simple design and consistent performance, making them integral for heavy-duty applications. However, the potential for sudden failures poses significant safety risks and economic burdens. A timely diagnosis of abnormal conditions in induction machines is crucial to avoid costly downtime and ensure reliable operations. These abnormal conditions encompass various types of faults: stator faults such as open circuits, short circuits, and inter-turn faults, rotor faults, mechanical

faults in the rotor, including bearing damage, shaft misalignment, and eccentricity, failures in different components of the power electronics control system, and cracking or breaking of the magnetic material in rotating electrical machines [1]. Detecting these diverse faults provides valuable insights for diagnostic purposes. The characteristics of multiple faults in induction machines manifest in various ways, including line currents, air-gap voltages, torque fluctuations, machine losses, acoustic emissions, and excessive heating. These signals contain crucial information for identifying and diagnosing abnormal conditions in induction machines [2].

Diverse types of abnormal conditions commonly occur in induction machines and are categorized as either mechanical or electrical faults. The distribution of faults in induction motors is as follows: approximately 38% of faults are related to stator winding, 40% are related to bearings, 10% are related to rotors, and the remaining 12% are miscellaneous flaws [3]. In steady-state conditions, condition monitoring can primarily detect broken rotor bar problems by analyzing harmonics. This involves monitoring fundamental supply surrounded by slip frequency sidebands in the current frequency spectrum. The amplitude of sidebands increases with the severity of consecutive broken rotor bars under the same pole. Statistics indicate that around 10% of induction machine issues are attributable to broken rotor bars [4]. Figure 1 illustrates the distribution of failures that commonly occur in rotating electrical machines.
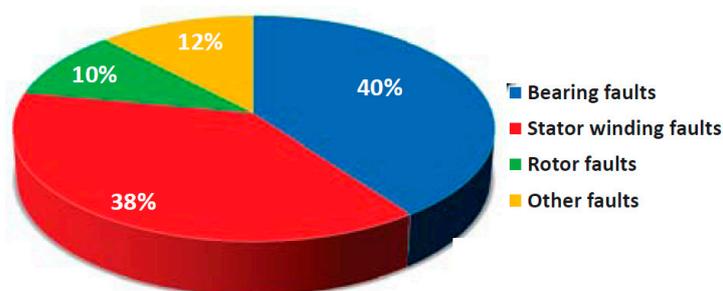
Figure 1. Distribution of abnormal conditions that occur in rotating electrical machines.

The conventional relays-based protection methods are gradually becoming obsolete, as advanced diagnostic algorithms capable of detecting faults at their incipient stages are emerging [5]. Among these faults, broken rotor bars (BRBs) pose a significant challenge in squirrel cage induction machines. The primary causes of fractured rotor bars in these machines include manufacturing defects, harsh duty cycles, and pulsating mechanical loads. These factors lead to machine vibration and torque fluctuations, resulting in mechanical, thermal, and magnetic stresses on the rotor bars. Induction motors, serving industrial, commercial, and domestic purposes, often operate at variable speeds [6]. Harmonics present in the motor line current can originate from three main sources: supply-side harmonics, internal harmonics, and motor faults. Principal and rotor slot harmonics are motor-generated harmonics resulting from non-sinusoidal winding distributions, slot openings, and material saturation [7]. The most prevalent method for identifying patterns in squirrel cage induction motor global signals is Motor Current Signature Analysis (MCSA) [8,9]. BRB defects introduce magnetic and geometric imbalances, leading to the appearance of sidebands in the stator current spectrum. The analysis ofsideband amplitudes and frequencies serves as valuable input for training machine learning models [10]. One notable advantage of MCSA is its non-invasiveness: sensor installation does not disturb motor operation [11].

Todiagnose broken bars in motors operating under indirect field-oriented control, artificial neural networks and the stator current spectrum derived from the Hilbert transform have been recommended [12,13]. These methods provide effective diagnostic capabilities while maintaining the operational integrity of the motor and enterprises. In [14], the proposal of an inverse thresholding technique applied to the current spectrogram aims to enhance the visibility of cracked rotor bars. Meanwhile, in [15], the authors presents an in-

tegrated design strategy focusing on preserving closed-loop functionality while leveraging active fault diagnostic and tracking control to detect early flaws. Authors in [16] advocate for the utilization of zonotopic observers and MANFIS models to ensure reliable detection of broken rotor bars (BRBs). However, [17] highlights a challenge where the fracture of two bars, situated a pole pitch apart, conceals the fault harmonics, rendering fault detection seemingly impossible. In scenarios where fault data are limited or unrepresentative, expert systems are often employed to augment defect identification. Domain knowledge-based problem signature identification systems, as discussed in [18], expedite outage correction by efficiently extracting fault features and patterns [19]. Deep learning methodologies offer promising avenues for classifying machine flaws, even in the absence of fault data, for both classification and regression analyses. Techniques such as neural networks, autoencoders, and LSTMs demonstrate efficacy in processing raw sensor data [20,21]. Transfer learning emerges as a solution to overcome data constraints [22,23], particularly in addressing BRB fault issues.

Recent advancements encompass intelligent diagnostics [24], real-time neural classifiers [25], and the integration of high-precision hybrid statistical and machine learning approaches [26]. Transfer learning combined with vibration imaging sheds light on fault-spectral behavior [27], while dictionary learning aids in damage identification through vibration analysis [28]. Furthermore, ongoing research emphasizes visual and time-series-based diagnostic solutions employing convolutional neural networks (CNNs). CNNs excel at capturing localized spatial correlations in both images and data, facilitating hierarchical feature learning across various scales [29]. Surface damage detection is garnering increasing attention, as evidenced by studies exploring node currents [30], convolutional neural networks [31], damage picture categorization [32], deep neural network feature extraction [33], and sensor fusion [31]. The main objectives and contributions of this paper can be summarized as follows:

- An enhanced methodology is implemented to accurately identify and classify the number of broken rotor bars (BRBs) in inverter-fed induction motors across various loading conditions by integrating sparse stacked autoencoders with LightGBM (SSAE–LightGBM) using stratified group k-fold cross-validation.
- This research introduces a novel cross-validation technique designed to balance the distributions of significant features in each fold, deviating from conventional stratification-based labeling. This refined approach exposes the model to diverse feature distributions, thereby augmenting its diagnostic capabilities and bolstering its reliability.
- Extensive datasets are utilized across various loading circumstances to collect a broad spectrum of signals and incorporating SSAE toreduce dimensionality and manage noise in motor fault data.
- The framework conducts an in-depth analysis of various operational parameters of the motor, including signatures of current, voltage, speed, and torque. By incorporating this multifaceted methodology, the framework enhances the breadth of fault detection and classification.
- Leveraging the sparse stacked autoencoder (SSAE), the framework effectively extracts complex features from comprehensive datasets, substantially improving the system's ability to identify BRB signatures across different loading conditions.
- Exclusive datasets collected from induction motors powered by inverters and operating under varied loading scenarios (0%, 25%, 50%, 75%, and 100%) are utilized to ensure a comprehensive analysis.

2. Identification of Abnormal Conditions in Broken Rotor Bars (BRBs)

Detection of damaged rotor bars using vibration signals, acoustics, current, magnetic flux, voltage, torque, and speed are prevalent in the literature [34]. A common approach in predictive maintenance involves segregating fundamental harmonics from faulty harmonics through spectral analysis [35]. In our study, we employed current, voltage, speed, and

torque harmonics to train the proposed algorithm. The following are the fundamental equations that wereused for the pre-analysis of the collected data.

2.1. Current Analysis

The motor current contains significant information about the health of the machine in the form of supply, winding, saturation, and fault-related harmonics [34]. The stator current modulates at a predetermined frequency in defective conditions. An induction motor with damaged rotor bars has the following frequency components in its stator current spectrum:

$$f_{brb} = (1 + 2sk)f_c, \; k = 1, \; 2, \; 3, \; 4 \ldots \tag{1}$$

where *s* represents the slip,f_{brb} denotes the frequency components of the current attributed to broken rotor bars in the frequency spectrum, f_C signifies the power supply frequency, and *k* stands for any positive integer.

2.2. Mechanical Torque Analysis

Flaws in rotor bars can induce torque pulsations due to magnetic field irregularities, resulting in harmonics in currents and voltages. Variations in speed resulting from torque inconsistencies can impact motor stability and performance [36]. Torque, being a function of electrical quantities, holds potential as a candidate for condition monitoring in electrical motors:

$$T = \frac{3 * V_{phase} * I_{phase} * \text{PF}}{\omega} \tag{2}$$

where V_{phase} and I_{phase} represent phase voltage and current, PF denotes the power factor, and ωsignifies the angular frequency. Detection oftorque pulses or deviations from expected values can serve as an indicator of rotor faults.

2.3. Voltage Analysis

Broken rotor bars can cause voltage spectrum abnormalities, especially in severe defects [34]. Detecting irregularities or distortions in the phase voltages can point out the BRBs in induction machines. Let V_a, V_b, andV_c be the phase voltages. The imbalances in the voltage can be calculated as:

$$Voltage \; imbalance(\%) = \frac{100 * \max(|v_a - v_b|, |v_b - v_c|, |v_c - v_a|)}{V_{average}} \tag{3}$$

2.4. Speed Analysis

In transient conditions, the detection of unanticipated changes in motor speed may indicate failures in BRBs. The synchronous speed N in an induction machine can be calculated as:

$$\omega_{syn} = \frac{120 * frequency}{No.of \; poles} \tag{4}$$

2.5. Difficulties of BRB Diagonistics in Inverter-Driven Induction Motors

The diagnosis of broken rotor bars is particularly challenging in inverter-driven systems due to the unique operational profiles these drives afford, including variable speed and load conditions. Traditional diagnostic methods often fall short under these conditions, necessitating innovative approaches to fault detection and diagnosis. Our experimental setup comprised two similar induction motors connected back-to-back: one serving as a test motor with deliberately induced rotor bar and bearing faults, and the other functioning as a healthy loading motor. This configuration allowed us to simulate operational conditions closely resembling real-world scenarios while maintaining control over the testing environment. Diagnosing rotor bar breaks in inverter-driven induction motors is tricky because these motors change speed and load frequently, making it hard to spot problems. In the frequency domain, the signals we use to detect faults can get mixed up with other signals

from the motor's control system, making it even harder to find issues. Sometimes, these faults are so small that normal diagnostic methods cannotdetectthem early on, leading to bigger problems later. In addition, the electrical noise from the inverter can hide fault signals in the motor electrical signals. Rotor bar breaks also occur in complex patterns, and they can happen when the motor is under different loads. To detectthese issues early on, we need to monitor the motor in real time and use advanced tools such asmachine learning to help us find and fix them quickly.

3. Improved Sparse Stacked Auto-Encoder (SSAE) with LightGBM

SAE (Sparse Autoencoder) is a hierarchical deep neural network comprising deep autoencoders. In basic autoencoder architecture, there are input, hidden, and output layers. The encoder is responsible for extracting hidden attributes from input data, while the decoder reconstructs input data using features from the hidden layer [37]. By closely matching input data, the autoencoder (AE) preserves fundamental properties. The input layer of the autoencoder receives the data, and the intermediate layers generate latent codes from the input data. Downsampling input data with principal component analysis (PCA) produces latent codes, the dimensionality of which depends on the nodes in the hidden layer [38]. The final layer of the model deciphers these latent codes to reconstruct the initial input. Autoencoders convert input data into latent codes and use them for reconstruction. The typical representation for AE inputs is $x = [x_1, x_2, \ldots, x_n]$. T in RD_x, where D_x is input dimension. The input x is translated from the input layer to the hidden feature vector h, which contains D_h neurons in the hidden layer, during encoding. This conversion process utilizes the activation function f. The encoding layer encodes network inputs, while the decoding layer decodes them [39]. Consequently, the neuron count in the decoding layer is determined by the input dimensionality. The primary objective of an autoencoder (AE) is to compute a reconstruction code h for an input instance x to accurately retrieve it. The formulation of a two-stage approximation function can be described as follows:

$$f_{dec}(f_{enc}(x)) = f_{\text{dec}}(\text{h}) = \hat{x} \approx X \tag{5}$$

where f_{enc} represents the function of the encoding layer and f_{dec} denotes the function of the decoding layer. The neuron count in the encoding layer is typically lower compared to the input dimensionality. Consequently, the network is encouraged to reduce input dimensionality in this layer to eliminate redundancy. The typical back propagation technique with random weight initialization is well-suited for training a single autoencoder because it is a shallow neural network. Figure 2 illustrates the workflow of the anticipated sparse stacked autoencoder with LightGBM architecture for BRBs diagnostics under various loading conditions.
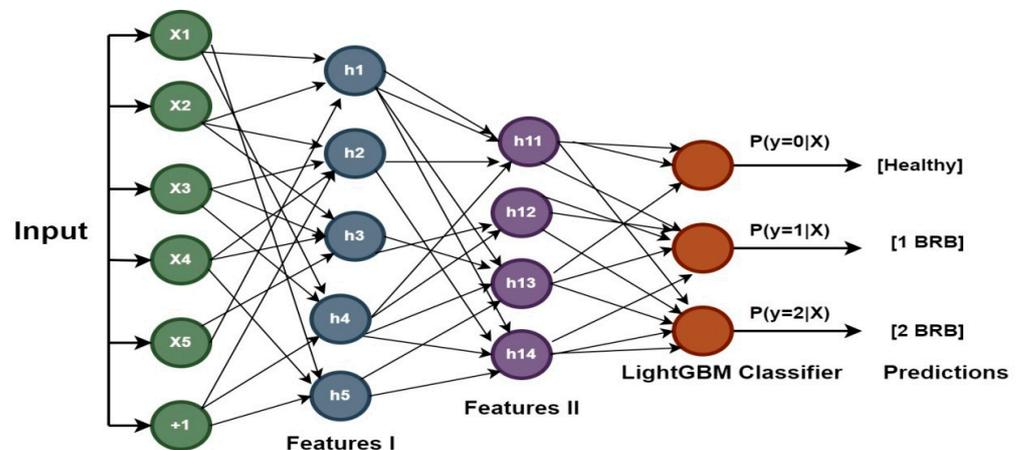

Figure 2. Anticipated sparse stacked autoencoder with LightGBM architecture for BRBs diagnostics.

Machine learning leverages the gradient-boosting framework LightGBM for tasks such as classification, regression, and ranking. Renowned for its ability to efficiently process massive datasets, LightGBM employs gradient boosting to construct a robust ensemble model from numerous weak learners, typically decision trees [40]. Notably, LightGBM has undergone enhancements for improved performance and efficiency. Unlike traditional boosting methods, LightGBM preserves eigenvalues within histograms, allowing for the simultaneous and distributed construction of decision trees. By iterating through discrete and continuous eigenvalues within histograms, LightGBM identifies optimal division points without the need for preordered boosting. Moreover, LightGBM employs depth-limited leaf growth, prioritizing the discovery of leaf nodes with the highest split gain over unnecessary split points. This approach contrasts with conventional level-wise decision trees, which may be less effective in certain classification tasks [41].

Consider the training sample set T of the LightGBM algorithm to be $\{(x_1,y_2)\,(x_2,y_2)\ldots(x_M,y_M)\}$ and the prediction output $\hat{y}_i$. Follow these procedures to determine the strong combination of weak decision trees for the *i*-th sample:

$$\hat{y}_i = \sum_{n=1}^{N} f_n\left(x_i, \{R_{ln}\}_{l=1}^{L}\right) \tag{6}$$

The decision tree is denoted as f_n, and *N* represents the number of trees at the *n*-th iteration. The feature space is divided into *L* non-overlapping regions denoted by $\{R_{ln}\}_{l=1}^{L}$. R_{ln} is the feature subset space that corresponds to the leaf node l of the *n*-th tree. To express f_n as an equation, it can be written as follows:

$$f_n\left(x_i, \{R_{ln}\}_{l=1}^{L}\right) = \sum_{l=1}^{L} \beta_{ln} I(x_i \epsilon R_{ln}) \tag{7}$$

The prediction score for the *i*-th leaf node is denoted by β_{ln}, whereas the piecewise function is depicted by I $(\cdot)$, which takes on a value of either 1 or 0. Let us consider the hypothetical value of the *i*-th sample at the *k*-th iteration, denoted as $\hat{y}_k^i$. At iteration *k*, the objective function $O^{(k)}$ can be estimated using the second-order Taylor expansion.

$$O^{(k)} \approx \sum_{i=1}^{M} \left(g_{ik} f_k\left(x_i, \{R_{lk}\}_{l=1}^{L}\right) + \frac{1}{2} h_{ik} f_k^2\left(x_i, \{R_{lk}\}_{l=1}^{L}\right) + \Psi(f_k) \right) \tag{8}$$

$$\Psi(f_k) = \alpha L + \frac{1}{2}\lambda \sum_{l=1}^{L} \beta_{lk}^2 \tag{9}$$

In which g_{ik} and h_{ik} represent the first and second derivatives of the logistic loss function with multiple classes. The effect of the regularisation term $\Psi(f_k)$ on the increase in leaf node quantity and quality is minimal. Additionally, the regularization factors α and λ are applied to the parameters. Determine the non-overlapping area $\{R_{lk}\}_{l=1}^{L}$ and the optimal leaf node score $\beta_{lk=1}^2$ for each tree to decrease $O^{(k)}$ during training. Rewriting the objective function $O^{(k)}$ as follows:

$$O^{(k)} \approx \sum_{i=1}^{L} \left(\left(\sum_{x_i \in R_{lk}} g_{ik} \right) \beta_{lk} + \frac{1}{2} \left(\sum_{x_i \in R_{lk}} h_{ik} + \lambda \right) \beta_{lk}^2 \right) + \alpha L \tag{10}$$

After the process of segmentation is completed, the sample sets RL and RR correspond to the left and right leaf nodes, respectively. The original nodes are represented by R = RL → RR. Superior tree growth is the consequence of increased benefit values. During the segmentation of leaf nodes, the gain associated with the split point of the candidate feature is calculated utilizing the given equation. The expansion of the leaf-wise growth method-

ology to include the utilization of a stacked autoencoder for the detection of anomalous conditions is illustrated in Figure 3.

$$Gain_{split} = \frac{1}{2}\left[\frac{\left(\sum_{x_i \epsilon R_L} g_i\right)^2}{\sum_{x_i \epsilon R_L} h_i + \lambda} + \frac{\left(\sum_{x_i \epsilon R_R} g_i\right)^2}{\sum_{x_i \epsilon R_R} h_i + \lambda} - \frac{\left(\sum_{x_i \epsilon R} g_i\right)^2}{\sum_{x_i \epsilon R} h_i + \lambda}\right] - \alpha \qquad (11)$$
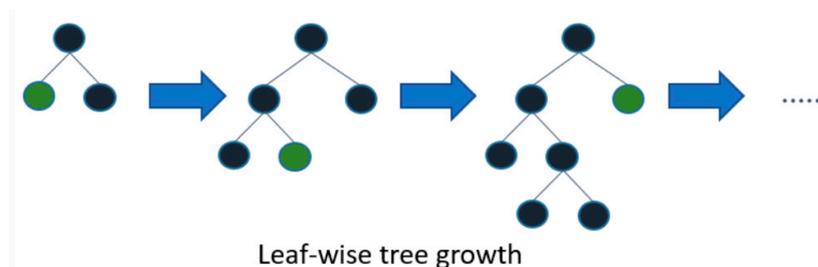

Leaf-wise tree growth

Figure 3. Methodology of tree growth based on leaves.

LightGBM divides the leaf with the largest split gain, and feature scores are computed by summing split operation gains or segmentation passes. Innovative algorithms such as exclusive feature bundling (EFB) and gradient-based one-side sampling (GOSS) are employed in LightGBM. GOSS targets poorly trained samples while preserving baseline data dispersion, enhancing model efficiency. On the other hand, EFB minimizes feature information while retaining original feature information, thereby improving algorithm efficiency [42]. Consequently, LightGBM enhances memory conservation and accelerates training.

3.1. Hyperparameter Setting of Improved SSAE and LightGBM

The optimal performance of an enhanced SSAE can be achieved through hyperparameter adjustment. Denoising, variational, or sparse autoencoders are some examples of changes that might be used in an enhanced SSAE. Take into account Table 1 shows the following critical hyperparameters of stacked sparse autoencoder SSAE (Table 1):

Table 1. Optimum hyperparameters of SSAE.

Hyperparameter	Layer 1	Layer 2	Layer 3
No. of neurons	300	250	200
No. of leaves	31	31	31
Learning rate	0.001	0.03	0.05
Dropout	0.3	0.3	0.2
Feature fraction	0.9	0.9	0.9
Batch size	64	64	32
Optimizer	Adam	Adam	SGD
Verbose	0	0	0
Activation function	ReLu	ReLu	Sigmoid
Loss function	Mean square error	Mean square error	Binary cross entropy

3.2. Bayesian Hyperparameters for Light GBM

The process of selecting hyperparameters is pivotal in the modeling process, and LightGBM provides a diverse set of hyperparameters to choose from. To enhance real-time efficiency in fault detection, only parameters with a significant impact on model performance were selected for hyperparameter tuning. Bayesian optimization was employed to effectively tune these hyperparameters, as it constructs a probabilistic model of the objective function and selects hyperparameters with the highest likelihood of success. Table 2 lists some of the most critical hyperparameters targeted for improvement using Bayesian optimization in the LightGBM gradient boosting framework:

Table 2. Optimum parameters for LightGBM.

| Hyperparameter	Values
Number of trees	300
Learning rate	0.0687
Max depth	7
Min child samples	13
Subsample	0.87
Feature fraction	0.76
Boosting type	gbdt
Iterations	721
Regularization parameters	0.54, 0.67
Bagging frequency	15
Min-split-gain	0.94

4. Test Bench Setup

For data collection, one, two, and three broken bar rotors were prepared by drilling radial holes, as shown in Figure 4. The monitoring of various parameters produces diverse datasets, resulting in the acquisition of voltage, current, torque, and speed signals. For measurements and experiments, we utilized two different motors, with their parameters provided in Table 3. Both motors were fixed to the same mechanical base and coupled with one another through shafts, as shown in Figure 5. The overall flow of the proposed method is shown in Figure 6. Both motors were supplied through industrial inverters; the load was changed by changing the torque of the loading machine working under direct torque control (DTC). A Dewetron transient recorder was utilized to measure the current, voltage, speed, and torque at a sampling frequency of 20 KHz. We divided the data into four categories: healthy, 1BRB ('1'), 2BRB ('2'), and 3BRB ('3'). The training dataset contained 1.2 million samples at 20 kHz for both scenarios, and the validation dataset contained 300,000 samples. We converted the data from the time domain to the frequency domain as well to train a better model for more accurate testing outcomes. For the measurement phase, both motors were driven by ABB industrial drives (loading: ACS600, testing: ACS800), employing scalar and direct torque control (DTC) modes to vary the operational conditions. A Dewetron data acquisition system, alongside Fukei400 current probes (pending confirmation), captured the electrical signals at a sampling frequency of 20 kHz. The experimental approach involved the creation of broken rotor bar conditions by drilling holes in the rotor slots and carefully matching the slot depth to ensure a consistent and reproducible fault condition. Tests were conducted under a range of loading conditions (from 0% to 100% of nominal load) and control modes to assess the diagnostic challenges and identify potential signatures indicative of broken rotor bars.


(a)   (b)   (c)   (d)

Figure 4. (**a**) Healthy BRB, (**b**) one BRB, (**c**) two BRB, (**d**) three BRB.

Table 3. Class labels and their associated conditions for IM dataset for BRB faults diagnosis.

Sr. No	Fault Type	Loading Condition	Sampling Frequency (KHz)	Data Points	Number of Samples	Fault Labels
1	Healthy	0%, 25%, 50%, 75% and 100%	20	417175	8250	Healthy
2	1 BRB	0%, 25%, 50%, 75% and 100%	20	413741	8250	1BRB
3	2 BRB	0%, 25%, 50%, 75% and 100%	20	413660	8250	2BRB
4	3 BRB	0%, 25%, 50%, 75% and 100%	20	415787	8250	3BRB
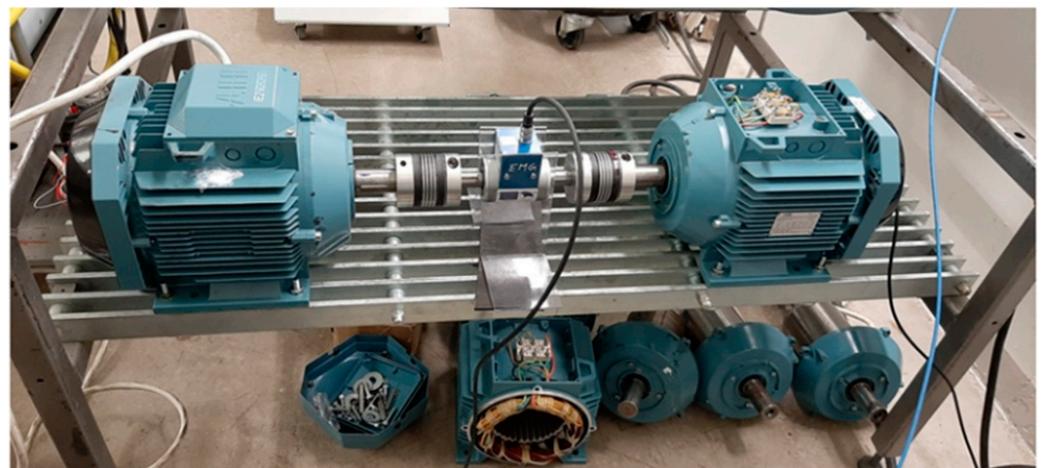
Figure 5. Practical setup with loading motor (**right side**) and testing motor (**left side**).
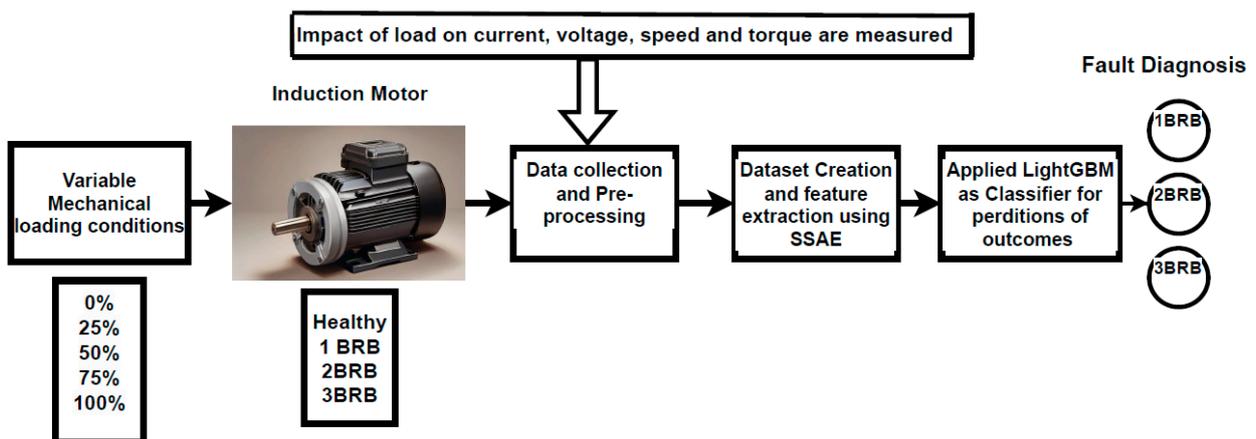
Figure 6. Proposed fault diagnosis architecture for fault classification in inverter-fed induction machines.

Table 4 shows the main parameters of the machines under investigation, including a four-pole, three-phase induction motor with a star/delta connection, 36 non-skewed stator slots, and 28 skewed rotor slots. The motors were rated at 7.5 kW power, with terminal voltages of 690 V/400 V at 50 Hz, and rated currents of 8.8 A/13.5 A.

Table 4. Specifications and electrical parameters of the 7.5 kW induction motor.

Sr.No	Parameter	Symbol	Value
1	Rated power	Pr	7.5 kw @ 50 Hz
2	Rated speed	Nr	1500 rpm @ 50 Hz
3	Rated-voltage	V	400 V @ 50 Hz
4	Connection	Δ	Delta (Δ)
5	No. of poles	P	4
6	Stator slots	Ns	36; non-skewed
7	Rotor bars	N_r	28; skewed
8	Terminal voltage	V	333 V, 50 Hz
9	Rated slip	S	0.0667
10	Rated current	I	8.8 A

5. Results and Discussion

It is important to analyze the harmonic components of current, voltage, torque, and speed and understand how these components change with varying loading circumstances. Increasing the load might cause the harmonic components to change their behavior, a phenomenon which may hide fault-induced harmonics or make their consequences worse. Thus, it is crucial to precisely differentiate fault-related harmonics from load-induced changes for successful fault identification. Furthermore, the existence of harmonics close to supply and slotting frequencies emphasizes the need to thoroughly analyze particular frequency ranges to distinguish fault patterns from surrounding noise. Utilizing sophisticated signal processing methods and strong diagnostic approaches can improve the dependability and precision of fault diagnostics in motor systems, reducing downtime and maximizing operating efficiency. To detect abnormal conditions, the harmonic components of the current, voltage, torque, and speed should be considered. Fourier transforms of the mentioned signals can identify these constituents in the frequency domain. These frequency components are the function of load. Here, the load was found to increase and change harmonic component behavior. The fault-based harmonics were distributed in the entire spectrum. They significantly appeared near supply and slotting harmonics. Figures 7 and 8 illustrate the frequency spectrum of current, voltage, speed, and torque with their FFTs at 100% loading conditions for the healthy and the three broken rotor bars, respectively.
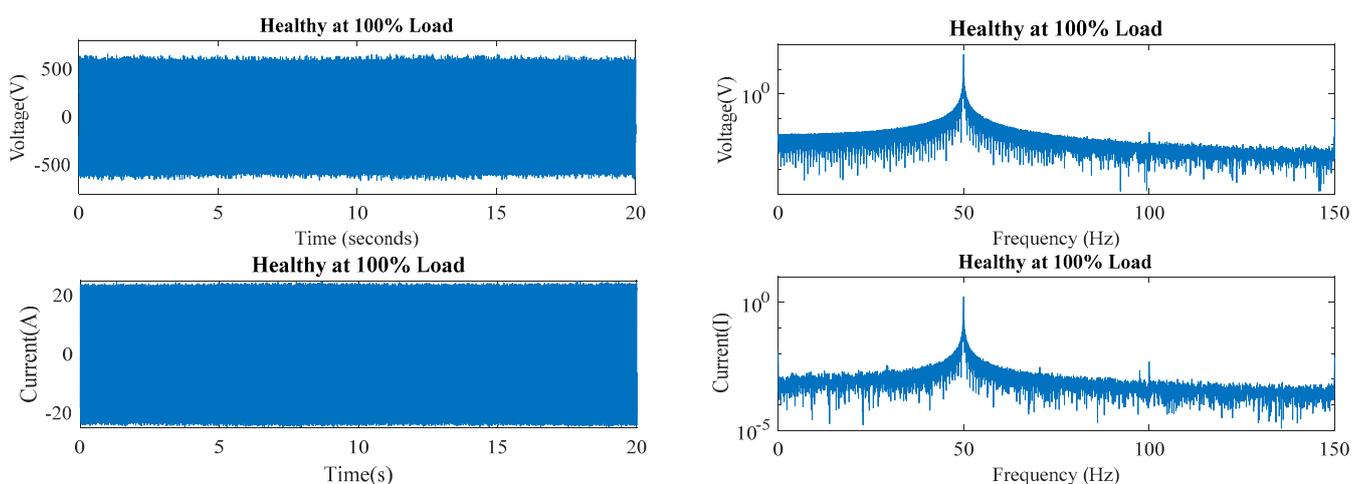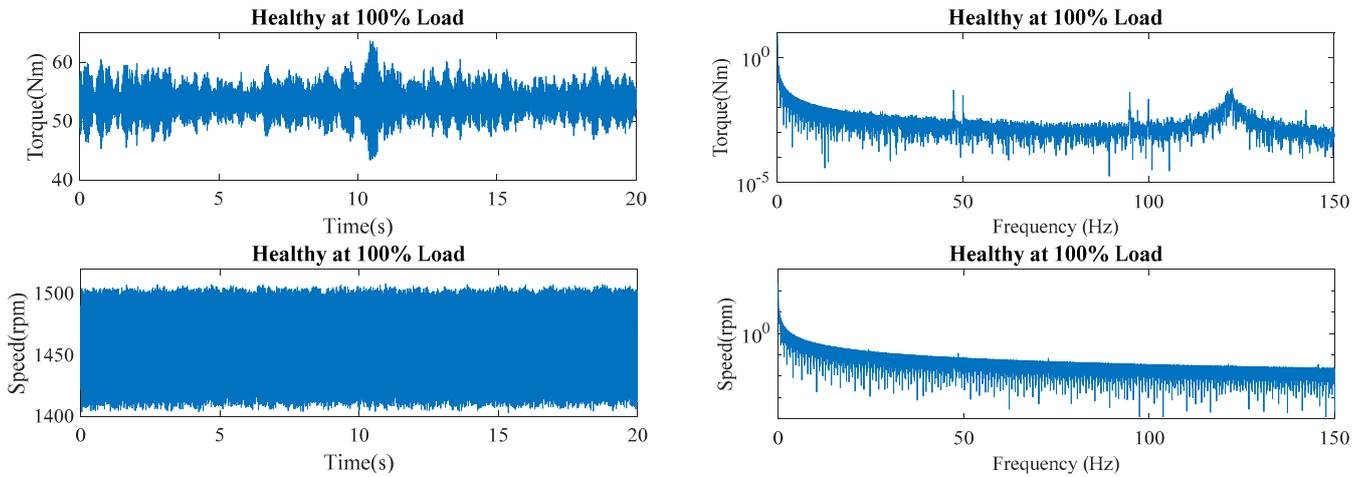

Figure 7. *Cont.*

Figure 7. Motor phase voltage, current, torque, and speed (from **top** to **bottom**) and their frequency spectrum for healthy broken rotor bars.
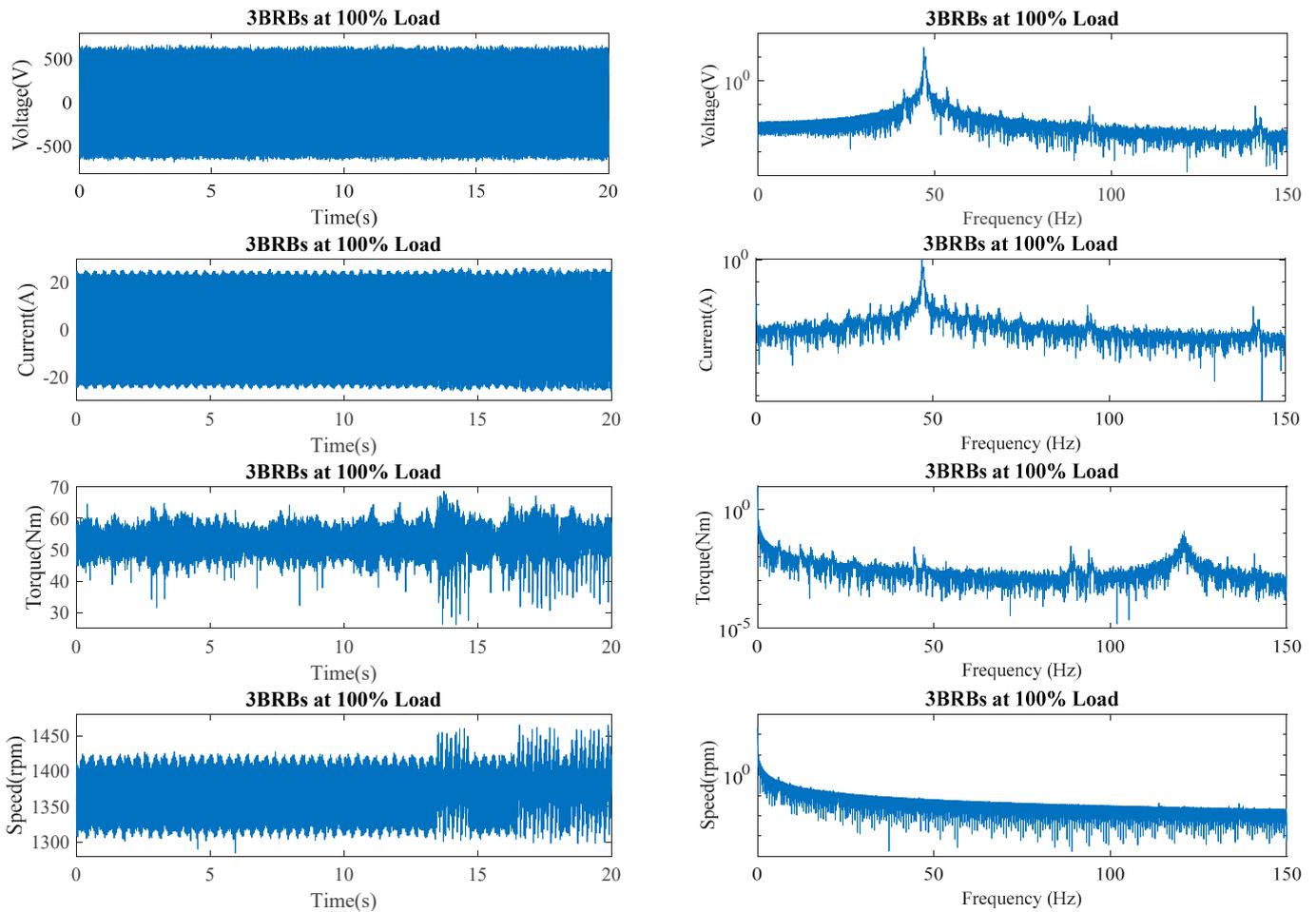

Figure 8. Motor phase voltage, current, torque, and speed (from **top** to **bottom**) and their frequency spectrum for three broken rotor bars.

We analyzed harmonic components and display signal frequency content in spectrograms under various operating parameters. By constructing spectrograms for normal operation and cases with one, two, or three broken rotor bars (BRBs), we can examine how

signal frequency properties vary under fault conditions. Spectrograms analyze frequency distribution and intensity, revealing event-specific fault patterns. We can evaluate rotor bar spectrum fluctuations to develop effective motor fault diagnostic methods by comparing spectrograms. The spectrograms of all the healthy and faulty conditions are shown below to detect the abnormal conditions in the healthy, one BRB, two BRBs, and three BRBs broken rotor bars. These spectrograms may reveal patterns connected to several faulty states, helping identify and understand the machine's health. When the load changes, these spectrograms demonstrate how the strength and emphasis of the frequency change. Figure 9 shows the spectrograms of healthy and faulty conditions at 100%. It is evident that, as the number of broken rotor bars increases, the spectral power density increases in the entire frequency range. This is because the fault-based harmonics and their copies spread in the wideband frequency range. The number and the strength of those harmonics depend on the type and the severity of the fault.
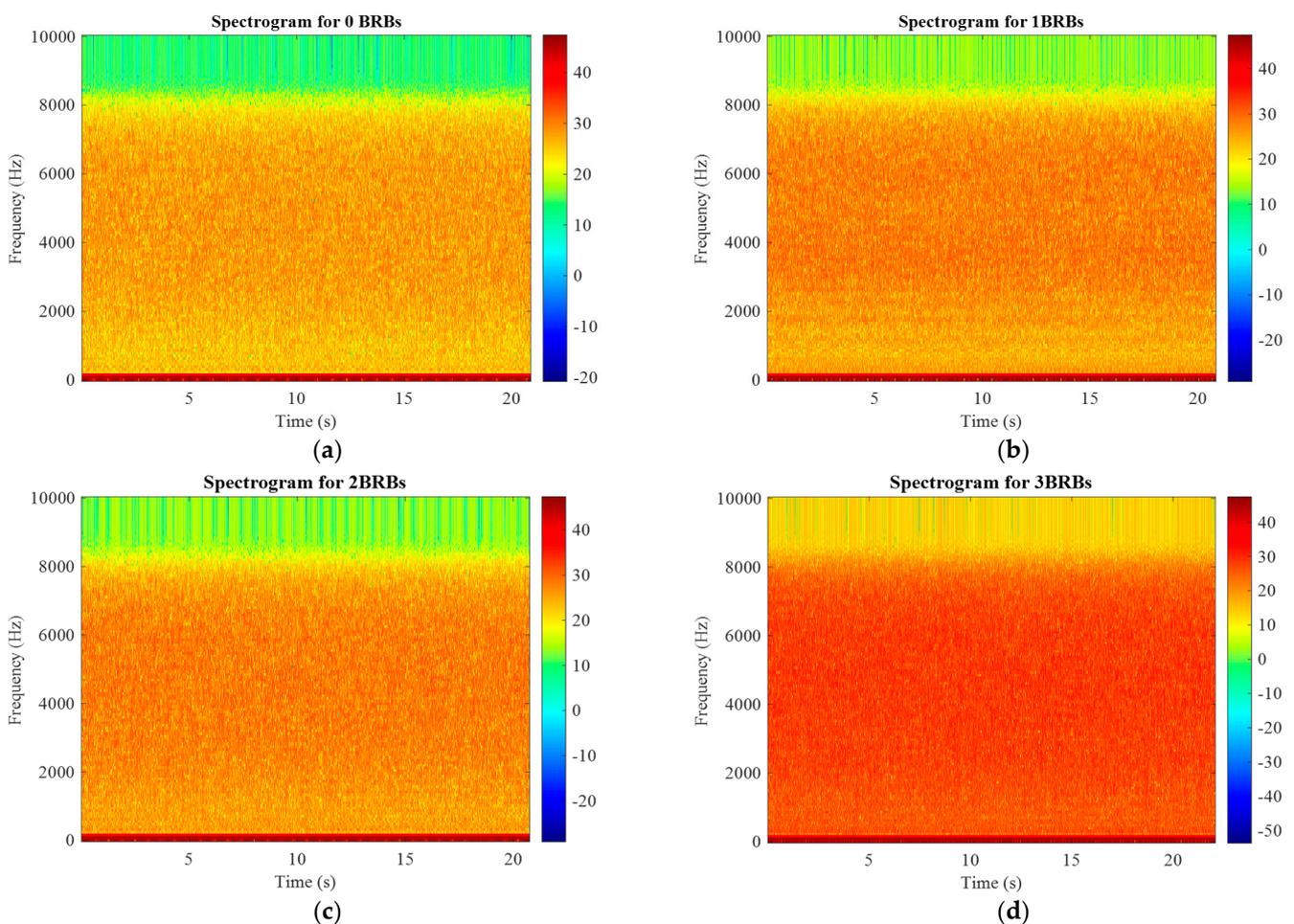
Figure 9. Spectrograms for (**a**) healthy, (**b**) 1 BRBs, (**c**) 2 BRBs, and (**d**) 3 BRBs broken rotor bars at 100% loadings conditions.

The suggested diagnostic framework (SSAE–LightGBM) incorporates a stratified feature-based cross-validation approach applied to the broken rotor bars dataset obtained from the test setup. The dataset was divided into folds, preserving the proportion of relevant features in each fold. This strategy facilitates a more accurate and comprehensive assessment of model performance, particularly beneficial for unbalanced or critical datasets. Leveraging the advanced SSAE–LightGBM method alongside strict stratified feature-based cross-validation enables thorough diagnostic investigations. The algorithm underwent fine-tuning through extensive experiments conducted on data generated at

various loading levels (0%, 25%, 50%, 75%, and 100%). The effectiveness with which the model was able to diagnose rotor bar health states is represented visually via the confusion matrices. Confusion matrices provide an evaluation of the performance of a classifier across different health states. The classifier's ability to resolve between rotor bars that were healthy and those that were damaged to various degrees is afforded by the comprehensive analysis this provides. It effectively maps out the efficiency and adaptability of the diagnostic process. The study displays confusion matrices illustrating the performance of the diagnostic framework, sparse stacked autoencoder (SSAE), and LightGBM under varying load conditions ranging from 0% to 100% for both healthy motors and those with damaged rotor bars. The matrices present the classification results, demonstrating the accuracy of our model in distinguishing between healthy and malfunctioning motor conditions across different operational loads. The confusion matrices at various loading levels demonstrate the resilience of our diagnostic method across a broad spectrum of operational scenarios. SSAE and LightGBM can effectively identify and classify BRB problems at various load levels, demonstrating the model's reliability and efficiency in real-world industrial settings. In Figure 10, the results of the confusion matrices for healthy broken rotor bars at 0%, 25%, 50%, 75%, and 100% loading conditions are presented. As such, the model's performance is depicted in different cases.
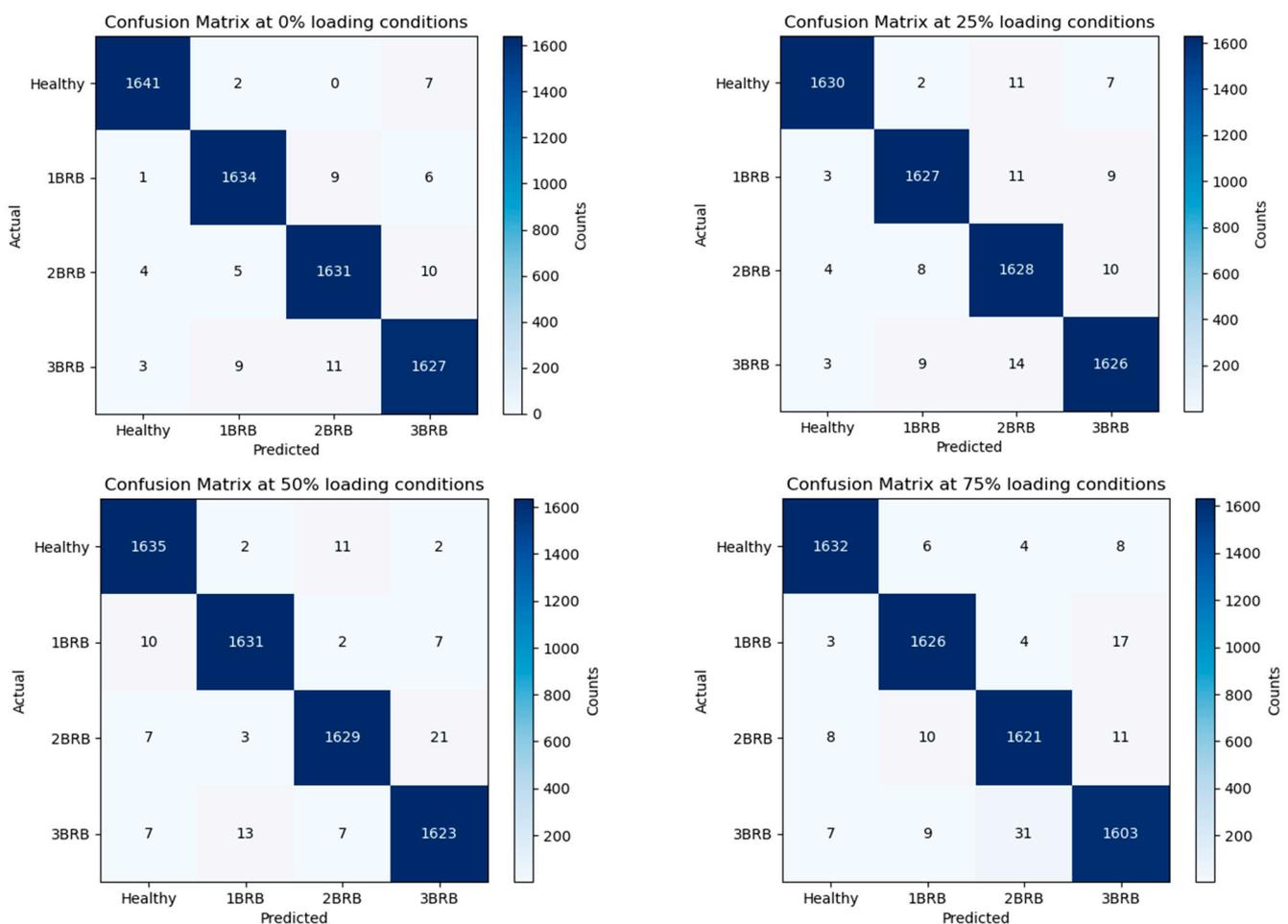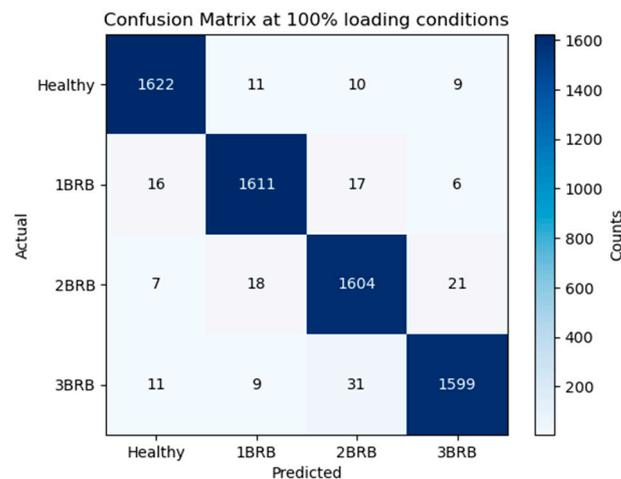

Figure 10. *Cont.*

Figure 10. Testing confusion matrices obtained from the proposed architecture (SSAE–LightGBM) for diagnosis of healthy and abnormal conditions in BRBs under various loading conditions.

Additionally, ROC curves beyond the confusion matrices offer nuanced insights regarding the discriminatory ability of the model. These illustrate the trade-off between true positives and false positives across multiple loading conditions, revealing the model's ability to discern BRB anomalies at differing operational thresholds. These ROC curves further reveal the model's capacity to discern accurately rotor bar faults over numerous loading conditions, presented in Figure 11 for comparison of abnormal conditions due to broken rotor bars in multiple loading conditions, providing further insights based on several operating scenarios.

The given Tables 5–8 provide diagnostic performance parameters for undamaged rotor bars and different fault levels under different loading circumstances. Accuracy, precision, recall, F1-score, and processing time are performance measurements. Each table shows fault scenarios with undamaged rotor bars and one, two, or three damaged bars. The diagnostic framework detects faults accurately under diverse loading conditions. Precision, recall, and F1-score measures show the framework's ability to detect and categorize rotor bar faults, with slight variances based on fault severity and operational load. The processing time also indicates the computational efficiency of the diagnostic procedure. The tables show the diagnostic framework's reliability and effectiveness in industrial fault diagnosis applications. Increased motor load can impact diagnostic accuracy in several ways. Excessive loads might mask trouble indicators, reducing precision due to noise or interference. Increased loads can reduce precision and recall as a result of signal intricacy and fault signature concealment. The F1-score, which considers both precision and recall, might vary based on loading conditions, reflecting the efficiency of the diagnostic system. Signal analysis necessitates additional computational resources, hence larger workloads might lead to longer processing times. Thorough testing is necessary to guarantee the reliability of industrial applications as fault detection algorithms need to be tuned to perform consistently under varying loading conditions. To develop effective defect detection algorithms and assure the reliability of motor systems in practical applications, it is crucial to comprehend these implications. Tables 5–8 provide the diagnostic performance measures of accuracy, precision, recall, F1-score, and performance time (that the proposed algorithm (SSAE–LightGBM) takes to predict the outcomes) for healthy and faulty BRBs at varying degrees under different loading situations.
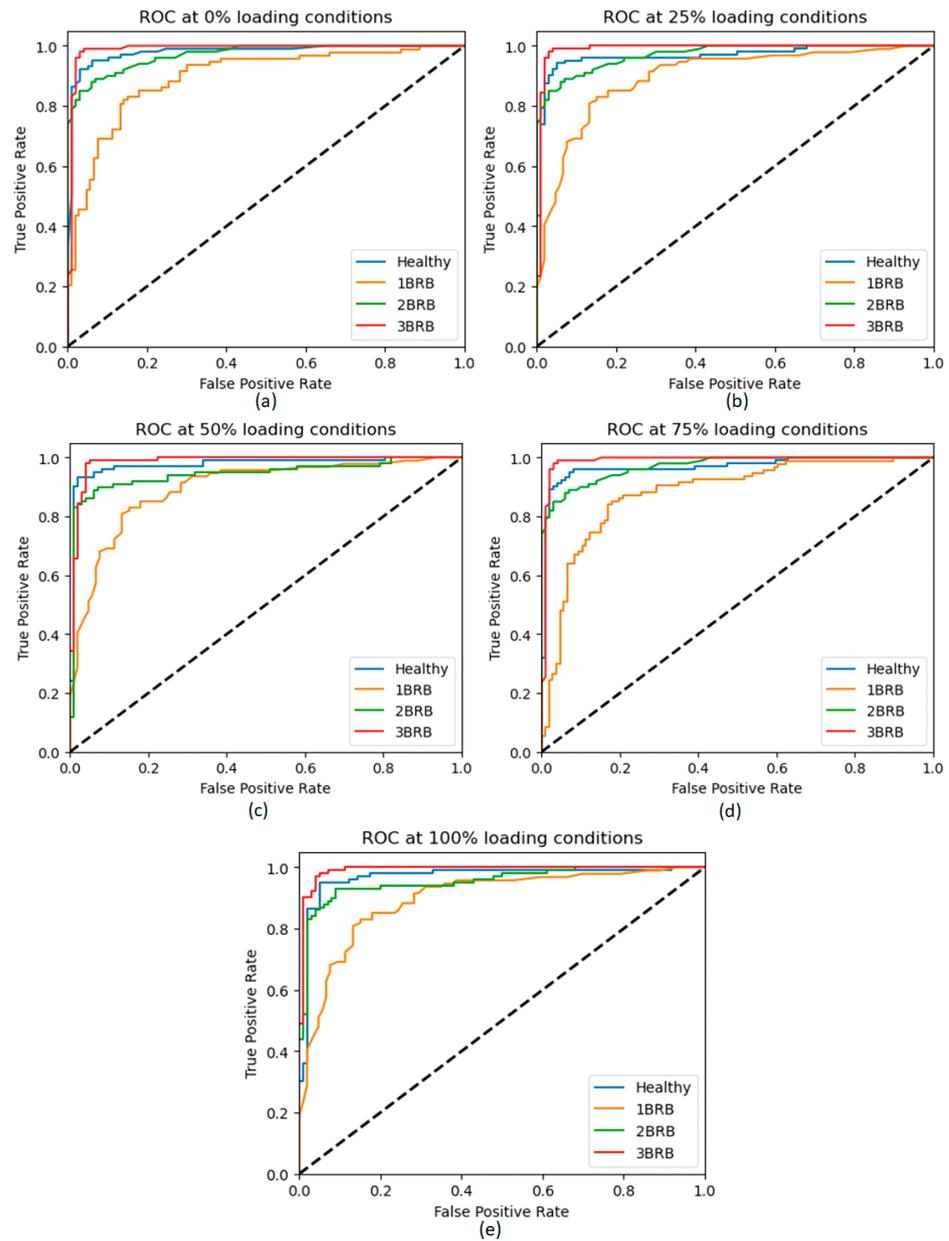
Figure 11. ROC curves from the proposed architecture (SSAE–LightGBM) for BRB abnormality diagnosis under varied loading circumstances (**a**) at 0% loading conditions (**b**) at 25% loading conditions (**c**) at 50% loading conditions (**d**) at 75% loading conditions (**e**) at 100% loading conditions.

Table 5. Diagnostic performance measures for healthy rotor bars.

Loading Condition	Accuracy	Precision	Recall	F1-Score	P.Time (s)
0%	0.9974	0.9951	0.9940	0.9930	211.973
25%	0.9951	0.9903	0.9878	0.9860	213.384
50%	0.9945	0.9903	0.9884	0.9860	234.345
75%	0.9948	0.9903	0.9881	0.9860	235.234
100%	0.9786	0.9876	0.9725	0.9713	287.234

Table 6. Diagnostic performance measures for one broken rotor bars.

Loading Condition	Accuracy	Precision	Recall	F1-Score	P.Time (s)
0%	0.9954	0.9845	0.9857	0.9845	211.973
25%	0.9664	0.9779	0.9668	0.9712	213.384
50%	0.9780	0.9967	0.9897	0.9723	234.345
75%	0.986	0.9932	0.9841	0.9756	235.234
100%	0.977	0.9823	0.9721	0.9701	287.234

Table 7. Diagnostic performance measures for two broken rotor bars.

Loading Condition	Accuracy	Precision	Recall	F1-Score	P.Time (s)
0%	0.9934	0.9841	0.9845	0.983	211.973
25%	0.9653	0.9745	0.9653	0.9623	213.384
50%	0.9735	0.9954	0.9823	0.9734	234.345
75%	0.9843	0.9922	0.9867	0.9789	235.234
100%	0.974	0.9842	0.9767	0.9711	287.234

Table 8. Diagnostic performance measures for three broken rotor bars.

Loading Condition	Accuracy	Precision	Recall	F1-Score	P.Time (s)
0%	0.9833	0.9863	0.9823	0.9827	211.973
25%	0.9623	0.973	0.964	0.9721	213.384
50%	0.9785	0.9923	0.9835	0.9775	234.657
75%	0.9832	0.9956	0.9845	0.9735	235.234
100%	0.9786	0.9821	0.9723	0.9735	287.234

6. Comparison with Previous Works

Table 9 offers an insightful comparison of various methods utilized to detect defects, focusing particularly on diagnosing BRBs in induction machines. Each row corresponds to a unique study, outlining the approach taken for feature extraction, the types of faults examined, and the achieved accuracy. The method proposed in this study employs sparse stacked autoencoder (SSAE) for feature extraction and LightGBM as a classifier, combined with stratified feature-based cross-validation. Remarkably, this approach achieves an impressive 99% accuracy rate when applied to healthy BRBs of varying types under different loading conditions.

Table 9. Analysis of the outcomes and distinctive attributes provided by the proposed study and past methodologies.

References	Suggested Approach	Fault type	Accuracy
[9]	Homogeneity analysis is used for feature extraction and a Gaussian probability is used as a classifier.	Half one and two BRBs	99%
[11]	Features are extracted using the MUSIC methodology and classified using the Bayes method.	One and two BRBs	100%
[13]	Fractal dimensions are used for feature extraction and a fuzzy logic classifier is employed.	Half one and two BRBs	95%

Table 9. *Cont.*

References	Suggested Approach	Fault type	Accuracy
[15]	The Wavelet and Hilbert transformations are employed to extract characteristics and the classifier is enabled by the linear discriminant technique.	One and two BRBs	100%
[19]	The extended Kalman filter MUSIC methodology is used for feature extraction and the MUSIC method is applied as the classifier.	Half and one BRB	100%
[21]	Signals are converted into pictures by the wavelet transform and a CNN is employed as both a classifier and a feature estimator.	Three BRBs	99%
[23]	Features are extracted using the Wavelet transform and the Pearson'scorrelation is used as the classifier.	Half one and two BRBs	99%
[27]	The Hilbert transform is used for feature extraction, and the Gaussian probability density function is used as the classifier.	Half one and two BRBs	98%
[28]	The signals are transformed into pictures using the Fourier transform and a CNN is used for feature estimation and classification.	Half one and two BRBs	100%
[29]	Frequency and time domain characteristics are utilized to identify features, with an Artificial Neural Network (ANN) employed as the classifier.	Healthy, three and six BRBs	95%
[30]	Six CNN architectures, including VGG16, NasNETMobile, Inception V4, ResNET152, VGG19, and SeNET154, are employed for diagnosing BRBs.	One, two, three, four, five and six BRBs	99%
[31]	The convolution layer is employed for feature extraction through a light gradient boosting machine and is also used as the classifier.	One and two BRBs	97.42%
[31]	A CNN is utilized to extract features and LightGBMis used as the classifier.	One and two BRBs	99.7%
Proposed work	For feature extraction, a sparse stacked autoencoder is employed, followed by LightGBM with stratified feature-based cross-validation.	Healthy, one, two, and three BRBs at varying loading conditions	Provides accuracy up to 99%

7. Conclusions

The paper proposes an advanced diagnostic approach for identifying and classifying broken rotor bars (BRBs) in inverter-fed induction motors. This method combines a sparse stacked autoencoder (SSAE) with LightGBM, a gradient-boosting framework, to analyze crucial motor operating data features such as current, voltage, speed, and torque signatures. A unique stratified group k-fold cross-validation technique is introduced to maintain the distribution of class labels and grouping variables across folds, enhancing the model's performance. SSAE is employed for feature extraction, automatically learning hierarchical representations of input data and identifying critical patterns, while LightGBM highlights important BRB features using the extracted traits. Extensive datasets exclusively sourced

from inverter-powered motor systems operating at various loading percentages are utilized for training and testing. Evaluation metrics including vibrations, spectrograms, confusion matrices, and receiver operating characteristic curves are employed to assess the model's performance. Bayesian optimization is utilized to fine-tune LightGBM hyperparameters, and the optimal feature subset is used for final classification. The experimental results demonstrate promising accuracy ranging from 91.55% to 99% across different loading scenarios, demonstratingthe efficacy of the proposed BRB diagnostic framework. Overall, the integration of advanced techniques such asSSAE and LightGBM enhances the robustness and effectiveness of the diagnostic architecture, offering significant improvements to BRB detection and classification in induction motors.

References

1. Ayman, E.R. Toward a sustainable more electrified future: The role of electrical machines and drives. *IEEE Electrif. Mag.* **2019**, *7*, 49–59.
2. Pawlus, W.; Birkeland, J.T.; Van Khang, H.; Hansen, M.R. Identification and experimental validation of an induction motor thermal model for improved drivetrain design. *IEEE Trans. Ind. Appl.* **2017**, *53*, 4288–4297. [CrossRef]
3. Merizalde, Y.; Hernández-Callejo, L.; Duque-Perez, O. State of the art and trends in the monitoring, detection and diagnosis of failures in electric induction motors. *Energies* **2017**, *10*, 1056. [CrossRef]
4. Choudhary, A.; Goyal, D.; Shimi, S.L.; Akula, A. Condition monitoring and fault diagnosis of induction motors: A review. *Arch. Comput. Methods Eng.* **2019**, *26*, 1221–1238. [CrossRef]
5. Gan, C.; Chen, Y.; Qu, R.; Yu, Z.; Kong, W.; Hu, Y. An overview of fault-diagnosis and fault-tolerance techniques for switched reluctance machine systems. *IEEE Access* **2019**, *7*, 174822–174838. [CrossRef]
6. Messaoudi, M.; Flah, A.; Alotaibi, A.A.; Althobaiti, A.; Sbita, L.; Ziad El-Bayeh, C. Diagnosis and fault detection of rotor bars in squirrel cage induction motors using combined Park's vector and extended Park's vector approaches. *Electronics* **2022**, *11*, 380. [CrossRef]
7. Ojaghi, M.; Sabouri, M.; Faiz, J. Performance analysis of squirrel-cage induction motors under broken rotor bar and stator inter-turn fault conditions using analytical modeling. *IEEE Trans. Magn.* **2018**, *54*, 8203705. [CrossRef]
8. Bradiya, J.R.; Jarial, R.K.; Sood, Y.R. Utilization of Signal Analysis Techniques in Investigation of Rotor Broken Bars and Misalignment Defects on low power Induction Machines. In Proceedings of the 2022 2nd Asian Conference on Innovation in Technology (ASIANCON), Ravet, India, 26–28 August 2022; pp. 1–9.
9. Li, H.; Feng, G.; Zhen, D.; Gu, F.; Ball, A.D. A normalized frequency-domain energy operator for broken rotor bar fault diagnosis. *IEEE Trans. Instrum. Meas.* **2020**, *70*, 3500110. [CrossRef]
10. Zaman, S.M.K.; Marma, H.U.M.; Liang, X. Broken rotor bar fault diagnosis for induction motors using power spectral density and complex continuous wavelet transform methods. In Proceedings of the 2019 IEEE Canadian Conference of Electrical and Computer Engineering (CCECE), Edmonton, AB, Canada, 5–8 May 2019; pp. 1–4.
11. Saidi, L.; Fnaiech, F.; Henao, H.; Capolino, G.-A.; Cirrincione, G. Diagnosis of broken-bars fault in induction machines using higher order spectral analysis. *ISA Trans.* **2013**, *52*, 140–148. [CrossRef] [PubMed]
12. Bessam, B.; Menacer, A.; Boumehraz, M.; Cherif, H. Detection of broken rotor bar faults in induction motor at low load using neural network. *ISA Trans.* **2016**, *64*, 241–246. [CrossRef]
13. Valtierra-Rodriguez, M.; Rivera-Guillen, J.R.; Basurto-Hurtado, J.A.; De-Santiago-Perez, J.J.; Granados-Lieberman, D.; Amezquita-Sanchez, J.P. Convolutional neural network and motor current signature analysis during the transient state for detection of broken rotor bars in induction motors. *Sensors* **2020**, *20*, 3721. [CrossRef] [PubMed]
14. Halder, S.; Bhat, S.; Dora, B.K. Inverse thresholding to spectrogram for the detection of broken rotor bar in induction motor. *Measurement* **2022**, *198*, 111400. [CrossRef]
15. Mhaskar, P.; Gani, A.; El-Farra, N.H.; Christofides, P.D.; Davis, J.F. Integrated fault-detection and fault-tolerant control of process systems. *AIChE J.* **2006**, *52*, 2129–2148. [CrossRef]

16. Barrera-Llanga, K.; Burriel-Valencia, J.; Sapena-Bañó, Á.; Martínez-Román, J. A Comparative Analysis of Deep Learning Convolutional Neural Network Architectures for Fault Diagnosis of Broken Rotor Bars in Induction Motors. *Sensors* **2023**, *23*, 8196. [CrossRef] [PubMed]

17. Zhu, Y.; Xie, B.; Wang, A.; Qian, Z. Fault diagnosis of wind turbine gearbox under limited labeled data through temporal predictive and similarity contrast learning embedded with self-attention mechanism. *Expert Syst. Appl.* **2023**, *245*, 123080. [CrossRef]

18. Edomwandekhoe, K.; Liang, X. Advanced feature selection for broken rotor bar faults in induction motors. In Proceedings of the 2018 IEEE/IAS 54th Industrial and Commercial Power Systems Technical Conference (I&CPS), Niagara Falls, ON, Canada, 7–10 May 2018; pp. 1–10.

19. Liboni, L.H.B.; Flauzino, R.A.; da Silva, I.N.; Costa, E.C.M. Efficient feature extraction technique for diagnosing broken bars in three-phase induction machines. *Measurement* **2019**, *134*, 825–834. [CrossRef]

20. Zhang, S.; Zhang, S.; Wang, B.; Habetler, T.G. Deep learning algorithms for bearing fault diagnostics—A compre-hensive review. *IEEE Access* **2020**, *8*, 29857–29881. [CrossRef]

21. Hakim, M.; Omran, A.A.B.; Ahmed, A.N.; Al-Waily, M.; Abdellatif, A. A systematic review of rolling bearing fault diagnoses based on deep learning and transfer learning: Taxonomy, overview, application, open challenges, weaknesses and recommendations. *Ain Shams Eng. J.* **2023**, *14*, 101945. [CrossRef]

22. Shao, S.; McAleer, S.; Yan, R.; Baldi, P. Highly accurate machine fault diagnosis using deep transfer learning. *IEEE Trans. Ind. Inform.* **2018**, *15*, 2446–2455. [CrossRef]

23. Al-Stouhi, S.; Reddy, C.K. Transfer learning for class imbalance problems with inadequate data. *Knowl. Inf. Syst.* **2016**, *48*, 201–228. [CrossRef] [PubMed]

24. Drakaki, M.; Karnavas, Y.L.; Karlis, A.D.; Chasiotis, I.D.; Tzionas, P. Study on fault diagnosis of broken rotor bars in squirrel cage induction motors: A multi-agent system approach using intelligent classifiers. *IET Power Appl.* **2020**, *14*, 245–255. [CrossRef]

25. Karvelis, P.; Georgoulas, G.; Tsoumas, I.P.; Antonino-Daviu, J.A.; Climente-Alarcón, V.; Stylios, C.D. A symbolic representation approach for the diagnosis of broken rotor bars in induction motors. *IEEE Trans. Ind. Inform.* **2015**, *11*, 1028–1037. [CrossRef]

26. Arabaci, H.; Mohamed, M.A. A knowledge-based diagnosis algorithm for broken rotor bar fault classification using FFT, principal component analysis and support vector machines. *Int. J. Intell. Eng. Inform.* **2020**, *8*, 19. [CrossRef]

27. Aydin, I.; Karakose, M.; Akin, E. A new method for early fault detection and diagnosis of broken rotor bars. *Energy Convers. Manag.* **2011**, *52*, 1790–1799. [CrossRef]

28. Defdaf, M.; Berrabah, F.; Chebabhi, A.; Cherif, B.D.E. A new transform discrete wavelet technique based on artificial neural network for induction motor broken rotor bar faults diagnosis. *Int. Trans. Electr. Energy Syst.* **2021**, *31*, e12807. [CrossRef]

29. Abdel-Basset, M.; Hawash, H.; Chang, V.; Chakrabortty, R.K.; Ryan, M. Deep learning for heterogeneous human activity recognition in complex IoT applications. *IEEE Internet Things J.* **2020**, *9*, 5653–5665. [CrossRef]

30. Sun, G.; Zhang, X.; Jia, X.; Ren, J.; Zhang, A.; Yao, Y.; Zhao, H. Deep fusion of localized spectral features and multi-scale spatial features for effective classification of hyperspectral images. *Int. J. Appl. Earth Obs. Geoinf.* **2020**, *91*, 102157. [CrossRef]

31. Deng, F.; Ding, H.; Yang, S.; Hao, R. An improved deep residual network with multiscale feature fusion for rotating machinery fault diagnosis. *Meas. Sci. Technol.* **2020**, *32*, 024002. [CrossRef]

32. Wang, S.; Lei, Y.; Lu, N.; Li, X.; Yang, B. A multi-sensor relation model for recognizing and localizing faults of machines based on network analysis. *Front. Mech. Eng.* **2023**, *18*, 20. [CrossRef]

33. Kumar, P.; Hati, A.S. Deep convolutional neural network based on adaptive gradient optimizer for fault detection in SCIM. *ISA Trans.* **2021**, *111*, 350–359. [CrossRef] [PubMed]

34. Garcia-Calva, T.A.; Morinigo-Sotelo, D.; Fernandez-Cavero, V.; Garcia-Perez, A.; Romero-Troncoso, R.d.J. Early detection of broken rotor bars in inverter-fed induction motors using speed analysis of startup transients. *Energies* **2021**, *14*, 1469. [CrossRef]

35. Romero-Troncoso, R.; Garcia-Perez, A.; Morinigo-Sotelo, D.; Duque-Perez, O.; Osornio-Rios, R.; Ibarra-Manzano, M. Rotor unbalance and broken rotor bar detection in inverter-fed induction motors at start-up and steady-state regimes by high-resolution spectral analysis. *Electr. Power Syst. Res.* **2016**, *133*, 142–148. [CrossRef]

36. Gyftakis, K.N.; Spyropoulos, D.V.; Arvanitakis, I.; Panagiotou, P.A.; Mitronikas, E.D. Induction motors torque analysis via frequency extraction for reliable broken rotor bar detection. In Proceedings of the 2020 International Conference on Electrical Machines (ICEM), Gothenburg, Sweden, 23–26 August 2020; Volume 1, pp. 1468–1474.

37. Ghorbani, A.; Fakhrahmad, S.M. A deep learning approach to network intrusion detection using a proposed su-pervised sparse auto-encoder and svm. Iranian Journal of Science and Technology. *Trans. Electr. Eng.* **2022**, *46*, 829–846.

38. Mushtaq, E.; Zameer, A.; Nasir, R. Knacks of a hybrid anomaly detection model using deep auto-encoder driven gated recurrent unit. *Comput. Netw.* **2023**, *226*, 109681. [CrossRef]

39. Lee, S.-W.; Sidqi, H.M.; Mohammadi, M.; Rashidi, S.; Rahmani, A.M.; Masdari, M.; Hosseinzadeh, M. Towards secure intrusion detection systems using deep learning techniques: Comprehensive analysis and review. *J. Netw. Comput. Appl.* **2021**, *187*, 103111. [CrossRef]

40. Lao, Z.; He, D.; Wei, Z.; Shang, H.; Jin, Z.; Miao, J.; Ren, C. Intelligent fault diagnosis for rail transit switch machine based on adaptive feature selection and improved LightGBM. *Eng. Fail. Anal.* **2023**, *148*, 107219. [CrossRef]

41. Zhang, C.; Kong, L.; Xu, Q.; Zhou, K.; Pan, H. Fault diagnosis of key components in the rotating machinery based on Fourier transform multi-filter decomposition and optimized LightGBM. *Meas. Sci. Technol.* **2020**, *32*, 015004. [CrossRef]

42. Wang, M.; Shen, K.; Tai, C.; Zhang, Q.; Yang, Z.; Guo, C. Research on fault diagnosis system for belt conveyor based on internet of things and the LightGBM model. *PLoS ONE* **2023**, *18*, e0277352. [CrossRef] [PubMed]