modelling

MDPI

Article

Rotor Imbalance Classification in Wind Turbines Using Multichannel Vibration Analysis and a DWT–LDA Framework

Oscar H. Sierra-Herrera [1], Mario Eduardo González Niño [2,3], Carlos E. Pinto-Salamanca [2], Wilman Alonso Pineda Muñoz [2] and Jersson X. Leon-Medina [2,*]

1 Grupo de Investigación I2E, Escuela de Ingeniería Electrónica, Facultad de Ingeniería, Universidad Pedagógica y Tecnológica de Colombia (UPTC), Avenida Central del Norte 39–115, Tunja 150003, Colombia; oscarhumberto.sierra@uptc.edu.co

2 Escuela de Ingeniería Electromecánica, Facultad Seccional Duitama, Universidad Pedagógica y Tecnológica de Colombia (UPTC), Carrera 18 con Calle 22, Duitama 150461, Colombia; marioeduardo.gonzalez@uptc.edu.co or mario.eduardo.gonzalez@upc.edu (M.E.G.N.); carlos.pinto@uptc.edu.co (C.E.P.-S.); wilman.pineda@uptc.edu.co (W.A.P.M.)

3 Control, Data, and Artificial Intelligence (CoDAlab), Department of Mathematics, Escola d'Enginyeria de Barcelona Est (EEBE), Universitat Politècnica de Catalunya (UPC), Campus Diagonal–Besòs, Eduard Maristany 16, 08019 Barcelona, Spain

* Correspondence: jersson.leon@uptc.edu.co

Abstract

Wind turbines are critical components in renewable energy systems, where early fault detection is essential to ensure reliable operation and reduce maintenance costs. Vibration-based monitoring using multichannel signals provides rich information about the dynamic behavior of the system, although it also introduces challenges related to high dimensionality and feature redundancy. This paper proposes a machine learning-based methodology for fault classification that combines Discrete Wavelet Transform (DWT) for time–frequency feature extraction with Linear Discriminant Analysis (LDA) for dimensionality reduction within a structured processing pipeline. The approach incorporates a Group K-Fold cross-validation strategy to prevent data leakage and ensure a reliable evaluation when working with segmented signals. Experimental results show that the proposed framework achieves high classification performance, reaching a mean accuracy of $98.84 \pm 1.16\%$ and a weighted F1-score of 0.9905 ± 0.0089 using a Support Vector Machine (SVM) classifier over five Group K-Fold splits. The results also indicate that dimensionality reduction plays a critical role in improving class separability, having a greater impact than the specific choice of wavelet transform. Findings demonstrate that the proposed DWT–LDA-based approach provides an effective solution for rotor imbalance detection in the laboratory-scale wind turbine evaluated in this study.

Keywords: wind turbine; fault diagnosis; vibration signals; discrete wavelet transform (DWT); linear discriminant analysis (LDA); machine learning; condition monitoring; feature extraction

Check for updates

Academic Editors: Margarita Terziyska and Miroslava Ivanova

Received: 19 May 2026
Revised: 26 June 2026
Accepted: 1 July 2026
Published: 7 July 2026

1. Introduction

Wind energy has been established as a sustainable and low-emission technology that plays a significant role in the energy transition. In recent years, the installed wind capacity in the world has exceeded 1100 GW [1,2]. However, the development of reliable, larger and more complex wind turbines, that are exposed to variable loads and demanding environmental conditions, is still a challenge [3]. Unexpected failures reduce availability, affect energy production, and increase maintenance costs, which may account for 20% to 30% of the cost of electricity in wind energy systems [2,4]. Therefore, condition monitoring and predictive maintenance are essential, although their implementation remains challenging when dealing with non-stationary, multichannel, and high-dimensional signals.

In wind turbines, the rotor, blades, bearings, gearbox, and drivetrain are critical components because they are exposed to aerodynamic loads, environmental effects, flow-induced vibrations, and degradation [3]. In particular, rotor imbalance modifies the dynamic response, increases vibration levels, and accelerates mechanical deterioration. Several studies have investigated blade and rotor faults using dynamic signals, including experimental datasets with cracks, surface erosion, mass imbalance, and blade twist [5–9]; unsupervised rotor imbalance detection using multichannel vibrations from a 3 kW Enair E30Pro wind turbine [6]; intelligent diagnosis using easy-to-measure electrical signals [10]; speed-invariant diagnostic features under variable operating conditions [11]; and detection, classification, and localization of rotor imbalances using drivetrain and SCADA signals [12,13]. These studies confirm the relevance of rotor imbalance and multi-sensor monitoring; however, they do not address compact multiclass classification based on multichannel vibration signals. SCADA systems allow the monitoring of operational, electrical, environmental, and thermal variables using already installed sensors [2]. Nevertheless, their low sampling frequency or storage as time-averaged measurements may limit the capture of transient phenomena and faults [2]. In contrast, high-resolution vibration signals provide more direct information about the dynamics of rotating components, although they are often non-stationary because of variable loads and operating conditions [14,15]. In this context, wavelet-based techniques are suitable for representing time–frequency information, since they analyze signals at different scales and extract localized transient frequency components [14]. The Discrete Wavelet Transform has been applied together with FFT for vibration-based blade fault detection [9], whereas other studies have used the Continuous Wavelet Transform and deep learning with time–frequency images [8]. In addition, recent networks based on multi-kernel wavelets, adaptive denoising, and contrastive learning have been proposed for wind turbine gearbox fault diagnosis [16–19]. However, several of these approaches rely on image-based representations, deep architectures, or more complex training procedures [8,20,21].

Artificial intelligence methods have shown strong potential for wind turbine fault diagnosis [22–24], including CNN [8], LSTM [25], SVM [22], ANN [7], and hybrid approaches applied to blades [23], bearings [26], generators [20], and gearboxes [21,22].

Lightweight spectrogram-based models [27], semi-supervised frameworks for blade surface defect detection using drone images [28], and multisensor fusion models with cross-attention and Dempster–Shafer theory for progressive crack detection using vibration signals [29] have also been proposed. In the drivetrain, studies have reported fault diagnosis in real 2 MW wind turbines through time- and frequency-domain vibration analysis [30], as well as approaches combining vibration indicators, SCADA data, and normal behavior models [31]; Sparse Autoencoder, KPCA, SVM, and Random Forest for gearbox faults [32]; SVM, Naive Bayes, and KNN with statistical vibration features [33]; VMD, envelope analysis, FFT, band energy, and KNN for bearing faults [34]; and compressive sensing, recurrence plots, and SqueezeNet for bearing and gear faults [35]. These studies demonstrate the usefulness of vibration signals and machine learning, but their focus is not directly on multi-sensor signals or rotor imbalance level classification using wavelet-based representation and discriminant projection.

In addition to feature extraction, dimensionality reduction is crucial for multichannel vibration signals, since multiple sensors increase dimensionality, redundancy, and computational complexity [36]. Linear Discriminant Analysis projects data into a lower-dimensional space aimed at maximizing class separation. The combination of wavelet representations with discriminant techniques has improved separability and efficiency in bearing and motor fault diagnosis [37–39]. However, in wind turbine rotor imbalance diagnosis, it remains necessary to evaluate how discriminant projection improves the representation obtained from multichannel vibration signals [6,37].

Recent advances in wind turbine condition monitoring have increasingly exploited time–frequency analysis to improve the diagnosis of mechanical faults from non-stationary vibration signals. Wavelet-based representations have proven particularly effective because they simultaneously capture temporal and spectral characteristics associated with fault evolution. Early studies demonstrated the potential of wavelet packet decomposition combined with deep learning for gearbox fault diagnosis, enabling multi-scale feature extraction from vibration signals [19]. More recently, Li and Gao [40] transformed nacelle vibration signals into time–frequency representations using wavelet analysis and employed convolutional neural networks to classify rotor imbalance conditions. Similarly, Guo et al. [41] proposed a dual-channel deep learning framework integrating continuous wavelet transform, attention mechanisms, and multi-scale feature aggregation to improve diagnostic performance under variable-speed and noisy operating conditions. Collectively, these studies demonstrate that time–frequency representations provide rich discriminative information for wind turbine fault diagnosis, particularly when combined with advanced learning algorithms.

Another important research direction has focused on exploiting information from multiple sensing channels to improve fault discrimination and robustness. Multi-channel vibration measurements allow complementary dynamic information from different structural locations to be incorporated into the diagnostic process, leading to more representative fault signatures than single-sensor approaches. Wang et al. [42] demonstrated that multi-channel convolutional neural networks can effectively exploit spatial correlations among vibration signals for wind turbine fault diagnosis. Likewise, Liang et al. [43] proposed a deep residual deformable subdomain adaptation framework capable of improving diagnostic performance across different operating domains, highlighting the importance of developing methods that remain reliable under varying working conditions. These advances indicate that combining information from multiple vibration measurements constitutes an effective strategy for enhancing fault diagnosis in complex rotating systems.

Despite these significant advances, several challenges remain regarding the development of reliable, interpretable, and computationally efficient diagnostic frameworks for wind turbine monitoring. Recent review studies have emphasized that robust condition monitoring systems must not only achieve high diagnostic accuracy but also demonstrate good generalization under realistic operating conditions while maintaining practical computational requirements [44,45]. Furthermore, the use of appropriate validation strategies has become increasingly important to avoid optimistic performance estimates caused by data leakage when multiple signal segments originate from the same experimental acquisition. Motivated by these challenges, the present work investigates a diagnosis framework that combines discrete wavelet transform (DWT) feature extraction, Linear Discriminant Analysis (LDA) for supervised dimensionality reduction, multichannel vibration measurements, and Group K-Fold cross-validation to provide an accurate and computationally efficient solution for rotor imbalance classification while ensuring a rigorous evaluation protocol.

There are several methodological challenges remaining. Many industrial datasets present class imbalance or difficulties in labeling early fault stages [22,46]. When signals are segmented into windows, samples originating from the same experiment may be highly correlated; if these windows are mixed between training and testing sets, information leakage and overly optimistic performance estimates may occur [24]. Therefore, preserving independence between training and testing is necessary [47].

To overcome these limitations, this work performs multiclass classification of rotor imbalance in wind energy systems using multichannel vibration signals. The proposed method combines the Discrete Wavelet Transform to extract time–frequency information, Linear Discriminant Analysis to reduce and compact the feature space, and supervised classification to distinguish different imbalance levels. The evaluation is performed using group-based cross-validation, ensuring that signal windows from the same experiment are not mixed between the training and testing sets, thereby preventing data leakage and providing a more realistic assessment of model generalization.

The remainder of this paper is organized as follows. Section 2 describes the materials and methods, including the dataset, the proposed methodology, signal processing, group-based cross-validation, and the experimental setup. Section 3 presents the results and discussion, considering the effect of the wavelet configuration, the impact of LDA, the feature-space analysis, and the performance of the classifiers. Finally, Section 4 summarizes the conclusions and outlines future research directions.

2. Materials and Methods

2.1. Dataset

Wind turbines are critical components in renewable energy generation, where structural integrity and fault detection play a key role in ensuring reliable operation. Among the different fault types, rotor imbalance in turbine blades is one of the most common and detrimental conditions, leading to increased vibration levels, mechanical stress, and reduced system lifetime. In this context, vibration-based condition monitoring has emerged as an effective approach for detecting such faults, as it provides non-invasive and information-rich measurements of the turbine's dynamic behavior.

The dataset used in this study was obtained from a laboratory-adapted wind turbine system (E30Pro, 3 kW, Enair, Alicante, Spain), instrumented for controlled experimental analysis [6]. To simulate realistic operating conditions, both healthy and faulty scenarios were considered. Rotor imbalance was introduced by adding calibrated masses to the turbine blades. The experimental setup is shown in Figure 1.

The dataset includes five operating conditions: (i) healthy condition (no added mass), (ii) imbalance level 1, (iii) imbalance level 2, (iv) imbalance level 3, and (v) imbalance level 4, representing increasing levels of rotor imbalance severity.

The turbine was equipped with eight triaxial accelerometers (356A17, PCB Piezotronics, Depew, NY, USA), resulting in 24 simultaneous vibration signals. Data acquisition was performed using (NI 9234, National Instruments, Austin, TX, USA), ensuring synchronized and high-resolution measurements across all channels.

The experimental campaign consisted of 35 independent tests conducted under controlled conditions. Each experiment was recorded over a 60 s duration at a sampling frequency of 1706.66 Hz, resulting in $N = 102{,}400$ samples per sensor. For each experiment, the acquired data are represented as $\mathbf{X}^{(i)}$,

$$\mathbf{X}^{(i)} \in \mathbb{R}^{N \times C} \tag{1}$$

where $i = 1, \ldots, 35$ denotes the experiment index and $C = 24$ is the number of channels.

Figure 2 presents representative vibration signals acquired from the X, Y, and Z axes of a triaxial accelerometer under healthy operating conditions. The figure illustrates the raw time-domain measurements and their corresponding frequency spectra, providing an overview of the characteristics of the acquired vibration data.

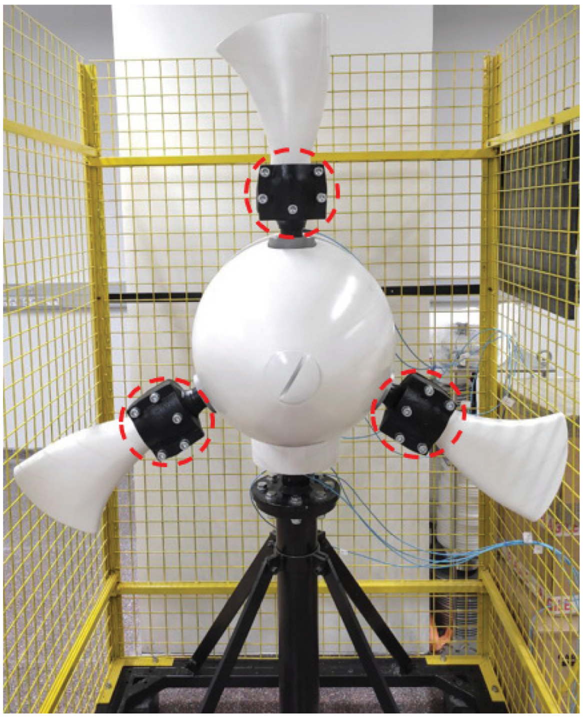


Figure 1. Experimental setup of the instrumented Enair E30Pro 3 kW wind turbine used for data acquisition. The red circles indicate the locations where additional masses were attached to the turbine blades to generate the different rotor imbalance conditions (reproduced from [6]).

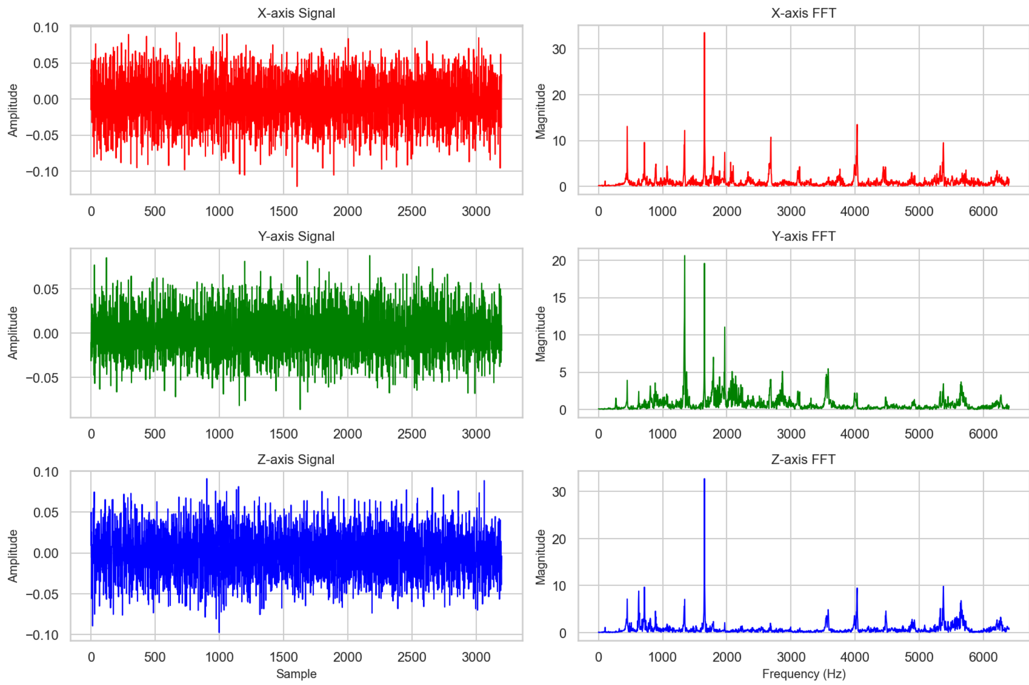


Figure 2. Time-domain vibration signals and corresponding FFT magnitude spectra for the three axes (X, Y, and Z) of a representative triaxial accelerometer under fault conditions. The left column shows the raw vibration signals, while the right column presents their frequency-domain representations obtained using the Fast Fourier Transform (FFT).

To enhance the extraction of meaningful temporal patterns, the signals are segmented into fixed-length windows of $W = 3200$ samples, corresponding to approximately 1.875 s. This window length is selected to capture relevant temporal dynamics associated with rotor imbalance.

2.2. Methodology

The proposed methodology, illustrated in Figure 3, follows a structured supervised learning workflow for rotor imbalance classification using multichannel vibration signals. The workflow starts with the acquired multichannel vibration signals, which are segmented into fixed-length windows. Each segmented multichannel window constitutes one sample for feature extraction and classification. Within each window, the DWT is independently applied to each of the 24 vibration channels, and the extracted subband energy features are concatenated into a single feature vector. Then, the Discrete Wavelet Transform (DWT) is applied to extract time–frequency features associated with the dynamic behavior of the system. These features are standardized using z-score normalization and subsequently projected into a lower-dimensional discriminative space through Linear Discriminant Analysis (LDA). Finally, supervised classification models are trained and evaluated using a Group K-Fold strategy. Feature normalization and LDA projection are fitted only with the training data within each cross-validation fold. In addition, all windows originating from the same experiment are kept within the same fold during training and testing. This strategy prevents data leakage and provides a more reliable performance assessment. The proposed methodology, illustrated in Figure 3, follows a structured supervised learning workflow for rotor imbalance classification using multichannel vibration signals. The workflow starts with the acquired multichannel vibration signals, which are segmented into fixed-length windows. Each segmented multichannel window constitutes one sample for feature extraction and classification. Within each window, the DWT is independently applied to each of the 24 vibration channels, and the extracted subband energy features are concatenated into a single feature vector. Then, the Discrete Wavelet Transform (DWT) is applied to extract time–frequency features associated with the dynamic behavior of the system. These features are standardized using z-score normalization and subsequently projected into a lower-dimensional discriminative space through Linear Discriminant Analysis (LDA). Finally, supervised classification models are trained and evaluated using a Group K-Fold strategy. Feature normalization and LDA projection are fitted only with the training data within each cross-validation fold. In addition, all windows originating from the same experiment are kept within the same fold during training and testing. This strategy prevents data leakage and provides a more reliable performance assessment.

2.3. Data Segmentation and Vectorization

Each signal $\mathbf{X}^{(i)}$ is divided into 32 non-overlapping windows of length $W = 3200$ samples:

$$M = \frac{N}{W} = \frac{102{,}400}{3200} = 32 \tag{2}$$

Each segment is defined as:

$$\mathbf{X}_k^{(i)} \in \mathbb{R}^{W \times C}, \quad k = 1, \ldots, M \tag{3}$$

where W denotes the window length, C the number of measurement channels, and M the total number of segments. The multichannel structure is preserved during feature extraction, and the Discrete Wavelet Transform (DWT) is applied independently to each channel.

In this work, the resulting dimensionality is:

$$W \cdot C = 3200 \times 24 = 76{,}800 \tag{4}$$

The segmentation and vectorization processes are illustrated in Figure 4.

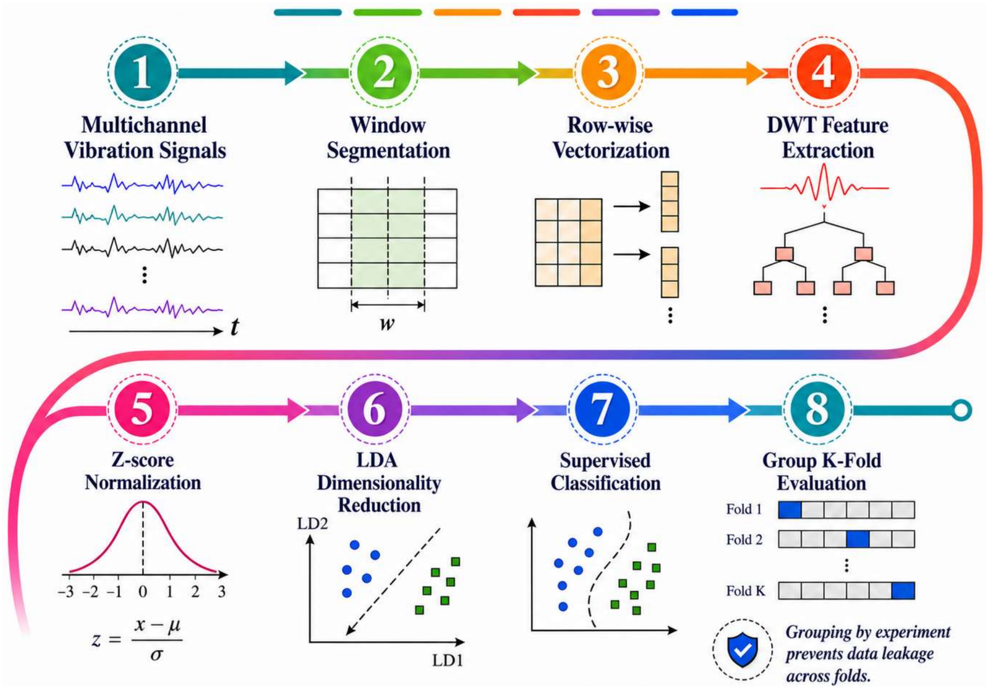


Figure 3. General workflow of the proposed supervised methodology for rotor imbalance classification using multichannel vibration signals.

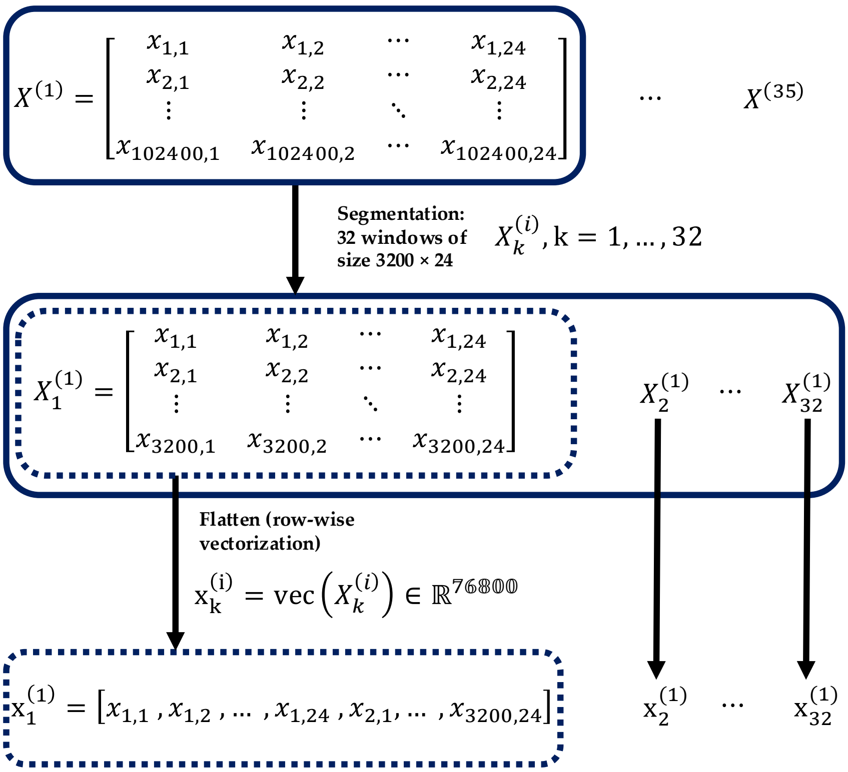


Figure 4. Segmentation and row-wise vectorization process used to transform each multichannel vibration window into a feature vector.

2.4. Labeling and Group Assignment

Each segmented window inherits the label of its corresponding experiment. Let $y^{(i)} \in \{0, 1, 2, 3, 4\}$ denote the class label, where 0 corresponds to the healthy condition and 1–4 indicate increasing imbalance levels, as described in the dataset.

Thus, each segment $\mathbf{X}_k^{(i)}$ is assigned:

$$y_k^{(i)} = y^{(i)} \quad (5)$$

To prevent data leakage, a grouping strategy is defined such that all segments originating from the same experiment are assigned to the same group:

$$g_k^{(i)} = i \quad (6)$$

This ensures that correlated samples remain within the same partition during cross-validation.

2.5. Cross-Validation Strategy

A 5-fold Group K-Fold cross-validation strategy is employed. Let the dataset be defined as

$$\mathbf{D} = \{\mathbf{x}_k^{(i)}, y_k^{(i)}, g_k^{(i)}\} \quad (7)$$

The dataset is partitioned into $K = 5$ folds such that all samples belonging to the same group are assigned to a single fold. During each iteration, a subset of groups (experiments) is used as the test set, while the remaining groups are used for training.

The data splitting strategy is illustrated in Figure 5, where samples from the same experiment are kept within the same fold to prevent data leakage.

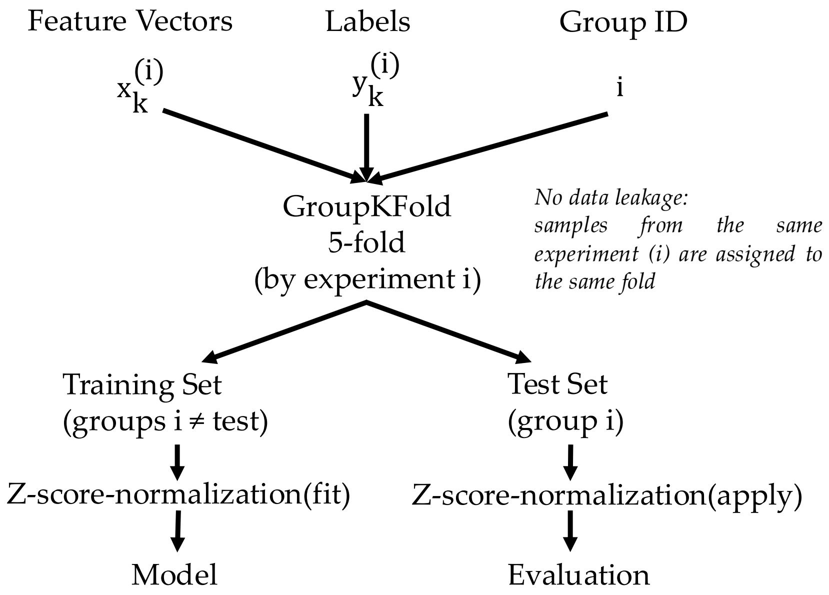


Figure 5. Group K-Fold cross-validation strategy.

2.6. Feature Normalization

Feature normalization is performed using z-score standardization. For each fold, the mean μ and standard deviation σ are computed from the training data and applied to both training and test sets:

$$\mathbf{x}' = \frac{\mathbf{x} - \mu}{\sigma} \quad (8)$$

This normalization is applied within each fold to prevent information leakage and is performed prior to dimensionality reduction.

2.7. Wavelet Transform and LDA

In this stage, feature extraction, dimensionality reduction, and classification are applied to each segmented signal $\mathbf{X}_k^{(i)}$.

2.7.1. Feature Extraction Using DWT

Each segment $\mathbf{X} * k^{(i)}$ consists of C channels and a window length of W samples. The Discrete Wavelet Transform (DWT) is applied independently to each channel signal $\mathbf{x} * k, c^{(i)}$, where $c = 1, \ldots, C$:

$$\mathbf{f} * k, c^{(i)} = \Phi(\mathbf{x} * k, c^{(i)}) \tag{9}$$

where $\Phi(\cdot)$ denotes the DWT. The DWT decomposes each channel signal into multiple frequency subbands, from which the energy of each subband is computed as:

$$E_{j,c} = \sum_n |c_{j,n}|^2 \tag{10}$$

where $c_{j,n}$ are the wavelet coefficients of the j-th subband for channel c. The feature vector $\mathbf{f}_k^{(i)}$ is constructed by concatenating the subband energies obtained from all channels. In the proposed configuration, each segmented multichannel window constitutes one sample for feature extraction and classification. Within each window, the DWT decomposition is independently applied to each of the 24 vibration channels, generating five subband energy features per channel. The extracted features are then concatenated into a single 120-dimensional feature vector (24 channels $\times$ 5 subbands), which is subsequently reduced to four discriminant components using LDA.

2.7.2. Dimensionality Reduction Using LDA

The extracted features are projected into a lower-dimensional space using Linear Discriminant Analysis (LDA):

$$\mathbf{z}_k^{(i)} = \mathbf{W}^T \mathbf{f}_k^{(i)} \tag{11}$$

where $\mathbf{W} \in \mathbb{R}^{d \times r}$ and $r = 4$. LDA is applied within each cross-validation fold using only the training data. The vibration signals are segmented into fixed-length windows of 3200 samples.

2.7.3. Classification

The reduced feature vectors $\mathbf{z}_k^{(i)}$ are used as inputs to classification models:

$$\hat{y}_k^{(i)} = \mathcal{C}(\mathbf{z}_k^{(i)}) \tag{12}$$

2.7.4. Processing Pipeline

The overall processing pipeline is summarized as

$$\mathbf{X}^{(i)} \to \mathbf{X}_k^{(i)} \to \mathbf{f}_k^{(i)} \to \mathbf{z}_k^{(i)} \to \hat{y}_k^{(i)} \tag{13}$$

where $\mathbf{f}_k^{(i)}$ denotes the feature vector obtained from the concatenated channel-wise DWT subband energies, and $\mathbf{z}_k^{(i)}$ the reduced feature representation.

2.8. Experimental Setup

The hyperparameters were selected based on standard values reported in the literature and empirical tuning. The selected parameters are listed in Table 1.

Table 1. Configurationof the proposed methodology and classification models.

Component	Configuration
Window size	3200 samples
Wavelet	sym5
Decomposition level	4
Feature type	Subband energy
LDA components	4
Cross-validation	Group K-Fold (5 splits)
SVM	kernel = RBF, C = 10, gamma = scale (scikit-learn default)
Random Forest	n_estimators = 150, random_state = 42
KNN	n_neighbors = 5
Logistic Regression	max_iter = 1000
Naive Bayes	default parameters

All experiments were implemented in Python 3.11 using Scikit-learn 1.7.2 and PyWavelets 1.8.0, together with NumPy 2.3.5, SciPy 1.16.3, Pandas 2.3.3, and Matplotlib 3.10.6. The computational experiments were executed on a MacBook Air M4 equipped with an Apple M4 processor and 8 GB of unified memory, running macOS Tahoe 26.5.1.

3. Results and Discussion

3.1. Effect of Wavelet Transform and Decomposition Level

Figure 6 presents the classification accuracy obtained without dimensionality reduction, comparing the performance of the Discrete Wavelet Transform (DWT) and Wavelet Packet Transform (WPT) using different wavelet families (haar, db4, and sym5) and decomposition levels.

In this analysis, feature normalization is not applied in order to evaluate the discriminative capability of the raw wavelet-based features.

The results show that accuracy ranges between approximately 0.93 and 0.96. Among the evaluated configurations, the sym5 wavelet consistently provides the best performance for both DWT and WPT, particularly at intermediate decomposition levels. This behavior suggests that the smoother and more symmetric structure of the sym5 wavelet is better suited to capture the underlying vibration patterns associated with rotor imbalance.

WPT achieves slightly higher accuracy than DWT in some configurations, especially at levels 2 and 3, indicating that its finer frequency decomposition can provide additional discriminative information. However, this advantage is not consistent across all levels, and in several cases DWT achieves comparable performance. This indicates that increasing frequency resolution does not necessarily translate into improved classification, particularly when redundant information is introduced.

It is also observed that increasing the decomposition level does not always improve performance. At higher levels, accuracy may decrease due to over-segmentation of the signal, which can generate redundant or less informative features. These results highlight the limitations of using raw wavelet-based features without dimensionality reduction.

These observations motivate the use of dimensionality reduction techniques to improve feature representation, as analyzed in the following subsection.

Table 2 presents the computational cost of the classification stage for DWT and WPT using the Symlet 5 wavelet at different decomposition levels. The reported values correspond to the mean execution time and standard deviation obtained over five Group K-Fold splits. As expected, the computational cost increased with the decomposition level for both methods due to the larger number of extracted features. However, DWT consistently required lower execution times than WPT, especially at higher decomposition levels. For example, at level 4, DWT required 0.0228 ± 0.0050 s without LDA, whereas WPT required

0.0473 ± 0.0089 s. The incorporation of LDA reduced the dimensionality of the DWT feature space from 120 features to only four discriminant components before classification. As a result, the execution time at level 4 decreased from 0.0228 s to 0.0103 s, corresponding to a reduction of approximately 55%. This result indicates that dimensionality reduction can decrease the computational cost of the classifier while preserving the relevant information for fault discrimination. Considering both classification performance and computational cost, the DWT-Sym5 configuration provided the most favorable trade-off. This representation achieved one of the highest classification accuracies while requiring lower execution times than the corresponding WPT configurations. Therefore, DWT-Sym5 was selected as the preferred representation method for the proposed fault diagnosis methodology.

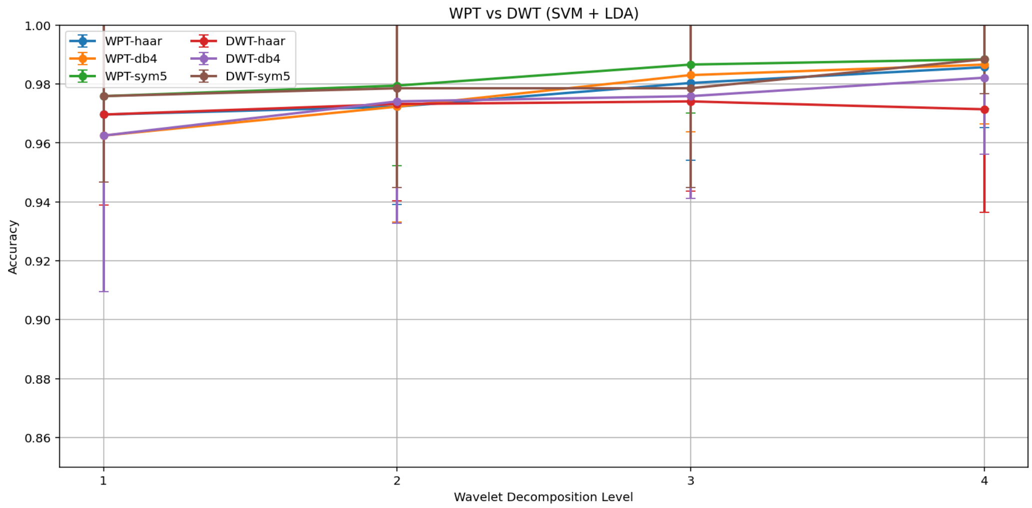


Figure 6. Comparison between WPT and DWT using SVM without LDA.

Table 2. Computational cost comparison of DWT and WPT with and without LDA using a Symlet 5 wavelet. Values are reported as mean ± standard deviation over 5 Group K-Fold splits.

Method	Level	No LDA Time (s)	LDA Time (s)
WPT	1	0.0086 ± 0.0038	0.0050 ± 0.0018
WPT	2	0.0110 ± 0.0027	0.0076 ± 0.0012
WPT	3	0.0173 ± 0.0033	0.0197 ± 0.0049
WPT	4	0.0473 ± 0.0089	0.0567 ± 0.0057
DWT	1	0.0081 ± 0.0027	0.0044 ± 0.0015
DWT	2	0.0097 ± 0.0023	0.0074 ± 0.0033
DWT	3	0.0147 ± 0.0024	0.0068 ± 0.0018
DWT	4	0.0228 ± 0.0050	0.0103 ± 0.0029

3.2. Comparison of Signal Representations

For comparison purposes, a baseline based on conventional time-domain statistical descriptors was implemented. The extracted features included mean, standard deviation, RMS, skewness, kurtosis, peak value, and crest factor computed for each vibration channel. Table 3 compares time-domain statistical features, FFT-based features, and DWT features using the same LDA and SVM classification framework. The DWT representation achieved the highest classification performance, with an accuracy of 0.9884 ± 0.0116 and an F1-score of 0.9905 ± 0.0089. In contrast, time-domain statistical features and FFT-based features obtained lower accuracies of 0.9366 ± 0.0546 and 0.9080 ± 0.0893, respectively. In addition, DWT exhibited the lowest variability across folds, indicating more stable performance. Regarding computational cost, DWT required 0.0151 ± 0.0012 s, which was comparable to the execution time of the statistical features (0.0157 ± 0.0035 s). In contrast, FFT required

4.2166 ± 0.1218 s, resulting in a substantially higher computational cost. These results indicate that DWT provides the best balance between classification performance and computational efficiency among the evaluated signal representations.

Table 3. Comparison of signal representations using LDA and SVM. Results are reported as mean ± standard deviation over five Group K-Fold splits.

Representation	Accuracy	F1 Score	Time (s)
Statistical	0.9366 ± 0.0546	0.9450 ± 0.0420	0.0157 ± 0.0035
FFT	0.9080 ± 0.0893	0.9141 ± 0.0810	4.2166 ± 0.1218
DWT	0.9884 ± 0.0116	0.9905 ± 0.0089	0.0151 ± 0.0012

3.3. Selection of the Dimensionality Reduction Method

Table 4 presents the performance of different dimensionality reduction techniques applied to the DWT feature vectors obtained using the Symlet 5 wavelet at decomposition level 4. The comparison includes the original feature space (No Reduction), PCA, Kernel PCA, and LDA, while maintaining the same SVM classifier and Group K-Fold validation strategy. The results show that LDA achieved the highest classification performance, reaching an accuracy of 0.9884 ± 0.0116 and an F1-score of 0.9905 ± 0.0089. In contrast, PCA and Kernel PCA significantly reduced the classification performance, obtaining accuracies of 0.6777 ± 0.0988 and 0.6491 ± 0.1119, respectively. The model without dimensionality reduction achieved an intermediate accuracy of 0.9446 ± 0.0772, which was also lower than that obtained with LDA. From a computational perspective, PCA provided the shortest execution time, whereas Kernel PCA exhibited the highest computational cost. Although LDA introduced an additional transformation stage, its execution time (0.0122 ± 0.0066 s) remained comparable to that of PCA (0.0106 ± 0.0021 s) and lower than the execution time obtained using the original feature space (0.0158 ± 0.0006 s). Therefore, LDA was selected as the dimensionality reduction technique for the proposed methodology because it provided the best trade-off between classification performance and computational efficiency. These results indicate that supervised dimensionality reduction is more suitable than unsupervised alternatives for the considered fault diagnosis problem, as the class information available during training can be exploited to maximize class separability.

Table 4. Comparison of dimensionality reduction methods using DWT (Sym5, level 4) and SVM classification. Results are reported as mean ± standard deviation over five Group K-Fold splits.

Method	Accuracy	F1 Score	Time (s)
No Reduction	0.9446 ± 0.0772	0.9500 ± 0.0663	0.0158 ± 0.0006
PCA	0.6777 ± 0.0988	0.6921 ± 0.0885	0.0106 ± 0.0021
Kernel PCA	0.6491 ± 0.1119	0.6637 ± 0.1080	0.0470 ± 0.0361
LDA	0.9884 ± 0.0116	0.9905 ± 0.0089	0.0122 ± 0.0066

3.4. Impact of LDA on Classification Performance

Figure 7 shows the impact of Linear Discriminant Analysis (LDA) on classification accuracy using both DWT and WPT with the sym5 wavelet and an SVM classifier.

Feature normalization is applied prior to dimensionality reduction and classification, following the proposed pipeline.

The incorporation of LDA leads to a clear improvement in performance across all configurations. Without LDA, accuracy values remain around 0.94–0.95, whereas the use of LDA increases performance to values above 0.97, reaching up to 0.98 at higher decomposition levels.

This behavior indicates that the original feature space contains redundant and overlapping information, which is effectively reduced by LDA through projection onto discriminant directions that maximize class separability. As a result, the transformed feature space becomes more compact and discriminative.

Additionally, after applying LDA, the performance difference between WPT and DWT becomes less pronounced. This suggests that dimensionality reduction has a greater impact on classification performance than the specific choice of wavelet transform.

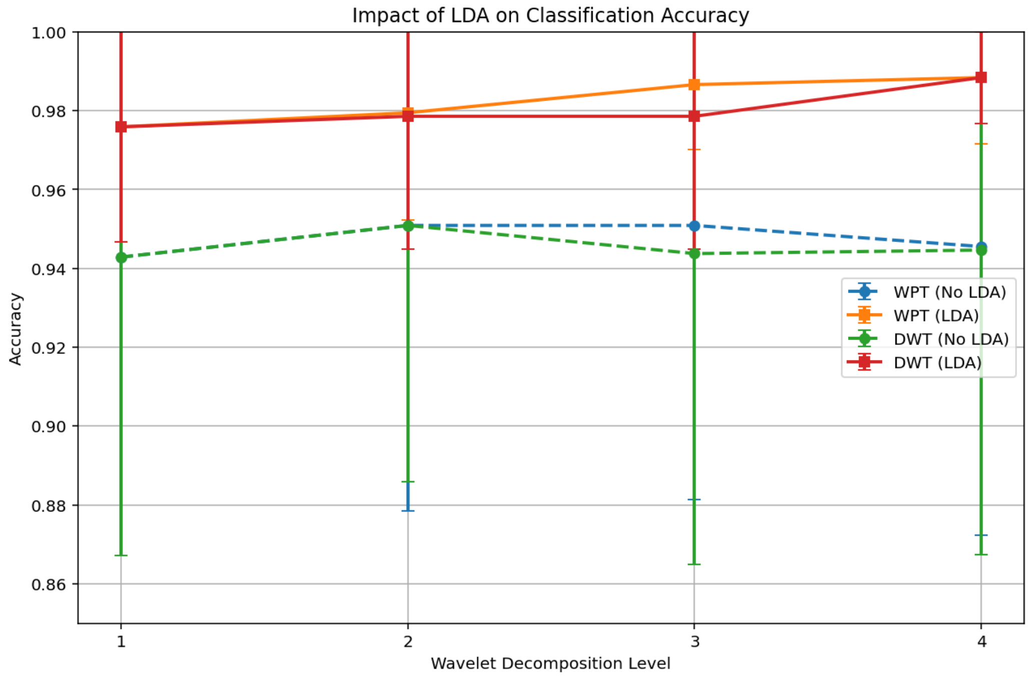


Figure 7. Impact of LDA on classification performance using DWT and WPT.

Although WPT achieves slightly higher accuracy in some cases, DWT with the sym5 wavelet at level 4 is selected due to its lower computational complexity while maintaining comparable performance, making it more suitable for practical applications.

3.5. Effect of LDA Dimensionality

Figure 8 illustrates the effect of the number of LDA components on classification accuracy using DWT with the sym5 wavelet at level 4. A significant improvement is observed when increasing the number of components from one to two, indicating that a single component is insufficient to capture the discriminative structure of the data. For higher numbers of components, performance continues to improve, although the gains become more gradual. This behavior suggests that additional LDA components capture complementary discriminative information that is not represented in lower-dimensional projections. However, beyond a certain number of components, the marginal gain becomes limited, indicating a trade-off between dimensionality reduction and information preservation. It should be noted that, for a classification problem with five classes, the maximum number of discriminant components provided by LDA is limited to $(K-1=4)$. Therefore, the value of four components does not correspond to an optimized hyperparameter but rather to the maximum dimensionality available in the LDA projection space. As shown in Figure 8, the highest classification performance is obtained when all four discriminant components are retained, indicating that each component contributes useful class-separating information for the considered fault diagnosis problem. Additionally, the reduction in standard deviation for higher numbers of components indicates improved stability across folds. From a comparative perspective, using only one LDA component provides the most compact representation but limits the amount of discriminative information available for separating the five operating conditions. Increasing the dimensionality to two and three components progressively improves the classification performance by incorporating additional discriminant directions. The four-component representation achieves the best trade-off between dimensionality reduction and classification accuracy, preserving the maximum discriminative information allowed by LDA for the five-class problem.

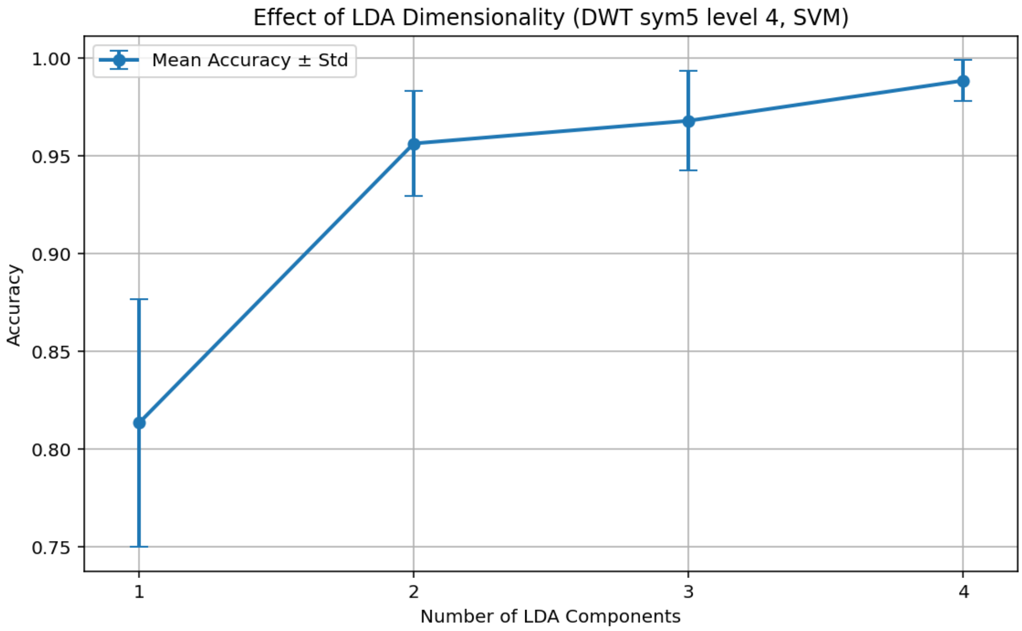


Figure 8. Effect of LDA dimensionality (DWT sym5 level 4).

The best configuration corresponds to DWT with the sym5 wavelet at level 4 and LDA with four components.

3.6. Feature Space Analysis

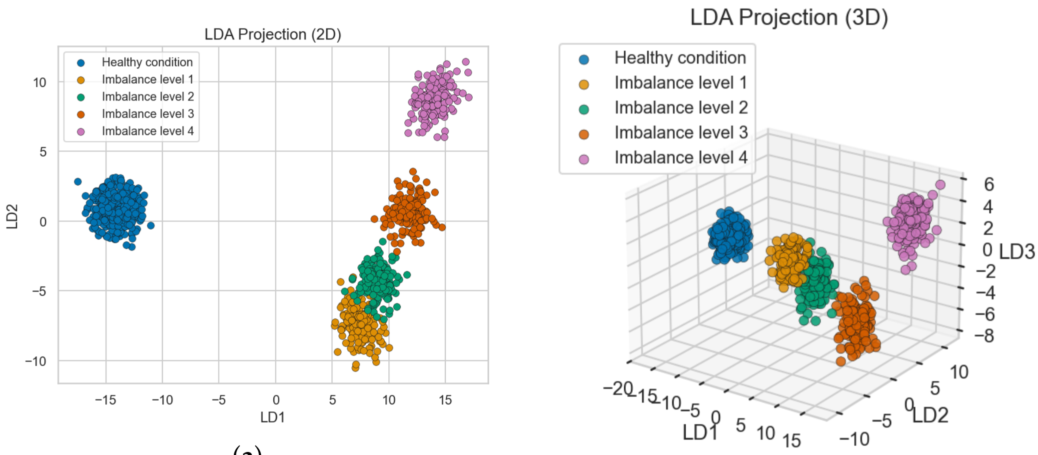
Figure 9a shows the two-dimensional projection of the dataset after applying DWT-based feature extraction, feature normalization, and LDA. Each point represents a segmented window, and colors indicate the corresponding class labels.

The results show well-defined and compact clusters with clear separation between most classes. This clustering behavior indicates that the extracted features effectively capture the underlying differences between operating conditions, particularly in terms of vibration patterns associated with different imbalance levels.

The first discriminant component (LD1) provides the primary separation between classes, while LD2 refines the distinction between neighboring conditions. The partial overlap observed between adjacent classes suggests that similar imbalance levels share comparable dynamic characteristics, which increases classification difficulty.

Figure 9b presents the three-dimensional projection. The inclusion of the third discriminant component (LD3) reveals additional structure in the data, reducing overlap between classes that appear closer in the two-dimensional representation.

This confirms that higher-dimensional LDA representations provide additional discriminative information that contributes to improved classification performance.



(**a**) (**b**)

Figure 9. LDA projections of the dataset: (**a**) two-dimensional projection and (**b**) three-dimensional projection.

3.7. Classification Performance

Figure 10 presents the aggregated confusion matrices obtained from the 5-fold Group K-Fold cross-validation.

All models exhibit strong diagonal dominance, indicating high classification accuracy. The SVM classifier achieves the most consistent performance, with minimal misclassifications across all classes.

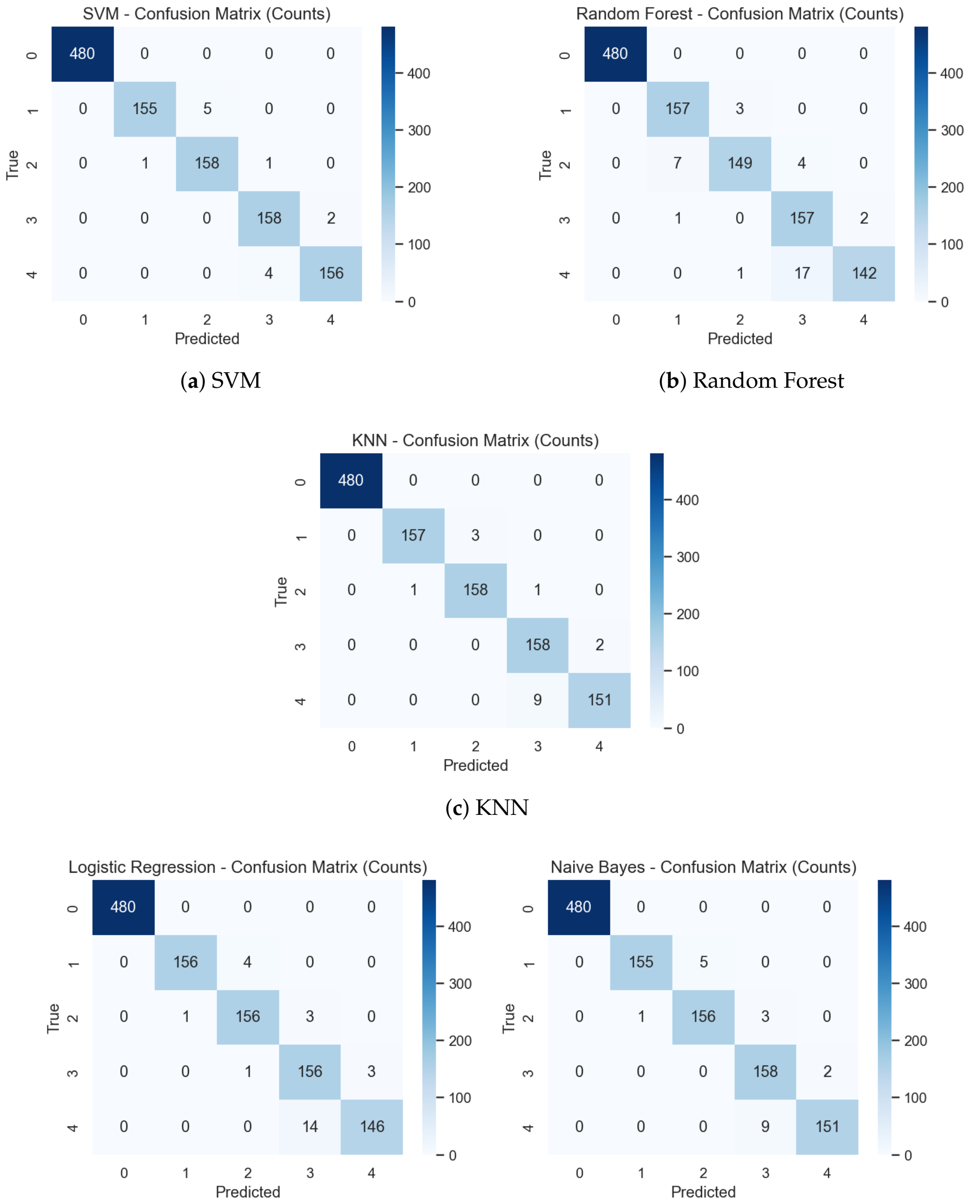
The perfect classification of class 0 across all models indicates that the healthy condition is clearly distinguishable from faulty states in the transformed feature space. In contrast, most misclassifications occur between neighboring classes, particularly between classes 1 and 2, and between classes 3 and 4. This behavior indicates that these conditions share similar vibration signatures, resulting in partial overlap in the feature space.

Table 5 summarizes the classification performance of the evaluated models using the selected configuration (DWT sym5, level 4, and LDA with four components).

Table 5. Performance comparison of classification models using DWT (sym5, level 4) and LDA (4 components). Results are reported as mean ± standard deviation over the five Group K-Fold splits.

Model	**Accuracy**	**Precision**	**Recall**	**F1 Score**
SVM	0.9884 ± 0.0116	0.9934 ± 0.0073	0.9884 ± 0.0116	0.9905 ± 0.0089
Random Forest	0.9688 ± 0.0479	0.9887 ± 0.0106	0.9688 ± 0.0479	0.9749 ± 0.0364
KNN	0.9857 ± 0.0203	0.9949 ± 0.0055	0.9857 ± 0.0203	0.9895 ± 0.0124
Logistic Regression	0.9768 ± 0.0332	0.9926 ± 0.0083	0.9768 ± 0.0332	0.9825 ± 0.0211
Naive Bayes	0.9821 ± 0.0238	0.9934 ± 0.0073	0.9821 ± 0.0238	0.9867 ± 0.0146

All evaluated classifiers achieved high classification performance, with accuracy values above 0.96, indicating that the proposed feature representation is robust across different learning algorithms. Among the evaluated models, SVM obtained the highest mean accuracy and weighted F1-score, followed closely by KNN, while Random Forest and Logistic Regression showed slightly lower performance. To assess the differences between classifiers, statistical significance tests were performed using the accuracies obtained over the five GroupKFold splits. As summarized in Table 6, the Friedman test did not reveal statistically significant differences among the five classifiers ($p = 0.1756$). The Wilcoxon signed-rank test was used to compare SVM with the remaining classifiers, yielding p-values greater than 0.05 in all cases. Results indicate that, under the adopted validation protocol, the observed performance differences were not statistically significant. SVM was selected for the subsequent experiments because it consistently achieved the highest mean accuracy and weighted F1-score across the evaluated folds.



(**a**) SVM (**b**) Random Forest (**c**) KNN (**d**) Logistic Regression (**e**) Naive Bayes

Figure 10. Confusion matrices of the evaluated classification models.

Table 6. Statistical comparison of the evaluated classifiers based on the accuracies obtained over the five GroupKFold splits.

Test	Statistic	p-Value
Friedman (all classifiers)	6.3333	0.1756
Wilcoxon (SVM vs. KNN)	1.0	1.0000
Wilcoxon (SVM vs. Logistic Regression)	0.0	0.5000
Wilcoxon (SVM vs. Naive Bayes)	0.0	1.0000
Wilcoxon (SVM vs. Random Forest)	2.0	0.3750

Table 7 presents the per-class performance metrics obtained with the SVM classifier using DWT-based features and LDA dimensionality reduction. The results show a consistently high classification performance across all operating conditions, with F1-scores ranging from 0.9783 to 1.0000. The healthy condition achieved perfect classification, obtaining precision, recall, and F1-score values of 1.0000. For the imbalance conditions, the classifier maintained F1-scores above 0.97, indicating a robust discrimination capability even for the minority classes. Although the healthy class contains a larger number of experiments than each imbalance level, the high and relatively balanced precision and recall values observed across all fault categories suggest that the overall performance is not solely driven by class imbalance. These results confirm that the proposed DWT–LDA representation provides highly discriminative features for distinguishing between healthy and imbalance conditions.

Table 7. Per-class performance metrics obtained with the SVM classifier using DWT features and LDA dimensionality reduction.

Class	Precision	Recall	F1-Score
Healthy	1.0000	1.0000	1.0000
Imbalance 1	0.9936	0.9688	0.9810
Imbalance 2	0.9693	0.9875	0.9783
Imbalance 3	0.9693	0.9875	0.9783
Imbalance 4	0.9873	0.9750	0.9811

3.8. Discussion

The results show that the combination of Discrete Wavelet Transform (DWT)-based feature extraction and Linear Discriminant Analysis (LDA) provides a clear separation of the multichannel vibration signals. The selected configuration (sym5, level 4, LDA with four components) captures the most relevant time–frequency information while reducing redundant features. The comparison with conventional statistical features and FFT-based features further demonstrated that the proposed DWT representation provides a better balance between classification performance and computational efficiency.

Although SVM achieved the highest mean accuracy and weighted F1-score, the statistical analysis did not reveal significant differences among the evaluated classifiers. Therefore, the similar results obtained with different classifiers suggest that the main improvement comes from the feature extraction and dimensionality reduction stages rather than from the classifier itself.

The use of Group K-Fold cross-validation is also important to obtain a realistic evaluation, since it avoids data leakage by ensuring that samples from the same experiment are not shared between training and testing. This is especially relevant in segmented time-series data, where samples from neighboring windows are naturally correlated.

Although the proposed methodology achieved high classification performance under controlled laboratory conditions, its applicability to more complex operating environments requires further investigation. The experiments were conducted using a single laboratory-scale wind turbine operating under fixed experimental conditions. In practical applications, factors such as variable rotational speeds, environmental noise, sensor degradation, and differences in turbine geometry may alter the statistical distribution of the extracted features, reducing the between-class separability on which LDA relies. Consequently, the discriminant projections obtained from a model trained under controlled conditions may not directly generalize to field deployments without additional adaptation or retraining. Future work will therefore focus on validating the proposed methodology using data collected under real operating conditions and evaluating its robustness under varying environmental and operational scenarios.

The obtained results suggest that the proposed DWT–LDA framework provides an effective and computationally efficient solution for rotor imbalance detection in the wind turbine system investigated in this study.

4. Conclusions

The present study proposed a fault classification methodology for wind turbine systems using multichannel vibration signals and a machine learning pipeline composed of Discrete Wavelet Transform (DWT)-based feature extraction, feature normalization, and Linear Discriminant Analysis (LDA) for dimensionality reduction. The framework was evaluated using Group K-Fold cross-validation in order to ensure a reliable assessment of the model generalization capability.

The results demonstrated that the proposed feature extraction strategy provides a clear separation between healthy and faulty operating conditions, achieving classification accuracies higher than 0.98 across different machine learning models. The comparison with conventional statistical features and FFT-based features demonstrated that the proposed DWT representation provides the best balance between classification performance and computational efficiency. Among the evaluated classifiers, the Support Vector Machine (SVM) achieved the highest mean accuracy and weighted F1-score. However, statistical significance tests did not reveal significant differences among the evaluated classifiers, indicating that the main contribution to the classification capability arises from the signal processing and feature representation stages rather than from the classifier selection itself.

The use of group-based validation proved to be important for avoiding data leakage in segmented time-series datasets, where adjacent samples are naturally correlated. By ensuring that samples originating from the same experiment were not simultaneously included in training and testing subsets, the proposed evaluation strategy provides a more realistic representation of real operating scenarios.

The analysis revealed that dimensionality reduction plays a more significant role in classification performance than the specific wavelet configuration employed. This finding highlights the importance of reducing feature redundancy and improving class separability within the latent feature space to enhance the robustness of vibration-based fault diagnosis methodologies.

Although the study was conducted under controlled laboratory conditions, the proposed DWT–LDA framework demonstrated strong potential as a computationally efficient and reliable solution for rotor imbalance detection in wind turbine systems. In this context, future work will focus on validating the methodology using real operating data and extending the framework toward real-time condition monitoring applications.

Author Contributions: Conceptualization, O.H.S.-H., J.X.L.-M. and W.A.P.M.; methodology, O.H.S.-H., J.X.L.-M. and W.A.P.M.; software, M.E.G.N., O.H.S.-H., C.E.P.-S. and W.A.P.M.; validation, M.E.G.N., O.H.S.-H., J.X.L.-M. and C.E.P.-S.; formal analysis, O.H.S.-H., M.E.G.N., C.E.P.-S. and W.A.P.M.; investigation, O.H.S.-H., M.E.G.N., C.E.P.-S. and W.A.P.M.; data curation, M.E.G.N., O.H.S.-H., C.E.P.-S. and W.A.P.M.; writing—original draft preparation, O.H.S.-H., M.E.G.N., J.X.L.-M., C.E.P.-S. and W.A.P.M.; writing—review and editing, O.H.S.-H., M.E.G.N., J.X.L.-M., C.E.P.-S. and W.A.P.M.; visualization, M.E.G.N., O.H.S.-H., C.E.P.-S. and J.X.L.-M.; supervision, J.X.L.-M. and W.A.P.M.; project administration, J.X.L.-M. and W.A.P.M.; funding acquisition, J.X.L.-M. and W.A.P.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Pedagogical and Technological University of Colombia, grant number DIN 02.

Data Availability Statement: The source code used in this study is publicly available at https://github.com/OscarHSierra/Rotor_Imbalance_Classification_Wind_Multichannel_Vibration_DWT--LDA (accessed on 22 June 2026). The dataset analyzed in this work is publicly available at https://doi.org/10.34810/data2304, originally published in the study Detection of Wind Turbine Rotor Imbalance Using Unsupervised Output-Only Vibration Data Analysis.

Acknowledgments: Oscar H. Sierra Herrera acknowledges the Vicerrectoría de Investigación y Extensión–Dirección de Investigaciones of the Universidad Pedagógica y Tecnológica de Colombia (UPTC) for the partial financial support provided toward his doctoral tuition under the institutional call "CONVOCATORIA VIE 02 DE 2026".

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

ANN	Artificial Neural Network
CNN	Convolutional Neural Network
DWT	Discrete Wavelet Transform
FFT	Fast Fourier Transform
KNN	K-Nearest Neighbors
KPCA	Kernel Principal Component Analysis
LDA	Linear Discriminant Analysis
LSTM	Long Short-Term Memory
NI	National Instruments
RBF	Radial Basis Function
SCADA	Supervisory Control and Data Acquisition
SVM	Support Vector Machine
VMD	Variational Mode Decomposition
WPT	Wavelet Packet Transform

References

1. Global Wind Energy Council. Global Wind Report 2024. 2024. Available online: https://www.gwec.net/hubfs/Website-2023/documents/GWEC-2024.pdf?hsLang=en (accessed on 7 May 2026).
2. Wang, S.; Vidal, Y.; Pozo, F. Recent advances in wind turbine condition monitoring using SCADA data: A state-of-the-art review. *Reliab. Eng. Syst. Saf.* **2026**, *267*, 111838. [CrossRef]
3. Naqash, T.M.; Alam, M.M. A State-of-the-Art Review of Wind Turbine Blades: Principles, Flow-Induced Vibrations, Failure, Maintenance, and Vibration Suppression Techniques. *Energies* **2025**, *18*, 3319. [CrossRef]
4. Carroll, J.; McDonald, A.; McMillan, D. Failure rate, repair time and unscheduled O&M cost analysis of offshore wind turbines. *Wind Energy* **2016**, *19*, 1107–1119. [CrossRef]
5. Ogaili, A.A.F.; Jaber, A.A.; Hamzah, M.N. Wind turbine blades fault diagnosis based on vibration dataset analysis. *Data Brief* **2023**, *49*, 109414. [CrossRef] [PubMed]
6. González, I.; Gómez, A.; Vidal, Y.; Tutivén, C. Detection of wind turbine rotor imbalance using unsupervised output-only vibration data analysis. *Energy AI* **2025**, *21*, 100565. [CrossRef]
7. Rangel-Rodriguez, A.H.; Centeno-Bautista, M.A.; Amezquita-Sanchez, J.P.; Valtierra-Rodriguez, M. Machine learning-based methodology for detecting cracks in wind turbine blades. In *Proceedings of the 2024 IEEE International Autumn Meeting on Power, Electronics and Computing (ROPEC24)*; IEEE: New York, NY, USA, 2024. [CrossRef]
8. Sethi, M.R.; Subba, A.B.; Faisal, M.; Sahoo, S.; Raju, D.K. Fault diagnosis of wind turbine blades with continuous wavelet transform based deep learning model using vibration signal. *Eng. Appl. Artif. Intell.* **2024**, *138*, 109372. [CrossRef]
9. Ogaili, A.; Hamzah, M.; Jaber, A.; Ghane, E. Application of Discrete Wavelet Transform for Condition Monitoring and Fault Detection in Wind Turbine Blades: An Experimental Study. *Eng. Technol. J.* **2024**, *42*, 104–116. [CrossRef]
10. Dias, A.L.; Ishioka, W.K.; Sestito, G.S.; Nicoletti, R. Intelligent fault diagnosis of rotor imbalance for small-scale wind turbines based on easy-to-measure signals. *Neural Comput. Appl.* **2026**, *38*, 239. [CrossRef]
11. Askari, A.R.; Gelman, L.; King, R.; Hickey, D.; Ball, A.D. A Novel Diagnostic Feature for a Wind Turbine Imbalance Under Variable Speed Conditions. *Sensors* **2024**, *24*, 7073. [CrossRef] [PubMed]

12. Mehlan, F.C.; Nejad, A.R. Rotor imbalance detection and diagnosis in floating wind turbines by means of drivetrain condition monitoring. *Renew. Energy* **2023**, *212*, 70–81. [CrossRef]
13. Mostafavi, A.; Friedmann, A. Wind turbine condition monitoring dataset of Fraunhofer LBF. *Sci. Data* **2024**, *11*, 1108. [CrossRef] [PubMed]
14. Peng, Z.; Chu, F. Application of the wavelet transform in machine condition monitoring and fault diagnostics: A review with bibliography. *Mech. Syst. Signal Process.* **2004**, *18*, 199–221. [CrossRef]
15. Liu, D.; Cui, L.; Cheng, W. Fault diagnosis of wind turbines under nonstationary conditions based on a novel tacho-less generalized demodulation. *Renew. Energy* **2023**, *206*, 645–657. [CrossRef]
16. Guan, S.; Li, J.; Shi, M. Research on wind turbine gearbox fault diagnosis method based on multi-kernel wavelet denoising network and contrastive learning. *Measurement* **2026**, *257*, 118982. [CrossRef]
17. Gao, H.; Zhang, X.; Gao, X.; Li, F.; Han, H. Multi-timescale attention residual shrinkage network with adaptive global-local denoising for rolling-bearing fault diagnosis. *Knowl.-Based Syst.* **2024**, *304*, 112478. [CrossRef]
18. Guo, J.; Liu, C.; Liu, S.; Liu, W. Graph contrastive learning for semi-supervised wind turbine fault diagnosis with few labeled SCADA data. *Measurement* **2025**, *245*, 116531. [CrossRef]
19. Huang, D.; Zhang, W.A.; Guo, F.; Liu, W.; Shi, X. Wavelet packet decomposition-based multiscale CNN for fault diagnosis of wind turbine gearbox. *IEEE Trans. Cybern.* **2023**, *53*, 443–453. [CrossRef] [PubMed]
20. Wang, Q.; Cui, S.; Li, E.; Du, J.; Li, N.; Sun, J. Deep Learning-Based Fault Diagnosis via Multisensor-Aware Data for Incipient Inter-Turn Short Circuits (ITSC) in Wind Turbine Generators. *Sensors* **2025**, *25*, 2599. [CrossRef] [PubMed]
21. Zhong, Q.; Liu, S.; Liu, C.; Liu, W.; Liu, S.; Zhao, Y.; Wu, Y. Fault Diagnosis of Wind Turbine Gearboxes Based on Multisource Signal Fusion. *IEEE Trans. Instrum. Meas.* **2025**, *74*, 3526413. [CrossRef]
22. Alagha, N.; Khairuddin, A.S.M.; Haitaamar, Z.N.; Al-Khatib, O.; Kanesan, J. Artificial Intelligence in Wind Turbine Fault Detection and Diagnosis: Advances and Perspectives. *Energies* **2025**, *18*, 1680. [CrossRef]
23. Badihi, H.; Zhang, Y.; Jiang, B.; Pillay, P.; Rakheja, S. A Comprehensive Review on Signal-Based and Model-Based Condition Monitoring of Wind Turbines: Fault Diagnosis and Lifetime Prognosis. *Proc. IEEE* **2022**, *110*, 754–806. [CrossRef]
24. Tavares, T.M.; Giesbrecht, M. Deep Learning-Based Fault Diagnosis in Wind Turbine Bearings and Gearboxes Using Vibration Signals: Survey, Challenges, and Recommendations. *IEEE Access* **2025**, *13*, 207013–207032. [CrossRef]
25. Wang, B.; Zhou, B.; Zhu, D.; Zou, M.; Rao, Z.; Luo, H. Feature alignment LSTM for detecting anomalies in wind turbine main bearing temperature based on SCADA data. *J. Eng. Res.* **2026**, *14*, 303–314. [CrossRef]
26. Li, X.; Wang, Y.; Yao, J.; Li, M.; Gao, Z. Multi-sensor fusion fault diagnosis method of wind turbine bearing based on adaptive convergent viewable neural networks. *Reliab. Eng. Syst. Saf.* **2024**, *245*, 109980. [CrossRef]
27. Zhu, Y.; Liu, X. A Lightweight CNN for Wind Turbine Blade Defect Detection Based on Spectrograms. *Machines* **2023**, *11*, 99. [CrossRef]
28. Ye, X.; Wang, L.; Huang, C.; Luo, X. Wind turbine blade defect detection with a semi-supervised deep learning framework. *Eng. Appl. Artif. Intell.* **2024**, *136*, 108908. [CrossRef]
29. Ma, T.; Fu, Y. A multi-sensor fused incremental detection model for blade crack with cross-attention mechanism and Dempster-Shafer evidence theory. *Adv. Eng. Inform.* **2024**, *62*, 102952. [CrossRef]
30. Tuirán, R.; Águila, H.; Jou, E.; Escaler, X.; Mebarki, T. Fault Diagnosis in a 2 MW Wind Turbine Drive Train by Vibration Analysis: A Case Study. *Machines* **2025**, *13*, 434. [CrossRef]
31. Jamil, F.; Peeters, C.; Verstraeten, T.; Helsen, J. Leveraging signal processing and machine learning for automated fault detection in wind turbine drivetrains. *Wind Energy Sci.* **2025**, *10*, 1963–1978. [CrossRef]
32. Felix, L.O.; de Sá Só Martins, D.H.C.; Monteiro, U.A.B.V.; Pinto, L.A.V.; Tarrataca, L.; Martins, C.A.O. Multiple Fault Diagnosis in a Wind Turbine Gearbox with Autoencoder Data Augmentation and KPCA Dimension Reduction. *J. Nondestruct. Eval.* **2024**, *43*, 114. [CrossRef]
33. Ogaili, A.A.F.; Mohammed, K.A.; Jaber, A.A.; Al-Ameen, E.S. Automated Wind Turbines Gearbox Condition Monitoring: A Comparative Study of Machine Learning Techniques Based on Vibration Analysis. *FME Trans.* **2024**, *52*, 471–485. [CrossRef]
34. Lakikza, A.; Cheghib, H.; Kahoul, N. Diagnosis of bearing faults in wind turbine systems using vibrational signal processing and machine learning. *Diagnostyka* **2024**, *25*, 2024307. [CrossRef]
35. Jian, T.; Cao, J.; Liu, W.; Xu, G.; Zhong, J. A novel wind turbine fault diagnosis method based on compressive sensing and lightweight SqueezeNet model. *Expert Syst. Appl.* **2025**, *260*, 125440. [CrossRef]
36. Wang, Y.; Ma, X.; Joyce, M.J. Reducing sensor complexity for monitoring wind turbine performance using principal component analysis. *Renew. Energy* **2016**, *97*, 444–456. [CrossRef]
37. Maliuk, A.S.; Ahmad, Z.; Kim, J.M. A Technique for Bearing Fault Diagnosis Using Novel Wavelet Packet Transform-Based Signal Representation and Informative Factor LDA. *Machines* **2023**, *11*, 1080. [CrossRef]
38. Van, M.; Kim, H.J. Bearing Defect Classification based on Individual Wavelet Local Fisher Discriminant Analysis with Particle Swarm Optimization. *IEEE Trans. Ind. Inform.* **2015**, *12*, 124–135. [CrossRef]

39. Jin, X.; Zhao, M.; Chow, T.W.S.; Pecht, M. Motor Bearing Fault Diagnosis Using Trace Ratio Linear Discriminant Analysis. *IEEE Trans. Ind. Electron.* **2014**, *61*, 2441–2451. [CrossRef]
40. Li, Z.; Gao, Y. Research on wind turbine unbalance fault diagnosis based on wavelet transform and convolutional neural network. *IEEE Access* **2024**, *12*, 176259–176269. [CrossRef]
41. Guo, H.; Guo, X.; Zhang, X.; Lu, F.; Liang, C. Fault diagnosis of wind turbine based on dual-channel feature aggregation network with attentional mechanism. *Eng. Appl. Artif. Intell.* **2025**, *161*, 112291. [CrossRef]
42. Wang, M.H.; Lu, S.D.; Hsieh, C.C.; Hung, C.C. Fault detection of wind turbine blades using multi-channel CNN. *Sustainability* **2022**, *14*, 1781. [CrossRef]
43. Liang, P.; Wang, B.; Jiang, G.; Li, N.; Zhang, L. Unsupervised fault diagnosis of wind turbine bearing via a deep residual deformable convolution network based on subdomain adaptation under time-varying speeds. *Eng. Appl. Artif. Intell.* **2023**, *118*, 105656.
44. Liu, H.; Wang, Y.; Zeng, T.; Wang, H.; Chan, S.C.; Ran, L. Wind turbine generator failure analysis and fault diagnosis: A review. *IET Renew. Power Gener.* **2024**, *18*, 3127–3148. [CrossRef]
45. Yang, W.; Tavner, P.J.; Crabtree, C.J.; Feng, Y.; Qiu, Y. Wind turbine condition monitoring: Technical and commercial challenges. *Wind Energy* **2014**, *17*, 673–693.
46. He, Q.; Pang, Y.; Jiang, G.; Xie, P. A spatio-temporal multiscale neural network approach for wind turbine fault diagnosis with imbalanced SCADA data. *IEEE Trans. Ind. Inform.* **2020**, *17*, 6875–6884. [CrossRef]
47. Leon-Medina, J.X.; Anaya, M.; Parés, N.; Tibaduiza, D.A.; Pozo, F. Structural Damage Classification in a Jacket-Type Wind-Turbine Foundation Using Principal Component Analysis and Extreme Gradient Boosting. *Sensors* **2021**, *21*, 2748. [CrossRef] [PubMed]