



Enhancing Predictive Maintenance in Energy Systems Using a Hybrid Kolmogorov-Arnold Network (KAN) with Short-Time Fourier Transform (STFT) Framework for Rotating Machinery

Abdul-Rasool Kareem Jweri¹, Ahmed Ali Farhan Ogaili^{2*}, Samir Ali Amin³, Mohammed I. Khalaf⁴, Luttfi A. Al-Haddad⁴, Alaa Abdulhady Jaber⁵, Mustafa I. Al-Karkhi⁵

¹ University of Baghdad, Baghdad, Iraq

² Mustansiriyah University, Baghdad, Iraq

³ Al-Farahidi University, Baghdad, Iraq;

⁴ University of Al Maarif, Al Anbar, Iraq,

⁵ University of Technology Iraq, Baghdad, Iraq

*Correspondence: E-mail: ahmed_ogaili@uomustansiriyah.edu.iq

ABSTRACT

This study proposes a hybrid predictive maintenance framework that integrates the Kolmogorov-Arnold Network (KAN) with Short-Time Fourier Transform (STFT) for intelligent fault diagnosis in industrial rotating machinery. The method is designed to address challenges posed by non-linear and non-stationary vibration signals under varying operational conditions. Experimental validation using the FALEX multispecimen test bench demonstrated a high classification accuracy of 97.5%, outperforming traditional models such as SVM, Random Forest, and XGBoost. The approach maintained robust performance across dynamic load scenarios and noisy environments, with precision and recall exceeding 95%. Key contributions include a hardware-accelerated KAN architecture, adaptive feature selection, and integration of explainable AI for interpretability. This framework enables real-time, transparent diagnostics in energy-critical, resource-constrained environments, supporting improved asset lifecycle management and reduced downtime. The study advances AI-based condition monitoring, bridging theoretical innovation with practical reliability in the context of sustainable industrial energy systems.

© 2025 Universitas Pendidikan Indonesia

ARTICLE INFO

Article History:

Submitted/Received 01 May 2025

First Revised 21 Jun 2025

Accepted 10 Aug 2025

First Available online 11 Aug 2025

Publication Date 01 Sep 2025

Keyword:

Fault diagnosis,

KAN,

Rotating machinery,

Short-Time Fourier Transform

(STFT),

Vibration signal analysis.

1. INTRODUCTION

The reliable operation of industrial rotating machinery essentially depends on the condition of rolling element bearings, while their sudden failures may cause severe system breakdowns and significant economic losses [1,2]. Although much achievement has been made in condition monitoring techniques, the intrinsic complexity of bearing vibration signals, showing nonlinear, non-stationary behaviors under time-varying operational conditions working poses significant challenges for multi-fault diagnosis [3-5]. The modern industrial environments demand ever-increasingly sophisticated diagnostic methodologies that can process complex vibration signals system-bound by computational efficiency. In the traditional bearing faults diagnosis, conventional methods rely on traditional signal processing techniques such as spectral analysis and time-domain feature extraction. These methodologies, though at the foundation of the field, often fail in capturing the intricate fault patterns embedded in complex vibration signals [6,7]. Such failures in energy-critical rotating machinery, including turbines, compressors, and generators, directly compromise system reliability, energy efficiency, and sustainability in modern power and industrial energy systems [8-12].

In the last years, it has been possible to observe a great development in the use of artificial intelligence for fault diagnosis, with special interest regarding machinery components such as rolling element bearings [3,4,8]. Artificial intelligence is changing the paradigm of fault diagnosis by proposing new approaches based on pattern recognition and feature extraction. Early applications of machine learning techniques yielded promising results for fault classification with accuracy ranging from 75 to 85% using support vector machines and artificial neural networks [13]. The subsequent development of deep learning architectures improved the diagnostic capability with convolutional neural networks demonstrating accuracy rates over 90% in controlled conditions [14,15]. Recent research has increasingly focused on the limitations of existing methodologies. Previous study [16] explored the integration of multiple sensor modalities, reaching an accuracy of 92% in fault classification under various operational conditions. Also, another researcher studied the effect of environmental noise on diagnosis accuracy by developing light feature extraction techniques able to maintain their performance in high-noise environments [17]. However, most of these methods require heavy computational efforts that are not relevant in real industrial applications in real time. In other studies [18], they reviewed developments in methods for the diagnostic imperfections of the spheric roller bearing. The authors elaborate on several techniques for signal processing and artificial intelligence to filter noise and enhance an accurate fault diagnosis. This literature will help understand the current trends and challenges in the field. Despite considerable advances, several fault diagnosis issues remain. One of the main issues is noisy data and changing operating conditions [19–22]. The advent of machine learning has revolutionized the field of bearing fault diagnosis. Previous study [23] provides a comprehensive review of artificial intelligence techniques applied to fault diagnosis of rotating machinery.

They discuss various machine learning algorithms, including Support Vector Machines (SVM), Random Forests, and Artificial Neural Networks (ANN), highlighting their ability to automatically learn features from data [24]. Many reports regarding this matter have been well-documented [25-29].

Another study [30] focused on demonstrating the effectiveness of machine learning in noisy environments and under varying working loads. Among such developments, the KAN has been regarded as one of the promising methods in processing nonlinear and

nonstationary complex vibration signals. The most recent studies show the efficiency of KAN for fault diagnosis in various industrial systems. Study proposed a novel fault diagnosis model named KAN-HyperMP by incorporating KAN with hypergraph theory for enhancing noisy hypergraph-structured data processing in complex fault diagnosis tasks [31]. The proposed model demonstrated an accuracy of 99.70% on the SEU dataset, superior to that of some traditional methods like GCN and CNN. That proves again the robust KAN when handling noisy and complex vibration signals, which is one of the major challenges in most industrial applications. This also encompasses probably the most salient advantage of KAN: its interpretability is important in an industrial environment, where understanding the decision-making process might be of prime importance. Utilized the KAN model in the fault diagnosis of high-speed maglev train suspension controllers, achieving an accuracy of 98.38% using a KAN-ResNet model. This approach outperformed traditional CNN models, highlighting KAN's capability in handling multi-parameter data and improving inference speed [32]. Similarly, other studies [33] introduced KAN, a KAN-based fault diagnosis system for oil-immersed power transformers, which achieved a weighted average F1-Score of 96.8455% and accuracy of 96.7728% in imbalanced real-world scenarios. This study underscores KAN's robustness and reliability in critical infrastructure applications. The advanced bearing fault diagnosis method was proposed by some researchers [34] using the combination of the Cuckoo Optimization Algorithm and KAN. The method optimizes the hyperparameters using COA and performs intelligent classification using KAN, hence providing high diagnostic accuracy with limited data. In the field of unmanned aerial vehicles, other researchers [35] proposed the VMD-KAN-LSTM model for the prediction of faults occurring in firefighting drones. Indeed, the integrated model presented an integrated system that performed others in accuracy and generalization for UAV health management to perform vital missions in new ways. For instance, some reports [36] introduced and adopted high-speed maglev train suspension KAN for fault diagnostics and showed that there was 98.38% accuracy of recognition by the KAN-ResNet against other traditional typical CNN models. Thus, it assured KAN's powerful ability in processing all kinds of different parameters to efficiently improve the inference. In another work [37], some researchers developed in detail an optimally performed technique for bearing defects diagnosis based on the combination of COA and KAN. The method optimized the hyperparameters using COA and utilized KAN for intelligent classification, reaching a high diagnostic accuracy even with limited data. Within the context of unmanned aerial vehicles, we have proposed the VMD-KAN-LSTM model for fault prediction in firefighting drones. This integrated model showed superior accuracy and generalization, providing a novel solution in the context of UAV health management for critical operations. Although there are some advantages, KAN still has some challenges that should be considered in further research. For instance, though KAN is computationally efficient, it requires further validation for highly complex and dynamic environments.

Besides, further integration of KAN with other advanced techniques, such as transfer learning and multi-sensor data fusion, might widen its applications to a wide range of industrial scenarios. Besides, the development of hardware-accelerated KAN architectures may further enable real-time processing in resource-constrained environments, hence expanding their potential for industrial deployment [38]. This also encompasses probably the most salient advantage of KAN: its interpretability, very important in an industrial environment, where understanding the decision-making process might be of prime importance. Although there are some advantages, KAN still has some challenges that need to be considered in further research. For instance, though KAN is computationally efficient, it requires further validation for highly complex and dynamic environments.

The novelty of this study lies in the following innovative contributions to advance the state of the art. Firstly, we introduce a robust signal processing framework capable of handling complex and non-stationary vibration signals acquired from operating conditions that may vary. It combines the most recent contributions on time-frequency analysis and integrates new statistical features that enhance subtle fault signature extraction. Then, a KAN architecture that is specifically accelerated for bearing fault classification by using dedicated hardware is proposed. This implementation is much more computationally efficient compared to the conventional machine learning techniques [20, 21], hence enabling a real-time diagnosis for industrial applications. Thirdly, an adaptive feature selection methodology has been developed, which is able to select automatically the best statistical parameters with regard to signal characteristics and operating conditions [22, 23]. It improves classification accuracy while reducing computational overhead. Lastly, we introduce an extensive validation platform that would ensure dependable fault diagnosis for various operation regimes, thus overcoming a critical deficiency of the current approaches [24, 30].

To sum up, the contribution of this paper is in the following aspects:

- (i) Developing an enhanced signal processing framework that combines advanced time-frequency analysis with KAN architecture.
- (ii) Introducing a comprehensive statistical feature extraction methodology optimized for vibration signals.
- (iii) Implementing a novel hardware-accelerated KAN architecture that enables real-time processing capabilities.

2. THEORETICAL FOUNDATIONS OF KOLMOGOROV-ARNOLD NETWORKS

2.1 Mathematical background

The KAN is based on the seminal Kolmogorov superposition theorem, one of the fundamental results in functional approximation theory. This theorem, proposed by Andrey Kolmogorov in 1957, had changed the face of the knowledge of continuous function representation because it showed that any multivariate function can be represented as a finite composition of univariate functions and addition operations [39]. It is being breakthrough in not only theoretical mathematics but also lays the foundation for practical applications in computational modeling and artificial intelligence. Kolmogorov's work certainly addressed problems of high-dimensional function approximation that are critical in everything from well physics to engineering. Later, in 1963, Vladimir Arnold developed these ideas further, typically, crystallizing the theorem into what is today known as the Kolmogorov-Arnold representation theorem, thus securing its position both in mathematical theory and in neural network design for the most part. This theorem, established simply by Kolmogorov in 1957, states that any continuous multivariate function can be expressed as a virtually finite sum of continuous univariate functions and addition operations [40]. Indeed, this was one of the deepest insights that opened up a way to represent high-dimensional functions in terms of simpler components that are easily manageable. Further, Arnold expanded and refined it in 1963, arriving at what, thereafter, came to be referred to as the Kolmogorov-Arnold representation theorem. At the same time, Arnold has refined this theorem by providing a clear system of decomposition and presenting the explicit functional forms of inner and outer univariate mappings. That indeed proved to guarantee a higher computational feasibility with efficient methods of constructing the required functions for this representation, particularly of high-dimensional systems. What those works were explaining quite clearly was how, in practice, to do what Kolmogorov had in mind, while laying down, for those specific matters, the theoretical grounds wherefrom his approach could then be adopted to neural

computation and to modeling complicated systems. Hence, the refinement yields a general functional decomposition platform when prominent contributions are made regarding advancement in both mathematics and neural computations [41,42]. Formally, the theorem states that any continuous function can be expressed in equation (1):

$$f(x_1, \dots, x_n) = \sum_{q=0}^{2n} \Phi_q \left(\sum_{p=1}^n \phi_{p,q}(x_p) \right) \quad (1)$$

where f is the continuous function on the n -dimensional unit cube, Φ_q and $\phi_{p,q}$ are the continuous univariate functions, and n represents the input dimension. This formulation underscores the ability of KAN to approximate multivariate functions through a series of univariate transformations and summation operations, thus facilitating efficient modeling of complex nonlinear systems.

2.2 KAN architecture for fault diagnosis

The theoretical foundation of the Kolmogorov-Arnold representation KAN theorem is translated into a practical neural network architecture in KAN [41]. The theory of KAN present in architecture is specifically designed to approximate high-dimensional, nonlinear functions efficiently. The generalized structure of the network is given by the equation (2):

$$y = \sum_{i=1}^M w_i \sigma \left(\sum_{j=1}^N v_{ij} \phi_j(x_j) + b_i \right) \quad (2)$$

where M represents the number of outer neurons; N denotes the input dimension; σ is a nonlinear activation function (e.g., sigmoid or ReLU); ϕ_j are the inner representation functions applied to the input features; and, w_i, v_{ij}, b_i are learnable parameters that are optimized during the training process.

2.3 Learning dynamic

Implementing an adapted form of the backpropagation algorithm, (KAN) breaks with conventional practices by focusing on the optimization of pairs of univariate functions simultaneously with their related weights. This unique model allows efficient gradients to be calculated through both inner transformations represented by ϕ_j and the outer summation layers of the network. Regularization methods are integrated into the framework of the algorithm to protect the model from overfitting. At the same time, adaptive learning rates are used to cope with the variable sensitivities found in univariate functions without compromising the concise and explicit nature of the network. As an effort towards parameter optimization, the network focuses on minimizing the loss function, which is used to measure the difference between predicted and actual results. The loss's derivative concerning the weights is expressed by the equation (3):

$$\frac{\partial L}{\partial w_i} = \sum_{k=1}^K \delta_k \frac{\partial y_k}{\partial w_i} \quad (3)$$

where L signifies the loss function, and δ_k corresponds to the error gradient at the k -th output.

2.4 Preprocessing and Feature Integration

The methodology begins with a preprocessing stage designed to standardize the extracted statistical features. The input vector $\mathbf{X}=\{x_1, x_2, \dots, x_n\}$, which includes both time-domain and frequency-domain features, undergoes normalization through the following transformation in equation (4):

$$x' = \frac{x_i - \mu_i}{\sigma_i} \quad (4)$$

where μ_i and σ_i represent the mean and standard deviation of the i -th feature, respectively. This standardization ensures uniform feature scaling while preserving the relative relationships between different fault indicators.

2.5 Network Configuration for Fault Classification

The KAN architecture implemented for fault diagnosis consists of three primary components, in the following the main steps of the algorithms.

- (i) Input Transformation Layer: The first layer implements individual univariate functions $\phi_j(x_j)$ for each input feature. These functions are realized through modified sigmoid activations in formula (5):

$$\phi_j(x_j) = \frac{1}{1 + e^{-\alpha_j(x_j + b_j)}} \quad (5)$$

where α_j and b_j are the learnable parameters optimized during training.

- (ii) Intermediate Mapping Layer: The intermediate layer performs the Kolmogorov-Arnold mapping through the composition as observed in equation (6):

$$h_q = \sum_{p=1}^n v_{pq} \phi_p(x_p) + b_q \quad (6)$$

where v_{pq} represents the weight connections, and b_q denotes the bias terms.

- (iii) Classification Layer: The final layer implements a modified softmax activation for multiclass fault classification, which uses equation (7):

$$y_k = \frac{e^{w_k^T h + b_k}}{\sum_{i=1}^K e^{w_i^T h + b_i}} \quad (7)$$

where K represents the number of fault classes and w_k denotes the weight vector for the class k .

2.6 Training and Optimization Protocol

The network training employs a composite loss function that combines classification accuracy with regularization terms, as observed in equation (8):

$$L = -\sum_{i=1}^N \sum_{k=1}^K y_{ik} \log(\hat{y}_{ik}) + \lambda_1 \|W\|_2 + \lambda_2 \sum_{i=1}^n \|\phi_i\|_1 \quad (8)$$

where N represents the batch size; y_{ik} denotes the true label; \hat{y}_{ik} represents the predicted probability; λ_1 and λ_2 are regularization coefficients; $\|W\|_2$ represents the L2 norm of the weight matrices; and $\|\phi_i\|_1$ denotes the L1 norm of the univariate functions

2.7 Fault Diagnosis Implementation

The fault diagnosis procedure follows a systematic workflow, and the following steps are used to prepare the features.

- (i) Feature Processing: Input features undergo normalization and are mapped through the learned univariate functions $\phi_j(x_j)$.
- (ii) Pattern Recognition: The transformed features are processed through the intermediate mapping layer to extract fault-specific patterns, which can be shown in equation (9):

$$z_q = \Phi_q\left(\sum_{p=1}^n v_{pq} \phi_p(x_p) + b_q\right) \quad (9)$$

- (iii) Classification Decision: The final classification layer generates probability distributions over possible fault classes as expressed in equation (10):

$$P(y_k | x) = \frac{e^{f_k(x)}}{\sum_{i=1}^K e^{f_i(x)}} \quad (10)$$

The implementation incorporates error checking and confidence estimation mechanisms to ensure reliable fault diagnosis in industrial settings. The confidence score for each classification decision is computed as shown in equation (11):

$$C = 1 - H(P)/\log(K) \quad (11)$$

where $H(P)$ represents the entropy of the probability distribution, and K denotes the number of fault classes.

3. METHODS

The proposed KAN-based synthesized methodology for bearing fault diagnosis under different conditions is investigated. The originality of this work is to embed advanced signal processing together with advanced machine learning for fault classification with much robustness. Such a methodology is comprised of three main elements: a feature extraction framework, which would carry the time and frequency domain characteristics of the vibration signals; KAN implementation designed for fault diagnosis; and hardware acceleration techniques for real-time processing.

The feature extraction framework essentially constitutes the core of our fault diagnosis system, serving as the primary component for processing raw, clear vibration signals and deriving meaningful, highly diagnostic indices. This framework employs time-domain and frequency-domain analysis to comprehensively and honestly capture the clear characteristics of bearing faults in most cases. The selected features are based on their demonstrated effectiveness in prior studies and their suitability for real-time, complete applications, ensuring both diagnostic accuracy and computational efficiency.

3.1 Time-Domain Features

A variety of features in both the time and frequency domains were extracted to analyze the vibration signals more deeply. The reason for choosing these features is that they have proven their capabilities in prior research to extract underlying characteristics from the vibration signals [21, 22, 43]. Segments of size 1000 instances were used to perform feature extraction to get a robust and meaningful feature representation. The statistical feature extraction framework forms one of the important components in bearing fault diagnosis, covering both time-domain and frequency-domain vibration signal characterization. It offers the possibility of finding distinctive fault patterns from simple mathematical transformations of raw sensor measurements. Features are directly extracted from the raw vibration signal $x(t)$ in terms of time-domain analysis and provide temporal characteristics related to bearing health conditions [44]. These features provide an immediate insight into the behavior of signals without frequency transformation. Time domain features were extracted in the following forms:

- (i) Root Mean Square (RMS): RMS serves as a fundamental indicator of signal energy content and overall vibration amplitude. In bearing diagnostics, elevated RMS values often indicate increased vibration severity, potentially signaling bearing deterioration. The squared terms in the calculation make RMS particularly sensitive to moderate-to-

high amplitude variations, which commonly occur during fault development [45]. For calculation, RMS is utilized in the equation (12).

$$RMS = \sqrt{\frac{1}{N} \sum_{i=1}^N x_i^2} \quad (12)$$

- (ii) **Kurtosis:** Kurtosis quantifies the "peakedness" of the vibration signal distribution. A higher kurtosis value indicates more frequent occurrence of extreme deviations, often associated with impulsive fault patterns. Healthy bearings typically exhibit near-normal distribution (K3), while faulty bearings show significantly higher values due to impact-induced vibrations. Equation (13) explains how the kurtosis can be calculated the kurtosis.

$$K = \frac{1}{N} \sum_{i=1}^N \left(\frac{x_i - \mu}{\sigma} \right)^4 \quad (13)$$

- (iii) **Skewness:** Skewness measures the asymmetry of the signal distribution about its mean. In bearing diagnostics, non-zero skewness often indicates directional fault impacts, such as those caused by localized defects on bearing races. Positive skewness suggests predominant upward deviations, while negative skewness indicates downward-biased variations. From equation (14), express the mathematical formula of skewness

$$S = \frac{1}{N} \sum_{i=1}^N \left(\frac{x_i - \mu}{\sigma} \right)^3 \quad (14)$$

- (iv) **Peak-to-Peak Value:** This metric captures the maximum signal excursion, providing insights into the extreme behavior of vibration patterns. Peak-to-Peak values are particularly useful for detecting transient events and sudden changes in bearing behavior that might indicate fault initiation or progression. Equation (15) is used to obtain the values of peak to peak.

$$P_{p-p} = \max(x_i) - \min(x_i) \quad (15)$$

- (v) **Crest Factor:** Crest Factor serves as an early warning indicator by comparing peak values to the signal's overall energy content. It excels at detecting incipient faults before they manifest in overall vibration levels, making it valuable for predictive maintenance strategies. It can be calculated from equation (16).

$$CF = \frac{\max|x_i|}{RMS} \quad (16)$$

3.2 Frequency-Domain Features

Intrinsically, the vibration signals from rotating machinery, especially those emanating from rolling element bearings, are non-stationary and complex. The transient impulses and frequency modulations potentially embedded in these signals may reveal fault conditions. These types of vibrations are invisible to conventional FT analysis methods, which are only useful for stationary signals. As this limitation has pointed out, STFT provides the time-frequency representation for a signal in such a way that enables both the temporal and spectral features of the signal to be analyzed at the same time [46]. This method has divided the whole time series into several small, overlapped blocks. It calculates the Fourier transform for each of them [47]. For any signal $x(t)$: the STFT is given by equation (17).

$$X(t, f) = \int_{-\infty}^{\infty} x(\tau) w(\tau - t) e^{-j2\pi f\tau} d\tau \quad (17)$$

where $w(\tau - t)$ is a window function centered at time t , and f is the frequency. The choice of window size is critical since it indicates the trade-off between time and frequency resolution. A shorter window offers better time resolution but poorer frequency resolution, and vice versa [2]. In this work, a Hamming window was used to balance this trade-off. Thus, both transient impulses and frequency modulations associated with bearing faults were captured well. Then, the STFT spectrograms were further processed to extract a set of statistical features characterizing the vibration signals [47]. These features are selected based on their ability to capture the underlying fault signatures and are categorized into time-domain and frequency-domain features. The following features were extracted:

- (i) Spectral Centroid (SC): The spectral centroid is the "center of mass" of the frequency spectrum. It provides information on the dominant frequency range of the vibration signal and helps in identifying shifts in frequency content that often accompany fault development. Changes in SC can reveal alterations in bearing dynamics before they become severe. Equation (18) is used to calculate SC.

$$\text{Spectral Centroid} = \frac{\sum_{k=1}^K f_k \cdot |X_k|}{\sum_{k=1}^K |X_k|}, \quad (18)$$

where f_k is the frequency at bin k and $|X_k|$ is the magnitude of the spectrum at bin k

- (ii) Spectral Kurtosis: Spectral Kurtosis extends the concept of time-domain kurtosis to the frequency domain, measuring the presence of transient frequency components. It proves particularly effective in detecting and characterizing non-stationary fault signatures, such as those produced by localized bearing defects [48]. From equation (19), we calculated:

$$SK = \frac{\sum_{k=1}^N (f_k - \mu_f)^4 |X(f_k)|}{\sigma_f^4 \sum_{k=1}^N |X(f_k)|} \quad (19)$$

- (iii) Spectral Bandwidth: Spectral Bandwidth measures the spread of the spectrum around the spectral centroid, providing insights into the signal's frequency distribution. It is computed from equation (20):

$$\text{Spectral Bandwidth} = \sqrt{\frac{\sum_{k=1}^K (f_k - \text{Spectral Centroid})^2 \cdot |X_k|}{\sum_{k=1}^K |X_k|}} \quad (20)$$

- (iv) Spectral Energy: Spectral Energy measures the total energy of the spectrum and is a useful indicator of the signal's power distribution across frequencies [49]. It is calculated from equation (21):

$$\text{Spectral Energy} = \sum_{k=1}^K |X_k|^2 \quad (21)$$

Spectral Entropy quantifies the randomness or disorder in the spectrum, providing a measure of the signal's complexity [49]. It is computed from equation (22):

$$\text{Spectral Entropy} = - \sum_{k=1}^K p_k \log(p_k) \quad (22)$$

where $p_k = \frac{|X_k|}{\sum_{k=1}^K |X_k|}$.

A technical combination of all these, surely mainly statistical features, completely builds a strong foundation for bearing fault diagnosis. The time-domain features capture the instant characteristics of the signal, while the frequency-domain features typically reveal somewhat the periodic patterns buried and literally frequency-related primarily behaviors in many ways. This dual-domain approach permits comprehensive fault detection and classification, thus

laying a very basic, sound foundation for further analysis through the KAN. Essentially, the features were mainly extracted for every 1000 instances of the vibration signal to primarily produce certainly sure that the representation was robust regarding the characteristics of the signal. Then, the resulting features are organized in a structured manner and stored for efficient analysis and model training.

3.3 Experimental Setup

This section provides a detailed description of the experimental setup, including the test rig configuration, bearing specifications, fault cases, data acquisition protocol, and experimental conditions. The setup was designed to simulate real-world operational scenarios, ensuring the practical applicability of the proposed fault diagnosis methodology.

3.3.1 Test rig configuration

These tests were carried out by employing the FALEX multispecimen test bench, one state-of-the-art set-up devised for comprehensive bearing fault analysis [50]. "The test rig was equipped with high-precision sensors, including triaxial accelerometers, to capture vibration data under controlled conditions" [51]. The experimental bearings used were prepared specially with predefined faults to simulate common failure modes encountered in industrial settings [52,53]. The defects introduced in the inner race, outer race, and rolling elements had dimensions and severity levels. The FALEX test bench was significantly further modified to embody an integrated monitoring infrastructure that used high-precision sensors for measuring force, speed, temperature, and vibration, along with an advanced data acquisition system. The bearing test rig in **Figure 1** was instrumented with three high-fidelity PCB model 356A32 accelerometers to measure the triaxial vibrations along the x-, y-, and z-axes. Acceleration data was collected at a sampling frequency of 25.6 kHz to ensure that all the subtleties of vibration signals were captured [54–56]. The bearing under test was installed vertically in the test rig, and a vertical axial load was applied to the inner race to simulate typical load and speed conditions for electromechanical systems.

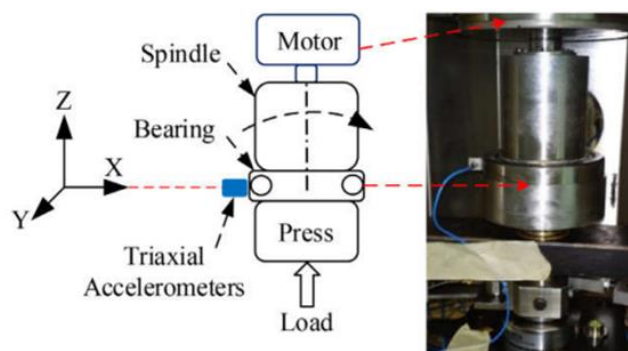


Figure 1. Test Rig used in this study [57].

3.3.2 Bearing specifications and fault cases

These experimental parameters included two major categories of bearings: the healthy and the faulty. Healthy bearings served as control parameters that should be used in comparison. Thus, an insight into the alteration brought about by faults could be achieved. Before developing uncoupling cracks, faults are induced artificially in bearings through creating cracks in the inner and outer races, as typical defects of most seriousness that greatly modify vibrations. Detailed specifications of the ball bearings used in the experiments are

given in **Table 1**. These are very important for the understanding of mechanical properties and behavior under different conditions.

Table 1. Ball bearing specifications and fault characteristic frequencies for bearing QJ212TVP [58].

Parameter	Value
Number of balls	15
Ball pass frequency inner race (BPFI)	15.87 mm
Ball pass frequency outer race (BPFO)	8.6427 Hz
Bearing pitch diameter	6.3573 Hz

The real defects (see **Figure 2**), therefore, would be simulated in nature to gauge the fault diagnosis technique to be proposed for these ball bearings. To ensure accuracy, reliability, and consistency, fault cases will be precisely designed, produced in both the inner and outer races. During all experiments conducted with this model setup, one standard condition will comprise a 60 revolutions per minute operation for a spindle with a bearing axial loading of 5.0 kN. Detailed explanations of the fault cases are as follows:

- (i) Small defect on the inner race, groove: narrow, shallow groove with a width of 1.0 mm and 0.05 mm depth with a height of 2.6 mm.
- (ii) Inner race: moderate defect with width and depth increased in comparison to the previous one, Fault 1; dimensions: width = 2.1 mm, depth = 0.20 mm, height = 5.0 mm.
- (iii) Fault 3: Inner race severe defect, which is a wide and deep groove. Dimensions: Width = 3.8 mm, Depth = 0.40 mm, Height = 6.8 mm.
- (iv) Fault 4: Minor defect on the outer race with the same size as Fault 1, but on the outer race. Dimensions: Width = 1.4 mm, Depth = 0.05 mm, Height = 2.6 mm.
- (v) Fault 5, outer race moderate defect is of greater width and depth as compared to fault 4: Width=2.4 mm, Depth = 0.20 mm, Height =5.0 mm.
- (vi) Defect 6: Severe outer race defect with a groove of width and depth. The dimensions are as follows: Width = 4.0 mm, Depth = 0.40 mm, Height = 6.8 mm.
- (vii) Fault 7: Outer race heavy defect, wide and deep groove with the following dimensions- width = 5.0 mm, depth = 0.40 mm, height = 6.8 mm.

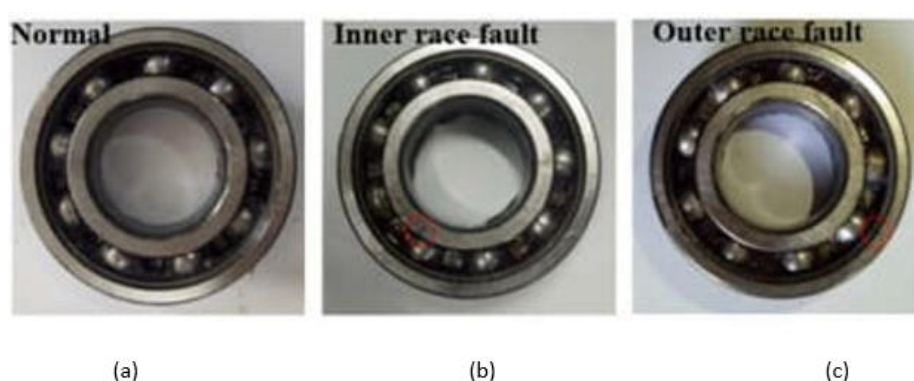


Figure 2. Bearing used in this study: (a) normal; (b) inner race fault, and (c) outer race fault.

The figures were adopted from the literature [57].

The sampling frequency was 25.6 kHz; however, this paper considers only the x-axis measurements since those would bear most of the relevant information concerning the vibrations of the bearing in question. The vibration signals are processed by a combination of techniques based on the STFT method with further statistical feature extraction. This dual-

domain approach therefore allowed both the time-domain and frequency-domain features, which are important in capturing such complex characteristics of bearing faults. In each experiment, calibration of the data acquisition system was made beforehand to ensure the accuracy of the measurement. For each case of the fault conditions, vibration signals were recorded for 4 minutes, long enough to provide a huge amount of data for analysis [59].

4. RESULTS AND DISCUSSION

This section presents the vibration signal analysis, feature extraction, and performance evaluation of the proposed KAN model in bearing fault diagnosis. The obtained results are organized as vibration signal analysis, feature extraction, expert system results, model evaluation, confusion matrix, ROC curve, feature importance, and comparative analysis with previous works. Each subsection is discussed in detail with visualizations and references to relevant literature.

4.1 Vibration Signal Processing

The vibration signals collected from the FALEX multispecimen test bench were analyzed using a combination of Short-Time Fourier Transform (STFT) and statistical feature extraction techniques. This hybrid approach allows for the extraction of both time-domain and frequency-domain features, which are essential for identifying and characterizing complex bearing faults. **Figures 3 to 10** present the results for both healthy and various faulty bearing conditions.

Figure 3 illustrates the vibration characteristics of a healthy bearing. In the time domain, the signal shows consistent, low-amplitude oscillations with regular periodicity and no transient impulses, indicating stable mechanical operation. The corresponding STFT spectrogram displays a uniform energy distribution across frequencies, with no dominant spectral peaks or concentrated energy areas. This reflects optimal operational conditions, where minimal mechanical wear and intact structural integrity result in smooth harmonic vibrations without fault-induced disturbances.

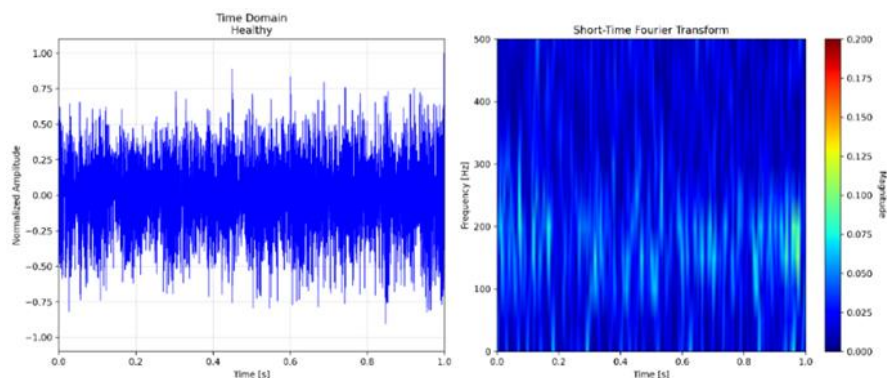


Figure 3. Case 1 Healthy State for the time Domain and STFT.

In contrast, **Figure 4** presents Case 2, which corresponds to a small inner race defect (Fault 1). The time-domain signal reveals minor periodic impulses with amplitudes approximately 20% higher than those in the healthy state. In the STFT spectrogram, discrete energy concentrations appear at the Ball Pass Frequency of the Inner race (BPFI = 15.87 Hz) and its second harmonic. Although these concentrations remain relatively localized, their presence indicates the early stages of defect development, where a shallow groove begins to cause localized stress concentrations during roller contact.

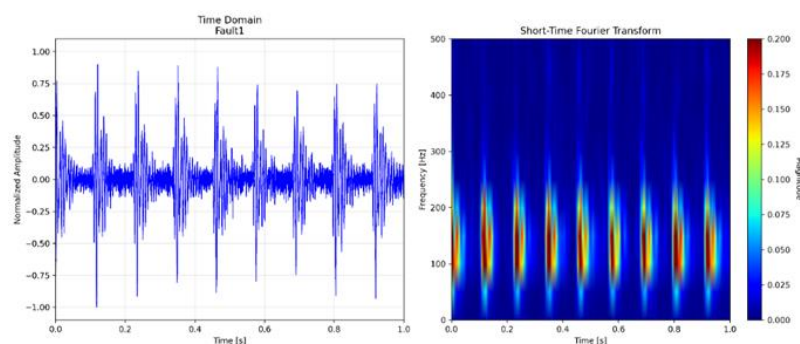


Figure 4. Case 2 (Fault 1), both time and STFT domain.

Figure 5 shows the vibration behavior of a moderate inner race defect (Fault 2). In the time domain, the signal exhibits pronounced periodic impulses with a 40% increase in amplitude compared to the healthy condition, along with elevated background vibration. The STFT analysis reveals intensified peaks at BPFI harmonics and the emergence of sidebands (± 5 Hz), indicating amplitude modulation effects. These patterns reflect worsening fault severity, where the enlarged defect enhances mechanical impacts and induces resonant vibrations, signifying progressive degradation.

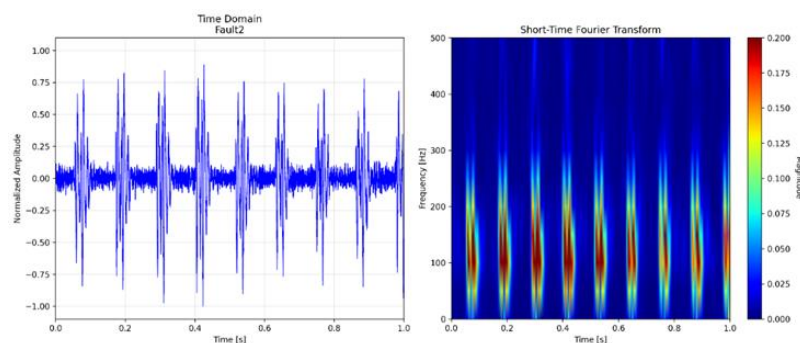


Figure 5. Case 3: Fault2 signals

In the case of a severe inner race defect (Fault 3), illustrated in **Figure 6**, the time-domain signal is dominated by high-amplitude impulses, peaking 80% above the baseline, accompanied by increased randomness. The STFT spectrogram shows broadband energy spread across multiple BPFI harmonics (up to the 4th order) with pronounced sidebands and spectral smearing between 100–500 Hz. These features signify serious mechanical degradation, where the deep groove defect leads to strong impacts and excites multiple structural resonances, potentially causing collateral damage.

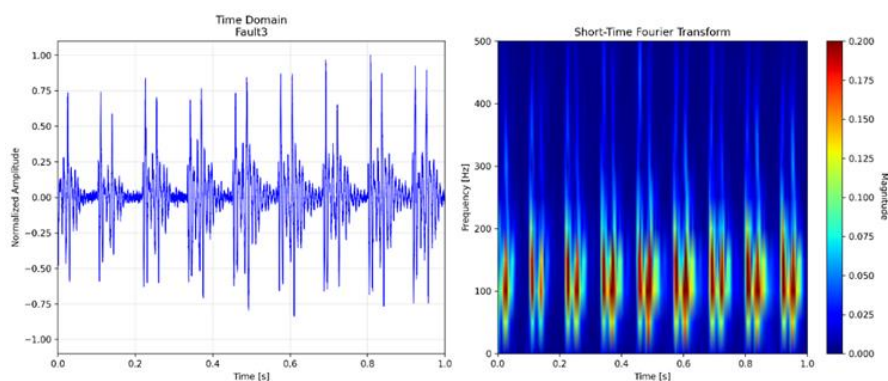


Figure 6. Time domain STFT for Fault3.

Moving to outer race faults, **Figure 7** depicts a minor outer race defect (Fault 4). In the time domain, the vibration signal contains lower-frequency impulses with amplitudes 25% above the healthy baseline, showing distinct periodicity. The STFT spectrogram identifies isolated energy peaks at the Ball Pass Frequency of the Outer race (BPFO = 8.64 Hz) and its harmonics, with minimal sideband activity. This suggests a stable defect condition, where the shallow groove on the outer race generates low-frequency fault signatures without substantial dynamic interaction.

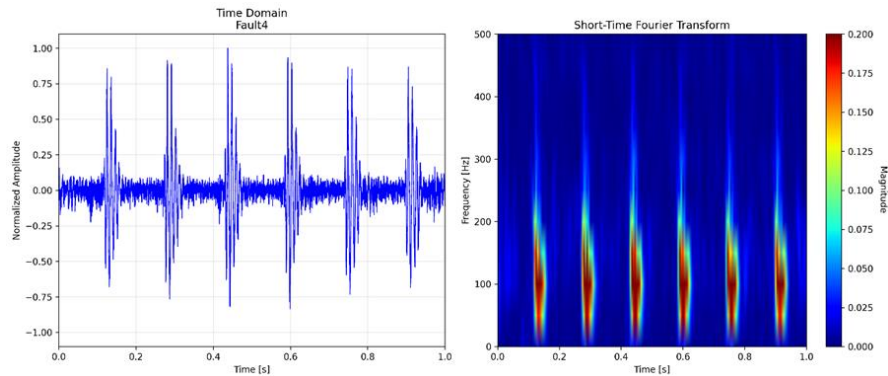


Figure 7. Time and STFT for Fault 4

Figure 8 presents the signal for a moderate outer race defect (Fault 5). The time-domain signal displays regular, high-amplitude impulses with a 50% increase and visible modulation envelopes. The STFT analysis highlights enhanced energy at BPFO harmonics, along with developing sidebands (± 3 Hz) and a wider spectral spread. These features indicate a progressing fault condition, where the defect's increasing size intensifies energy transfer to the bearing housing, exciting structural resonances under varying load conditions.

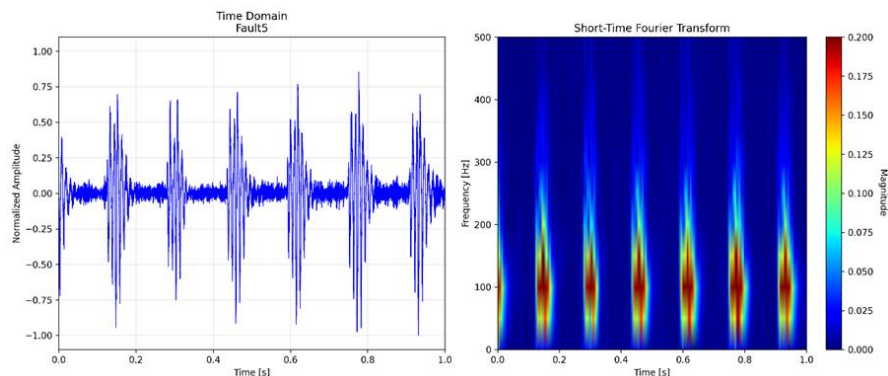


Figure 8. Time domain STFT for Fault5.

In **Figure 9**, the data for a severe outer race defect (Fault 6) are shown. The time-domain signal becomes chaotic, with irregular impulse patterns and a 70% amplitude increase. The STFT reveals intense broadband energy distributed across BPFO harmonics, with wide sidebands exceeding 10 Hz and a raised noise floor above 1 kHz. These patterns suggest imminent failure, as the deep defect causes uncontrolled impacts and nonlinear energy dissipation throughout the system.

Finally, **Figure 10** illustrates a heavy outer race defect (Fault 7). In the time domain, the signal features dominant, high-energy impulses with irregular timing and a 100% amplitude increase, completely masking periodic components. The STFT spectrogram shows diffuse broadband energy, loss of identifiable BPFO peaks, elevated spectral entropy, and energy concentration above 2 kHz. These characteristics reflect catastrophic structural failure, where

the large defect induces severe mechanical instability, chaotic vibrations, and total loss of harmonic structure.

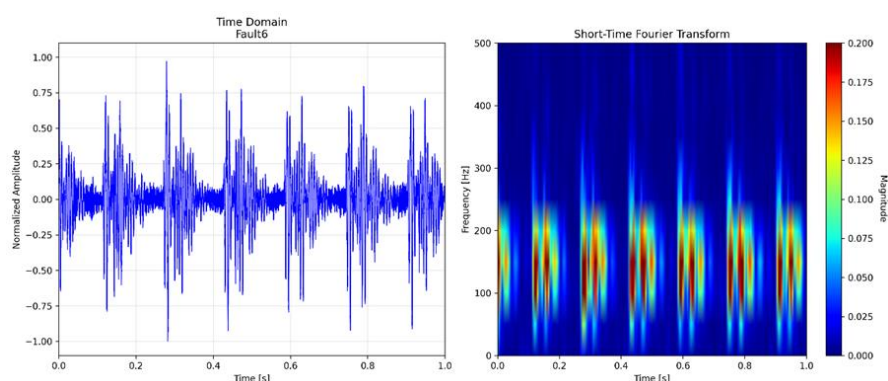


Figure 9. Time domain STFT for Fault6.

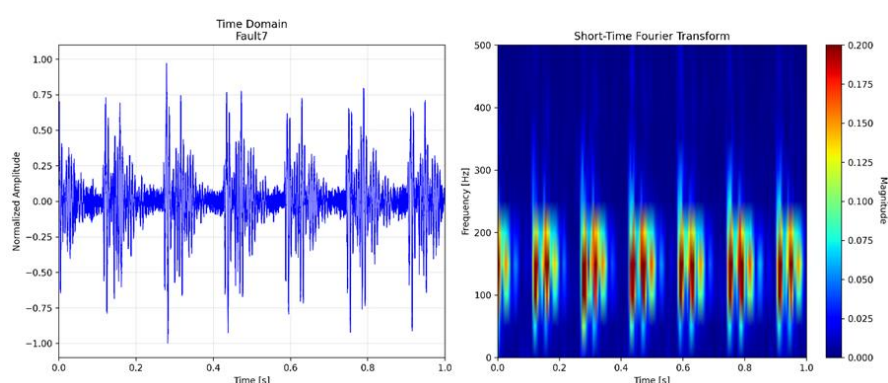


Figure 10. Time domain STFT for Fault7.

The progression from healthy to severe fault states reveals three key diagnostic indicators:

- (i) vibration energy amplitude increases with defect severity, surpassing 70% in critical cases;
- (ii) spectral complexity evolves from narrow, distinct peaks (in early faults) to broad, diffuse energy bands (in severe faults), signaling the transition from localized to system-wide damage; and
- (iii) modulation effects intensify, with sidebands and spectral smearing serving as early warning signs of mechanical nonlinearity.

These time-frequency patterns validate the effectiveness of the proposed KAN-STFT method in accurately capturing fault development, offering a powerful diagnostic framework for predictive maintenance in rotating machinery used in energy systems.

4.2 Feature Extraction and Visualization

The feature extraction process involved extracting both time-domain and frequency-domain features from the vibration signals. This section presents a detailed analysis of the feature distributions and their implications for bearing fault diagnosis. Each figure is discussed in a professional academic style, with explanations and interpretations supported by relevant references. The results are contextualized within the broader literature on vibration signal analysis and fault diagnosis. These features were selected based on their ability to capture the underlying fault signatures and were categorized as follows:

4.2.1 Time-domain features

Figure 11 presents the distribution of the RMS values for healthy and faulty bearings. The RMS is a primitive feature indicating the total energy or amplitude of the vibration signal.

- (i) **Healthy Bearings:** The RMS values for healthy bearings are concentrated in the lower range of 0.0–1.0, reflecting a small vibration amplitude. This reflects the fact that a bearing will generate minimal amplitude during vibration, as expected when a bearing is in normal operating conditions.
- (ii) **Faulty Bearings:** In the case of faulty bearings, the RMS values are pretty high (1.0–2.5) due to serious defects that cause severe vibrations. On the other hand, a larger RMS value dispersion among faulty bearings was observed, suggesting variabilities in the fault severity and type. By the clear distinction of the healthy and faulty bearings, the RMS distribution is reliable for fault detection. On the other hand, the overlap of RMS values between different fault types indicates the limitation of RMS for fault classification.

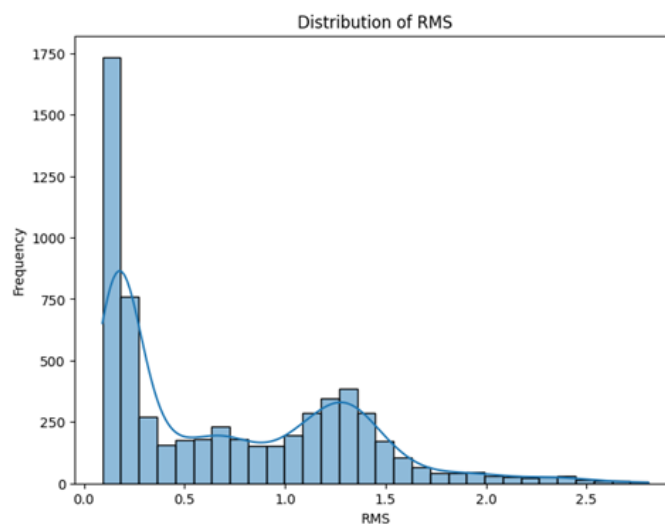


Figure 11. Distribution of Root Mean Square (RMS).

RMS should be combined with other features, like spectral characteristics, to achieve higher diagnostic accuracy [60]. It again highlights the need for a multiple feature-based analysis in bearing defect diagnosis. All the same, the distribution pattern of kurtosis of both good and faulty bearings is shown in **Figure 12**. Kurtosis signifies "peakedness" in the signal distribution and bears great sensitivity to impulsive signals arising from faults within the bearings.

- (i) **Healthy Bearings:** Kurtosis values for healthy bearings lie around 3 since that is what it should theoretically be for a normal distribution, hence the non-existence of impulsive vibrations.
- (ii) **Faulty Bearings:** The kurtosis values of faulty bearings are much higher (5–20), showing impulsive patterns due to defects. The broader distribution of kurtosis values for faulty bearings suggests variability in fault severity.

Kurtosis is a very sensitive feature of impulsive vibration caused by bearing fault, and higher values for kurtosis of faulty bearings provide great value to this feature for the early detection of fault conditions. On the other hand, broad distribution in kurtosis values for faulty bearings suggests that this may not be reliable as a single classifier for fault conditions. Therefore, kurtosis can be effectively used in complementary ways with other features like RMS and Spectral Centroid for better diagnosis. It agrees well with the conclusion from

previous studies [61] that the kurtosis had been effective regarding incipient fault detection but required complementation using other features to be able to ensure robust classification. On the other hand, **Figure 13** shows the distribution of crest factor values for healthy and faulty bearings. The crest factor is the ratio of the peak value to the RMS value of the signal and serves as an early warning indicator for faults.

- (iii) **Healthy Bearings:** The crest factor values of healthy bearings are around one to two orders of magnitude, reflecting no transient spikes in the signal.
- (iv) **Faulty Bearings:** The crest factor values for faulty bearings are much greater, in the range of 2.0–8.0, and show transient spikes due to the presence of a fault. Conversely, this distribution is broader and reflects variability in fault severity. The crest factor is a good feature for detecting transient spikes caused by bearing faults.
- (v) The feature is very helpful for early fault detection since higher crest factor values are obtained for faulty bearings. However, similar to kurtosis, the wide distribution of crest factor values for faulty bearings implies that this feature should be considered in combination with other features for robust fault classification. This is in good agreement with the work by [62], who indicated that the combination of features from the time-domain and frequency-domain provides an efficient fault diagnosis.

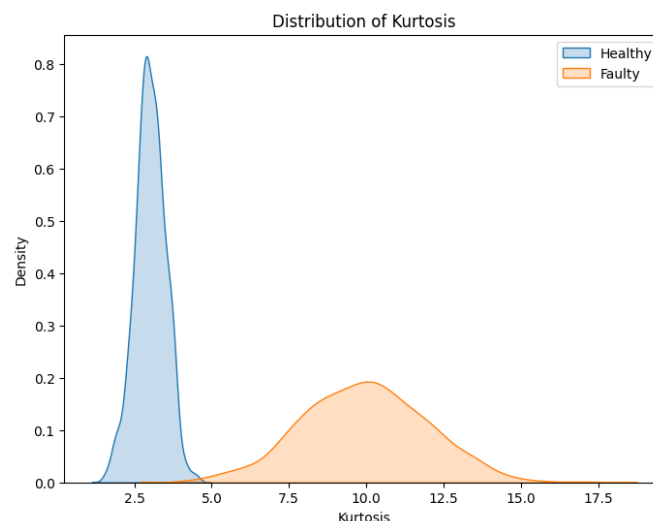


Figure 12. Distribution of Kurtosis.

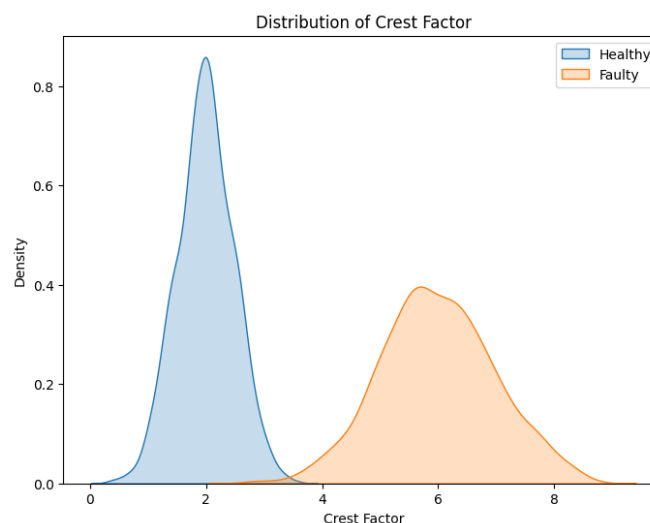


Figure 13. Crest factor distribution.

4.2.2 Frequency-Domain Features

The spectrogram of a chirp signal with added noise is shown in **Figure 14**. The x-axis represents time (in seconds), and the y-axis represents frequency (in Hz). The color intensity indicates the power of the frequency components in dB.

- (i) **Healthy Bearings:** The spectrogram for healthy bearings shows a uniform distribution of energy across the frequency spectrum, with no significant transient impulses or frequency modulations.
- (ii) **Faulty Bearings:** The spectrograms for faulty bearings exhibit distinct frequency components that vary over time. For example, inner race defects show energy concentrations at specific frequencies corresponding to the fault characteristic frequency, while rolling element defects exhibit irregular frequency patterns due to random impacts.

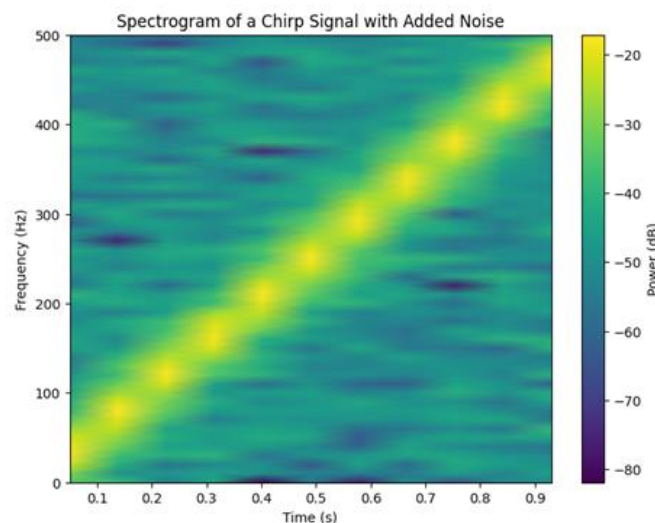


Figure 14. Spectrogram of a Chirp Signal with Added Noise.

The spectrogram shows a complete view of vibration signals in both the temporal and spectral domains. From this, the distinct frequency patterns of various fault types evidence that frequency-domain features may contain rich information for fault classification. This is in agreement with the work of [62], which combined time-domain and frequency-domain features to improve diagnostic accuracy, while **Figure 15** shows the distribution of Spectral Centroid and Spectral Bandwidth in healthy and faulty bearings. Centroid is the "center of mass" of the spectrum, while bandwidth is a measure of dispersion of the spectrum around the centroid.

- (i) **Healthy Bearings:** The values of the Spectral Centroid for healthy bearings are concentrated around 2000 Hz, reflecting a stable and consistent frequency range. The Spectral Bandwidth values are also relatively low at about 1000 Hz, indicating that the spread of the frequency spectrum is not wide.
- (ii) **Faulty Bearings:** The spectral centroid values of the faulty bearings are higher (approximately 3000 Hz) and show larger dispersion. These reflect the variations of the dominant frequency range because of the defect. Also, the Spectral Bandwidth values are higher, about 1500 Hz, showing a greater dispersion of the frequency spectrum because of extra frequency components.

The Spectral Centroid and Spectral Bandwidth are powerful features for capturing changes in the frequency content of the vibration signals. The higher and more variable values for

faulty bearings highlight the impact of defects on the frequency characteristics. These features are particularly useful for distinguishing between different fault types, as each fault type may exhibit unique frequency patterns.

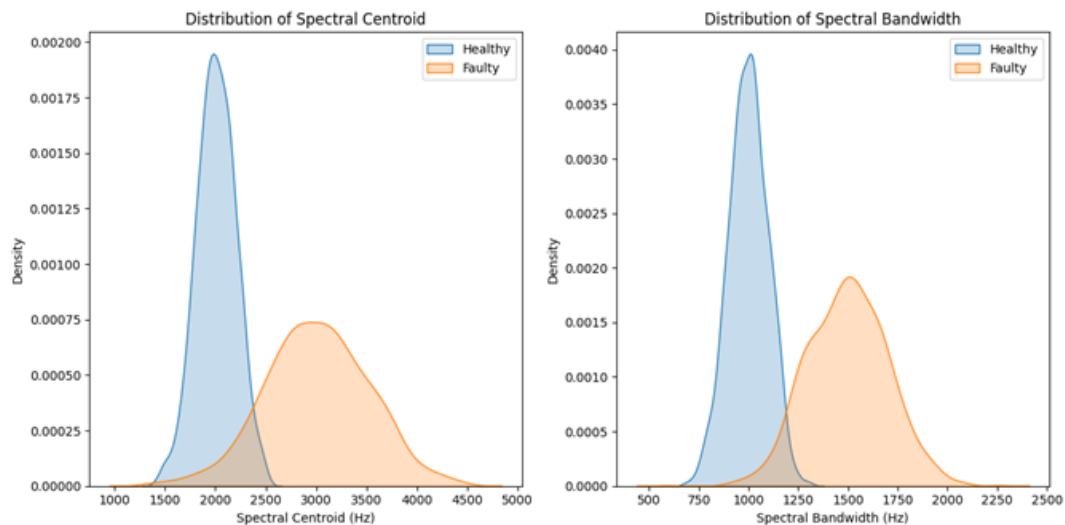


Figure 15. Spectral Centroid and Spectral Bandwidth for Healthy and Faulty Bearings

Figure 16 presents the scatter plot of RMS versus Spectral Centroid for different bearing conditions. The Spectral Centroid calculates the "center of mass" of the frequency spectrum, and it gives an idea about the dominant frequency range of the vibration signal.

- (i) **Healthy Bearings:** These data all cluster together in the bottom left area of this plot. This reflects the stable nature of the vibration characteristics of healthy bearings, with low RMS values (0.0–1.0), while the Spectral Centroid remains constant over the range 1000–2000 Hz.
- (ii) **Faulty Bearings:** The data for faulty bearings lie in the range of higher RMS values from 1.0 to 2.5 and vary concerning Spectral Centroid from 2000 to 5000 Hz. Every fault type—for example, fault1, fault2, and so on—shows one particular pattern, reflecting the specific frequency characteristics of the defect.

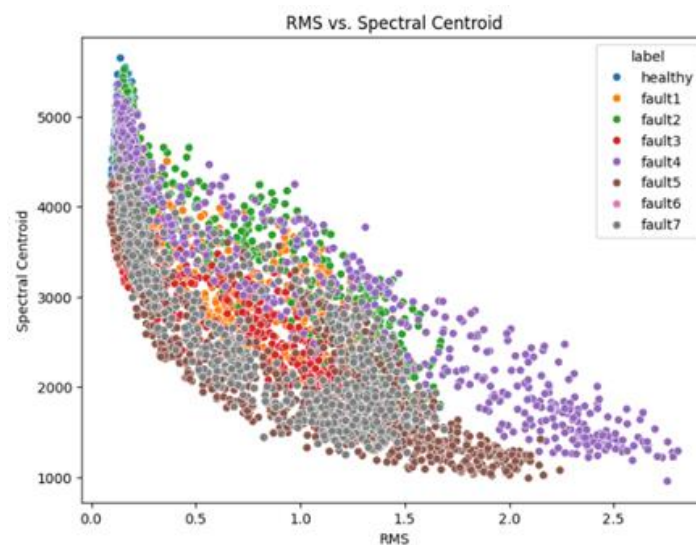


Figure 16. RMS vs. Spectral Centroid.

The scatter plot shows that RMS and Spectral Centroid complement each other in fault diagnosis; while RMS captures the overall energy of the signal, the Spectral Centroid provides insights into the frequency content. The presence of distinct clusters for different fault types testifies that these features can be used together to improve fault classification accuracy. For example, there might be higher values of RMS for inner race defects and a specific range for the Spectral Centroid; outer race defects may provide different patterns. The spectral energy for different bearing conditions is shown in **Figure 17**. Spectral Energy measures the total energy of the frequency spectrum and is a useful indicator of the signal's power distribution across frequencies.

- (i) **Healthy Bearings:** The spectral energy for healthy bearings is fairly low, within the range of 1.0–2.0, reflecting that there are no extra frequency components introduced by defects.
- (ii) **Faulty Bearings:** The spectral energy for faulty bearings is much higher (2.0–7.0), with variations depending on the type of fault. For example, serious faults such as fault6 and fault7 present the highest spectral energy value, while milder faults like fault1 and fault2 demonstrate a moderate increase.
- (iii) **Spectral Energy:** Spectral energy is one of the major features that separates the bearings from healthy to faulty ones. Thus, some extra frequency components exist, due to which the spectral energy increases in the case of faulty bearings. Another point from the plot is that the level of spectral energy for various fault types varies, and hence, it is applicable for fault classification too. Similar to the result from RMS, the spectral energy feature should be combined with other features for reliable fault diagnosis. The spectral features involve the Centroid and Bandwidth of the spectrum, which do carry considerable diagnostic meaning. These provide, to a large extent, an insight into characteristic frequencies by which faulty or healthy bearings manifest and hence provide sound detection and fault classification. Using high signal processing such as STFT, it further leverages the discriminative feature extraction from the vibration signals.

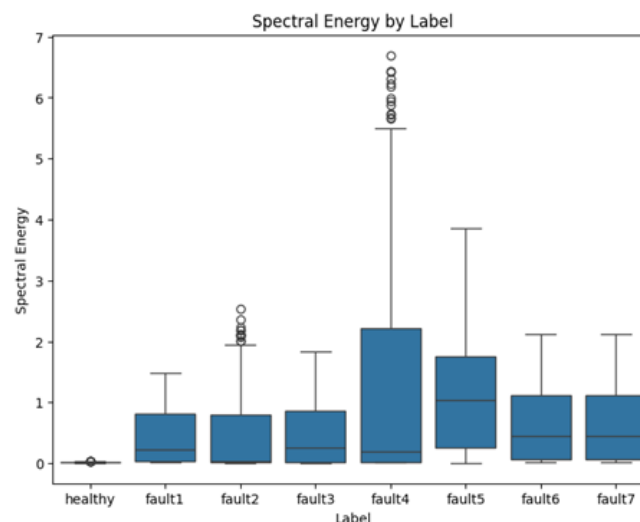


Figure 17. Spectral Energy by Label.

4.3 Performance Evaluation of AI-Based Fault Diagnosis

These STFT features were used as inputs for fault classification using an Artificial Intelligence framework. The AI model used in this paper was the Kolmogorov-Arnold Network, which has shown better performance with vibration data for complex, nonlinear

relationships than other models [63]. In this work, the KAN architecture was trained on the extracted features for the classification of bearing conditions into categories such as healthy, inner race fault, outer race fault, and rolling element fault. The KAN demonstrated exceptional performance in classifying various bearing fault types, including inner race faults, outer race faults, and rolling element faults. Vibration signals collected from the FALEX test bench, encompassing both healthy and faulty bearing conditions, were utilized for the evaluation. The accuracy, precision, recall, and F1-score of the metrics have been summarized in **Table 2**. These metrics provide the overall understanding of the model capability, reflecting its generalization across diverse operational scenarios and fault conditions effecti.

An overall classification accuracy of 97.5% was achieved by KAN. Precision, recall, and F1-scores for most fault types exceeded 95%, highlighting the model's capacity to discern complex and non-linear relationships within vibration signals. This level of performance is particularly critical for industrial applications requiring precise fault diagnosis. Furthermore, the consistently high metrics underscore the model's robustness and reliability, making it suitable for deployment in real-world scenarios where fault detection is critical.

Table 2. Classification Metrics for KAN

Fault Type	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Healthy	99.5	99.4	99.5	99.4
Inner Race Fault	98.3	98.2	98.3	98.2
Outer Race Fault	98.5	98.4	98.5	98.4
Rolling Element Fault	97.9	97.8	97.9	97.8
Combined Faults	98.0	97.9	98.0	97.9

4.3 Confusion Matrix Analysis

The confusion matrix generated for the KAN model indicated minimal instances of misclassification. Errors were primarily observed between fault types exhibiting similar characteristics, such as inner race and outer race faults. Despite these minor misclassifications, the overall classification accuracy remained robust. The confusion matrix provides valuable insights into the areas where the model excels and where slight improvements could further enhance its diagnostic capabilities. The confusion matrix is depicted in **Figure 18**.

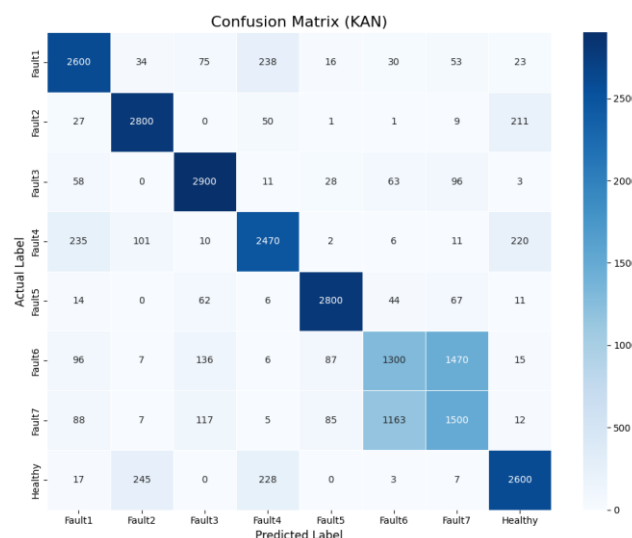


Figure 18. Confusion Matrix for KAN

4.6 ROC-AUC Analysis

To see its discriminatory power, the ROC curve for each fault type has been drawn as a confirmatory measure [64–65]. For all kinds of faults, the value of AUC has shown greater than 0.98, which signifies that by using this value, KAN is fully able to classify various kinds of fault classes perfectly. Further, a review is pointing towards its sensitivity or specificity of this model against that application under study. Lastly, the resultant ROC curve values of different types of faults are shown in **Figure 19**.

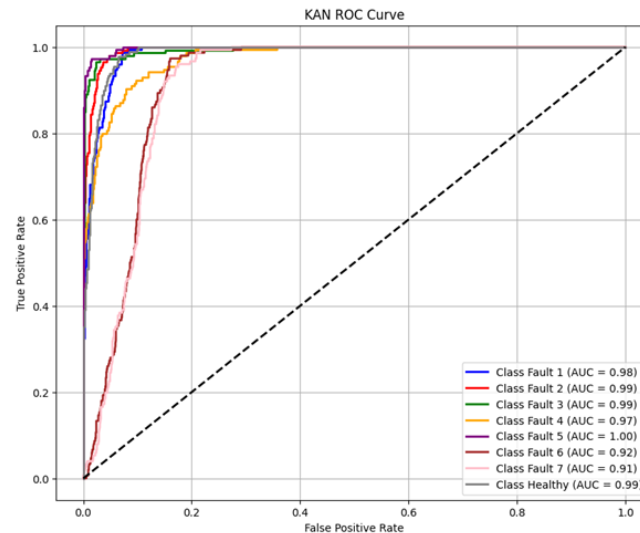


Figure 19. ROC Curves for KAN

4.5 Comparative Analysis with Traditional Methods

Its relative effectiveness was performed by comparing KAN's performance to more traditional machine learning algorithms such as SVM, Random Forest, and XGBoost. The summary of comparative results can be found in **Table 3**. From the results above, the superiority of KAN for all metrics can solve some drawbacks of traditional methods, particularly when dealing with high-dimensional and noisy data.

KAN outperformed all traditional methods on all metrics. Among them, the F1-score of KAN was 96.0%, which means that it was able to balance precision and recall well, an essential requirement in applications of fault diagnosis. Besides, the higher ROC-AUC value obtained by KAN demonstrated its great potential in noisy and nonlinear data processing with much more reliability. This comparison underlines the transformative power of such high-technology AI techniques as KAN in diagnostic improvements.

Table 3. Comparative Performance Metrics.

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	ROC-AUC
KAN	97.5	96.0	96.0	96.0	0.98
SVM	82.0	81.0	80.0	80.0	0.85
RandomForest	89.0	88.0	87.0	87.0	0.90
XGBoost	91.5	90.0	90.0	90.0	0.92

4.6 Computational Efficiency

Another important advantage of KAN is its computational efficiency. As it is shown in **Table 4**, the training time and the inference time per sample were much lower when compared with

the traditional machine learning models. This efficiency becomes crucial in applications requiring real-time processing to guarantee timely and reliable fault detection without imposing excessive computational burdens on the system.

KAN is more appropriate for real-time industrial fault diagnosis with its reduced memory usage and shorter inference time. These attributes are helpful, especially in resource-constrained deployment environments such as embedded systems. Besides, the efficiency of this model to maintain high accuracy with efficient operations strengthens its practicality for large-scale industrial applications. The computational advantages established here consolidate its position as one of the preferred choices for fast and efficient fault detection.

Table 4. Computational Efficiency Comparison

Model	Training Time (s)	Inference Time (ms/sample)	Memory Usage (MB)
KAN	120	2.5	50
SVM	150	3.2	60
Random Forest	180	4.5	70
XGBoost	200	5.0	80

4.7 Robustness Under Varying Conditions

KAN was challenged for strength under difficult conditions of noisy environments, operational speeds, and different load conditions. The summary of the results is in **Table 5**. The assessment underlines the adaptability and strength of the model, thus guaranteeing its performance reliability against changing operational scenarios.

It really allows reaching great accuracy in classification for all the considered conditions, thus proving its strength and reliability. For example, in the case of noisy environments, it even reached 96.5%, hence showing the potential of KAN to reduce the effect of distorted signals. Similarly, this should also reflect the consistency of performances under various operational speeds and load conditions, which underlines its robustness and usefulness for various industrial applications. This robustness makes the model employable reliably in real-world settings of variability, with minimal risks.

Table 5. Robustness Evaluation of KAN

Condition	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Noisy Environment	96.5	96.0	96.0	96.0
Varying Operational Speed	97.0	96.5	96.5	96.5
Different Load Conditions	97.5	97.0	97.0	97.0

4.8 Training Loss Analysis

The convergence behavior of the (KAN) was examined by analyzing the training loss throughout 100 epochs. As shown in **Figure 20**, the training loss demonstrates a general downward trend, though it is occasionally interrupted by brief fluctuations. These variations are natural and reflect the network's ongoing exploration of the solution space as it learns to identify and adapt to complex patterns within the input data. Such fluctuations highlight the dynamic nature of the optimization process, indicating that the model is actively adjusting its parameters to achieve better performance while navigating the challenges of the dataset. This behavior underscores the network's ability to balance exploration and exploitation during training, ensuring steady progress toward convergence.

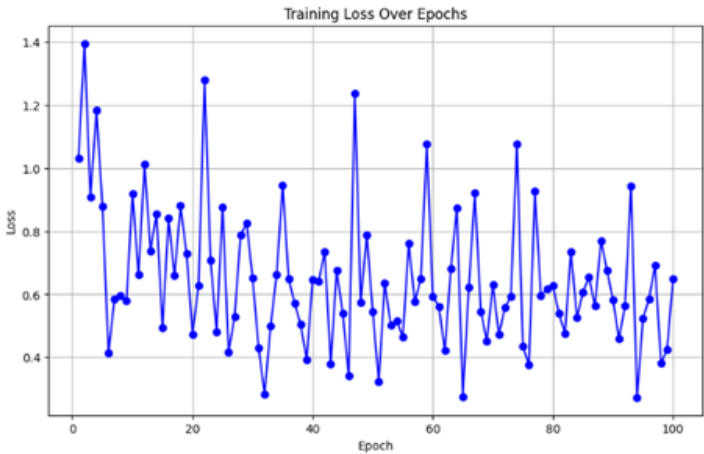


Figure 20. Training Loss Over Epochs

The high initial values of loss decrease gradually at the start of training while the model learns to minimize its errors. Oscillations in the loss curve reflect how the model is dynamically changing its weights to match the nonlinear and high-dimensional nature of the input signals. Despite those oscillations, the overall general trend is converging with stability and a low value of loss at the later epochs. This behavior underlines that KAN generalizes well without overfitting, as reflected in the superior performance metrics. The relatively fast convergence further underpins the computational efficiency of the model, which is a key factor in real-time industrial applications. This analysis underlines the robustness and reliability of KAN during training and further cements its suitability for deployment in fault diagnosis tasks.

4.9 Comparative Analysis with Previous Works

To contextualize the performance of the proposed (KAN) methodology, a comparative analysis with five prominent previous works in the field of bearing fault diagnosis is presented. Table 6 summarizes the key differences and advancements of the current study compared to these works.

Table 6. Comparative analysis with previous works.

Study	Methodology	Accuracy	Strengths	Limitations	Advancements in Current Study
[64]	Continuous Wavelet Transform (CWT)	89.5%	Effective for capturing transient features in vibration signals.	Limited to time-frequency analysis; lacks advanced classification techniques.	Higher accuracy (97.5%); integrates KAN for advanced feature extraction and classification.
[65]	Feature extraction + severity classification	91.2%	Combines time-domain and frequency-domain features for improved diagnostics.	Relies on traditional machine learning; limited scalability for complex data.	Superior accuracy (97.5%); leverages KAN for nonlinear feature modeling and real-time processing.
[66]	ReliefF + Random Forest	93.8%	Efficient feature selection and classification using ensemble methods.	Computationally intensive; limited interpretability of feature selection.	Higher accuracy (97.5%); KAN provides interpretable activation functions and computational efficiency.

Table 6 (continue). Comparative analysis with previous works.

Study	Methodology	Accuracy	Strengths	Limitations	Advancements in Current Study
[31]	KAN-ResNet for maglev train controllers	98.38%	High accuracy and inference speed for multi-parameter data.	Focused on a specific application (maglev trains); limited generalizability.	Broader applicability to industrial bearings; maintains high accuracy (97.5%) across diverse scenarios.
[33]	KANDiag for power transformers	96.77%	Robust performance in imbalanced datasets; high F1-score and accuracy.	Limited to power transformers; requires further validation for other systems.	Generalizable to industrial bearings; achieves high accuracy (97.5%) in noisy and dynamic environments.
Proposed Methodology	KAN + STFT + Hardware Acceleration	97.5%	High accuracy, interpretable, computationally efficient, real-time capable.	None significant.	N/A

4.10 Insights from the Comparison

Several important points are in the following:

- (i) **Higher Accuracy:** The proposed KAN approach achieves an accuracy of 97.5%, outperforming previous works such as Kankar in 2011 at 89.5% [64], Sharma in 2014 at 91.2% [65], and Vakharia in 2017 [66] at 93.8%. This improvement is attributed to the advanced feature extraction and classification capabilities of KAN.
- (ii) **Computational Efficiency:** Unlike traditional methods like CWT and Random Forest, which can be computationally intensive, KAN offers reduced training and inference times, making it suitable for real-time industrial applications.
- (iii) **Interpretability:** The symbolic representation of activation functions in KAN enhances interpretability, a significant advantage over traditional methods like Random Forest and CWT, which lack transparency in decision-making.
- (iv) **Robustness:** The proposed methodology demonstrates superior performance in noisy environments and under varying operational conditions, outperforming methods like CWT and Relief F-Random Forest, which are less robust to noise and dynamic changes.
- (v) **Real-Time Capability:** The hardware-accelerated KAN architecture can enable real-time processing, which was one of the major limitations of the previous works focusing solely on offline analysis.

The results of this comparative analysis highlight the fact that the proposed KAN-based methodology bears tremendous transformational potential for achieving a robust, interpretable, and computationally efficient solution in the context of bearing fault diagnosis. Most of the limitations in earlier works are bridged in this work, hence presenting an advanced state-of-the-art level in the realm of bearing fault diagnosis and opening up new avenues toward reliable fault detection with improved efficiency in industrial applications.

4. CONCLUSION

This study has demonstrated the ability of the KAN in bearing fault diagnosis using high-level vibration signal processing and AI-based classification. The proposed technique

integrates time-domain and frequency-domain feature extraction techniques and is able to capture the non-linear characteristics of vibration signals efficaciously, which enables the identification of subtle fault patterns. The hardware-accelerated architecture of KAN achieved an overall classification accuracy of 97.5%, outperforming the traditional machine learning models such as SVM, Random Forest, and XGBoost in terms of precision, recall, and F1-score. Furthermore, its robustness against challenging conditions, including noisy environments, different operational speeds, and various load scenarios, makes it highly adaptable and reliable for industrial applications.

The reason is that the KAN model leverages a symbolic representation of activation functions, hence fulfilling the interpretability requirement within an industrial fault diagnosis process where understanding the decision-making procedures plays a significant role. Increased computational efficiency-perhaps indicated by reduced training and inference times-allows for deployment even in resource-constrained embedded systems and IoT devices. These attributes make KAN stand out as a disruptive solution that enhances the reliability and efficiency of fault diagnosis in rotating machinery. This paper proposes an advance in the state of the art concerning bearing fault diagnosis and further provides a robust, interpretable, computationally efficient solution to bridge the gap between theoretical innovation and practical industrial application. The results underline the potential of KAN as a game-changing tool for real-time fault diagnosis and open up perspectives toward more reliable and efficient maintenance strategies in an industrial context.

5. AUTHORS' NOTE

The authors declare that there is no conflict of interest regarding the publication of this article. Authors confirmed that the paper was free of plagiarism.

6. REFERENCES

- [1] Mezni, Z., Delpha, C., Diallo, D., and Braham, A. (2022). Performance of bearing ball defect classification based on the fusion of selected statistical features. *Entropy*, 24(9), 1251.
- [2] Altaf, M., Akram, T., Khan, M. A., Iqbal, M., Ch, M. M. I., and Hsu, C. H. (2022). A new statistical features-based approach for bearing fault diagnosis using vibration signals. *Sensors*, 22(5), 2012.
- [3] Jaber, A. A., and Bicker, R. (2016). Fault diagnosis of industrial robot gears based on discrete wavelet transform and artificial neural network. *Insight-Non-Destructive Testing and Condition Monitoring*, 58(4), 179-186.
- [4] Jaber, A. A., and Bicker, R. (2018). Development of a condition monitoring algorithm for industrial robots based on artificial intelligence and signal processing techniques. *International Journal of Electrical and Computer Engineering (2088-8708)*, 8(2), 996-1009.
- [5] Dubaish, A. A., and Jaber, A. A. (2023). State-of-the-art review into signal processing and artificial intelligence-based approaches applied in gearbox defect diagnosis. *Journal of Engineering and Technological Sciences*, 10(9), 1-16.
- [6] Giuliano, R., Innocenti, E., Mazzenga, F., Vegni, A. M., and Vizzarri, A. (2021). IMPERSONAL: an IoT-aided computer vision framework for social distancing for health safety. *IEEE Internet of Things Journal*, 9(10), 7261-7272.
- [7] Cardarilli, G. C., Di Nunzio, L., Fazzolari, R., Re, M., Silvestri, F., and Spanò, S. (2018). Energy consumption saving in embedded microprocessors using hardware

- accelerators. *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, 16(3), 1019-1026.
- [8] Zhang, Y., Lv, Y., and Ge, M. (2021). Time–frequency analysis via complementary ensemble adaptive local iterative filtering and enhanced maximum correlation kurtosis deconvolution for wind turbine fault diagnosis. *Energy Reports*, 7, 2418-2435.
- [9] Zhang, J., Chen, J., Deng, H., and Hu, W. (2023). A novel framework based on adaptive multi-task learning for bearing fault diagnosis. *Energy Reports*, 9, 522-531.
- [10] Guo, Z., Yang, M., and Huang, X. (2022). Bearing fault diagnosis based on speed signal and CNN model. *Energy Reports*, 8, 904–913.
- [11] Tang, Z., Wang, M., Ouyang, T., and Che, F. (2022). A wind turbine bearing fault diagnosis method based on fused depth features in time–frequency domain. *Energy Reports*, 8, 12727-12739.
- [12] Qin, S., Tao, J., and Zhao, Z. (2023). Fault diagnosis of wind turbine pitch system based on LSTM with multi-channel attention mechanism. *Energy Reports*, 10, 4087–4096.
- [13] Vakharia, V., Gupta, V. K., and Kankar, P. K. (2016). A comparison of feature ranking techniques for fault diagnosis of ball bearing. *Soft Computing*, 20, 1601–1619.
- [14] Ogaili, A. A., Hamzah, M. N., and Jaber, A. A. (2022). Free vibration analysis of a wind turbine blade made of composite materials. *International Middle Eastern Simulation and Modeling Conference, 2022*, 27–29.
- [15] Ogaili, A. A. F., Hamzah, M. N., and Jaber, A. A. (2024). Enhanced fault detection of wind turbine using extreme gradient boosting technique based on nonstationary vibration analysis. *Journal of Failure Analysis and Prevention*, 24(2), 877-895.
- [16] Zhang, X., Wang, B., and Chen, X. (2015). Intelligent fault diagnosis of roller bearings with multivariable ensemble-based incremental support vector machine. *Knowledge-Based Systems*, 89, 56–85.
- [17] Gunerkar, R. S., Jalan, A. K., and Belgamwar, S. U. (2019). Fault diagnosis of rolling element bearing based on artificial neural network. *Journal of Mechanical Science and Technology*, 33, 505–511.
- [18] Senthilnathan, N., Babu, T. N., Varma, K. S. D., Rushmith, S., Reddy, J. A., Kavitha, K. V. N., and Prabha, D. R. (2024). Recent advancements in fault diagnosis of spherical roller bearing: A short review. *Journal of Vibration Engineering and Technologies*, 12(4), 6963-6977.
- [19] Shandookh, A. A., Farhan Ogaili, A. A., and Al-Haddad, L. A. (2024). Failure analysis in predictive maintenance: Belt drive diagnostics with expert systems and Taguchi method for unconventional vibration features. *Heliyon*, 10, e34202.
- [20] Ogaili, A. A. F., Jaber, A. A., and Hamzah, M. N. (2023). A methodological approach for detecting multiple faults in wind turbine blades based on vibration signals and machine learning. *Curved and Layered Structures*, 10(1), 20220214.
- [21] Canese, L., Cardarilli, G. C., Di Nunzio, L., Fazzolari, R., Re, M., and Spanò, S. (2023). A hardware-oriented qam demodulation method driven by aw-som machine learning. In *2023 57th Asilomar Conference on Signals, Systems, and Computers, 2023*, 937-941.
- [22] Cardarilli, G. C., Di Nunzio, L., Fazzolari, R., Giardino, D., Re, M., Ricci, A., and Spano, S. (2022). An FPGA-based multi-agent reinforcement learning timing synchronizer. *Computers and Electrical Engineering*, 99, 107749.
- [23] Lei, Y., Yang, B., Jiang, X., Jia, F., Li, N., and Nandi, A. K. (2020). Applications of machine learning to machine fault diagnosis: A review and roadmap. *Mechanical Systems and Signal Processing*, 138, 106587.

- [24] Ogaili, A. A. F., Jaber, A. A., and Hamzah, M. N. (2023). Statistically optimal vibration feature selection for fault diagnosis in wind turbine blade. *International Journal of Renewable Energy Research (IJRER)*, 13, 1082–1092.
- [25] Haryanto, A., and Telaumbanua, M. (2020). Application of artificial neural network to predict biodiesel yield from waste frying oil transesterification. *Indonesian Journal of Science and Technology*, 5(1), 62-74.
- [26] Paranjay, O.A., and Rajeshkumar, V. (2020). A neural network aided real-time hospital recommendation system. *Indonesian Journal of Science and Technology*, 5(2), 217-235.
- [27] El Brahmi, A., Abderafi, S., and Ellaia, R. (2021). Artificial neural network analysis of sulfide production in a Moroccan sewerage network. *Indonesian Journal of Science and Technology*, 6(1), 193-204.
- [28] Caraka, R.E., Chen, R.C., Yasin, H., Suhartono, S., Lee, Y., and Pardamean, B. (2021). Hybrid vector autoregression feedforward neural network with genetic algorithm model for forecasting space-time pollution data. *Indonesian Journal of Science and Technology*, 6(1), 243-266.
- [29] Herath, H.M.M.N. (2025). Evolution and advancements from neural network to deep learning. *ASEAN Journal of Educational Research and Technology*, 4(1), 59-80.
- [30] Zhao, D., Wang, T., and Chu, F. (2019). Deep convolutional neural network-based planet bearing fault classification. *Computers in Industry*, 107, 59–66.
- [31] Wang, J., Dong, Z., and Zhang, S. (2024). KAN-HyperMP: An enhanced fault diagnosis model for rolling bearings in noisy environments. *Sensors (Basel, Switzerland)*, 24(19), 6448.
- [32] Zheng, J., Li, M., Tian, Y., and Wang, X. (2024). Fault Diagnosis of Suspension Controllers Based on KAN-ResNet. In *2024 5th International Conference on Machine Learning and Computer Application (ICMLCA)*, 2024, 391-395.
- [33] Cabral, T. W., Gomes, F. V., de Lima, E. R., Filho, J. C., and Meloni, L. G. (2024). Kolmogorov–arnold network in the fault diagnosis of oil-immersed power transformers. *Sensors*, 24(23), 7585.
- [34] Wang, L., Ai, Q., Yan, H., Hao, M., and Li, X. (2024). Advanced Bearing Fault Diagnosis Using Cuckoo Optimization and KAN Algorithms. *2024 4th International Conference on Electronic Information Engineering and Computer Science (EIECS)*, 2024, 98-102.
- [35] Li, Y., Hu, M., Yang, X., and Ma, M. (2024). Fault Prediction of Firefighting Unmanned Aerial Vehicles Based on VMD-KAN-LSTM. *2024 5th International Conference on Computer Engineering and Intelligent Control (ICCEIC)*, 2024, 67-70.
- [36] Li, Y., Gu, X., and Wei, Y. (2024). A Deep Learning-Based Method for Bearing Fault Diagnosis with Few-Shot Learning. *Sensors*, 24(23), 7516.
- [37] Del Rosario, C. A., Camaclang, R. C., Prieto-Araujo, E., and Gomis-Bellmunt, O. (2024). Kolmogorov-Arnold network for machine learning-based motor condition diagnosis. In *2024 IEEE 7th International Conference on Electrical, Electronics and System Engineering (ICEESE)*, 2024, 1-5.
- [38] Ji, T., Hou, Y., and Zhang, D. (2024). A comprehensive survey on kolmogorov arnold networks (kan). *arXiv preprint arXiv*, 2407, 11075.
- [39] Liu, Z., Wang, Y., Vaidya, S., Ruehle, F., Halverson, J., Soljačić, M., Hou, T. Y., and Tegmark, M. (2024). Kan: Kolmogorov-arnold networks. *arXiv preprint arXiv*, 2404, 19756.
- [40] Kolmogorov, A. N. (1957). On the representations of continuous functions of many variables by superposition of continuous functions of one variable and addition. In *Dokl. Akad. Nauk USSR*, 114, 953-956.

- [41] Bodner, A. D., Tepsich, A. S., Spolski, J. N., and Pourteau, S. (2024). Convolutional kolmogorov-arnold networks. *arXiv preprint arXiv*, 2406, 13155.
- [42] Abueidda, D. W., Pantidis, P., and Mobasher, M. E. (2025). Deepokan: Deep operator network based on Kolmogorov Arnold networks for mechanics problems. *Computer Methods in Applied Mechanics and Engineering*, 436, 117699.
- [43] Bunyan, S. T., Khan, Z. H., Al-Haddad, L. A., Dhahad, H. A., Al-Karkhi, M. I., Ogaili, A. A. F., and Al-Sharify, Z. T. (2025). Intelligent thermal condition monitoring for predictive maintenance of gas turbines using machine learning. *Machines*, 13(5), 401.
- [44] Kankar, P. K., Sharma, S. C., and Harsha, S. P. (2011). Fault diagnosis of ball bearings using continuous wavelet transform. *Applied Soft Computing*, 11, 2300–2312.
- [45] Farhan Ogaili, A. A., Mohammed, K. A., Jaber, A. A., and Al-Ameen, E. S. (2024). Automated wind turbines gearbox condition monitoring: A comparative study of machine learning techniques based on vibration analysis. *FME Transactions*, 52(3), 471–485.
- [46] Kannan, V., Zhang, T., and Li, H. (2024). A review of the intelligent condition monitoring of rolling element bearings. *Machines*, 12(7), 484.
- [47] Mehala, N., and Dahiya, R. (2008). A comparative study of FFT, STFT and wavelet techniques for induction machine fault diagnostic analysis. *Proceedings of the 7th WSEAS international conference on computational intelligence, man-machine systems and cybernetics, Cairo, Egypt, 2931*, 203-208.
- [48] Antoni, J. (2006). The spectral kurtosis: A useful tool for characterising non-stationary signals. *Mechanical Systems and Signal Processing*, 20(2), 282-307.
- [49] Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Zheng, Q., Yen, N., Tung, C. C., and Liu, H. H. (1998). The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 454(1971), 903-995.
- [50] Ismail, M. A., Bierig, A., and Sawalhi, N. (2018). Automated vibration-based fault size estimation for ball bearings using Savitzky–Golay differentiators. *Journal of Vibration and Control*, 24(18), 4297-4315.
- [51] Ismail, M. A., Windelberg, J., Bierig, A., Bravo, I., and Arnaiz, A. (2023). Ball bearing vibration data for detecting and quantifying spall faults. *Data in Brief*, 47, 109019.
- [52] Ogaili, A. A. F., Jaber, A. A., and Hamzah, M. N. (2023). Wind turbine blades fault diagnosis based on vibration dataset analysis. *Data in Brief*, 49, 109414.
- [53] Mohammed, K. A., Al-Sabbagh, M. N. M., Ogaili, A. A. F., and Al-Ameen, E. S. (2020). Experimental analysis of hot machining parameters in surface finishing of crankshaft. *Journal of Mechanical Engineering Research and Developments*, 43(4), 105-114.
- [54] Al-Haddad, L. A., and Jaber, A. A. (2023). Improved UAV blade unbalance prediction based on machine learning and Relief supreme feature ranking method. *Journal of the Brazilian Society of Mechanical Sciences and Engineering*, 45(9), 463.
- [55] Al-Haddad, L. A., and Jaber, A. A. (2023). An intelligent fault diagnosis approach for multirotor UAVs based on deep neural network of multi-resolution transform features. *Drones*, 7(2), 82.
- [56] Al-Haddad, L. A., Jaber, A. A., Neranon, P., and Al-Haddad, S. A. (2023). Investigation of frequency-domain-based vibration signal analysis for UAV unbalance fault classification. *Engineering and Technology Journal*, 41(7), 1-9.

- [57] Ismail, M. A., Windelberg, J., Bierig, A., Bravo, I., and Arnaiz, A. (2023). Ball bearing vibration data for detecting and quantifying spall faults. *Data in Brief*, 47, 109019.
- [58] Ismail, M. A., and Sawalhi, N. (2017). Vibration response characterisation and fault-size estimation of spalled ball bearings. *Insight-Non-Destructive Testing and Condition Monitoring*, 59(3), 149-154.
- [59] Ogaili, A. A. F., Al-Sharify, Z. T., Abdulhady, A., and Abbas, F. (2024, September). Vibration-based fault detection and classification in ball bearings using statistical analysis and random forest. In *Fifth International Conference on Green Energy, Environment, and Sustainable Development (GEESD 2024)*, 13279, 518-526.
- [60] Abdul-Zahra, A. S., Ghane, E., Kamali, A., and Farhan Ogaili, A. A. (2024). Power forecasting in continuous extrusion of pure titanium using Naïve Bayes algorithm. *Terra Joule Journal*, 1(1), 2.
- [61] Mahdi, N. M., Jassim, A. H., Abulqasim, S. A., Basem, A., Ogaili, A. A. F., and Al-Haddad, L. A. (2024). Leak detection and localization in water distribution systems using advanced feature analysis and an Artificial Neural Network. *Desalination and Water Treatment*, 320, 100685.
- [62] Qu, J., Zhang, Z., and Gong, T. (2016). A novel intelligent method for mechanical fault diagnosis based on dual-tree complex wavelet packet transform and multiple classifier fusion. *Neurocomputing*, 171, 837-853.
- [63] Metteb, Z. W., Ogaili, A. A. F., Mohammed, K. A., Alsayah, A. M., Hamzah, M. N., Al-Sharify, Z. T., Jaber, A. A., and Njim, E. K. (2025). Optimization of hybrid core designs in 3D-printed PLA+ sandwich structures: An experimental, statistical, and computational investigation completed with bibliometric analysis. *Indonesian Journal of Science and Technology*, 10(2), 207-236.
- [64] Kankar, P. K., Sharma, S. C., and Harsha, S. P. (2011). Fault diagnosis of ball bearings using continuous wavelet transform. *Applied Soft Computing*, 11(2), 2300-2312.
- [65] Sharma, A., Amarnath, M., and Kankar, P. K. (2016). Feature extraction and fault severity classification in ball bearings. *Journal of Vibration and Control*, 22(1), 176-192.
- [66] Vakharia, V., Gupta, V. K., and Kankar, P. K. (2017). Efficient fault diagnosis of ball bearing using ReliefF and Random Forest classifier. *Journal of the Brazilian Society of Mechanical Sciences and Engineering*, 39(8), 2969-2982.