AUT Journal of Electrical Engineering
10.22060/EEJ.2025.24182.5658

Motor Bearing Fault Diagnosis Based on Vibration Signal, Wavelet Denoising and CNN

Riyadh Abduljaleel mhalhal[1], Heydar Toossian Shandiz[2]*, Naser Pariz[3], Alaa Abdulhady Jaber[4]

[1]PhD candidate, Engineering faculty, Ferdowsi University of Mashhad, Mashhad, Iran
[2]Associate Professor, Engineering faculty, Ferdowsi University of Mashhad, Mashhad, Iran
htoosian@um.ac.ir
[3]Professor, Engineering faculty, Ferdowsi University of Mashhad, Mashhad, Iran
[4]Professor, Engineering faculty, University of Technology, Baghdad, Iraq

Abstract

Induction motors are broadly applied in different industries because of low cost, high efficiency, reliability. Although, failures in such motors could cause in significant issues like decreased efficiency, unexpected shutdown, ruin to other system sections. Diagnosing bearing failures is critical to reducing maintenance costs and operational failures. Bearing failures are a major cause of machine vibrations. Unfortunately, existing methods are optimized for controlled environments, and disregard realistic conditions such as variable load, time-varying rotational speeds, and non-stationary nature of vibration. This study presents an integration of time analysis and deep learning techniques to diagnose bearing failures under time-varying speeds and varying noise levels. In this study, we present an approach to diagnosing bearing failures employing vibration signals and convolutional neural networks (CNN) with Pre-processing of the vibration signal by using discrete wavelet transform (DWT) to remove the effect of Variable Frequency Drive (VFD) which causes odd harmonics. The experimental outcomes show that presented technique surpasses conventional techniques in the two computational efficiency as well as accuracy to diagnose bearing failures. This work paves the way for further research in the field of bearing fault diagnosis and provides a promising solution for real-world applications.

Keywords:

Induction motor bearing fault, Vibration signal, DWT, CNN, Time analysis.

List of abbreviations

fs	power supply frequency
fr	rotor shaft frequency
n	the number of rolling elements
Db	ball diameter
Dc	the pitch diameter
β	ball contact angle
fv	one of the characteristic vibration frequencies (*fi*, *fo*, *fc*)
f_i	Inner Race Fault Frequency (also known as BPFI - Ball Pass Frequency of Inner Race)
f_C	Cage Fault Frequency (also known as FTF - Fundamental Train Frequency).
f_o	Outer Race Fault Frequency (also known as BPFO - Ball Pass Frequency of Outer Race)
f_B	Ball Spin Frequency or Rolling Element Fault Frequency.
ψ (t)	function of zero mean
b	translation parameter
a	scale parameter
g[n]	low-pass filters
k^l_{ij}	convolution kernels
M	total number of input channels
b_{lj}	bias conforming to the j^{th} kernel
f( )	activation function
(Cr)	Crest Factor
(Kur)	Kurtosis
(RMS)	Root Mean Square
(Ske)	skewness
(Sta)	standard deviation
(Var)	variance
H	Entropy
h[n]	high-pass filters

1. Introduction

Induction motors are the backbone of many industrial systems, perhaps one of the most widely applied tools in various industrial applications such as domestic appliances, aerospace, petrochemical, and

chemical sectors. This widespread use is due to their high efficiency, ruggedness, and low cost. However, they are susceptible to mechanical, electrical, thermal, and environmental stresses that can lead to catastrophic failures [1]. Bearings play a crucial role in guiding and supporting shafts in rotating machinery, often operating in extreme environments. As essential components of these machines, bearings are a major source of failure due to harsh working conditions and increased pressures [2, 3]. Extensive research has demonstrated that bearing failure is a primary cause of most mechanical failures in rotating machinery, contributing to over 40% of induction motor failures [4]. In functional applications, rotating machinery bearing failures pose significant risks to safety and reliability, and can also lead to substantial production and equipment losses. While the early onset of bearing failures might not cause immediate catastrophic failure, the gradual degradation over time inevitably leads to major machinery breakdowns, making their maintenance more costly [5]. This underscores the critical importance of condition monitoring and fault diagnosis to ensure system safety and reliability [6]. In recent years, considerable effort has been dedicated to bearing health monitoring, with several approaches considered for preventing bearing failure. Consequently, monitoring health status and detecting faults in bearings have become a crucial area of industrial research. Various bearing health monitoring techniques have been developed, including current analysis, temperature monitoring, vibration analysis, and noise analysis [7, 8, 9, 10].

Bearing Fault Diagnosis Techniques Diagnosing bearing faults relies on utilizing various indicators and features to accurately and effectively analyze potential issues. These indicators include parameters such as vibration, speed, noise, acceleration, and temperature, with a particular focus on vibration signal processing and analysis, which has been extensively studied for fault detection [11]. Vibration signals are considered one of the most significant and valuable sources of accurate information for understanding phenomena associated with bearing problems [5, 12]. They provide precise information about the equipment's operational condition in real-time without interrupting the production line. Vibration monitoring is one of the most effective strategies because of its ability to diagnose and distinguish the locations and types of different defects from their early stages, before they worsen and lead to severe damage [13].

However, the effectiveness of vibration monitoring largely depends on the signal-processing approach employed. Different techniques have been developed for extracting fault detection attributes from raw data, including time-domain, frequency-domain, and time-frequency-domain analysis. Due to their relevance in various health monitoring applications, time-domain statistical analyses are frequently applied [14, 15]. These methods utilize various time-based indicators such as peak value, skewness [16], root mean square [13, 17], and kurtosis [17]. The benefits of time-domain analysis include effective and simplified computations. However, a primary deficiency of this method is its low sensitivity to early-stage faults and inability to diagnose deeply embedded defects.

Frequency analysis, or spectrum analysis, is one of the conventional methods for diagnosing faults in rotating machines. It allows the conversion of time signals into the frequency domain, which provides more comprehensive information about the machine's condition compared to time-domain analysis.

Techniques used include bearing fault frequency analysis [18], envelope spectrum [19], and Hilbert transform [20]. While these are accurate methods for determining the location and type of bearing faults, they depend on factors such as the size and speed of the bearing. However, noise can obscure valuable information, which necessitates careful frequency range selection. Time-frequency domain analysis represents a significant advancement due to its ability to provide information about both stationary and non-stationary signals [21]. Techniques in this domain include short-time Fourier transform (STFT) [22, 23], wavelet transforms [24, 25], and empirical mode decomposition [26]. In systems with variable speeds and loads, simple evaluation of monitoring indicators often does not provide reliable information about the machine's condition [5, 11].

Advancements in AI and Deep Learning for Fault Detection With the rapid advancement of AI technology, machine learning methods have been widely applied for detecting mechanical element faults [21]. Roller bearings exhibit unique vibration features that change based on the operating area, directly impacting machine reliability and stability. Consequently, researchers have proposed various intelligent methods based on Machine Learning (ML) and Artificial Intelligence (AI) models for efficiently detecting bearing faults. Pattern recognition uses several techniques for effectively grouping

signal patterns indicative of bearing faults. Examples of such techniques include convolutional neural networks (CNN), k-nearest neighbor classifiers (KNN) [27], support vector machines (SVM), Bayesian classifiers, and back-propagation neural networks (BPNN). Patil et al. [24] presented a fault conditioning and classification system to monitor the health of induction motor bearings using vibration signal analysis. Discrete wavelet transform (DWT), incorporating different wavelet families such as Sym5, DB4, and DB8, was initially used for signal decomposition. Various statistical features, including RMS, kurtosis, and skewness, were extracted from the third DWT decomposition level. These extracted features were then used as input for a 3-layer artificial neural network (ANN) model for classifying bearing faults. The improved system achieved a classification accuracy of 98.7% for inner race, outer race, and ball bearing faults. Other research has also offered KNN, Support Vector Machine (SVM), and Decision Tree (DT) mechanisms to detect various induction motor faults, including bearing defects [20]. These investigations relied on Hilbert–Huang Transform (HHT) for extracting fault-based features from signal analysis. Subsequently, the most significant attributes were selected using various dimensionality reduction mechanisms and presented to the employed ML strategies. The comprehensive inference from this study indicates that the feature-selection stage can remarkably improve fault detection system classification accuracy. Additionally, for localizing defects and identifying restrictions in roller bearings, some statistical vibration signal features were extracted from both time and frequency domains [28].

Currently, deep learning has garnered broad attention in the mechanical fault detection domain. An increasing number of researchers are relying on deep learning methods for achieving fault detection and machine life prediction under complicated working conditions [29]. Researchers in [30] propose an innovative system based on IoT and deep learning for fault diagnosis and a rectification scheme for induction motors, utilizing vibration signals. This system is notable for its ability to identify faults with high accuracy and detect fake data injection attacks. It has been successfully tested in a pilot environment, with results showing an accuracy of up to 99.84%, making it more effective compared to conventional methods [30]. This system allows for improved decision-making and identification of engine conditions with high reliability. This paper addresses the limitations of exploiting original

signals' temporal association features, the high cost of parameter setting, and the complexity of obtaining diverse training data in various operation situations. It achieves this by proposing a developed CNN-BiLSTM diagram for bearing fault detection**.** This model extracts accurate features from raw vibration signals by applying a Particle Swarm Optimization (PSO) mechanism for training parameter optimization. The pre-trained model can then be transferred to novel working situations with limited training instances, improving fault detection accuracy. Extensive tests demonstrated the presented model's efficiency in effectively addressing model training and fault detection issues in new working situations with data scarcity.

While previous research has made significant strides in bearing fault diagnosis, the work presented here contributes several novel aspects to the state-of-the-art. Firstly, many existing studies have focused on motors with a direct power supply, neglecting the widespread use of Variable Frequency Drives (VFDs) that introduce unique noise characteristics. Secondly, much of the literature has emphasized one-dimensional vibration analysis or employed complex computational models requiring substantial processing resources. This research addresses these gaps by specifically targeting fault diagnosis in VFD-driven motors using a three-dimensional vibration analysis approach combined with tailored wavelet denoising techniques**.** This approach effectively isolates characteristic fault frequencies below 500 Hz by removing higher-frequency VFD-induced noise components. Furthermore, the selective feature reduction methodology using information gain preserves diagnostic accuracy while significantly reducing computational requirements, making the solution more practical for real-time industrial monitoring applications.

This article remnant is regularized as: Section below proposes bearing fault features the Wavelet Transform and CNN deep learning mechanisms' summary developed. ''Experimental Setup'' part labels experimental rig, bearing dataset obtain technique in present study, vibration signal pre-processing for eliminating variable frequency drive (VFD) effect. Statistical features' extracting process in time-domain is discussed in ''Statistical Features'' section. ''Feature Selection'' part checks developed methods to choose statistical features. Data processing and Model training technique is defined in ''Training of Model'' part. ''Results and Discussion'' parts presents findings and discussion,

such as confusion matrix analysis, comparative assessment, classification results, feature significance rankings. This investigation contributions and summarizes are concluded in ‘‘Conclusion’’ part.

2. Methodology

2.1 Bearing Fault Characteristic

Bearing defects produce distinct vibration frequencies that vary depending on the affected bearing surface. Each type of defect has a unique frequency, which can be calculated based on the bearing's structure and dimensions. Figure 1 illustrates the geometric structure of a bearing. These characteristic frequencies include the inner race fault frequency (BPFI)**,** outer race fault frequency (BPFO)**,** cage fault frequency (FTF)**,** and ball spin frequency (BSF)**.** These frequencies can be predicted by recording the machine's vibration spectrum, as shown in equations (1)-(4) [31, 32, 33, 34]. Inner race faults, outer race faults, and cage faults are among the most common bearing defects. Localized defects cause periodic pulsations in the vibration signal, which are influenced by the shaft rotation speed and the bearing's location and dimensions, leading to periodic vibrations during engine operation

$$f_i = \frac{nf_r}{2}\left(1 + \frac{Db}{Dc}cos\beta\right) \quad (1)$$

$$f_o = \frac{nf_r}{2}\left(1 - \frac{Db}{Dc}cos\beta\right) \quad (2)$$

$$f_C = \frac{f_r}{2}\left(1 - \frac{Db}{Dc}cos\beta\right) \quad (3)$$

$$f_B = \frac{D}{2d}\left[1 - \left(\frac{d}{D}cos\beta\right)2\right] \quad (4)$$

Where: *fr* is rotor shaft frequency, *n* is the number of rolling elements, *Db* is ball diameter *Dc* is the pitch diameter, β is ball contact angle.

10.22060/EEJ.2025.24182.5658

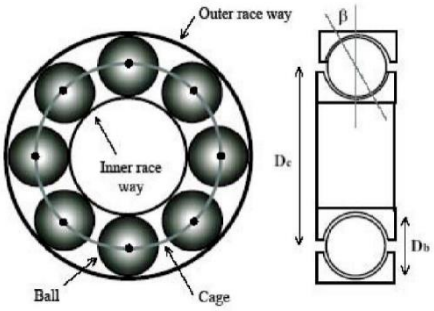

Fig. 1. Geometry structure of the bearing[35]

Such mechanical vibrations generate anomalies in the air gap flux density, which in turn cause **s**tator current modulation. These frequencies can also be computed as shown in equation (5) [31, 32, 33, 34]

$$f_{Bng=}|f_s \pm mf_v| \tag{5}$$

Where *fs* is the power supply frequency, *fv* is one of the characteristic vibration frequencies (*fi*, *fo*, *fc*), and *m*= 1, 2, 3, ...

2.2 Wavelet Transform

Signals from faulty components often exhibit non-stationary behavior. When the Fourier transform is used to analyze these signals, it provides an average frequency configuration over the entire signal period [6, 7**].** Time-frequency analysis techniques, such as wavelet transforms (WT)**,** are more suitable for analyzing these non-stationary signals [18]. The wavelet transform has been widely used in many applications to improve the accuracy of signal analysis compared to Fourier transforms [12]. In a wavelet series, the parent wavelet is modulated and superimposed on the signal. The inner product of the signal is then calculated with a series of child wavelets using scale and translation (offset) parameters. This transformation process is performed by shifting the wavelet along the x-axis to cover the entire signal, ensuring that the measured wavelet power density matches the original parent wavelet. Mathematically, this can be expressed as follows [10]:

$$CWf(a,b) = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{+\infty} f(t)\psi^* \left(\frac{t-b}{a}\right) dt \tag{6}$$

Where: a is the scale parameter, b is a translation parameter, ψ (t) is a function of zero mean.

While the Continuous Wavelet Transform (CWT) improves signal processing accuracy, it comes at a high cost in terms of energy, computation time, and memory [3], making it impractical in many cases.

10.22060/EEJ.2025.24182.5658

To overcome these limitations, the Discrete Wavelet Transform (DWT**)** was developed. In DWT, wavelets are measured and modified at discrete points within the signals. DWT is employed to decompose the original signal into multiple sub-signals, each possessing a specific bandwidth. A key advantage of DWT is its ability to evaluate data at different levels by utilizing filters with varying cutoff frequencies**. A** high-pass filter (HP) is used for analyzing high frequencies, while a low-pass filter (LP) is used for analyzing low frequencies [5, 23]. The use of Discrete Wavelet Transform (DWT) allows time-domain signals to be segmented and decomposed into various frequency bands with differing resolutions [12]. DWT employs wavelet and scaling functions associated with low-pass (LP) and high-pass (HP) filters. The original signal is passed through these filters at the first level, generating two signals with similar sampling lengths as the original signal. To maintain the number of samples, the number of samples is reduced by a factor of two, keeping one out of every two samples. Detail coefficients (cD1**)** represent the high-frequency information of the signal, while approximation coefficients (cA1) represent low-frequency information. Mathematically expressed as follows [9]:

$$cD_1[k] = \sum_n x[n] * h_n\varphi[2k - n] \tag{7}$$

$$cA_1(k) = \sum_n x[n] * g_n\psi[2k - n] \tag{8}$$

Where g[n] and h[n] represent the low-pass and high-pass filters, respectively. After achieving the first level of decomposition, the method can be used to further decompose cA1 into detail coefficients and an additional approximation, as illustrated in equations (9) and (10). This process is repeated until the desired decomposition level is reached.

$$cD_l[k] = \sum_n cD_{1-1}[n] * h_n\varphi[2k - n] \tag{9}$$

$$cA_l(k) = \sum_n cA_{1-1}[n] * g_n\psi[2k - n] \tag{10}$$

Where $cD_l[k]$ and $cA_l[k]$ are the DWT coefficients at level l, and cA_{l-1} is the approximate coefficient at level $l-1$. It is important to note that filtering and subsampling at each level reduce the number of instances by half (halving the temporal precision) and halve the frequency spectrum (doubling frequency accuracy). As a result of the repeated sampling by a factor of two, the total number of

instances in the processed signal must be a power of two. Figure 2 shows a graphical representation of the multilevel DWT process [19]

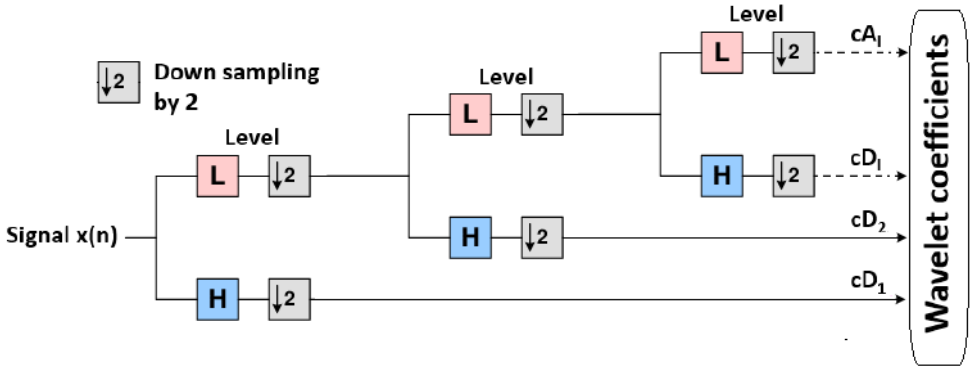

Fig. 2. Graphical representation of the multilevel DWT process

No comprehensive theoretical strategy for choosing the best wavelet family has been presented, leading investigators to use several families for decomposing similar wavelet signals [12, 13]. In many cases, the choice is made through trial and error [17]. In fact, when there is a high similarity between the applied sample signal and the parent wavelet, the wavelet function is considered appropriate to assess the signal in question [18]. Symlet and Daubechies families are recognized for their excellent performance in analyzing vibration signals and cover a broad range of wavelet orders. Therefore, the fourth-order Daubechies (db4) was employed in this study.

2.3. Convolution neural network

A Convolutional Neural Network (CNN) is a multi-layer architectural model [19] that typically includes a convolution layer, a pooling layer, a fully connected (FC) layer, and a classifier. The process begins with an image entering the network through the input layer. The convolution layer then extracts important local features using a convolution kernel. Subsequently, the pooling layer reduces feature dimensionality, maintains feature invariance, and helps prevent overfitting. In the fully connected layer, all 2D features are concatenated into 1D features, which then serve as input. Finally, a classifier (e.g., Softmax) is used in the output layer to obtain the classification results [10]. CNNs are particularly effective for analyzing and classifying grid-like data, such as images and videos [11]. The network relies on convolutional layers, which perform dot product operations between input data, filter weights,

10.22060/EEJ.2025.24182.5658

and local patches, allowing more abstract representations to be learned. Typically, a convolutional layer is followed by a pooling layer to reduce data dimensionality while preserving dominant attributes. The pooled feature maps are then transformed into a 1D vector using a flattening layer and subsequently fed into a fully connected block for classification. Standard CNNs handle 2D inputs. However, for processing sequential data, such as signal data from an induction motor, a 1DCNN framework has been developed. This framework utilizes similar standard CNN concepts but applies them to 1D input data, resulting in a 1D kernel and a 1DCNN output, as illustrated in Figure 3. The mathematical equation for a convolutional layer in a 1D convolutional neural network (1DCNN) can be formulated as shown in Equation (11) [15]:

$$x_j^l = f\left(\sum_{i=1}^{M} x_i^{l-1} * k_{ij}^l + b_j^l\right) \tag{11}$$

Where k_{ij}^l symbolizes the convolution kernels, j indicates the number of kernels, M denotes the total number of input channels, b_{ij} is the bias conforming to the j^{th} kernel, $f(\ )$ is the activation function and $*$ is the convolution factor.

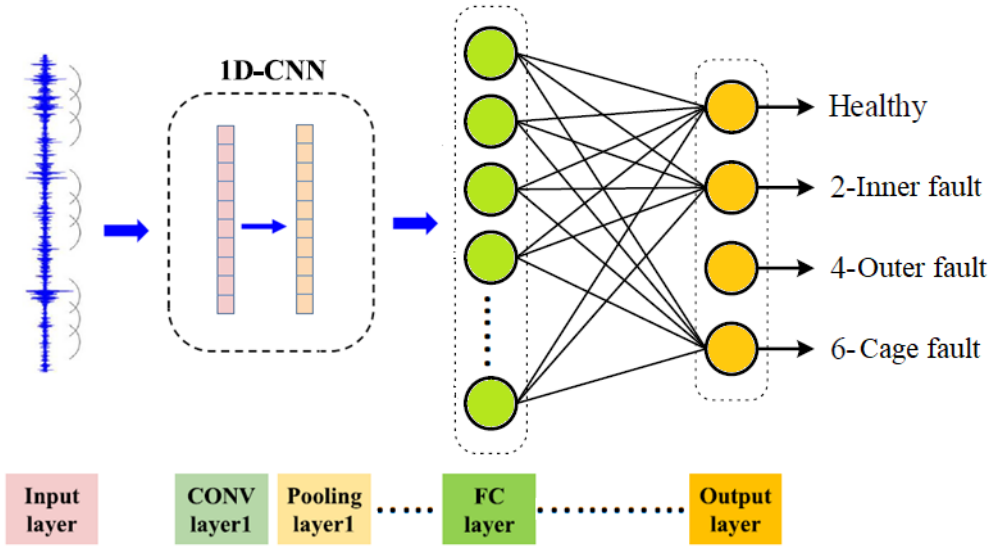

Fig. 3. One-dimensional CNN (1DCNN)

The network starts with a 32-filter convergence layer, followed by a max-pooling layer for dimensionality reduction, another 64-filter convergence layer, and finally a dense layer for final sorting. The loss value decreased steadily on both the training and validation sets as the number of epochs increased, confirming that the model learned effectively and did not suffer from overfitting. Especially

compared to deeper CNN models, our decision to use the current 1D-CNN architecture (as shown in Figure 3) was based on achieving the optimal balance between accuracy and computational efficiency. While deeper models may result in a slight increase in accuracy, they often require more computing resources and longer training time. Our model delivers robust performance with low complexity, making it a practical and effective solution for real-time equipment monitoring. The current model represents a well-thought-out trade-off between performance and complexity, a key consideration in machine learning system engineering.

3. Experiment

3.1. Rig description

The experimental setup used in this study is shown in Figure 4. This setup employs a 0.74-kW, 230-V, 3-phase, 4-pole induction motor. The motor is directly powered by a Variable Frequency Drive (VFD), specifically a Nflixin 9600D (3-phase, 50 Hz, 360V input/output). To apply a load, the rotor shaft is connected to a mechanical load, chosen to provide a consistent friction force. Throughout all tests, the machine is connected to the VFD. Bearing vibration signals are collected by an ADXL335 accelerometer, which features three sensors to measure vibration along the X, Y, and Z axes. This accelerometer is mounted at the front of the induction motor, as depicted in Figure 5. The accelerometer measures vibration in millivolts per gravity (mV/g). These signals using an IN-6009 data acquisition system (14-bit resolution, 48 kS/s). Table 1 provides the bearing parameters, which were obtained from the datasheet. This paper's experiments were conducted on four bearings: one in a healthy (undamaged) condition, and three with simulated faults. Inner race and outer race defects of 2 mm were simulated using the Electrical Discharge Machining (EDM) method.

Fig. 4 The rig experimental of use in this

Additionally, a 4-mm hole was drilled in one of the bearing cages, as illustrated in Figure 6, to simulate a cage fault.

Fig. 5. ADXL335 sensor

Fig. 6. Deformed bearings

Table 1. The bearing parameters

Type	Outside diameter	Inside diameter	NP	Dc	Db	Cos
6204	47mm	20mm	8	7.938 mm	33.5 mm	1

3.2. The bearing dataset capture method

Experiments were conducted under four load conditions: 300W, 400W, 500W, and 600W. To acquire the datasets, which consist of vibration signals captured by a bearing accelerometer at a sampling rate of 2.048 kHz, the analog output of the ADXL335 accelerometer is converted to acceleration in g by dividing it by 0.3, based on the sensor's sensitivity. To remove any constant offset and center the signal around zero, the arithmetic mean (AM) of the data is calculated and subtracted. This continuous calibration step, shown in Equation (12) [14], improves accuracy by highlighting actual acceleration changes such as vibrations or abnormal motion. A specialized program was developed in NI LabVIEW 2020 to process and extract time-domain features, as illustrated in Figure 7. Input analog voltage signals were integrated and then divided by the vibration sensor's sensitivity to convert them into g components. The signal was then separated into three output signals: X, Y, and Z.

$$AM = \frac{1}{N}\sum_{i}^{N} X_i \tag{12}$$

3.3. Pre-processing of vibration signal

While Variable Frequency Drives (VFDs) offer precise control over induction motors, they can introduce electrical phenomena that negatively affect bearing health and complicate vibration-based

10.22060/EEJ.2025.24182.5658

fault detection [54]. Implementing appropriate mitigation strategies is essential to ensure accurate monitoring and prolong the lifespan of motor bearings. To mitigate the VFD effect, the vibration signal was decomposed into 8 levels. It was then reconstructed after removing levels D1 and D2, which represent high frequencies that do not impact the characteristics of fault frequencies. This is because fault frequencies do not typically exceed 500 Hz, according to Equations (1-4).

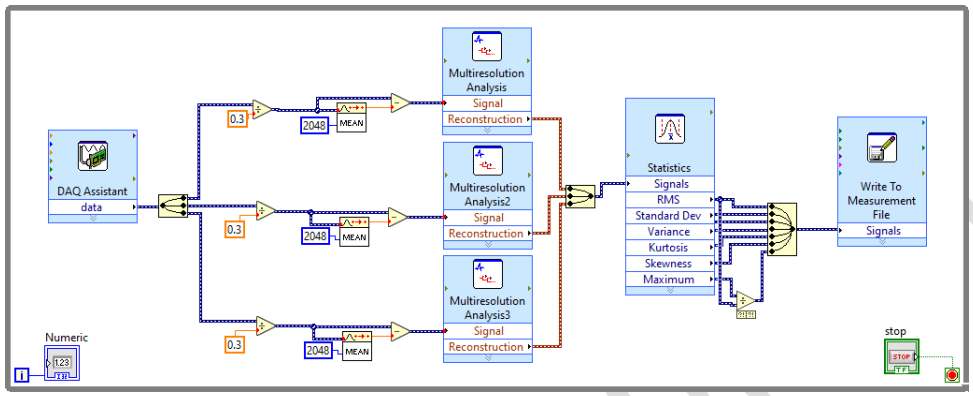

Fig. 7. Sample code from the overall block diagram to extract statistics adopting on time-domain

The table 2 illustrates the signal decomposition levels using Discrete Wavelet Transform (DWT), as shown in Figure 8. Figure 9 displays the X, Y, and Z vibration signals at 300 watts in a healthy state, both before and after processing. Figure 10, on the other hand, shows the Fast Fourier Transform (FFT) of the signal before and after the removal of higher frequencies. Each reconstructed signal will have the same number of instances as the primary input signal but will cover a distinct frequency range. This is achieved by resampling the detail (or approximation) coefficients by a factor of 2, as they were initially created by sampling coefficients by 2 and generating them using low-pass and high-pass filters. For instance, to reconstruct the first approximation wave signal (A1 level), only the approximation coefficients at that level are needed, while a vector of zeros is supplied in place of the detail coefficients. Figure 8 illustrates the concept of signal synthesis.

Table 2. Level of DWT frequencies

Level	1	2	3	4	5	6	7	8	9
Frequency Range (Hz)	4096-2048	2048 –1024	1024–512	512 –256	256 –124	124–64	64–32	32–16	16–0

10.22060/EEJ.2025.24182.5658

The primary purpose of using the Discrete Wavelet Transform (DWT) is not to change the amplitude of the main peaks associated with faults, but rather to filter out random high-frequency noise that is not related to the mechanical vibrations of the bearing. In an industrial environment, such as the one where the data was collected, other motors or power sources (like Variable Frequency Drives, VFDs) can generate electromagnetic or mechanical noise that contaminates the signal.

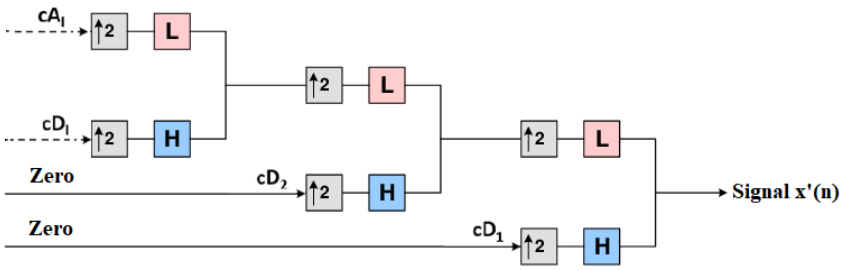

Fig. 8. The concept of signal synthesis

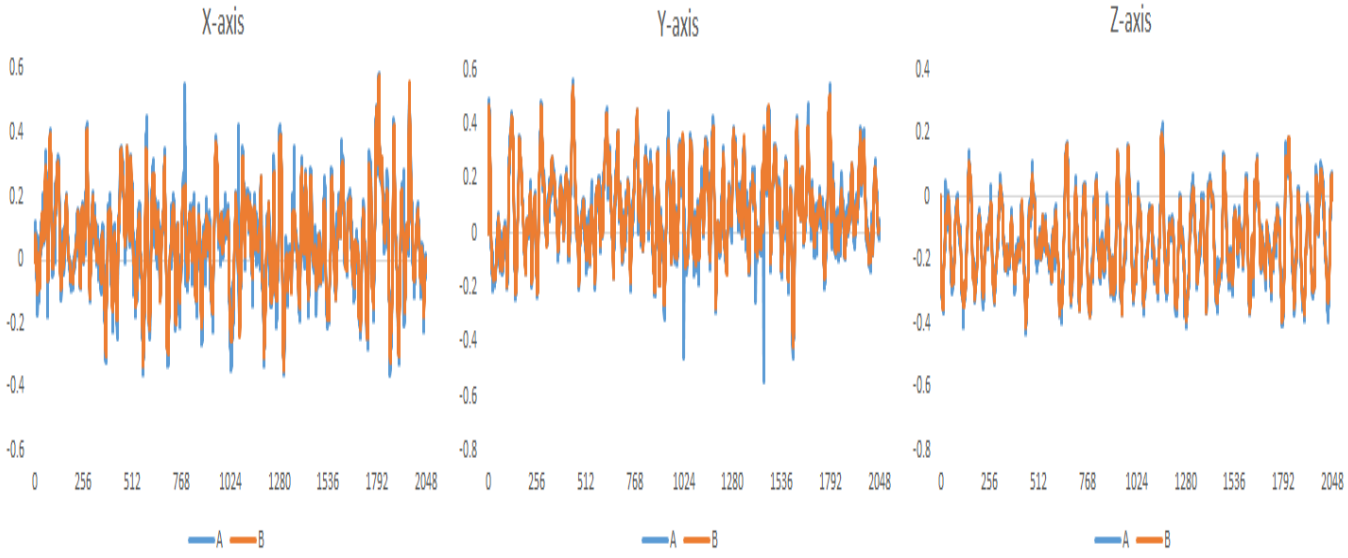

Fig. 9. The vibration signals blue color (A) before processing and orange color (B) after processing

The Figure 10 shows is a general reduction in the background noise level, especially in the high-frequency ranges. This reduction, even if not dramatic in these graphs, has a significant impact on the performance of the Convolutional Neural Network (CNN) model. By providing a "cleaner" signal to the model, we ensure that it learns from the true characteristics of vibrations caused by faults, instead of learning random noise features. This reduces the probability of overfitting and increases the model's

ability to generalize to new data. The signals shown in Figure 9 are the result of a pre-processing step to remove this DC offset, which allows for the isolation of the actual vibration signal centered around zero. We have clarified this procedure in the manuscript. We have also correctly added the axis units, with the horizontal axis representing Time (in seconds) and the vertical axis representing Acceleration

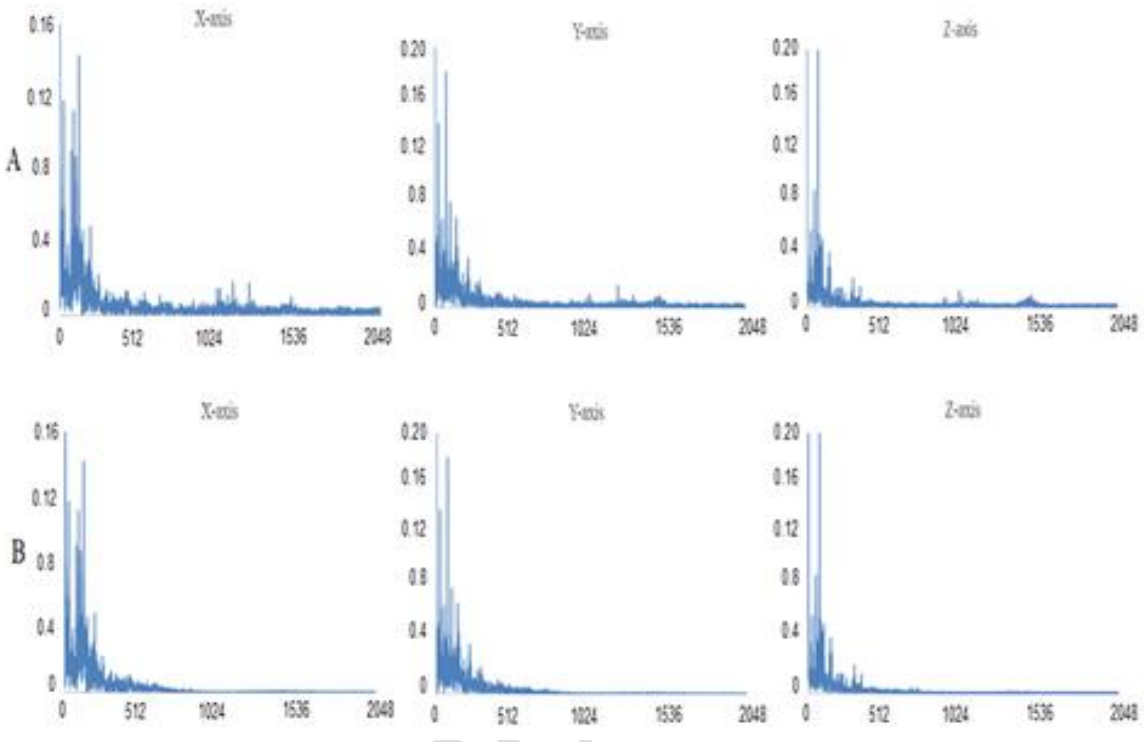

Fig. 10. The Fast Fourier Transform (FFT) of the signal, (A) before processing and (B) after processing

(ing). The discrepancy in scale between Figure 9 and Figure 10 was a result of the data visualization method.

4. Statistical Features

Time-domain analysis is commonly employed to monitor the state of an induction motor. Statistical features such as Crest Factor (Cr), Kurtosis (Kur), Root Mean Square (RMS), Skewness (Ske), Standard Deviation (Sta), and Variance (Var) are applied in this research's time domain to assess the health of the induction motor. These statistical attributes significantly impact pattern recognition capabilities and were chosen due to their proven efficacy in previous investigations [4], [9], [15]. The statistical feature equations below were derived and utilized to detect early faults in induction motors:

$$Cr = \frac{Max}{RMS} \tag{13}$$

10.22060/EEJ.2025.24182.5658

$$Kur = \frac{\frac{1}{N}\sum_{i=1}^{N}(X_i - AM)^4}{RMS} \quad (14)$$

$$RMS = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(X_i)^2} \quad (15)$$

$$Ske = \frac{1}{N-1}\sum_{i=1}^{N}(X_i - AM)^3 \quad (16)$$

$$Sta = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(X_i - AM)^3} \quad (17)$$

$$Var = \frac{1}{N}\sum_{i=1}^{N}(X_i - AM)^2 \quad (18)$$

This means there are six features for each output vibration signal axis utilized in the calculations.

5. Features Selection

Choosing the right features is crucial for training machine learning algorithms for several reasons. First, it significantly impacts performance by enabling the algorithm to better understand data patterns, leading to more accurate predictions and fewer errors [16]. Second, it reduces complexity by selecting only significant features, thereby lessening the data size and making the model less complex and more efficient. This, in turn, speeds up the training process and minimizes resource consumption. Third, it helps avoid overfitting, as unnecessary or redundant features can cause the model to over-learn the training data, reducing its ability to generalize to new data [17]. Fourth, it promotes interpretability; by employing a smaller set of significant features, illustrating the model's decisions becomes simpler, which is essential for applications requiring transparency [18]. Fifth, it reduces noise, as unnecessary features can obscure true patterns, and removing them improves model quality. Finally, it enhances computing performance, as smaller datasets demand fewer computing resources, making the training operation more efficient [9]. In summary, selecting good features is essential to improve the model's performance and ensure robust and efficient model training [6]. The presented method includes the use of a ranking technique identified as information gain (IG) for feature classification.

10.22060/EEJ.2025.24182.5658

Entropy, denoted as H, is a measure that quantifies the level of impurity in a given dataset. Building on this concept, one can define the information gain (IG) metric, which quantifies the supplementary information concerning class Y that is provided by attribute X [11, 12, 13]. This measure accurately quantifies the degree to which variable Y's entropy decreases when variable X is considered. The formula to calculate this measure is as follows [14]:

$$IG = H(Y) - H(Y|X) \_ H(X) - H(X|Y) \quad (19)$$

The proposed model's performance was then verified, and the number of componential features was reduced by 44.44%. Figure 11 depicts the results of the Entropy model. Ten features were selected based on these figures: Ske1, Sta2, Var2, RMS2, Crist1, Kur1, Ske2, Kur3, Ske3, and Sta1. The remaining features were removed from the dataset matrix due to lower gain. The suggested model's precision decreased from 99.5% to 98.64%, which is an insignificant percentage compared to the substantial reduction in the number of features, and it does not affect the overall performance of the model.

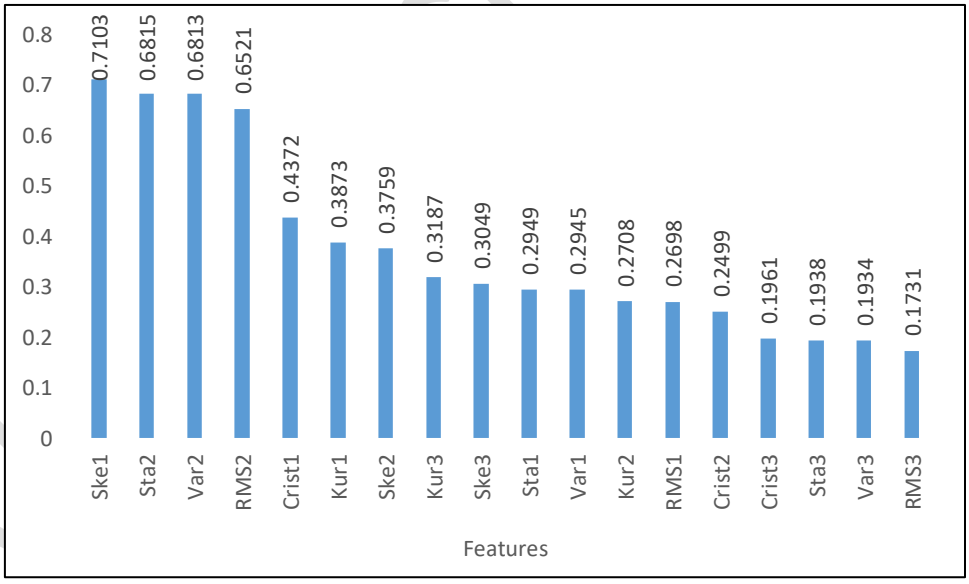

Fig. 11. The top 10 statistical features

6. Training of Model

All data points from three defect data integrations were used in the present test. A baseline set of 10 statistical attributes (Ske1, Sta2, Var2, RMS2, Crist1, Kur1, Ske2, Kur3, Ske3, Sta1) was included, with

10.22060/EEJ.2025.24182.5658

each attribute set comprising 3200 instances. For every healthy instance, race cage, outer, and inner faults each had 800 instances. Seventy percent of the statistical attributes were chosen as training instances, while the remaining 30% were used as test instances. To confirm that the dataset representations are effective in fault detection and classification, a Convolutional Neural Network (CNN) was employed for the four basic fault levels. The network generates predictions based on grouped faults. This includes a 1D convolutional layer followed by a pooling layer. The 1D convolutional layer has 32 filters, each with a size of 2. All such layers apply Rectified Linear Unit (ReLU) activation functions. The convolutional layer is followed by a Max-Pooling layer with a pooling size of 1. The outputs are then flattened and fully connected to a dense layer with a ReLU activation function and 80 neurons after the last layer. Finally, an output layer is generated using a Softmax activation function that possesses four neurons, representing the four fault levels. Model weights are initialized randomly, and the Adam Optimizer is applied for updating model parameters with a learning rate of 0.0005 and 80 epochs.

7. RESULTS AND DISCUSSION

After extracting and selecting the relevant statistical time-domain features, the dataset was divided using a 70:30 split ratio for training and testing. A total of 2,240 feature vectors were used for training, while 960 vectors were reserved for testing, representing the following class distributions: Healthy (218 samples), Inner race fault (244 samples), Outer race fault (237 samples), and Cage fault (261 samples).

Precision measures the proportion of positive cases correctly identified out of all cases predicted as positive. Recall measures the proportion of positive cases correctly identified out of all actual positive cases. The F1-score provides a balanced assessment of precision and recall by calculating their harmonic mean, effectively combining these two metrics into a single value, ranging from 0 (lowest) to 1 (highest) To evaluate the model's performance, several classification metrics were computed, including Accuracy, Precision, Recall, and the F1-score, following the definitions and procedures outlined in prior works [64]–[66]. Table 3 presents a summary of these evaluation metrics, while Figure 12 illustrates the overall accuracy achieved. Accuracy measures the percentage of correct predictions

10.22060/EEJ.2025.24182.5658

out of the total number of instances assessed, where higher percentages indicate better performance. Precision, on the other hand, quantifies the correct identification of positive patterns from all predicted positive patterns. Recall rates the proportion of positive patterns that are accurately classified. The F1-score offers a balanced assessment of Precision and Recall by computing their harmonic average, effectively combining these two metrics into a single value, achieving its lowest value at 0 and its highest value at 1. Precision and Recall values were above 0.97 for all classes. The F1-score was particularly strong: 1.00 for healthy, and 0.98–0.99 for all fault types. The macro and weighted averages for all metrics were both 0.99, indicating balanced performance regardless of class imbalance.

The figure 12 is a classification report that displays the precision, recall, f1-score, and support for each individual class (Healthy, Inner race, Outer race, Cage), along with the overall accuracy, macro average, and weighted average of these metrics. The "accuracy" value of 0.99 (or 99%) shown in the "accuracy" row of this specific Figure 12 refers to the overall accuracy of the model across all classes based on the aggregated data in this report. This 0.99 (99%) is essentially a rounded value of the more precise overall accuracy of 98.64% that you mention in your text and Table 3. Therefore, Figure 12 directly relates to the overall accuracy (98.64%) by providing the detailed breakdown of performance across individual classes and then summarizing it with an overall accuracy value (rounded to 0.99). The 98.64% is the precise numerical value that this figure conceptually represents as the "accuracy" of the entire model.

Table 3. Assessment measures of the used CNN learning model at feature selection

Indicator	Percent
Accuracy	**98.64**
Precision	99
Recall	99
F1-score	99

10.22060/EEJ.2025.24182.5658

	precision	recall	f1-score	support
Healthy	1.00	1.00	1.00	218
Inner race	0.99	0.98	0.98	244
Outer race	1.00	0.98	0.99	237
Cage	0.97	0.99	0.98	261
accuracy			0.99	960
macro avg	0.99	0.99	0.99	960
weighted avg	0.99	0.99	0.99	960

Fig. 12. The accuracy measures

The figure 13 illustrates the training and validation accuracy across 80 epochs. The model quickly converged after around 15 epochs and continued to improve steadily, reaching a training accuracy close to 100% and a validation accuracy of approximately 99%, indicating minimal overfitting. The corresponding loss curves further support this, with both training and validation loss sharply decreasing early and stabilizing around epoch 30.

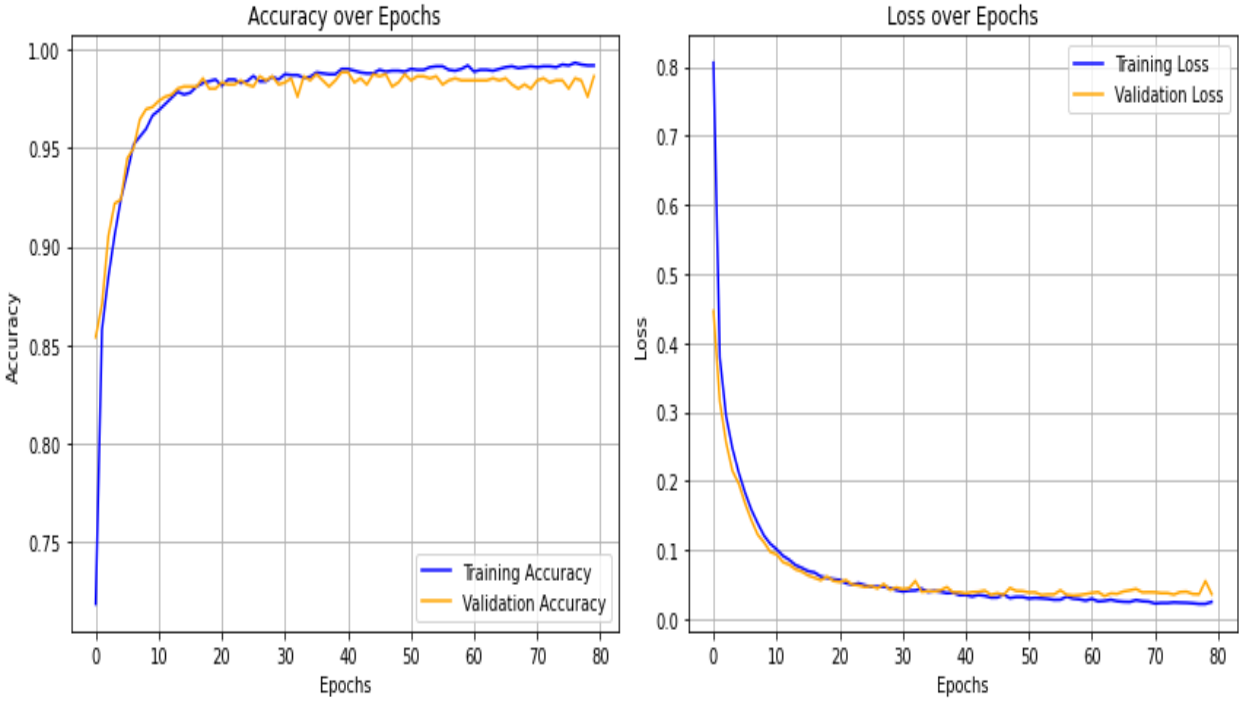

Fig. 13. Performance evaluation using accuracy and loss

Additionally, to better comprehend the correctly and incorrectly classified states, confusion matrices were presented in Figure 14. This confusion matrix for the CNN offers valuable insights into the correctly and incorrectly classified states. These numbers are used to track performance and determine

the need for model modifications to achieve higher accuracy and lower loss, contributing to the development of a more efficient and accurate model. Healthy samples were identified without a single error (218/218), while minor misclassifications were observed in the fault classes especially for the cage fault, where three samples were incorrectly labeled as inner race faults. The total model accuracy reached 98.64%

The confusion matrix (Figure 14) evaluates the CNN model's performance in classifying healthy and faulty conditions. The numbers along the diagonal (218, 238, 233, 258) represent correctly classified samples, while the off-diagonal numbers (1, 3, 3, 6) indicate errors. For instance, the model perfectly classified all 218 healthy samples but incorrectly predicted 6 inner race faults as cage faults.

Healthy	**218**	**0**	**0**	**0**
Inner race	**0**	**238**	**0**	**6**
Outer race	**1**	**0**	**233**	**3**
Cage	**0**	**3**	**0**	**258**

Fig. 14. Confusion matrix of CNN

To further validate the classifier, a multi-class ROC curve was analyzed, as shown in Figure 15. Each class achieved an AUC (Area Under the Curve) of 1.00, reflecting perfect separability between fault categories based on the learned features. These results collectively highlight the strength of the proposed method. The combination of preprocessing, temporal statistical feature extraction and selection, and a 1D CNN model proved highly effective. Notably, preprocessing helped mitigate the influence of the Variable Frequency Drive (VFD), enhancing signal clarity and enabling the network to detect subtle fault signatures with high confidence. The high classification performance particularly in distinguishing between similar faults such as inner and outer race failures makes this approach promising for real-world industrial applications, where early and accurate diagnosis is critical to prevent costly downtime. While the current setup performs strongly under controlled conditions, future work should explore more

10.22060/EEJ.2025.24182.5658

complex load scenarios, cross-device generalization, and real-time implementation, especially in noisy or variable-speed environments.

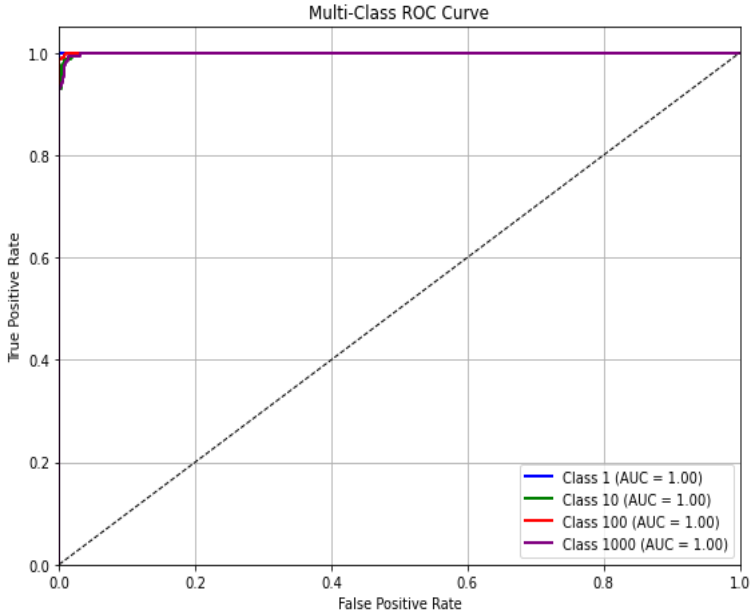

Fig. 15. Muti-Class AUC

An AUC (Area Under the Curve) of 1.00 signifies perfect separability. In the context of this application, it means the model can perfectly distinguish between the different classes (Healthy, Inner race, Outer race, and Cage faults). The classifier is able to set a threshold such that it correctly identifies all positive cases and all negative cases without any overlap. This result reflects a flawless ability to separate the fault categories based on the features the model learned from the data.

7.1. Comparative Analysis

Table 4 shows a comparative abstract of the best models achieved from the present article, along with a list of recent models. Some strategies were considered during this research regarding classifier intricacy and computational effort. Rule-based classifiers, like PART classifiers developed by Grover and Turk [16], show adequate computational efficiency and accuracy using fewer features, while Hjorth parameters effectively diagnose bearing faults. Neural Networks, especially those based on architectures like Back-Propagation Neural Networks (BPNN) and Convolutional Neural Networks (CNN), such as those described in Huang et al. [15], are adept at modeling non-linear relationships in data and handling complex patterns. However, their training process may require substantial

computational resources due to their numerous connections and multiple layers. The use of Support Vector Machine (SVM) and statistical methods such as one-way ANOVA and the Kruskal-Wallis test, as reported in the study by Cascales-Fulgencio et al. [16], is an effective way to classify the importance of features in high-dimensional data. However, their computational intricacy may rise as the number of features increases. The kNN mechanism, an axiomatic and simple mechanism applied by Jamil et al. [67], offers low computational difficulty in the training step. Although, large datasets might require more computational resources during the prediction process. Naive Bayes classifiers, including the kernel Naive Bayes applied by Alonso-Gonzalez et al. [18], are known for their speed and simplicity, making them computationally effective when coping with high-resolution data. As Rajput et al. [19] introduced, a Fuzzy Convolutional Neural Network is an effective model for complex pattern recognition. However, its multi-layer architecture and convolution operations require high computational complexity and longer training time to achieve high performance. The KNN classifier and the FCBF feature selection method applied by Jaber, Alaa. Abdulhady [11] demonstrate a balance between classification performance and computational efficiency. The simple CNN model used here, which contains 10 single time-domain statistical features with 3 dimensions, was applied for appraising CNN mechanisms' performance in bearing fault diagnostics. As presented in Table 3, the CNN model obtained an accuracy of 98.58%, which ranks near the Fuzzy-CNN model at 16 attributes that achieved the most proper performance at 99.87%. Although the model used in this study needs low computational power and less complexity, it provided a very high degree of accuracy, thus indicating its high efficiency in fault diagnosis. Our model delivers an excellent accuracy of 98.58%, which is sufficient and even very suitable for many real-world scenarios. The 1.29% accuracy gap is acceptable for applications such as product or content recommendation systems, where an error may lead to a suboptimal recommendation but not a catastrophe. In non-critical image classification or preliminary data analysis, the focus is often on processing speed and the ability to handle large amounts of data at a reasonable cost, and this is where our model clearly excels. The cost-benefit balance becomes the cornerstone of decision-making. If a less accurate model delivers the same practical value at a much lower cost, it is often the smarter choice.

10.22060/EEJ.2025.24182.5658

To ensure the reliability and reproducibility of the results, we applied k-fold cross-validation to the dataset. The data was divided into five equal subsets, and the model was trained and evaluated five times, with one subset used as the test data each time.

The final results in the Performance section have been updated to include the average accuracy and standard deviation across all cross-validation rounds. For example, the results now represent 98.64%±0.25%. This reporting of variance significantly strengthens the credibility of our claims about the model's performance, showing that the model achieves high and stable accuracy, and that the variance in performance across validation rounds was low, confirming that the results are not accidental but rather reflect the model's true performance.

Table 4. Models ML performance comparison to bearing fault diagnosis

Author	Classifier	Feature Type	Number of features	Accuracy (%)
Grover and Turk [26]	Rule-based	Time-domain	3	93.82
Huang, et al. [36]	BPNN	Time-domain features	4	91.6
Cascales, et al. [37]	SVM	Envelope Spectrum features	16	84.7
Jamil, et al. [38]	KNN	Time & Frequency-domain	9	96.2
Alonso, et al. [39]	Kernel Naive Bayes	Time-domain features	5	94.4
Rajput, et al. [40]	Fuzzy-CNN	Raw vibration signals	16	99.87
J.Abdulhady [11]	FCBF-kNN	Time-domain features	7	97
Current work	CNN	Time-domain features	10	98. 64

While Rajput et al. [69] achieved a slightly higher accuracy (99.87%), their work did not address the specific challenge of VFD noise. Our methodology features a crucial preprocessing step using the Discrete Wavelet Transform (DWT), which specifically targets this noise, ensuring that the data fed to the CNN is clean and truly representative of the bearing's actual vibrations. Furthermore, our model achieves a very high accuracy of 98.64% using only 10 features, which indicates high efficiency in computational resource usage. Secondly, when compared to traditional methods using classifiers like SVM, kNN, and BPNN, our approach demonstrated a clear superiority in accuracy (98.64% compared to 96.2%, 94.4%, and others). This superiority is due to the CNN's ability to automatically learn

complex geometric features from the signals, rather than relying solely on traditional statistical features. Therefore, the true innovation of this work lies in the synergistic integration of three key steps that have not been combined with the same effectiveness in the referenced works: dedicated preprocessing using DWT to remove VFD noise, efficient feature selection based on only 10 statistical features, and robust classification using a CNN that performs deep feature learning. This integration highlights the practical relevance of our methodology as a comprehensive and effective solution for diagnosing bearing faults in real-world environments challenged by noise.

8. CONCLUSION

This paper highlights the successful performance of a Convolutional Neural Network (CNN) for accurate classification of various bearing faults using time-domain information. We achieved this by employing a three-axis vibration sensor. The captured signals were then pre-processed using Discrete Wavelet Transform (DWT)**,** specifically by removing levels D1 and D2. These levels represent high frequencies that do not affect the characteristics of fault frequencies, as fault frequencies typically do not exceed 500 Hz. This step was crucial for mitigating the effect of the Variable Frequency Drive (VFD). Next, statistical features were extracted from the three axes, and an optimal selection of these features was made using the information gain (Entropy) method**.** These selected features then served as the input to the CNN. For training the CNN model, a dataset of 3200 samples was used. Each of the healthy, inner race fault, outer race fault, and cage fault conditions contributed 800 samples. Seventy percent of the statistical features were randomly selected as training instances, while the remaining 30% were used as test instances. The study achieved an impressive accuracy of 98.64% in fault classification. This paper underscores the capability of time-signal representation as a feature transformation technique for simplifying bearing fault detection in this domain and paves the way for future studies in this area. Ultimately, the current study presents a satisfactory solution for real-life applications in bearing fault detection.

9. Reference

[1] P. Gangsar and R. Tiwari, "Multifault Diagnosis of Induction Motor at Intermediate Operating Conditions Using Wavelet Packet Transform and Support Vector Machine," *Journal of Dynamic Systems, Measurement and Control, Transactions of the ASME*, vol. 140, no. 8, Aug. 2018, doi: 10.1115/1.4039204.

[2] A. Boudiaf, A. Moussaoui, A. Dahane, and I. Atoui, "A Comparative Study of Various Methods of Bearing Faults Diagnosis Using the Case Western Reserve University Data," *Journal of Failure Analysis and Prevention*, vol. 16, no. 2, pp. 271–284, Apr. 2016, doi: 10.1007/s11668-016-0080-7.

[3] J. Liu, L. Xue, L. Wang, Z. Shi, and M. Xia, "A new impact model for vibration features of a defective ball bearing," *ISA Trans*, vol. 142, pp. 465–477, Nov. 2023, doi: 10.1016/j.isatra.2023.08.014.

[4] T. Ali, D. #1, A. Abdulhady, and J. #2, "Bearing Fault Diagnosis Using Motor Current Signature Analysis and the Artificial Neural Network," vol. 10, no. 1, 2020.

[5] C. Abdelkrim, M. S. Meridjet, N. Boutasseta, and L. Boulanouar, "Detection and classification of bearing faults in industrial geared motors using temporal features and adaptive neuro-fuzzy inference system," *Heliyon*, vol. 5, no. 8, Aug. 2019, doi: 10.1016/j.heliyon.2019.e02046.

[6] C. Wang, M. Wang, B. Yang, K. Song, Y. Zhang, and L. Liu, "A novel methodology for fault size estimation of ball bearings using stator current signal," *Measurement (Lond)*, vol. 171, Feb. 2021, doi: 10.1016/j.measurement.2020.108723.

[7] P. Guo, J. Fu, and X. Yang, "Condition monitoring and fault diagnosis ofwind turbines gearbox bearing temperature based on kolmogorov-smirnov test and convolutional neural network model," *Energies (Basel)*, vol. 11, no. 9, Sep. 2018, doi: 10.3390/en11092248.

[8] H. Zhao, H. Liu, Y. Jin, X. Dang, and W. Deng, "Feature Extraction for Data-Driven Remaining Useful Life Prediction of Rolling Bearings," *IEEE Trans Instrum Meas*, vol. 70, 2021, doi: 10.1109/TIM.2021.3059500.

[9] M. Irfan *et al.*, "A Comparison of Machine Learning Methods for the Diagnosis of Motor Faults Using Automated Spectral Feature Extraction Technique," *J Nondestr Eval*, vol. 41, no. 2, p. 31, Jun. 2022, doi: 10.1007/s10921-022-00856-3.

[10] H. Shi, Y. Li, X. Bai, K. Zhang, and X. Sun, "A two-stage sound-vibration signal fusion method for weak fault detection in rolling bearing systems," *Mech Syst Signal Process*, vol. 172, p. 109012, Jun. 2022, doi: 10.1016/j.ymssp.2022.109012.

[11] A. A. Jaber, "Diagnosis of Bearing Faults Using Temporal Vibration Signals: A Comparative Study of Machine Learning Models with Feature Selection Techniques," *Journal of Failure Analysis and Prevention*, vol. 24, no. 2, pp. 752–768, Apr. 2024, doi: 10.1007/s11668-024-01883-0.

[12] R. Liu, B. Yang, E. Zio, and X. Chen, "Artificial intelligence for fault diagnosis of rotating machinery: A review," *Mech Syst Signal Process*, vol. 108, pp. 33–47, Aug. 2018, doi: 10.1016/j.ymssp.2018.02.016.

[13] V. V. Rao and C. Ratnam, "Estimation of Defect Severity in Rolling Element Bearings using Vibration Signals with Artificial Neural Network," 2015.

[14] L. A. Al-Haddad and A. A. Jaber, "An Intelligent Fault Diagnosis Approach for Multirotor UAVs Based on Deep Neural Network of Multi-Resolution Transform Features," *Drones*, vol. 7, no. 2, Feb. 2023, doi: 10.3390/drones7020082..

[15] T. T. Vo, M. K. Liu, and M. Q. Tran, "Harnessing attention mechanisms in a comprehensive deep learning approach for induction motor fault diagnosis using raw electrical signals," *Eng Appl Artif Intell*, vol. 129, Mar. 2024, doi: 10.1016/j.engappai.2023.107643.

[16] S. Agrawal and V. K. Giri, "Improved mechanical fault identification of an induction motor using Teager-Kaiser energy operator," *Journal of Electrical Engineering and Technology*, vol. 12, no. 5, pp. 1955–1962, Sep. 2017, doi: 10.5370/JEET.2017.12.5.1955.

[17] M. Savolainen and A. Lehtovaara, "Development of damage detection parameters over the lifetime of a rolling element bearing," *Tribologia - Finnish Journal of Tribology*, vol. 38, no. 3–4, Dec. 2021, doi: 10.30678/fjt.112092.

[18] H. Shi, Z. Liu, X. Bai, Y. Li, and Y. Wu, "A theoretical model with the effect of cracks in the local spalling of full ceramic ball bearings," *Applied Sciences (Switzerland)*, vol. 9, no. 19, Oct. 2019, doi: 10.3390/app9194142.

[19] P. Wang *et al.*, "Vibration characteristics of rotor-bearing system with angular misalignment and cage fracture: Simulation and experiment," *Mech Syst Signal Process*, vol. 182, p. 109545, Jan. 2023, doi: 10.1016/j.ymssp.2022.109545.

[20] C.-Y. Lee and W.-C. Lin, "Induction Motor Fault Classification Based on ROC Curve and t-SNE," *IEEE Access*, vol. 9, pp. 56330–56343, 2021, doi: 10.1109/ACCESS.2021.3072646.

[21] X. Liu, H. Huang, and J. Xiang, "A personalized diagnosis method to detect faults in a bearing based on acceleration sensors and an fem simulation driving support vector machine," *Sensors (Switzerland)*, vol. 20, no. 2, Jan. 2020, doi: 10.3390/s20020420.

[22] H. Liu, L. Li, and J. Ma, "Rolling Bearing Fault Diagnosis Based on STFT-Deep Learning and Sound Signals," *Shock and Vibration*, vol. 2016, 2016, doi: 10.1155/2016/6127479.

[23] C. T. Alexakos, Y. L. Karnavas, M. Drakaki, and I. A. Tziafettas, "A Combined Short Time Fourier Transform and Image Classification Transformer Model for Rolling Element Bearings Fault Diagnosis in Electric Motors," *Mach Learn Knowl Extr*, vol. 3, no. 1, pp. 228–242, Mar. 2021, doi: 10.3390/make3010011.

[24] A. B. Patil, J. A. Gaikwad, and J. V. Kulkarni, "Bearing fault diagnosis using discrete Wavelet Transform and Artificial Neural Network," in *2016 2nd International Conference on Applied and Theoretical Computing and Communication Technology (iCATccT)*, IEEE, 2016, pp. 399–405. doi: 10.1109/ICATCCT.2016.7912031.

[25] F. He and Q. Ye, "A Bearing Fault Diagnosis Method Based on Wavelet Packet Transform and Convolutional Neural Network Optimized by Simulated Annealing Algorithm," *Sensors*, vol. 22, no. 4, Feb. 2022, doi: 10.3390/s22041410.

[26] C. Grover and N. Turk, "Rolling Element Bearing Fault Diagnosis using Empirical Mode Decomposition and Hjorth Parameters," in *Procedia Computer Science*, Elsevier B.V., 2020, pp. 1484–1494. doi: 10.1016/j.procs.2020.03.359.

[27] Y. Hou *et al.*, "Acoustic feature enhancement in rolling bearing fault diagnosis using sparsity-oriented multipoint optimal minimum entropy deconvolution adjusted method," *Applied Acoustics*, vol. 201, p. 109105, Dec. 2022, doi: 10.1016/j.apacoust.2022.109105.

[28] M. Altaf, T. Akram, M. A. Khan, M. Iqbal, M. M. I. Ch, and C. H. Hsu, "A New Statistical Features Based Approach for Bearing Fault Diagnosis Using Vibration Signals," *Sensors*, vol. 22, no. 5, Mar. 2022, doi: 10.3390/s22052012.

[29] S. Tang, S. Yuan, and Y. Zhu, "Deep learning-based intelligent fault diagnosis methods toward rotating machinery," *IEEE Access*, vol. 8, pp. 9335–9346, 2020, doi: 10.1109/ACCESS.2019.2963092.

[30] B. Song, Y. Liu, J. Fang, W. Liu, M. Zhong, and X. Liu, "An optimized CNN-BiLSTM network for bearing fault diagnosis under multiple working conditions with limited training samples," *Neurocomputing*, vol. 574, Mar. 2024, doi: 10.1016/j.neucom.2024.127284.

[31] J. Zarei and J. Poshtan, "Bearing fault detection using wavelet packet transform of induction motor stator current," *Tribol Int*, vol. 40, no. 5, pp. 763–769, May 2007, doi: 10.1016/j.triboint.2006.07.002.

[32] A. Ibrahim, M. El Badaoui, F. Guillet, and F. Bonnardot, "A new bearing fault detection method in induction machines based on instantaneous power factor," *IEEE Transactions on Industrial Electronics*, vol. 55, no. 12, pp. 4252–4259, 2008, doi: 10.1109/TIE.2008.2003211.

[33] *Proceedings, IECON 2012 : 38th Annual Conference of the IEEE Industrial Electronics Society : École de Technologie Supérieure de Montréal, Université du Québec, Montreal, Canada, 25-28 October, 2012*. IEEE, 2012.

[34] G. Niu, X. Dong, and Y. Chen, "Motor Fault Diagnostics Based on Current Signatures: A Review," 2023, *Institute of Electrical and Electronics Engineers Inc.* doi: 10.1109/TIM.2023.3285999.

[35] A. Ibrahim', M. El Badaoui', F. Guillet', and M. Zoaeter2, "Using the Cyclostationarity of Electrical Signal for Bearing Fault Detection in Induction Machine."

[36] M. Huang, Z. Liu, and Y. Tao, "Mechanical fault diagnosis and prediction in IoT based on multi-source sensing data fusion," *Simul Model Pract Theory*, vol. 102, Jul. 2020, doi: 10.1016/j.simpat.2019.101981.

[37] D. Cascales-Fulgencio, E. Quiles-Cucarella, and E. García-Moreno, "Computation and Statistical Analysis of Bearings' Time- and Frequency-Domain Features Enhanced Using Cepstrum Pre-Whitening: A ML- and DL-Based Classification," *Applied Sciences (Switzerland)*, vol. 12, no. 21, Nov. 2022, doi: 10.3390/app122110882.

[38] M. A. Jamil, M. A. A. Khan, and S. Khanam, "Feature-based performance of SVM and KNN classifiers for diagnosis of rolling element bearing faults," in *Vibroengineering Procedia*, EXTRICA, Dec. 2021, pp. 36–42. doi: 10.21595/vp.2021.22307.

[39] M. Alonso-Gonzalez, V. G. Diaz, B. Lopez Perez, B. Cristina Pelayo G-Bustelo, and J. P. Anzola, "Bearing Fault Diagnosis With Envelope Analysis and Machine Learning Approaches Using CWRU Dataset," *IEEE Access*, vol. 11, pp. 57796–57805, 2023, doi: 10.1109/ACCESS.2023.3283466.

[40] G. S., M. G.E., and S. M., "Fault Prediction Using Fuzzy Convolution Neural Network on IOT Environment with Heterogeneous Sensing Data Fusion," *Bonfring International Journal of Networking Technologies and Applications*, vol. 11, no. 1, pp. 01–05, Jan. 2024, doi: 10.9756/bijnta/v11i1/bij24001.

AUT Journal of Electrical Engineering
10.22060/EEJ.2025.24182.5658