Feature Optimization for Machine Learning Based Bearing Fault Classification

Mohammad Mohiuddin[1], Md Saiful Islam[2],Jia Uddin[3]
[1,2]Department of Electronics & Telecommunication Engineering, Chittagong University of Engineering & Technology (CUET). Chittagong, Bangladesh
[3]AI and BigData Department, Endicott College, Woosong University, Daejeon, South Korea

Article Info

Article history:

Received Jun 29, 2024
Revised Aug 31, 2024
Accepted Sep 11, 2024

Keyword:

Bearing fault
Vibration signals
Wavelet Transform
Shannon Entropy Criterion
Artificial Neural Networks

ABSTRACT

The most critical and essential parts of rotating machinery are bearings. The main problem of the bearing fault classification is to select the fault features effectively because all extracted features are not useful, and the high-dimensional features give poor performances and slow down the training process. Due to the effective feature selection problem, the bearing fault diagnosis method does not achieve a satisfactory result. The main goal of this paper is to extract the effective fault features with an optimization technique to classify the bearing faults using machine learning algorithms. Since wavelet entropy can determine complexity and degree of order of a vibration signal, this research uses it in features optimization. The proposed wavelet entropy-based optimization technique reduces the dimensionality of input, elapsed time and raises the learning process. Four Machine learning algorithms (naïve Bayes, support vector machine, artificial neural network and KNN) are applied to classify the bearing faults using the optimized features. To evaluate the proposed method, Case Western Reserve University's (CWRU's) bearing dataset is used which consists of three types of bearing faults. Based on the experimental data, it was shown that the proposed system reached 99.5% accuracy. The accuracy and robustness of the bearing fault classification are tested by adding noise to the vibration raw signals at various levels of Signal-to-Noise Ratio (SNR). Experimental results show that the proposed method is very highly reliable in detecting bearing faults compared to the conventional methods.

Corresponding Author:

Jia Uddin
AI and BigData Department, Endicott College, Woosong University
Daejeon, South Korea
Email: jia.uddin@wsu.ac.kr

1. INTRODUCTION

Bearings are critical elements of rotating machinery. Bearing conditions have a significant impact on machines. Vibration analysis has been widely employed for bearing condition monitoring for many decades. Vibration signals caused by bearing problems have been intensively studied, and several diagnostic approaches have been used in the past [1]. Bearing faults diagnostics may be approached in three ways: predictive maintenance, reactive maintenance, and preventive maintenance. Real-time monitoring and diagnostics of bearings are the foundation of predictive maintenance. Reactive maintenance relies on repair activities conducted after a bearing failure has already been identified. Preventive maintenance relies on time bound approaches and best practices for planning and scheduling repair actions when the actual status of the bearing is unknown [2].

Artificial Neural Network (ANN) has been used to detect and diagnosis machine conditions, which have been considered classification problems according to learning patterns [3]. An issue with classifying machine faults is divided into two sections. The first section deals with feature extraction from vibration signals, which is applied to extract some features exhibiting fault information. The second section is classification, which employs various artificial intelligence approaches to diagnose faults using the extracted features [4].

The features of vibration can be obtained using time-domain analysis, frequency-domain analysis, and combining time-frequency analysis [5], [6]. The time-domain features [3], [7], [8] are root means square (RMS), standard deviation, crest-factor, kurtosis, variance, envelop spectrum, estimation and crest value. Feature extraction via time-frequency analysis is commonly employed for nonlinear and non-stationary signals such as Empirical mode decomposition (EMD), wavelet packet transform, and short-time Fourier transform, which has demonstrated its high analytical capabilities for these signal types [9].

It is exceedingly challenging to extract the features of bearing signals using the typical feature extraction process [10] because of the complicated vibration signals. The timefrequency domain analytic technique known as the wavelet transform has recently become quite effective for the fault classification of bearings [11]. As an advanced signal processing technique, Saida Dahmane et al. [12] proposed the discrete wavelet transform (DWT) with the envelope spectrum (ENV), combining a machine learning approach based on random forest classifier. By combining these methods, the dataset may be better organized and features can be extracted, which increases the random forest classifier's accuracy and raises its classification rate. The weakness of this method is high-dimensional feature set which increases training time, and slows down the learning process of the classifier.

In order to extract both temporal and spatial characteristics enhanced by Hilbert transform 2D images, a hybrid DTL architecture consisting of lengthy short-term memory layers and a deep convolutional neural network is described. In a variety of settings, the suggested model performed more accurately than the most advanced models [13]. But it requires 2D images, where the vibration signal is 1D data.

A compact fault diagnostic model is presented by Zabin et al. [14]. It combines a hybrid texture representation method using gammatone spectrogram (GS) filter and empirical mode decomposition (EMD) with a self-attention SqueezeNet architecture. Based on the experimental data, it was shown that the SqueezeNet self-attention mechanism reached 97% accuracy. Its accuracy is not satisfactory as compared tp others so the accuracy needs to be improved. Purushotham et al. [15] and Prabhakar et al. [16] employed Discrete Wavelet Transforms (DWT) to identify bearing faults. Saravanan et al. [17] highlight the use of wavelet-based features for gear fault diagnostics using support vector machine (SVM) and Proximal Support Vector Machine (PSVM). It performs better than the Fourier analysis for processing non-stationary vibration signals [4]. Using a high-dimensional feature set reduces performance, increases training time, and slows down the learning process of the classifier [18]. Because of the large dimensionality of the feature set, the dimensionality of the feature set must be lowered following feature extraction because not all extracted features are equally meaningful [19], [20].

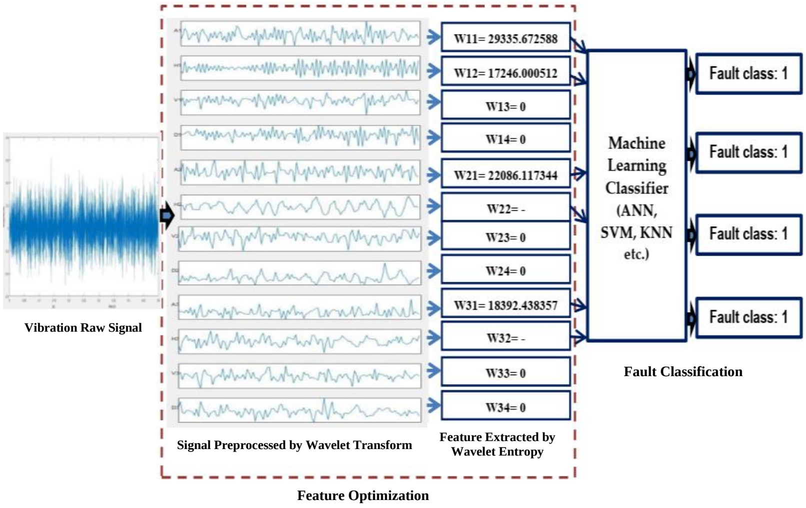
In order to solve this issue, this work focuses on an effective feature selection-based optimization technique. First, The vibrational signal is decomposed using the discrete wavelet transforms (DWT) utilizing the Haar wavelet to yield both approximations and details. To ensure that no information is lost, every detail is then recreated. The minimum Shannon entropy criterion is then used to the reconstructed details that were collected to identify the fault frequencies, which are regarded as novel features and then using an optimization process. This combination of approaches allows effective feature extraction and structuring. These optimized features are fed into a feed-forward artificial neural network to train and test the model to classify the bearing fault diagnosis. To analyze the proposed system, the Case Western Reserve University (CWRU) dataset is used, which consists of three types of faulty bearing data. After successfully training the ANN, it will be ready to classify samples for the fault class of bearing diagnosis. Finally, it achieves the best accuracy, which shows that the proposed system is quite reliable in identifying bearing faults.

The main contributions of this research are as follows:

- An effective preprocessing technique is developed using wavelet transform and entropy to reduce input dimensionality and elapsed time and raise the learning process.
- Wavelet entropy is introduced in this research to determine the significant fault signatures, i.e., optimized features. The impacts of the selected optimized features are investigated.
- The proposed model is verified against four machine learning algorithms: naïve Bayes, support vector machine, artificial neural network, and KNN.
- Furthermore, perform a comparative analysis of the proposed system to study the improvements in efficiency and performance.

2. PROPOSED FAULT CLASSIFICATION WITH OPTIMIZATION TECHNIQUE

Due to the problem with the sizeable dimensional feature set, the dimensionality of the feature set needs to be reduced after the feature extraction because all extracted features are not equally valuable to detect the bearing faults. This proposed method focuses on detecting and classifying of bearing faults using optimization technique and machine learning algorithms. This proposed bearing fault detection method consists of two key steps: feature optimization and classification. The flowchart for the proposed system is shown in Figure 1.



2.1. Dataset preparation

The Case Western Reserve University (CWRU) test bench is used to collect vibration signals (acceleration) to evaluate the effectiveness of the proposed system [21]. One of the most extensively used and freely available datasets is the CWRU dataset. It provides data on both normal and faulty bearings. There are mainly three types of bearing data in faulty data: inner race fault, ball fault, and outer race fault. Figure 3 shows the schematic diagram of the rotor-bearing system, which includes a dynamometer, an encoder and torque transducer, and an electric motor. There is a single-point fault in each faulty bearing to test bearings, which have sizes of 0.007 inch, 0.014 inch, and 0.021 inch. Accelerometers are used to collect data on vibration signals. 12 kHz and 48 kHz were chosen as the sample rates for the data collection. MATLAB (.mat) format is used to store all of the data files. Table 1 shows the information of the bearing fault data file with three types of fault: inner race fault, ball fault, and outer race fault. Each type of fault data consists of four diameters of fault bearing 0.00 inch (Class: 1), 0.007 inch (Class: 2), 0.014 inch (Class: 3) and 0.021 inch (Class: 4). Figure 4 depicts the four vibration signals, which are normal, inner race fault, ball fault, and outer race fault. Figure 3 shows how training and validation data are split for the ANN architectures.

Table 1. Information of bearing faults

Bearing Fault Type	**Fault Diameter (inch)**	**Fault Class**	**Motor Speed (RPM)**	**Length of Data File**
Fault in Inner Race	0.00 (Normal)	Class: 1	1797	243938
	0.007	Class: 2	1797	121535
	0.014	Class: 3	1797	121351
	0.021	Class: 4	1797	121168
Fault in Ball	0.00 (Normal)	Class: 1	1797	243938
	0.007	Class: 2	1797	121168
	0.014	Class: 3	1797	122086
	0.021	Class: 4	1797	121351
Fault in Outer Race	0.00 (Normal)	Class: 1	1797	243938
	0.007	Class: 2	1797	121168
	0.014	Class: 3	1797	120984
	0.021	Class: 4	1797	121351

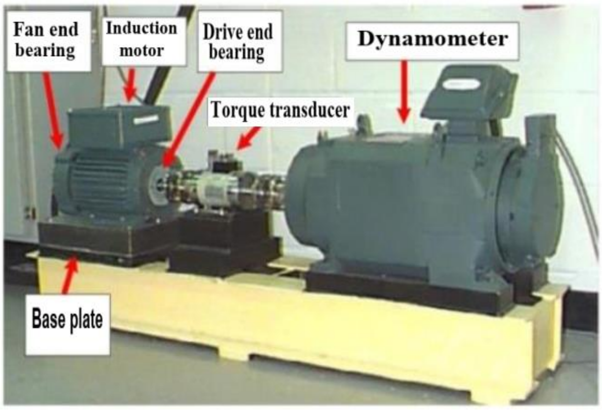


Figure 2. Experiental Setup [21]

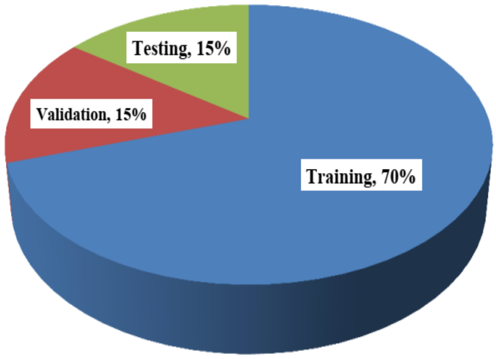


Figure 3. Data Split for Training, Validation and

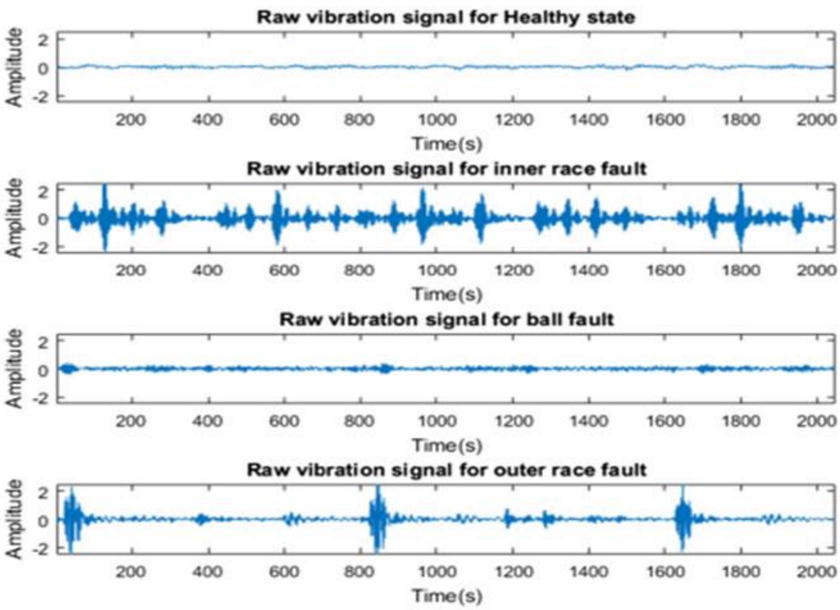


Figure 4. Raw Vibration Signal

2.2. Feature Optimization

Dimensional reduction must be used in order to choose important and discriminating features that improve diagnosis system performance [14], [19]. To address this point, this paper focuses on an effective feature selection-based optimization technique, which reduces the dimensional of input features and raises the learning process, also increasing the diagnosis process efficiency with less elapsed time. At first, the vibration signal is decomposed by WT, and then entropy is applied to determine whether the decomposed part contains significant components of the fault features.

2.2.1. Signal Decomposing by DWT

The bearing's vibration signal frequently comprises components from other elements. Background noise is also present in the vibration signals, which makes extracting fault features difficult. A well-known technique for the analysis of signals is the wavelet transforms [9]. A wavelet transform in either the time domain or the frequency domain can represent the complexity of a non-stationary signal. The discrete wavelet transforms (DWT) is a signal decomposition technique that divides a signal into a number of separate, spatially directed frequency channels. Two filters are applied to the original signal, and two new signals—details and approximation—are produced. This procedure is known as signal decomposition or analysis. To ensure that no information is lost, every detail is then recreated. The breakdown signal's components can be reconstructed again into the original raw signal without losing any information. This is known as synthesis or reconstruction [10]. This paper focuses on wavelet transform to preprocess the vibration raw signals. This paper used the one-dimensional wavelet decomposition up to three levels for the vibration raw signals: normal (healthy), inner race fault, ball fault, and outer race fault. In Figure 5, it shows the output of wavelet decomposition up to three levels. Moreover, it shows the three levels of detail (D1-D3), Vertical (V1-V3), Horizontal (H1-H3), and approximation (A1-A3) which are chosen for each signal.

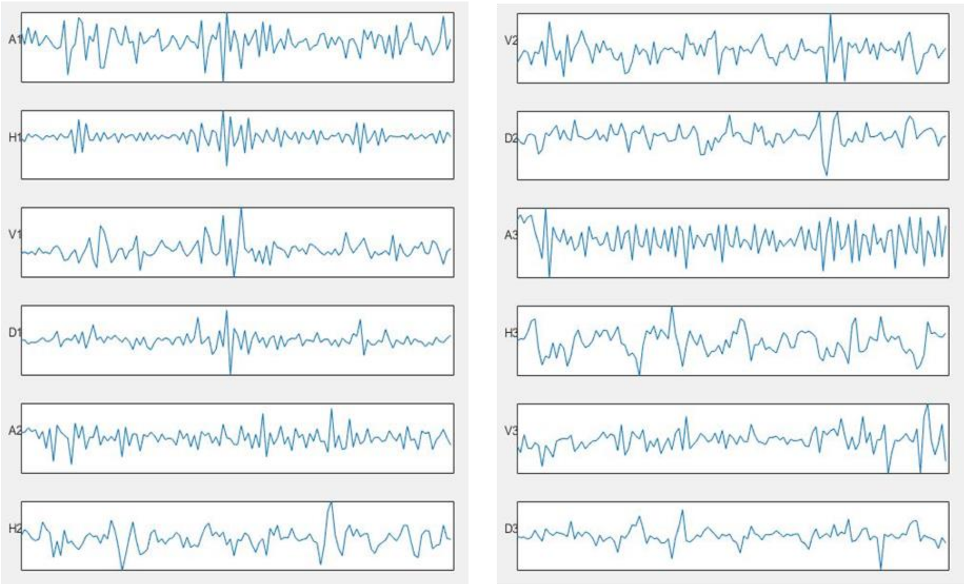


Figure 5. Output Result of Wavelet Transform

2.2.2. Optimum Feature Extraction by Wavelet Entropy

Entropy measures complexity and disorder of a signal. Vibration signals are adjusted appropriately when a single-point defect, also known as an incipient fault, occurs in a bearing. Changes in complexity values from vibration signals can be effectively connected with the incipient fault's growing rate [11]. The wavelet entropy can determine both the complexity and degree of order of a signal. As such, it is capable of offering the required information. Let x[n] be a discrete signal changed at instant K and scale j. It composes two parts: the high-frequency component coefficient, H_j [n], and the low-frequency component coefficient, L_J [n]. The signal x[n] is defined as the total of these components as following equations [12].

$$[n] = \sum_{j=1}^{J} H_j[n] + L_J[n] \qquad (1)$$

Non-normalized Shannon entropy can be described as E_{jk}, where E_{jk} is the wavelet energy spectrum at instant k and scale j.

$$E_j = -\sum_k E_{jk} log E_{jk} \qquad (2)$$

Where

$$E_{jk} = \left|D_j(k)\right|^2 \qquad (3)$$

The minimum Shannon entropy criterion can be provided as an efficient method for generating feature sets. The Shannon Entropy of the associated wavelet coefficients to extract features from the decomposed signal and get four features from each level of the wavelet decomposition. W_{11}, W_{12}, W_{13}, and W_{14} are four features from first-level decomposition, W_{21}, W_{22}, W_{23}, and W_{24} are four features from second-level decomposition, and W_{31}, W_{32}, W_{33}, W_{34} are four features from third-level decomposition. 12 features are extracted from three decomposition levels: W_{11}, W_{12}, W_{13}, W_{14}, W_{21}, W_{22}, W_{23}, W_{24}, W_{31}, W_{32}, W_{33}, and W_{34} as shown in Table 2.

Table 2. Entropy result with feature

Feature (W)	Fault in Inner	Fault in Ball	Fault in Outer
W_{11}	29335.672588	24979.440333	55554.992716
W_{12}	17246.000512	28846.154124	2850.4850231
W_{13}	0	0	0
W_{14}	0	0	0
W_{21}	22086.117344	13850.856714	70336.370370
W_{22}	-34348.580889	26268.966535	34980.320309
W_{23}	0	0	0
W_{24}	0	0	0
W_{31}	18392.438357	10416.198884	62089.722435
W_{32}	-27012.849409	13848.516554	49998.225610
W_{33}	0	0	0
W_{34}	0	0	0

Table 2 shows the values of six like W_{13}, W_{14}, W_{23}, W_{24}, W_{33}, and W_{34} which are zero. It implies that these components do not contain any fault features. Therefore, we can omit these six features and take the other six components as effective features, which are W_{11}, W_{12}, W_{21}, W_{22}, W_{31}, and W_{32}. Table 3 represents the six effective features fed into the ML classifier's input to train the network for classifying the bearing fault.

Table 3. Optimized features

Feature (W)	Fault in Inner	Fault in Ball	Fault in Outer
W_{11}	29335.672588	24979.440333	55554.992716
W_{12}	17246.000512	28846.154124	2850.4850231
W_{21}	22086.117344	13850.856714	70336.370370
W_{22}	-34348.580889	26268.966535	34980.320309
W_{31}	18392.438357	10416.198884	62089.722435
W_{32}	-27012.849409	13848.516554	49998.225610

2.3. Machine Learning (ML) Classifier

In machine learning, a classifier is an algorithm that automatically sorts or groups data into one or more "classes." The proposed model is verified against four machine learning algorithms: naïve Bayes, support vector machine, artificial neural network, and KNN. In this work, the optimized features are applied as the input to various machine learning classifiers, including artificial neural network (ANN), k-nearest neighbor (k NN), support vector machine (SVM), and naïve Bayes to select the best classifier.

2.3.1. ANN Classifier

The most typical challenge in this sector is connected to decision-making, when the problem is difficult to solve using typical computation. ANNs can combine the processing capacity of digital computers with the capability to make rational decisions and learn through regular experience, much like people [22]. The procedure of training an ANN is carried out using the most widely used algorithm known as back-propagation. The ANN has three layers, including an input layer, a hidden layer, and an output layer. This work designs a simpler ANN model which can detect and correctly classify the bearing fault diagnosis, which is shown in Figure 6, where the six features (W_{11}, W_{12}, W_{21}, W_{22}, W_{31}, W_{32}) in the four-fault classes, respectively, are taken as each input of the ANN. The hidden layer consists of 30 nodes, and the output is the four fault classes accordingly, i.e., normal condition, inner race fault, ball fault, and out-race fault. So, the final ANN structure consisted of three layers: the input layer with 6 nodes, the hidden layer with 30 nodes, and the output layer with 4 nodes. Each class is trained by 200 samples and the iteration is 42. The learning speed of the ANN training algorithm is 0.06, and the network is trained indefinitely until it achieves convergence. The network's parameters are given in Table 4, where "mse" is used as a performance function, "trainlm'" is used as a training function, the maximum number of iterations is 1000, and the number of epochs is 30. The fault classes are given as follows:

Class 1—Fault bearing with 0.00-inch diameter (normal bearing) [0 0 0]; Class —2 fault bearing with 0.007-inch diameter [1 0 0]; Class 3— fault bearing with 0.014-inch diameter [0 1 0]; and Class 4— fault bearing with 0.021-inch diameter [0 0 1].

All test samples are correctly recognized by applying a trained ANN to the bearing fault test samples.

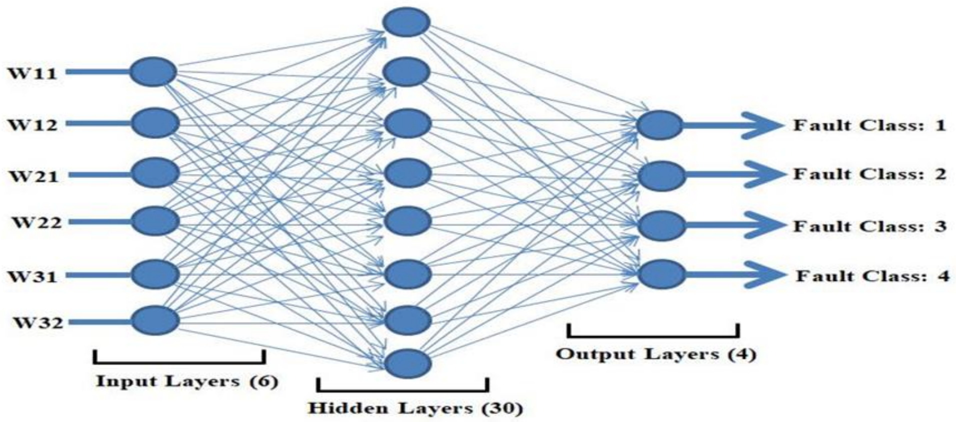


Figure 6. Architecture of ANN

Table 4. Proposed ANN hyper-parameters

Parameters	Values
Epochs	30
Maximum Iterations	1000
Performance Function	'mse'
Training Function	'trainlm'

2.3.2. SVM classifier

Figure 7 shows the SVM architectures for bearing fault classification where the six features (W_{11}, W_{12}, W_{21}, W_{22}, W_{31}, W_{32}) in the four-fault classes are taken as each input of the SVM. There are four outputs in the SVM classifier: class 1, class 2, class 3, and class 4. The network's parameters of SVM are given in Table 5. In the parameter of SVM, the value of C is 1.0, Kerner is "rbf", degree is 3, tol value is 0.001, max_iter is –1, and decision_function_shape is "ovr".

Table 5. SVM classifier's parameters

SVM Parameters	Size/ Value
C	1.0
kernel	rbf
degree	3
gamma	scale
oefo	0.0
Shrin king	true
probability	false
tol	0.001
Cache_size	200
Class_weight	none
verbose	false
Max_iter	-1
Decision_function_shape	ovr
Break_ties	false
Random_state	none

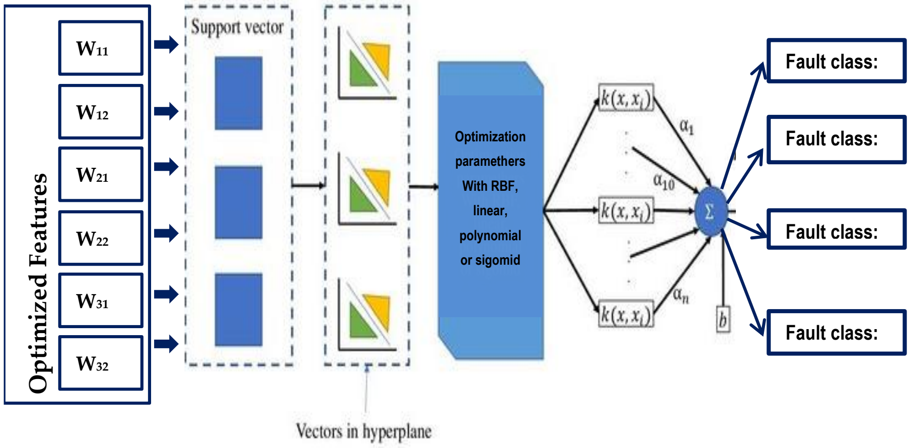


Figure 7. Architecture of SVM classifier

2.3.3. KNN classifier

Figure 8 shows the KNN architectures for bearing fault classification where the six features (W_{11}, W_{12}, W_{21}, W_{22}, W_{31}, W_{32}) in the four-fault classes are taken as each input of the KNN. There are four outputs in the KNN classifier: class 1, class 2, class 3, and class 4. The network's parameters of KNN are given in Table 6. In the parameter of KNN, the n_neighbors is 5, weights are "uniform", algorithm is auto, leaf_size is 30, P is 2, and metric is "minkowski".

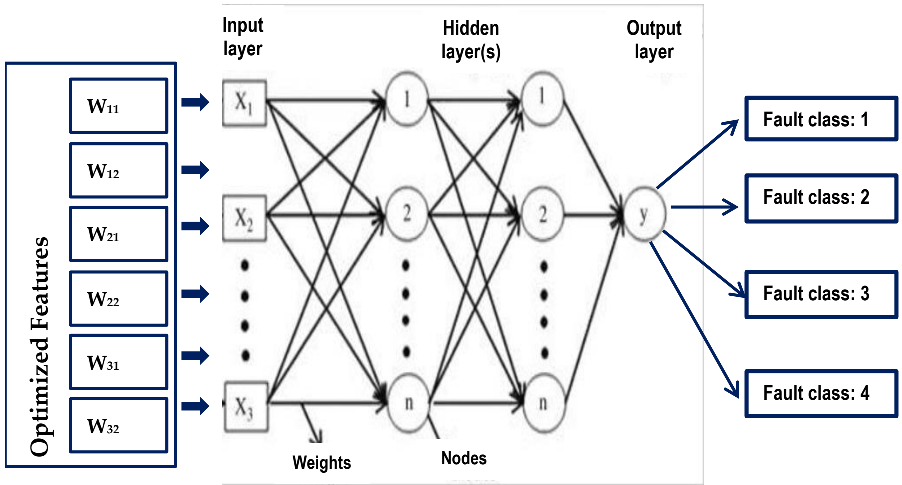


Figure 8. Architecture of KNN classifier

Table 6. KNN classifier's parameters

KNN Parameters	Size/Value
n_neighbors	5
weights	uniform
algorithm	auto
leaf_size	30
p	2
metric	'minkowski'
metric_params	None
n_jobs	None

3. RESULTS AND DISCUSSION

This section represents the fault diagnosis capabilities of the proposed method using various machine learning classifiers.

3.1. ML classifier accuracy

Optimized effective features are applied as the input to various machine learning models like ANN, SVM, KNN, Naïve Bayes, etc. Table 7 shows the effect of the feature optimization process on accuracy and training time on other classifiers. Five classifiers such as Linear SVM, ANN, Cubic SVM, KNN, and Naïve Bayes, are considered to observe the feature optimization effect. Table-7 indicates that the performance accuracy of the feature optimization process is significantly improved compared to reduce training times with the exception of feature optimization. It can be seen that the proposed ANN model provided 99.50% accuracy with minimum training time. Therefore, the ANN model has been selected for further analyzing the model.

Table 7. Effect of feature optimization on ML classifiers

	Except Optimization		**With Optimization**	
Classifier	**Accuracy**	**Training Time**	**Accuracy**	**Training Time**
Linear SVM	99%	39.703 s	99.7%	34.768 s
ANN	**98.10%**	**7.0 s**	**99.50%**	**4.10**
Cubic SVM	99.9%	28.64 s	100%	27.655 s
KNN	99.4%	10.705 s	100%	6.727 s
Naïve Bayes	98.76%	5.41 s	99%	8.646 s

3.2. Performance analysis of proposed ANN model

After the proposed ANN architectures are successfully implemented, the proposed system is trained for CWRU datasets. 70% of the data in this network is used for training, 20% for validation, and 10% for testing. Proposed ANN uses “mse” as a performance function as well as “trainlm” as a training function. The learning rate is 0.002, the number of epoch is 30, and the maximum iterations are 1000. The simulation provides worthwhile values. Then, the accuracies are observed for CWRU datasets. This section considers five scenarios to verify the proposed ANN model:

3.2.1. Performances of hidden layer of proposed ANN model

Table 8 shows the effect of various hidden layers of the ANN model, which also shows that the accuracy was increased when number of the hidden layers was increased. However, when the number of hidden layers reached a certain threshold, its accuracy did not improve much than the previous. Therefore, the optimum number of the hidden layer must be set.

Table 8. comparison results with tuning hidden layer

Hidden Layers	**No of iterations**	**Accuracy of fault in inner race**	**Accuracy of fault in ball**	**Accuracy of fault in outer race**
15	1000	98.3%	98.1%	85.6%
20	1000	98.8%	97.9%	96.8%
25	1000	98.6%	98.0%	97.9%
30	**1000**	**99.5%**	**98.8%**	**99.%**
35	1000	98.3%	74.4%	97.0%

3.2.2. Performances of training function of proposed ANN model

Table 9 represents the effect of changing the train function of the ANN model. Eight training functions are used to find best function for proposed model. It can be seen that the “Traingdm” and “traingda’’ train functions provide less accuracy which is only 30.9% and 72.3% but other train functions provide higher accuracy. This experiment observation confirms that the “trainlm” train function provides the best accuracy (99.5%) than other train functions for this ANN model. Therefore, the proposed model uses the “trainlm” train function.

Table 9. comparison result with tuning train function

Train Function	**Performance Function**	**Hidden Layers**	**Accuracy race**	**Rate of Error**
traingdm	mse	30	30.9%	69.1%
trainggda	mse	30	72.3%	27.7%
trainggdx	mse	30	93.3%	6.7%
trainr	mse	30	97.3%	2.7%
traingrp	mse	30	99.1%	0.9%
traingoss	mse	30	99.3%	0.7%
traingscg	mse	30	99.3%	0.7%
trainglm	**mse**	**30**	**99.5%**	**0.5%**

3.2.3. Performances of performance function of proposed ANN model

After selecting the “trainlm” function as a train function, this research observes three performance functions to find the best one. Table 10 shows the effect of changing the performance function of the ANN model, which shows that the sum of squared errors (SSE) performance function provides only 50% accuracy which is very low, the mean squared error and mean squared weight and bias (MSEREG) performance function provides 98.5% accuracy and the mean square errors (MSE) performance function provides 99.5% accuracy. This experiment observation confirms that the "MSE" as performance function provides the highest accuracy for this ANN model.

Table 10. comparison results tuning performance function

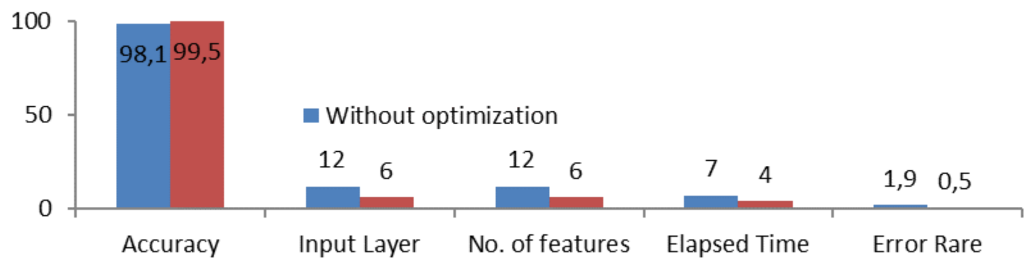
Performance Function	Train Function	Hidden Layers	Accuracy race	Rate of Error
SSE	trainlm	30	50%	50%
MSEREG	trainlm	30	98.5%	1.5%
MSE	**trainlm**	**30**	**99.5%**	**0.5%**

3.2.4. Performances of optimized features for elapsed time and error rate

Figure 9 shows the performance without and with feature optimization; the blue color of the graph indicates the performance without feature optimization and the red color shows the performance with feature optimization. Table 11 demonstrates that there are without-feature optimization methods for classifying bearing faults. The proposed system with the feature optimization method is best and capable of achieving 99.5% accuracy which is high enough for bearing fault classification. This accuracy is enhanced compared to the without feature optimization method. Table 11 also clearly displays that the proposed system with a feature optimization process takes only 4 seconds to obtain the testing accuracy. On the other hand, without a feature optimization process, the system takes more time than the proposed system with a feature optimization process. The error rate is also reduced to only 0.5% after using feature optimization. This experiment observation confirms that the proposed system with a feature optimization process reduces the dimensionality of input features and raises the learning process, also increasing the diagnosis process efficiency with less elapsed time.

Table 11. Performance analysis on feature optimization

Particulars	Without optimization	With optimization
Input Layer	12 (More)	6 (Less)
Number of features	12 (More)	6 (Less)
Elapsed Time	7sec	4sec
Accuracy	98.1%	99.5%
Error Rare	1.9%	0.5%



3.2.5. Accuracy for each fault diameters

Table 12 depicts the overall result of the neural network respective to the inner race fault, the ball fault, and the outer race fault, with various fault diameters which shows a summary of the results of classification for four classes on the CWRU dataset. The first row of in table 12 shows as depicted in the accuracy result of the inner race fault of bearing. Thus, the classification accuracy for class 1, class 2, class 3 and class 4 are 97.1%, 100%, 100% and 100%, respectively. The overall accuracy of the inner fault of the bearing is 99.3%. The second row of in Table 12 shows as depicted in the accuracy confusion matrix of ball fault of bearing. Thus, the classification accuracy for class 1, class 2, class 3, and class 4 are 100%, 96.1, 99.5, and 99.5%, respectively. The overall accuracy of ball fault of the bearing is 98.8%. The third row in Table 12 shows as depicted in the confusion matrix accuracy of outer race fault is shown in the third row of table 12 of bearing. Thus, the classification accuracy for class 1, class 2, class 3 and class 4 is 96.2%, 100%, 100% and

100%, respectively. The testing accuracy of the proposed system for bearing fault classification is recorded in the following Table 12. This table is presented to help explore the best result.

Table 12. Classification accuracy of the system for various fault diameters

Type of Fault	Diameter of Fault	Accuracy	Overall Accuracy
Fault in Inner Race	0.00 inch (Normal)	97%	99.5%
	0.007 inch	100%	
	0.014 inch	100%	
	0.021 inch	100%	
Fault in Ball	0.00 inch (Normal)	100%	98.8%
	0.007 inch	96.1%	
	0.014 inch	95.5%	
	0.021 inch	95.5%	
Fault in Outer Race	0.00 inch (Normal)	96.2%	99.0%
	0.007 inch	100%	
	0.014 inch	100%	
	0.021 inch	100%	

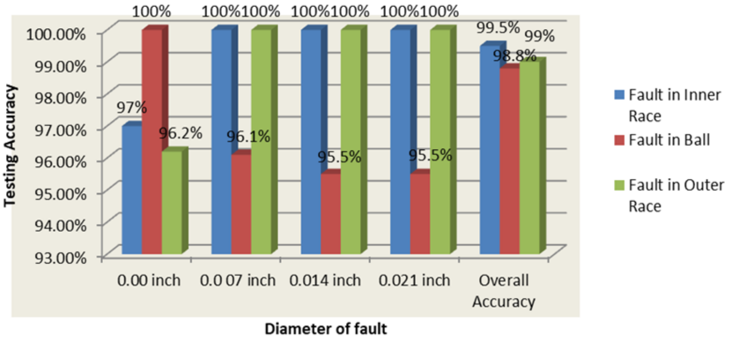


Figure 10. Graphical representation of the testing accuracy of the proposed system

Figure 10 shows the testing performance of the proposed system, four-fault diameter and among them, the blue color of the graph indicates the fault in inner race accuracy, which shows in 0.00 inch diameter of a fault, the accuracy is less (97%). However, when the diameter of fault increased by 0.007 inch, 0.014 inch and 0.021 inch, the accuracy increases up to 100%. The red color indicates the fault in the ball, and the green color indicates the fault in the outer race of bearing fault classification accuracy, which shows accuracy, is 100% except for 0.00 inch diameter of the fault.

3.3. Evaluate under different noisy environments

In the real environments of industry, signals are affected by noise. Noise is another problem due to the change in working conditions, which reduces the performance of bearing fault diagnosis [22]. This part analyses the proposed system's performance under various noisy environments. The Additive Gaussian White Noise (AGWN) is added to the original raw signal to create the noisy signals. Figure 11 shows the process of creating noisy signals.

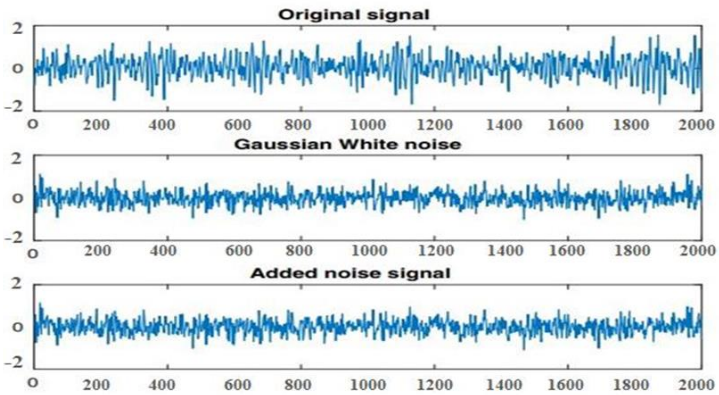


Figure 11. Process of creating noisy signal process

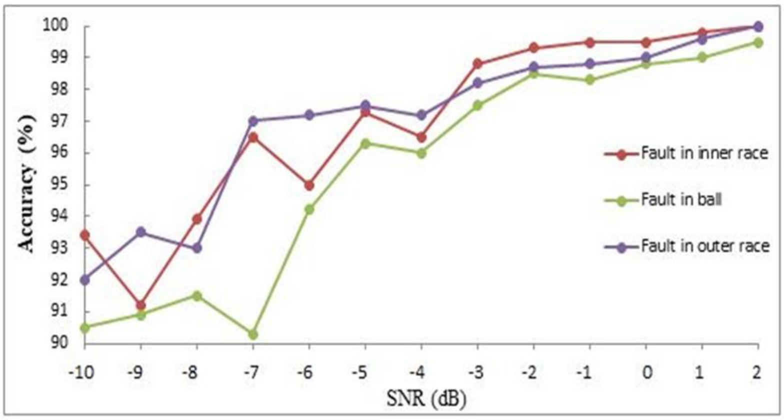


Figure 12. Performance of the proposed model under different noisy

In the environments of real industry, signals are affected by noise. Noise is a big problem due to the change in working conditions, which reduces the performance of bearing fault diagnosis. This section examines the effectiveness of the proposed system under noisy conditions with different values of SNR. The performance of the proposed model is shown in figure 12 under noisy conditions with varying SNR values from -10dB to +2 dB. In SNR=-3 dB, the proposed model still achieves 98.8%, 97.5%, and 98.2 % accuracy, respectively, inner, ball, and outer fault. SNR= -10 dB is the more considerable noise power, where it is more difficult to classify the bearing fault diagnosis. However, the proposed system provides 93.4%, 90.5%, and 92% accuracy, respectively, for inner, ball, and outer race faults. The proposed system's performance in this experiment exhibits its high reliability in defect diagnosis, even in noisy situations.

3.4. Comparison of this work with some previous research using CWRU bearings datasat

This section compares the performance of the proposed system with several reported results in the literature. In Table 13, it shows the comparative studies on the CWRU dataset. In [24], To improve feature extraction, the Short-Time Fourier Transform (STFT) is applied as a preprocessing technique. Gammatone Transformation and raw data are also taken into consideration in order to establish the optimal preprocessing method. Few-shot learning (FSL) is used to overcome data scarcity, which is one of the main barriers to fault detection in the industrial context and the testing accuracy was 94.82%. According to [25], an Attention Mechanisms was used to evaluate the efficacy of the feature extraction approach based on EMD and GS filter and the testing precision was 97%. In [26], Various machine learning models was also employed to evaluate the efficiency of the feature reduction approach based on Frequency domain vibration analysis & envelope analysis and the accuracy is 94.4%. A Lite and Efficient Deep Learning Model were employed in [27] to determine the efficacy of the feature reduction technique based on spectrograms and the testing accuracy was 98.66%. It can be seen from Table 10 that the proposed method provides a greater accuracy (99.5%) than Ref. [24-27].

Table 13. comparison result with several works

Reference	Method	Preprocessing	Published Year	Accuracy
[24]	A Few-Shot Learning	Gammatone Transformation	2023	94.82%
[25]	Attention Mechanisms	EMD and GS filter	2024	97%
[26]	Various machine learning models	Frequency domain vibration analysis & envelope analysis	2023	94.4%
[27]	Lite and Efficient Deep Learning Model	spectrograms	2023	98.66%
Proposed system	ANN, DWT and Wavelet Entropy with feature optimization technique	DWT & Entropy	-	99.5%

The results show that the proposed system regularly outperforms other recent research regarding diagnostic accuracy. This indicates that the proposed system efficiently addresses the difficulty of extracting in-depth features from complex vibration utilizing DWT and minimum Shannon entropy signal analysis approaches, as well as fulfilling the effective features requirements for ANN training while showing outstanding reliability.

4. CONCLUSION

Bearing is a machinery device that constrains relative motion to only the desired motion and reduces friction between moving parts. The health of bearings has a significant impact on machines. Earlier detection and classification of bearing fault detection is an important aspect of machine health monitoring. A typical intelligent diagnosis method includes two main parts feature extraction and fault classification. The feature part needs signal processing techniques, which requires specialist knowledge, and human labor. The bearing fault diagnosis method does not produce a satisfactory result due to the effective feature selection difficulty. Additionally, because the working load is constantly changing and noise from the place of operation is unavoidable, the efficiency of intelligent fault diagnosis techniques suffers great reductions.

To address the above problem of bearing fault detection and classification this research proposed a feature optimization based bearing fault classification method that can effectively solve the problem. The proposed method consists of two main parts including feature optimization and feature classification using machine learning classifier. Firstly, The vibrational signal is decomposed using the discrete wavelet transforms (DWT) utilizing the Haar wavelet to yield both approximations and details. To ensure that no information is lost, every detail is then recreated. The minimum Shannon entropy criterion is then used to the reconstructed details that were collected to identify the fault frequencies, which are regarded as novel features and then using an optimization process. This combination of approaches allows effective feature extraction and structuring. Finally those features are applied as a input of four machine learning algorithms: naïve Bayes, support vector machine, artificial neural network, and KNN. Because of the success rate of ANN classifier is higher compared to other machine learning classifier, The ANN is chosen for further analysis.

The CWRU bearing dataset, which includes three different types of bearing faults, is employed to evaluate the proposed system. To verify the proposed ANN model five scenarios are considered like performances of hidden layer, performances of training function, performances of performance function, performances of optimized features for elapsed time and error rate, and accuracy for each fault diameters. Finally, the proposed system was tested satisfactorily and obtained 99.5%, 98.8%, and 99% accuracy for the inner race fault, the ball fault, and the outer race fault, respectively.
In the real environments of industry, signals are affected by noise. Noise is another problem due to the change in working conditions, which reduces the performance of bearing fault diagnosis. The performance in noisy situations was investigated in this research using various SNR values. The results of this experiment highlight the great reliability of the proposed system in fault diagnosis even in noisy environments. Future research can make an effort to employ an automated feature extraction-based deep learning approach, which will analyze the performance in different load conditions, and noisy conditions.

ACKNOWLEDGEMENTS

This research is funded by Woosong University Academic Research 2024.

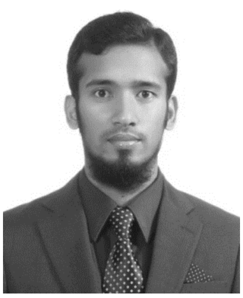
REFERENCES

[1] R. B. Randall and J. Antoni, “Rolling element bearing diagnostics—a tutorial,” Mechanical systems and signal processing, vol. 25, no. 2, pp. 485–520, 2011.
[2] 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.
[3] B. Samanta, K. Al-Balushi, and S. Al-Araimi, “Artificial neural networks and support vector machines with genetic algorithm for bearing fault detection,” Engineering applications of artificial intelligence, vol. 16, no. 7-8, pp. 657–665, 2003.
[4] P. K. Kankar, S. C. Sharma, and S. P. Harsha, “Rolling element bearing fault diagnosis using wavelet transform,” Neurocomputing, vol. 74, no. 10, pp. 1638–1645, 2011.
[5] B. Nayana and P. Geethanjali, “Effective time domain features for identification of bearing fault using lda and nb classifiers,” Int J Mech Product Eng Res Dev, vol. 8, pp. 1135–1150, 2018.
[6] T. P. K. Nguyen, A. Khlaief, K. Medjaher, A. Picot, P. Maussion, D. Tobon, B. Chauchat, and R. Cheron, “Analysis and comparison of multiple features for fault detection and prognostic in ball bearings,” in Fourth european conference of the prognostics and health management society 2018, 2018, pp. 1–9.
[7] Kabir, M.H.; Kabir, M.A.; Islam, M.S.; Mortuza, M.G.; Mohiuddin, M. Performance Analysis of Mesh Based Enterprise Network Using RIP, EIGRP and OSPF Routing Protocols. Eng. Proc. 2021, 10, 47. https://doi.org/10.3390/ecsa-8-11285.
[8] J. Shiroishi, Y. Li, S. Liang, T. Kurfess, and S. Danyluk, “Bearing condition diagnostics via vibration and acoustic emission measurements,” Mechanical systems and signal processing, vol. 11, no. 5, pp. 693–705, 1997.
[9] M. Mohiuddin and M. S. Islam, “Rolling element bearing faults detection and classification technique using vibration

signals," Engineering Proceedings, vol. 27, no. 1, p. 53, 2022.
[10] S. Shanjida, M. S. Islam and M. Mohiuddin, "MRI-Image based Brain Tumor Detection and Classification using CNN-KNN," 2022 IEEE IAS Global Conference on Emerging Technologies (GlobConET), Arad, Romania, 2022, pp. 900-905, doi: 10.1109/GlobConET53749.2022.9872168.
[11] P. W. Tse, Y. Peng, , and R. Yam, "Wavelet analysis and envelope detection for rolling element bearing fault diagnosis—their effectiveness and flexibilities," J. Vib. Acoust., vol. 123, no. 3, pp. 303–310, 2001.
[12] Saida Dahmane, Fouad Berrabah, Mabrouk Defdaf, Saad Salah, "Diagnosis and Monitoring Method for Detecting and Localizing Bearing Faults". IJEEI, vol. 12, no. 1, http://dx.doi.org/10.52549/ijeei.v12i1.4612.
[13] Zabin, M., Choi, HJ. & Uddin, J. Hybrid deep transfer learning architecture for industrial fault diagnosis using Hilbert transform and DCNN–LSTM. J Supercomput 79, 5181–5200 (2023). https://doi.org/10.1007/s11227-022-04830-8.
[14] Zabin, M.; Choi, H.-J.; Kabir, M.K.; Kabir, A.N.B.; Uddin, J. Machine Fault Diagnosis: Experiments with Different Attention Mechanisms Using a Lightweight SqueezeNet Architecture. Electronics 2024, 13, 3112. https://doi.org/10.3390/electronics13163112
[15] V. Purushotham, S. Narayanan, and S. A. Prasad, "Multi-fault diagnosis of rolling bearing elements using wavelet analysis and hidden markov model based fault recognition," Ndt & E International, vol. 38, no. 8, pp. 654–664, 2005.
[16] S. Prabhakar, A. R. Mohanty, and A. Sekhar, "Application of discrete wavelet transform for detection of ball bearing race faults," Tribology International, vol. 35, no. 12, pp. 793–800, 2002.
[17] N. Saravanan, V. K. Siddabattuni, and K. Ramachandran, "A comparative study on classification of features by svm and psvm extracted using morlet wavelet for fault diagnosis of spur bevel gear box," Expert systems with applications, vol. 35, no. 3, pp. 1351–1366, 2008.
[18] Y. Yu, C. Junsheng et al., "A roller bearing fault diagnosis method based on emd energy entropy and ann," Journal of sound and vibration, vol. 294, no. 1-2, pp. 269–277, 2006.
[19] M. Mohiuddin, M. S. Islam, and M. H. Kabir, "Performance analysis of bearing fault diagnosis using convolutional neural network," in 2021 IEEE 4th International Conference on Computing, Power and Communication Technologies (GUCON). IEEE, 2021, pp. 1–6.
[20] S. Shanjida, M. S. Islam and M. Mohiuddin, "Hybrid model-based Brain Tumor Detection and Classification using Deep CNN-SVM," 2024 6th International Conference on Electrical Engineering and Information & Communication Technology (ICEEICT), Dhaka, Bangladesh, 2024, pp. 1467-1472, doi: 10.1109/ICEEICT62016.2024.10534376.
[21] V. Vakharia, V. Gupta, and P. Kankar, "A multiscale permutation entropy based approach to select wavelet for fault diagnosis of ball bearings," Journal of Vibration and Control, vol. 21, no. 16, pp. 3123– 3131, 2015.
[22] I. A. Basheer and M. Hajmeer, "Artificial neural networks: fundamentals, computing, design, and application," Journal of microbiological methods, vol. 43, no. 1, pp. 3–31, 2000.
[23] Mohiuddin, M.; Islam, M.S.; Islam, S.; Miah, M.S.; Niu, M.-B. Intelligent Fault Diagnosis of Rolling Element Bearings Based on Modified AlexNet. Sensors 2023, 23, 7764. https://doi.org/10.3390/s23187764
[24] Siraj, F.M.; Ayon, S.T.K.; Uddin, J. A Few-Shot Learning Based Fault Diagnosis Model Using Sensors Data from Industrial Machineries. Vibration 2023, 6, 1004-1029. https://doi.org/10.3390/vibration6040059
[25] Zabin, M.; Choi, H.-J.; Kabir, M.K.; Kabir, A.N.B.; Uddin, J. Machine Fault Diagnosis: Experiments with Different Attention Mechanisms Using a Lightweight SqueezeNet Architecture. Electronics 2024, 13, 3112. https://doi.org/10.3390/electronics13163112.
[26] M. Alonso-González, V. G. Díaz, B. L. Pérez, B. C. P. G-Bustelo and J. P. Anzola, "Bearing Fault Diagnosis With Envelope Analysis and Machine Learning Approaches Using CWRU Dataset," in IEEE Access, vol. 11, pp. 57796-57805, 2023, doi: 10.1109/ACCESS.2023.3283466.
[27] Yoo, Y.; Jo, H.; Ban, S.-W. Lite and Efficient Deep Learning Model for Bearing Fault Diagnosis Using the CWRU Dataset. Sensors 2023, 23, 3157. https://doi.org/10.3390/s23063157.

BIOGRAPHIES OF AUTHORS

Mohammad Mohiuddin, He received his B. Sc. in Engineering degree from the Department of Electrical & Electronics Engineering (EEE). He obtained his M. Sc. in Engineering degree from Chittagong University of Engineering and Technology (CUET) in 2023. He is very much enthusiastic about Machine Learning, Artificial intelligence, Internet of Things (IoT), Signal Processing, Industrial Fault Diagnosis, Wireless Communication, and Automation related researches. The above research lines have produced more than 20 publications on international journals and presentations within international conferences. He can be contacted at email: muhin.cuet.bd@gmail.com.



Md. Saiful Islam is an Associate Professor of dept. of Electronics and Telecommunication Engineering at Chittagong University of Engineering and Technology, Chittagong, Bangladesh. He received the BSc degree in electrical and electronic engineering from Chittagong University of Engineering and Technology in 2010. He obtained MS and PhD degree from University of Ulsan, Korea. His research interests include Artificial intelligence, signal processing, image and video processing, fault detection and diagnosis. The above research lines have produced more than 50 publications on international journals, presentations within international conferences, and book chapters. He also acts as a referee in several highly reputed journals and international conferences.. He can be contacted at email: saiful05eee@cuet.ac.bd

Dr Jia Uddin is as an Assistant Professor, Department of AI and Big Data, Endicott College, Woosong University, Daejeon, South Korea. He received Ph.D. in Computer Engineering from University of Ulsan, South Korea, M.Sc. in Electrical Engineering (Specialization: Telecommunications), Blekinge Institute of Technology, Sweden, and B.Sc. in Computer and Communication Engineering, International Islamic University Chittagong, Bangladesh. He was a visiting faculty at School of computing, Staffordshire University, United Kingdom. He is an Associate Professor (now on leave), Department of Computer Science and Engineering, Brac University, Dhaka, Bangladesh. His research interests are Industrial Fault Diagnosis, Machine Learning/Deep Learning based prediction and detection using multimedia signals.