Using MLP-GABP and SVM with wavelet packet transform-based feature extraction for fault diagnosis of a centrifugal pump

Al Tobi, Maamar; Bevan, Geraint; Wallace, Peter; Harrison, David; Okedu, Kenneth Eloghene

[Link to publication in ResearchOnline](#)

Using MLP-GABP and SVM with wavelet packet transform-based feature extraction for fault diagnosis of a centrifugal pump

Maamar Al Tobi[1] [ID] | Geraint Bevan[2] | Peter Wallace[2] | David Harrison[2] |
Kenneth Eloghene Okedu[3,4] [ID]

[1]Mechanical & Industrial Engineering, College of Engineering, National University of Science and Technology, Muscat, Oman

[2]School of Engineering and Built Environment, Glasgow Caledonian University, Glasgow, UK

[3]Electrical and Computer Engineering, College of Engineering, National University of Science and Technology, Muscat, Oman

[4]Electrical and Electronic Engineering, Nisantasi University, Istanbul, Turkey

Correspondence
Kenneth Eloghene Okedu, Electrical and Computer Engineering, College of Engineering, National University of Science and Technology, Muscat, Oman.
Email: okedukenneth@nu.edu.om

Abstract

This paper explores artificial intelligent training schemes based on multilayer perceptron, considering back propagation and genetic algorithm (GA). The hybrid scheme is compared with the traditional support vector machine approach in the literature to analyze both fault and normal scenarios of a centrifugal pump. A comparative analysis of the performance of the variables was carried out using both schemes. The study used features extracted for three decomposition levels based on wavelet packet transform. In order to investigate the effectiveness of the extracted features, two mother wavelets were investigated. The salient part of this work is the optimization of the hidden layers numbers using GA. Furthermore, this optimization process was extended to the multilayer perceptron neurons. The evaluation of the model system performance used for the study shows better response of the extracted features, and hidden layers variables including the selected neurons. Moreover, the applied training algorithm used in the work was able to enhance the classifications obtained considering the hybrid artificial intelligent scheme been proposed. This work has achieved a number of contributions like GA-based selection of hidden layers and neuron, applied in neural network of centrifugal pump condition classification. Furthermore, a hybrid training method combining GA and back propagation (BP) algorithms has been applied for condition classification of a centrifugal pump. The obtained results have shown the good ability of the proposed methods and algorithms.

KEYWORDS

back propagation (BP), centrifugal pump, genetic algorithm (GA), multilayer feedforward perceptron (MLP), support vector machine (SVM), wavelet packet transform (WPT)

1 | INTRODUCTION

Various techniques have been applied for fault detection of centrifugal pumps regarding condition monitoring, such as analysis based on time[1] and frequency domain, considering fast Fourier transform (FFT).[2,3] Also, a technique known as wavelet that is multi-resolution and powerful in nature has been employed for detection of faults in rotating machinery. This technique also has the capability of non-stationary signal analysis in machines.[3-5]

Wavelet transform (WT) is a method that has the ability to deal with non-stationary signals as it is known for its

time-frequency domain resolution. However, looking for a more advanced and automatic fault diagnosis method, then, artificial neural network (ANN) is a promising technique[6,7] that has proved its ability as a fault classifier. ANN has been employed for the automatic classification of fault diagnosis in centrifugal pumps in references.[8-10] The main concept of WT is based on presenting both the time and frequency domains, in which it possess the ability to analyze the non-stationary signals that are not easy to be analyzed using Fourier transform. Besides, it is a multi-powerful resolution method that requires fast calculation process.[11] Among the different WT methods, the wavelet packet transform (WPT) is one of the promising type that has been applied along with AI in machines fault diagnosis. Reference,12 used three different classifications, namely, multilayer feedforward perceptron (MLP), support vector machine (SVM), and radial basis function (RBF) to diagnose a fault of rub impact. WPT with db4 wavelet function is used for the feature extraction, and the classification rates are 82%, 99.3%, and 98.6% for MLP, SVM, and RBF, respectively.

For machinery components like gears and bearings,[13,14] genetic algorithm (GA) may be used and even extended to hidden layers of neuron numbers[15] of MLP-ANN. Applications of GA for centrifugal pump have been analyzed with two WT methods; CWT with best classification rates of 99.5% and 94.64%, using MLP and SVM, respectively in,[16] and DWT with best classification rates of 100% and 99.8%, using SVM and MLP, respectively in.[17] It is also applied in training MLP combined with back propagation (BP) using both CWT- and DWT-based feature extraction, and the best rates are 88.5% and 89%, respectively. It is also concluded that MLP-BP and SVM in combination with WPT would help achieve classification rates that are better than continuous wavelet transform (CWT).[18]

ANN and SVM have gained significant interest in the area of machine fault diagnosis due to the successful performance and implementation by these classifiers.[19] Moreover, there can be more algorithms and methods to be further examined in the future including the optimization algorithms. Though there are other optimization algorithms that have been recently implemented in engineering applications that are similar in the area of condition monitoring, making them promising for future works regarding machinery fault diagnosis. Multi-strategies quantum-inspired differential evolution (QDE) is an evolutionary algorithm that can be used for solving complex optimization problems as reported in.[20] Also, differential evolution (DE) was successfully applied to estimate the parameters of the photovoltaic (PV) models in.[21]

In this paper, the extracted features were based on WPT considering a three-level decomposition. The variables were decomposed considering both low- and high-frequency-based approximations. The decomposition tree was used to obtain the detail coefficients of the schemes. Classification performance of the MLP-BP based on GA and conventional SVM scheme for both fault and normal conditions of a centrifugal pump operation was considered as a case study. Data collection, data preprocessing based on feature extractions, and classification of faults are the three main stages employed in this work. The db4 and rbio1.5 were the two mother wavelets considered in this paper, for effective feature extraction considering WPT. Besides, the centrifugal pump working conditions were classified and diagnosed using both MLP and SVM artificial intelligent schemes. MLP was analyzed with respect to back-propagation algorithm and compared with MLP-GABP hybrid training algorithm. Manual selection was used for both the hidden layers network and neurons, in order to obtain an optimized GA performance. Figure 1 shows the training algorithm and methods of fault diagnosis employed for the AI schemes in this study. The hidden layers numbers, neural network neurons, features extracted, training and kernel strategies were some of the metrics used to judge the performance of the system. The novel contribution of this work can be emphasized with the GA-based selection of hidden layers and neuron. This has been applied for the neural network of centrifugal pump condition classification; a hybrid training method combining GA and BP algorithms for condition classification of a centrifugal pump; signal decompositions (approximations and details) have also been investigated based on features extracted from centrifugal pump data using DWT and WPT; and WPT has been applied for feature extraction from centrifugal pump data using two mother wavelet functions: db4 and rbio1.5. In this work, a number of contributions and innovations were realized and some of the main ones are a new rig has been designed, built, tested, and used to investigate mechanical and hydraulic faults in a centrifugal pump; WPT has been applied for feature extraction from centrifugal pump data using two mother wavelet functions of db4 and rbio1.5; and signal decompositions (approximations and details) have been investigated based on features extracted from centrifugal pump data using WPT.

The following are the sections of the paper. In Section 2, AI systems including MLP-NN, SVM classifiers, and GA brief overview were given. The experimental model employed in the work is given in Section 3. The applied strategies and procedures considering WPT for extracted features were presented in Section 4. Section 5 presents the classification methods. The evaluation of the system performance was carried out in Section 6, while a conclusion was drawn in Section 7.

2 | BRIEF ON ARTIFICIAL INTELLIGENCE SYSTEMS

The use of computational system[22] could help in mental abilities of replicating and automatic detection of fault in AI
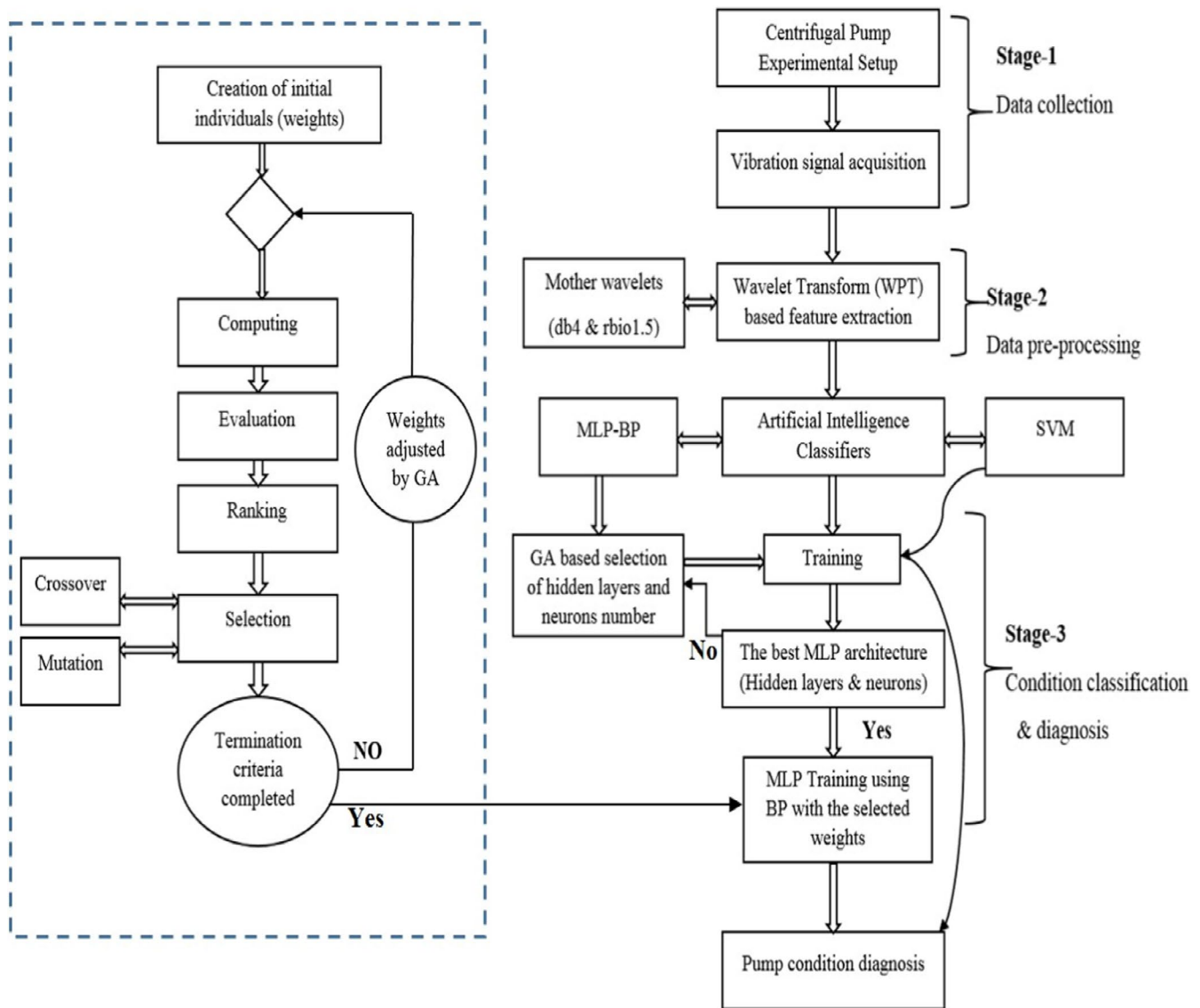
FIGURE 1 Algorithm for diagnosis and training techniques

schemes. References23,24 introduced artificial neural network (ANN) and fuzzy logic techniques, respectively.

Artificial intelligence systems have been applied in the area of computational systems,[22-24] for centrifugal pump fault diagnosis,[10,11] fast Fourier transforms (FFT),[25-28] and time-frequency method.[29-33] In reference,[34] ANN and a fuzzy neural network were applied for faults diagnosis of a centrifugal pump, where statistical methods were used for extracting the features based on time and spectral analysis.

There are many categories of AI in the literature used for diagnosing faults automatically in various types of rotating machines. Back propagation artificial neural network (BP-ANN) or multilayer MLP,[13-16] and SVM,[35-40] were some of the common schemes considered in the reported works.

MLP consists basically of three neuron layers: input, hidden, and output. Between the input and output layers, there may be several hidden layers. While the generalizability of

the system depends on the neuron numbers, its efficiency rely on the neurons and hidden layers. The training data could be over fitted with a larger number, leading to weak generalization of new data. Based on this fact, some strategies like genetic algorithm[41] are imperative to effectively consider the appropriate layers that are hidden and also the neurons. Generally, based on the fault classifications, the output layer could be more than one layer. The role of each hidden layers with number of neurons is to determine the inputs weighted sum and execute it as an activation function. Back propagation algorithm is normally employed in the training of AI MLP scheme,[42] and its performance has been proven via comparative studies with other ANN schemes.[16,41] However, the MLP has shortcoming in the sense that it is slow in training and requires longer time of computation compared with other schemes.[38,40] On the other hand, this weakness can be mitigated by reducing the number of input features.[35]

SVM[43] is another type of AI scheme, which utilizes a recognition based on pattern, as a new solution, using non-linear projections of features input relative to a much higher dimensional pattern area. The basic working principle and conditions of SVM are shown in Figure 2,[17] to classify the various conditions of operation namely: class A and class B. The separator which is the optimal hyper plane separates a margin, between the two classes for better linear classification. The hyper plane, which is a linear classifier, is expressed mathematically in Figure 2, where W is the weight vector and b is the bias.[18]

SVM has proven high efficiency in the literature over other AI schemes like MLP (ANN-BP)[12,13,15] and RBF.[37] This paper uses SVM scheme and compares its performance with MLP, using MATLAB platform for data training and testing. Also, classification is implemented for the different conditions.

GA is a form of optimization technique applied to complex functions[44] and is another AI scheme that is based on the Darwinian-type fitness concept for survival, to enable individuals of the desired problem, for effective competition and matching of variables. This type of AI scheme has similarity with human chromosomes that are represented by linear strings.[45] Figure 3 displays the concept of GA by initiating individual population known as chromosomes that are computed, evaluated, and ranked on the basis of fitness, according to survival rates. Crossover and mutation are the two main operators of this type of AI scheme, in order to help generate new chromosomes that are sent to the iteration process of healthy ones.[45]

This paper uses GA to optimize hidden layers numbers and neurons for optimal architecture of the selection of neural network using MATLAB platform. The range of constraints and parameters space from 1 to 4 layers is considered with about 30 neurons in each layer, to produce around 20 generations, based on a population size of 10 chromosomes, in order to avoid unnecess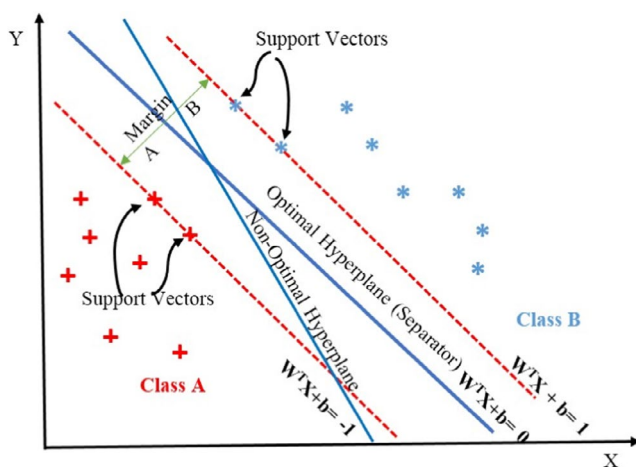ary time of computation. Furthermore, this AI scheme is used to train MLP based on BP employing 1000 generations and size of population.

Practically, GA has been applied using MATLAB with a function handle code and a main one with "ga" function. The "ga" function has to be linked with the handle function to optimize the neural network (weights and biases) and then to minimize mean square error (MSE). The function handle code contains the objective function which represents MSE, where it returns MSE for the given variables (ie, weights, biases, and the network with its inputs and targets, and the main code has been identified with the GA's parameters such as the size of population and number of generations). Using "gaoptimset" to produce genetic algorithm options structure, the following options have been selected and identified:

Number of generations, size of population, function tolerance "TolFun," mutation option, and crossover option have been adjusted after many attempts/tests and found that optimization-based training are with good performance using the following values:

No. of generations = 1000;
Size of population = 1000;
TolFun = 1e−60; mutation option = using Gaussian function "mutationgaussian"; "which is a good option for the nonlinear problems."
Crossover option = using function "crossoverscattered"; "which is also a good option for the nonlinear problems."

For the optimization-based neural network hidden neurons and layers selection with good performance the following values were used:

No. of generations = 40; to avoid long computational time
Size of population = 10; to avoid long computational time
TolFun = 1e−60; mutation option = using Gaussian function "mutationgaussian"; "which is a good option for the nonlinear problems." Crossover option = using function "crossoverscattered"; "which is also a good option for the nonlinear problems."

3 | EXPERIMENTAL SETUP

Figure 4 shows the experimental model system of study for the centrifugal pump with bearing and the faulty impeller, coupled to a motor of specifications as follows: (Saer company, Italy, model: NCBZ-2P-50-125C, 2.2 kW, 3-phase, 420 V, head 8-17 m, and flow rate 500-1000 L/M). The system has control panel with speed controller (Schneider model VFD with speed controller and display screen, switch (OFF/ON), and it is embedded with an emergency shutdown). There is a digital turbine flow meter with specifications; (USA-TM model, 2-inch diameter) pressure gauges, vacuum pump and
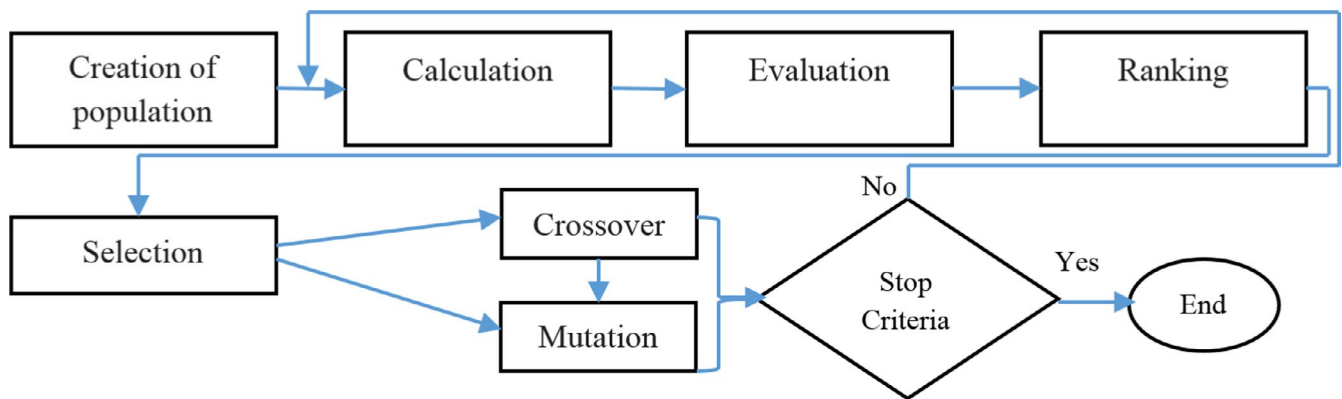

FIGURE 2 Working principle of support vector machine

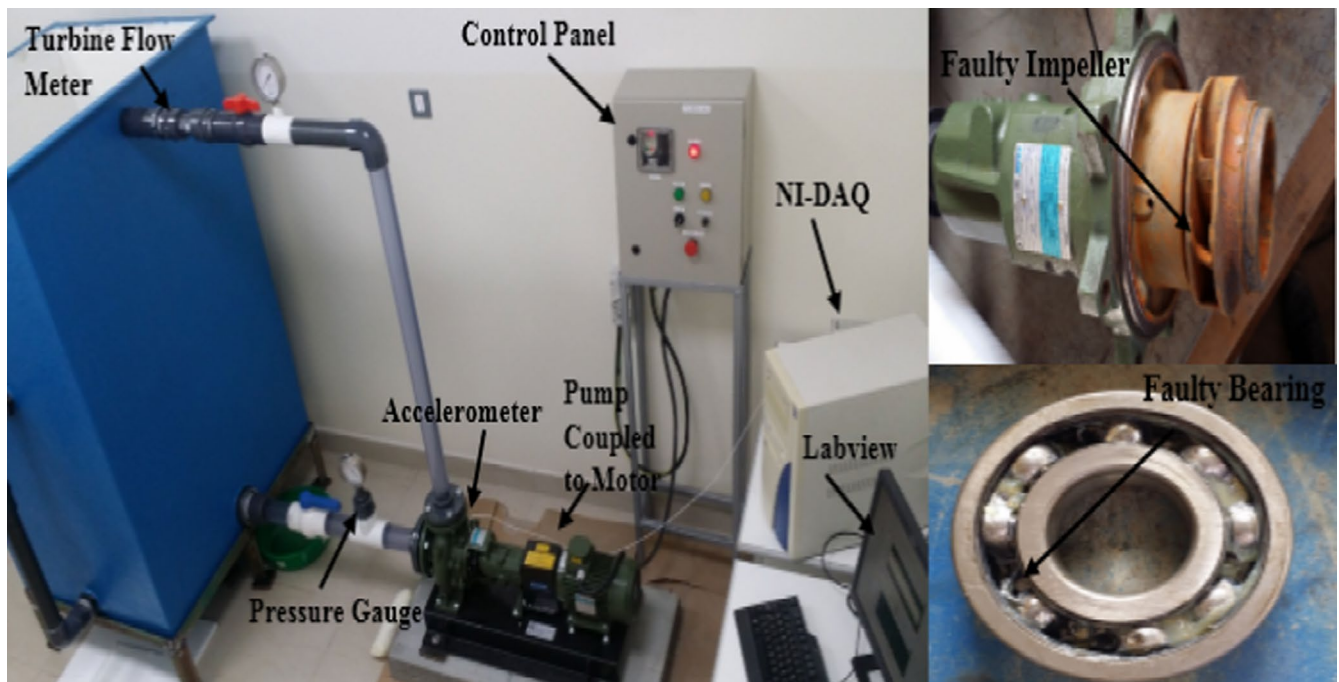**FIGURE 3** Flow chart of genetic algorithm

FIGURE 4 Experimental model of the study

clear PVC pipes; and spare parts: a rolling element bearing, mechanical seal, gasket, and impeller. The model system uses data acquisition system (DAQ) that is made of SCXI-1000 and SCXI-1530 models and National Instruments (NI) accelerometers IMI 621B40, having frequency range 3.4 Hz to 18 kHz for ($\pm$10%) and 1.6 Hz to 30 kHz for ($\pm$3 dB), and sensitivity of 10 mV/g.

Two conditions, healthy and faulty, were used to measure the vibration signals. During healthy conditions of the pump, the normal condition signal is acquired, without any faults. There are two categories of faulty conditions: mechanical faults (bearing, misalignment, unbalance, impeller, and looseness) and hydraulic fault (cavitation). The faults were created and simulated in this study, intentionally as real ones, and instrumentations including accelerometer and data acquisition device (DAQ) with LabVIEW

have been verified with different samples and acquired signals were compared with the norms/standards of the frequency according to the previous works. An accelerometer mounted on its bearing housing is used for signals acquisition from the pump. This sensor transfers the vibrational data to the DAQ, where the signals have to be amplified and noise filtered out and then moved to a computer which is equipped with a digital/analog converter card (D/A), in order to convert the analog signals to digital. In this study, the data acquisition sampling rate is 16 kHz and 2.4 seconds as a sampling time with 38 400 number of samples. A LabVIEW software is used for capturing the signals, and the raw signals are saved in order to use them in the second stage for further processing. A speed of 20 Hz (1200 RPM) was considered in the data acquisition of the pump conditions. Moreover, such faults are found to be the

most common ones that occur with centrifugal pumps in industries (as per some references and interview with some engineers from industries). However, the main purpose was to test how AI classifiers work with different faults.

4 | FEATURE EXTRACTION

Feature extraction is very important as the main aim is to extract from the vibration signals, some characteristics that the neural network would implement. If there is good feature extraction and selection, there is bound to be weak classification performance in the system.[46] Reference47 recommended that the extracted features have to be strongly relevant to the machine faults. However, for vibration signals having strong noise that conceals important information, feature extraction becomes difficult. This drawback gave rise to the application of wavelet transform analysis to achieve noise cancellation for feature extraction.[48]

WPT was introduced by[49] and is a multi-stage filtering method that decomposes a signal into packets or levels of approximation which are denoted with A, and details coefficients which are denoted with D, as illustrated in Figure 5.[34,50,51] The WPT is defined as follows:

$$W_{(j,k)}(t) = 2^{\frac{j}{2}} w \left(2^j x - k \right) j \in Z \tag{1}$$

WPT is similar to DWT except WPT provides higher and finer decomposition tree, where both approximation (A) and detail (D) can produce pairs of packets (second level of approximation and detail), but DWT does not have such ability (ie, the next or second level of approximation and detail can be split by the approximation (A) only).[52] WPT has been applied for other types of rotating machinery..[46,51-55] In this work, WPT using two mother wavelets (db4 and rbio1.5) is applied for the preprocessing and feature extraction. Three cases are considered, and they are as follows:
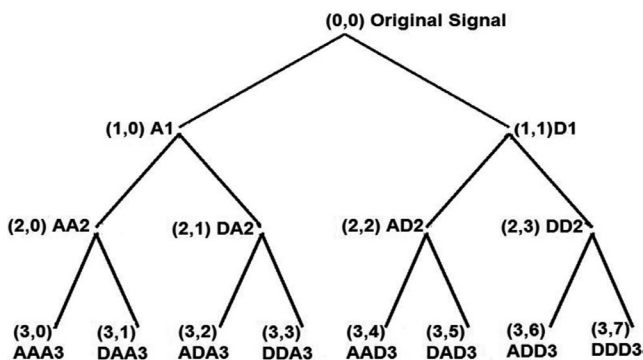

FIGURE 5 WPT tree decomposition to three levels: (A and D denote the approximation and detail, respectively)

4.1 | Case 1

There are 60 features of both approximation and detail used to train the MLP-BP, where the desired number of features (60) are completed considering 14 features from each segment (except the fifth segment, as the total number of 60 features are extracted and discarding the last 10 features from the fifth segment. It is also considered that the network is trained with the best three approximations per each segment, except the fifth segment (due to the intention of considering the required number of features only, discarding the unnecessary ones, where the first two approximations from the fifth segment are considered only to have in total, 14 features from each condition.

The decomposition of the signals is into 3 levels for the purpose of feature extraction, where the approximation and detail coefficients are extracted from 7 different pump cases. A signal length of 38 400 samples was recorded in each case. These signals were each divided into five segments, of length 7680 samples in order to have more divisions and segments extracted from the main one that will allow production of a number feature sets, which will then be used as inputs for the classifier. The five segments produced a total of 60 features. From these 60 features, 6 parameters (Kurtosis, RMS, Peak, Crest Factor, Shape Factor and Impulse Factor) are computed for the signal from each case.

Figure 5 shows the description of the WPT tree decomposition to three levels, where A denotes the approximation, and D refers to the detail. Figure 6 illustrates the third-level tree decomposition of imbalance condition using the db4 function, where the general sinusoidal pattern of the signals has better representation with the approximation decomposition, and it is also preserved in successive approximation levels (but not the detail levels).

It is also remarked that approximation reveals successively, less noisy signals by reducing the high-frequency information, which could be resulting in extracting better features than the ones from the detail. The best approximation decomposed signals (A1, AA2, and AAA3) are considered for this work, as they have successively less noise, and detail signals have been discarded due to the best selection from the approximation ones.

4.2 | Case 2

The signals are also decomposed to 6 levels with db4 only, and from 5 segments of each signal, the first 3 approximation packets of each level are selected. The total features per condition and parameter are 30.

4.3 | Case 3

The signals are finally again decomposed in 3 levels, and each signal is divided into 8 segments with a length of 4800
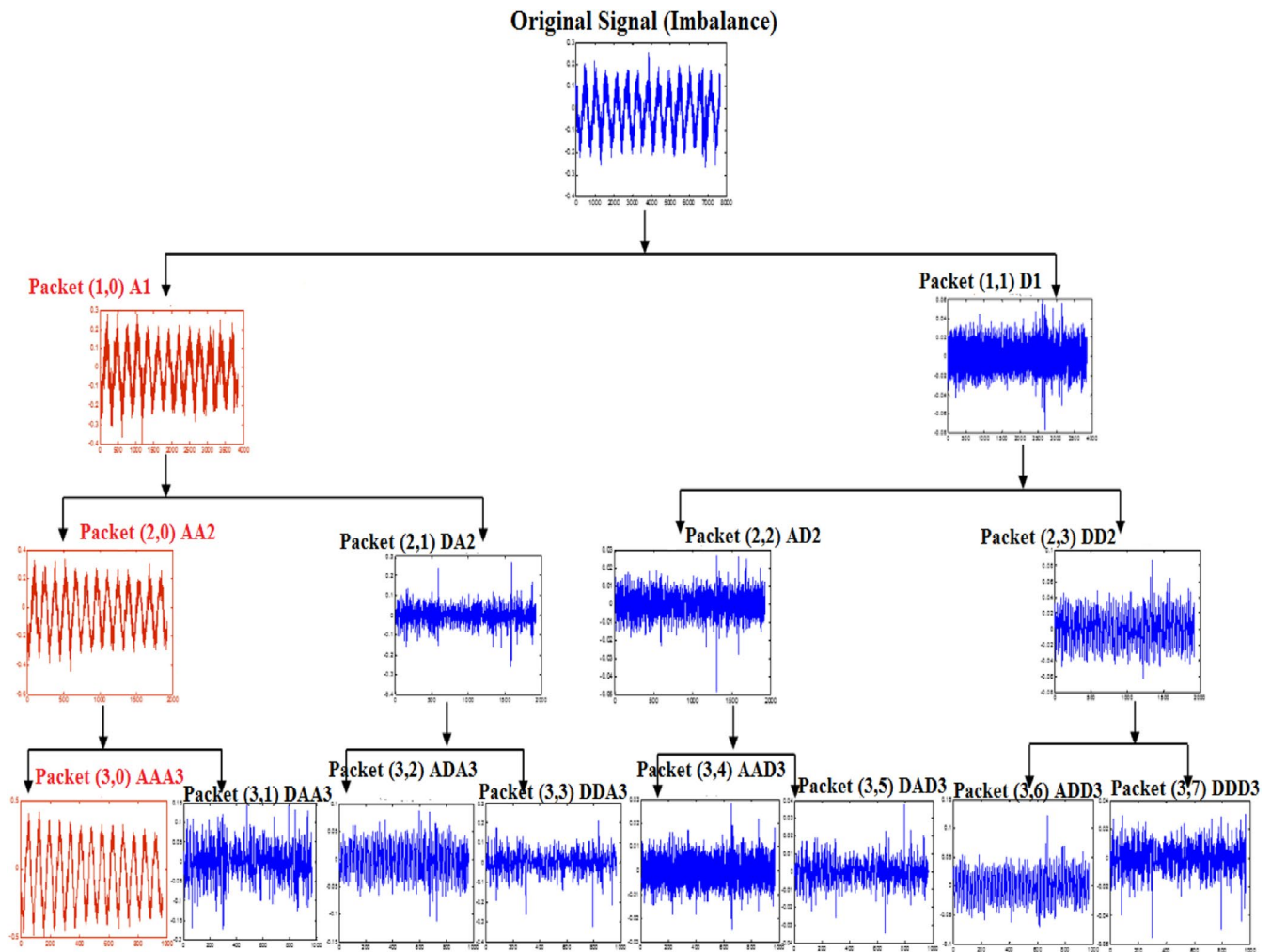
FIGURE 6 The third level tree decomposition of imbalance condition using db4 function

samples. The first 3 approximation packets of each level and each decomposed signal (segment) are selected. The total features per condition and parameter are 24. It is also considered that the network is trained with the best three approximations per each segment, except the fifth segment (due to the intention of considering the required number of features only while discarding the unnecessary ones, where the first two approximations from the fifth segment are considered only to have 14 features in total from each condition). These three cases are analyzed to determine the best number of features and types of coefficients for classification accuracy.

For the SVM, one case is considered, using 2 parameters and 14 features. The extracted features are normalized. The effectiveness (sensitivity) of each parameter is plotted in Figure 7, for all conditions. Normally, when healthy (blue +) is the lowest, it indicates the good effectiveness of the parameter. Therefore, peak and RMS are selected for SVM due to their ability in distributing and distinguishing the conditions effectively.

5 | CLASSIFICATION METHODS

In this study, the neural network classifier and SVM were fed based on the features extracted as input vectors. The MLP has input, hidden, and output layers. There are 6 neurons representing the features that are extracted and normalized, for each of the parameters, considering the WPT pre-processed data. GA was used to select and optimize the hidden layers numbers and neurons. There are 7 neurons in the output layer: tested pump condition is one; healthy condition is one neuron; and six different fault conditions for six neurons. The Levenberg-Marquardt (LM) function was used to train the network using algorithm based on back propagation, in order to update the weights and biases. As shown in section 4, there are three cases for the extracted features-based decomposition levels and number of signal segments, 60, 30, 24, and 14 features (60 and 14 features are used as normalized and non-normalized) per condition with a total of 420, 210, 168, and 98 input features for all
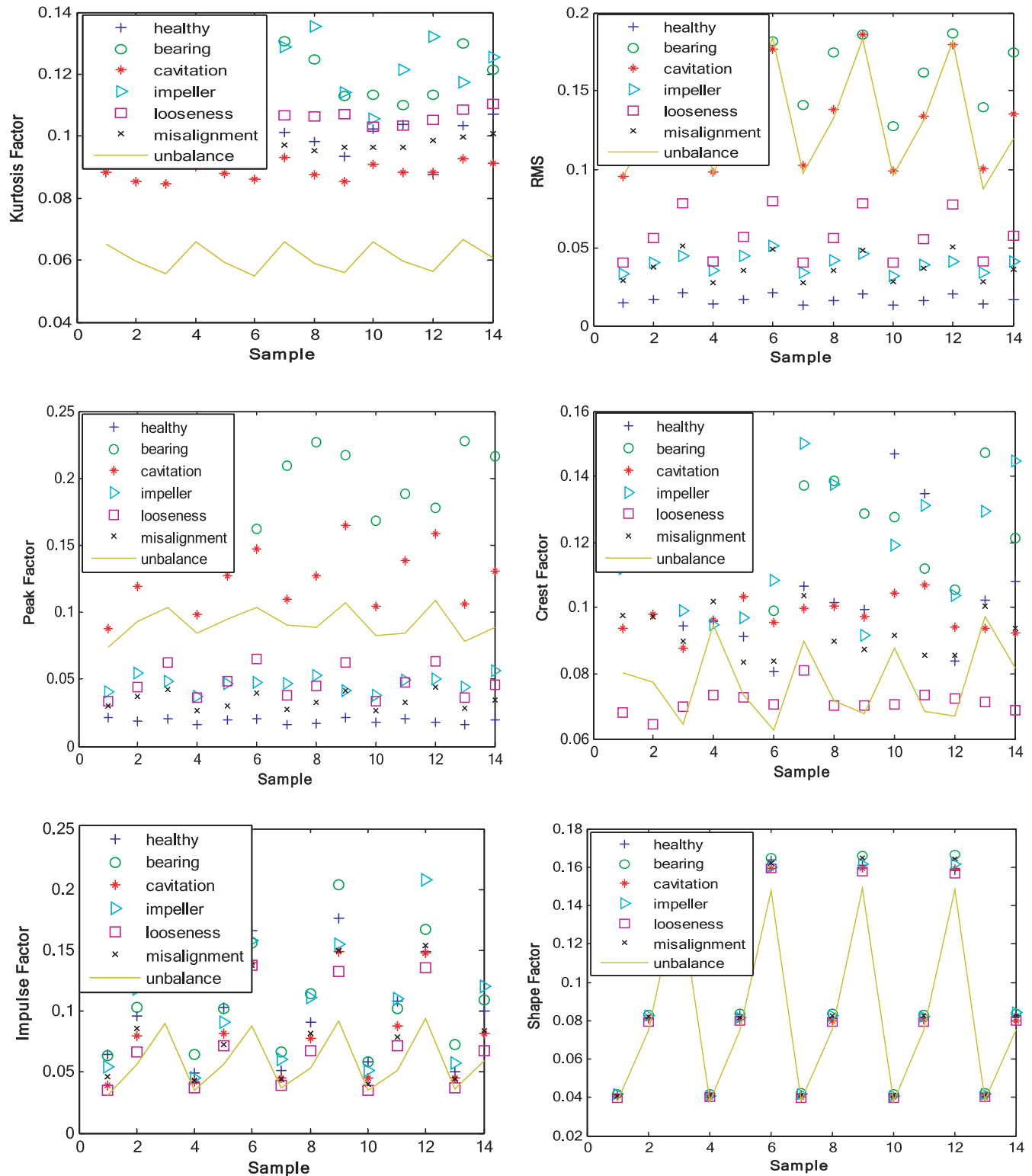
FIGURE 7 The effectiveness of each parameter against all conditions (db4 mother function)

conditions per parameter are forwarded to the MLP-ANN, which results in a matrix of size $[6 \times 420]$, $[6 \times 210]$, $[6 \times 168]$, and $[6 \times 98]$, respectively. There are three divisions of the datasets (70% is training, 15% is test, and 15% is validation). The Boolean matrix of size $[7 \times 420]$ (60

features), $[7 \times 210]$ (30 features), $[7 \times 168]$ (24 features), and $[7 \times 98]$ (14 features); is the training target with 7 conditions representing the rows: the digit represents the block size (1×60), (1×30), (1×24), and (1×14), respectively, given as follows:

1. Healthy	[1 0 0 0 0 0 0]
2. Bearing fault	[0 1 0 0 0 0 0]
3. Cavitation	[0 0 1 0 0 0 0]
4. Impeller fault	[0 0 0 1 0 0 0]
5. Misalignment	[0 0 0 0 1 0 0]
6. Looseness	[0 0 0 0 0 1 0]
7. Imbalance	[0 0 0 0 0 0 1]

Neural network details and structure are shown in Figure 8 and Table 1, respectively.

SVM classification has been applied using 14 features (normalized) which are representing the best three-level approximations. The results using the considered seven conditions were classified, tested, and compared with each other.

The polynomial kernel was used to investigate the SVM. The width of the parameter C is set to 3, and the input data sets are randomly selected and the training and test data-set divided. Both conditions used the RMS and peak parameters, because they are well distinguished based on the conditions.

6 | RESULTS AND DISCUSSION

The strengths and drawbacks of the three considered AI schemes, in relation to the effect possessed by the selection of the mother wavelet, using approximation detail features, and normalized or non-normalized features were investigated in this section.

6.1 | MLP-BP

There are four layers that are hidden, having [24 21 24 23] neurons. Four layers were considered after many selection trials for the best number of layers, and through GA, 4 layers were identified as the best option with the best classification outputs. In addition, the number of neurons in each section affects the generalization ability of the system, while the number of neurons and hidden layers affects the efficiency of the system. With larger number, there is a possibility of over-fitting the training data and weak generalization of new data. Therefore, some methods might be used to select the proper number of hidden layers and neurons such as genetic algorithm. The rates of classifications for all the cases that have db4 mother wavelet are based on GA. Using 14 approximation normalized and non-normalized features, presented overall classification rates of 100% and 98%, respectively. It is remarked that the test classification is successfully conducted for 6 cases out of 7. To avoid such misclassification, higher number of features are used. Classification rates of 75.5% using 60 normalized features, 71.2% using 60 non-normalized features, 97.6% using 30 normalized 6 level approximation features, and 100% using 24 normalized features were obtained. However, only 6 out of 7 cases are classified (test) with 14 approximation normalized features of 100, and 14 approximation non-normalized features of 98%.

Based on the results obtained using db4, classification rates with rbio1.5 mother functions are conducted using 14 and 24 approximation normalized features only, and classification rates are 100% for 6 out of 7 cases (validation) and
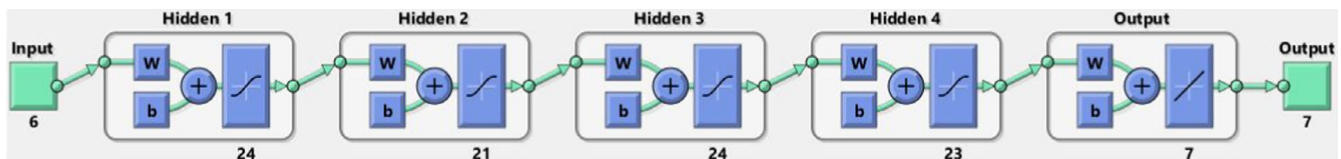

FIGURE 8 Network structure

TABLE 1 Multilayer feed-forward perceptron neural network architecture parameters

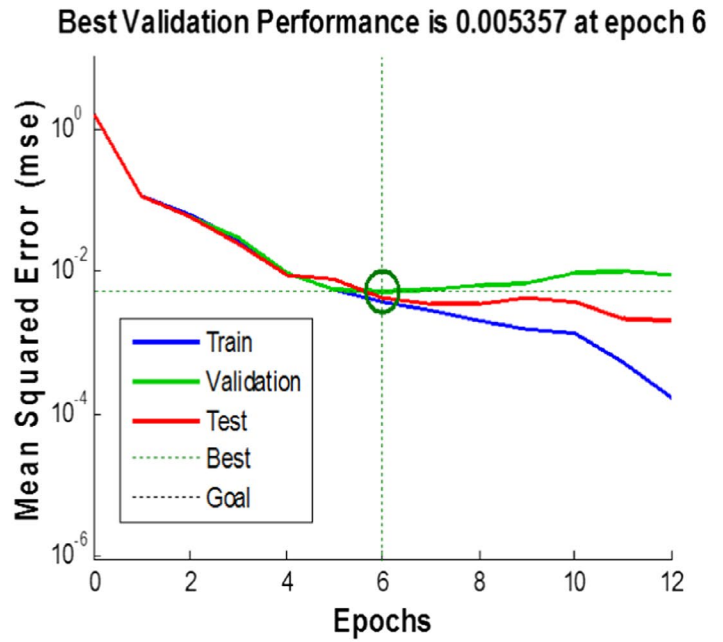| MLP-ANN details				
Transfer function			**No. of hidden layers and neurons (GA-based selection)**	**No. of output neurons**
Hidden layer	**Output layer**	**No. of input neurons**		
Sigmoid	Linear	6	4 layers: [24 21 24 23]	7
ANN training parameters				
Training algorithm	**Learning rate**		**Training stop criteria**	
LM (BP) & GA	0.56		Max. epoch	MSE
			1000	10E−20

100% for all 7 cases, respectively. Therefore, the best accuracy rate is achieved using 24 approximation normalized features; the overall confusion matrix for training, testing, and validation is illustrated in Figure 9. The overall classification rates are shown by the right lower square in blue color. From Figure 9, there are 0% incorrect classifications, and

(A) The overall confusion matrices for classification accuracy rate

(B) Performance of MLP-BP using 24 normalized features (rbio1.5 mother wavelet)

FIGURE 9 Validation of MLP-BP scheme

(A) db4 function

(B) rbio1.5 function

FIGURE 10 Overall classification accuracy rate using 24 normalized approximation features of GA-based hidden neuron and layer selection with MLP-GABP (WPT)

TABLE 2　Artificial Intelligence methods overall performance

Method	No. of features per condition and parameter	No. of hidden layers & neurons	Training Computational time (hh:mm:ss)	Test rate (%)	Validation rate (%)	Training rate (%)	Overall classification rate (%)	Remarks
SVM (rbio1.5)	14 approx. (normalized)	-	-	-	-	-	100	
SVM (db4)	14 approx. (normalized)	-	-	-	-	-	100	
MLP-BP with GA-based selection (db4)	24 approx. (normalized)	[24 21 24 23]	00:01:29	100	100	100	100	
MLP-BP with GA-based selection (rbio1.5)	24 approx. (normalized)	[24 21 24 23]	00:00:10	100	100	100	100	
MLP-BP with GA-based selection (db4)	14 approx. (normalized)	[24 21 24 23]	00:00:13	100	100	100	100	6 classes are classified (test)
MLP-BP with GA-based selection(rbio1.5)	14 approx. (normalized)	[24 21 24 23]	00:00:16	100	100	100	100	6 classes are classified (validation)
MLP-GABP with GA based selection (db4)	14 approx. (normalized)	[24 21 24 23]	MLP: 00:00:18, GA: 08:00:00	100	100	100	100	6 classes are classified (test)
MLP-GABP with GA-based selection (rbio1.5)	24 approx. (normalized)	[24 21 24 23]	MLP: 00:00:12, GA: 13:00:00	100	100	99.2	99.4	6 classes are classified (validation)
MLP-BP with GA-based selection (db4)	14 approx. (non-normalized)	[24 21 24 23]	00:00:10	100	86.7	100	98	6 classes are classified (test)
MLP-GABP with GA-based selection (rbio1.5)	14 approx. (normalized)	[24 21 24 23]	MLP: 00:00:24, GA: 15:00:00	100	100	97.1	98	6 classes are classified (test & validation)
MLP-BP with GA-based selection(db4)	30 approx. (normalized)	[24 21 24 23]	00:00:20	100	96.9	97.3	97.6	
MLP-GABP with GA-based selection (db4)	24 approx. (normalized)	[24 21 24 23]	MLP: 00:00:10, GA: 12:00:00	96	92	95.8	95.2	6 classes are classified (validation)
MLP-BP with GA-based selection (db4)	60 (normalized)	[24 21 24 23]	00:00:54	58.7	60.3	82.3	75.5	
MLP-BP with GA-based selection (db4)	60 (non-normalized)	[24 21 24 23]	00:00:43	63.5	57.1	75.9	71.2	

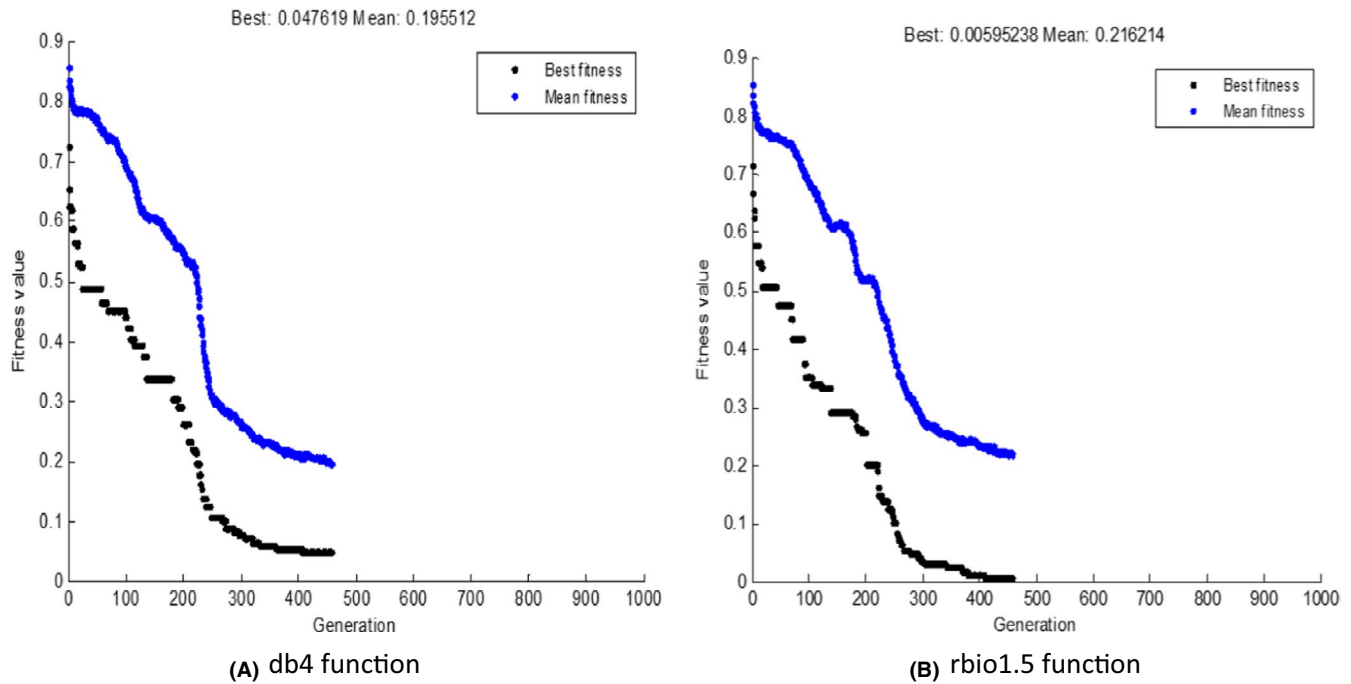**(A)** db4 function   **(B)** rbio1.5 function

FIGURE 11   Best score and mean score values vs generation based on MLP-GABP with WPT using 24 normalized approximate features

classifications in green color are 100%. Considering the individual accuracy rate, which is known as the pump condition, the classification rates are (healthy), known as case 1, having 100% accuracy rate, case 2 (bearing fault) with accuracy rate of 100%, case 3 (cavitation) has 100%, case 4 (impeller fault) with 100% accuracy, case 5 (misalignment) with 100% accuracy, case 6 (mechanical looseness) with 100% accuracy, and case 7 (imbalance) having 100% accuracy rate.

6.2 | MLP-GABP

Two cases of 14 and 24 approximations with db4 mother normalized features were implemented using MLP-GABP. The rbio1.5 shows poor performance compared with MLP-BP regarding time of computation and overall accuracy rates, and number of classified cases, where using 14 features, overall rates of 100% for db4 and 98% for rbio1.5, but only 6 cases out of the 7 cases are classified with test and validation classifications using rbio1.5, and with test using db4. Although using 24 normalized features, accuracy rates with rbio1.5 and db4 are presented an overall of 99.4% and 95.2%, respectively, but only 6 cases out of 7 are classified in validation for both wavelets as shown in Figure 10.

The layers containing [24 21 24 23] neurons are the four hidden layers, and they are used in MLP, considering GA selection and neural network weights were adjusted. In the case of using 14 features, GA-based optimization and training using db4 and rbio1.5 are terminated after 576 and 403 generations with best fitness functions of 0.020482 and 0, respectively. Whereas using 24 features, terminations with db4 and rbio1.5

are 457 and 548 and with best fitness functions of 0.047619 and 0.00595238, respectively, as shown in Figure 11. The best fitness function denotes the best minimized MSE.

6.3 | SVM

Classification accuracy rates using polynomial kernel function of the cases of db4 and rbio1.5 mother wavelets (approximation features) 14 normalized are 100% for both cases. Table 2 gives the overall performance of the AI methods employed in this work.

7 | CONCLUSION

In this study, pump conditions classification and feature extraction considering MLP-BP were conducted successfully for all 7 cases using WPT with normalized 60 and 30 features of 75.3% and 97.6%, respectively. It has been observed that using 6 levels approximations only (30 features) rather than using all decomposed 3 levels approximations and details (60 features) provided a better classification rate. On the other hand, with reduced number of features, MLP-BP has successfully achieved training, validation of 100% for all the seven cases, but in test, 100% of accuracy is achieved for the classification of 6 cases out of the seven cases. It can be remarked that MLP-BP can loss its ability in classifying all the cases using insufficient features. Therefore, the number of features has to be carefully selected and reduced. However, it has been found that using 3 more level approximations (24

features) achieved an overall accuracy rate of 100% by using both db4 and rbio1.5 for all 7 cases, which outperformed the 6 level approximations (30 features). It was observed from the study that in the SVM, both db4 and rbio1.5 have reduced feature numbers, (14 normalized features) resulting in an overall classification rate of 100%. When the features were extracted as approximations, enhanced classification rates were achieved than approximation and detail. The good selection of approximation features has a positive impact on the classification rate using all features. Both MLP and SVM gave better performance considering the features and fewer normalized parameters. GA has good optimization ability in hidden layers and neurons selection. This study showed that the use of 4 hidden layers having 24, 21, 24, and 23 neurons, gave the best performance. However, the drawback of GA is that the computational time is longer and in local minimum has the risk of been stuck. On the other hand, a combination of GA with BL MLP training-based, gave a good performance, but slightly lower compared with MLP-BP, using 14 approximate features of 100% and 98% as an overall rate with db4 and rbio1.5, respectively, based on 6 classified cases. In addition, MLP-GABP with 24 approximate features classified all 7 cases with an accuracy rate of 99.4% and 95.2%, using rbio1.5 and db4, respectively, but 6 cases only are classified with validation using both wavelets.

This study proved that the accuracy classification of MLP-BP can be achieved if the architecture of the neural network is optimized using GA, and a suitable mother wavelet for wavelet transform-based feature extraction. Furthermore, a good selection for the approximation features is achieved in this study with fewer number of features.

CONFLICT OF INTEREST
The authors declare that they have no conflict of interest.

ORCID
Maamar Al Tobi https://orcid.org/0000-0002-5520-5703
Kenneth Eloghene Okedu https://orcid.org/0000-0002-9635-1029

REFERENCES
1. Al-Tubi MAS, Al-Raheem KF. Rolling element bearing faults detection, a time domain analysis. *Caledonian J Eng*. 2010;6:1–10.
2. Al-Tubi MAS, Al-Raheem KF, Abdul-Karem W. Rolling element bearing element faults detection, power spectrum and envelope analysis. In: *International Conference on Applications and Design in Mechanical Engineering. Penang, Malaysia*; 2012.
3. Mehala N, Dahiya R. A comparative study of FFT, STFT and wavelet techniques for induction machine fault diagnostic analysis. In: *International Conference of Computational Intelligence, Man-machine Systems Cybernetics (CIMMACS '08). India*; 2008.
4. Peng ZK, Chu FL. Application of the wavelet transform in machine condition monitoring and fault diagnostics: a review with bibliography. *Mech Syst Sign Process*. 2004;18:199-221.

5. Al-Tobi M, Al-Raheem KF. Rolling element bearing faults detection, wavelet de-noising analysis. *Univ J Mech Eng*. 2015;5:47-51.
6. Elsken T, Metzen JH, Hutter F. Neural architecture search: a survey. *J Mach Learn Res*. 2019;20(2019):1-21.
7. van Stein B, Wang H, Thomas B. Neural network design: learning from neural architecture search. Preprint; 2020.
8. Farokhzad S. Vibration based fault detection of centrifugal pump by fast Fourier transform and adaptive neuro-fuzzy inference system. *J Mech Eng Technol*. 2013;1:82-87.
9. Sakthivel NR, Nair BB, Elangovan M, Sugumaran V, Saravanmurugan S. Comparison of dimensionality reduction techniques for the fault diagnosis of mono block centrifugal pump using vibration signals. *Eng Sci Technol, Int J*. 2014;17:30-38.
10. Al TM, Bevan G, Wallace P, Harrison D, Okedu KE. Faults diagnosis of a centrifugal pump using multilayer perceptron genetic algorithm back propagation and support vector machine with discrete wavelet transform-based feature extraction. *Comput Intellig*. 2020;37(3):1-26.
11. Sui K, Kim H. Research on application of multimedia image processing technology based on wavelet transform. *J Image Video Proc*. 2019;2019:24. https://doi.org/10.1186/s13640-018-0396-1
12. Yanjun L, Fanlong M, Yibo L. Research on rub impact fault diagnosis method of rotating machinery based on wavelet packet and support vector machine. In: *International Conference on Measuring Technology and Mechatronics Automation. Changchun, IEEE*; 2009.
13. Samanta B. Gear fault detection using artificial neural networks and support vector machines with genetic algorithms. *Mech Syst Sign Process*. 2004;18:625-644.
14. Yang Z, Hoi WI, Zhong J. Gearbox fault diagnosis based on artificial neural network and genetic algorithms. In: *International Conference on System Science and Engineering. Macau, China*; 2011.
15. Al-Raheem KF, Abdul-Karem W. Rolling bearing fault diagnostics using artificial neural networks based on laplace wavelet analysis. *Int J Eng Sci Technol*. 2010;2:278-290.
16. Altobi MAS, Bevan G, Wallace P, Harrison D, Ramachandran KP. *Faults Diagnosis of a Centrifugal Pump Using MLP-GABP and SVM with CWT Based Feature Extraction*; 2017.
17. Altobi MAS, Bevan G, Wallace P, Harrison D, Ramachandran KP. *Faults Diagnosis of a Centrifugal Pump Using MLP-GABP and SVM with DWT Based Feature Extraction*; 2017.
18. ALTobi MAS, Bevan G, Wallace P, Harrison D, Ramachandran KP. Fault diagnosis of a centrifugal pump using MLP-GABP and SVM with CWT. *Eng Sci Technol, Int J*. 2019;22(3):854-861. https://doi.org/10.1016/j.jestch.2019.01.005
19. Dai J, Tang J, Huang S, Wang Y. Signal-based intelligent hydraulic fault diagnosis methods: review and prospects. *Chin J Mech Eng*. 2019;32:75. https://doi.org/10.1186/s10033-019-0388-9
20. Deng W, Xu J, Gao XZ, Zhao H. An enhanced MSIQDE algorithm with novel multiple strategies for global optimization problems. *IEEE Trans Syst Man Cybernet: Syst*. 2020:1–10. https://ieeexplore.ieee.org/abstract/document/9247966
21. Song Y, Wu D, Wagdy Mohamed A, Zhou X, Zhang B, Deng W. *Enhanced Success History Adaptive DE for Parameter Optimization of Photovoltaic Models*. Complexity; 2021.
22. Charniak E, McDermott D. *Introduction to Artificial Intelligence*. Boston, USA: Addison Wesley Longman Inc.; 2000.
23. McCulloch WS, Pitts W. A logical calculus of the ideas immanent in nervous activity. *Bull Math Biophys*. 1943;5:115-133.

24. Zadeh LA. Fuzzy sets. *Inf Control*. 1965;8:338-353.

25. Nasiri MR, Mahjoob MJ, Vahid-Alizadeh H. Vibration signature analysis for detecting cavitation in centrifugal pump using neural networks. In: *IEEE International Conference on Mechatronics (ICM). Istanbul, Turkey, IEEE*; 2011.

26. Farokhzad S, Ahmadi H, Jafary A. Fault classification of centrifugal water pump based on decision tree and regression model. *J Sci Today's World*. 2013;2:170-176.

27. Farokhzad S, Ahmadi H, Jaefari A. Artificial neural network based classification of faults in centrifugal water pump. *J Vibroeng*. 2012;14:1734–1744.

28. Liu L-S, Yang Y-H, Li Z-Y, Yu W. Condition monitoring for helicopter main gearbox based on wavelet packet transform and wavelet neural network. In: *International Conference on Quality, Reliability, Maintenance, and Safety Engineering. Xi'an, IEEE*; 2011.

29. Muralidharan V, Sugumaran V. Selection of discrete wavelets for fault diagnosis of monoblock centrifugal pump using the j48 algorithm. *Appl Artif Intellig*. 2013;27:1-19.

30. Wang H, Chen P. Intelligent diagnosis method for a centrifugal pump using features of vibration signals. *Neural Comput Appl*. 2009;18:397-405.

31. Wang HQ, Chen P. Fault diagnosis of centrifugal pump using symptom parameters in frequency domain. *Agric Eng Int: CIGR Ejournal*. 2007;IX:1–13.

32. Muralidharan V, Sugumaran V, Shanmugam P, Sivanathan K. Artifical neural network based classification for monoblock centrifugal pump using wavelet analysis. *Int J Mech Eng*. 2010;1:28-37.

33. Iiott PW, Griffiths AJ. Fault diagnosis of pumping machinery using artificial neural networks. *J Process Mech Eng*. 1997;211:185-194.

34. Zouari R, Sieg-Zieba S, Sidahmed M. Fault detection system for centrifugal pumps using neural networks and neuro-fuzzy techniques. *SURVEILLANCE 5 CETIM Senlis*; 2004.

35. Nasiri MR, Mahjoob MJ, Vahid-Alizadeh H. Vibration signature analysis for detecting cavitation in centrifugal pump using neural networks. In: *IEEE International Conference on Mechatronics (ICM). Istanbul, Turkey, IEEE*; 2011.

36. Zhang Z, Wang Y, Wang K. Intelligent fault diagnosis and prognosis approach for rotating machinery integrating wavelet transform, principal component analysis, and artificial neural networks. *Int J Adv Manuf Technol*. 2013;68:763-773.

37. Jedlinski Ł, Jonak J. Early fault detection in gearboxes based on support vector machines and multilayer perceptron with a continuous wavelet transform. *Appl Soft Comput*. 2015;30:636-641.

38. Wang DY, Zhang WZ, Zhang JG. Fault bearing identification based on wavelet packet transform technique and artificial neural network. In: *International Conference on System Science, Engineering Design and Manufacturing Information, Yichang, IEEE*; 2010.

39. Sui W-T, Zhang D. Rolling element bearings fault classification based on SVM and feature evaluation. In: *Proceedings of the Eighth International Conference on Machine Learning and Cybernetics. Baoding, IEEE*; 2009.

40. Liu Y, Liu T. Rotating machinery fault diagnosis based on support vector machine. In: *International Conference on Intelligent Computing and Cognitive Informatics. Kuala Lumpur, IEEE*; 2010.

41. Samanta B, Al-Balushi KR, Al-Araimi SA. Artificial neural networks and genetic algorithm for bearing fault detection. *Soft Comput*. 2006;10:264-271.

42. Rumelhart DE, Hinton GE, William RJ. Learning representations by back-propagating errors. *Nature*. 1986;323:533-536.

43. Cortes C, Vapnik V. Support-vector networks. *Mach Learn*. 1995;20:273-297.

44. Holland JH. *Adaptive in Natural and Artificial Systems*. Ann Arbor: The University of Michigan Press; 1975.

45. Rooij AJFV, Jain LC, Johnson RP. *Neural Network Training Using Genetic Algorithm*. Singapore: World Scientific Publishing Co. Pte. Ltd.; 1996.

46. Rafiee J, Tse PW, Harifi A, Sadeghi MH. A novel technique for selecting mother wavelet function using an intelligent fault diagnosis system. *Expert Syst Appl*. 2009;36:4862-4875.

47. Saberi M, Azadeh A, Nourmohammadzadeh A, Pazhoheshfar P. Comparing performance and robustness of SVM and ANN for fault diagnosis in a centrifugal pump. In: *19th International Congress on Modelling and Simulation. Perth, Australia*; 2011.

48. Alsberg BK, Woodward AM, Kell DB. An introduction to wavelet transforms for chemometricians: a time-frequency approach. *Chemomet Intellig Lab Syst*. 1997;37:215-239.

49. Coifman RR, Meyer Y, Wickerhauser V. Size properties of wavelet packets. In: Ruskai MB, editor. *Wavelets and Their Applications*. Boston: Jones and Bartlett; 1992.

50. Sakthivel NR, Nair BB, Sugumaran V. Soft computing approach to fault diagnosis of centrifugal pump. *Appl Soft Comput*. 2012;12:1574-1581.

51. Bendjama H, Gherfi K, Idiou D, Boucherit MS. Condition monitoring of rotating machinery by vibration signal processing methods. In: *International Conference on Industrial Engineering and Manufacturing. Batna University Algeria*; 2014.

52. Al Tobi M, Bevan G, Wallace P, Harrison D, Okedu KE. Faults diagnosis of a centrifugal pump using multilayer perceptron genetic algorithm back propagation and support vector machine with discrete wavelet transform-based feature extraction. *Comput Intellig*. 2021;37(1):21-46.

53. Zhong J, Yang Z, Wong SF. Machine condition monitoring and fault diagnosis based on support vector machine. In: *International Conference of Industrial Engineering Management. Macao, IEEE*; 2010.

54. Wuming H, Peiliang W, Qiangguo Y. Fault diagnosis of traction machine for lifts based on wavelet packet algorithm and RBF neural network. In: *International Conference on information Sciences. Chengdu, IEEE*; 2010.

55. Chen C-H, Shyu R-J, Ma C-K. Rotating machinery diagnosis using wavelet packets-fractal technology and neural networks. *J Mech Sci Technol*. 2007;21:1058-1065.