Sound Signal Based Fault Classification System in Motorcycles Using Hybrid Feature Sets and Extreme Learning Machine Classifiers

T. Jayasree[1,*] and R. Prem Ananth[2]

[1]Department of Electronics & Communication Engineering, Government College of Engineering, Tirunelveli, Tamil Nadu, India
[2]Department of Electronics & Communication Engineering, Francis Xavier Engineering College, Tirunelveli, Tamil Nadu, India
*Corresponding Author: T. Jayasree. Email: jayasree@gcetly.ac.in

Abstract: Vehicles generate dissimilar sound patterns under different working environments. These generated sound patterns signify the condition of the engines, which in turn is used for diagnosing various faults. In this paper, the sound signals produced by motorcycles are analyzed to locate various faults. The important attributes are extracted from the generated sound signals based on time, frequency and wavelet domains which clearly describe the statistical behavior of the signals. Further, various types of faults are classified using the Extreme Learning Machine (ELM) classifier from the extracted features. Moreover, the improved classification performance is obtained by the combination of feature sets in different domains. The simulation results clearly demonstrate that the proposed hybrid feature set together with the ELM classifier gives more promising results with higher classification accuracy when compared with the other conventional methods.

Keywords: Extreme learning machine; wavelet transform; feature; audio

1 Introduction

Asia has the world's most astounding centralization of mechanized two and three-wheeled vehicles. Motorcycles are the most favored means of travel frequently used by sparingly middle-class citizens due to their inexpensiveness, mileage, street conditions, and less demanding mobility. Motorcycle acquires the automobile market above 75% virtually of the total vehicle sales in the Asian countries. The Engine is the heart of automobiles and may get a complaint if the rider doesn't make continuous checking of the engine and other particles. Viewing from the side of engine complaints, the sound of the engine would provide major information about the engine problem [1]. Vehicles of a particular type, in various working environments, produce different sound patterns. Each sound pattern is regarded as an aural signature. Sounds of moving vehicles afford hints of their behaviors such as possible faults make and concerts of subsystems. In such circumstances, stalking of faults manually befits difficult and automatic acoustic investigation assists easy monitoring of assured conditions of the vehicles and future significances [2]. The problematic sound might be produced due to the hydraulic tap fit failures, float needle problem, high acceleration problem, air screw problem, alteration of silencer, air filter problem, etc. Due to the rapid development of transportation technology, fault detection and diagnostics in the motorcycles based on the

audio signatures is one of the major issues in the day to day life. Fault diagnostic can be done once the fault is detected properly.

Anami et al. [3] employed an acoustic signal-based recognition technique for the detection of faults in motorcycles using time domain-based features based and ANN classifiers. But the training time required for such type of classifier was found to be more. Paulraj et al. [4] adopted the autoregressive modeling algorithm for extracting features from the vehicle sound signals. But the authors focused only on finding the distance between the vehicles and the type of moving vehicle. Pan et al. [5] introduced a fault diagnosis method using the combination of entropy and mutual information as features derived from the vibration signals. But all the time domain-based features do not provide proper information, because most of the data might be embedded in the frequency content of the signal. For feature extraction, frequency domain-based signal processing techniques such as Discrete Fourier Transform (DFT) and short Time Fourier Transform (STFT) is used for detecting faults in the automobiles [6]. Anami et al. [7] adopted spectral based techniques by exploiting the variations in spectral behavior together with the chain code of the pseudo spectrum of the sound signal. Zhou et al. [8] adopted Fourier Transform based Mel-Frequency Cepstral Coefficients (MFCC) for recognizing the faults. Since the audio signals are non-stationary in nature, the Fourier transform based methods do not provide appropriate results because of their fixed window size. However, DFT and STFT have been realistic only to stationary signals. The wavelet Transform (WT) is localized in both time and frequency domain whereas the Fourier Transform (FT) is localized only in the frequency domain. Generally, by using the Wavelet Transform the frequency components can be separated based on the number of levels. Then number of sub bands is directly proportional to the number of levels of decomposition. In the wavelet Transform, various frequency components can be separated and stored and further processed for the classification purpose. Wavelet domain with various sub bands can produce detailed frequency information when compared to the Fourier Transform based analysis. Moreover, Wavelet Transform is well suited for the analysis of non-stationary signals, whereas, Fourier Transform is suitable only for the analysis of stationary signals. Since most of the signals such as speech, sound, music biomedical signals etc. are non-stationary in nature, WT based methods are used for the analysis of such type of signals. Furthermore, the local behavior of the signal can be studied well using Wavelet Transform when compared to Fourier Transform [10].

Wavelet Transform (WT) with variable window size is used for obtaining features from the non-stationary audio signals. Because of their inherent multi-resolution nature, wavelet coding schemes are especially suitable for applications where scalability and tolerable degradation are important. Wavelet Transform decomposes the signal into a set of basic functions [9]. These basis functions are called wavelets. Wavelets are obtained from a single prototype wavelet called mother wavelet by dilations and shifting. WT is computed separately for different segments of the time domain signal at different frequencies. But in WT based methods, only low frequency components called approximations are considered for analysis. This may cause problems, since important information may be located in the high frequency components. The aforementioned problem can be overcome by the application of Wavelet Packet Transform (WPT), in which the signal under consideration is divided into both low frequency components called approximations and high frequency components called details and both are considered for analysis. Wang et al. [10] employed the Wavelet Packet Transform (WPT) based approach in which distribution of energies in the first five sub bands of wavelet packet decomposition levels are used as features. A two stage Artificial Neural Network (ANN) classifier is deployed for the recognition of faults. But the computation time required was more and the classification rate obtained was less. Berredjem et al. [11] used wavelet packet coefficients to extract features from faulty bearings. Huseyin et al. [12] adopted wavelet packet analysis to extract features from the sound signals recorded from the cars and used multilayer perceptron classifiers for classification. Despite the advantages of the feature sets in all the three domains, the classification accuracy attained was still less, which can be further improved by

using the combination of feature sets in the three domains. This work mainly focuses on the application of hybrid feature sets.

Classification is one of the major tasks in fault identification in the vehicles. Many of the authors suggested Feed Forward Neural Network (FFNN) classifiers. Nevertheless, such type of classifiers produced less classification rate and Mean Square Error (MSE). ELM is a comparatively novel learning technique used for fault classification. Certainly, ELM models are highly adaptable and are much faster to train than standard neural network models. The main idea behind ELM is the projection of the original input into a high dimensional feature space. The uniqueness is that this new space is entirely fixed before observing the data. Thus, the actual learning consists of a simple linear regression that can be calculated proficiently in closed form. In this paper, the classification of faults in motorcycles based on acoustic signals using hybrid feature sets and ELM has been addressed. The rest of the paper is organized as follows: Section 2 discusses the proposed methodology. Feature extractions in three domains are discussed in Section 3. Results and discussions are discussed in Section 4 followed by conclusion Section 5.

2 Proposed Methodology

The objective of this study is to classify different faults in motorcycles based on the sound signals. The sound signals generated by the vehicles in practice are in the time domain. But most distinguished information may be hidden in the frequency content of the signal. The proposed methodology for fault detection incorporates the advantages of both times, frequency and wavelet domain-based features. Fig. 1 illustrates the block diagram of the proposed framework. The important stages include signal preprocessing, feature extraction and classification.

Initially, the signals undergo preprocessing which includes low pass filtering and normalization. Low pass filter is used for limiting the signal values below a predefined level and normalization is done for bringing all the signal values to an offset value. The appropriate features are extracted from the preprocessed signal. The feature vectors reflect the statistical characteristics of the signal in time, frequency and wavelet domains. The Extreme Learning Machine (ELM) is used for classifying the faults. The classification performance is analyzed by using a different combination of feature sets. The forthcoming sections describe the detailed description of each stage.

3 Feature Extraction

The extraction of features from input signals is vital to resolve the accuracy of classification. The usage of all coefficients of the input signal into the classifier may increase the computational load and decrease the classification rate. Thus, it is necessary to reduce the size of the input signal by extracting the relevant characteristics of the signal.

3.1 Time Domain-Based Features

The time-domain features give the basic information of data that is stored in a one-dimensional format [13]. The important statistical features extracted from the audio signals are explained below:

(a) Mean (M): Mean is the arithmetic average of the signal $x(n)$ and is computed by adding all the values in the signal and dividing it by the total number of values in the signal.

$$M = \frac{\sum x(n)}{n} \tag{1}$$

where n is the number of samples.

(b) Variance (V): Variance of a signal uses the power of the signal as a feature. Generally, the variance is the mean value of the square of the deviation of that variable. However, the mean of the signal is close to zero [18].
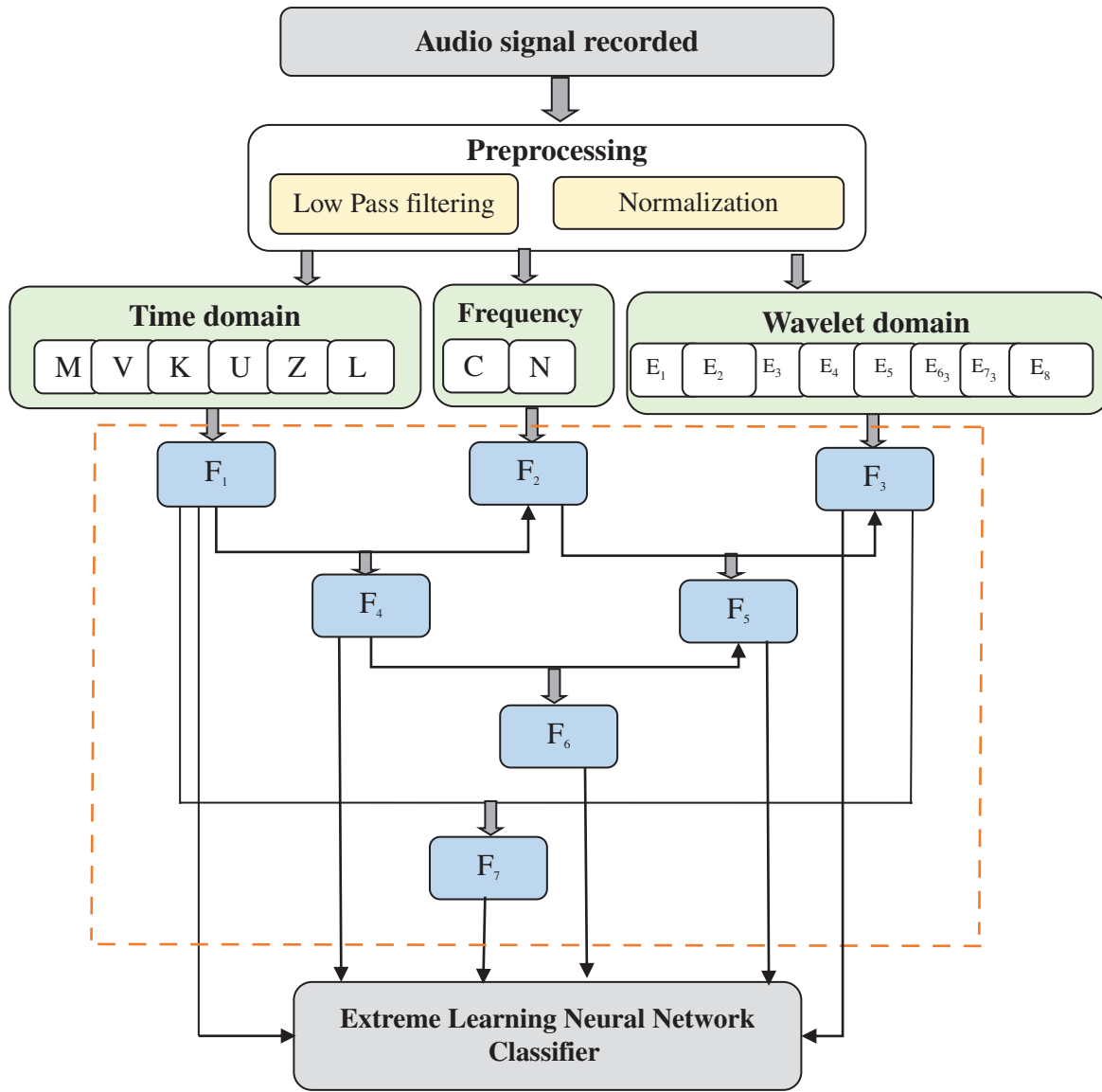
Figure 1: Acoustic signal-based fault classification system

$$V = \frac{\sum \left| x(n) - \overline{x(n)} \right|^2}{n - 1} \tag{2}$$

The variance represents the power of the variation of the standard deviation from the mean.

(c) Kurtosis (K): Kurtosis is a statistical measure that describes how profoundly the tails of distribution vary from the tails of a normal distribution. Fig. 2 shows the curves for Normal distribution and kurtosis.

The kurtosis of a distribution is defined as

$$K = \frac{E(x(n) - \mu)^4}{\sigma^4} \tag{3}$$

where μ is the mean of $x(n)$, σ is the standard deviation of $x(n)$, and $E(n)$ represents the expected value of the quantity n.
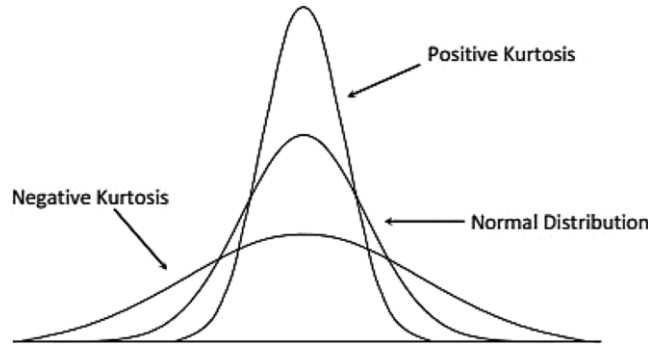
Figure 2: Normal distribution and kurtosis

(d) Entropy (U): The average rate at which information is produced by a stochastic source of data is called Entropy [14]. Entropy related with each promising data value is the negative logarithm of the probability mass function for the value:

$$U = -\sum (\text{p.} * \log 2(\text{p}))$$

(4)

where p contains the normalized sample number counts.

There are some features that describe the information about the structure of the envelope of the signal. They are zero Crossing (Z) and waveform length(L).

(e) Zero-Crossing (Z): A zero-crossing is a point where the sign of a function changes (e.g., from positive to negative), represented by an intercept of the axis (zero value) in the graph of the function. Thus, the zero-crossing rate is the rate of sign-changes along with a signal, i.e., the rate at which the signal changes from positive to negative or vice–versa [15].

$$Z = \frac{1}{2(M-1)} \sum_{m=1}^{M} |sign(x(m)) - sign(x(m-1))|$$

(5)

M is the total number of samples in a processing window. x(m) is the value of the m$^{\text{th}}$ sample.

(f) Waveform Length (L): It is the cumulative length of the waveform over the segment. The resultant values indicate a measure of waveform amplitude, frequency, and duration all within a single parameter [12].

$$L = \sum_{i=1}^{N-1} |x_{i+1} - x_i|$$

(6)

where x_i is the value of each part of the segment k, and N is the length of the segment N.

3.2 Frequencydomain Based Features

The property of the signal can also be described from the frequency content of the signal. The visual representation of the spectrum of frequencies of the signal can be analyzed from the audio spectrogram. Spectrograms are occasionally called as sonographs [16]. The estimate of the spectrogram for the particular fault is shown in Fig. 3. The important features extracted from the spectrogram are Mean frequency and Median frequency

(a) Mean Frequency (C): The mean frequency of an audio spectrum is calculated as the sum of the product of the audio spectrogram intensity (in dB) and the frequency divided by the total sum of audio spectrogram intensity. It can be expressed as follows:
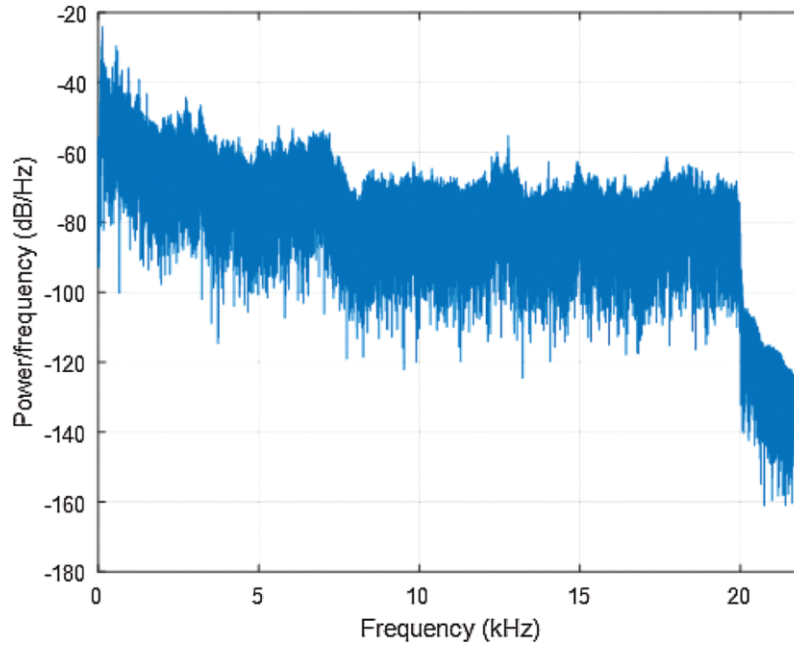
Figure 3: Spectrogram of the audio signal for a particular fault

$$C = \frac{\sum_{i=0}^{n} I_i f_i}{\sum_{i=0}^{n} I_i} \tag{7}$$

n = number of frequency bins in the audio spectrum

f_i = frequency of audio spectrum at bin i of n

I_i = Intensity (dB scale) of the audio spectrum at bin i of n

   (b) Median Frequency (N): It is a frequency at which the audio spectrum is divided into two regions with an equal integrated power [19]. It can be expressed as

$$\sum_{j=1}^{M} P_j = \sum_{j=MDF}^{M} P_j = \frac{1}{2} \sum_{j=1}^{M} P_j \tag{8}$$

   Figures 4a and 4b show the Mean frequency and median frequency estimates of the signal.

3.3 Wavelet Packet Transform (WPT)

   Wavelet Packet Transform provides a multi-resolution and sparse representation of the audio signal. The wavelet packets are more sensitive to small changes in the audio signal. In the DWT, each level of decomposition is evaluated by passing only the proceeding wavelet approximation coefficients through discrete-time low and high pass quadrature mirror filters. Nevertheless, in the wavelet packet decomposition, both the detail and approximation coefficients are decomposed for creating the full binary tree. For n levels of decomposition, 2^n sets of coefficients are produced in WPT [17]. The energy values distributed in the approximations and detailed coefficients of the wavelet packet sub bands are revealed as features. Fig. 5 shows the Wavelet packet decomposition structure.

   Wavelet packets are precise linear combinations of wavelets. They form bases that preserve many smoothness, and localization properties of the orthogonality, of their parent wavelets [18]. The packet is given by the following equation
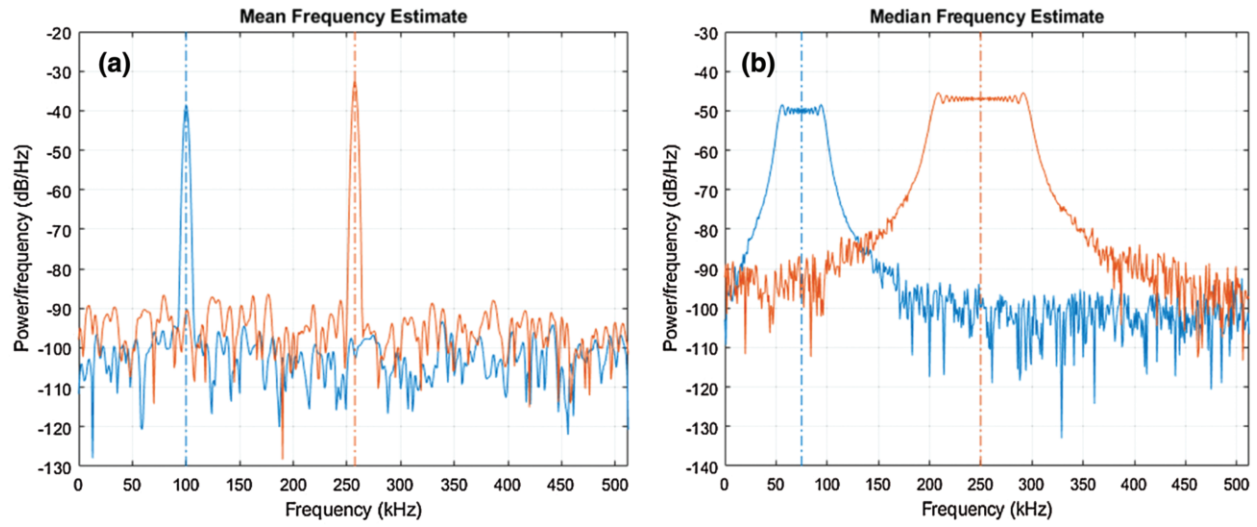
Figure 4: (a) Mean Frequency estimate (b) Median frequency estimate
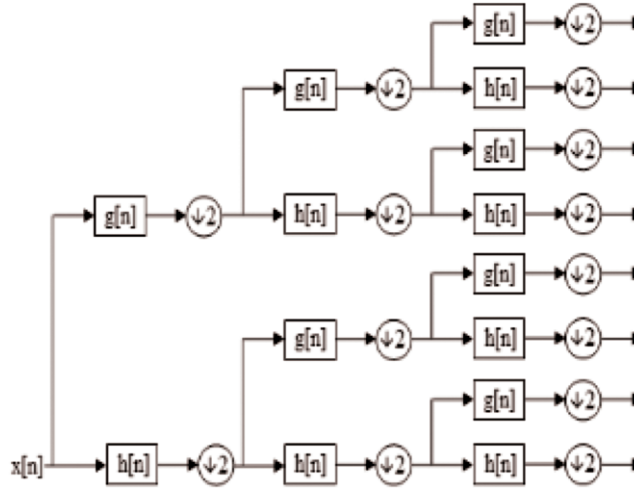
Figure 5: Wavelet packet decomposition structure for 3 levels

$$U_{2n}^{(j)}(t) = \sqrt{2} \sum_{k} h(k) U_n^{(j)}(2t - k) \tag{9}$$

$$U_{2n+1}^{(j)}(t) = \sqrt{2} \sum_{k} g(k) U_n^{(j)}(2t - k) \tag{10}$$

where $n = 0, 1, 2, \ldots\ldots$ and $k = 0, 1, 2, \ldots\ldots, m$

The scaling function is $\phi(t) = u_0^{(0)}(t)$ and the wavelet function is $\Psi(t) = u_1^{(0)}(t)$

3.4 Hybrid Feature Sets

The feature combinations in various domains (time, frequency and wavelet) have been introduced which make use of the combined characteristics of all the feature sets. Different combinations of feature sets are given below. Here F4, F5, F6 and F7 represent a hybrid set of features. The feature vectors are

$$F_1 = [M, V, K, U, Z, L]$$

$$F_2 = [C, N]$$

$$F_3 = [E_1, E_2, E_3, E_4, E_5, E_6, E_7, E_8]$$

$$F_4 = [M, V, K, U, Z, L, C, N]$$

$$F_5 = [C, N.E_1, E_2, E_3, E_4, E_5, E_6, E_7, E_8]$$

$$F_6 = [M, V, K, U, Z, L, E_1, E_2, E_3, E_4, E_5, E_6, E_7, E_8]$$

$$F_7 = [M, V, K, U, Z, L, C, N, E_1, E_2, E_3, E_4, E_5, E_6, E_7, E_8]$$

Algorithm 1 briefs the process of feature extraction

Algorithm 1: Feature extraction

Input: Load the input audio signal

Output: Total feature set to be classified

Feature Extraction:

1: for i= 1 to N do

2: Calculate the feature sets {F1}, {F2}, {F3} using equations

3: Form the hybrid feature sets {F4, F5, F6, F7}

5: end

*6: Compose the total feature vector*

---

4 Extreme Learning Machine Based Classifier

ELM is the single hidden layer Feed Forward Neural network (FFNN) used for a broad range of non-linear applications. The ELM structure comprising of three layers (input, hidden, output) is shown in Fig. 6. In this structure, the weights and biases of the input layer are arbitrary while only the outputs are calculated using the ELM algorithm. The learning time required for ELM is extremely less. Furthermore, the ELM structure has precise generalization aptitude compared to the FFNN based conventional learning algorithm [19]. AN ELM is a model of the form

$$f(x) = \sum_{i=1}^{L} h_i(x)\beta_i = h^T(x)\beta \tag{11}$$

where $h(x) = [h_1(x), \ldots\ldots h_L(x)]^T$ is referred to as an ELM feature vector, and β denotes the vector of expansion coefficients. Eq. (11) represents a network with two layers, where the input is projected first into a M dimensional space, in which the linear combination is accomplished.

The basic ELM frame consists of M number of hidden nodes and operates with $g(x)$ activation function. The other important characteristics of ELM are the randomly chosen input weights and hidden unit biases
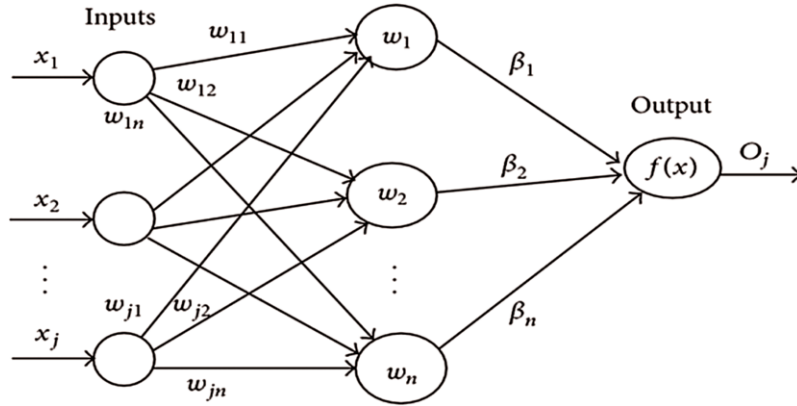
Figure 6: Structure of ELM

[20, 21]. Furthermore, the weights and biases are adjusted and the network is trained by finding the least square solutions which analytically determine output weights. The hidden layer parameters are related by the following equation.

$$H\beta = T, \tag{12}$$

$$\begin{bmatrix} g(w_1 \cdot X_1 + b_1) & \cdots & g\left(w_{\widetilde{N}} \cdot X_1 + b_{\widetilde{N}}\right) \\ \vdots & \cdots & \vdots \\ g(w_N \cdot X_N + b_N) & \cdots & g\left(w_{\widetilde{N}} \cdot X_N + b_{\widetilde{N}}\right) \end{bmatrix}_{N \times \widetilde{N}} \tag{13}$$

$$\beta = \begin{bmatrix} \beta_1^T \\ \vdots \\ \beta_{\widetilde{N}}^T \end{bmatrix}_{\widetilde{N} \times M} \quad \text{And } T = \begin{bmatrix} t_1^T \\ \vdots \\ t_{\widetilde{N}}^t \end{bmatrix}_{N \times M} \tag{14}$$

$\widetilde{N}$ = number of Hidden Units, N = number of observations

The hidden layer parameters are randomly selected, before giving the training data [22]. The ELM not only achieves minimum training error but also the smallest norm of the output weights. Tab. 1 shows the parameters used for ELM Training.

Table 1: ELM training parameters

Parameter	Specification
ELM Type	ELM Classification
Kernel Type	Radial basis function kernel
Wavelet Feature	1 × 16384 samples
Time-domain Feature	1 × 43 samples
Frequency Domain Feature	1 × 2 samples
Regularization coefficient	1

5 Result and Discussion

5.1 Acquisition of Sound Samples

The different sound signals of motorcycles (Royal Enfield Bullet) are recorded using the digital voice recorder, under the supervision of experienced mechanics. Acoustic features are acquired from motor cycle engine with single cylinder, 346cc, air cooled petrol engine, four stroke, maximum torque of 28Nm @4000 and 5 speed manual gear box. The signals are sampled with the frequency 44.1 kHz. The sampled signals are further low pass filtered and normalized. The database consists of a total of 4500 samples which includes 500 samples of healthy motorcycles and 4000 samples of faulty motorcycles are show in Tab. 2. The faulty motorcycles produce problematic sounds like an air filter, air screw problem, air silencer; float needle problem, high acceleration problem hydraulic tap pit failure and pushrod problem.

Table 2: Experimental data for ELM training and testing

S. No.	Type	Sample Size	ELM Training	ELM Testing
1	Motorcycle in good condition	500	370	130
2	Air filter problem	570	425	145
3	Air screw problem	570	425	145
4	Alter silencer problem	580	435	145
5	High Acceleration problem	570	425	145
6	Float needle problem	570	425	145
7	Hydraulic tap pit problem	570	425	145
8	Push rod problem	570	425	145

(a) Alter silencer problem

There may be sparks firing in the silencers. This ensures when the exhaust/silencer is tampered to achieve better performance. The inner circumference of the silencer would be surfaced with exhaust residue, mostly unburnt fuel which is pushed from the cylinder. And this unburnt fuel is sparked by the heat exhaust fumes along with the silencer and it appears as sparks and sounds like gunfire. Fig. 7a shows the air silencer and the generated sound signal due to air silencer problem.

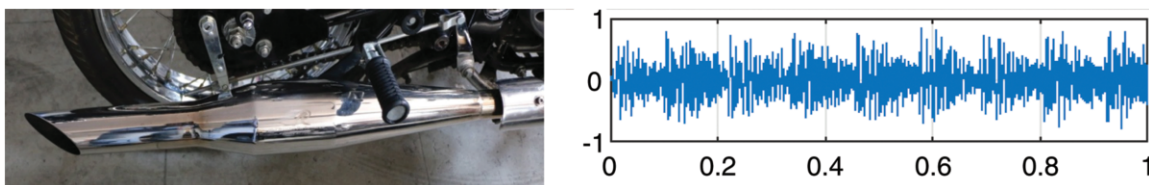**Figure 7a:** Alter silencer and the generated sound signal due to alter silencer problem

(b) Air filter problem

An air filter is a pleated multi-layered, oil impregnated mesh prepared from surgical cotton or other fabric. It is used for cleaning air going to the intake system of the engine, thereby reducing its turbulence, while limiting the foreign particles in that air. The motorcycle with a problem in the air filter produces a

different sound. i.e., a high flow air filter allows more air than a normal air filter. Fig. 7b shows the air filter and the generated sound signal due to air filter problem
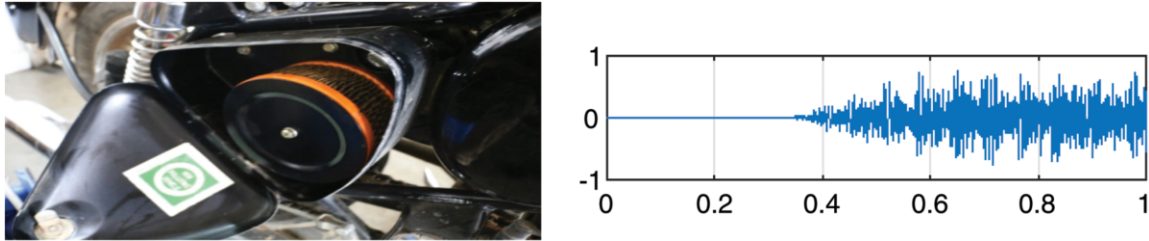
Figure 7b: Air filter and the generated sound signal due to air filter problem

(c) Pushrod problem

A bad push rod (usually this means it's bent) should produce unwanted sound. Fig. 7c shows Pushrod and the generated sound signal due to pushrod problem
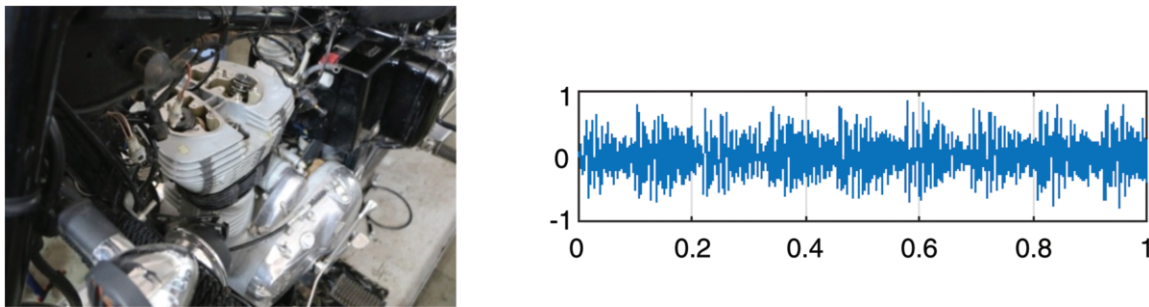
Figure 7c: Pushrod and the generated sound signal due to pushrod problem

(d) Hydraulic tappet

Due to the damaged hydraulic tappets, tappet noise should be produced. When dirt pushes inside the tappet cylinder its hydraulic property is lost and thus resulting in some clearance with the pushrods. As a consequence, the unwanted sound might be produced. Fig. 7d shows Hydraulic tappet and the generated sound signal due to Hydraulic tappet problem

Figure 7d: Hydraulic tappet and the generated sound signal due to Hydraulic tappet problem

(e) Airscrew problem

When the carburetor is out of adjustment, the air/fuel screw and the balance between two or more carburetors need to be adjusted. The incorrect adjustment can produce unwanted sounds.

Figure 7e: Airscrew and the generated sound signal due to air screw problem

(f) Float needle problem

The floats in a carburetor are typically made from either brass or plastic. Most of the floats in small engines do not have a metal tang to adjust them. The problems in the float needle produce different sounds.

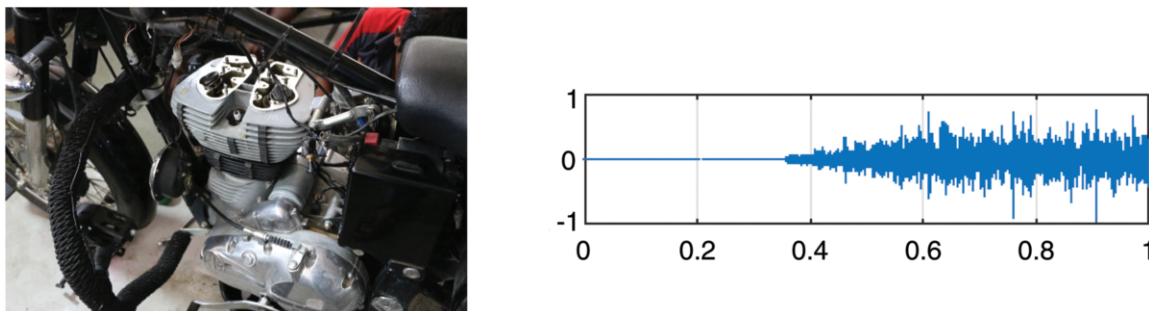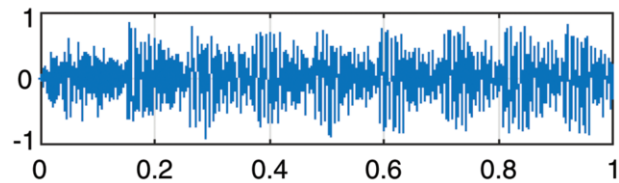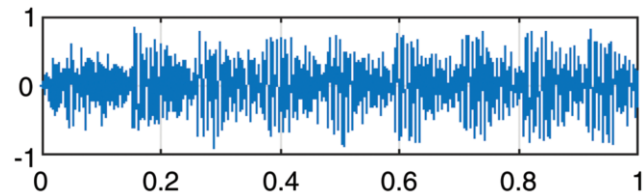**Figure 7f:** Float needle and the generated sound signal due to float needle problem

(g) High acceleration problem:

Increasing the speed of the accelerator produces different sounds. This is called a high acceleration problem. Figs. 7g and 7h show the generated sound signal due to the bike with normal speed and high speed.
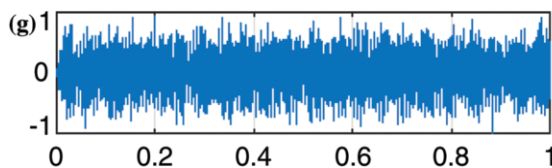
Figure 7g: Sound signal due to high acceleration problem
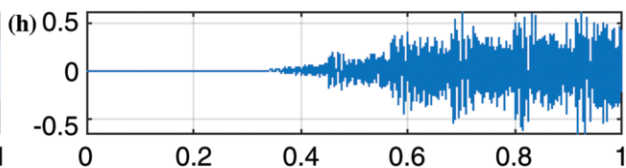
Figure 7h: Sound signal due to the bike with good condition (normal Speed)

Figure 7: Shows the sample waveform for the faults which are mentioned above

The pertinent features are extracted from the recorded sound signals after preprocessing. The features are further normalized and randomized. 75% of the dataset is used for training and the remaining 25% are used for testing.

These features are given as input to the FFNN, ELM classifiers. In the hidden layer, the hyperbolic tangent sigmoidal activation function is used and in the output layer, the logistic sigmoidal activation function is used. The total number of output neurons is fixed as 8 and the number of the epoch is fixed as

5000. The output of the ELM designates whether the input feature vector matches to the healthy or faulty motorcycle. In our experiments, numbers of hidden neurons are 225 and 20 for ELM and FFNN respectively.

The different classes are given below:

C_1 – Motorcycle in good condition

C_2 – Air filter problem

C_3 – Air screw problem

C_4 – Alter silencer problem

C_5 – High acceleration problem

C_6 – Float needle problem

C_7 – Hydraulic tap pit problem

C_8 – Push rod problem

The performance of the network is tested using the test datasets and validated by calculating different performance metrics. A 10-fold cross-validation algorithm is adopted to get enhanced test performance.

5.2 Performance Assessment

The performance of the proposed methodology is analyzed by evaluating different performance metrics given in Eqs. (18) to (20).

$$Accuracy = \frac{t_p + t_n}{t_p + f_p + t_n + f_n} \tag{18}$$

$$Sensitivity = \frac{t_p}{t_p + f_n} \tag{19}$$

$$specificity = \frac{t_n}{t_n + f_p} \tag{20}$$

$$Precision = \frac{t_p}{t_n + f_p} \tag{21}$$

$$FalseMeasure = \frac{f_p}{f_p + t_n} \tag{22}$$

$$F1 - score = \frac{2t_p}{2t_p + f_p + f_n} \tag{23}$$

where t_p -True-positive samples,

t_n - True-negative samples

f_p - False-positive samples

f_n - False-negative samples

Tables 3 and 4 show the classification accuracy using ELM classifier for the individual feature set (F_1, F_2 and F_3). The diagonal elements of the Table correspond to the correctly classified and the off-diagonal elements correspond to misclassified in percentage. It is observed that, for the time domain-based feature

Table 3: Classification accuracy for F_1 and F_2 feature set and ELM

| | Time-domain features (F_1) | | | | | | | | Frequency domain-based feature set (F_2) | | | | | | | |
	C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8	C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8
C_1	91	1	1	2	0	1	2	1	92	1	1	0	3	1	1	1
C_2	3	77	4	1	3	5	4	3	3	83	2	3	2	4	2	1
C_3	0	1	92	2	1	2	1	1	2	1	92	0	1	0	3	1
C_4	1	2	3	86	1	2	2	3	1	2	2	90	1	1	0	3
C_5	3	6	4	4	68	5	4	6	1	2	2	3	86	2	1	3
C_6	2	1	1	2	0	91	2	1	5	1	3	4	2	78	4	3
C_7	5	0	4	1	4	1	79	6	0	4	2	2	3	4	83	2
C_8	4	0	2	3	2	4	2	83	1	1	2	5	1	4	3	82

set, the average classification rate achieved is 83.4% and for the frequency domain, 85.5% is attained, i.e., only 2% improvement is accomplished.

Similarly, the classification rate for the wavelet domain-based features is found out and the results are shown in Tab. 4.

The features $F_3 = [E_1, E_2, E_3, E_4, E_5, E_6, E_7, E_8]$ obtained from the energy distribution of the detailed and approximation coefficients of wavelet packet decomposition are used as input to the classifier.

Table 4: Classification accuracy for F_3 feature set and ELM

| | Wavelet domain features (F_3) | | | | | | | |
	C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8
C_1	93	1	1	2	0	1	0	2
C_2	1	79	4	1	3	5	4	3
C_3	3	0	86	2	0	2	4	3
C_4	1	3	4	82	1	3	2	4
C_5	1	3	0	2	89	2	1	2
C_6	0	2	1	2	0	91	3	1
C_7	2	2	3	1	1	1	88	2
C_8	1	1	2	0	1	0	2	93

From Tab. 4, it is evident that the classification accuracy obtained is approximately 87.62%. ELM and FFNN have the same form as the SLFN structure. Hence, the proposed method is compared with the classic FFNN classifier. Tab. 5 summarizes the comparison result of various performance metrics for both the classifiers.

Table 5 clearly shows that the performance of the ELM classifier gives better results when compared to FFNN classifier. Additionally, the WPT based feature sets also provide higher performance values.

Table 5: Performance results for different feature sets

Classifier	Accuracy (%)	Specificity	Sensitivity	Precision	False Measure	F1-Score	Time (s)	
							Training	Testing
Time-domain features (F1)								
ELM	**83.37**	**0.8921**	**0.9012**	**0.9521**	**0.12**	**0.85**	**0.3982**	**0.4731**
FFNN	82.4	0.8723	0.8891	0.942	0.16	0.84	0.5012	0.5029
Frequency domain-based feature set (F2)								
ELM	**85.75**	**0.9045**	**0.9126**	**0.945**	**0.36**	**0.91**	**0.3366**	**0.4067**
FFNN	83.12	0.9144	0.9258	0.932	0.23	0.89	0.4203	0.6755
Wavelet packet-based feature set(F3)								
ELM	**87.62**	**0.9166**	**0.9216**	**0.9321**	**0.132**	**0.89**	**0.2934**	**0.3210**
FFNN	86.23	0.9275	0.9311	0.9213	0.235	0.87	0.3012	0.3211

Hybrid feature sets: The performance of the classifier can be further enhanced by using the combination of feature sets (hybrid features), i.e., fusing the features in the three domains. The classification accuracy for the hybrid feature sets (F_4, F_5, F_6 and F7) is shown in

Table 6: Classification accuracy for hybrid feature sets and ELM

	Hybrid Feature set (F_4)								Hybrid Feature set (F_5)							
	C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8	C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8
C_1	92	1	1	2	0	1	2	1	95	1	1	0	1	1	0	1
C_2	1	89	2	1	3	1	2	1	1	95	0	0	3	0	0	1
C_3	1	1	94	0	1	0	2	1	1	1	93	0	1	0	0	0
C_4	1	0	2	93	1	1	0	2	1	1	0	96	1	1	0	0
C_5	1	1	0	0	92	2	1	3	1	0	2	0	94	2	1	0
C_6	1	1	1	1	1	92	2	1	0	1	1	0	1	96	0	1
C_7	2	1	2	1	1	1	90	2	0	0	1	1	1	0	97	0
C_8	1	1	2	1	1	0	0	94	0	1	2	1	1	0	0	95
Hybrid Feature set (F_6)									Hybrid Feature set (F_7)							
	C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8	C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8
C_1	97	1	1	0	0	1	0	0	98	0	0	0	0	1	0	1
C_2	1	96	1	0	2	0	0	1	0	99	1	0	0	0	0	0
C_3	0	0	98	0	0	0	1	1	0	1	99	0	0	0	0	1
C_4	1	0	0	97	1	1	0	0	1	1	0	98	0	0	0	0
C_5	1	0	0	1	98	0	0	0	1	0	0	0	97	2	0	0
C_6	1	0	1	1	0	96	0	1	0	0	1	0	0	99	1	1
C_7	0	1	1	0	1	1	96	0	0	0	0	1	0	0	99	0
C_8	1	1	2	0	0	0	2	94	0	0	1	0	0	0	1	98

The average classification accuracy for F_4 feature set is 94.3%. For feature set F_5, 95.34% is attained. For F6, 96.5% is achieved. For F7 feature set, 98.5% is attained. To validate the performance of the ELM classifier, the proposed hybrid feature extraction technique with FFNN classifier is also investigated and the results are shown in Tab. 7. Thus, it is proved that the increase in classification accuracy is attained by the application of a hybrid feature set.

Table 7: Performance results using hybrid feature sets

Classifier	Accuracy (%)	Specificity	Sensitivity	Time (s)	
				Training	Testing
Hybrid feature set (F4)					
ELM	**94.3**	**0.8921**	**0.9012**	**0.3982**	**0.4731**
FFNN	82.4	0.8723	0.8891	0.5012	0.5029
Hybrid feature set (F5)					
ELM	**95.34**	**0.9045**	**0.9126**	**0.3366**	**0.4067**
FFNN	88.12	0.9144	0.9258	0.4203	0.6755
Hybrid feature set (F6)					
ELM	**96.51**	**0.9166**	**0.9216**	**0.2934**	**0.3210**
FFNN	89.23	0.9275	0.9311	0.3012	0.3211
Hybrid feature set (F7)					
ELM	**98.51**	**0.9312**	**0.9291**	**0.1037**	**0.2609**
FFNN	91.54	0.9293	0.9367	0.2082	0.2207

A classification rate of 98.5% was achieved for the hybrid feature set. Fig. 8 shows the performance of the ELM classifier with hybrid dataset for variable number of hidden nodes.

It is found that the best performance is achieved by increasing the number of a hidden node with a maximum of 100.

Thus, in the above discussions and from the comparison results, it is concluded that the combination of the features sets in all three domains, i.e., time, frequency and wavelet domains provide better performance.
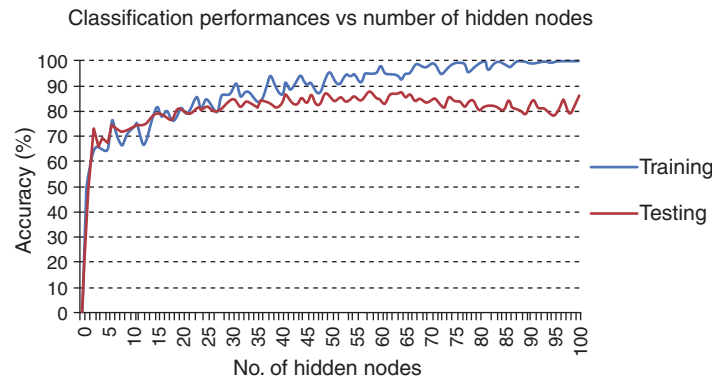

Figure 8: Performance of the neural network training

6 Conclusion

Fault classification using ELM classifier is proposed for the recognition of various faults in the motorcycle. Features from time, frequency and wavelet domains are extracted from the sound signals captured generated by the motorcycles. These feature sets give important attributes of the signal. The high performance is obtained by the application of ELM classifier, because of its unique characteristics. The proposed methodology is tested in the real environment by collecting the sound signals which is recorded from the motorcycles with different faults. Thus, the ELM classifier with hybrid feature sets can be used for predicting, identifying and classifying various faults in the motor cycles at the early stage. By identifying the faults automatically, we can avoid situations such as engine failure, serious accidents and other components from getting damaged. Consequently, the faulty components can be replaced or get repaired at the early stages, thereby improving the life time of other engine components. For example, if the problem in the air filter is automatically identified first, it can be replaced immediately and by replacing the faulty component, the mileage of the motor cycle can be improved very much. The experimental results and discussions clearly demonstrate that the proposed system provides better performance when compared to the conventional classification systems.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

1. Kemalkar, A. K., Bairagi, V. K. (2016). Engine fault diagnosis using sound analysis. *Proceedings of the International Conference on Automatic Control and Dynamic Optimization Techniques (ICACDOT)*, *Pune* 943–946.

2. Goksu, H. (2018). Vehicle speed measurement by on-board acoustic signal processing. *Measurement and Control, 51(5–6),* 138–149.

3. Anami, B. S., Pagi, V. B. (2013). Acoustic signal based detection and localisation of faults in motorcycles. *IET Intelligent Transport Systems, 8(4),* 345–351. DOI 10.1049/iet-its.2012.0193.

4. Paulraj, M. P., Adom, A. H., Hema, C. R., Sathishkumar, S. (2012). Acoustic signature recognition of moving vehicles using Elman neural network. *KarpagamJcs, 6.*

5. Pan, S., Han, T., Tan, A. C. C., Lin, T. R. (2016). Fault diagnosis system of induction motors based on multiscale entropy and support vector machine with mutual information algorithm. *Shock and Vibration, 2016,* 1–12. DOI 10.1155/2016/5836717.

6. Schneider, W. H., TarmoSavolainen, P., Boxel, D. V., Beverley, R. (2012). Examination of factors determining fault in two-vehicle motorcycle crashes. *Accident Analysis and Prevention, 45,* 669–676.

7. Anami, B. S., Pagi, V. B. (2014). Acoustic signal-based approach for fault detection in motorcycles using chaincode of the pseudospectrum and dynamic time warping classifier. *IET Intelligent Transport Systems, 8(1),* 21–27. DOI 10.1049/iet-its.2012.0086.

8. Zhou, Q., Shen, H., Zhao, J., Liu, X., Xiong, X. (2019). Degradation state recognition of rolling bearing based on K-means and CNN algorithm. *Shock and Vibration, 2019,* 1–9. DOI 10.1155/2019/8471732.

9. Pagi, V. B., Wadawadagi, R. S., Anami, B. S. (2015). Fault localization in motorcycles using wavelet packet energy distribution. *Conference on Soft-Computing and Networks Security, IEEE, Coimbatore.*

10. Wang, X., Xiang, J., Hu, H., Xie, W., Li, X. (2015). Acoustic emission detection for mass fractions of materials based on wavelet packet technology. *Journal of Ultrasonics, 60,* 27–32. DOI 10.1016/j.ultras.2015.02.001.

11. Berredjem, T., Benidir, M. (2018). Bearing fault diagnosis using Fuzzy expert system relying on an improved range overlaps and similarity measures. *Expert Systems with Applications, 15,* 34–142.

12. Goksu, H. (2018). Engine speed-independent acoustic signature for vehicles. *Measurement and Control, 51(3–4),* 94–103.

13. Patel, K. (2016). A review on feature extraction methods. *International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, 5(2),* 823–827.

14. Kemalkar, A. K., Bairagi, V. K. (2016). Engine fault diagnosis using sound analysis. *Proceedings of the International Conference on Automatic Control and Dynamic Optimization Techniques, Pune.*

15. Plaza, E. G., Núñez López, P. J. N. (2018). Application of the wavelet packet transform to vibration signals for surface roughness monitoring in CNC turning operations. *Mechanical Systems and Signal Processing, 98,* 902–919. DOI 10.1016/j.ymssp.2017.05.028.

16. Bianchi, D., Mayrhofer, E., Gröschl, M., Betz, G., Vernes, A. (2015). Wavelet packet transform for detection of single events in acoustic emission signals. *Mechanical Systems and Signal Processing, 64-65,* 441–451. DOI 10.1016/j.ymssp.2015.04.014.

17. DongSik, G., Kim, J., Kelimu, T., Huh, S. C., Choi, B. K. (2012). Evaluation of the use of envelope analysis and DWT on AE signals generated from degrading shafts. *Materials Science and Engineering, 177(19),* 1683–1690. DOI 10.1016/j.mseb.2012.03.001.

18. Darwish, H. A., Hesham, M., Taalab, A-M. I., Mansour, N. M. (2010). Close accord on DWT performance and real-time implementation for protection applications. *IEEE Transactions on Power Delivery, 25(4),* 2174–2183. DOI 10.1109/TPWRD.2009.2036401.

19. Rodriguez, P. H., Alonso, J. B., Ferrer, M. A., Travieso, C. M. (2014). Review of automatic fault diagnosis systems using audio and vibration signals. *IEEE Transactions on Systems, Man, and Cybernetics: Systems, 44(5),* 642–652.

20. Chen, Q. X., Wu, C. F. (2017). Research on engine fault diagnosis technology based on ANFIS. *Proceedings of the IEEE Conference on Chinese Automation Congress, Jinan.*

21. Anam, K., Al-Jumaily, A. (2017). Evaluation of extreme learning machine for classification of individual and combined finger movements using electromyography on amputees and non-amputees. *Neural Networks, 85,* 51–68. DOI 10.1016/j.neunet.2016.09.004.

22. Liang, L., Liu, F., Li, M., He, K., Xu, G. (2016). Feature selection for machine fault diagnosis using clustering of non-negation matrix factorization. *Measurement, 94,* 295–305. DOI 10.1016/j.measurement.2016.08.003.