End of the Assembly Line Gearbox Fault Inspection Using Artificial Neural Network and Support Vector Machines

Prasad V. Kane and Atul B. Andhare

Department of Mechanical Engineering, Visvesvaraya National Institute of Technology, Nagpur, Maharashtra, India.

Gear fault diagnosis is important not only during the routine maintenance of machinery, but also during the inspection of newly manufactured gearboxes at the end of the assembly line. This paper discusses the application of an artificial neural network (ANN) and a support vector machine (SVM) for identifying faults in the gearbox, using the psychoacoustic and conventional statistical features extracted from acoustics and vibration signals. It is observed that at the end of the assembly line, the gearbox is tested by mounting it on a test bench and driving it by an electric motor. Based on the sound emitted while running on the test bench, the operator decides on the acceptance of the gearbox for further assembly on a vehicle or machine. This method of acceptance or rejection of the gearbox involves subjectivity and it is not reliable. Hence, it is important to have a reliable and objective fault detection and diagnosis method. To eliminate subjectivity, psychoacoustic features, which are derived from the science of listening in human beings, are proposed to be used as features, along with ANN and SVMs as classifiers. To ascertain the ability of the psychoacoustic features to classify faults, laboratory experiments are carried on a test setup by simulating faults like a gear shaft misalignment, a profile error of a gear tooth, a crack at the root of the tooth, and a broken tooth. ANN and SVM are trained with the psychoacoustic features extracted from the acoustic signal and other statistical features from the acoustics and vibration signals. The trained SVM and ANN are tested for fault classification for these features and their accuracy is compared. Fault classification accuracy is found to be 95.65% for ANN and 93.44% for SVM with psychoacoustic features and is found to be better than pure statistical features obtained from the vibration and acoustic signals. With the optimised ANN and SVM architecture, SVM is found to be performing better than ANN. It is concluded that the psychoacoustic features, along with the ANN and SVM method, could be adopted at the end of assembly line inspection to make the inspection process more objective.

1. INTRODUCTION

The gearbox plays an important role in industrial applications, such as power transmission machinery and rotating machinery. It is one of the most important elements in almost all vehicles. Identification of gear faults is an important task in maintenance functions. It is also very important during the inspection of the newly manufactured gearbox at the end of the assembly line. Many researchers have contributed to the field of gear fault diagnosis and its ramifications are wide and are interdisciplinary in nature. Earlier, fault diagnosis was based on vibration techniques in which time and frequency domain analysis was carried out. Summary of such methods has been reported in many books and review papers.[1–4] Worden et al. had elaborated on, in detail, the application of different soft computing algorithms to the fault diagnosis of mechanical systems describing all the steps involved, i.e. data collection with sensors, feature extractions, and application of different artificial intelligence techniques used for pattern recognition.[5] The complexity of gearbox technology led researchers to develop innovative and advanced techniques for fault diagnosis to enhance accuracy and reliability. Zhang, et al.[6] and Sharma, et al.[7] proposed a time synchronous averaging method to diagnose faults under fluctuating speed conditions. Innovative

techniques like empirical mode decomposition, Hilbert-Haung transforms, short-time Fourier transforms, wavelets, etc. are applied for feature extraction and signal processing.[8–11] These methods figure out the information content of the signal due to the transient effect generated by fault and inherent nature of amplitude and frequency modulated vibration signal emitted by the gearbox. Envelop analysis is also preferred by many researchers to extract the information content in the signal.[12] Other approaches like cepstrum analysis and bispectrum analysis are also used by researchers to detect and localise fault in the gearbox.[13,14] Though vibration-based techniques are widely studied and accepted, the acoustic technique is also emerging and efforts have been made by researchers to diagnose gearbox and bearing faults using it.[15,16] Due to the complexity involved in analysing signals obtained in the time domain or the frequency domain, researchers have focused their attention on complementing vibration and acoustic methods with the artificial intelligence technique for fault diagnosis. Vakharia, et al. applied SVM as a supervised machine learning technique and self-organising maps to diagnose bearing faults with the features extracted from the vibration signal.[17] Singh, et al. demonstrated a rotor fault diagnosis using ANN and SVM.[18] ANN based on a multilayer feedforward backpropagation algorithm is the most accepted pattern classifier based on

the human biological system, which can be trained to map all the non-linearity involved in the features.[19,20] SVM belongs to a different pedigree. It is based on the statistical learning technique where the features are classified using separating hyperplanes for a multidimensional vector feature space. It is found to work in a robust way, avoiding overtraining and generalisation issues.[21,22] Many researchers have used these techniques along with the features extracted from the vibration signal.[23,24]

The work discussed in this paper focuses on end of the line inspection of the automobile gearbox, where it is observed that the gearbox is tested manually by a human operator on the basis of his judgment. The gearbox is operated by a motor on a test bench and the human auditor listens the sound emitted by the gearbox to judge the quality of the gearbox based on his experience to discriminate between good and faulty gearboxes. The identification of the quality of the gearbox is based on the subjective decision of operator and his experience of testing gearboxes. Very rare work is reported in the literature for fault diagnosis problems for newly manufactured gearboxes. Shang, et al. have reported on this issue and have proposed vibration-based technique with time synchronous averaging and an ANN technique to address this issue by applying a genetic algorithm for feature selection.[25] Cook, et al. have elaborated on trends and perspectives of the end of line inspection in a technical note and proposed that the ability of psychoacoustic features to objectively inspect the devices can overcome the limitations of human auditors.[26] Hence, attempts are made to develop a technique to identify faults using psychoacoustic features and artificial intelligence technique to help the operator to objectively make a decision about the acceptability of the gearbox. Kane and Andhare have reported the use of psychoacoustics for a single fault, i.e. a crack at the root of the gear tooth for the diagnosis of a gear fault with ANN.[27] In the present work, multiple faults like gear shaft misalignment with a healthy gear, a gear tooth with profile error (e.g. a scuffed tooth), a crack at the root of the gear tooth, and a broken tooth are simulated, and each condition is diagnosed using ANN and SVM. Experiments are carried out in a laboratory by seeding the faults in a gearbox; the details of the experimental setup and methodology are discussed in following section.

2. EXPERIMENTAL SETUP AND METHODOLOGY

2.1. Experimental Setup

Experiments were carried out on a Spur gearbox test rig. The layout and the photograph are shown in Fig. 1(a) and 1(b). It consists of a single stage spur pinion and gear. The shafts are supported by flanges with ball bearings. A single-phase DC motor with a speed controller operates the input shaft. The output shaft is connected to the loading arrangement with a pulley and rope for applying the required load. A data acquisition (DAQ) card by National Instruments (NI9234) was used to acquire the vibration and acoustic signals with the accelerometer and the microphone. The microphone was located in a free field at a 1 m distance from the gearbox and the accelerometer was placed on the bearing housing. It was connected to the DAQ card and the data were acquired and stored on a computer. LabVIEW software was used to acquire the vibration

Table 1. Expressions to compute statistical features of acoustic and vibration signals.

Sr. No	Statistical Indicator	Formula
1	Root Mean Square (RMS)	$\sqrt{\frac{\sum_{n=1}^{N}(x(n))^2}{N}}$
2	Kurtosis	$\frac{\sum_{n=1}^{N}(x(n)-\mu)^4}{N\sigma^4}$
3	Skewness	$\frac{\sum_{n=1}^{N}(x(n)-\mu)^3}{N\sigma^3}$
4	Maximum	$\max[x]$
5	Minimum	$\min[x]$
6	Range	$\max[x] - \min[x]$
7	Crest Factor	$\frac{Peak\ Value}{RMS\ Value}$
8	Form Factor	$\frac{RMS\ Value}{Mean\ Value}$
9	Mean	$\frac{\sum_{n=1}^{N}(x(n))}{N}$
11	Variance	$\frac{\sum_{n=1}^{N}(x(n)-\mu)^2}{N}$

Where $x(n)$ = amplitude of the nth digitized point in the time domain, N = number of points in time domain and μ = mean of the N points, σ = standard deviation.

and acoustic signals and extract various psychoacoustic and signal statistical features. The vibration and acoustics signals and their features were obtained at a constant speed of 720 rpm, and at a no-load condition, and with a partial load condition. The various gear faults with the increasing severity of faults were introduced and the assembly error for the gear shaft misalignment was also introduced in the setup. In the experiment, 50 samples of the signals of each condition of gear were acquired and its features were extracted. Thus, there were 500 sample signals. The sampling frequency selected for acquiring data was set at 44 kHz for the acoustic signal and of 20 kHz for the vibration data.

2.2. Methodology

Various gear tooth faults like a scuffed tooth (i.e., a tooth profile error), a crack at the root, and a broken tooth were introduced. Other faults (e.g. an assembly error of shaft misalignment with a healthy gear) are also introduced. Various features of acoustic and vibration signal were extracted by simulating these faults in test setup. The pattern classifiers, such as ANN and SVM, are applied on the extracted features to diagnose the presence of a fault as a two-class problem. The methodology followed in this study is depicted in the flow chart shown in Fig. 2. The details of the features extracted from the vibration and acoustic signal are discussed in next section.

3. SIGNAL PROCESSING AND FEATURE EXTRACTION

3.1. Statistical Features of Vibration and Acoustic Signal

Statistical features characterise the signal by obtaining the values such as the mean, rms, kurtosis, skewness, etc. from the vibration and acoustics time domain signals, and various statistical features were computed using the LabVIEW programme. A brief description of these features is presented in Table 1 and the sample values obtained are shown in Table 2 and 3.

RMS is the normalized second statistical moment of the signal. It is used to describe the overall health of the machine. Kurtosis is the normalized fourth statistical moment of the sig-
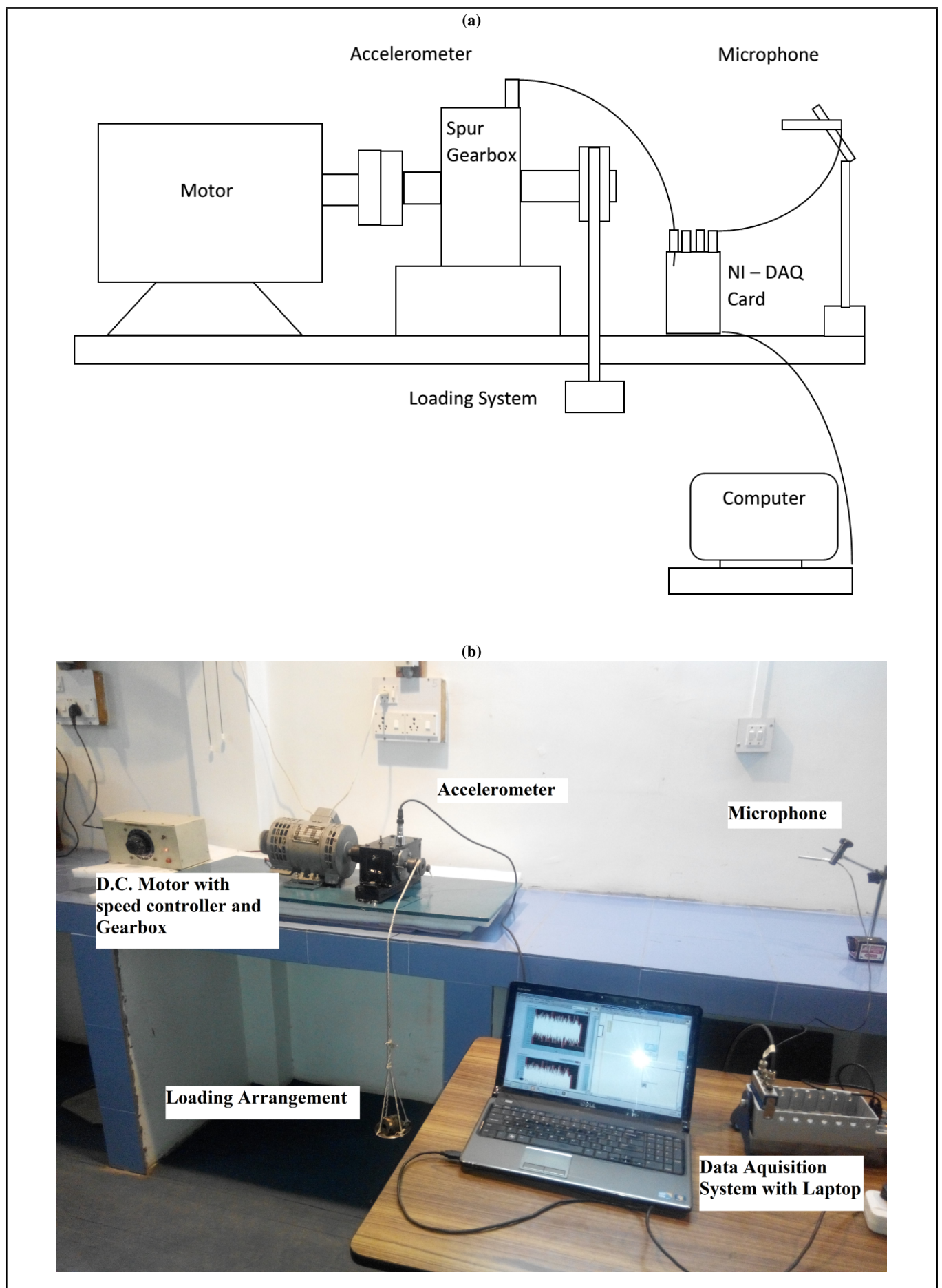
(a)

(b)

Figure 1. (a) Layout of experimental setup. (b) Photograph of experimental setup.
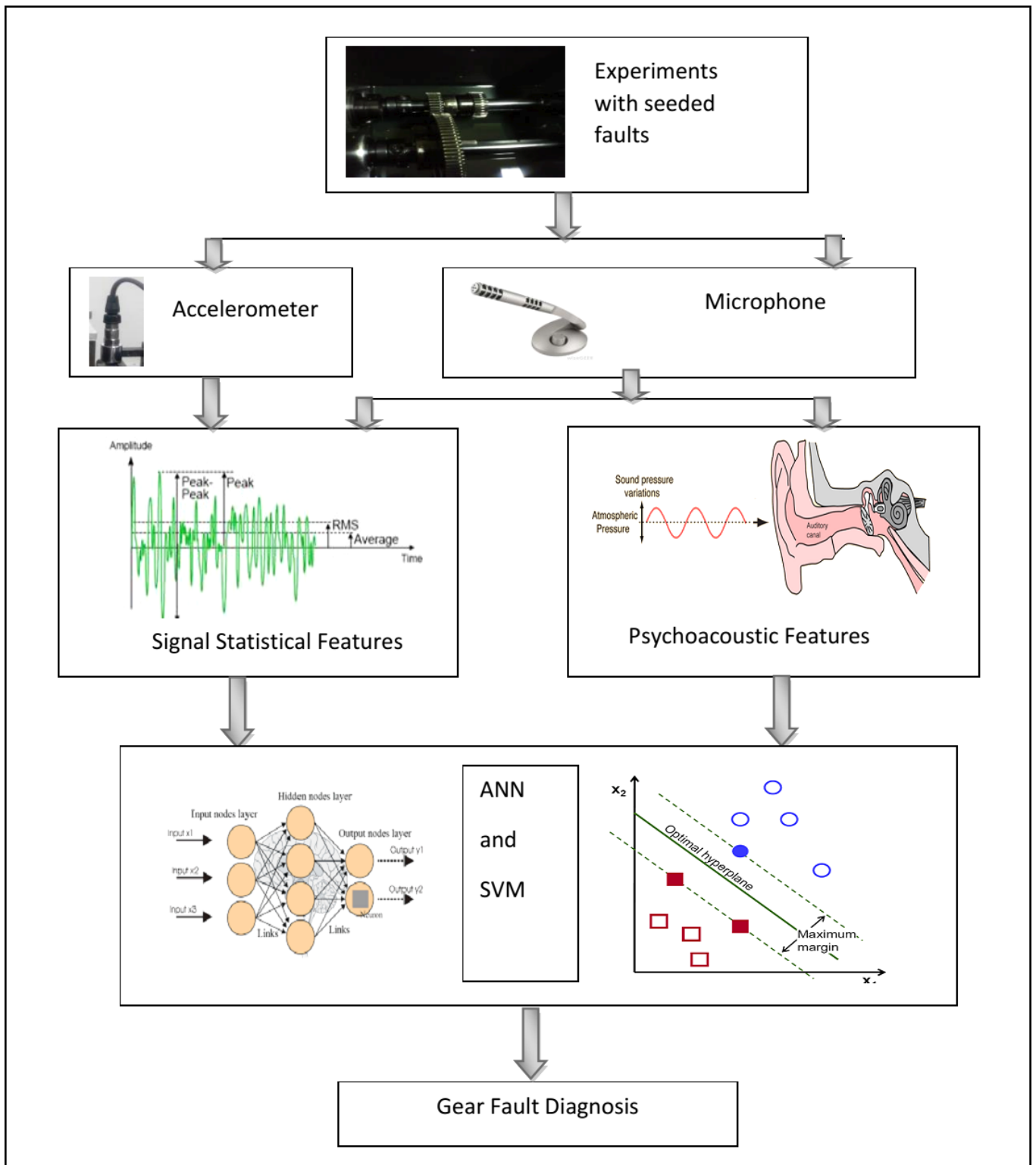
Figure 2. Methodology.

Table 2. Sample values of statistical features extracted from vibration signal.

Status	Load	RMS	Variance	Kurtosis	Skewness	Max	Min	Range	Mean	Form Factor	Crest Factor
Healthy	no load	0.149	0.022	4.826	0.061	0.893	−1.052	1.945	0.010	14.670	5.994
Healthy	with load	0.344	0.118	3.415	−0.030	1.843	−1.799	3.642	0.002	186.395	5.358
Crack at root	no load	0.200	0.040	5.995	0.024	1.435	−1.673	3.107	0.002	118.074	7.156
Crack at root	with load	0.723	0.523	3.608	−0.016	3.453	−3.345	6.798	0.002	461.447	4.776
Profile error	no load	0.555	0.308	7.913	−0.011	3.765	−4.013	7.779	0.003	186.349	6.781
Profile error	with load	0.803	0.645	3.369	0.003	3.415	−3.719	7.133	0.002	437.071	4.250
Broken tooth	no load	0.191	0.037	5.351	0.033	1.353	−1.207	2.559	0.003	65.266	7.079
Broken tooth	with load	0.725	0.526	7.079	0.014	5.015	−5.856	10.870	0.002	334.341	6.914
Misalignment	no load	0.415	0.172	5.892	0.001	3.050	−2.984	6.034	0.002	240.989	7.351
Misalignment	with load	0.149	0.022	4.826	0.061	0.893	−1.052	1.945	0.010	14.670	5.994

Table 3. Sample values of statistical features extracted from acoustic signal.

Status	Load	RMS	Variance	Kurtosis	Skewness	Max	Min	Range	Mean	Form Factor	Crest Factor
Healthy	no load	0.060	0.004	3.055	−0.039	0.230	−0.254	0.484	0.008	7.849	3.852
Healthy	with load	0.067	0.004	3.041	−0.028	0.261	−0.273	0.533	0.003	25.794	3.913
Crack	no load	0.061	0.004	2.938	0.018	0.247	−0.255	0.502	0.002	27.491	4.065
Crack	with load	0.091	0.008	3.057	−0.022	0.411	−0.384	0.795	0.002	53.881	4.523
Profile error	no load	0.107	0.011	3.091	−0.027	0.467	−0.466	0.933	0.003	36.957	4.384
Profile error	with load	0.087	0.008	3.049	0.008	0.388	−0.376	0.764	0.004	20.066	4.469
Broken tooth	no load	0.056	0.003	3.011	−0.054	0.259	−0.233	0.491	0.006	9.922	4.636
Misalignment	no load	0.062	0.004	3.012	0.009	0.267	−0.308	0.575	0.006	10.115	4.293
Misalignment	with load	0.059	0.003	3.097	−0.014	0.256	−0.249	0.506	0.002	25.190	4.371

nal. It provides the measure of the pickiness or impulsive nature of the signal. Skewness is a measure of symmetry, or more precisely, the lack of symmetry. Peak value or maximum value is the highest point in a set of values. Range indicates the difference between maximum and minimum value of amplitude of the signal. Crest factor is the ratio of peak level to RMS level. It indicates the presence of high amplitude peaks caused by local damages. Form factor is the ratio of the RMS value to mean value which indicates the overall status of signal. Mean is the average of all the amplitude of the digitized points sampled. Variance indicates the spread of the amplitude of the values from its mean.[28–30]

3.2. Psychoacoustic Features of Acoustic Signal

Psychoacoustics is the scientific study of the sound perception. It involves the study of the psychological and physiological responses associated with sound. It is an interdisciplinary field, including psychology, acoustics, electronic engineering, physics, biology, physiology, and computer science. Fastle, Zwicker, and Aures have significant contributions in this field. The contributions of these scientists have helped to develop the methods to objectively quantify the sound that the human being perceives. The way a human being extracts the information contained in the acoustic signals using the natural senses is mimicked in the algorithms that have been accepted by the International Organisation for Standardisation (ISO). As per ISO532B, the objective indices specified are stationary loudness, time varying loudness, roughness, sharpness, tonality, and fluctuation strength.[31] Loudness is a term referring to the human perception of sound volume, expressed in the units of sone corresponds to 40 dB sound at 1 kHz tone. Roughness is another algorithm used to determine the subjective judgement to correlate noticeable sound as heard by the human ear. Roughness is a hearing sensation related to the loudness modulation at frequencies too high to discern separately, such as a modulation frequency greater than 30 Hz. It is the algorithm developed to measure the energy in 24 barks that computes

and filters the envelope of signal in each band by measuring the amplitude modulation of each envelope. It then weights the level in each band with the frequency dependent weighting function.[32] This algorithm returns the roughness spectrum versus critical band and then integrates the roughness spectrum to measure the roughness. Sharpness is a hearing sensation related to frequency and it is independent of loudness. It corresponds to the sharp, painful, high energy sound and is the comparison of the amount of high frequency energy to the total energy.[33] Tonality is used to determine whether a sound consists mainly of the tonal component of broadband noise. The algorithm for tonality measures the relative strength of the signal, as compared to the overall signal.[34] For each time block, this algorithm first varies the frequency resolution according to human frequency selectivity, then searches the frequencies of likely tones and then compares the loudness of the sound. Fluctuation strength is a hearing sensation related to the loudness modulation at low frequencies that is perceptible individually. It uses a similar method to roughness versus time analysis, except that it focuses specifically on the signal variation with very low modulation frequencies.[35, 36] These psychoacoustics features like loudness, sharpness, roughness, and tonality were extracted by using the modules in LabView, which follows expressions given in Table 4, and the sample values are tabulated and shown in Table 5 for the various faults simulated in the setup.

4. APPLICATION OF ARTIFICIAL NEURAL NETWORK AND SUPPORT VECTOR MACHINE

4.1. Artificial Neural Network

ANN is an interconnected network model based on the biological learning process of the brain and finds numerous applications in data analysis, pattern recognition, and control.[37] Among the different types of neural networks, a feedforward backpropagation multilayer perceptron neural network is used for the present work. It consists of an input layer of source

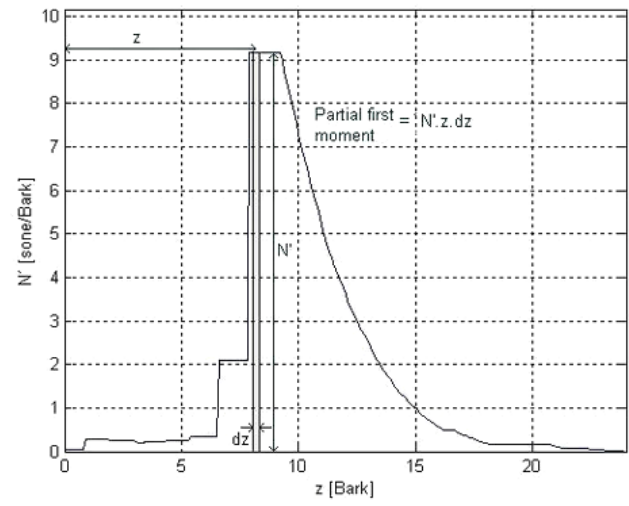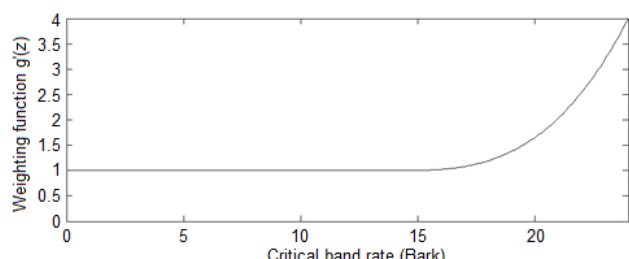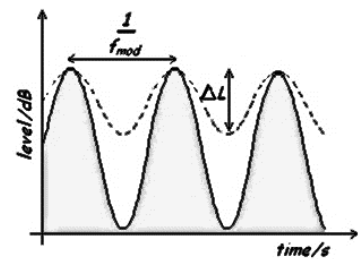**Table 4.** Psychoacoustic features and related expressions.[27]

Sr. No	Psycho-acoustic Feature	Formula	Description
1	Loudness (sone)[31]	$N = \int_0^{24\ \text{Bark}} N' dz$ (Refer Fig. 3)	N = loudness in sones N' = specific loudness z = Bark scale
2	Loudness (phon)[31]	$N = 2^{\frac{LL-40}{10}}$ $LL = 40 + 10\frac{\log(N)}{\log(2)}$	LL = loudness level in phons **Figure 3.** Critical band rate vs. specific loudness.[32]
3	Sharpness (acum)	$S = 0.11\frac{\int_0^{24\ \text{Bark}} N' g(z) dz}{\int_0^{24\ \text{Bark}} N' dz}$ $z < 14 \rightarrow g'(z) = 1$ $z > 14 \rightarrow g'(z) = 0.00012z^4 - 0.0056z^3 + 0.1z^2 - 0.81z + 3.51$ (Refer Fig. 4 for z value[31])	 **Figure 4.** Weighing $g'(z)$ vs. critical band rate.[32]
4	Roughness (asper)	$R = f_{mod} \int_0^{24\ \text{Bark}} \nabla Lg(z) dz$ (Refer Fig. 5)	 **Figure 5.** Modulation frequency (f_{mod}) and perceived modulation depth (ΔL).[32]
5	Fluctuation Strength (vacil)[31]	$F = \frac{\Delta L}{\frac{f_{mod}}{4\ \text{Hz}} + \frac{4\ \text{Hz}}{f_{mod}}}$	ΔL = perceived modulation depth f_{mod} = modulation frequency
6	Tonality (Tu) Aures Model[33]	$T = c\,W_T^{0.29} * W_{Gr}^{0.79}$ $W_T = \{\sum_{i=1}^M [W_1(\Delta z_i) * W_2(f_i) * W_3(\Delta L_i)]^2\}^{0.5}$ $W_1(\Delta z_i) = \left[\frac{0.13}{\Delta z_i + 0.13}\right]^{1/0.29}$ $W_2(f_i) = \left\{\frac{1}{1 + 0.2*\left(\frac{f_i}{700\ \text{Hz}} + \frac{700\ \text{Hz}}{f_i}\right)^2}\right\}^{0.5}$ $W_3(\Delta L_i) = 1 - \exp\left(-\frac{\Delta L_i}{15\ \text{dB}}\right)$ $\Delta L_i = L_i - 10\log_{10}\left[\sum_{k \neq i}^n A_E(f_i)\right]^2 + E_{Gr}(f_i) + E_{HS}(f_i)$	c = calibrating constant and is adjusted so that the value of 1 kHz tonal component at 60 dB will be 1 W_{Gr} = weight function signifies the loudness to tonal ratio of tonal element W_T = coherence function on the incitement of annoyance by tonal elements W_1 = the width of an individual tonal element W_2 = centre frequency W_3 = related to the value of tonal size Δz = width of tonal element f = center frequency in the unit Hz ΔL_i = value of calibration size of the tonal element in the unit of dB $A_E(f_i)$ = is the effect of tonal element close to tonal element $E_{Gr}(f_i)$ = intensity of noise in the critical band including the ith tonal element.

Table 5. Sample values of psychoacoustic features extracted from acoustic signal.

Status	Load	Roughness (asper)	Sharpness (acum)	Loudness (sone)	Loudness (phon)	Fluctuation Strength (vacil)	Tonality (Tu)
Healthy	no load	0.367	2.011	15.687	79.715	0.189	0.069
Healthy	with load	0.357	2.696	22.724	85.062	0.245	0.027
Crack	no load	0.442	2.128	18.309	81.945	0.401	0.000
Crack	with load	0.342	3.122	28.403	88.280	0.179	0.084
Profile error	no load	0.390	3.320	30.237	89.183	0.225	0.049
Profile error	with load	0.450	2.662	23.100	85.298	0.405	0.057
Broken tooth	no load	0.511	2.338	18.620	82.188	0.575	0.037
Broken tooth	with load	0.516	3.471	29.754	88.950	0.316	0.003
Misalignment	no load	0.348	2.795	22.049	84.626	0.317	0.034
Misalignment	with load	0.294	2.793	23.420	85.497	0.215	0.026

nodes, two hidden layers of computation neurons, and the output layer. The input layer nodes represent the normalized feature extracted from the measured signal. The number of input nodes is six, for the six psychoacoustic features used. Similarly, ten input nodes are used for the ten statistical features of the acoustic signal and the vibration signal. For selecting the number of hidden layers and the number of neurons in it, the rule of thumb and a few suggestions as given by Palit and Popovic are referred to, which suggests that two hidden layers are sufficient to solve the complex non-linear problems and number of neurons in the hidden layer should be in the neighbourhood of 75% of the number of network inputs, or between 0.5 and 3 times the number of inputs.[38] The geometric pyramid rule suggests assigning a hidden node as per the formula,

$$N_h = \alpha \sqrt{N_i \times N_0};$$

where N_h is hidden neuron where N_i is the number of input nodes in input layer, N_0 is the number of nodes in the output layer, and α is the multiplication factor, the value of which should be selected in the range of 0.5 to 2. Efforts are made in this work to observe the best performance of the ANN architecture by varying the number of layers and number of nodes in it and the architecture with the best results is selected and has been converted into a generalised program, which selects the number of hidden nodes as equal to the number of input nodes for the two hidden layers. Therefore, the two hidden layers are selected and the number of neurons in each hidden layer is taken equal to the number of inputs in every architecture. The output node is one in all architectures. The target value of the output node is having binary value 0 and 1, representing good and faulty condition, respectively. In the ANN, the activation function of tan-sigmoid (tanh) and logistic (log-sigmoid) were used in the hidden layer and output layer, respectively. The ANN was created, trained, and implemented using code written in MATLAB with the training algorithm of Levenberg-Marquardt.[39] Out of the 500 sample signals acquired, 40% were used for training, 30% for testing, and 30% for validation. The ANN was trained iteratively to minimize the performance function of the Mean Square Error (MSE) between the network output and corresponding target values. At each iteration, the gradient of performance function MSE was used to adjust the network weights and biases. In this work, an MSE of 10^{-5}, a minimum gradient of 10^{-10}, and maximum iteration number of 5000 were used. The training process got terminated when the error converged to a specified condition within the specified iteration. The initial weights and biases of the network were generated automatically by the program.

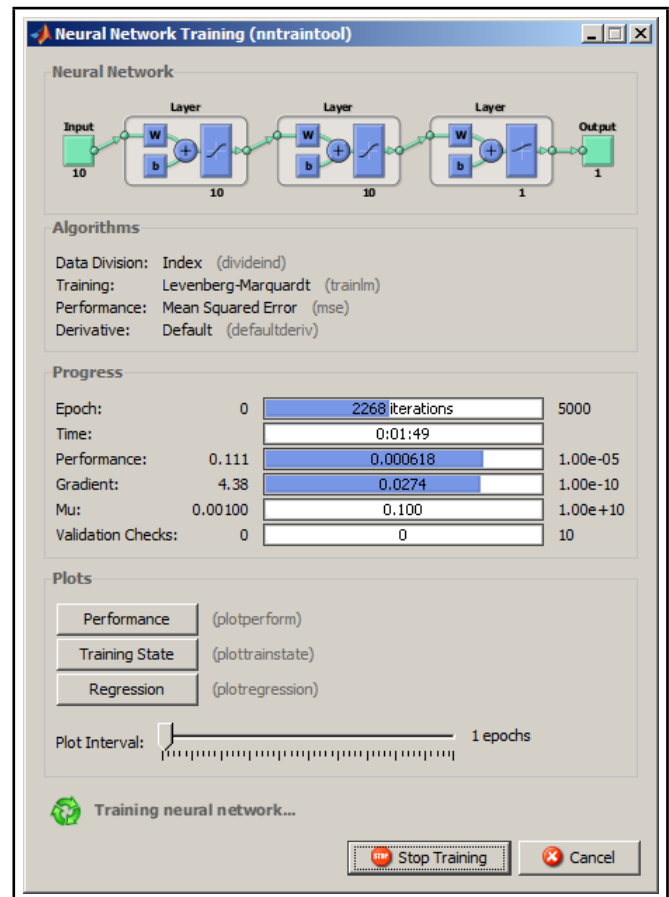The sample values of the input features for the training net-

Figure 6. ANN architecture with signal statistical features.

work are given in the Tables 2, 3, and 5. The training of ANN is carried with these features for different faults versus healthy gear condition as a two-class problem. The network architectures used in MATLAB for statistical and psychoacoustic features, respectively, are shown in Figs. 6 and 7. Figures 8(a) and 8(b), 9(a) and 9(b), to Figs. 10(a) and 10(b) are the sample regression plots of the ANN training, representing the correlation between the network output and target values for different faults and features. The value of correlation coefficient R is 0.97 for statistical features of acoustics signal and 0.99 for training with psychoacoustic features of acoustic signal as well as vibration signal. The value of the correlation coefficient greater than 0.9 indicate a good fit of the data and the training is perfect.

Similar plots were obtained for the other fault conditions. Similarly, the performance plots for all the cases obtained are observed to check the issue of overtraining and generalisation and it's observed that there is no chance of over-fitting and that
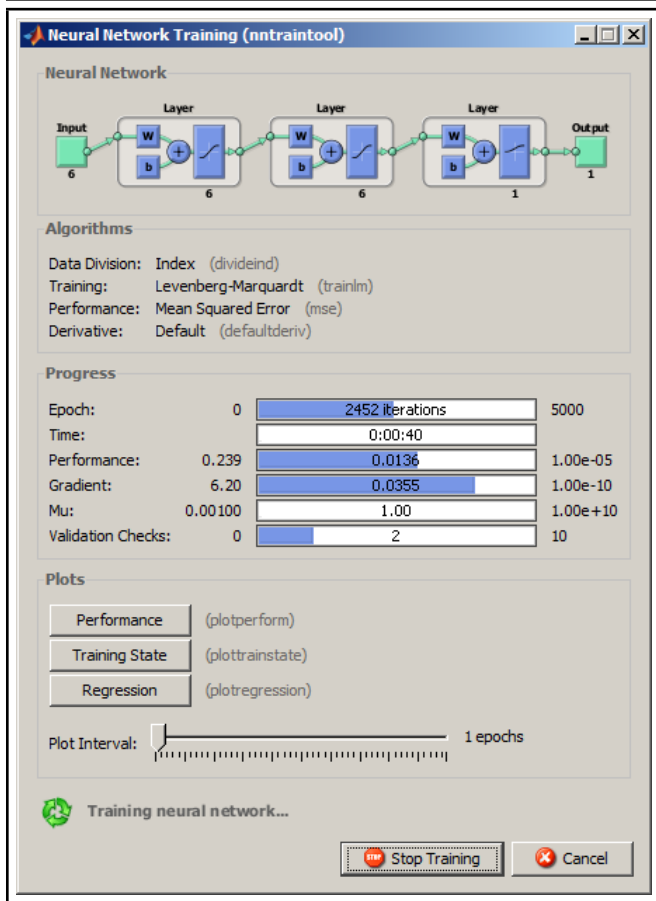
Figure 7. ANN architecture with psychoacoustic features.

generalisation is good.

4.2. Support Vector Machine

The Support Vector Machine (SVM) has gained popularity as a pattern recognition technique which is based on statistical learning theory as it overcomes the over-fitting and generalisation issue, as compared to the popular classifier such as ANN. Yang et al.,[40] Sarvan et al.,[41] Cheng et al.,[42] and Saidi et al.[24] have successfully demonstrated the application of SVMs for gear and bearing fault classification, using features extracted from the vibration signal such as higher order spectral features, wavelet coefficients, and features obtained by decomposing the signal using empirical mode decomposition etc. SVM has proved successful in other fields of pattern recognition such as finger print recognition, face detection, etc., as it can work with small sample size.

The data obtained from the healthy and other all faulty conditions in the gearbox are classified using SVM with a radial basis function, such as kernel. Healthy gear condition was represented by the target value of 0 and other all faults like a crack, profile error, broken tooth, and misalignment were assigned the target value of 1. It's not possible to pictorially present the scatter of the feature points used for the classification in multidimensional feature space, and hence the pair wise clustering is demonstrated by applying the SVM classifier. Pair wise comparison of the features is made and the pictorial representation showing clustering of data for classification purpose is presented for few features in Figs. 11, 12, and 13. From Fig. 11(a-c) and Fig. 13(a-c), it can be inferred that the statisti-

Table 6. Comparison of classification accuracy of ANN and SVM.

Gear Condition	Features	ANN			SVM		
		Training	Testing	Validation	Training	Testing	Validation
Crack	VSF	100	95.83	91.66	100	92.83	95.42
	ASF	94.59	94.64	88.21	97.29	98.21	92.32
	PF	98.90	95.65	95.65	97.53	93.44	95.08
Profile error	VSF	100	100	100	100	100	100
	ASF	98.75	100	95.08	100	100	100
	PF	100	100	96.72	100	100	100
Broken tooth	VSF	98.70	94.91	100	100	98.30	100
	ASF	100	90.16	93.44	100	91.80	91.80
	PF	100	100	100	100	100	100
Mis-align-ment	VSF	98.88	94.20	89.85	97.77	98.55	88.40
	ASF	88.88	89.85	85.42	94.44	89.85	85.50
	PF	97.10	97.80	96.10	98.90	94.20	95.65

Where VSF = Vibration Statistical Features, ASF = Acoustical Statistical Features, and PF = Psychoacoustic Features.

Table 7. Network statistics for ANN for misalignment error.

Features	No. of Hidden Layer	No. of Hidden Nodes	Training Correlation Coefficient	Training Accuracy	Validation Accuracy	Testing Accuracy	Corr	RMSE
VSF	1	10	0.96	97.77	89.85	97.10	0.94	0.17
	2	9	0.93	98.88	92.75	97.89	0.92	0.20
ASF	1	18	0.83	92.22	91.30	96.20	0.94	0.170
	2	10	0.82	91.11	92.75	97.10	0.89	0.24
PF	1	15	0.92	98.90	97.10	98.50	0.97	0.12
	2	09	0.98	98.95	97.20	98.55	0.98	0.11

cal features of the vibration signal and psychoacoustic features give better clustering of the features and hence, better diagnosis between the healthy and faulty gearboxes, compared to the statistical features of the acoustics signal (Fig. 12(a-c)). Few statistical features of the acoustics signal do not show distinct clustering in the two-dimensional representation for faults, like misalignment.

5. RESULTS AND DISCUSSION

ANN and SVM are trained to classify data with the psychoacoustic and statistical feature vectors of acoustic and vibration signal, as discussed earlier. The trained ANN and SVM functions are tested for the test data and the accuracy of classification, without performing the optimisation that is compared and shown in Table 6.

The Table 6 gives the comparison of the testing accuracy for data from four different faults with different features on which ANN and SVM are operating. Among the various features used, psychoacoustic features are giving the best results (93.44% to 100%) for various fault classification compared to statistical features of vibration (88.40% to 100%) and statistical features of acoustic signals (85.50% to 100%). These results are in line with the clustering observed for the statistical

Table 8. Optimised SVM for misalignment fault.

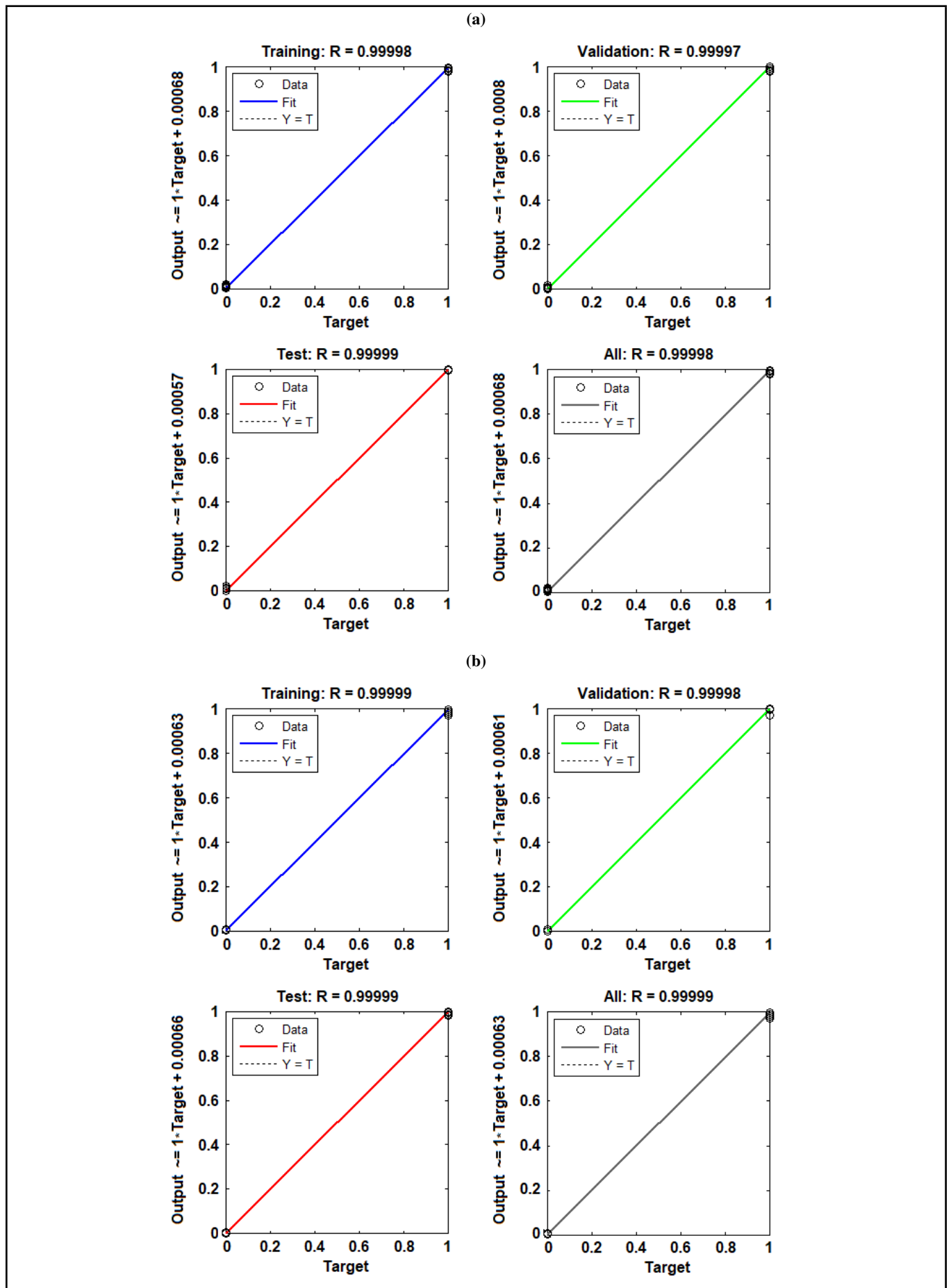| Features	RBF_Sigma	C	Training Accuracy	Validation Accuracy	Testing Accuracy	Corr	RMSE
VSF	0.0001	0.1	100	100	100	1	0
ASF	0.00001	0.01	100	100	100	1	0
PF	0.000001	0.01	100	100	100	1	0

Figure 8. Regression plot for: (a) statistical features of healthy and misaligned gear shaft from vibration signal; (b) statistical features of healthy and scuffed gear from vibration signal.
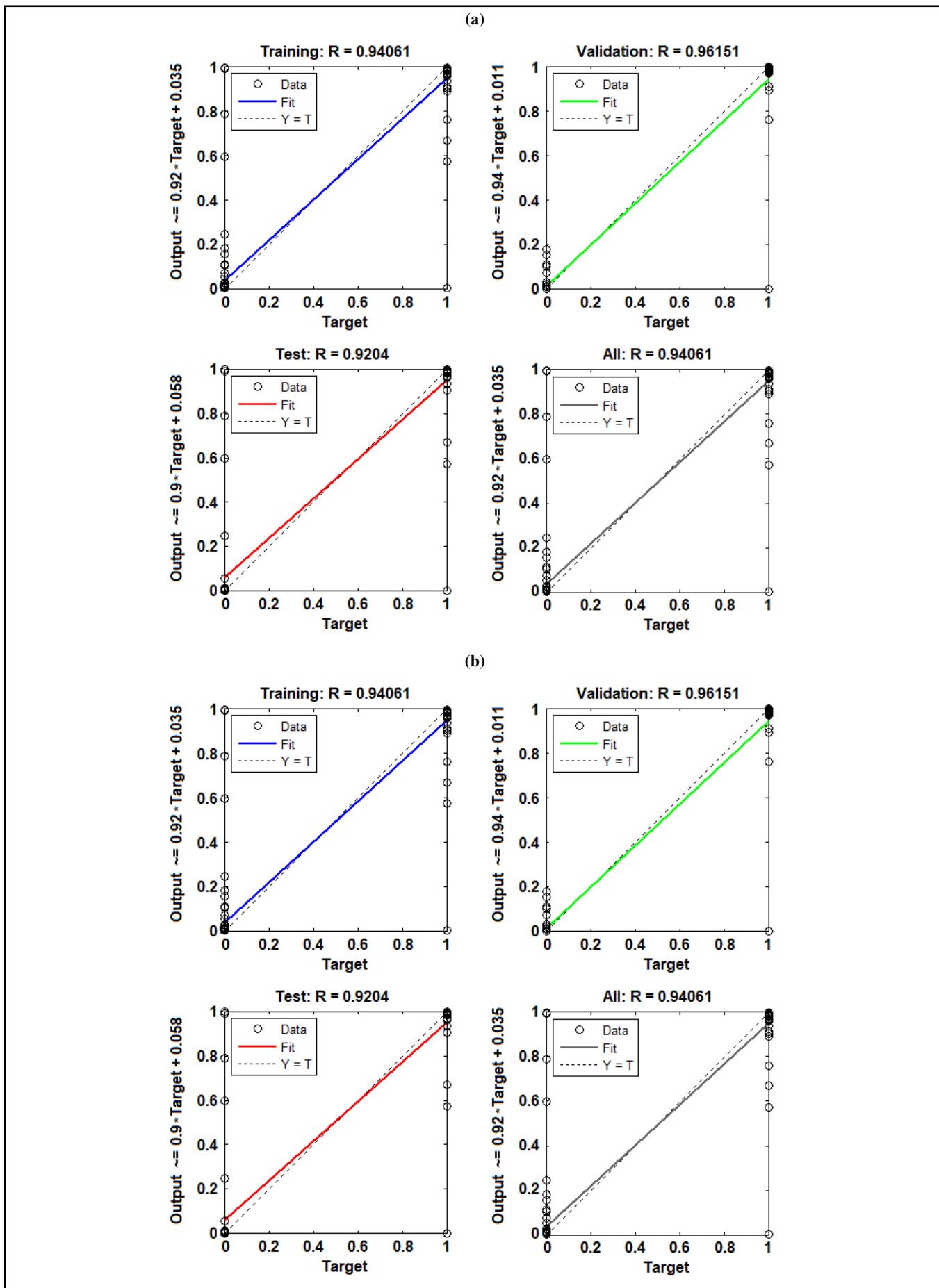
Figure 9. Regression plot for: (a) statistical features of acoustic signal of healthy and misaligned gear shaft; (b) statistical features of acoustic signal of healthy and scuffed gear.
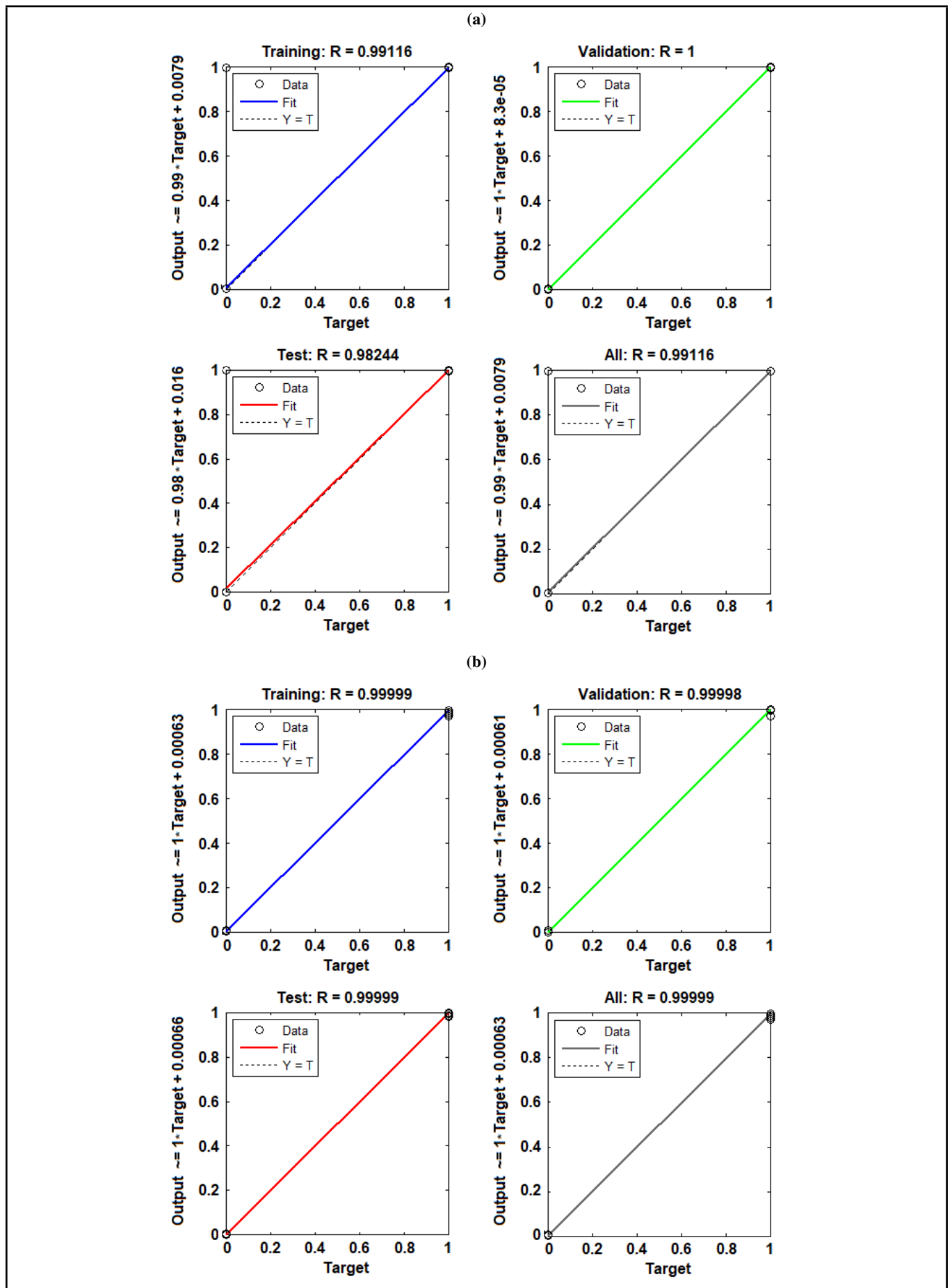
Figure 10. Regression plot for: (a) psychoacoustic features of healthy and misaligned gear shaft; (b) psychoacoustic features of healthy and scuffed gear.
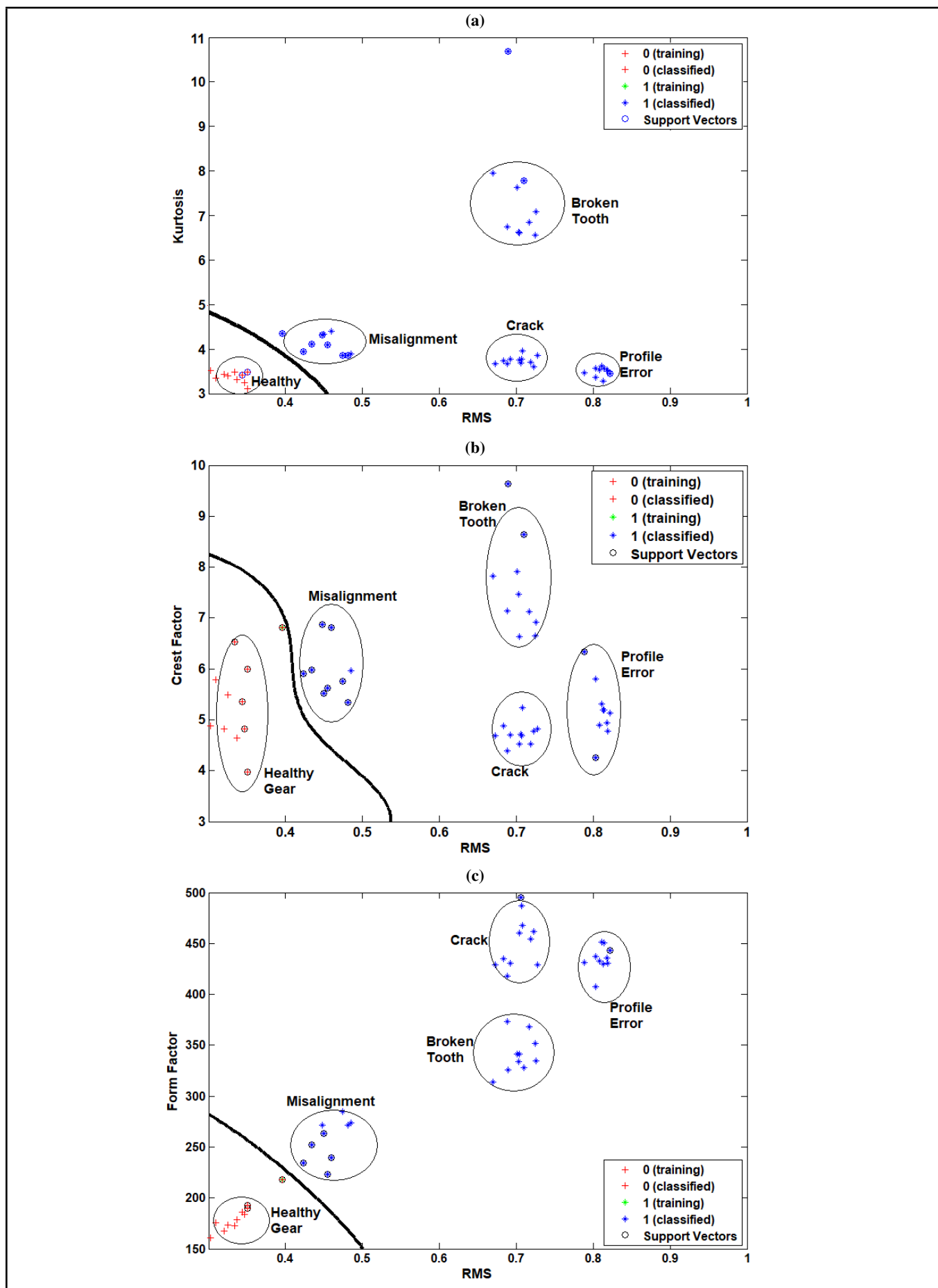
Figure 11. Pair wise clustering of statistical features of vibration signal with SVM classification: (a) kurtosis vs. RMS; (b) crest factor vs. RMS; (c) form factor vs. RMS.
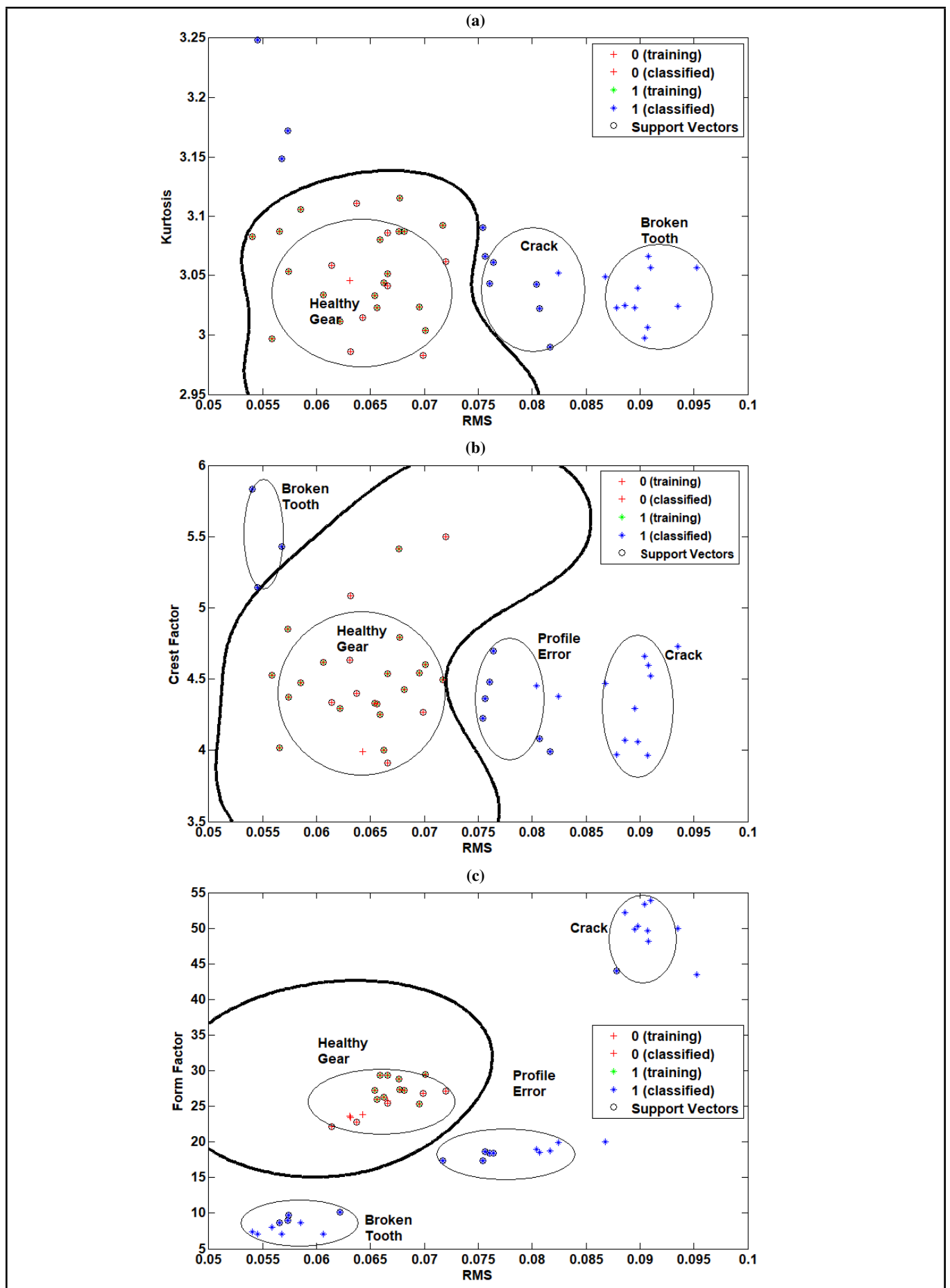
Figure 12. Pair wise clustering of statistical features of acoustic signal with SVM classification: (a) kurtosis vs. RMS; (b) crest factor vs. RMS; (c) form factor vs. RMS.
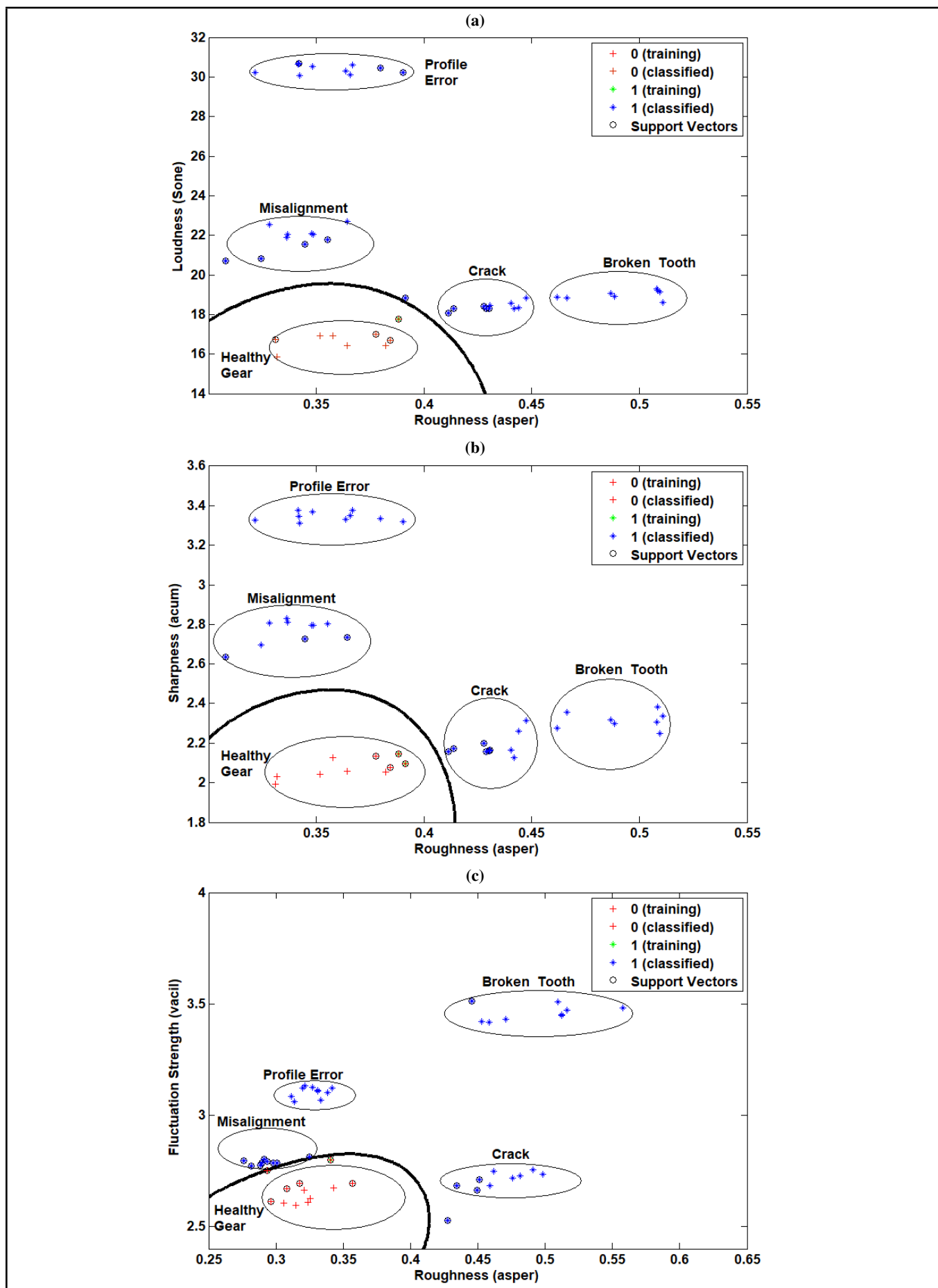
Figure 13. Pair wise clustering of psychoacoustic features of acoustic signal with SVM classification: (a) loudness vs. roughness; (b) sharpness vs. roughness; (c) fluctuation strength vs. roughness.

features of the vibration and acoustic signal and psychoacoustic features, presented in Figs. 11, 12, and 13. These results are based on the generalised program developed for ANN and SVM, as discussed in sections 4.1 and 4.2, which provides the basis to compare the classification ability of the different features' datasets.

To check whether these results can be further improved, the ANN architecture and SVM architecture were varied. This variation is done in the following manner. The ANN architecture was varied by changing the number of hidden layers and number of nodes in it and the fault classification accuracy was evaluated. The accuracy of the single hidden layer and the two hidden layers with number of hidden nodes varied from 1 to 20 is evaluated. Table 7 shows the sample of best results for the different features' datasets obtained for the fault of misalignment. In Table 7, the training correlation coefficients indicates the correlation between the input and target values during ANN training. Classification accuracy indicates the correctness with which the ANN can identify the target values. The values of Pearson's linear correlation coefficient, indicated by '*corr*' and root mean square error indicated by *RMSE*, is obtained for testing the datasets in Table 7 and these are obtained by following expressions

$$corr = \frac{\sum_{i=1}^{N}(t_p - \bar{t})(O_p - \bar{O})}{\sqrt{\sum_{i=1}^{N}(t_p - \bar{t})^2 \sum_{i=1}^{N}(O_p - \bar{O})^2}};$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(t_p - O_p)^2};$$

where N is the amount of data, t_p is the target values, O_p is the predicated values, $\bar{t}$ and $\bar{O}$ are the averages of the target and predicted values, respectively.

Similarly, the SVM architecture was also modified by varying variable C and RBF_Sigma (also termed as gamma in literature). A standard SVM seeks to find the margin that separates all the positive and negative data points and the results of a standard SVM are demonstrated in Table 6. However, this can lead to poorly fit models if any data points are mislabelled or extremely unusual. To account for this, in 1995, Cortes and Vapnik proposed the ideas of a "soft margin" SVM that allows some data points to be ignored or placed on the wrong side of margin; this innovation often leads to a better overall fit. C is the parameter for the soft margin cost function, which controls the influence of each individual support vector; this process involves the trading error penalty for stability. C is a cost of classification and a large C gives low bias and high variance. Low bias as the cost of misclassification is penalised a lot. RBF_Sigma is the parameter of a Gaussian Kernel, which controls the position of data points in higher dimensions. Small RBF_Sigma gives low bias and high variance and large RBF_Sigma gives higher bias and low variance. The best values of the C and RBF_Sigma hyperparameters are obtained using the Grid-Search method.[43,44] Table 8 shows the values of the parameters RBF_Sigma and C, for which the best classification accuracy results are observed for the datasets of a misalignment fault compared to a healthy gearbox condition. The training, validation, and testing accuracies are obtained within the range of 62.31% to 100% for the values of C and

RBF_Sigma varying from 0.00001 to 10000 on a logarithmic scale for the feature sets presented in Table 8.

From the results obtained in Table 7, it can be observed that there is a marginal change in the results of the ANN by changing its architecture compared with the results of Table 6 for the generalised program developed to compare the fault classification accuracy of different feature datasets. From the results of Table 8, reasonable improvement in fault classification accuracy is obtained by varying the C and RBF_Sigma parameters, compared to the standard SVM.

ANN and SVM with psychoacoustic features are performing the best as compared to the other datasets. Psychoacoustic features are derived from the variation in time as well as the frequency domain and are analogous to the human ability to listen. Different faults manifest differently in the form of sound waves and psychoacoustic features can characterize them effectively, compared to the simple statistical features. Loudness is based on amplitude of the frequency content of the acoustic signal, sharpness indicates a variation in loudness at high frequencies, while roughness characterises temporal deviations due to the frequency modulation of 20 Hz to 300 Hz present in sound pressure. Fluctuation strength characterises the temporal deviation due to low frequency modulation in 0.25 Hz to 20 Hz. The psychoacoustic features can characterise the change in the health of a gearbox, analogous to human hearing, as they are based on science of hearing. Therefore, the psychoacoustic features are performing better.

6. CONCLUSIONS

This paper attempts to provide a solution to the issue of hearing-based end of assembly line inspection during the manufacturing of a gearbox. Various faults in gear assembly are simulated in a laboratory on a test setup. Vibration and acoustic signals are acquired from the gearbox under various conditions. A novel idea of applying psychoacoustics features is attempted for fault identification. Psychoacoustics and signal statistical features are extracted from the acquired signals of the gearbox. These features are then used as inputs for ANN and SVM for fault classification. The accurate results of classification of ANN and SVM are compared with the proposed psychoacoustic features and conventional statistical features of the vibration and acoustic signal. An attempt was made to improve the performance of ANN and SVM by changing their architecture. However, it resulted in marginal improvement in ANN and significant improvement in performance of SVM. From the comparison of various features used, it is observed that for the combinations of datasets (four faulty and one good gear condition), the classification with psychoacoustic features has an edge over the others. Hence, it can be concluded that the psychoacoustic features can capture the variation in the acoustic signal more effectively than the statistical signal features, and it can objectively ascertain the presence of a fault along with ANN and SVMs. The proposed method can overcome the issue of the subjectivity involved in the inspection process and can help the operator to make the decision objectively during the end of the line inspection in the manufacturing environment. The process of fault inspection can be automated using artificial intelligence techniques, like ANN or SVM, along

with the psychoacoustic features to facilitate more objective and intelligent decision making.

REFERENCES

[1] Mohanty, A. R. *Machinery Condition Monitoring Principles and Practices*, CRC Press, Taylor & Francis Group, (2015).

[2] Heng, A., Zhang, S., Tan, A. C. C., and Mathew, J. Rotating machinery prognostics: State of the art, challenges and opportunities, *Mechanical Systems and Signal Processing*, **23** (3), 724–739, (2009). https://dx.doi.org/10.1016/j.ymssp.2008.06.009

[3] Jardine, A. K. S., Lin, D., and Banjevic, D. A review on machinery diagnostics and prognostics implementing condition-based maintenance, *Mechanical Systems and Signal Processing*, **20** (1), 1483–1510, (2006). https://dx.doi.org/10.1016/j.ymssp.2005.09.012

[4] Bajrić, R., Sprečić, D., and Zuber, N. Review of vibration signal processing techniques towards gear pairs damage identification, *International Journal of Engineering & Technology*, **11** (4), 124–128, (2011).

[5] Worden, K., Staszewski, W. J., and Hensman, J. J. Natural computing for mechanical systems research: A tutorial overview, *Mechanical Systems and Signal Processing*, **25** (1), 4–111, (2011). https://dx.doi.org/10.1016/j.ymssp.2010.07.013

[6] Zhang, X., Wen, G., and Wu, T. A new time synchronous average method for variable speed operating condition gearbox, *Journal of Vibroengineering*, **14** (4), 1766–1774, (2012).

[7] Sharma, V. and Parey, A. Gear crack detection using modified TSA and proposed fault indicators for fluctuating speed conditions, *Measurement*, **90**, 560–575, (2016). https://dx.doi.org/10.1016/j.measurement.2016.04.076

[8] El-morsy, M. S., Abouel-seoud, S., and Rabeih, E. Gearbox damage diagnosis using wavelet transform technique, *International Journal of Acoustics and Vibration*, **16** (4), 173–179, (2011). https://dx.doi.org/10.20855/ijav.2011.16.4294

[9] Peng, Z. K. and Chu F. L. Application of the wavelet transform in machine condition monitoring and fault diagnostics: a review with bibliography, *Mechanical Systems and Signal Processing*, **18**, 199–221, (2004). https://dx.doi.org/10.1016/s0888-3270(03)00075-x

[10] Cheng, G., Cheng, Y., Shen, L., Jin-bo Q., and Zhang, S. Gear fault identification based on Hilbert–Huang transform and SOM neural network, *Measurement*, **46** (3), 1137–1146, (2013). https://dx.doi.org/10.1016/j.measurement.2012.10.026

[11] Jinlu, S., Shaojiang, D., and Zhu, L. Bearing fault diagnosis based on intrinsic time scale decomposition and improved Support vector machine model, *Journal of Vibroengineering*, **18** (2), 849–859, (2016).

[12] Cheng, J., Yang, Y., and Yu, D. The envelope order spectrum based on generalized demodulation time–frequency analysis and its application to gear fault diagnosis, *Mechanical Systems and Signal Processing*, **24** (2), 508–521, (2010). https://dx.doi.org/10.1016/j.ymssp.2009.07.003

[13] Nacib, L., Midzodzi Pekpe, K., and Sakhara, S. Detecting gear tooth cracks using cepstral analysis in gearbox of helicopters, *International Journal of Advances in Engineering & Technology*, **5** (2), 139–145, (2013).

[14] Jena, D. P. and Panigrahi, S. N. Gear fault diagnosis using bispectrum analysis of active noise cancellation-based filtered sound and vibration signals, *International Journal of Acoustics and Vibration*, **18** (2), 58–70, (2013). https://dx.doi.org/10.20855/ijav.2013.18.2320

[15] Ulus, S. and Erkaya, S. An experimental study on gear diagnosis by using acoustic emission technique, *International Journal of Acoustics and Vibration*, **21** (1), 103–111, (2016). https://dx.doi.org/10.20855/ijav.2016.21.1400

[16] Tandon, N. and Choudhury, A. A review of vibration and acoustic measurement methods for the detection of defects in rolling element bearings, *Tribology International*, **32** (8), 469–480, (1999). https://dx.doi.org/10.1016/s0301-679x(99)00077-8

[17] Vakharia, V., Gupta, V. K., and Kankar, P. K. Ball bearing fault diagnosis using supervised and unsupervised machine learning methods, *International Journal of Acoustics and Vibration*, **20** (4), 244–250, (2015). https://dx.doi.org/10.20855/ijav.2015.20.4387

[18] Singh, S. and Kumar, N. Rotor faults diagnosis using artificial neural networks and support vector machines, *International Journal of Acoustics and Vibration*, **20** (3), 153–159, (2015). https://dx.doi.org/10.20855/ijav.2015.20.3379

[19] Samanta, B. Gear fault detection using artificial neural networks and support vector machines with genetic algorithms, *Mechanical Systems and Signal Processing*, **18** (3), 625–644, (2004). https://dx.doi.org/10.1016/s0888-3270(03)00020-7

[20] Shen, C., Wang, D., Kong, F., and Tse, P. W. Fault diagnosis of rotating machinery based on the statistical parameters of wavelet packet paving and a generic support vector regressive classifier, *Measurement*, **46** (4), 1551–1564, (2013). https://dx.doi.org/10.1016/j.measurement.2012.12.011

[21] Saravanan, N., Siddabattuni, V. N. S. K., and Ramachandran, K. I. Fault diagnosis of spur bevel gear box using artificial neural network (ANN), and proximal support vector machine (PSVM), *Applied Soft Computing*, **10** (1), 344–360, (2010). https://dx.doi.org/10.1016/j.asoc.2009.08.006

[22] Kumar, H., Kumar, T. A. R., Amarnath, M., and Sugumaran, V. Fault diagnosis of antifriction bearings through sound signals using support vector machine, *Journal of Vibroengineering*, **14** (4), 1601–1606, (2012).

23 Mohammad, H., Hadi, H., Hossein, G.,and Ali, H. Fault diagnosis of gearboxes using wavelet support vector machine, least square support vector machine and wavelet packet transform, *Journal of Vibroengineering*, **18** (2), 860–875, (2016).

24 Saidi, L., Ali, J. B., and Fnaiech, F. Application of higher order spectral features and support vector machines for bearing faults classification, *ISA Transactions*, **54**, 193–206, (2015). https://dx.doi.org/10.1016/j.isatra.2014.08.007

25 Shang, W., Zhou, X., and Yuan J. An intelligent fault diagnosis system for newly assembled transmission, *Expert Systems with Applications*, **41** (9), 4060–4072, (2014). https://dx.doi.org/10.1016/j.eswa.2013.12.045

26 Cook, V. G. C. and Ali, A. End-of-line inspection for annoying noises in automobiles: Trends and perspectives, *Applied Acoustics*, **73** (3), 265–275, (2012). https://dx.doi.org/10.1016/j.apacoust.2011.06.019

27 Kane, P. V. and Andhare, A. B. Application of psychoacoustics for gear fault diagnosis using Artificial Neural Network, *Journal of Low Frequency Noise, Vibration and Active Control*, **35** (3), 207–220, (2016). https://dx.doi.org/10.1177/0263092316660915

28 Sharma, V. and Parey, A. A review of gear fault diagnosis using various condition monitoring indicators, Proceedings of ICOVP 2015, *Procedia Engineering*, **144**, 253–263, (2016). https://dx.doi.org/10.1016/j.proeng.2016.05.131

29 Večeř, P., Kreidl, M., and Šmíd, R. Condition indicators for gearbox condition monitoring systems, *ActaPolytechnica*, **45**, 35–43, (2005).

30 Lebold, M., McClintic, K., Campbell, R., Byington, C., and Maynard, K. Review of vibration analysis methods for gearbox diagnostics and prognostics, *Proceedings of the 54th Meeting of the Society for Machinery Failure Prevention Technology*, Virginia, VA, (2000).

31 Zwicker, E. and Fastl, H. *Psychoacoustics: Facts and models*, Second Edition, Springer-Verlag Berlin Heidelberg, New York, (2007). https://dx.doi.org/10.1063/1.1387599

32 NI-Tutorial-1526, White paper on Measurement of Sound Quality, June 13, 2013. Retrieved from http://www.ni.com/white-paper/1256/en/ (Accessed April 12, 2014)

33 Technical note, An Introduction to Sound Quality Testing, Acoustic Research Centre, School of Computing, Science and Engineering, University of Salford, Manchester, UK, October 10, 2015. Retrieved from http://www.salford.ac.uk/computing-science-engineering/research/acoustics/psychoacoustics/sound-quality-making-products-sound-better/sound-quality-testing/roughness-fluctuation-strength (Accessed December 15, 2014)

34 Kim, E. Y., Lee, Y. J., and Lee, S. K. Sound metric design for evaluation of tonal sound in laser printer, *International Journal of Precision Engineering and Manufacturing*, **13** (8), 1349–1358, (2012). https://dx.doi.org/10.1007/s12541-012-0178-0

35 Fastl, H. Psychoacoustic basis of sound quality evaluation and sound engineering, *Proceedings of The Thirteenth International Congress on Sound & Vibration*, July 2–6, Vienna, Austria, (2006).

36 Fastl, H. The psychoacoustics of sound quality evaluation, *Acta Acustica united with Acustica*, **83** (5), 754–764, (1997).

37 Munakata, T. *Fundamentals of the New Artificial Intelligence*, Second edition, Springer-Verlag London Limited, (2008). https://dx.doi.org/10.1007/978-1-84628-839-5

38 Palit, A. K. and Popvik, D. *Computational Intelligence in Time Series Forecasting. Theory and Engineering Applications*, First edition, Springer-Verlag London Limited, (2005). https://dx.doi.org/10.1007/1-84628-184-9

39 Demuth, H. and Beale, M. *Neural Network Toolbox for use with MATLAB*, Version 4.0, Math Work Inc., (2002).

40 Yang, J., Zhang, Y., and Zhu, Y. Intelligent fault diagnosis of rolling element bearing based on SVMs and fractal dimension, *Mechanical Systems and Signal Processing*, **21** (5), 2012–2024, (2007). https://dx.doi.org/10.1016/j.ymssp.2006.10.005

41 Saravanan, N., Siddabattuni, V. N. S. K., and Ramachandran, K. I. 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*, **35** (3), 1351–1366, (2008). https://dx.doi.org/10.1016/j.eswa.2007.08.026

42 Cheng, J., Yu, D., and Yang, Y. A fault diagnosis approach for gears based on IMF AR model and SVM, *EURASIP Journal on Advances in Signal Processing*, **2008**, 1–7, (2008). https://dx.doi.org/10.1155/2008/647135

43 Lin, S. W., Lee, Z. J., Chen, S. C., and Tseng, T. Y. Parameter determination of support vector machine and feature selection using simulated annealing approach, *Applied Soft Computing*, **8** (4), 1505–1512, (2008). https://dx.doi.org/10.1016/j.asoc.2007.10.012

44 A technical note on RBF SVM Parameters. Retrieved from http://scikit-learn.org/stable/auto_examples/svm/plot_rbf_parameters.html