

Article

A New Statistical Features Based Approach for Bearing Fault Diagnosis Using Vibration Signals

Muhammad Altaf ^{1,†}, Tallha Akram ^{1,†}, Muhammad Attique Khan ², Muhammad Iqbal ¹,
M Munawwar Iqbal Ch ³ and Ching-Hsien Hsu ^{4,5,6,*}

¹ Department of Electrical and Computer Engineering, COMSATS University Islamabad, Wah 47000, Pakistan; mohammadaltaf@gmail.com (M.A.); tallha@ciitwah.edu.pk (T.A.); miqbal1976@gmail.com (M.I.)

² Department of Computer Sciences, HITEC University Taxila, Taxila 47080, Pakistan; attique.khan@hitecuni.edu.pk

³ Institute of Information Technology, Quaid-i-Azam University, Islamabad 44000, Pakistan; mmic@qau.edu.pk

⁴ Department of Computer Science and Information Engineering, Asia University, Taichung 400-439, Taiwan

⁵ Department of Medical Research, China Medical University Hospital, China Medical University, Taichung 400-439, Taiwan

⁶ Guangdong-Hong Kong-Macao Joint Laboratory for Intelligent Micro-Nano Optoelectronic Technology, School of Mathematics and Big Data, Foshan University, Foshan 528000, China

* Correspondence: chh@cs.ccu.edu.tw

† These authors contributed equally to this work.



Citation: Altaf, M.; Akram, T.; Khan, M.A.; Iqbal, M.; Ch, M.M.I.; Hsu, C.-H. A New Statistical Features Based Approach for Bearing Fault Diagnosis Using Vibration Signals. *Sensors* **2022**, *22*, 2012. <https://doi.org/10.3390/s22052012>

Academic Editors: Rafal Burdzik, Minvydas Ragulskis, Maosen Cao, Radosław Zimroz, Chaari Fakher, Łukasz Konieczny and Giovanni Betta

Received: 21 November 2021

Accepted: 23 December 2021

Published: 4 March 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: In condition based maintenance, different signal processing techniques are used to sense the faults through the vibration and acoustic emission signals, received from the machinery. These signal processing approaches mostly utilise time, frequency, and time-frequency domain analysis. The features obtained are later integrated with the different machine learning techniques to classify the faults into different categories. In this work, different statistical features of vibration signals in time and frequency domains are studied for the detection and localisation of faults in the roller bearings. These are later classified into healthy, outer race fault, inner race fault, and ball fault classes. The statistical features including skewness, kurtosis, average and root mean square values of time domain vibration signals are considered. These features are extracted from the second derivative of the time domain vibration signals and power spectral density of vibration signals. The vibration signal is also converted to the frequency domain and the same features are extracted. All three feature sets are concatenated, creating the time, frequency and spectral power domain feature vectors. These feature vectors are finally fed into the K- nearest neighbour, support vector machine and kernel linear discriminant analysis for the detection and classification of bearing faults. With the proposed method, the reduction percentage of more than 95% percent is achieved, which not only reduces the computational burden but also the classification time. Simulation results show that the signals are classified to achieve an average accuracy of 99.13% using KLDA and 96.64% using KNN classifiers. The results are also compared with the empirical mode decomposition (EMD) features and Fourier transform features without extracting any statistical information, which are two of the most widely used approaches in the literature. To gain a certain level of confidence in the classification results, a detailed statistical analysis is also provided.

Keywords: vibration signal analysis; condition based maintenance; time domain analysis; frequency domain analysis; machine learning; classification

1. Introduction

Rotating machinery using bearings is one of the most important components of a wide range of mechanical setups from small motors to turbines, compressors and heavy ground and air vehicles [1,2]. Different faults arise during the mechanical and industrial process, generating vibration and Acoustic Emission (AE) signals [3,4]. These signals have different characteristics due to the nature of faults, the complexity of the underlying industrial

setup and the correlation between different mechanical components [5,6]. Preemptive measure and real time monitoring of these signals can avoid severe losses and disastrous failures [7,8] and hence has received considerable attention [9]. Different methods of condition based monitoring are being developed and used, including but not limited to oil debris, vibration, acoustic emission, electrostatic and temperature analysis [10–13].

Early detection and localisation of mechanical faults, such as misalignment, gear faults mass unbalance and cracks along with their propagation in rotating shafts and gear wheels are possible by analysing vibration [14] and acoustic emission [15] signals, with AE having a much wider range of frequencies as compared to vibration signal [16]. The AE signals are generally transient elastic waves resulting from fast strain energy discharge due to damage on or within the material surface [17].

Both the AE and vibration signals can effectively be used for the detection and localization of defects in rotating machinery. However, the AE signal outperforms the vibration signal in case early and preemptive detection is required and also in fault detection in low speed rotating machines due to the limited efficiency of vibration signals as compared to AE signals [10]. These signals are non-stationary in nature and are complicated to analyse due to the heavy background noise of industrial set up [18]. Therefore, state-of-the-art signal processing algorithms and pattern recognition techniques are combined to monitor these signals for the detection and classification of faults. The signal processing is used to extract different time domain, frequency domain and time-frequency domain features of these signals for use with Artificial Intelligence (AI) techniques [8,18,19].

Statistical analysis of time domain AE signals have detected shaft cracks using features like energy, duration count and average signal level [20]. A similar approach is used in [21], where the amplitude and energy of the AE signals are used for defects in roller bearing. Abdullah et al. [22] detected the bearing defects and its sizes from both the time domain AE and vibration signals, using its amplitude and the root mean square (RMS) values. Time domain and wavelet domain kurtosis of the vibration signal, along with the Largest Lyapunov Exponent (LLE), were used in [19] for the analysis of slew bearing defects. These features were fed to kernel based regression for detection of the damage and estimation of the useful life of the bearings under testing. Antoni et al. [23] used a fast kurtogram for the detection of transient faults with a similar computational complexity to that of Fast Fourier Transform (FFT). Kurtosis and its different variations, such as kurtogram, spectral kurtosis, adaptive spectral kurtosis, and Short Term Fourier Transform (STFT) based kurtosis, have been used extensively by the research community for the analysis of vibration signals from rotating machinery; interested readers are referred to [24].

Feature extraction is one of the most key factors in reducing the classification errors; however, recent developments in the field of Deep Neural Networks that have the ability to extract features from original signals are being used in fault diagnosis [25–28]. That is why deep learning is swiftly paving the way for condition based maintenance with reproducible results. Mingyong et al. used Convolutional Neural Networks (CNN) with time domain vibration signals for fault diagnosis, with 96% accuracy on Case Western Reserve University (CWRU) data, and classified the faults into rolling element, inner and outer ring faults [29]. Inner and outer race faults are detected using the root mean square measure of the vibration signal and adaptive neuro-fuzzy inference system with an accuracy of 98.37% [30]. The authors proposed a feature learning approach in [31] using vibration and current signal with deep CNN, claiming an accuracy of 98% if the same machine is used for feature learning and testing, and an accuracy of 92% if the model trained on one machine is used for testing another machine. The algorithm can also provide other information such as rotational speed and number of balls and so forth.

The Fourier transform is used for analysis of signals in frequency domain [32]; however, it lacks time information. Therefore, to provide local feature information, a time-frequency technique like STFT is used to calculate the local features using the windowing approach [33]. In the time-frequency domain, the Wavelet Transform is one of the most appropriate approaches for the detection and localisation of faults [34–37]. Other time-frequency approaches to detect abnormal conditions such as misalignment, rotor-stator rub

and shaft cracks in the vibration signals, including but not limited to Continuous Wavelet Transform (CWT) and Hilbert Huang Transform (HHT). The latter is also used to effectively monitor defects in bearing [38], while the former is used for early stage fault detection in the outer race [6]. Envelop detection with autocorrelation is used to detect faulty patterns at the inception stage in low SNR (Signal to Noise Ratio) AE signals [6], with wavelet transform for de-noising of the AE signals. Spectral components of the Intrinsic Mode Functions (IMF) are used to analyse both the AE and vibration signals for the detection of bearing defects, broken bar and unbalanced load distribution in the indoor motors [39]. However, it needs the a priori knowledge of the number of modes in which the signal is required to be decomposed. Features from raw vibration signals were extracted by [8], using Empirical Mode Decomposition algorithm and CNN. These features were fed to SVM and CNN training algorithms for the classification of faults into the outer race, inner race and ball faults. In another study, Khan et al. used EMD with KNN to classify the different machine states into normal, cracking, offset pulley and wear states [40]. In [41], the EMD is combined with deep neural networks to classify the faults into roller, inner and outer races with an accuracy of 98.5%. The authors have detected and classified faults using ensemble empirical mode decomposition (EEMD) and SVM. The EMD is used to decompose the vibration signal, isolating and denoising high frequency IMFs using Pearson correlation coefficients and the wavelet semi-soft threshold, respectively. The Eigen vector of the signal is used as a feature vector for the SVM to classify faults into inner race, outer race and rolling elements, with an accuracy of 100% [42]; however, the 100% accuracy seems to be on the higher end. Delprete et al. [43] used orthogonal empirical mode decomposition analysis (a time-frequency) to detect faults in the inner and outer raceway of the bearings using vibration signals.

Informative signatures can be extracted from vibration signals in the form of wavelets transform. Morlet wavelet transform is used to extract features from vibration signals that were then used with artificial neural networks and SVM to classify the signals into ball fault, inner and outer race fault, with 95.271% accuracy for SVM and 87.25% for ANN [44]. The authors of [45] combined features like vibration severity, dyadic wavelet energy time-spectrum and coefficients power spectrum of maximum wavelet energy level and fed it to SVM for classification into a normal state, eccentric axle fault (EAF), bearing pedestal fault (BPF), and sealing ring wear fault. The SVM parameters were optimized using the modified shuffled frog-leaping algorithm, claiming a maximum accuracy of 96% [45]. A Deep Fault Diagnosis (DFD) method is proposed for rotating machinery with scarce data labels. In this procedure, discriminative STFT data are obtained from the spectrogram of the vibration signals. Several SVM models are trained with different features selected from the pool with scarce labels and the most discriminative features and the best SVM models were selected, hence forming an augmented training set. This augmented training set was then forwarded to a 2-D deep CNN. The proposed algorithm classified different faults with an accuracy of 98.4% [46].

The research community has also used image analysis for diagnosing faults; here, image sparse representation is proposed to extract meaningful features from redundant information present in images using orthogonal matching pursuit and K-singular value decomposition algorithms along with 2-D PCA. The features are used with a minimum distance to classify into the inner race, outer race and ball faults with an average accuracy of 99.92% [47]. In [48], Moussa et al. proposed an algorithm for bearing fault diagnosis based on the probability of image recognition techniques under constant and variable speed conditions using average PCA. The paper used vibration spectrum imaging of the vibration signal obtained from faulty and normal bearings with CNN for classification. In [49], the time domain vibration signal was first segmented using a time-moving segmentation window and was then transformed into a spectral image for training and testing with CNN. The proposed scheme provided good accuracy at different levels of noises and speeds. Similarly, the time domain, frequency domain and time-frequency domain parameters are used to detect faults in axial piston pumps with deep belief networks [50].

The literature review as discussed above does not cover the complete review of the topic; however, it provides quite a clear picture of the work being contributed by the research community. The statistical approach, frequency domain and time–frequency domain provide different levels of accuracy with conventional machine learning techniques and deep neural networks. In this research work, the statistical and frequency domain techniques are analysed to detect and classify the faults in rotating machinery using conventional machine learning algorithms. The vibration signal from the faulty and healthy rotating machinery was recorded. From these vibration signals, the statistical features, such as Skewness, Kurtosis, Average and root mean square (RMS) values of time domain vibration signals, are considered. The same features were then extracted from the second derivative of the time domain vibration signals, making a vector of eight features. For the second and third steps, these features were calculated by first taking the Fourier Transform and Power Spectral Density of the vibration signals. All three sets of feature vectors were concatenated creating time domain, frequency domain and spectral power domain feature vectors. SVM, KNN and KLDA classifiers are used for the classification of signals into outer race fault, inner race fault, ball fault and healthy signal. Simulation results showed that the KLDA resulted in an accuracy of 99.13% for our proposed method using PSD, followed by the statistical features with KLDA giving an accuracy of 98.275%, showing the strength of our proposed algorithms.

These results were then compared with those of EMD, which has a very good accuracy as given in the literature review discussed above. In this case, the accuracy of the EMD feature vector was 97.01% with SVM. It is important to note that the size of the feature vector for EMD is $14 \times 160,000$ in this particular case and that of the Statistical, FT and PSD is 8×228 . This also shows the effectiveness of the proposed approach in terms of size of the feature vector and hence less computation.

The primary contributions of this research are enumerated below:

1. We exploit the behaviour of feature extraction based on the performance of selected classifiers;
2. We propose to utilise the statistical features including kurtosis, skewness, average and root mean square of the selected window;
3. To extract the more meaningful information, we extracted the same features from the second derivative;
4. We utilise both frequency and time domain signal information for feature extraction—employing the moving window concept;
5. We propose a system which generates an accuracy of more than 95%—utilising less than 5% of information and achieving the reduction percentage of 95%.

In the rest of the paper, signal processing techniques for extracting the proposed features and calculating FFT and PSD and so forth are discussed in Section 2. The discussion of these features for fault detection is given Section 3. The vibration signals are analysed using signal graphs of different faults. A discussion on the classifier and classification of faults into outer race, inner race, and ball faults is presented in Section 4, with the concluding remarks given in Section 5.

2. Feature Extraction

In this section, the digital signal processing (DSP) techniques that are used for feature extraction are discussed. These techniques can be categorised under three main headings, the statistical features of time domain signal, the statistical features of signal in Fourier domain and the statistical features of signal's Power Spectral Density. The raw vibration data was used to extract statistical features like maximum value, minimum value, standard deviation, mean, median, variance, skewness, kurtosis, range, Fisher Information Ratio [51,52], Petrosian Fractal Dimension [51], and entropy. The results of skewness, kurtosis and standard deviation and second derivative and so forth, to name a few. The features that were selected are mean value, standard deviation, skewness, kurtosis and second derivative. These features were calculated for time domain vibration signal, for the signal in Fourier

domain and for the same signal after calculating its PSD. These features were then fed to the SVM, KNN and KLDA algorithms for classification into outer race fault, inner race fault, and ball fault as given in the block diagram of Figure 1. As is shown, the raw data are extracted, pre-processed and its statistical features like skewness, kurtosis, average and RMS values are calculated. It is important to note that extracting these statistical features in the Fourier domain and from the PSD of the vibration signal is not used previously, to the best of our knowledge. This proposed method has also reduced the number of data points and hence computational requirements as discussed in Section 4. The vibration signal of healthy and faulty bearings were recorded, using vibration sensors, shaft rotating at a rate of 800 revolutions per minute and a sampling rate of 40,000 samples per second. Figure 2 shows the specifications and block diagram of the test rig with ball bearing model, fault types and specifications of the data acquisition board.

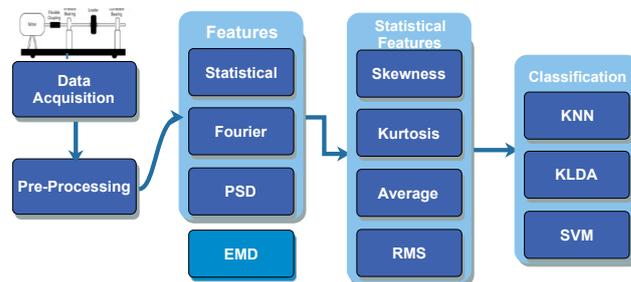
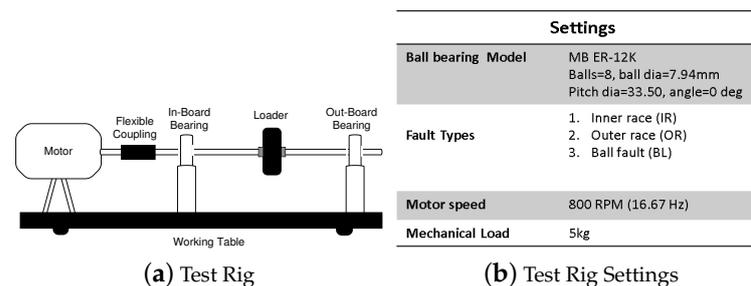


Figure 1. Block Diagram of Proposed Procedure.



(a) Test Rig

(b) Test Rig Settings

Instrumentation Settings	
DAQ board	NI PCI-4472B, 24-bit
Sampling rate	40000 Samples/Sec
Computer	i7 Computer , 8GB RAM
Op. System	Windows 7

(c) Instrumentation Settings

Figure 2. Test rig block diagram and specifications.

The equations for skewness, kurtosis, standard deviation and second derivative are given below. The skewness is a measure of symmetry or the lack of symmetry, and is zero for the data with normal distribution and should be zero for any symmetric data. The kurtosis is a measure of whether the data are heavy-tailed or light-tailed relative to a normal distribution. Both provide important statistical characteristics of a population. The mean and the standard deviation provide the central value; however, some times it is required to find how far the data are spread, and that is measured by variance and standard deviation. The mathematical representation of all these statistical measures is given here for ready reference. For a 1-Dimensional data Y_1, Y_2, \dots, Y_N , the mathematical representation of these statistical measures is given as:

$$\text{Mean} = \bar{X} = \frac{1}{n} \sum_{i=1}^n Y_i$$

$$\text{Std} = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (Y_i - \bar{Y})^2}$$

$$\text{Skewness} = \frac{\sum_{i=1}^N (Y_i - \bar{Y})^3 / N}{s^3}$$

$$\text{kurtosis} = \frac{\sum_{i=1}^N (Y_i - \bar{Y})^4 / N}{s^4}.$$

In the above equations, the \bar{Y} is the mean, s is the standard deviation, and N is the number of data points (NOTE: please see if all symbols are correctly defined here?). The equation below shows the first difference as implemented in *Matlab*^T. Taking the same equation twice will give an approximation of second derivative. The first order derivative gives non-zero values along the ramp while the second order derivative gives non-zero and a sign change at the onset and end of a ramp, and a much aggressive response at the spikes along with a sign change. Thus, it gives the features that can be easily differentiated. The output vector of the second difference is then used to calculate the statistical measures as given above.

$$\text{Diff} = [Y(2) - Y(1) \quad Y(3) - Y(2) \dots Y(N) - Y(N-1)].$$

In a second attempt, the Fourier Transform and Power Spectral Density of the vibration signal are calculated and the statistical values of the spectral components were used as a feature vector for the classification giving the concept of spectral mean, skewness, kurtosis, standard deviation and second order difference. Fourier transform is calculated using the Fast Fourier Transform (FFT) algorithm as in Equation (1), while for PSD the Fourier Transform of the autocorrelation function is required. For in depth details the interested readers are referred to [53].

$$y[k] = \sum_{n=0}^{N-1} x[n] e^{-j\frac{2\pi n k}{N}} \quad k = 0, 1, 2, 3, \dots, N-1, \quad (1)$$

where $y[k]$ is the Fourier transform and $x[n]$ is the signal under test.

3. Signal Analysis for Fault Detection

The vibration data collected from machine under test with the specifications given in Figure 2 were used for extracting different features as discussed in Section 2. The vibration signal was recorded at a sampling rate of 40,000 samples per second and was divided into chunks of one second data for calculating these features. Figures 3–6 show candidate feature vectors using the three different set of features. Time domain features are shown in Figure 3 and it can be seen that the average and standard deviation features show clear segregation between different types of faults; however, the skewness and kurtosis cannot be easily interpreted visually. Similar behaviour can be seen in the frequency domain representation of Average, Kurtosis, Skewness and Standard Deviation as shown in Figure 4. These features are obtained by first taking the Fourier Transform of the vibration signal, dividing it into chunks of one second data. The Average, Kurtosis, Skewness and Standard Deviation of the Fourier Transform of that one second data is taken. Similarly, the PSD of one second vibration data is calculated and its Average, Kurtosis, Skewness and Standard Deviation is shown in Figure 5. Figure 6 shows the Average values of time domain, Fourier domain and of PSD of the vibration signal; however, here these values are calculated after taking the second derivative of the signal. These features are much clearer as compared to the other features given in Figures 3, 4 and 5d; however, to properly

analyse these features and classify them into different types of faults, these features are forwarded to the SVM, KNN and KLDA. The results are discussed in the next Section.

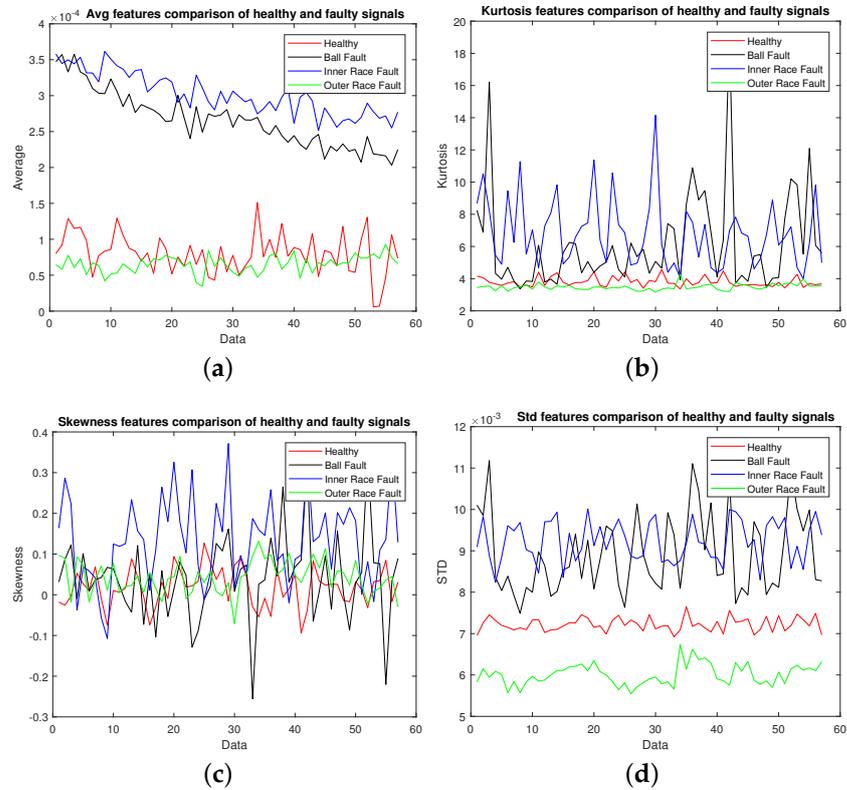


Figure 3. Statistical signal demonstration: (a) Average, (b) Kurtosis, (c) Skewness, (d) Standard Deviation.

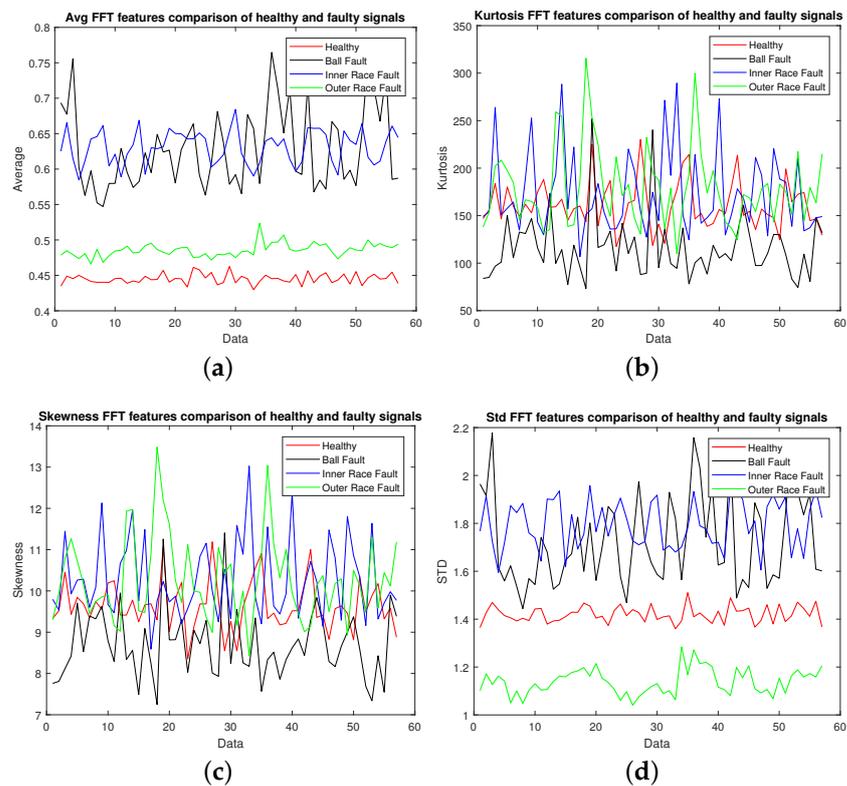


Figure 4. FFT signal demonstration: (a) Average, (b) Kurtosis, (c) Skewness, (d) Standard Deviation.

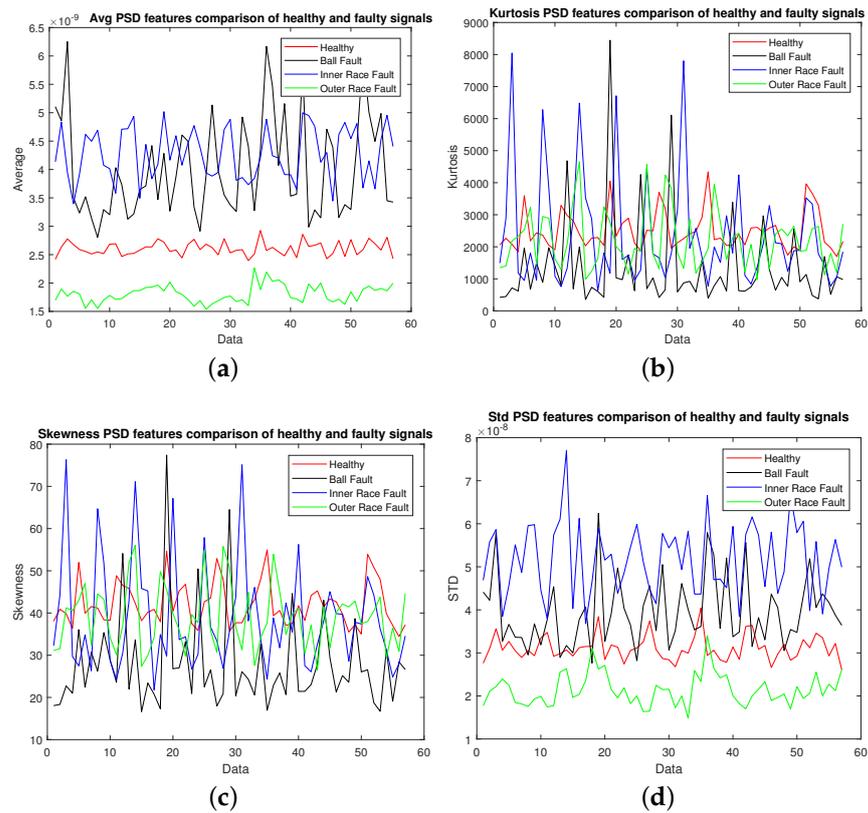


Figure 5. PSD signal demonstration: (a) Average, (b) Kurtosis, (c) Skewness, (d) Standard Deviation.

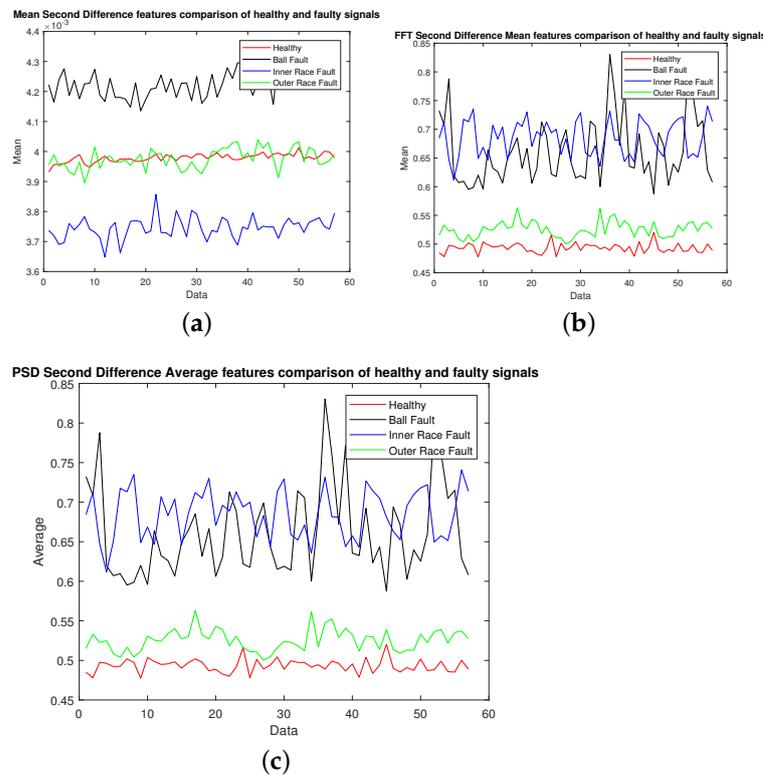


Figure 6. Analysis of mean values of second derivative: (a) Average (Time Domain), (b) Average FFT, (c) Average PSD.

4. Classification of Faults

Let $\mathbf{s} = [s_1, s_2, \dots, s_p]^\top \in \mathbb{R}^p$ is a p dimensional feature vector representing statistical and frequency domain features of the vibration signal. The training data matrix is given by $\mathbf{V} = \{\mathbf{s}_j\}_{j=1}^v \in \mathbb{R}^{p \times v}$, where v is the normalised feature vectors and c is the number of classes with the discrete class labels represented by $\mathbf{Y} = \{y_j\}_{j=1}^v$. The fault classification problem estimates the label y_t of a test feature vector $\mathbf{s}_t \in \mathbb{R}^p$ given the labelled training data \mathbf{V} . In order to analyse the effectiveness of our proposed features for fault classification, the feature vectors are used with classical supervised learning algorithms like Support Vector Machines (SVM), Nearest Neighbour (NN), and Kernel Linear Discriminant Analysis (KLDA).

KNN is a popular algorithm based on distance measure between two feature vectors \mathbf{s}_i and \mathbf{s}_j , using the Mahalanobis distance which is defined as [54]:

$$d = (\mathbf{s}_i - \mathbf{s}_j)^\top \mathbf{C}^{-1}(\mathbf{s}_i - \mathbf{s}_j), \quad (2)$$

where $\mathbf{C} \in \mathbb{R}^{p \times p}$ represents the covariance matrix obtained from training feature vectors. The \mathbf{C} can be computed as a diagonal matrix in case of small training sample sizes, with the feature variances as the diagonal elements.

Support Vector Machine was designed as supervised learning method for two class classification problems ($y_j \in \{1, -1\}$), but was subsequently developed for multiple class approach. In this research work, the vibration signal is required to be classified into four classes therefore, the binary SVM is extended to multi-class SVM via a one-versus-all strategy. Interested readers are referred to [55,56], for proper discussion on SVM. Here, the Lib-SVM [57] library is used to compute the parameters of the hyper-plane. It uses the following objective function with the training data provided, to optimise the hyper-plane using the training data:

$$\min_{\mathbf{w}, b, \xi} \left(\frac{1}{2} \mathbf{w}^\top \mathbf{w} + C \sum_j \xi_j \right) \quad (3)$$

$$s.t. \quad y_j(\mathbf{w}\mathbf{s}_j + b) \geq 1 - \xi_j, \xi_j \geq 0,$$

where C is the regularization constant, the hyperplane is represented by \mathbf{w} and b , while the non-separable cases are incorporated by ξ_j . For non-linear separation, the constraint $y_j(\mathbf{w}\phi(\mathbf{s}_j) + b) \geq 1 - \xi_j, \xi_j \geq 0$ can be introduced to perform the computation in an implicit higher dimensional space. The label y_t of a test feature vector \mathbf{s}_t is determined using the sign of $\frac{\mathbf{w}\mathbf{s}_t + b}{\|\mathbf{w}\|}$, once the parameters of the optimal hyper-plane are calculated.

KLDA uses supervised dimensionality reduction to represent data more efficiently, suppressing the less useful features for classification. It is applied to non-linearly separable classes that transform the p dimensional training feature vectors to $c - 1$ dimensional vectors which are then used for classification using machine learning techniques like SVM and so forth. KLDA uses Kernel matrix, computed using the function: $\mathbf{K}(i, j) = k(\mathbf{s}_i, \mathbf{s}_j)$. Given an input kernel \mathbf{K} , KLDA solves the following objective function [58]:

$$\alpha_{opt} = \arg \max \frac{\alpha^\top \mathbf{K} \mathbf{W} \mathbf{K} \alpha}{\alpha^\top \mathbf{K} \mathbf{K} \alpha}, \quad (4)$$

where $\alpha = [\alpha_1, \dots, \alpha_g]^\top$. $\mathbf{W} \in \mathbb{R}^{g \times g}$ is a block-diagonal matrix: $\mathbf{W} = \text{diag}\{\mathbf{W}_1, \mathbf{W}_2, \dots, \mathbf{W}_c\}$, where $\mathbf{W}_j \in \mathbb{R}^{m_j \times m_j}$ have every elements equal to $\frac{1}{m_j}$ (m_j represent the number of samples in class j). The largest eigenvectors of $(\mathbf{K} \mathbf{K} + \epsilon \mathbf{I})^{-1} (\mathbf{K} \mathbf{W} \mathbf{K}) \alpha = \lambda \alpha$ gives the optimal solution. The $(c - 1)$ dominant eigenvectors ($\mathbf{\Lambda} = [\alpha_1, \dots, \alpha_{c-1}] \in \mathbb{R}^{p \times (c-1)}$) are used to construct the transformation matrix and the training data matrix is projected on $\mathbf{\Lambda}$ to perform dimensionality reduction.

The Average, Kurtosis, Skewness and Standard Deviation vectors of each domain were concatenated before giving to SVM, KNN and KLDA for classification into bearings

with outer race fault, inner race fault, and ball fault. The results are shown in Table 1. The first column shows the feature vector along with its dimensions. In the first column Statistical_P, Fourier_P and PSD_P show our proposed feature vectors. The Statistical_P is a concatenation of Average, Kurtosis, Skewness and Standard Deviation of the vibration signal along with the same features calculated after taking the second difference of the vibration signals. Thus making a total of eight feature vectors of a size of two hundred and twenty eight each. The vectors of Fourier_P and PSD_P were arranged in a similar way, while the EMD, Fourier and PSD were calculated without taking the Average, Kurtosis, Skewness, and Standard Deviation and were used for comparison. The highest accuracy is given by the PSD_P for KLDA, which is 99.13%, followed by Statistical_P, which is 98.257% using KLDA. The Fourier is third in the row with 98.0% using KLDA followed by EMD with 97.01% using KNN. It is clear that our proposed feature vectors outperformed some of the most widely used feature vectors referenced in the literature. To further authenticate the performance of the proposed method, a reduction percentage is presented in Table 2. It can be observed that with the proposed method, more than 95% reduction level is achieved, which clearly decreases the computational complexity and also the classification time.

Table 1. Classification of faults.

Classifier	SVM	KNN	KLDA
Statistical_P	66.11%	93.53%	98.27%
Fourier_P	59.16%	95.52%	95.70%
PSD_P	62.25%	95.64%	99.13%
EMD	80.49%	93.01%	66.06%
Fourier	60.71%	94.13%	97.04%
PSD	65.65%	94.99%	85.47%

Table 2. Data dimensions and their reduction percentage.

Original Data Vector	Feature Vector	Final Feature Vector	Reduction Percentage
$1 \times 2,420,000$	Statistical (8×228)	(8×228)	99.924
	Fourier ($1 \times 40,000$)	(8×228)	95.44
	PSD ($1 \times 40,000$)	(8×228)	95.44
	EMD ($14 \times 160,000$)	($14 \times 160,000$)	7.438

Statistical Significance

Models are mostly evaluated by utilising the resampling methods such as k-fold cross-validation or hold-out from which the mean scores are determined and later compared. Though, the approaches are simple and valid, but could be misleading as it is quite hard to recognise that the difference between the scores is real or the results are statistically not stable. Therefore, statistical tests offer this confidence by quantifying the likelihood of the samples following the same distribution. The main motivation behind this statistical analysis is to acquire an absolute level of confidence in the implemented scheme. In this work, we utilise the analysis of variance (ANOVA) [59] statistical model to validate the significance of accuracy intervals—simply by analysing the selected samples for the differences among means.

For the proposed scenario, we are considering two different classifiers (KNN^q & $KLDA^r$)—selected on the basis of their improved classification results. In ANOVA, multiple tests are performed for the presumption of homogeneity of variance and normality. A Bartlett's test is performed to verify the homogeneity of variance, whereas the Shapiro–Wilk test is performed to check normality. The significance level α is selected to be 0.05, which

indicates if the test value found greater than α then data would be normally distributed and vice versa. The means for the given samples are represented by \bar{x}_1, \bar{x}_2 , which are calculated by performing the Monte-Carlo simulations to identify the lower and upper bounds after simulated 500 times. The null hypothesis H_0 claims that for the given means $\bar{x}_1 = \bar{x}_2$, whereas the alternative hypothesis H_a claims the rejection, $\bar{x}_1 \neq \bar{x}_2$. To test the null hypothesis, H_0 , p -value is calculated—conceded that the rejection fulfils the relation, $p \leq \sigma$, otherwise, the Bonferroni post-hoc test will be performed [60].

In this work, three different scenarios are considered, which are the calculation of statistical parameters on; (1) time-series data, (2) Fourier descriptors, and (3) PSD. In the first case (time-series data), by using the selected classifiers (KNN^q , & $KLDA^r$), the Shapiro–Wilk test generated the p -values, $p^q = 0.3203$, and $p^r = 0.9792$. Similarly, by applying Bartlett’s probability test, the associated Chi-squared probabilities are: $c^q = 2.277$, and $c^r = 0.420$. For the second case (Fourier descriptors), the Shapiro–Wilk test generated the p -values, $p^q = 0.9206$, and $p^r = 0.9724$. Similarly, by applying Bartlett’s probability test, the associated Chi-squared probabilities are $c^q = 0.1668$, and $c^r = 0.0698$. Finally, for the last case (PSD), the associated p -values are $p^q = 0.7386$, and $p^r = 0.1616$. By performing the Bartlett’s probability test, the calculated probabilities are $c^q = 0.606$, and $c^r = 3.645$.

It can be observed from the calculated p -values, considering all the cases and using the selected classifiers, are greater than σ . Thus, from the both probability test results (equality of variances, and normality), we fail to reject the null hypothesis H_0 , and therefore certain about the claim that the test data is normally distributed, and with the homogeneous variances. In Tables 3–5, we present a few statistical parameters to verify the authenticity of the proposed method based on the classification results. The ANOVA test parameters include degree of freedom (df), sum of squared deviation (SS), mean squared error (MSE), F-statistics and Prob-F value. Similarly, the confidence interval of all the selected classifiers and the selected cases are depicted in Figures 7–9. where the horizontal label 1 represents KNN classifiers, and 2 represents KLDA classifier.

The classification time of the proposed method is also provided in Figure 10, where it can be observed that the maximum time to classify any feature set is 12.62 s after achieving the reduction percentage of more than 95%. Whereas, Figure 11 depicts the classification time of original feature vectors in minutes. One can observe the time taken using the EMD features, which is more than 80 min, in comparison to PSD and Fourier, which are 11.34 min and 13.65 min, respectively. The classification time comparison is also provided in Figure 10.

Table 3. ANOVA test on statistical time-series data based on the selected classifiers.

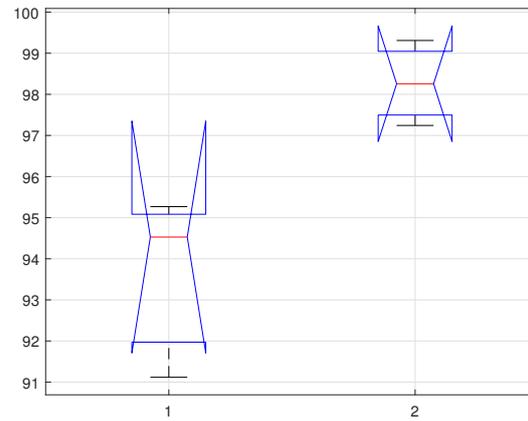
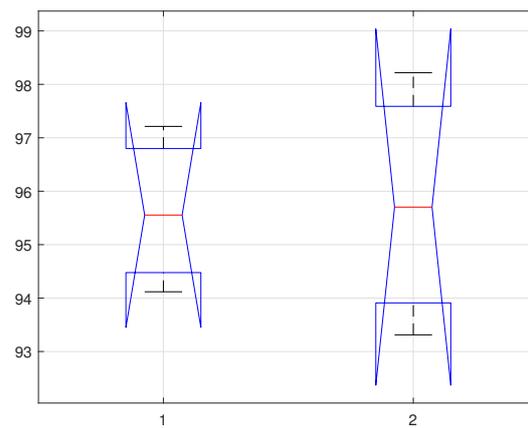
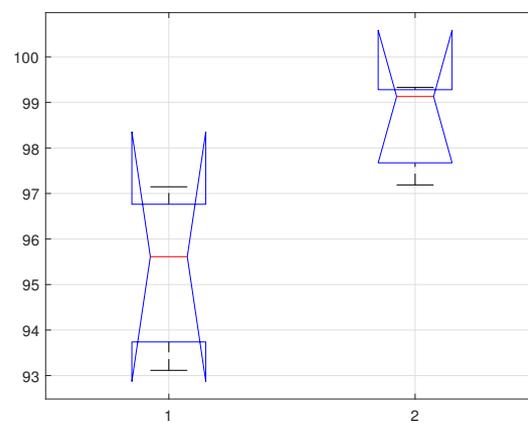
Variance Source	SS	df	MS	F	Prob > F
Columns	32.1785	1	32.1785		0.0304
Error	11.9377	4	2.9844	-	-
Total	44.1162	5			

Table 4. ANOVA test on statistical Fourier data based on the selected classifiers.

Variance Source	SS	df	MS	F	Prob > F
Columns	0.0201	1	0.02007	0.001	0.9483
Error	16.8298	4	4.20746	-	-
Total	16.8499	5			

Table 5. ANOVA test on statistical PSD data based on the selected classifiers.

Variance Source	SS	df	MS	F	Prob > F
Columns	15.9218	1	15.9218	5.75	0.0746
Error	11.0826	4	2.7706	-	-
Total	27.0044	5			

**Figure 7.** Confidence Interval of selected classifiers using statistical features of time-series data.**Figure 8.** Confidence Interval of selected classifiers using statistical features of fourier descriptors.**Figure 9.** Confidence Interval of selected classifiers using statistical features of PSD data.

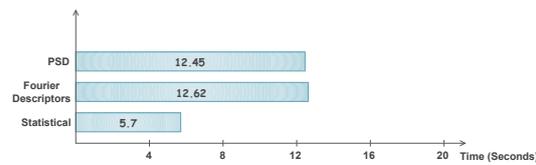


Figure 10. Classification time of three feature sets using proposed technique in seconds.

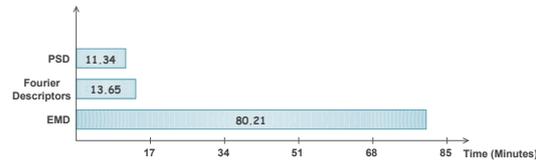


Figure 11. Classification time of three feature sets using proposed technique in minutes.

5. Conclusions

In this work, the vibration signals were analysed to detect and classify faults in rotating machinery. The signal was recorded and its statistical features, such as Average, Kurtosis, Skewness and RMS, were calculated in the time domain and the frequency domain. These features were also calculated by first finding the second derivative of the raw time domain signal. The features were then fed to different machine learning algorithms and were analysed for different patterns due to different faults and were used to train these machine learning models, resulting in successful detection and classification into ball, inner race and outer race faults. The Power Spectral Density features showed the best results for KLDA, followed by the statistical features using KLDA. This result was compared with that of the EMD, Fourier Transform and Power Spectral Density, in which the former one is time-frequency while the latter two are frequency domain representation. It is also important to note that the sizes of our proposed features are much less than those of the EMD, Fourier and Power Spectral Density, showing the computational efficiency of our proposed techniques. The proposed technique can be extended to time-frequency analyses like Short Term Fourier Transform and Wavelet Transform and so forth; also other bearing faults can be added, such as cage fault, which is not addressed here.

Author Contributions: Conceptualization, Visualization, Methodology, and validation, M.A. and T.A.; Original article Writing, validation, Supervision, and Software, M.A.K. and M.M.I.C.; Funding acquisition, Project administration, review and editing, and Validation, M.I. and C.-H.H. All authors have read and agreed to the published version of the manuscript.

Funding: No funding was received for this work.

Institutional Review Board Statement: Not Applicable.

Informed Consent Statement: Not Applicable.

Data Availability Statement: Not Applicable.

Conflicts of Interest: The authors declare no conflict of interests.

References

1. Moliner-Heredia, R.; Bruscas-Bellido, G.M.; Abellán-Nebot, J.V.; Peñarrocha-Alós, I. A Sequential Inspection Procedure for Fault Detection in Multistage Manufacturing Processes. *Sensors* **2021**, *21*, 7524. [[CrossRef](#)] [[PubMed](#)]
2. Lu, L.; Wang, W. Fault Diagnosis of Permanent Magnet DC Motors Based on Multi-Segment Feature Extraction. *Sensors* **2021**, *21*, 7505. [[CrossRef](#)] [[PubMed](#)]
3. Lin, S.L. Intelligent Fault Diagnosis and Forecast of Time-Varying Bearing Based on Deep Learning VMD-DenseNet. *Sensors* **2021**, *21*, 7467. [[CrossRef](#)] [[PubMed](#)]
4. He, J.; Wu, P.; Tong, Y.; Zhang, X.; Lei, M.; Gao, J. Bearing Fault Diagnosis via Improved One-Dimensional Multi-Scale Dilated CNN. *Sensors* **2021**, *21*, 7319. [[CrossRef](#)] [[PubMed](#)]
5. Kafeel, A.; Aziz, S.; Awais, M.; Khan, M.A.; Afaq, K.; Idris, S.A.; Alshazly, H.; Mostafa, S.M. An Expert System for Rotating Machine Fault Detection Using Vibration Signal Analysis. *Sensors* **2021**, *21*, 7587. [[CrossRef](#)] [[PubMed](#)]
6. Chacon, J.L.F.; Kappatos, V.; Balachandran, W.; Gan, T.H. A novel approach for incipient defect detection in rolling bearings using acoustic emission technique. *Appl. Acoust.* **2015**, *89*, 88–100. [[CrossRef](#)]

7. Lei, X.; Sandborn, P.A. PHM-based wind turbine maintenance optimization using real options. *Int. J. Progn. Health Manag.* **2016**, *7*, 1–8.
8. Xie, Y.; Zhang, T. Fault Diagnosis for Rotating Machinery Based on Convolutional Neural Network and Empirical Mode Decomposition. *Shock Vib.* **2017**, *2017*, 12. [[CrossRef](#)]
9. Li, C.J.; Ma, J. Wavelet decomposition of vibrations for detection of bearing-localized defects. *NDT E Int.* **1997**, *30*, 143–149.
10. Rai, A.; Upadhyay, S. A review on signal processing techniques utilized in the fault diagnosis of rolling element bearings. *Tribol. Int.* **2016**, *96*, 289–306. [[CrossRef](#)]
11. Mba, D.; Rao, R.B.K.N. Development of Acoustic Emission Technology for Condition Monitoring and Diagnosis of Rotating Machines: Bearings, Pumps, Gearboxes, Engines, and Rotating Structures. *Shock Vib. Dig.* **2006**, *38*, 3–16. [[CrossRef](#)]
12. Guo, P.; Infield, D.G.; Yang, X. Wind Turbine Generator Condition-Monitoring Using Temperature Trend Analysis. *IEEE Trans. Sustain. Energy* **2012**, *3*, 124–133. [[CrossRef](#)]
13. Kumar, M.; Mukherjee, P.S.; Misra, N.M. dvan cement and current status of wear debris analysis for machine condition monitoring: A review. *Ind. Lubr. Tribol.* **2012**, *65*, 3–11. [[CrossRef](#)]
14. Gomaa, F.R.; Khader, K.M.; Eissa, M.A. Fault Diagnosis of Rotating Machinery based on vibration analysis. *Int. J. Adv. Eng. Glob. Technol.* **2016**, *4*, 1571–1586.
15. Scheer, C.; Reimche, W.; Wilhelm Bach, F. Early Fault Detection at Gear Units by Acoustic Emission and Wavelet Analysis. *J. Acoustic Emission* **2007**, *25*, 331–340.
16. Gu, D.; Kim, J.; An, Y.; Choi, B. Detection of faults in gearboxes using acoustic emission signal. *J. Mech. Sci. Technol.* **2011**, *25*, 1279–1286. [[CrossRef](#)]
17. Othman, M.S.; Nuawi, M.Z.; Mohamed, R. Experimental comparison of vibration and acoustic emission signal analysis using kurtosis-based methods for induction motor bearing condition monitoring. *Prz. Elektrotechniczny* **2016**, *92*, 208–212. [[CrossRef](#)]
18. Chen, J.; Li, Z.; Pan, J.; Chen, G.; Zi, Y.; Yuan, J.; Chen, B.; He, Z. Wavelet transform based on inner product in fault diagnosis of rotating machinery: A review. *Mech. Syst. Signal Process.* **2016**, *70–71*, 1–35. [[CrossRef](#)]
19. Caesarendra, W.; Tjahjowidodo, T.; Kosasih, B.; Tieu, A.K. Integrated Condition Monitoring and Prognosis Method for Incipient Defect Detection and Remaining Life Prediction of Low Speed Slew Bearings. *Machines* **2017**, *5*, 11. [[CrossRef](#)]
20. Elforjani, M.; Mba, D. Detecting natural crack initiation and growth in slow speed shafts with the Acoustic Emission technology. *Eng. Fail. Anal.* **2009**, *16*, 2121–2129. [[CrossRef](#)]
21. Al-Dossary, S.; Hamzah, R.R.; Mba, D. Observations of changes in acoustic emission waveform for varying seeded defect sizes in a rolling element bearing. *Appl. Acoust.* **2009**, *70*, 58–81. [[CrossRef](#)]
22. Al-Ghamd, A.M.; Mba, D. A comparative experimental study on the use of acoustic emission and vibration analysis for bearing defect identification and estimation of defect size. *Mech. Syst. Signal Process.* **2006**, *20*, 1537–1571. [[CrossRef](#)]
23. Antoni, J. Fast computation of the kurtogram for the detection of transient faults. *Mech. Syst. Signal Process.* **2007**, *21*, 108–124. [[CrossRef](#)]
24. Wang, Y.; Xiang, J.; Markert, R.; Liang, M. Spectral kurtosis for fault detection, diagnosis and prognostics of rotating machines: A review with applications. *Mech. Syst. Signal Process.* **2016**, *66–67*, 679–698. [[CrossRef](#)]
25. Tagawa, T.; Tadokoro, Y.; Yairi, T. Structured Denoising Autoencoder for Fault Detection and Analysis. In Proceedings of the Sixth Asian Conference on Machine Learning, Nha Trang City, Vietnam, 26–28 November 2015; Volume 39, pp. 96–111.
26. Sakurada, M.; Yairi, T. Anomaly Detection Using Autoencoders with Nonlinear Dimensionality Reduction. In Proceedings of the MLSDA 2014 2nd Workshop on Machine Learning for Sensory Data Analysis, MLSDA'14, Gold Coast, QLD, Australia, 2 December 2014; pp. 4:4–4:11.
27. Verma, N.K.; Gupta, V.; Sharma, M.; Sevakula, R.K. Intelligent condition based monitoring of rotating machines using sparse auto-encoders. In Proceedings of the 2013 IEEE Conference on Prognostics and Health Management (PHM), Gaithersburg, MD, USA, 24–27 June 2013; pp. 1–7.
28. Yan, B.; Weidong, Q. Aero-engine sensor fault diagnosis based on stacked denoising autoencoders. In Proceedings of the 35th Chinese Control Conference, (CCC'16), Chengdu, China, 27–29 July 2016; pp. 6542–6546.
29. Li, M.; Wei, Q.; Wang, H.; Zhang, X. Research on fault diagnosis of time-domain vibration signal based on convolutional neural networks. *Syst. Sci. Control. Eng.* **2019**, *7*, 73–81. [[CrossRef](#)]
30. Abdelkrim, C.; Meridjet, M.S.; Boutasseta, N.; Boulanouar, L. Detection and classification of bearing faults in industrial geared motors using temporal features and adaptive neuro-fuzzy inference system. *Heliyon* **2019**, *5*, e02046. [[CrossRef](#)]
31. González-Muñiz, A.; Díaz, I.; Cuadrado, A.A. DCNN for condition monitoring and fault detection in rotating machines and its contribution to the understanding of machine nature. *Heliyon* **2020**, *6*, e03395. [[CrossRef](#)] [[PubMed](#)]
32. Liu, X.; Shi, J.; Sha, X.; Zhang, N. A general framework for sampling and reconstruction in function spaces associated with fractional Fourier transform. *Signal Process.* **2015**, *107*, 319–326. [[CrossRef](#)]
33. Giv, H.H. Directional short-time Fourier transform. *J. Math. Anal. Appl.* **2013**, *399*, 100–107. [[CrossRef](#)]
34. Baccar, D.; Söffker, D. Wear detection by means of wavelet-based acoustic emission analysis. *Mech. Syst. Signal Process.* **2015**, *60–61*, 198–207. [[CrossRef](#)]
35. Jedliński, L.; 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. [[CrossRef](#)]
36. Han, L.; Hong, J.; Wang, D. Fault diagnosis of aeroengine bearings based on wavelet package analysis. *Tuijin Jishu/J. Propuls. Technol.* **2009**, *30*, 328–341.

37. Deriche, M. Bearing fault diagnosis using wavelet analysis. In Proceedings of the 2005 1st International Conference on Computers, Communications and Signal Processing with Special Track on Biomedical Engineering, Kuala Lumpur, Malaysia, 14–16 November 2005.
38. Pandya, D.; Upadhyay, S.; Harsha, S. Fault diagnosis of rolling element bearing with intrinsic mode function of acoustic emission data using APF-KNN. *Expert Syst. Appl.* **2013**, *40*, 4137–4145. [[CrossRef](#)]
39. Delgado-Arredondo, P.A.; Morinigo-Sotelo, D.; Alfredo, R. Methodology for fault detection in induction motors via sound and vibration signals. *Mech. Syst. Signal Process.* **2017**, *83*, 568–589. [[CrossRef](#)]
40. Khan, M.U.; Imtiaz, M.A.; Aziz, S.; Kareem, Z.; Waseem, A.; Akram, M.A. System Design for Early Fault Diagnosis of Machines using Vibration Features. In Proceedings of the 2019 International Conference on Power Generation Systems and Renewable Energy Technologies (PGSRET), Istanbul, Turkey, 26–27 August 2019; pp. 1–6.
41. Han, D.; Liang, K.; Shi, P. Intelligent fault diagnosis of rotating machinery based on deep learning with feature selection. *J. Low Freq. Noise, Vib. Act. Control.* **2019**, *39*, 1461348419849279. [[CrossRef](#)]
42. Ge, J.; Niu, T.; Xu, D.; Yin, G.; Wang, Y. A Rolling Bearing Fault Diagnosis Method Based on EEMD-WSST Signal Reconstruction and Multi-Scale Entropy. *Entropy* **2020**, *22*, 290. [[CrossRef](#)]
43. Delprete, C.; Brusa, E.; Rosso, C.; Bruzzone, F. Bearing Health Monitoring Based on the Orthogonal Empirical Mode Decomposition. *Shock Vib.* **2020**, *2020*, 8761278. [[CrossRef](#)]
44. Malla, C.; Rai, A.; Kaul, V.; Panigrahi, I. Rolling element bearing fault detection based on the complex Morlet wavelet transform and performance evaluation using artificial neural network and support vector machine. *Noise Vib. Worldw.* **2019**, *50*, 313–327. [[CrossRef](#)]
45. Paudyal, S. Classification of Rotating Machinery Fault Using Vibration Signal. Master's Thesis, Dehradun Institute of Technology, University of North Dakota, Grand Forks, ND, USA, 2014.
46. Zhang, J.; Zhang, D.; Yang, M.; Xu, X.; Liu, W.; Wen, C. Fault Diagnosis for Rotating Machinery with Scarce Labeled Samples: A Deep CNN Method Based on Knowledge-Transferring from Shallow Models. In Proceedings of the 2018 International Conference on Control, Automation and Information Sciences (ICCAIS), Hangzhou, China, 24 – 27 October 2018; pp. 482–487.
47. Tong, Z.; Li, W.; Jiang, F.; Zhu, Z.; Zhou, G. Bearing fault diagnosis based on spectrum image sparse representation of vibration signal. *Adv. Mech. Eng.* **2018**, *10*, 1687814018797788. [[CrossRef](#)]
48. Hamadache, M.; Lee, D.; Mucchi, E.; Dalpiaz, G. Vibration-Based Bearing Fault Detection and Diagnosis via Image Recognition Technique Under Constant and Variable Speed Conditions. *Appl. Sci.* **2018**, *8*, 1392. [[CrossRef](#)]
49. Youcef Khodja, A.; Guersi, N.; Saadi, M.N.; Boutasseta, N. Rolling element bearing fault diagnosis for rotating machinery using vibration spectrum imaging and convolutional neural networks. *Int. J. Adv. Manuf. Technol.* **2020**, *106*, 1737–1751. [[CrossRef](#)]
50. Wang, S.; Xiang, J.; Zhong, Y.; Tang, H. A data indicator-based deep belief networks to detect multiple faults in axial piston pumps. *Mech. Syst. Signal Process.* **2018**, *112*, 154–170. [[CrossRef](#)]
51. Bao, F.S.; Liu, X.; Zhang, C. PyEEG: An open source python module for EEG/MEG feature extraction. *Comput. Intell. Neurosci.* **2011**, *2011*, 406391. [[CrossRef](#)] [[PubMed](#)]
52. James, C.J.; Lowe, D. Extracting multisource brain activity from a single electromagnetic channel. *Artif. Intell. Med.* **2003**, *28*, 89–104. [[CrossRef](#)]
53. Oppenheim, A.V.; Schaffer, R.W. *Discrete-Time Signal Processing*, 3rd ed.; Prentice Hall Press: Upper Saddle River, NJ, USA, 2009.
54. Theodoridis, S.; Koutroumbas, K. *Pattern Recognition*; Academic: Boston, MA, USA, 2010.
55. Cristianini, N.; Shawe-Taylor, J. *An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods*; Cambridge University: Cambridge, UK, 2000.
56. Scholkopf, B.; Smola, A.J. *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*; MIT: Cambridge, MA, USA, 2001.
57. Chang, C.; Lin, J. LIBSVM: A Library for Support Vector Machines. *ACM Trans. Intell. Syst. Tech.* **2011**, *2*, 1–27. [[CrossRef](#)]
58. Baudat, G.; Anouar, F. Generalized Discriminant Analysis Using a Kernel Approach. *Neural Comput.* **2000**, *12*, 2385–2404. [[CrossRef](#)] [[PubMed](#)]
59. Akram, T.; Laurent, B.; Naqvi, S.R.; Alex, M.M.; Muhammad, N. A deep heterogeneous feature fusion approach for automatic land-use classification. *Inf. Sci.* **2018**, *467*, 199–218. [[CrossRef](#)]
60. Aziz, S.; Khan, M.U.; Alhaisoni, M.; Akram, T.; Altaf, M. Phonocardiogram Signal Processing for Automatic Diagnosis of Congenital Heart Disorders through Fusion of Temporal and Cepstral Features. *Sensors* **2020**, *20*, 3790. [[CrossRef](#)]