Enhancing the Classification Performance of Machine Learning Techniques by Using Hjorth's and Other Statistical Parameters for Precise Tracking of Naturally Evolving Faults in Ball Bearings

Sameera Mufazzal and S. M. Muzakkir
Department of Mechanical Engineering, Jamia Millia Islamia, New Delhi-25, India.

Sidra Khanam
Department of Mechanical Engineering, ZHCET, Aligarh Muslim University, Aligarh, India.
E-mail:sidra.khanam10@gmail.com

(Received 15 November 2021; accepted 28 March 2022)

The research on identification of artificially induced faults in bearing is available in abundance in the past literature, however, the diagnosis becomes more challenging when the fault evolves naturally inside the bearing, and especially when its stages need to be precisely tracked. The conventional statistical features used commonly in the past literature do not uniquely characterize the fault status, and yield satisfactory results only in limited cases, like those for artificial faults. In this work, a new combination of fault descriptors, including three Hjorth's parameters, three statistical features and an entropy measure, is proposed and its effectiveness has been analyzed on classification performances of k-Nearest Neighbor (k-NN) and Support Vector Machine (SVM). Two datasets comprising the signals of run-to-failure tests were taken from Intelligent Maintenance Systems (IMS) and the Paderborn university repository. The data were categorized into large number of classes to closely indicate the actual fault type and size. Compared to conventional statistical features, the new combination was able to enhance the classification accuracy of k-NN and SVM, respectively, from 91.3% to 99.9%, and from 94.8% to 99.7% in the case of IMS dataset, and from 94.1% to 98.5%, and from 94.7% to 98.4% in the case of Paderborn dataset. In addition to accuracy, the performance metrics, including precision, recall, and F1-score were also improved using the proposed combinatory features.

1. INTRODUCTION

Rolling element bearings are used in a majority of the rotary machines in industries, which develop faults due to fatigue under running condition. The failure in bearings can cause unexpected breakdowns or interruption in services, and hence condition monitoring of these bearings has become a crucial task. The application of Machine Learning (ML) techniques is increasingly becoming common for automated condition monitoring tasks in modern industries due to its easy implementation and relatively reasonable accuracy.[1] This involves a sequence of operations starting from data acquisition, to signal processing and subsequent fault classification. Some of the basic ML algorithms include Decision Tree (DT), Support Vector Machines (SVM), Naïve Bayes, Neural Network, Random Forest (RF), k-Nearest Neighbor (k-NN), Artificial Neural Networks (ANN), Deep Believe Network (DBN), Hidden Markov Model, Convolutional Neural Network (CNN). With extensive research in this field, a variety of other algorithms have been developed to date, whose applications in bearing fault diagnosis have been widely reported. By deploying an ML technique, it is not only possible to detect the fault in a component, but by an efficient classification scheme, the location as well as the extent of the defect can also be determined. The principle of supervised ML models lies in the training of an algorithm to learn different class labels according to a set of input features, and then assigning a label for given values of fault features of a new data, based on its prior learning. Thus, the feature vector strongly determines the accuracy and the effectiveness of the learned model. In most of the earlier and some of the recent ML models, the basic statistics of the signal in time or frequency or in both the domains were used for training and testing of supervised algorithms or otherwise for assisting informative mode selection in some of the un-supervised learning algorithms.[2–7]

Since the occurrence of faults in the bearing give rise to frequent impulses in the vibration signal, which can be estimated by the means of kurtosis (or Kurtogram), and hence, kurtosis had remained a useful tool for capturing the cyclo-stationarity in the signal for fault identification in much of the earlier literature.[8–10] However, as established later, the kurtosis is highly sensitive to outliers and can produce misleading results where the signal is strongly contaminated by non-Gaussian noise, or it cannot reveal fault impulses for signals containing highly repetitive impulses as in the case of well-advanced faults. As a consequence, newer health indicators of bearings were derived by various researchers to accomplish reliable feature extraction from the signal. Barszcz and Jabłoński[11] proposed Protrugram as the kurtosis of envelope spectrum of the demodulated signal. This method could better detect the fault related transients in the signals with low signal-to-noise ratio. Mo et al.[12] categorized all these indices into three, as a measure of: i) fault impulse magnitude, like kurtosis, L2/L1 norm,[13] L-kurtosis,[14] complexity, smoothness index,[15] Gini index,[16,17] ii) fault impulse periodicity or cyclo-stationarity, like correlated kurtosis,[16] harmonic-to-noise ratio,[12,18] and iii) both magnitude and cyclo-stationarity, like squared envelope Infogram,[19] harmonic spectral kurtosis.[20]

Another common method of determining the bearing health status is complexity analysis, which is useful especially for complex and non-linear bearing signals. Complexity measures the chaotic behavior of the signal, and provides information related to the intrinsic oscillatory components of the signal. Earlier methods of finding complexity of rotary machines, as reported in the literature, include correlation dimension,[21] and Lempel-Ziv complexity.[22] Recently, the entropy-based method and its extensions have been more commonly used for detecting the malfunctions in the rotary system. The entropy of different frequency bands of the vibration signal provides the knowledge of energy distribution, which is altered with the occurrence of fault. Yu et al.[23] derived entropy-based fault feature extraction from the different empirical modes for training of ANN. Lei et al.[24] constructed the feature vector for Adaptive Neuro-Fuzzy Inference System (ANFIS), using the statistics of time and frequency domain signals along with the energy entropies of empirical modes. Yan and Gao[25] investigated the usefulness of 'approximate entropy' in diagnosing faults in a rotary machine by quantifying the regularity in a time series. Zhao and Yang[26] used sample entropy with Ensemble Empirical Mode Decomposition (EEMD) to extract the fault related information from different scales of the vibration signal. Zheng et al.[27] used a modified fuzzy entropy function to take into account the ambiguous nature of the state boundaries of the different fault classes. Yan et al.[28] investigated the dynamic changes in rotary machines due to fault in rolling element bearings, using Permutation Entropy (PE). The PE is considered more efficient for describing non-linear dynamic alterations in the signal, induced by various tribo-mechanical events, like slip, friction, non-linear contact stiffness, and damping in the bearing. Wu et al.[29] created the feature vector using Multiscale PE (MPE) by coarse-graining method. Zhang et al.[30] combined PE with EEMD and optimized SVM for identification of fault in a motor bearing. Tian et al.[31] proposed a real-time fault monitoring framework by computing MPE for different mono-components, obtained by self-adaptive decomposition. Zhou et al.[32] used a more generalized form of entropy, called Rényi entropy to assist the classification process of the neural network.

More recently, the applications of Hjorth's statistical parameters for feature construction and fault diagnosis have been increasingly reported. The Hjorth's parameters are comprised of three metrics, namely, activity, mobility and complexity, which are calculated from variance and its subsequent derivatives.[33] Earlier, these parameters were more common in biomedical fields,[34] however, with recognized potential of fault characterization, the Hjorth's parameters have been successfully applied in some of the recent research works.[35–41]

Caesarendra and Tjahjowidodo[35] studied the changes in statistical features, including Hjorth's parameters, as the natural fault develops in slow speed slew bearings, and found that impulse factor, margin factor, approximate entropy and largest Lyapunov exponent possess the fault indicative potentials. Among Hjorth's three parameters, only activity was found sensitive to the occurrence of fault. Grover and Turk[36] performed fault diagnosis of ball bearings with artificial defects, by training some rule-based classifiers using Hjorth's three parameters in time domain. Grover and Turk[37] next improved the classification performance by using Hjorth's parameters and negative log likelihood for Gaussian mixture model. Mufazzal et al.[38] further improved the classification results for CWRB datasets by training the k-NN algorithm using features, like root mean square (RMS), kurtosis, spectral entropy, and Hjorth's mobility and complexity. A total of twelve classes were formed corresponding to different fault locations and severity, for which a maximum classification accuracy of 99.6% was attained. Liang and Lu[39] performed Variational Mode decomposition (VMD), and extracted the features including Rényi entropy, singular value and Hjorth's parameters, from the best mode. Recently, Cocconcelli[41] devised a new health indicator called 'Detectivity' by combining three Hjorth's parameters to enhance the fault detection potential of Hjorth's parameters. The derived parameter was interpreted as an electric combination of three components in series: with activity and complexity as two amplifiers and mobility as an attenuator. The detectivity was capable of detecting naturally occurring emerging faults in ball bearings under constant load and running conditions.

The literature survey revealed that there are plenty of research papers on intelligent classification bearing faults, however, the research on naturally occurring faults is very limited. Moreover, the reported research works either deal with methods which are complex in implementation or disclose ML algorithms which generates relatively less accurate results.[3,42–49] Even though some methods based on neural network[50–53] were found to yield comparative or even better results, however, these methods have been validated only against a fewer number of fault classes, i.e., three to four (compared to eight in the present work) in the first dataset while only up to seven number of classes (compared to nine in the present work) in second data set, which remain unrealistic to precisely track the growth of spalls with operation. To ensure economic replacement of damaged bearings, it is extremely essential that the bearings being replaced should have completed its expected lifetime, while on the same hand, it should not compromise safety. This requires precise monitoring of the bearings and progressive faults, so that it cannot exceed the acceptable limit. This objective can be met by an automated fault diagnosis scheme only if it is reliable and robust under different operating conditions. The present work aims at proposing a ML based fault diagnosis framework with higher classification accuracy by applying a combination of Hjorth's parameters with other simple statistics, for careful identification of naturally occurring faults and their stages. In order to examine the effectiveness of the proposed combinatory features, they were used to classify the faults into large categories, i.e., 8 and 9, using standard ML algorithms. Since the proposed feature vector involves simple statistics, which are not only computationally efficient but also omits human involvement, it is suitable for real time condition monitoring applications.

The remainder of this article has been structured into the following sections: Section 2 provides the background of ML techniques and explains a set of fault features. Section 3 describes the analyzed datasets and fault characteristics of different classes considered in this study. Section 4 presents the results and relevant discussions. Finally, Section 5 concludes the present work by highlighting its main findings, and associated limitations and the potential future scopes.

2. BACKGROUND

2.1. ML Classifiers

The purpose of a classifier is to predict the category of a new observation with the least prediction errors. In the present study, two different supervised ML techniques, namely k-NN and SVM, have been used for fault classification. A brief description of each is provided below.

2.1.1. k-Nearest Neighbors (k-NN)

k-NN is the simplest of the supervised ML technique, which follows nearest neighbor serial algorithm to iteratively group the samples into clusters, depending on their relative distance. It can be used both for regression and classification problems. It generates reasonably accurate results, and it is robust to noise. There are a variety of distance methods for quantifying the similarity between the neighboring samples, which includes, but not limited to, Manhattan, Euclidian, Minkowski, cosine distances etc. In an Euclidean based algorithm, which is also used in the present work, the distance (d_E) is calculated as the square root of the sum of the squares of the differences between the points (x_i) of the original dataset (X) and datapoints (y_i) of a new observation (Y), that is to be predicted,[34] i.e.:

$$d_E(x, y) = \sqrt{\sum_{i=1}^{N} |x_i - y_i|^2}. \tag{1}$$

The term 'k' in k-NN refers to the number of nearest neighboring points relative to the new data point, between which the distances are calculated for classification. After computation of the distance between individual samples, these are arranged according to their similarities and samples where high similarities are grouped to that class or cluster. In a modified form, called weighted k-NN, different neighbors of the new observation are assigned different weights, typically, monotonously decreasing with distance. A more comprehensive detail on weighted k-NN can also be found in the work of Xu.[54]

In the present work, 'weighted' kernel function was used for classification and the number of neighbors (k) was set between 3–10.

2.1.2. Support Vector Machine (SVM)

SVM is another supervised ML technique which defines a hyperplane to divide the two classes. An optimal hyper-plane is obtained by maximizing its distance with the nearest data of the two classes, i.e., for N training samples having feature vector ($x_i \in R$; $i = 1, 2, 3, \ldots, N$), and a corresponding output vector ($y_i \in \{+1, -1\}$; $i = 1, 2, 3, \ldots, N$), the best hyperplane is obtained by solving the decision function, expressed by Eq. (2), in which the average classification error of the complete training sample is minimized:

$$\text{Minimize:} \quad \mathbf{w}^T\mathbf{w} + C\sum_{i=1}^{N} \xi_i; \tag{2}$$

$$\text{Subjected to:} \quad y_i\left(\mathbf{w}^T K(x_i) + b\right) \geq 1 - \xi_i,\ \xi_i \geq 0,\ \forall i; \tag{3}$$

where $\mathbf{w}$ and b represent the coefficients (weight vector) and bias of the hyperplane, respectively. ξ is a nonnegative relaxation factor or slack variable which permits misclassification with some cost, and C is the penalty factor which penalizes misclassification (for hard margin type SVM, C is kept zero), $K(\cdot)$ is the kernel function which transforms the input data to a high dimensional feature space by performing non-linear mapping. The transformation converts the non-separable data in input space to linearly separable data in feature space.[55] The common kernel functions include linear, polynomial, radial basis function, Gaussian, sigmoid, Laplacian, Bessel, etc. The samples which satisfy Eq. (3) and possess minimum classification error are called the support vectors.

The multi-class SVM problems can be solved by two approaches: in the first approach, called 'single machine' approach, a single optimization problem is formed and solved, while in the second approach, called 'divide and conquer', the whole multiclass problem is solved by decomposing it into a set of binary subproblems, and solving each of them. This can be accomplished in two ways, namely 'One-Against-All (OAA)' or 'One-Against-One (OAO)'. In OAA classification strategy, one SVM is constructed per each class so as to distinguish the samples in a particular class from the samples in the remaining classes, thereby transforming an n-class SVM problem into an 'n' numbers of binary class problems. In the second strategy, viz. OAO, one SVM is constructed per each pair of classes, so as to distinguish the samples in a particular class from the samples in the paired class, thereby forming a total of '$n(n-1)/2$' SVM classifiers.[34] Since lesser number of classifiers is involved in OAA method, it is computationally more efficient. For this reason, the OAA method has been used in the present work. Other details on SVM can be found in the works of Mammone, et al.[55] and Zoppis, et al.[56]

2.2. Fault Indicators / Features

Since in supervised learning process, the features of the signal are used for training and testing of the algorithm, the classification performance is strongly determined by the characteristics of these features. It is thus important that the features used for classification should be highly descriptive of the classes of interest. For bearing fault diagnosis, the classification features should be indicative of the different types of faults, so that they can uniquely represent each fault. A majority of the past research attempts have relied on the basic statistical descriptors of time and frequency domain signals as bearing health indicators.[2–7,42,44] This typically includes Root Mean Square (RMS) value, mean value, peak value, Standard Deviation (SD) / variance, kurtosis, skewness, etc., and some relatively newer parameters including Crest Factor (CrF), Clearance Factor (ClF), Shape Factor (SF), and Impulse Factor (IF). These parameters

Table 1. Feature matrices used in past research.

References	Features
[2]	16 time domain features, 13 frequency domain features, while energy and amplitude energy and instantaneous permutation entropy of time-frequency domain sub-signals of VMD.
[3]	10 statistical parameters (RMS, kurtosis, skewness, peak to peak value, CrF, SF, IF, margin factor (MF), and two other logarithmic features), while RMS, mean and entropy measures of intrinsic modes obtained after EMD.
[4][42]	9 statistical features corresponding to each wavelet packet coefficient including absolute mean, variance, skewness, kurtosis, CrF, IF, SF, ClF, square root amplitude.
[5]	10 statistical features including RMS, mean value, variance, skewness, kurtosis, CrF, IF, SF, MF, median, range.
[7]	9 features corresponding to each time, frequency and frequency domain signal: kurtosis, skewness, CrF, ClF, SF, impulse indicator, variance, square of the averaged square roots of absolute amplitude, and mean of absolute amplitude.
[44]	16 time domain, and 13 frequency domain features of modes obtained after EMD.
[57]	Energy entropy of the first 12 intrinsic modes of EMD.
[58]	9 statistical features from time, frequency, and time-frequency domain signals obtained by dual tree wavelet packet transform: kurtosis, skewness, CrF, SF, IF, ClF, frequency center, energy and entropy of wavelet packets.

Table 2. Sets of fault features used for classification.

Feature set	Names (Number of features)
Set 1	RMS, mean value, peak value, SD, kurtosis, skewness, SF, IF, ClF, and CrF (10)
Set 2	Hjorth's activity, mobility, and complexity (3)
Set 3	RMS, mean value, SF, SE, activity, mobility, and complexity (7)

can yield satisfactory results only in a limited scenario, where the contribution of noise / uncertainty in the signal is least. Hence, these parameters were used extensively for classifying mostly artificial faults. However, when real faults are concerned, the signal behavior lacks certainty, especially from one case to other. In the recent past, researchers have used Hjorth's parameters: activity, mobility, and complexity, either directly or combined with other indicators, for diagnosing real faults, and highly reasonable results were achieved. However, these research works considered only a very few numbers of fault classes, three or four, and as a result it becomes impossible to gain precise knowledge of the fault, its stage and exact location, altogether. Table 1 lists the different feature matrices used in some of the past research works for automated fault diagnosis of ball bearings by applying supervised ML techniques.

The feature vector in the present work is formed by combining Hjorth's parameters with a few other statistical features for a confident fault diagnosis. The proposed feature set containing three Hjorth's parameters, along with RMS, mean, spectral entropy, and shape factor, was obtained after conducting several studies towards establishing exceptionally reasonable classification results with a large number of fault categories, closely resembling the actual fault stage. Further, to identify the effectiveness of the proposed combinatory features, two other feature sets, as detailed in Table 2, were also taken for comparative investigation of the ML classification.

Set 1 contains ten conventional statistical features, while Set 2 contains only three Hjorth's parameters, which are derived from variance of the signal and its subsequent derivatives. Set 3 is composed of seven features, combining Hjorth's parameters and a few conventional indicators. A brief description of the parameters used in the proposed feature combination is given below.

A. Time domain statistics:

RMS: It is the most common descriptor used in industries to identify the emergence of a fault in bearings. It is a measure of the energy content of the signal waveform and can indicate any changes in the vibration amplitude due to occurrence of defects.

Mean: It is also an indicator of change in vibration levels due to a fault in the system, and hence, mean can be an informative indicator of the onset of bearing faults. The mean can be both positive and negative.

Shape factor: It is expressed as the ratio of RMS value to the mean of the absolute values of the signal. It is relevant to the shape of the signal, without its strength, and hence it can measure the relative energy distribution of the signal.

Hjorth's parameters: These are also statistical parameters, which are derived from the variance of the signal and its subsequent derivatives, and were previously common in clinical research applications for evaluating EEG signals. However, due to its extraordinary potential of fault classification, its applications are now emerging in machinery fault diagnosis. The first of the three parameters is called $activity$, which is defined as the variance of a signal, i.e. $activity(x) = Var(x)$. Hence, it indicates the spread of the data points around the average value. For signals with near zero mean, the $activity$ gives the measure of the average power of the signal. The second parameter is called $mobility$, which is defined as the square root of the ratio of the $activity$ of the derivative of the signal to the $activity$ of the signal itself, i.e., $mobility(x) = \sqrt{\frac{Var(\dot{x})}{Var(x)}}$. The $mobility$ gives an estimate of the mean frequency of the signal. With the occurrence of the fault, the frequency components of the signal get affected, as sensed by the mobility parameter. The last parameter of Hjorth is called $complexity$, which is defined as the ratio of the $mobility$ of the derivative of the signal to the $mobility$ of the signal itself, i.e., $complexity(x) = \frac{mobility(\dot{x})}{mobility(x)} \Rightarrow \sqrt{\frac{Var(\ddot{x})\,Var(x)}{Var(\dot{x})^2}}$. It gives an estimate of the bandwidth of the signal and indicates the resemblance of the signal with a pure sinusoid. A more detailed description of the three Hjorth's parameters can be found in Hjorth,[33] and Cocconcelli, et al.[41]

B. Frequency domain statistics:

Spectral entropy (SE): Entropy or information (Shannon) entropy is a measure of uncertainty in an information. The spectral entropy is the Shannon entropy of a signal in frequency domain, which is used to indicate the spectral power distribution. Let the discrete time series of a signal be given by $x = [x_1, x_2, x_3, \ldots, x_n]$, and its Fourier transformation be given by $X = [X_1, X_2, X_3, \ldots, X_N]$, then the power is calculated as the square of frequency

Table 3. Details of fault classes and associated natural fault characteristics of bearing dataset I.

Class	Fault details			
	Bearing code		Fault type	Cause
C1	K001		No fault	–
C2	KA04	Outer race	Single point OR spall (2mm length)	Fatigue
C3	KA16	Outer race	Two point-spalls on OR with 2mm and 3mm lengths	Fatigue
C4	KA30	Outer race	Random indents on OR shorter than 1mm in length	Particle inclusions
C5	KB23	Mixed	Multiple points spalls: Two on IR with 1.5 mm and 3 mm lengths, and one on OR with 3 mm length	Fatigue
C6	KB24	Mixed	Multiple fault: Extended spall on IR with 9.4 mm length, and random indents on OR of length shorter than 1 mm	Fatigue and particle inclusions
C7	KI16	Inner race	Extended spall on IR with 6 mm length	Fatigue
C8	KI18	Inner race	Single point spall on IR with 2.5 mm length	Fatigue

Table 4. Different load conditions applied during accelerated life test of bearing (dataset I).

S. No.	Designation	Shaft rotational speed (rpm)	Driving torque (Nm)	Static radial load (N)
1	N09_M07_F10	900	0.1	1000
2	N15_M01_F10	1500	0.7	1000
3	N15_M07_F04	1500	0.1	400
4	N15_M07_F10	1500	0.1	1000

sequences as: $E_i = |X(f)|^2$. The information (uncertainty) can be obtained using probability distribution as: $p_i = \frac{E_i}{\sum_i E_i}$. Next, the spectral entropy is calculated as: $H = \sum_{m=1}^{N} p_i \log_2 \left(\frac{1}{p_i}\right)$.

3. TEST SETUPS, DATASETS AND FAULT FEATURES

3.1. Description of Test Rig and Datasets

3.1.1. Set I: Paderborn Datasets

The Paderborn bearing dataset, which was provided by the Chair of Design and Drive Technology, Paderborn University (Germany), contains the vibration and motor current signals of a total of 32 run-failure tests, conducted on FAG 6203 single row, deep groove ball bearings, out of which 6 experiments belonged to healthy test bearing, 12 with artificially induced defects, while 14 with natural defects generated on inner race (IR) and outer race (OR) during experiments.[59] In this work, only seven of the total datasets of run-failure test have been used, representing eight different fault categories, as provided in Table 3.

In naturally generated faults, only the outer and the inner races were found damaged at the end of the accelerated life tests. Fault class C1 represents the normal bearing, fault classes C2 to C4 represent bearings with OR defects of varying severity. C4 comprises of extremely small defects, of less than an mm which is in the form of plastic deformation, caused by particle inclusions, and is distributed throughout the OR. C2 contains single point spall while C3 contains two point-spalls. Fault classes C7 and C8 indicate IR faults, with C8 as small size, while C7 as extended type fault. C5 and C6 represent mixed type fault, which contain defects of varying sizes on IR and OR, where C6 is more severe than C5.

The Paderborn data contain vibration acceleration of the test bearing housing in two radial directions, X and Y, collected at 64 kHz under four different operational conditions as listed in Table 4.

From the total dataset of natural faults, the acceleration components of only eight bearings (one normal and other faulty), with the codes mentioned in Table 3, were used in the analysis. Additionally, nearly equal number of samples were extracted from each set for constructing eight classes, forming a total of 4000 samples for training and testing purposes.

Table 5. Details of fault classes and associated natural fault characteristics of bearing dataset II.

Class	Fault details			Dataset
	Faulty Bearing	Fault Location	Fault severity	No.
C1 or Normal	None	–	–	Set 1
C2	Bearing 3	Inner race	Stage 1	Set 1
C3	Bearing 3	Inner race	Stage 2	Set 1
C4	Bearing 4	Rolling element	Stage 1	Set 1
C5	Bearing 4	Rolling element	Stage 2	Set 1
C6	Bearing 1	Outer race	Stage 1	Set 2
C7	Bearing 1	Outer race	Stage 2	Set 2
C8	Bearing 3	Outer race	Stage 1	Set 3
C9	Bearing 3	Outer race	Stage 2	Set 3

3.1.2. Set II: IMS Datasets

In the previous dataset, only one bearing is tested at a time by conducting a run-failure test, thereby making the diagnosis procedure relatively easy. However, when there are multiple bearings in an assembly, and each has an equal susceptibility to the occurrence of these faults, the vibration characteristics due to faults at different bearings interfere and the procedure for identifying the fault, and the corresponding bearing, is further complicated. In order to study the effectiveness of the proposed feature combinations on a bearing assembly, the second dataset has been obtained from the repository of the Intelligent Maintenance Systems (IMS), which was provided by NASA.[60] These data correspond to run-failure test of Rexnord ZA-2115 double row ball bearings for the duration of 30–35 days. Four such bearings were installed on an assembly and three datasets have been prepared. Each file in the dataset contains signals of 1-sec duration (with 20480 data points), which was recorded at 20 kHz sampling frequency. At the end of the tests, spalls were observed on bearings at inner race (IR), outer race (OR) and rolling element (RE). Based on the location of these spalls, and their possible stages, nine classes were formed in the present work, as shown in Table 5, where stage 1 represents the incipient (or onset of the) spall while stage 2 indicates the extended fault.

By correctly classifying the features into relevant fault class, it will be possible to not only identify the type of fault but also to locate the exact bearing with that fault. Further, by knowing the stage of a particular fault, a right decision on the replacement of the damaged bearing can be taken at an appropriate time, thereby meeting both safety and economic requisites.

Table 6. Classification results of SVM and k-NN models for combined load conditions (dataset I).

Feature set	Quadratic SVM		Weighted k-NN ($k = 5$)	
	Classification accuracy (%)	Classification time (s)	Classification accuracy (%)	Classification time (s)
Set 1	94.7	18.036	94.1	1.1118
Set 2	93.2	205.48	94.0	0.8679
Set 3	98.4	9.8993	98.5	1.4270

3.2. Fault Classification Methodology

The signals were sorted as normal, with incipient faults, and with critical (extended) faults, according to their code for dataset I, and date of acquisition for dataset II. All the features, as listed in Section 2.2, were extracted to create the feature vector, and fed as input to both the classifiers, i.e., SVM and k-NN, in MATLAB 2019a environment with 10-fold classification scheme. The results were found best for Quadratic kernel function in case of SVM (with OAA classification procedure), while for weighted kernel functions in the case of k-NN (with 5 nearest neighbors and Euclidean distance method), in the maximum number of cases. The performance of the two classifiers with different feature vectors are presented in the next section.

4. EXPERIMENTAL VALIDATION

This section briefs the results of SVM and k-NN models in terms of classification accuracy and computational time, corresponding to the three sets of feature vectors, as discussed in Section 2.2. Further, to assess the robustness of the proposed scheme, other performance metrics, namely precision, recall and F1-score were also evaluated, and the class-wise performance of the algorithms examined.

4.1. Influence of Fault Features

4.1.1. Dataset I (Paderborn) Results

Table 6 depicts the results of the two classification models trained for dataset I, in which the samples were taken from all the four load conditions during fault feature extraction process.

The results clearly indicate that among the three sets, the proposed feature vector is capable of producing the highest classification accuracy, within a reasonable training time. The accuracy of the models using Hjorth's parameters is similar to that of trained using general statistics. But, when Hjorth's parameters are combined with RMS, mean value, shape factor and spectral entropy, the accuracies of the models are improved by over 4%. The confusion matrices for both the SVM and k-NN models, corresponding to the three feature vectors are shown in Fig. 1, where the green color boxes represent the correctly identified fault classes, while red and peach color boxes indicate incorrect classification.

The wrong classification highlighted in peach color represent that the classifier has predicted the correct location of the fault, but it could not differentiate the severity of the fault, like between classes C2, C3, and C4. Further, it is important to notice that in some cases, C4 has been classified as C1. This is reasonable according to the fact that the vibration characteristics of normal bearings are highly similar to those with incipient OR faults.[61] The bright red color boxes represent unacceptably wrong classification, where the classifier failed to correctly predict both the fault location and severity. There are some fault cases represented in a dull red color, in which the fault is either on IR or OR, but the classifier categorized them in a mixed type fault containing both IR and OR defects, and vice-versa.

Since the accuracy alone cannot justify the overall performance of the ML algorithm, it is important to assess other parameters also. For this, three other performance metrics, namely precision, recall and F1-score have been evaluated for examining the reliability of the trained algorithms. These metrics were calculated corresponding to different classes. The average and the minimum values (among all the classes) of precision, recall, and F1-score associated with Fig. 1 corresponding to the three feature sets are given in Table 7.

It can be observed from Table 7 that compared to other two sets, both the average and the minimum values of precision, recall and F1-score are the highest in the case of the proposed feature set (viz. Set 3). Since the minimum precision and recall of all the classes correspond to the performance of the algorithm in the worst case, their higher magnitudes in the case of feature Set 3 indicate reasonably good performance. The minimum precision and recall values are relatively less than the average values, but as these values correspond to fault classes C7 and C8, it is acceptable because both these classes represent fault on IR, however with a different severity. For other discrete fault classes, these metrics are very high. The results provided in Table 6 and Table 7 jointly imply a considerable enhancement in the performance of ML using feature Set 3 required for reliable fault diagnosis.

Even though the results provided in Table 6 were obtained when the classifier was trained irrespective of the loading conditions, it is known that the dynamic behavior of the rotary system, and hence the vibration characteristics of the bearing is strongly determined by multiple factors including the operating conditions. It is highly recommended that the training and testing should be performed according to the applied loading conditions, to achieve better accuracy. Table 8 depicts the results of the two classification models trained separately using the samples at different loading conditions.

In all the cases (except for SVM under load condition 4), the classification accuracy is highest for feature Set 3. The average classification accuracies of SVM and k-NN (as 99.375% and 99.225%, respectively, using third feature set) were found to be better than those obtained in combined load conditions (98.4% and 98.5%, respectively).

4.1.2. Dataset II (IMS) Results

Table 9 depicts the results of dataset II, for different classification models trained using samples with 2048 data points.

It is evident from the classification results given in Table 9 that the first feature vector containing the conventional statistical measures of the signal has produced the least accuracy of all the models, while the second feature set, which contains three Hjorth parameters, resulted in better accuracy. This provides a clue that the Hjorth parameters have the potential to truly manifest the unique fault characteristics. Further, if Set 3 is used for classification, the accuracy is highest as well as the computational time is least.

The results reveal that the accuracy is highest corresponding to Set 3 for both SVM and k-NN models. Moreover, k-NN is more accurate (over 0.2%) as well as nearly fifteen times more

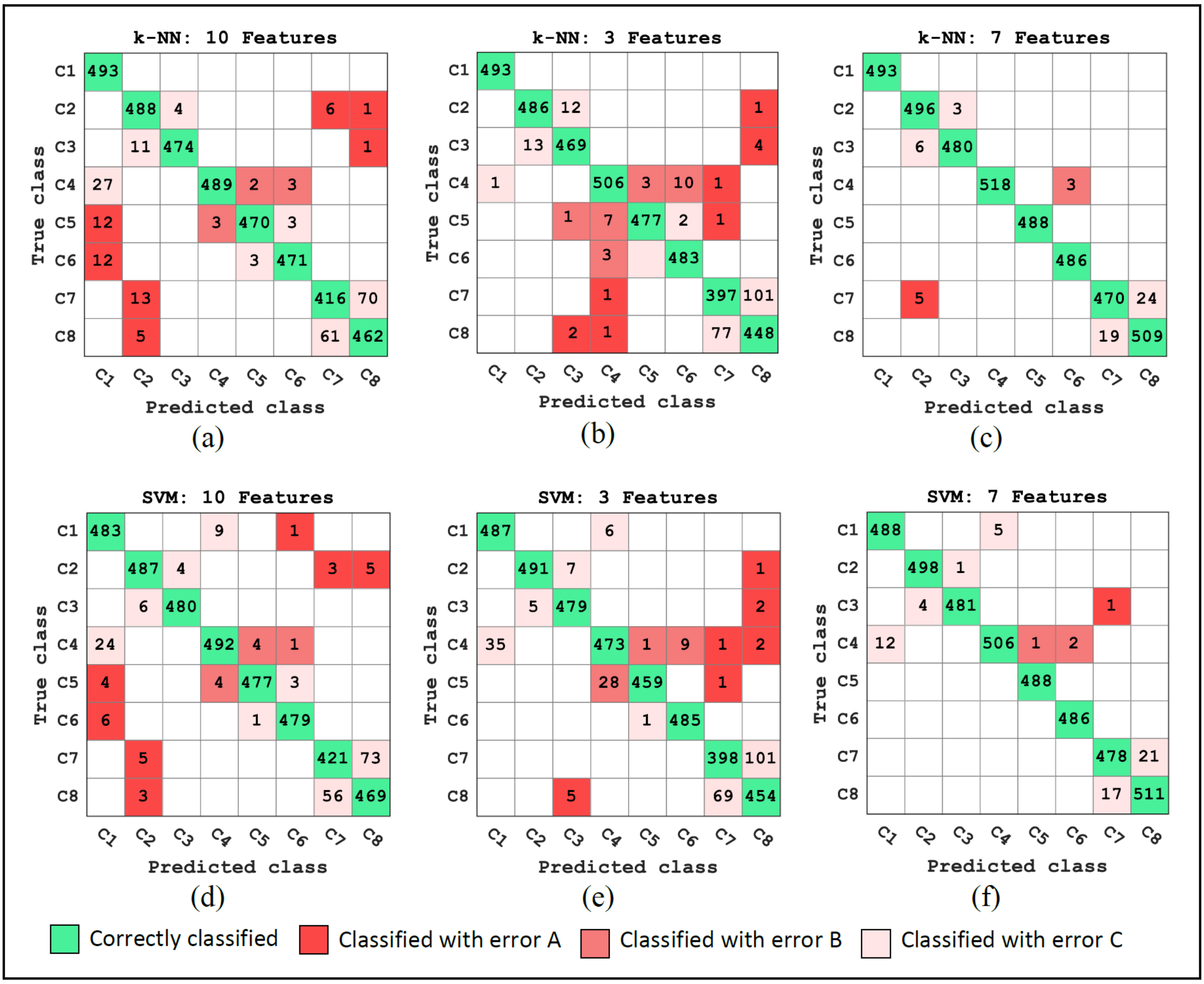

Figure 1. Confusion matrices of k-NN and SVM models for different feature sets (dataset I).

Table 7. Precision, recall and F1-score of Paderborn data set corresponding to different feature sets.

	Feature set	Average			Minimum		
		Precision	Recall	F1-score	Precision	Recall	F1-score
k-NN	1	0.942	0.942	0.942	0.765	0.765	0.765
	2	0.941	0.941	0.941	0.855	0.865	0.860
	3	0.985	0.985	0.985	0.995	0.995	0.995
SVM	1	0.948	0.948	0.948	0.877	0.844	0.860
	2	0.933	0.933	0.933	0.811	0.797	0.804
	3	0.984	0.984	0.984	0.961	0.958	0.964

efficient in computation than SVM. These two characteristics makes k-NN favorable for real time monitoring.

The confusion matrices for k-NN model, trained using 2048 sample size data with different feature vectors, are shown in Fig. 2, where green and red color boxes indicate correct and incorrect predictions, respectively.

The number of red color boxes, representing wrong classification and their frequencies are at a maximum in the case of feature Set 2, while in the case of feature Set 3, except for two of a total of 1800 samples, all fault classes have been correctly predicted. The incorrectly classified RE defect samples are due to vibration signals associated with RE defect are weak and get strongly modulated by the cage frequency. It is further attenuated due to mechanical compliance, while passing through the outer race to housing, thereby leading to inaccurate classification. This small-scale inaccuracy is reasonable due to simultaneous existence of faults in multiple bearings, which more or less, causes signal interference. The average and minimum values of precision, recall, and F1-score associated with Fig. 2 corresponding to the three feature sets are given in Table 10.

Compared to other two sets, the precision, recall and F1-scores are the highest in the case of the proposed feature set, implying very high potential of feature Set 3 for accurate fault classification. Moreover, the minimum values of these performance metrics are also considerably high, viz. very close to 1, for feature Set 3, whereas they are least, and nearly unacceptable for Set 1, followed by Set 2.

Figure 3 depicts the values of all the fourteen features studied in the present analysis, for every 200 samples of dataset II, corresponding to different fault types (C1 to C9).

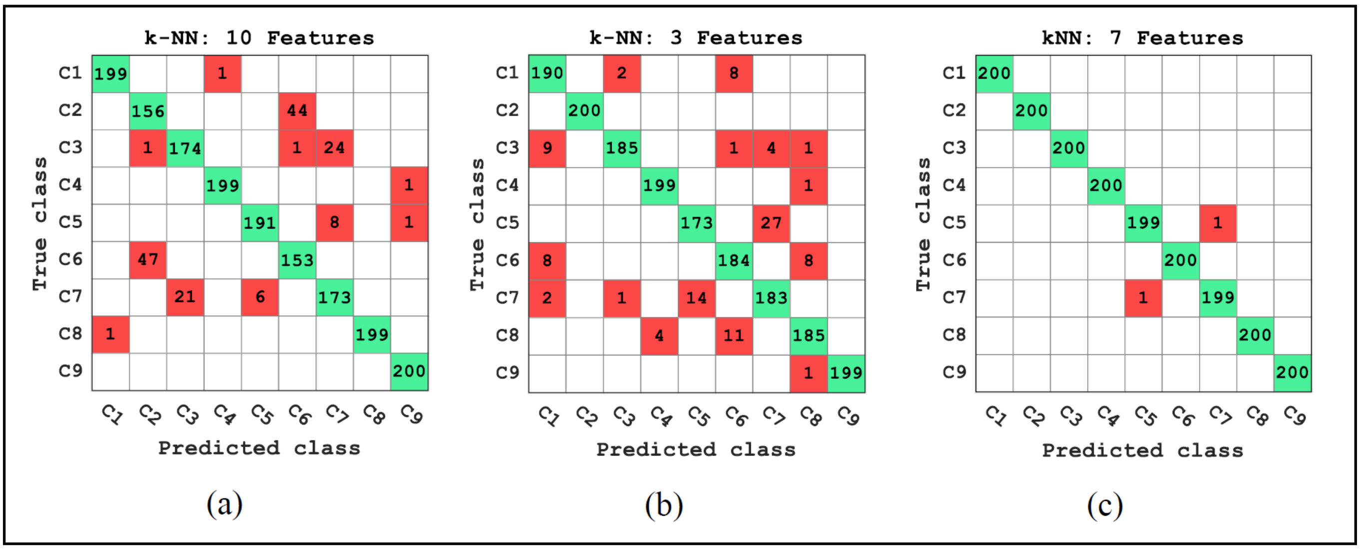

Figure 2. Confusion matrix of k-NN corresponding to different feature sets and samples with 2048 data points.

Table 8. Classification results for different sets of training features (bearing dataset I).

Load condition	Feature set	Quadratic SVM		Weighted k-NN ($k = 5$)	
		Classification accuracy (%)	Classification time (s)	Classification accuracy (%)	Classification time (s)
1	Set 1	95.7	6.1422	91.7	0.9566
	Set 2	96.5	11.789	95.5	0.9564
	Set 3	**99.0**	5.1619	**98.9**	1.0428
2	Set 1	99.2	4.9688	99.1	0.9376
	Set 2	98.2	4.958	97.7	0.8494
	Set 3	**99.8**	5.5814	**99.6**	0.9963
3	Set 1	98.9	6.3911	97.7	1.0364
	Set 2	92.1	45.353	91.3	0.8156
	Set 3	**99.0**	5.1613	**98.6**	0.8915
4	Set 1	**99.8**	5.6426	99.4	0.9559
	Set 2	98.3	5.0930	98.5	0.8581
	Set 3	99.7	5.3763	**99.8**	0.8952
Average		99.375		99.225	

Table 9. Classification results for different sets of training features (2048 sample size).

	Features	Accuracy (%)	Classification time (s)
SVM (Quadratic)	Set 1	94.8	7.8964
	Set 2	96.4	9.5465
	Set 3	99.7	7.6484
k-NN (Weighted)	Set 1	91.3	1.0213
	Set 2	94.3	0.92887
	Set 3	99.9	0.50885

The following observations, pertinent to IMS dataset, can be drawn from Fig. 3:

- RMS values can distinguish between different fault stages: normal state, impending fault and well-advanced faults, whereas the mean values can more clearly differentiate between the type (location) of the faults.
- RMS value, standard deviation, and activity follow the same distribution, hence according to "minimum correlation and maximum relevance", either of these three features will suffice for the classification for the present set of data.
- Standard deviation and Hjorth's activity follow the same distribution as these are correlated features. Hence, consideration of any of these two features will convey the same meaning.
- The distributions of RMS and peak values are also similar; however, the data are more dispersed for severe fault types (C3, C5 and C7), in the case of peak plot.
- Skewness shows a large variation in the data for all the fault classes, which makes it impractical for the bearing fault classification.
- Crest factor, clearance factor and impulse factor possess nearly the same distribution for different fault types, and hence these three features convey the same information. Besides this, these values are highly scattered, like skewness, which make them difficult to distinguish different fault types.
- Since activity refers to variance, which is square times the standard deviation, the values of activity (being fractional) are more concentrated than the standard deviation, and hence, activity (or variance) will thus be a more robust feature.
- Hjorth's complexity plot indicates that it is the highest for normal bearing, while it decreases with the occurrence of faults, and becomes less as the fault grows. On the other, Hjorth's mobility exhibits the converse behavior.
- The distribution of Hjorth's mobility values is highly discriminative for different faults, which is useful for accurate fault classification.
- The distribution of spectral entropy values is the same as that of shape factor, and hence, inclusion of spectral entropy will not significantly affect the results as far as shape factor is considered.

4.2. Comparison with Existing Literature

The results obtained using Set 3 feature vectors for training of k-NN learner were compared with the results of other methods, reported in the past literature, using the same Paderborn

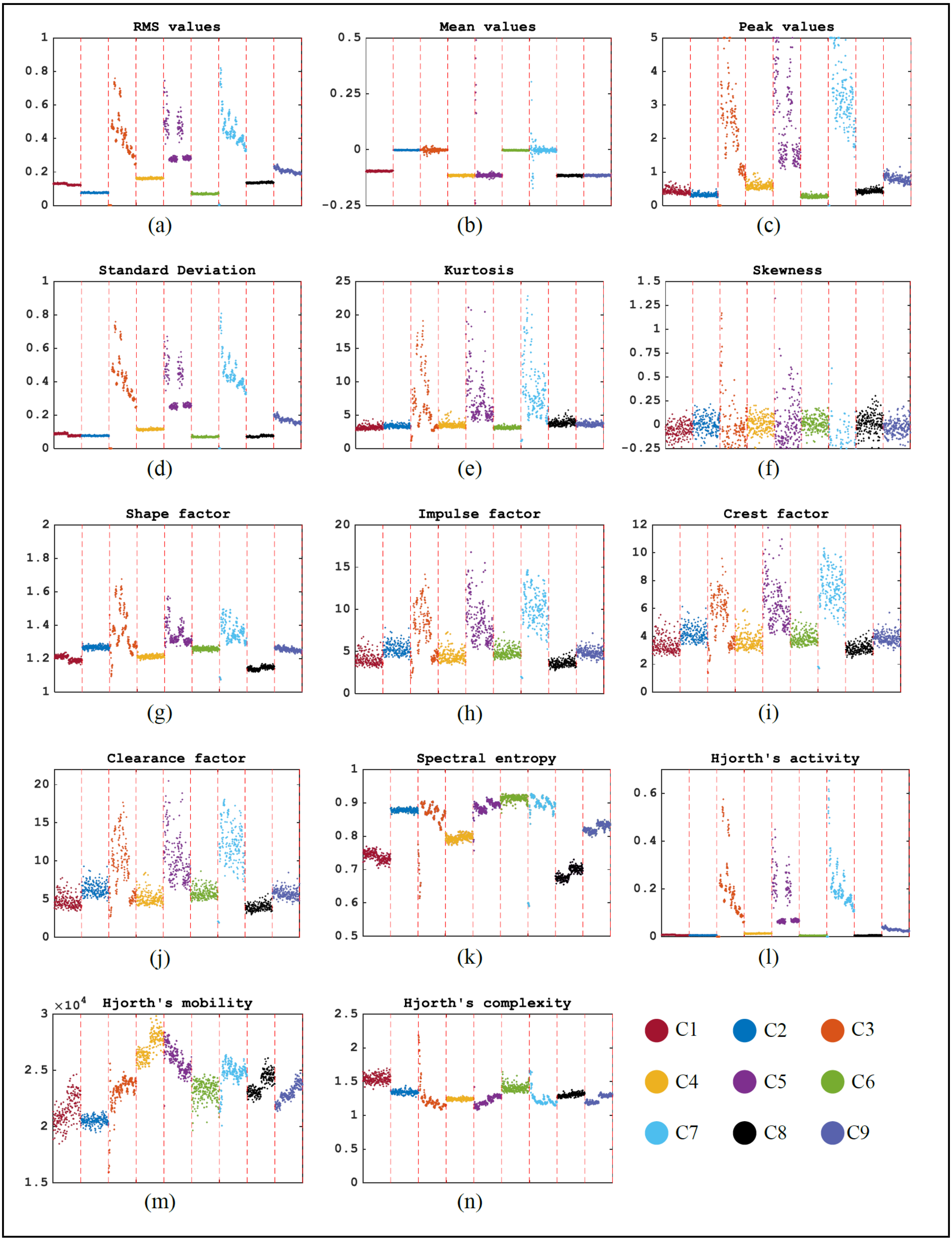

Figure 3. Scatter plots of different statistical features of all the nine fault classes (dataset II).

Table 10. Precision, recall and F1-score of IMS data set corresponding to different feature sets.

Feature set	Average			Minimum		
	Precision	Recall	F1-score	Precision	Recall	F1-score
1	0.914	0.913	0.913	0.765	0.780	0.772
2	0.944	0.943	0.944	0.855	0.865	0.860
3	899/900 $\leftrightarrow$ 1.0	899/900 $\leftrightarrow$ 1.0	899/900 $\leftrightarrow$ 1.0	0.995	0.995	0.995

Table 11. Comparison of the dataset I (Paderborn) results with other methods in the existing literature.

Reference	Classifier	No. of classes	Maximum Accuracy (%)
[48]	Modified CNN	Three	96.55%
[52]	Attention stream network	Three	99.10% (both artificial and natural faults)
[49]	CNN	Four	97.15% with Inception block
[53]	Supervised DNN	Four	100% considering temporal coherence
Present work	SVM and k-NN	Eight	98.5% and 98.4% for k-NN and SVM, respectively

and IMS datasets. The comparative results for the first dataset are provided in Table 11.

It is observed that the results of neural network based methods[52,53] are capable of producing accuracy similar to, or even higher than, that obtained in the present work, nevertheless it is also noteworthy that these approaches have only considered three to four fault classes, derived only on the basis of the fault location in the bearing. Such validation cannot meet the requirement of the economic diagnostic scheme, wherein the information of the fault severity is equally demanded. Besides this, the application of deep learning methods requires selection of appropriate structure and parameters of the network, for the algorithm to generate robust results.

Table 12 juxtaposes the results of IMS dataset obtained using feature Set 3 with the results of other methods. The previous works are found to deal with only a maximum of seven classes, corresponding to a normal bearing, and two states for incipient and extended faults, each on OR, IR and RE. These classes are, however, not sufficient to reveal the exact bearing containing that particular fault. Hence, in the present work, the classes have been formed both with respect to fault stage and bearing number, thereby totaling up to nine classes. This will assist the monitoring team to locate the actual bearing at fault, and perform the replacement of that bearing only, making the process safe and economical.

It is interesting to note that the present work, in which only seven parameters were used for training and testing purposes, is not only capable of classifying the maximum number of faults but also with exceptionally reasonable classification accuracy, reaching up to 99.9% for k-NN model. The method is also favorable due to its quick computational capability by virtue of a lesser number of faults characterizing features and relatively better computational efficiency compared to neural networks. Also, unlike neural networks, there is no need to search appropriate values of structural parameters, like number of layers, weights and bias. Further, it should be noted that for the neural networks, producing 100% accuracy,[50,51] the training and testing times are very high, which make them unamenable for real-time tasks.

5. CONCLUSIONS

The influence of three sets of features on the classification accuracy of naturally occurring faults in ball bearings have been investigated. The results revealed that compared to conventional statistical features, the addition of Hjorth parameters facilitates the true potential of differentiating the faults. If a feature vector is formed of Hjorth's parameters, combined with statistics like RMS value, mean value, shape factor, and the spectral entropy, the accuracy is significantly enhanced. Also, in addition to accuracy, the potential of the aforementioned combinatory features, as proposed in this work, for producing reliable diagnosis results was evaluated in terms of other ML performance metrics, including precision, recall, and F1-score, and their values were also found to be reasonably high.

It was revealed from the first part of the paper that both SVM and k-NN were able to produce highly accurate results, which were found to outperform the results reported in the literature. Although, the accuracy of SVM and k-NN are very close, the high computational efficiency of k-NN makes it more favorable for real-time health monitoring applications. There are a few cases in the literature where the accuracy of the neural network-based method was found to be better than the proposed work, however, these accuracies pertained to only a small number of fault classes, which are not sufficient to supply a reasonable level of information for economic maintenance and replacement operations. Further, it is observed that the implementation of the proposed framework involves simple extraction of features, and training of standard supervised ML classifiers, which eliminates the requirement of parameters, like those needed in complex deep learning methods. It is established from the results that the proposed combination of seven features can produce confident classification outputs, without demanding much effort on the ML techniques. Moreover, these classification results can be further improved by applying some competent signal processing tools, for noise removal and fault feature enhancement.

ACKNOWLEDGEMENT

The first author of this manuscript acknowledges the financial assistance of the Department of Science & Technology (DST)-India for supporting this research in terms of the Junior Research Fellowship (JRF)-INSPIRE.

REFERENCES

1 Arevalo, F., Diprasetya, M. R., and Schwung, A. A cloud-based architecture for condition monitoring based on machine learning, *Proc. of the IEEE 16th International Conference on Industrial Informatics, INDIN 2018*, 163–168, (2018). https://doi.org/10.1109/INDIN.2018.8471970

2 Yan, X. and Jia, M. A novel optimized SVM classification algorithm with multi-domain feature

Table 12. Comparison of the dataset II (IMS) results with other methods in the existing literature.

Reference	Classifier	No. of classes	Maximum Accuracy (%)
[42]	Fuzzy inference logic	Three	96.08%
[43]	Semi-supervised deep generative model	Three	92.01%
[44]	ANN	Three	97.5% with Multi-Layer Perceptron
[50]	1D CNN	Three	100% with 6458 sec and 40 sec of training and testing time, respectively
[51]	Convolutional Long-Short-Term-Memory Recurrent NN	Four	100% training and 97.13% testing (with 419 sec computational time)
[45]	CNN	Four	93.9%
[3]	Neural network	Seven	93%
[46]	Deep learning	Seven	99.66%
[47]	SVM	Seven	99.21%–100%
Present work	SVM and k-NN	Nine	99.9% and 99.7% for k-NN and SVM, respectively

and its application to fault diagnosis of rolling bearing, *Neurocomputing*, **313**, 47–64, (2018). https://doi.org/10.1016/j.neucom.2018.05.002

[3] Ali, J. B., Fnaiech, N., Saidi, L., Chebel-Morello, B., and Fnaiech, F. Application of empirical mode decomposition and artificial neural network for automatic bearing fault diagnosis based on vibration signals, *Applied Acoustics*, **89**, 16–27, (2015). https://doi.org/10.1016/j.apacoust.2014.08.016

[4] 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**, 1551–1564, (2013). https://doi.org/10.1016/j.measurement.2012.12.011

[5] Tahir, M. M., Khan, A. Q., Iqbal, N., Hussain, A., and Badshah, S. Enhancing fault classification accuracy of ball bearing using central tendency based time domain features, *IEEE Access*, **5**, 72–83, (2017). https://doi.org/10.1109/ACCESS.2016.2608505

[6] Liu, J., Xu, Z., Zhou, L., Nian, Y., and Shao, Y. A statistical feature investigation of the spalling propagation assessment for a ball bearing, *Mechanism and Machine Theory*, **131**, 336–350, (2019). https://doi.org/10.1016/j.mechmachtheory.2018.10.007

[7] Li, C., Sánchez, R. V., Zurita, G., Cerrada, M., and Cabrera D. Fault diagnosis for rotating machinery using vibration measurement deep statistical feature learning, *Sensors*, **16**, (2016). https://doi.org/10.3390/s16060895

[8] Antoni, J. and Randall, R. B. The spectral kurtosis: Application to the vibratory surveillance and diagnostics of rotating machines, *Mechanical Systems and Signal Processing*, **20**, 308–331, (2006). https://doi.org/10.1016/j.ymssp.2004.09.002

[9] Antoni, J. Fast computation of the kurtogram for the detection of transient faults, *Mechanical Systems and Signal Processing*, **21**, 108–124, (2007). https://doi.org/10.1016/j.ymssp.2005.12.002

[10] Wang, Y. and Liang, M. An adaptive SK technique and its application for fault detection of rolling element bearings, *Mechanical Systems and Signal Processing*, **25**, 1750–1764, (2011). https://doi.org/10.1016/j.ymssp.2010.12.008

[11] Barszcz, T. and Jabłoński, A. A novel method for the optimal band selection for vibration signal demodulation and comparison with the Kurtogram, *Mechanical Systems and Signal Processing*, **25**, 431–451, (2011). https://doi.org/10.1016/j.ymssp.2010.05.018

[12] Mo, Z., Wang, J., Zhang, H., and Miao, Q. Weighted cyclic harmonic-to-noise ratio for rolling element bearing fault diagnosis, *IEEE Transactions on Instrumentation and Measurement*, **69**, 432–442, (2020). https://doi.org/10.1109/TIM.2019.2903615

[13] Antoni, J. The spectral kurtosis: A useful tool for characterising non-stationary signals, *Mechanical Systems and Signal Processing*, **20**, 282–307, (2006). https://doi.org/10.1016/j.ymssp.2004.09.001

[14] Liu, S., Hou, S., He, K., and Yang, W. L-Kurtosis and its application for fault detection of rolling element bearings, *Measurement*, **116**, 523–532, (2018). https://doi.org/10.1016/j.measurement.2017.11.049

[15] Bozchalooi, I. S. and Liang, M. A smoothness index-guided approach to wavelet parameter selection in signal de-noising and fault detection, *Journal of Sound and Vibration*, **308**, 246–267, (2007). https://doi.org/10.1016/j.jsv.2007.07.038

[16] Rai, A. and Kim, J. M. A novel health indicator based on the Lyapunov exponent, a probabilistic self-organizing map, and the Gini-Simpson index for calculating the RUL of bearings, *Measurement*, **164**, 108002, (2020). https://doi.org/10.1016/j.measurement.2020.108002

[17] Miao, Y., Wang, J., Zhang, B., and Li, H. Practical framework of Gini index in the application of machinery fault feature extraction, *Mechanical Systems and Signal Processing*, **165**, 108333, (2022). https://doi.org/10.1016/j.ymssp.2021.108333

[18] Xu, X., Zhao, M., Lin, J., and Lei, Y. Envelope harmonic-to-noise ratio for periodic impulses detection and its application to bearing diagnosis, *Measurement*, **91**, 385–397, (2016). https://doi.org/10.1016/j.measurement.2016.05.073

[19] Antoni, J. The infogram: Entropic evidence of the signature of repetitive transients, *Mechanical Systems and Signal Processing*, **74**, 73–94, (2016). https://doi.org/10.1016/j.ymssp.2015.04.034

[20] Zhang, K., Chen, P., Yang, M., Song, L., and Yonggang, X. The Harmogram: A periodic impulses detection method and its application in bearing fault diagnosis, *Mechanical Systems and Signal Processing*, **165**, (2022). https://doi.org/10.1016/j.ymssp.2021.108374

[21] Logan, D. and Mathew, J. Using the correlation dimension for vibration fault diagnosis of rolling element bearings—I. Basic concepts, *Mechanical Systems and Signal Processing*, **10**, 241–250, (1996). https://doi.org/10.1006/mssp.1996.0018

[22] Hong, H. and Liang, M. Fault severity assessment for rolling element bearings using the Lempel-Ziv complexity and continuous wavelet transform, *Journal of Sound and Vibration*, **320**, 452–468, (2009). https://doi.org/10.1016/j.jsv.2008.07.011

[23] Yu, Y., YuDejie, and Junsheng, C. A roller bearing fault diagnosis method based on EMD energy entropy and ANN, *Journal of Sound and Vibration*, **294**, 269–277, (2006). https://doi.org/10.1016/j.jsv.2005.11.002

[24] Lei, Y., He, Z., and Zi, Y. A new approach to intelligent fault diagnosis of rotating machinery, *Expert Systems with Applications*, **35**, 1593–1600, (2008). https://doi.org/10.1016/j.eswa.2007.08.072

[25] Yan, R. and Gao, R. X. Approximate Entropy as a diagnostic tool for machine health monitoring, *Mechanical Systems and Signal Processing*, **21**, 824–839, (2007). https://doi.org/10.1016/j.ymssp.2006.02.009

[26] Zhao, Z. H. and Yang, S. P. Sample entropy-based roller bearing fault diagnosis method, *Journal of Vibration and Shock*, **31**, (2012).

[27] Zheng, J., Cheng, J., and Yang, Y. A rolling bearing fault diagnosis approach based on LCD and fuzzy entropy, *Mechanism and Machine Theory*, **70**, 441–453, (2013). https://doi.org/10.1016/j.mechmachtheory.2013.08.014

[28] Yan, R., Liu, Y., and Gao, R. X. Permutation entropy: A nonlinear statistical measure for status characterization of rotary machines, *Mechanical Systems and Signal Processing*, **29**, 474–484, (2012). https://doi.org/10.1016/j.ymssp.2011.11.022

[29] Wu, S. D., Wu, P. H., Wu, C. W., Ding, J. J., and Wang, C. C. Bearing fault diagnosis based on multiscale permutation entropy and support vector machine, *Entropy*, **14**, 1343–1356, (2012). https://doi.org/10.3390/e14081343

[30] Zhang, X., Liang, Y., Zhou, J., and Zang, Y. A novel bearing fault diagnosis model integrated permutation entropy, ensemble empirical mode decomposition and optimized SVM, *Measurement*, **69**, 164–179, (2015). https://doi.org/10.1016/j.measurement.2015.03.017

[31] Tian, Y., Wang, Z., and Lu, C. Self-adaptive bearing fault diagnosis based on permutation entropy and manifold-based dynamic time warping, *Mechanical Systems and Signal Processing*, **114**, 658–673, (2019). https://doi.org/10.1016/j.ymssp.2016.04.028

[32] Zhou, H., Zhong, F., Shi, T., Lai, W., Duan, J., and Zhang, Y. Class-information–incorporated kernel entropy component analysis with application to bearing fault diagnosis, *Journal of Vibration and Control*, **27** (5–6), 543–555, (2021). https://doi.org/10.1177/1077546320932030

[33] Hjorth, B. EEG analysis based on time domain properties, *Electroencephalography and Clinical Neurophysiology*, **29**, 306–310, (1970). https://doi.org/10.1016/0013-4694(70)90143-4

[34] Safi, M. S. and Safi, S. M. M. Early detection of Alzheimer's disease from EEG signals using Hjorth parameters, *Biomedical Signal Processing and Control*, **65**, 102338, (2021). https://doi.org/10.1016/j.bspc.2020.102338

[35] Caesarendra, W. and Tjahjowidodo, T. A review of feature extraction methods in vibration-based condition monitoring and its application for degradation trend estimation of low-speed slew bearing, *Machines*, **5**, (2017). https://doi.org/10.3390/machines5040021

[36] Grover, C. and Turk, N. Rolling element bearing fault diagnosis using empirical mode decomposition and Hjorth parameters, *Procedia Computer Science*, **167**, 1484–1494, (2020). https://doi.org/10.1016/j.procs.2020.03.359

[37] Grover, C. and Turk, N. Optimal statistical feature subset selection for bearing fault detection and severity estimation, *Shock and Vibration*, **2020**, (2020). https://doi.org/10.1155/2020/5742053

[38] Mufazzal, S., Muzakkir, S. M., and Khanam, S. Intelligent evaluation of ball bearing health degradation using wavelet packet transform, *Proc. of the 2nd International Congress on Advances in Mechanical and Systems Engineering*, Jalandhar, India, (2021).

[39] Liang, T. and Lu, H. A novel method based on multi-island genetic algorithm improved variational mode decomposition and multi-features for fault diagnosis of rolling bearing, *Entropy*, **22**, (2020). https://doi.org/10.3390/e22090995

[40] Roy, S. S., Dey, S., and Chatterjee, S. Autocorrelation Aided Random Forest Classifier-Based Bearing Fault Detection Framework, *IEEE Sensors Journal*, **20**, 10792–10800, (2020). https://doi.org/10.1109/JSEN.2020.2995109

[41] Cocconcelli, M., Strozzi, M., Cavalaglio Camargo Molano, J., and Rubini, R. Detectivity: A combination of Hjorth's parameters for condition monitoring of ball bearings, *Mechanical Systems and Signal Processing*, **164**, 108247, (2022). https://doi.org/10.1016/j.ymssp.2021.108247

[42] Berredjem, T. and Benidir, M. Bearing faults diagnosis using fuzzy expert system relying on an Improved Range Overlaps and Similarity method, *Expert Systems with Applications*, **108**, 134–142, (2018). https://doi.org/10.1016/j.eswa.2018.04.025

[43] Zhang, S., Ye, F., Wang, B., and Habetler, T. G. Semi-supervised bearing fault diagnosis and classification using variational autoencoder-based deep generative models, *IEEE Sensors Journal*, **21**, 6476–6486, (2021). https://doi.org/10.1109/JSEN.2020.3040696

[44] Shah, A. K., Yadav, A., and Malik, H. EMD and ANN based intelligent model for bearing fault diagnosis, *Journal of Intelligent & Fuzzy Systems*, **35**, 5391–5402, (2018). https://doi.org/10.3233/JIFS-169821

[45] Eren, L., Ince, T., and Kiranyaz, S. A generic intelligent bearing fault diagnosis system using compact adaptive 1D CNN classifier, *Journal of Signal Processing Systems*, **91**, 179–189, (2019). https://doi.org/10.1007/s11265-018-1378-3

[46] Zhang, Y., Li, X., Gao, L., and Li, P. A new subset based deep feature learning method for intelligent fault diagnosis of bearing, *Expert Systems with Applications*, **110**, 125–142, (2018). https://doi.org/10.1016/j.eswa.2018.05.032

[47] Ambika, P. S., Rajendrakumar, P. K., and Ramchand, R. Vibration signal based condition monitoring of mechanical equipment with scattering transform, *Journal of Mechanical Science and Technology*, **33**, 3095–3103, (2019). https://doi.org/10.1007/s12206-019-0604-7

[48] Liu, H., Yao, D., Yang, J., and Li, X. Lightweight convolutional neural network and its application in rolling bearing fault diagnosis under variable working conditions, *Sensors*, **19**, 1–20, (2019). https://doi.org/10.3390/s19224827

[49] Zhu, Z., Peng, G., Chen, Y., and Gao, H. A convolutional neural network based on a capsule network with strong generalization for bearing fault diagnosis, *Neurocomputing*, **323**, 62–75, (2019). https://doi.org/10.1016/j.neucom.2018.09.050

[50] Ozcan, I. H., Devecioglu, O. C., Ince, T., Eren, L., and Askar, M. Enhanced bearing fault detection using multichannel, multilevel 1D CNN classifier, *Electrical Engineering*, **104**, (2021). https://doi.org/10.1007/s00202-021-01309-2

[51] Khorram, A., Khalooei, M., and Rezghi, M. End-to-end CNN + LSTM deep learning approach for bearing fault diagnosis, *Applied Intelligence*, **51**, 736–751, (2021). https://doi.org/10.1007/s10489-020-01859-1

[52] Karnavas, Y. L., Plakias, S., and Chasiotis, I. D. Extracting spatially global and local attentive features for rolling bearing fault diagnosis in electrical machines using attention stream networks, *IET Electric Power Applications*, **15**, 903–915, (2021). https://doi.org/10.1049/elp2.12063

[53] Zhang, R., Peng, Z., Wu, L., Yao, B., and Guan, Y. Fault diagnosis from raw sensor data using deep neural networks considering temporal coherence, *Sensors*, **17**, 1–17, (2017). https://doi.org/10.3390/s17030549

[54] Xu, J. Multi-Label Weighted k-Nearest Neighbor Classifier with Adaptive Weight Estimation, in: Lu BL., Zhang L., Kwok J. (eds) *Neural Information Processing. ICONIP 2011. Lecture Notes in Computer Science*, **7063**, Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-24958-7_10

[55] Mammone, A., Turchi, M., and Cristianini, N. Support vector machines, *WIREs Computational Statistics*, **1**, 283–289, (2009). https://doi.org/10.1002/wics.49

[56] Zoppis, I., Mauri, G., and Dondi, R. Kernel Methods: Support Vector Machines, in *Encyclopedia of Bioinformatics and Computational Biology, Volume 1: Methods*, Elsevier, 503–510, (2019).

[57] Ali, J. B., Saidi, L., Mouelhi, A., Chebel-Morello, B., and Fnaiech, F. Linear feature selection and classification using PNN and SFAM neural networks for a nearly online diagnosis of bearing naturally progressing degradations, *Elsevier Engineering Applications of Artificial Intelligence*, **42**, 67–81, (2015). https://doi.org/10.1016/j.engappai.2015.03.013

[58] Shao, H., Jiang, H., Wang, F., and Wang, Y. Rolling bearing fault diagnosis using adaptive deep belief network with dual-tree complex wavelet packet, *ISA Transactions*, **69**, 187–201, (2017). https://doi.org/10.1016/j.isatra.2017.03.017

[59] KAt-DataCenter, Chair of Design and Drive Technology, Paderborn University, https://mb.uni-paderborn.de/kat/forschung/datacenter/bearing-datacenter/ (Accessed August 15, 2021).

[60] Lee, J., Qiu, H., Yu, G., and Lin, J. Bearing Data Set, Rexnord Tech. Serv. IMS, Univ. Cincinnati. "Bearing Data Set", NASA Ames Progn. Data, NASA Ames Res. Center, Moffett Field, CA Repos. (2007). https://ti.arc.nasa.gov/tech/dash/groups/pcoe/prognostic-data-repository/ (accessed August 19, 2021).

[61] Mufazzal, S., Muzakkir, S. M., and Khanam, S. Theoretical and experimental analyses of vibration impulses and their influence on accurate diagnosis of ball bearing with localized outer race defect, *Journal of Sound and Vibration*, **513**, 116407, (2021). https://doi.org/10.1016/j.jsv.2021.116407