



Article citation info:

Basaran M, Fidan M. Induction motor fault classification via entropy and column correlation features of 2D represented vibration data. *Eksploracja i Niezawodność – Maintenance and Reliability* 2021; 23 (1): 132–142, <http://dx.doi.org/10.17531/ein.2021.1.14>.

Induction motor fault classification via entropy and column correlation features of 2D represented vibration data

Indexed by:



Murat Basaran^a, Mehmet Fidan^a

^aVocational School of Transportation, Department of Motor Vehicles and Transportation Technologies, Eskisehir Technical University, 26140 Odunpazari, Eskisehir, Turkey

Highlights

- Induction motor bearing defects are classified due to the defect type as well as depth.
- Normalised and partitioned 1D vibration signals are converted into 2D greyscale images.
- 2D discrete wavelet transform is applied to greyscale images to create four sub-images.
- Novel entropy and column correlation-based features are extracted from these sub-images.
- Comparative classification success of the proposed feature extraction method is shown.

Abstract

Due to long-term use under challenging conditions, the sub-elements of induction motors may suffer certain defects over time. Such defects impair the vibration characteristics of the motors in different ways, depending on the type of defect. Therefore, the change in vibration characteristic provides indicators about the fault type and can be used in preventive maintenance strategies to ensure safe operation of the system. In this work, discrete-time vibration data were transformed into 2-dimensional grey-level images and decomposed into individual components by the Wavelet decomposition method. Features based on entropy and column correlation were extracted from these components and used to classify motor faults by using the Support Vector Machine method implemented by using the Sequential Minimal Optimisation algorithm. When the selected classifier is compared with other popular Machine Learning algorithms, it is observed that motor faults are more successfully classified, and these observations are presented in detail with comparative classification performance results.

Keywords

This is an open access article under the CC BY license (<https://creativecommons.org/licenses/by/4.0/>)

entropy, fault diagnosis, support vector machines, wavelet transforms.

1. Introduction

Electric motors consume nearly half of the electricity supply and may therefore be considered the workhorses of industry. Among all electric motors, induction motors — ranging from fractional horsepower to higher levels of industrial scale — are the most preferred electric motors in industrial applications due to their simple rugged construction, cost-effective price and ease of maintenance. Industrial companies are continuously pushed to continue their operations more reliably and efficiently, and therefore, applying predictive maintenance strategies such as vibration monitoring are essential to ensuring the continuity and quality of the production process.

Bearings contribute to the proper mechanical rotation of motors via their set of spherical or cylindrical rolling elements located between two circular rings called ‘races’ — one inner and one outer. Due to non-ideal operating conditions and ageing, incipient types of bearing defects may occur, which may further deteriorate and propagate on to the races and the rolling elements of the bearings.

The vibration frequencies caused by bearing-related defect types are formulated as functions of rotational frequency and bearing geometry metrics [34]. Bearing-related defects can be broadly classified as outer race defects (ORD), inner race defects (IRD), and ball defects (BD). The characteristic vibration frequencies related to these main defect types are:

$$f_{ORD} = (N_b / 2) f_r \left[1 - d_b \cos(\beta) / d_p \right] \quad (1)$$

$$f_{IRD} = (N_b / 2) f_r \left[1 + d_b \cos(\beta) / d_p \right] \quad (2)$$

$$f_{BD} = d_p f_r / 2 d_b \left[1 - \left(d_b \cos(\beta) / d_p \right)^2 \right] \quad (3)$$

where f_r is the rotational speed, N_b is the number of balls between the races, β is the contact angle of the ball with the races, and d_b and d_p are the ball diameter and the pitch diameter, respectively, as depicted in Fig. 1.

These characteristic frequencies are widely used to construct models of impulse trains induced by bearing related defects and the vibration responses are used in various fault classification techniques. However, these characteristic frequencies derived by kinematical relationship based on simple rolling motion and smooth rolling assumptions. But in reality, in the presence of a load between a rolling ball and the races, the contact surface is formed and ball rotates relative to the deformed surface of the races. Due to the loading conditions, the impulse trains show stochastic character rather than being strictly

E-mail addresses: M. Basaran - murath@eskisehir.edu.tr, M. Fidan - mfidan@eskisehir.edu.tr

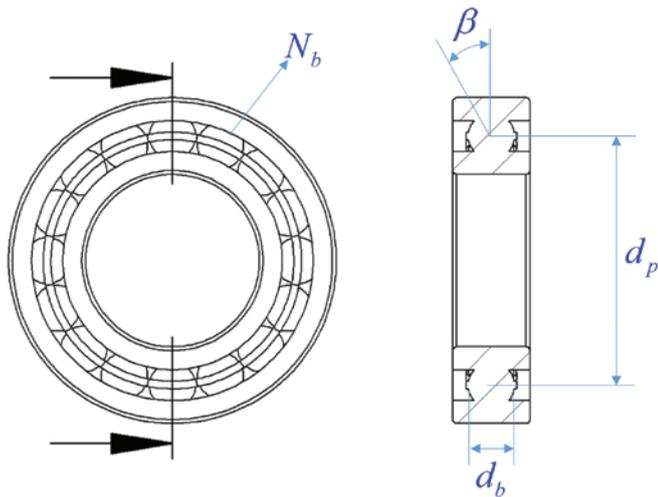


Fig. 1. Ball bearing geometry

periodical. Consequently, the characteristic defect frequencies are strongly depending on bearing metrics, loading conditions and lubrication level that has an important role in the contact angle condition and the relative slippage between balls and the raceway [26]. Therefore, more complex models not solely based on kinematic constraints but considering cyclostationary behaviour of the vibration signals is needed to enhance the effectiveness of the diagnosis.

Monitoring vibration signals is a widely used technique to detect incipient types of bearing defects. In the existing literature, numerous methods have been applied to detect and classify induction motor faults. In one of the first studies on this subject, Ocak and Loparo used the frequency of the fundamental harmonics of the vibration signals to predict motor fault type [27]. In another research study, Ocak and Loparo used the Hidden Markov Model (HMM) to separate motors with inner race defects, outer race defects and ball defects from healthy motors by using the same vibration data used in our research, and obtained maximally 100% accuracies for both class problems [28]. Nandi, Toliyat and Li investigated motor current harmonics instead of vibration, using Fast Fourier Transform (FFT) and clustering the motor faults by using this harmonic knowledge on an Artificial Neural Network (ANN); their results achieved accuracies as high as 93% with an uneasy current-measuring process [25]. Trajin, Regnier and Faucher compared the vibration-based motor fault diagnosis method with the motor stator current-based fault diagnosis method, and concluded that the vibration method is more appropriate under constant speed mode, while the current method is more appropriate for the variable speed mode of induction motors [40]. Immovilli, Bellini, Rubini and Tassoni also compared the vibration-based method with the current-based method, and concluded that although the current-based method is suitable only for low-frequency working conditions, the vibration-based method is suitable for both low- and high-frequency working conditions [17]. Lei and Meng proposed the Symplectic Entropy method with Radial Basis Function (RBF) classifier for vibration-based defect classification for four conditions: normal condition, outer race defect, rolling element defect and inner race defect, and achieved 99.8% test accuracy with 6,000 vibration samples [19]. In more recent research, Li, Wang, Si and Huang applied an entropy-based defect classification method to the same data used in our research, and achieved 98.75% accuracy for ball defect detection and 100% accuracy for inner and outer race defect detection [20]. In another recent study, Zhao, Liu and Meng proposed switchable normalisation semi-supervised generative adversarial networks (SN-SSGAN) for 1D representation of the same data used in our research, and achieved 99.93% test accuracy for four problem classes (normal, inner race defect, outer race defect and ball defect) while splitting data as 80% training to 20% testing [49]. Gan, Zhao and Chow also utilized vibration signals in electrical and mechanical motor fault detection under different frequency conditions

using genetic algorithms and achieved a maximum of 93.96% test accuracy for electrical faults under 35 Hz frequency and a maximum of 96.9% test accuracy for mechanical faults under 45 Hz frequency [9]. One of the most recent studies using vibration signals for diagnostics was presented by Tabaszewski and Szymański, which proposes a set of binary tree-based classification for three valve clearance classes, listed as: tight, optimum and excess [39]. In [39], the classification accuracy achieved was 99%.

Support Vector Machines (SVM) are also applied for motor fault classification problems [5, 10, 13, 23]. Banerjee and Das proposed a hybrid method of Short-Time Fourier Transform (STFT) and Linear SVM (LSVM) for motor fault detection with multiple sensors and achieved 95% test accuracy [5]. Glowacz used LSVM for the classification of healthy motors, motors with faulty rotor bar and motors with two faulty rotor bars by using acoustic data, and achieved a 96.66% total efficiency of recognition of acoustic signal (TEoRoAS) [13]. Gangsar and Tiwari applied one-vs-one multiclass SVM (MSVM) with RBF kernel for classification of nine different electrical and mechanical faults of an induction motor by using vibration and current data both separately and together, and achieved 98.3% test accuracy by using vibration and current data together under the condition of no-load and 20 Hz working frequency [10]. In a more recent study, Mao and Wang proposed an SVM method supported by Multi-Objective Particle-Swarm-Optimisation (MOPSO) and applied it to separate inner race defect data from normal condition data, outer race defect data from normal condition data, and ball defect data from normal condition data — and obtained 99.47%, 100% and 100% respective accuracies for the corresponding two-class problems [23].

One-dimensional time domain vibration signals can be converted into 2D via quantisation of actual values between 0 to 255 for greyscale image representation by using non-overlapping segments as rows of 2D matrices [11]. Representation in 2D has some significant benefits over the regular 1D signal representation [41]. Do and Chong proposed the Scale-Invariant Feature Transform (SIFT) algorithm to detect faults using 2D representation of the vibration signals, and achieved 98.1% accuracy for the eight-class classification problem, wherein the classes are listed as: angular misalignment, bowed rotor shaft, broken rotor bar, faulty bearing, motor unbalance, normal motor, parallel misalignment and phase unbalance [6]. If a priori knowledge of the classes to be recognised exists, texture-based methods can be applied for the classification of the patterns caused by indicators of vibration signals [18, 36, 48]. In one of these texture-based methods, Khan and Kim calculated Local Binary Patterns (LBP) from 2D representation of the vibration data used in our research, derived global histogram of these LBPs, and used the values in the global histogram as input of k-NN classifier to achieve an average classification accuracy of 99.74% [18]. 2D texture analysis based on local binary patterns (LBP) is also applied to vibration signals. Vibration data converted into 2D greyscale images are used as a medium to find discriminating texture features by employing an LBP operator [36]. In the other texture-based research, Zhang, Peng and Li used Convolutionary Neural Network (CNN) for the classification of motor faults from the 2D representation of the vibration data used in our research, and obtained 99.95% accuracy for separations of 30,000 training and 7,500 test data samples, and 98.17% accuracy for separations of 1,500 training and 7,500 test data samples, where each data sample contains 2,400 data points [48]. In one of the most recent studies, Sun and Cao integrated curvature filtering, Histogram of Gradients (HOG) and one-vs-one MSVM for classification of motor faults by using 2D representation of vibration data and achieved a maximum of 98.48% accuracy by using RBF kernel for MSVM [38]. In another recent study, Ma et al. applied Transfer Learning CNN (TLCNN) to 2D representation of the same data used in our research and achieved 99.71% accuracy [22]. Zimnickas, Vanagas, Dambrauskas and Kalvaitis built a test workbench for collecting vibration signals from several different induction motors whose conditions can be listed as: bad bearings, loose mounting, rotor eccentricity, lost phase to motor

and short circuit in stator winding, and they applied a hybrid method of Continuous Wavelet Transform (CWT) and CNN to 2D representations of their collected data [50]. In [50], 97.53% classification accuracy was achieved. Artificial Neural Network (ANN)-based bearing defect classification via vibration spectrum imaging is another application of greyscale to binary image conversion, showing the spectral contents of the translation-variant time-segmented vibration signal, transformed into a spectral image [3].

This article proposes a novel feature extraction method for column correlation and entropy features from Wavelet decomposition of 2D represented vibration. The vibration data of the induction motors used in the experiments are explained in detail in the introduction of Section 2. In Section 2.1, the construction of 2D greyscale images from 1D vibration data is explained. In Section 2.2, the 2D Haar Wavelet decomposition method is explained. Calculations of column correlation features and entropy features are explained in Section 2.3 and Section 2.4, respectively, and the combined feature vector representation is introduced in Section 2.5. In Section 3, the chosen MSVM classifier with Pearson VII (PUK) kernel is introduced. Section 4 presents comparative results for different classifiers using proposed features, and in Section 5 the presented comparative results are analysed.

2. Material and feature extraction method

In this work, the proposed feature extraction technique is tested on a publicly available seeded fault data set from Case Western Reserve University (CWRU) Bearing Data Center [1]. The test bench included a 2-horsepower (hp) induction motor, a torque transducer and a dynamometer used to acquire vibration signals using accelerometers, which were attached to the housing with magnetic bases. Single-point defects were deliberately introduced to the test bearings with three different defect diameters. In this work, 12 kHz drive-end vibration data of induction motors with defective bearings with defect diameters of 0.18, 0.36 and 0.54 millimetres of inner race defect, outer race defect located at the 6:00 position and ball defect bearing are used. Vibration data was taken under four different loading conditions: 0, 1, 2 and 3 hp motor loads. Class labels of defect types with different defect diameters are given in Table 1.

Table 1. Defect Descriptions for Classes

Class Label	Defect Description	Defect Depth Level
Class 1	Inner Race Defect	0.18 mm
Class 2	Inner Race Defect	0.36 mm
Class 3	Inner Race Defect	0.54 mm
Class 4	Ball Defect	0.18 mm
Class 5	Ball Defect	0.36 mm
Class 6	Ball Defect	0.54 mm
Class 7	Outer Race Defect	0.18 mm
Class 8	Outer Race Defect	0.36 mm
Class 9	Outer Race Defect	0.54 mm

All loading conditions of motors are included into the classification to propose a classification method that is effective regardless of loading.

Sample 1D vibration data related to inner race defect, ball defect and outer race defect with different defect diameters under 2 hp load are plotted in Figs. 2, 3 and 4, respectively.

2.1. Construction of 2D greyscale images

The first step of the proposed feature extraction method is to convert 1D discrete vibration data to 256-grey-level 2D images. The discrete 1D vibration data can be represented by the vector V_i as in (4) for the i^{th} experiment, where v_j denotes the j^{th} sample of the i^{th} data.

$$V_i = [v_i[0] \ v_i[1] \ v_i[2] \ \dots \ v_i[L-1]] \quad (4)$$

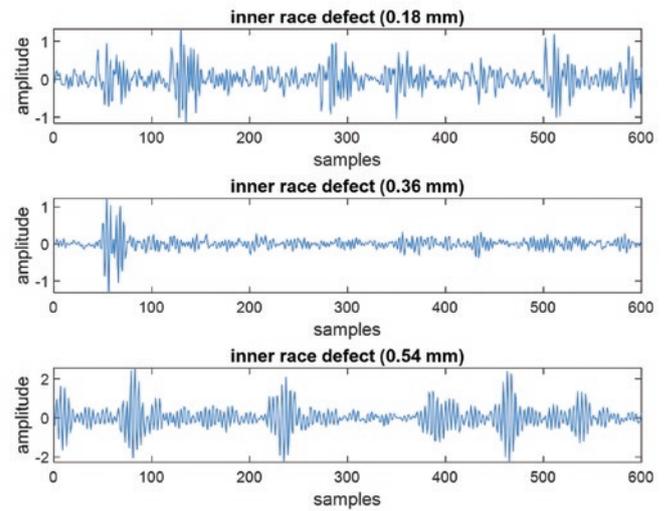


Fig. 2. Sample 1D inner race defect vibration data under 2 hp motor load

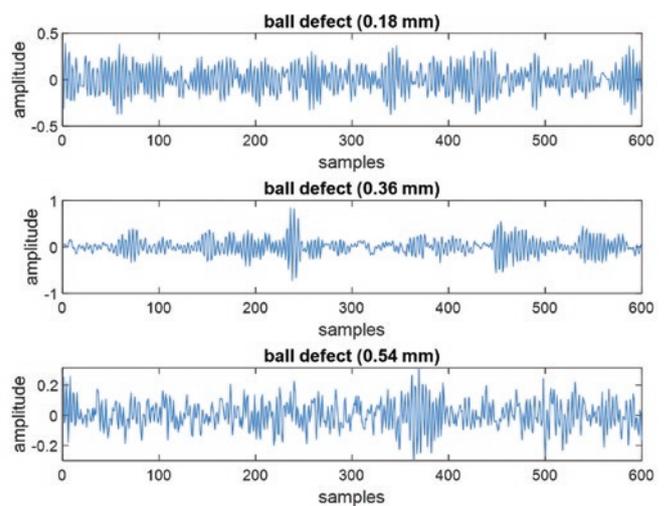


Fig. 3. Sample 1D ball defect vibration data under 2 hp motor load

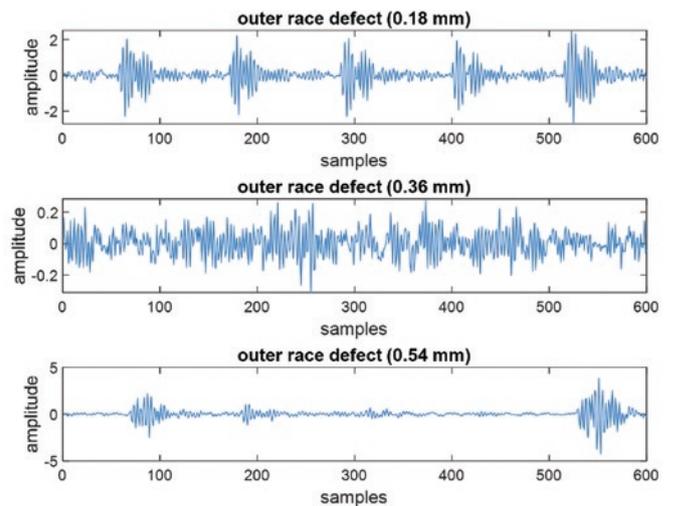


Fig. 4. Sample 1D outer race defect vibration data under 2 hp motor load

where: \hat{V}_i denotes the normalised vector for the i^{th} 1D vibration data, which is calculated as in (5). The subtraction of minimum valued element of the vector from each sample guarantees that the values will be nonnegative, and the division of each subtraction result by the difference of maximum and minimum elements guarantees that the values will fall within the interval [0,1]:

$$\hat{V}_i = \frac{V_i - \min(V_i)}{\max(V_i) - \min(V_i)} \quad (5)$$

The size selection of the vectors is another crucial step in the construction of 2D greyscale images:

$$L = M \times N \quad (6)$$

In (6), L denotes the number of samples in a single measurement vector, which is taken as 6,000. M is the number of rows in the image representation, which is taken as 30, which correspond to 30 cycles. N is the number of columns in the image representation, which is taken as 200. Since the sampling frequency is 12 kHz, the length N is taken as 200. The related experiments were carried out at a network frequency of 60 Hz. Thus, each part of the vibration data comprising 200 samples will correspond to a single period of the AC voltage applied at the motor input. In other words, if the sampling frequency is 12 kHz, by setting M and N values to 30 and 200 respectively, each greyscale image represents a vibration data recorded for 30 cycles, corresponding to half seconds of data, and the pixels in each row belong to a single period of that part. The normalised vector is converted to 2D 8-bit grey-level images as in (7) and (8).

$$I_i[m, :] = \text{Round}\left(255 \cdot \hat{V}_i[(mN) : ((m+1)N - 1)]\right) \quad (7)$$

$$I_i = \begin{bmatrix} I_i[0, :] \\ I_i[1, :] \\ \vdots \\ I_i[(M-1), :] \end{bmatrix} \quad (8)$$

It should be noted that 60 Hz is the fundamental supply frequency in all experiments, and the f_{ORD} , f_{IRD} and f_{BD} are the frequencies expressed in kinematical equations in (1), (2) and (3) respectively, are the approximate frequencies of the harmonics those may occur due to the relevant fault type. It should be noted that these harmonics do not always occur in a deterministic way and show stochastic behaviour due to external conditions. In addition, the depth of the defects may cause different harmonics in the vibration spectrum, as a result, the type of fault that cannot be explained solely by the characteristic equations related to fault types. For this reason, the N parameter used in 2D image representation should be calculated based on the fundamental frequency, not the approximate harmonic frequencies. The effect of harmonic frequencies depending on the fault type on the

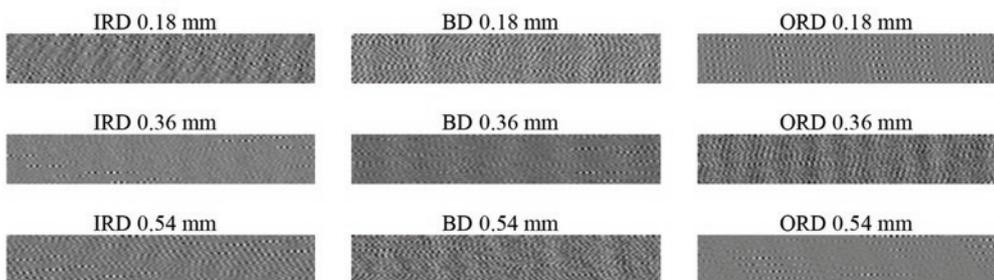


Fig. 5. Sample 2D image representations of vibration data

vibration data will be better understood with the texture structure that will be occurred specific to the fault type in 30x200 pixels 2D images. In this work, column correlation and entropy features which described detailly in Sections 2.3 and 2.4 respectively, suggested to employed together to detect texture type corresponds to fault type. Sample 2D images constructed from vibration data recorded under 2 hp motor loads for inner race defect, ball defect and outer race defect classes with different defect diameters are given in Fig. 5.

2.2. 2D Haar Wavelet Decomposition of constructed images

After constructing 20 images per data file downloaded from the CWRU database, single-level 2D Discrete Wavelet Transformation (DWT) is applied to these greyscale images in order to obtain the sub-band images needed for the proposed feature extraction technique [2, 4]. In DWT, High-Pass (HP) and Low-Pass filters (LP) and the down-sampling by two operations are applied along rows of 30 to 200 pixels images. Haar filters are selected for the Low-Pass and High-Pass filters [14, 37]. The High-Pass Haar filter is given in (9) and the Low-Pass Haar Filter is given in (10):

$$H_{Haar} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \quad (9)$$

$$L_{Haar} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \quad (10)$$

The same procedure is applied along the columns of the two sub-images obtained, with a nuance — this time, for the columns, the Low-Pass filter is applied before the High-Pass filter. The outputs at the second step are four greyscale images having sizes of 15 to 100 pixels, corresponding to single-level DWT of the image outputs. The output images are called the vertical, diagonal, approximate and horizontal sub-images, as shown in Fig. 6.

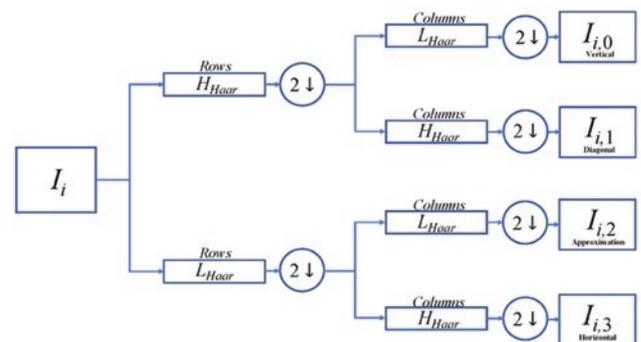


Fig. 6. 2D Haar wavelet decomposition

The order of the sub-images is determined by the strength of the information obtainable from the sub-image, with stronger information for the classification to the sub-images with weaker information.

Sample sub-images for the nine classes of vibration data are presented in Fig. 7 in order of vertical, diagonal, approximate and horizontal components. The approximate image resembles the original image, while remaining images contain

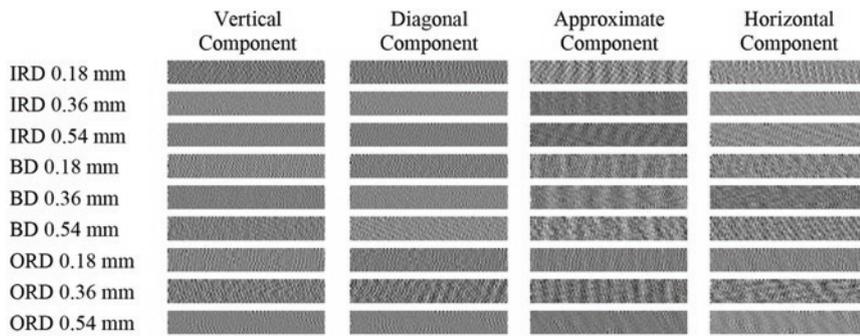


Fig. 7. Sample sub-images obtained for the nine classes of vibration data

detail textures — all of which have some close relationship with the indicators of motor faults.

2.3. Extraction of column correlation features

The correlations between two neighbouring columns are evaluated over the sub-images, and the means of these calculated values are used as column correlation features. The detailed analysis of the images obtained by wavelet decomposition reveals that the strongest information is captured in their columns, since each row corresponds to single period of vibration data. Consequently, there are insignificant variations between rows due to the periodicity. The column correlation information contains the potential distinguishing indicators of the fault types, which would be more informative. The n^{th} column of the j^{th} sub-image for the i^{th} image can be represented as in (11):

$$c_{i,j}[n] = I_{i,j}(0:(M_j - 1), n) \quad (11)$$

where $M_j = M/2$.

The Pearson correlation coefficient can be expressed as in (12), which can be computed by the covariance of two random variables and the variances of each random variable [30, 46]:

$$\rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X) \cdot \text{Var}(Y)}} \quad (12)$$

The column correlation is generated by the mean of Pearson correlations of the neighbouring columns in the sub-images, which can be expressed as in (13):

$$r_{\text{column}}(i, j) = \frac{\sum_{n=0}^{N_j-1} \rho(c_{i,j}[n], c_{i,j}[n+1])}{N_j - 1} \quad (13)$$

where $N_j = N/2$ [12].

2.4. Extraction of entropy features

Entropy is a strong measure for the randomness and texture of the images [15, 45]. However, for the measurement of entropy, the data samples or the pixels of the images should take a countable number of values. After the wavelet transform, the pixels of the sub-images take floating-point values, thereby making each value unique, which is not conducive to entropy calculation. To overcome this problem, the pixels of wavelet sub-images are rounded to the nearest integer value as in (14):

$$Q_{i,j} = \text{Round}(I_{i,j} - \min(I_{i,j})) \quad (14)$$

After this, a histogram is constructed as in (15) and (16):

$$h_{i,j}(k) = \text{count}(s_{i,j}[k]) \quad (15)$$

$$H_{i,j} = [h_{i,j}(0) \ h(1) \cdots \ h(NS_{i,j} - 1)] \quad (16)$$

where $s_{i,j}[k]$ is the k^{th} state in $Q_{i,j}$ and $NS_{i,j}$ is the total number of states observed in $Q_{i,j}$.

The histogram is an array of numbers of each state, which are generalised as in (17):

$$L_j = M_j \times N_j \quad (17)$$

The probabilities of each state are calculated by (18):

$$P(s_{i,j}[k]) = \frac{H_{i,j}[k]}{L_j} \quad (18)$$

According to the probability of each state, the entropy of each of the wavelet sub-images can be calculated as in (19):

$$\text{entropy}(i, j) = - \sum_{k=0}^{NS_{i,j}-1} [P(s_{i,j}[k]) \log_2(P(s_{i,j}[k]))] \quad (19)$$

2.5. Feature vector representation

After the calculation of column correlation and entropy features from sub-band images, features are combined to construct a feature vector of size eight. The overall feature extraction process can be summarised with the flowchart depicted in Fig. 8. For vibration data samples of size 6,000, 20 sets of feature vectors are extracted from each data file downloaded from the CWRU data set. Each set can be considered as a separate experiment and the feature vector for the i^{th} experiment can be represented by (20):

$$F_i = \begin{bmatrix} r_{\text{column}}(i, 0) \\ r_{\text{column}}(i, 1) \\ r_{\text{column}}(i, 2) \\ r_{\text{column}}(i, 3) \\ \text{entropy}(i, 0) \\ \text{entropy}(i, 1) \\ \text{entropy}(i, 2) \\ \text{entropy}(i, 3) \end{bmatrix} \quad (20)$$

Feature vectors are normalised as in (21)–(23):

$$\hat{F}_i[k] = \frac{F_i[k] - \min_k}{\max_k - \min_k} \quad (21)$$

where:

$$\min_k = \min([F_0[k] : F_{K-1}[k]]) \quad (22)$$

$$\max_k = \max([F_0[k] : F_{K-1}[k]]) \quad (23)$$

The scatter plots of selected normalised feature vectors for all experiments are presented in Fig. 9 for the Three-Class case, where

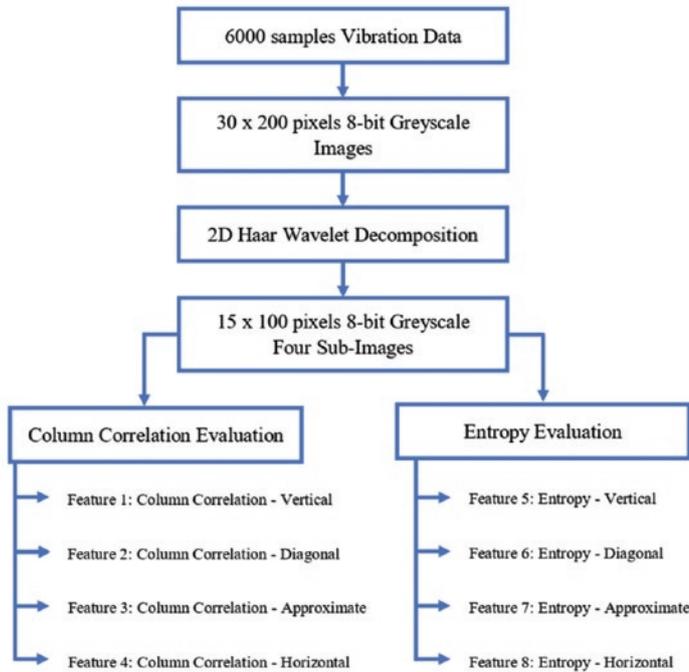


Fig. 8. Feature extraction flowchart

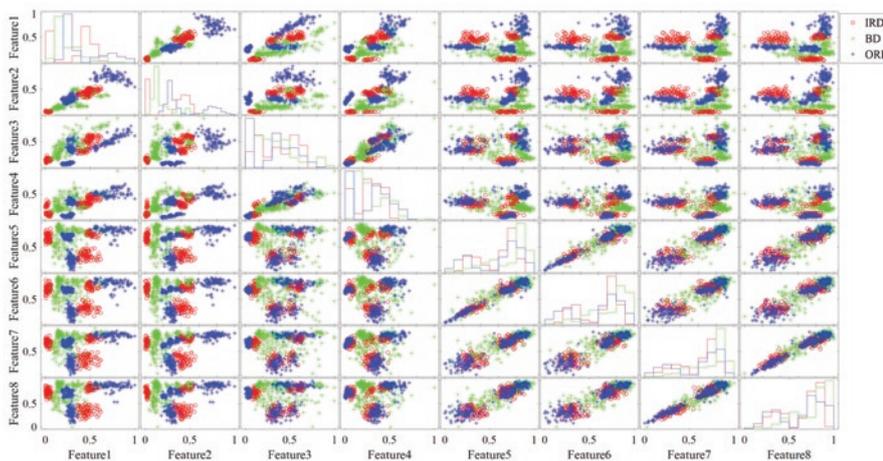


Fig. 9. Scatter plot of normalised features for Three-Class case

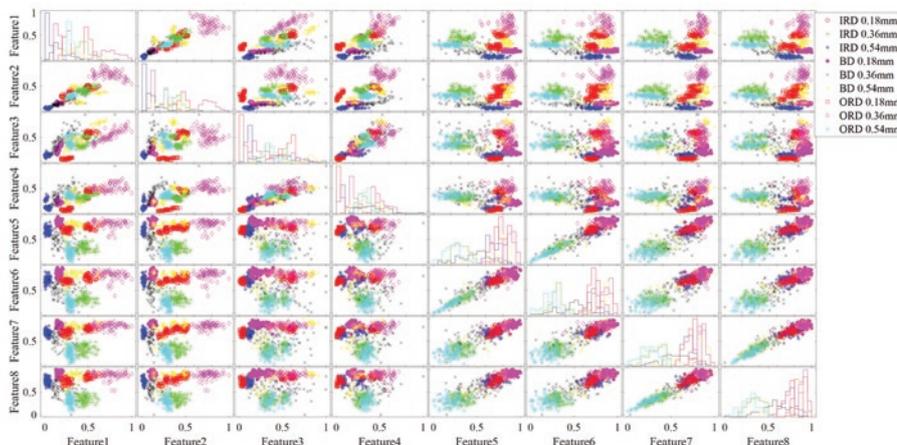


Fig. 10. Scatter plot of normalised features for Nine-Class case

defect diameters are ignored. Similarly, scatter plots of selected normalised feature vectors for all experiments for the Nine-Class case, where each defect diameter was considered as a separate class, as described in Table 1, are presented in Fig 10.

The contribution of column correlation features to the Three-Class problem can be observed by examining the 16 sub-figures in the upper-left quadrant of Fig. 9, which show that in particular, the column correlation coefficients group ball defect and outer race defect better than inner race defect. However, upon examining the 16 sub-figures in the upper-right quadrant and the 16 sub-figures in the lower-left quadrant of Fig. 9, it is seen that the inner race defect can be grouped significantly with the contribution of entropy features.

As seen in the 16 sub-figures in the upper-left quadrant of Fig. 10, it is understood that the most significant contributions of column correlation features to the Nine-Class problem are grouping with 0.18, 0.36, and 0.54 millimetres outer race defects and 0.54 millimetres inner race defects. When the 16 sub-figures in the lower-left quadrant are examined, it can be understood that 0.18, 0.36, and 0.54 millimetres ball defects can be grouped better when column correlation features are used together with entropy features.

When the 16 sub-figures in the lower-right quadrant of Fig. 10 are examined, it is seen that when entropy features alone are used, 0.18, 0.36, and 0.54 millimetres outer race defects can be largely distinguished from each other. However, entropy features can distinguish ball defects from inner race and outer race defects only with the help of column correlation features as seen in the 16 sub-figures in the upper-right quadrant of Fig. 10.

3. Support Vector Machine Classifier

The purpose of the Support Vector Machine (SVM) method is to obtain a suitable hyperplane in an N-dimensional space that distinctly separates the classes [35, 43]. The chosen hyperplane should have the maximum distance from both classes.

The original method was developed for two-class problems. However, the method has been adapted to multi-class problems using one of three potential strategies: one vs. one, one vs. all, and the non-heuristic method [24, 44]. In this study, the one vs. one strategy is chosen.

The algorithm is fastened by the Sequential Minimal Optimization (SMO) algorithm [32]. SVM is defined in terms of a quadratic programming (QP) problem. SMO decomposes this QP problem into several QP sub-problems and fastens the algorithm by choosing smallest possible optimisation problem at each step.

The classification accuracy of SVM is improved by the application of the Pearson VII Kernel (PUK) method [29, 31]. Üstün, Melssen and Buydens presented the accuracy of SVM with PUK on various test data for classification [42]. Zhang and Ge used SVM with PUK for the classification of halophilic and non-halophilic proteins [47]. The main reason for using PUK is the nonlinearity of the hyperplanes between classes. The PUK can be expressed as in (24) for two feature vectors:

$$K(\hat{F}_i, \hat{F}_j) = 1 / \left[1 + \left[\left(2\sqrt{|\hat{F}_i - \hat{F}_j|^2 \cdot \sqrt{2^{1/\omega} - 1}} \right) / \sigma \right]^2 \right]^\omega \quad (24)$$

The best classification performance is achieved when ω is taken as 1 and σ is taken as 0.5 in (24).

4. Comparative results

For the implementation of the chosen classifiers, the 3.8.2 version of the Workbench for Machine Learning (WEKA) is used [16]. The performance of the proposed feature extraction method with the chosen classifiers is compared with the Bayesian Network classifier [8], Naïve Bayes classifier [33], Naïve Bayes classifier with Kernel Density Estimation (KDE) [21], K-Nearest Neighbours (KNN) Classifier [7] and SMO with Polynomial Kernel [35].

4.1. Comparison metrics

Classification accuracies are compared using the following performance metrics: Accuracy, Macro-Precision, Macro-Recall and Macro-F1 Score.

These performance metrics are calculated from a confusion matrix. A confusion matrix for a multi-class problem can be generalised as in (25):

$$CM \begin{array}{c|cccc} & \hat{P}_1 & \hat{P}_2 & \dots & \hat{P}_N \\ \hline P_1 & TP_{11} & F_{12} & \dots & F_{1N} \\ P_2 & F_{21} & TP_{22} & \dots & F_{2N} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ P_N & F_{N1} & F_{N2} & \dots & TP_{NN} \end{array} \quad (25)$$

In (25), \hat{P}_i denotes the total number of samples predicted as i^{th} class and P_i denotes the exact number of samples in i^{th} class. TP_i denotes the number of correctly classified samples as i^{th} class (number of True Positive samples) and F_{ij} denotes the number of samples in i^{th} class that are wrongly classified as j^{th} class. The Accuracy metric is calculated from a confusion matrix as in (26):

$$Accuracy = \frac{\sum_{i=1}^N TP_i}{\sum_{i=1}^N P_i} \cdot 100\% \quad (26)$$

The calculations of Precision, Recall and F1 Score need False Positive (FP_i : the total number of samples that are wrongly classified as i^{th} class) and False Negative (FN_i : the total number of samples in i^{th} class that are wrongly classified as one of the other classes) numbers of each class. The calculation of FP_i is given in (27), and that of FN_i is given in (28):

$$FP_i = \sum_{\substack{j=1 \\ j \neq i}}^N F_{ji} \quad (27)$$

$$FN_i = \sum_{\substack{j=1 \\ j \neq i}}^N F_{ij} \quad (28)$$

The Precision of the classifier for a specific i^{th} class is the fraction of the number of samples correctly classified as i^{th} class out of the total number of samples classified as i^{th} class, which is defined by (29):

$$Precision_i = \frac{TP_i}{\hat{P}_i} \cdot 100\% = \frac{TP_i}{TP_i + FP_i} \cdot 100\% \quad (29)$$

The Recall of the classifier for a specific i^{th} class is the fraction of the number of samples correctly classified as i^{th} class among the total number of samples in i^{th} class, which is defined by (30):

$$Recall_i = \frac{TP_i}{P_i} \cdot 100\% = \frac{TP_i}{TP_i + FN_i} \cdot 100\% \quad (30)$$

The F1 Score of the classifier for a specific i^{th} class is the geometric mean of the Precision and Recall metrics for that class as shown in (31):

$$F1_i = \frac{2 \cdot Precision_i \cdot Recall_i}{(Precision_i + Recall_i)} \quad (31)$$

The data set used in experiments consists of equal numbers of samples belonging to each specific fault class. Therefore, the classification problem dealt with is a balanced multi-class problem. Since the classes are balanced, it is appropriate to use the macro average of the class-based metrics in measuring the overall classification performance.

The overall precision, recall and F1 score performances of the classifiers are measured by Macro-Precision, Macro-Recall and Macro-F1 Score metrics, which are equal to the arithmetic mean of the class-specific precision, recall and F1-scores as shown in (32), (33) and (34) respectively:

$$Macro-Precision = \frac{\sum_{i=1}^N Precision_i}{N} \quad (32)$$

$$Macro-Recall = \frac{\sum_{i=1}^N Recall_i}{N} \quad (33)$$

$$Macro-F1 = \frac{\sum_{i=1}^N F1_i}{N} \quad (34)$$

4.2. Results for Three-Class case

Table 2 shows overall the Three-Class classification success of the benchmarked classifiers for proposed features. Table 3 and Table 4 show detailed Three-Class performance of the benchmarked classi-

Table 2. Three-Class Classification Accuracy

Classifier	10-fold cross validation	80% split
Bayes Net with Naïve Bayes search	95.1389%	95.1389%
Naïve Bayes without kernel	51.9444%	58.3333%
Naïve Bayes with kernel	95.5556%	95.1389%
KNN	99.0278%	98.6111%
SMO with polykernel	70.5556%	77.7778%
SMO with PUK (Pearson VII kernel)	99.5833%	100.000%

Table 3. Three-Class Detailed Classification Performance Statistics with 10-Fold Cross Validation

Classifier	Macro-Precision	Macro-Recall	Macro-F1
Bayes Net with Naïve Bayes search	95.17%	95.14%	95.14%
Naïve Bayes without kernel	51.85%	51.94%	51.78%
Naïve Bayes with kernel	95.72%	95.56%	95.50%
KNN	99.04%	99.03%	99.02%
SMO with polykernel	69.53%	70.56%	69.59%
SMO with PUK (Pearson VII kernel)	99.58%	99.58%	99.58%

Table 4. Three-Class Detailed Classification Performance Statistics with 80% Split

Classifier	Macro-Precision	Macro-Recall	Macro-F1
Bayes Net with Naïve Bayes search	95.36%	95.15%	95.23%
Naïve Bayes without kernel	58.90%	58.93%	58.90%
Naïve Bayes with kernel	95.17%	95.54%	95.22%
KNN	98.55%	98.77%	98.64%
SMO with polykernel	76.77%	76.84%	75.88%
SMO with PUK (Pearson VII kernel)	100%	100%	100%

fiers according to the comparison metrics for 10-fold cross validation and a split of 80% training set and 20% test set, respectively.

4.3. Results for Nine-Class case

Table 5 shows overall Nine-Class classification success of the benchmarked classifiers for proposed features. Table 6 and Table 7 show detailed Nine-Class performance of the benchmarked classifiers according to the comparison metrics for 10-fold cross validation and a split of 80% training set and 20% test set, respectively.

Table 5. Nine-Class Classification Accuracy

Classifier	10-fold cross validation	80% split
Bayes Net with Naïve Bayes search	95.1389%	96.5278%
Naïve Bayes without kernel	95.2778%	97.9167%
Naïve Bayes with kernel	96.2500%	99.3056%
KNN	99.0278%	98.6111%
SMO with polykernel	91.6667%	91.6667%
SMO with PUK (Pearson VII kernel)	99.4444%	100.000%

4.4. Evaluation of the Results

When the literature is re-examined, there are similar studies that include feature extraction from the 2D representations of vibration data taken from the CWRU database or a self-designed test rig [3, 6, 18, 22, 36, 38, 48]. In the studies that use CWRU data, image representation of vibration signals differs in some aspects.

In [48], pixels of 60 to 40 sized images were directly used as 2400 inputs of a batch-normalized CNN and 99.95% accuracy for a separation of 30,000 training and 7,500 test data samples and 98.17% accu-

Table 6. Nine-Class Detailed Classification Performance Statistics with 10-fold Cross Validation

Classifier	Macro-Precision	Macro-Recall	Macro-F1
Bayes Net with Naïve Bayes search	95.28%	95.14%	95.11%
Naïve Bayes without kernel	95.46%	95.28%	95.34%
Naïve Bayes with kernel	96.48%	96.27%	96.29%
KNN	99.05%	99.03%	99.02%
SMO with polykernel	93.14%	91.67%	90.90%
SMO with PUK (Pearson VII kernel)	99.45%	99.44%	99.44%

Table 7. Nine-Class Detailed Classification Performance Statistics with 80% Split

Classifier	Macro-Precision	Macro-Recall	Macro-F1
Bayes Net with Naïve Bayes search	96.14%	96.76%	96.26%
Naïve Bayes without kernel	97.92%	98.36%	97.99%
Naïve Bayes with kernel	99.21%	99.47%	99.32%
KNN	98.25%	98.89%	98.50%
SMO with polykernel	91.34%	92.02%	90.19%
SMO with PUK (Pearson VII kernel)	100%	100%	100%

racy is achieved for a separation of 1,500 training and 7,500 test data samples, where each data sample contains 2400 data points.

In [18], average classification accuracy was obtained as 99.74% by LBP features, which is similar to the classification accuracy of our proposed method.

In [3], spectral imaging and filtering with appropriate threshold selection — which is more complicated than the proposed technique — was used to discriminate the same type of faults, and 96.9% average classification accuracy was obtained. However, it focused only on the fault types and neglected the defect depths.

In [38], the sizes of 2D representation of the vibration signal are 100 x 128, which has more data points than our proposed 2D representation (30 x 200) and obtained 98.48% accuracy by using one-vs-one MSVM with RBF kernel, which is less than the accuracy of our proposed method.

Ma et al. achieved 99.71% accuracy by using TLCNN [22]; however, in our proposed method 100% accuracy can be achieved for both Three-Class and Nine-Class problems without the complex structure of CNN and additional knowledge used for transfer learning.

Classifying bearing-related faults, as well as classifying the same type of faults with different defect depths, makes the classification process more challenging when compared to the studies using a self-designed test rig where the fault types are very distinct from each other. In [6], the classification accuracy of the SIFT algorithm, which yields feature vectors of size 128, remained at 98.1%. In [36], the LBP technique, which uses histogram bins of size 256 as feature vectors, was used to discriminate the same fault types as in [6], and 100% accuracy was measured by 4-fold cross validation. Even though the classification accuracy of [36] is similar to that of our proposed method, the feature vector size and computational effort are higher than in our method. On the other hand, we proposed only 8 features based on 2D Discrete Wavelet Transform to discriminate similar bearing-related defects with different depths and achieved remarkably high classification accuracies up to 100%.

5. Discussion

The Three-Class classification problem comprises challenge of the ball defect separation from the inner and outer race defects. The difficulty of this separation problem can be observed in Fig. 9. According to the figures, it can be emphasised that the ball defect data is spread over a larger area in the feature space than the data of the other two classes, which may potentially cause confusion. On the other hand, the inner race data bunch along three narrow regions in the feature space, and this situation expedites the distinction of the inner race defect from other defect types. The outer race data bunch along three other narrow regions different from the inner race regions; therefore, the outer race defect type is also distinct from the inner race defect type. The inner race and outer race defects can be separated by linear hyperplanes. However, the ball defect needs a nonlinear hyperplane. As also seen in (3), the characteristic vibration frequency of the ball defect is more complicated than that of other defect types. Therefore, it is understandable that a ball defect is harder to discriminate from other defect types.

According to the overall accuracies of the classifiers on Three-Class classification, the worst performance is observed for the Naïve Bayes without KDE. The simple structure of pure Naïve Bayes is sufficient to avoid both the confusion between ball defect and the other classification types, as well as the confusion between inner and outer race defects. This confusion problem is mostly solved by application of KDE to Naïve Bayes. After the application of KDE, the overall success of the Naïve Classifier increased from 58.33% to 95.14% for the 80% training – 20% testing strategy. In addition, confusion is observed only between ball defect and others after KDE. Bayesian Network shows similar performance with Naïve Bayes with KDE. SMO with simplest polynomial kernel with an exponent of 1 shows relatively worse performance than Naïve Bayes with KDE.

The SMO with polynomial kernel shows confusions mostly between inner race defect and the others. The cause of this confusion could be bunching along three narrow regions which cannot be separated by first-degree polynomial kernel. The performance of SMO is boosted by choosing PUK, which shows 100% sufficiency for Three-Class classification for the 80% train – 20% test strategy. The closest performance to SMO with PUK is observed for KNN classifier, which offers little confusion on ball defect.

References

1. Case Western Reserve University Bearing Data Center Website. <https://csegroups.case.edu/bearingdatacenter/pages/welcome-case-western-reserve-university-bearing-data-center-website>, (accessed 12-22-2019).
2. Akansu AN, Haddad R A. Multiresolution Signal Decomposition: Transforms, Subbands and Wavelets. Academic Press, 2001: 396-401.
3. Amar M., Gondal I., and Wilson C. Vibration Spectrum Imaging: A Novel Bearing Fault Classification Approach. IEEE Transactions on Industrial Electronics 2015; 62(1): 494-502, <https://doi.org/10.1109/tie.2014.2327555>.
4. Arivazhagan S, Ganesan L. Texture classification using wavelet transform. Pattern Recognition Letters 2003; 24(9-10): 1513-1521, [https://doi.org/10.1016/S0167-8655\(02\)00390-2](https://doi.org/10.1016/S0167-8655(02)00390-2).
5. Banerjee TP, Das S. Multi-sensor data fusion using support vector machine for motor fault detection. Information Sciences 2012; 217: 96-107, <https://doi.org/10.1016/j.ins.2012.06.016>.
6. Do V, Chong UP. Signal Model-Based Fault Detection and Diagnosis for Induction Motors Using Features of Vibration Signal in Two-Dimension Domain. Stroj Vestn-J Mech E 2011; 57(9): 655-666, <https://doi.org/10.5545/sv-jme.2010.162>.
7. Duda RO, Hart PE. Pattern Classification and Scene Analysis. The Library Quarterly 1973; 44(3): 258-259, <https://doi.org/10.1086/620282>.
8. Friedman N, Geiger D, Goldszmidt M. Bayesian Network Classifiers. Machine Learning 1997; 29(2-3): 131-163, <https://doi.org/10.1023/A:1007465528199>.
9. Gan Z, Zhao M-B, Chow T. W. Induction machine fault detection using clone selection programming. Expert Systems with Applications 2009; 36(4): 8000-8012, <https://doi.org/10.1016/j.eswa.2008.10.058>.
10. Gangsar P, Tiwari R. Comparative investigation of vibration and current monitoring for prediction of mechanical and electrical faults in induction motor based on multiclass-support vector machine algorithms. Mechanical Systems and Signal Processing 2017; 94: 464-481, <https://doi.org/10.1016/j.ymssp.2017.03.016>.
11. Gerek ON, Ece DG. 2-D analysis and compression of power-quality event data. IEEE Transactions on Power Delivery 2004; 19(2): 791-798, <https://doi.org/10.1109/tpwr.2003.823197>.
12. Germen E, Basaran M, Fidan M. Sound based induction motor fault diagnosis using Kohonen self-organizing map. Mechanical Systems and Signal Processing 2014; 46(1): 45-58, <https://doi.org/10.1016/j.ymssp.2013.12.002>.
13. Glowacz A. Recognition of acoustic signals of induction motor using FFT, SMOFS-10 and LSVM. Eksploatacja i Niezawodność -

If the Nine-Class performances of the classifiers with proposed feature extraction methods are analysed, the first pattern that stands out is that the confusions in the Three-Class classification are mostly avoided, especially for the worst classifiers in Three-Class classification, as Naïve Bayes without Kernel and SMO with polykernel. The one-vs-one hyperplanes in the Nine-Class problem have simple structures than the one-vs-one hyperplanes in Three-Class problem, because each level of the inner race and outer race defects are bunched at specific locations, which can be easily separated from other type of defects as seen in Fig. 10. According to Tables 5–7, despite the remarkable improvement in scores, the SMO with polykernel becomes the worst classifier in Nine-Class classification. Naïve Bayes without KDE achieves the success of Bayesian Network, and Naïve Bayes with KDE exceeds the success of Bayesian Network. The best performance in the Nine-Class classification is also observed in chosen SMO with PUK classifier, which is followed by KNN classifier.

If the success of the proposed features in both Three-Class and Nine-Class problems is evaluated over the precision, recall and F1 metrics, there is no significant difference between the precision and recall values, as can be seen in Tables 3, 4, 6 and 7. Because of the closeness of precision and recall performances, the F1 scores are remarkably close to the accuracy scores. The precision, recall and F1 metrics prove that the proposed features are sufficient to obtain consistent and reliable fault classification results.

6. Conclusions

To sum up, the proposed feature extraction technique provides strong information, not only for the Nine-Class fault classification problem, but also for the Three-Class fault classification problem. In addition, if these features are supported by a kernel-based classifier with a suitable kernel, they provide great improvement — up to 100% correct classification. The type of kernel is crucial, as observed in the comparison of polykernel and PUK.

In future work, the method can be improved upon in two different ways. First, some additional features can be proposed, especially for separating ball defects from other types of defects. Alternatively, a different kernel can be proposed that is much more suitable for the hyperplanes between separated classes.

- Maintenance and Reliability 2015; 17(4): 569-574, <https://doi.org/10.17531/ein.2015.4.12>.
14. Haar A. Zur Theorie der orthogonalen Funktionensysteme. *Mathematische Annalen* 1910; 69(3): 331-371, <https://doi.org/10.1007/bf01456326>.
 15. Haralick RM, Shanmugam K, Dinstein I. Textural Features for Image Classification. *IEEE Transactions on Systems, Man, and Cybernetics* 1973; SMC-3(6): 610-621, <https://doi.org/10.1109/tsmc.1973.4309314>.
 16. Holmes G, Donkin A, Witten IH. Weka: A machine learning workbench. *Second Australia and New Zealand Conference on Intelligent Information Systems* 1994, In: *Proceedings of ANZIIS '94*, <https://doi.org/10.1109/ANZIIS.1994.396988>.
 17. Immovilli F, Bellini A, Rubini R, Tassoni C. Diagnosis of bearing faults in induction machines by vibration or current signals: A critical comparison. *IEEE Transactions on Industry Applications* 2010; 46(4): 1350-1359, <https://doi.org/10.1109/TIA.2010.2049623>.
 18. Khan SA., Kim J-M. Automated Bearing Fault Diagnosis Using 2D Analysis of Vibration Acceleration Signals under Variable Speed Conditions. *Shock and Vibration* 2016; 2016: 1-11, <https://doi.org/10.1155/2016/8729572>.
 19. Lei M, Meng G, Dong GM. Fault Detection for Vibration Signals on Rolling Bearings Based on the Symplectic Entropy Method. *Entropy* 2017; 19(11), <https://doi.org/10.3390/e19110607>.
 20. Li Y, Wang X, Si S, Huang S. Entropy Based Fault Classification Using the Case Western Reserve University Data: A Benchmark Study. *IEEE Transactions on Reliability* 2019: 1-14, <https://doi.org/10.1109/tr.2019.2896240>.
 21. Liu B, Yang Y, Webb GI, Boughton J. A Comparative Study of Bandwidth Choice in Kernel Density Estimation for Naive Bayesian Classification. *Lecture Notes in Computer Science* 2009; 5476: 302-313, https://doi.org/10.1007/978-3-642-01307-2_29.
 22. Ma P, Zhang H, Fan W, Wang C, Wen G, Zhang X. A novel bearing fault diagnosis method based on 2D image representation and transfer learning-convolutional neural network. *Measurement Science and Technology* 2019; 30(5): 055402, <https://doi.org/10.1088/1361-6501/ab0793>.
 23. Mao W, Wang L, Feng N. A New Fault Diagnosis Method of Bearings Based on Structural Feature Selection. *Electronics* 2019; 8(12), <https://doi.org/10.3390/electronics8121406>.
 24. Mayoraz E, Alpaydin E. Support Vector Machines for Multi-Class Classification. *IWANN99* 1999, In: *Engineering Applications of Bio-Inspired Artificial Neural Networks*, *Lecture Notes in Computer Science* 1607, <https://doi.org/10.1007/BFb0100551>.
 25. Nandi S, Toliyat HA, Li XD. Condition monitoring and fault diagnosis of electrical motors - A review. *IEEE Transactions on Energy Conversion* 2005; 20(4): 719-729, <https://doi.org/10.1109/Tec.2005.847955>.
 26. Niu L, Cao H, He Z, Li Y. A systematic study of ball passing frequencies based on dynamic modeling of rolling ball bearings with localized surface defects. *Journal of Sound and Vibration* 2015; 357: 207-232, <https://doi.org/10.1016/j.jsv.2015.08.002>.
 27. Ocak H, Loparo KA. Estimation of the running speed and bearing defect frequencies of an induction motor from vibration data. *Mechanical Systems and Signal Processing* 2004; 18(3): 515-533, [https://doi.org/10.1016/S0888-3270\(03\)00052-9](https://doi.org/10.1016/S0888-3270(03)00052-9).
 28. Ocak H, Loparo KA. HMM-Based Fault Detection and Diagnosis Scheme for Rolling Element Bearings. *Journal of Vibration and Acoustics* 2005; 127(4): 299-306, <https://doi.org/10.1115/1.1924636>.
 29. Pearson K. Contributions to the Mathematical Theory of Evolution. II. Skew Variation in Homogeneous Material. *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character* 1895; 186: 343-414, <https://doi.org/10.1098/rsta.1895.0010>.
 30. Pearson K. Mathematical Contributions to the Theory of Evolution. III. Regression, Heredity, and Panmixia. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 1896; 187: 253-318, <https://doi.org/10.1098/rsta.1896.0007>.
 31. Pearson K. Mathematical Contributions to the Theory of Evolution. X. Supplement to a Memoir on Skew Variation. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 1901; 197(287-299): 443-459, <https://doi.org/10.1098/rsta.1901.0023>.
 32. Platt JC. Sequential Minimal Optimization: A Fast Algorithm for Training Support Vector Machines. Microsoft, MSR-TR-98-14, 1998.
 33. Rish I. An Empirical Study of the Naive Bayes Classifier. *IJCAI Workshop on Empirical Methods in AI*. 2001.
 34. Schoen RR, Lin BK, Habetler TG, Schlag JH, Farag S. An unsupervised, on-line system for induction motor fault detection using stator current monitoring. *IEEE Transactions on Industry Applications* 1995; 31(6): 1280-1286, <https://doi.org/10.1109/28.475698>.
 35. Schölkopf B, Smola AJ. *Learning with Kernels*. The MIT Press, Cambridge, Massachusetts, London, England 2001, <https://doi.org/10.7551/mitpress/4175.001.0001>.
 36. Shahriar MR, Ahsan T, Chong U. Fault diagnosis of induction motors utilizing local binary pattern-based texture analysis. *Eurasip J Image Video* 2013, <https://doi.org/10.1186/1687-5281-2013-29>.
 37. Stanković RS, Falkowski BJ. The Haar wavelet transform: its status and achievements. *Computers & Electrical Engineering* 2003; 29(1): 25-44, [https://doi.org/10.1016/s0045-7906\(01\)00011-8](https://doi.org/10.1016/s0045-7906(01)00011-8).
 38. Sun W, Cao X. Curvature enhanced bearing fault diagnosis method using 2D vibration signal. *Journal of Mechanical Science and Technology* 2020; 34(6): 2257-2266, <https://doi.org/10.1007/s12206-020-0501-0>.
 39. Tabaszewski M, Szymański G. Engine valve clearance diagnostics based on vibration signals and machine learning methods. *Eksplotacja i Niezawodność - Maintenance and Reliability* 2020; 22(2): 331-339, <https://doi.org/10.17531/ein.2020.2.16>.
 40. Trajin B, Régner J, Faucher J. Comparison between vibration and stator current analysis for the detection of bearing faults in asynchronous drives. *IET electric power applications* 2010; 4(2): 90-100, <https://doi.org/10.1049/iet-epa.2009.0040>.
 41. Tuceryan M, Jain AK. *Texture Analysis*. The Handbook of Pattern Recognition and Computer Vision. World Scientific Publishing Co., 1998: 207-248.
 42. Üstün B, Melsens WJ, Buydens LMC. Facilitating the application of Support Vector Regression by using a universal Pearson VII function based kernel. *Chemometrics and Intelligent Laboratory Systems* 2006; 81(1): 29-40, <https://doi.org/10.1016/j.chemolab.2005.09.003>.
 43. Vapnik V. The Support Vector method. *ICANN'97* 1997, In: *Artificial Neural Networks*, *Lecture Notes in Computer Science* 1327: 261-271, <https://doi.org/10.1007/BFb0020166>.
 44. Weston J, Watkins C. Multi-class Support Vector Machines. Royal Holloway University of London, Egham, Surrey TW20 0EX, England, Technical Report CSD-TR-98-04, 1998.
 45. Weszka JS, Rosenfeld A. An application of texture analysis to materials inspection. *Pattern Recognition* 1976; 8(4): 195-200, [https://doi.org/10.1016/0031-3203\(76\)90039-x](https://doi.org/10.1016/0031-3203(76)90039-x).
 46. Wright S. Correlation and Causation. *Journal of Agricultural Research* 1921; 20(7): 557-585.

47. Zhang G. and Ge H. Support vector machine with a Pearson VII function kernel for discriminating halophilic and non-halophilic proteins. *Comput Biol Chem* 2013; 46: 16-22, <https://doi.org/10.1016/j.compbiolchem.2013.05.001>.
48. Zhang W, Peng G, Li C. Bearings Fault Diagnosis Based on Convolutional Neural Networks with 2-D Representation of Vibration Signals as Input. *MATEC Web of Conferences* 2017; 95, <https://doi.org/10.1051/mateconf/20179513001>.
49. Zhao D, Liu F, Meng H. Bearing fault diagnosis based on the switchable normalization SSGAN with 1-D representation of vibration signals as input. *Sensors* 2019; 19(9): 2000, <https://doi.org/10.3390/s19092000>.
50. Zimmnickas T, Vanagas J, Dambrauskas K, Kalvaitis A. A Technique for Frequency Converter-Fed Asynchronous Motor Vibration Monitoring and Fault Classification, Applying Continuous Wavelet Transform and Convolutional Neural Networks. *Energies* 2020; 13(14): 3690, <https://doi.org/10.3390/en13143690>.