

ORIGINAL ARTICLE

Open Access



Connected Components-based Colour Image Representations of Vibrations for a Two-stage Fault Diagnosis of Roller Bearings Using Convolutional Neural Networks

Hosameldin O. A. Ahmed¹ and Asoke K Nandi^{1,2*}

Abstract

Roller bearing failure is one of the most common faults in rotating machines. Various techniques for bearing fault diagnosis based on faults feature extraction have been proposed. But feature extraction from fault signals requires expert prior information and human labour. Recently, deep learning algorithms have been applied extensively in the condition monitoring of rotating machines to learn features automatically from the input data. Given its robust performance in image recognition, the convolutional neural network (CNN) architecture has been widely used to learn automatically discriminative features from vibration images and classify health conditions. This paper proposes and evaluates a two-stage method RGBVI-CNN for roller bearings fault diagnosis. The first stage in the proposed method is to generate the RGB vibration images (RGBVIs) from the input vibration signals. To begin this process, first, the 1-D vibration signals were converted to 2-D grayscale vibration images. Once the conversion was completed, the regions of interest (ROI) were found in the converted 2-D grayscale vibration images. Finally, to produce vibration images with more discriminative characteristics, an algorithm was applied to the 2-D grayscale vibration images to produce connected components-based RGB vibration images (RGBVIs) with sets of colours and texture features. In the second stage, with these RGBVIs a CNN-based architecture was employed to learn automatically features from the RGBVIs and to classify bearing health conditions. Two cases of fault classification of rolling element bearings are used to validate the proposed method. Experimental results of this investigation demonstrate that RGBVI-CNN can generate advantageous health condition features from bearing vibration signals and classify the health conditions under different working loads with high accuracy. Moreover, several classification models trained using RGBVI-CNN offered high performance in the testing results of the overall classification accuracy, precision, recall, and F-score.

Keywords: Bearing fault diagnosis, Image representation of vibrations, Deep learning, Convolutional neural networks

1 Introduction

In most production procedures in manufacturing, roller bearings need to be maintained in a healthful condition to guarantee the steadiness of production. Thus, it is essential to monitor the health condition of roller

bearings to avoid machine breakdowns. Bearings may be categorised into two key types: (i) plain (sliding) bearings; and (ii) rolling bearings. Of these, rolling bearings are commonly used in most applications of rotating machinery [1]. Vibration-based condition monitoring has been extensively studied and has become a well-accepted method for planned maintenance management as various typical features can be observed from vibration signals. In general, with these features, machine learning classifiers can be utilised to identify machine health

*Correspondence: asoke.nandi@brunel.ac.uk

¹ Department of Electronic and Electrical Engineering, Brunel University London, Uxbridge UB8 3PH, UK

Full list of author information is available at the end of the article

conditions. However, the extracted features are typically distorted with noise and measurement errors that make it practically challenging to obtain distinguishable data that are well generalised. Therefore, considerable literature can be found around the theme of vibration signals feature extraction and feature selection for machine fault diagnoses.

It is now well established from a variety of studies, that vibration signal analysis can be performed in three main groups - time domain, frequency domain, and time-frequency domain. Various time domain-based techniques are used for vibration signal analyses. For instance, most of the time domain techniques are used to extract features from the raw vibration signals for bearings fault diagnoses using statistical functions as well as some other advanced functions [2–11]. A considerable amount of literature has been published on the practice of frequency-domain techniques to extract various spectrum features from vibration signals that can efficiently represent a bearing's health condition. These studies showed that the frequency domain analysis techniques can reveal information from vibration signals that are not easy to be observed in the time domain. For example, Fourier analysis including Fourier series, Discrete Fourier Transform (DFT), and Fast Fourier Transform (FFT) techniques are used to transform time-domain vibration signals to the frequency domain [12–18]. Moreover, various techniques are used to extract different spectrum features to represent a bearing's health condition. For instance, envelop analysis that is also called high-frequency resonance is evaluated for detecting incipient faults of bearings [19]. Furthermore, various frequency domain features based on high-order spectra techniques are utilised to represent the bearing's health condition [20, 21].

The time-frequency domain-based methods such as short-time Fourier transform (STFT), wavelet transform (WT), Hilbert-Huang transform (HHT), local mean decomposition (LMD), empirical mode decomposition (EMD), which are introduced for nonstationary waveform signals, are used to extract features from vibration signals for bearing fault diagnosis [22–32]. Several classification methods, such as logistic regression (LR), artificial neural networks (ANNs), and support vector machines (SVMs), can be utilised to classify different vibration signals based on the extracted features [1]. In case the features are sensibly formulated, and the parameters of the classification methods are wisely tuned, it is possible to achieve high classification accuracy. Nevertheless, extracting useful features from such a huge and noisy vibration dataset, which may also contain measurement errors, is usually a challenging task. Recently, several lines of evidence suggest that feature-learning methods that can automatically learn

representations of the vibration dataset can be a solution to address this challenge. Deep learning (DL) that usually learns representations of the data using a hierarchical multi-layer data processing architecture has been attracting a lot of interest. For example, Autoencoder-based Deep neural networks (DNNs) methods are used for bearings fault diagnosis in several studies [33–39].

Moreover, the literature on the application of DL-based techniques for machine fault diagnosis has highlighted several studies describing the use of deep belief networks (DBNs) for bearings fault diagnosis [40–44]. Furthermore, the application of recurrent neural networks (RNNs)-based techniques in bearings fault diagnosis was investigated by several researchers [45–48]. In the same vein, several studies used convolution neural networks (CNNs)-based algorithms to process vibration signals for bearings fault diagnosis [49–54]. Most of these studies applied pre-processing techniques such as FFT, WT, time-domain statistical functions, spectral kurtosis, to extract features from the raw vibration signals and used them as the input to the targeted DL technique, while others used the raw vibration signals directly as the input to the targeted DL technique. However, all the previously mentioned methods suffer from some serious limitations. For example (1) feature extraction from fault signals requires expert prior information and human labour; (2) it is sometimes hard to recognize faults features using only time-domain features, only frequency domain features, and only time-frequency domain features; and (3) the CNNs deep architecture was originally modeled for 2-D signals such as images and their application to 1-D signals such as vibrations was not straightforward.

Lately, researchers have shown increased interest in transforming the 1-D vibration signal into a 2-D image, which can often offer more discriminative descriptions of the vibration signals and allows direct usage of the CNN for fault diagnosis. For instance, Chong proposed a method for induction motors utilizing features of vibration signals in the two-dimension domain. In this method, the 2-D features of the vibration signal are achieved using the scale-invariant feature transform (SIFT) [55]. In Ref. [56] an ANN classifier with vibration spectrum imaging (VSI) is used for bearing fault classification where the vibration signal is first divided into time segments and adapted into an image. Then, the spectral contents of each image are computed and normalized to form a spectral image using FFT. Afterwards, to enhance features of the obtained spectral images, an average filter and binary threshold techniques are used to retain featured patterns and remove noise patterns. Finally, ANN is used as a fault classifier using these enhanced features of the faults.

Moreover, Kang and Kim presented a method for fault diagnosis of multiple induction motor faults using a 2-D representation of Shannon Wavelet. In this method, first, wavelet coefficients deduced from the Shannon wavelet function with dilation and translation parameters are used to create 2-D gray-level images. Then, the texture features of the created images are utilised as inputs to a multi-class support vector machine (SVM) classifier to identify faults in the induction machine [57]. Li et al. [58] presented a method for bearing fault using spectrum images of vibration signals. In this method, first, the FFT is used to obtain the spectrum images then each image is processed with 2-D principal component analysis (2DPCA) to reduce the dimensions. Finally, a minimum distance method is employed to classify bearing faults. Lu et al. [59] proposed a fault diagnosis method for rotating machinery using image processing. In this method, first, the bi-spectrum technique is used to transform the vibration signal into a bi-spectrum contour map. Then, the speeded-up robust features (SURF) detector and descriptor technique is employed to extract automatically features from the transformed bi-spectrum contour map. After, the t-Distributed Stochastic Neighbor Embedding technique is used to reduce the dimensionality of the generated feature vectors. Finally, with these reduced features, the probabilistic neural network is used for fault identification. Verstraete et al. [60] presented a method for rolling element bearing fault diagnosis using time-frequency representations and CNN. In this method, to validate the ability of the proposed CNN model to accurately diagnose bearings fault, three time-frequency techniques, i.e., STFT, WT, and HHT, are used to generate different representations of the raw signal. Then, these representations are separately fed into a CNN architecture for fault classification. The classification accuracy results of the three representations are compared to study their representation effectiveness.

Additionally, a vibration imaging and deep learning-based feature engineering technique for rotor systems fault diagnosis is proposed. In this technique, first, vibration signals are collected from sensors in the rotor systems then vibration images are prepared to be used as input to deep learning architecture. The vibration images are generated by first producing signals from virtual vibration sensors then the individual vibration signals are stacked to form the vibration images based on a phase synchronization rule. After, the vibration images are enhanced using the histogram of oriented gradients (HOG) descriptor technique. Then, the pretraining of the DBN is used to extract high-level features from the generated vibration images. Finally, a fault classifier that is based on fine-tuning the pre-trained DBN by combining it with a multilayer perceptron (MLP) is used for fault

diagnosis [61]. Zhang et al. [62] presented a technique for bearing fault diagnosis using CNN with a 2-D representation of vibration signals. In this technique, first, the raw vibration signal is divided into n equal parts and each part is aligned as the row of the 2-D image representation in sequence. Then, the obtained 2-D representations of the vibration signals are used as input to a CNN architecture for fault classification [62]. Hasan and Kim [63] proposed a method for bearing fault diagnosis under variable rotational speeds using Stockwell transform-based imaging and transfer learning techniques. In this method, the discrete orthonormal Stockwell transform (DOST)-based vibration imaging is used as a preprocessing step to generate health patterns. Then, a CNN-based transfer learning approach is used for fault diagnosis.

In 2019, Wang et al. [64] proposed a fault recognition technique based on multi-sensor data fusion and bottleneck layer optimized CNN (MB-CNN). In this technique, the vibration signals of several sensors are fused in the feature maps. Then, MB-CNN is used to extract features and deal with the fault recognition of rotating machinery. Yan et al. [65] proposed a fault diagnosis method for an active magnetic bearing-rotor system using vibration images. In this method, three features, histogram of vibration image (HVI), the histogram of oriented vibration image (HOVI), and 2-D FFT of vibration image (2DFFT) are designed based on signal amplitude, signal phase, and frequency domain, respectively. Then, a feature fusion technique called two-layer AdaBoost is introduced to train the fault recognition model. Moreover, Hoang and Kang presented a method for rolling element bearing fault diagnosis using CNN and vibration image. In this method, the amplitude of the sample in the vibration signal is normalized into the range $[-1, 1]$ and then each normalized amplitude becomes the intensity of the corresponding pixel in the corresponding vibration image. After, with these vibration images, a CNN architecture is used for fault diagnosis [66].

Furthermore, Zhu et al. [67] proposed a method for bearing fault diagnosis using CNN based on a capsule network with an inception block (ICN). In this method, first, the raw data are changed from a one-dimensional signal to a two-dimensional graph using STFT. Then, the ICN, which is proposed to address the problem of poor generalization of CNN models in diagnosing bearing faults under different loads, is used to deal with the fault diagnosis. Ma et al. [68] presented a bearing fault diagnosis method using 2-D image representation and transfer learning-based CNN (TLCNN). In this method, first, the time-domain raw signals are reconstructed to a fault signal component using the frequency slice wavelet transform (FSWT), then the reconstructed signals are converted into a 2-D time-frequency image. With these

time-frequency images, the TLCNN model is used to extract features and achieve the classification conditions of bearing health. Zhu et al. [69] proposed a method for rotor fault diagnosis using CNN with symmetrized dot pattern (SDP) images. In this method, vibration signals are transformed into SDP images then the graphical features of the SDP images of different vibration states are learned and identified using a CNN architecture.

Recently, Zhang et al. [70] presented an enhanced CNN for bearing fault diagnosis method based on time-frequency image. In this method, the STFT is used to obtain the input images. Then, the obtained images are used as input to a CNN architecture with the scaled exponential linear unit (SELU). Kaplan et al. [71] presented a feature extraction method for bearing fault diagnosis using texture analysis with local binary patterns (LBP). In this method, first, bearing vibration signals are converted to grayscale images. Then, the LBP technique is employed to obtain texture features. Finally, the obtained features are used as input to different classifiers such as K-nearest neighbor (K-NN), random forest (RF), Naive Bayes, Bayes Net, and ANN to deal with the classification problem.

Two important themes emerge from the studies discussed above: (1) Given its robust performance in image recognition, the CNN deep learning architecture has been used in most of the studies discussed above; and (2) taken together, the fault diagnosis results from these studies indicate that they may be further enhanced by considering two main factors: (a) how well the vibration images are generated from the vibration signals; and (b) how efficiently the generated vibration images reveal different patterns of each vibration signal status. On the question of image feature extraction techniques that can be applied to the generated vibration images, several types of features can be extracted from the generated vibration images such as texture, colour, shape, pixel intensity, etc. Hence, the more characteristics we have within the generated vibration images the more robust features can be extracted, and consequently, the more accurate learned classification model can be achieved.

This paper proposes a two-stage method for bearing fault diagnosis based on RGB vibration image representation and CNN (RGBVI-CNN). In the first stage, a technique for generating image representation from vibration signals that can successfully represent the bearing health condition is proposed. This technique uses image analytic techniques to generate efficient vibration image signals with rich characteristics using three main steps: (1) convert the 1-D vibration signals to 2-D gray-level vibration Images; (2) find the region of interest (ROI) in the binary images of the converted 2-D gray-level vibration images; and (3) generate RGBVIs based on connected components of the ROI of each vibration image that

demonstrate useful characteristics of the targeted vibration signals. In the second stage, the RGBVIs with their texture and colour features are used as inputs to a CNN architecture to further learn useful features, which are used to obtain an accurate classification model for bearing faults.

This study aims to contribute to this growing area of research by exploring the efficacy of the two-stage RGBVI-CNN method in generating vibration images with useful characteristics that can distinguish different bearing health conditions using vibration signals. The contributions of this paper are summarized as follows:

- 1) A new three-step approach of image analytics techniques is proposed in the first stage of the RGBVI-CNN, which produces 2-D RGBVIs with advantageous texture and colour features of bearing health conditions from the 1-D time-series vibration signals. The approach does not require any prior knowledge or any programmed parameters.
- 2) The visualized texture and colour features of the RGBVIs that are generated in the first stage of the RGBVI-CNN method can visually offer discriminative patterns of bearing health conditions.
- 3) A CNN-based deep learning architecture with three feature learning blocks is proposed to automatically learn features from the RGBVIs and to achieve improved classification accuracy for bearing health conditions.

The rest of this paper is structured as follows. Section 2 is dedicated to a description of the proposed method. Section 3 is devoted to a description of the performed experiments and datasets of the two case studies of bearing fault classification and the corresponding experimental results. Finally, Section 4 draws some conclusions from this study.

2 The Proposed Method

This section presents the details of the RGBVI-CNN method for bearing fault diagnosis. The RGBVI-CNN is a two-stage method that generates RGB vibration images with useful features that can be used as inputs to a CNN architecture for bearing fault diagnosis. The flow chart of the proposed two-stage RGBVI-CNN fault diagnosis method is shown in Figure 1. In the first stage, a proposed three-step based approach of image analytic techniques is utilised to generate efficient connected components based RGB vibration image signals with rich characteristics that can efficiently represent bearing health condition. In the second stage, a CNN architecture is applied to accomplish two main tasks: (1) to learn further features from the RGB vibration images generated in the

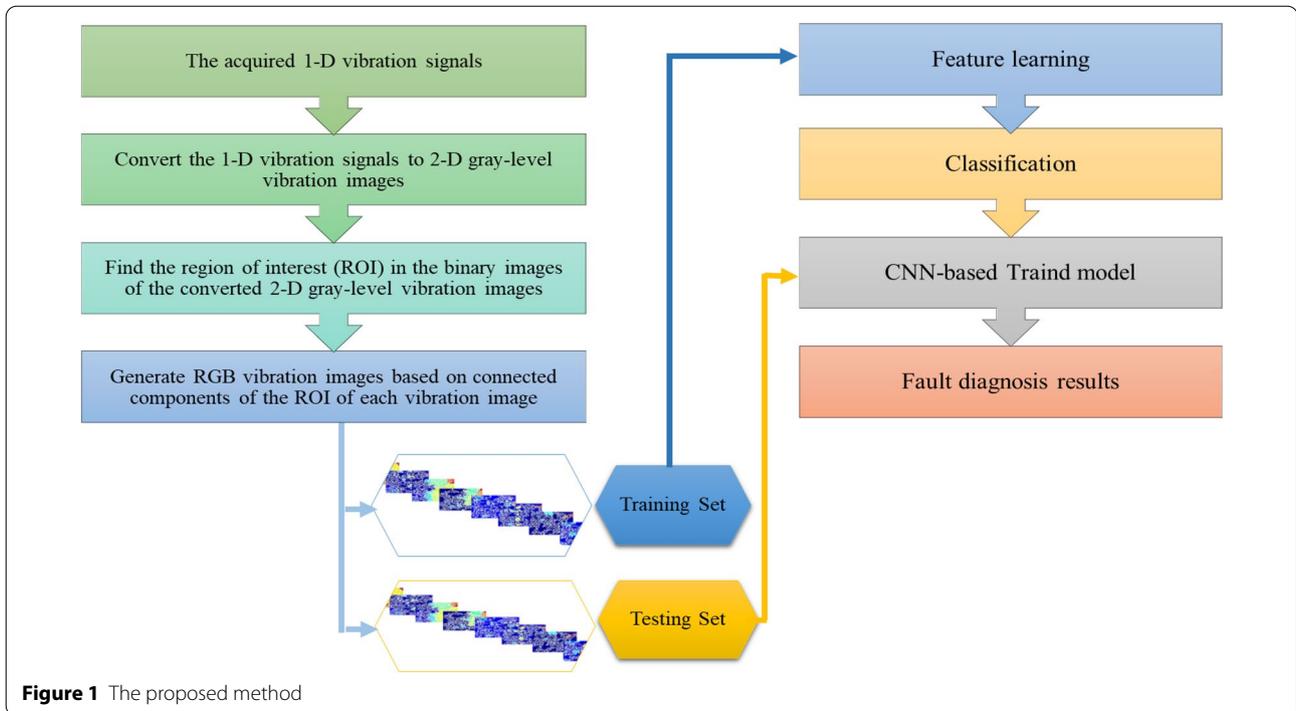


Figure 1 The proposed method

first stage; and (2) to classify bearing health condition utilizing the learned features. The subsequent subsections describe these two stages in more detail.

2.1 First Stage: RGBVI Generation

In this stage, RGBVI-CNN generates RGB based vibration images (RGBVIs) using a three-step approach of image analytic techniques as follows.

2.1.1 Conversion of 1-D Vibration Signal to 2-D Vibration Image

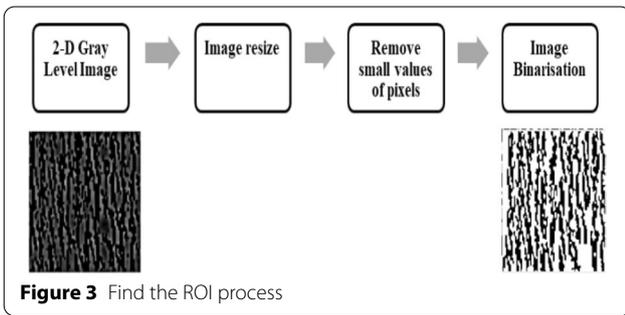
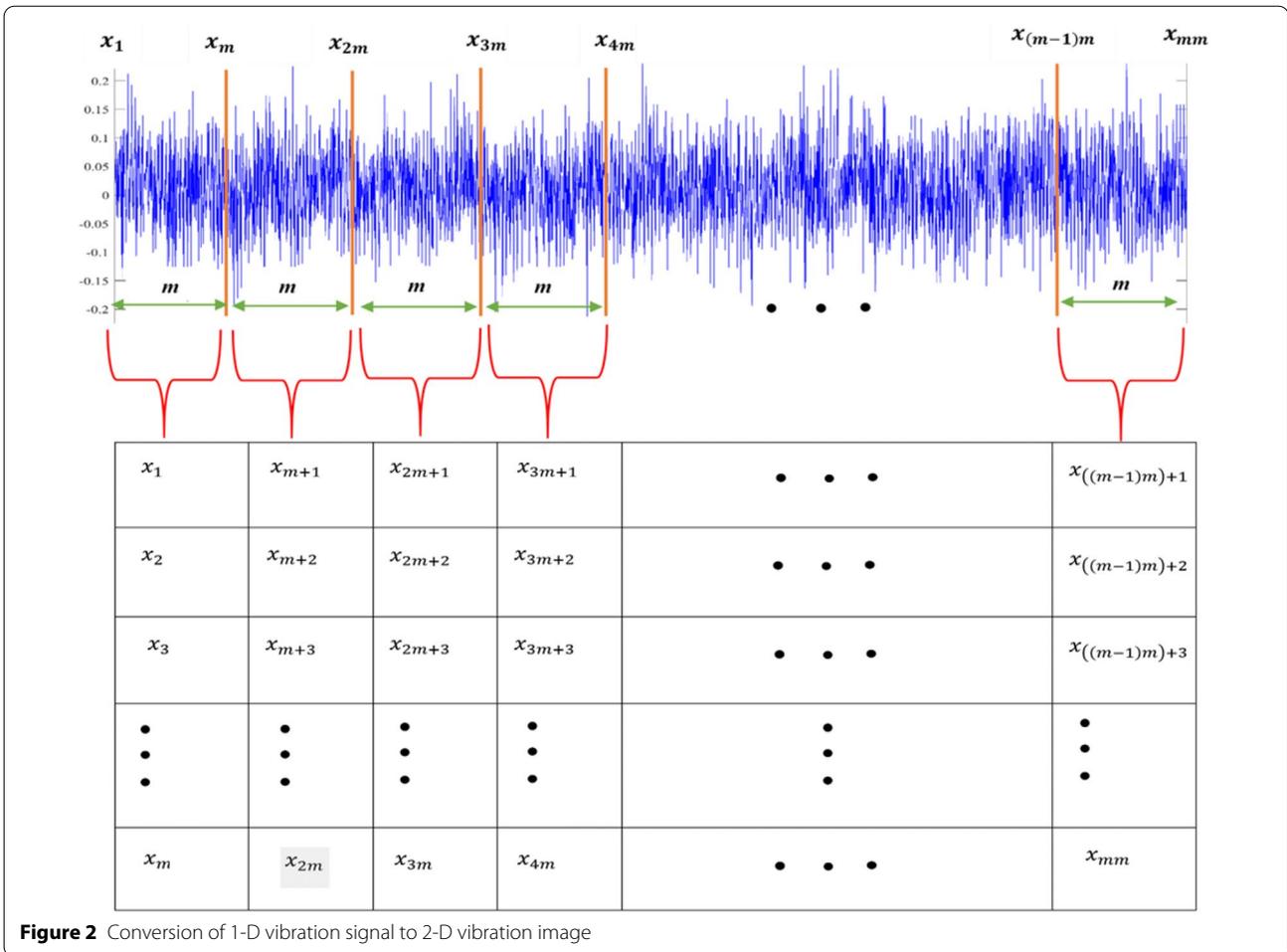
As described in Section 1, various techniques have been used to convert the 1-D vibration signals into 2-D grayscale image representations, ranging from simple methods that directly reshape the 1-D vibration signal into a 2-D image, to transformation based techniques, such as FFT, STFT, WT, HHT, SIFT, DOST, FSWT, which transform the 1-D time series vibration signal to frequency or time-frequency domain first and then construct the 2-D image representation. In this step of the first stage of the proposed method, we used a technique to directly reshape the 1-D time series vibration signal into a 2-D grayscale vibration image representation. In this technique, the raw vibration signal is divided into m equal segments and each segment is aligned as the column of an $m \times m$ 2-D image representation. Briefly, we describe this technique as follows. Assume that we have the original 1-D vibration signal vector $x(t)$ where $x(t) \in R^{1 \times n}$. As

shown in Figure 2, to construct a 2-D grayscale vibration image representation of $m \times m$ matrix dimensions $I \in R^{m \times m}$ from $x(t) \in R^{1 \times n}$ where n denotes the length of the 1-D vibration signal. We need to segment x into m equal segments where $m = \sqrt{n}$ (i.e., $n = m^2$) and each segment is associated as the column of the image I such that

$$I = \begin{bmatrix} x(t)_1 & \cdots & x(t)_{((m-1)m)+1} \\ \vdots & \ddots & \vdots \\ x(t)_m & \cdots & x_{mm} \end{bmatrix}. \tag{1}$$

2.1.2 Find the Region of Interest in the Converted 2-D Vibration Image

It is desirable to find the region of interest (ROI) in the converted 2-D grayscale vibration images that can offer useful characteristics from which different bearing health conditions can be classified. Hence, in this step of the first stage of RGBVI-CNN, an effective algorithm for finding ROI is used. As shown in Figure 3, the algorithm starts with the image resizing step where the size of the converted 2-D vibration image is increased so that more pixel information will be added. Then, for easy processing, the 2-D vibration images are converted to double type. Afterwards, to smooth the converted 2-D vibration image and to alleviate noises and enhance its feature, we applied an average filter, which is also called the mean filter. Here, the pixel value in the image matrix of the



converted 2-D vibration image is replaced by the average value of the targeted pixel and neighbors' pixels values, which reduce the intensity variation between one pixel and its neighbors' pixels. Compared to other types of filters such as median filters, the average filter is simple and easy to implement for image smoothing based on a kernel to represent the shape and size of the neighborhood to be sampled while the mean is computed. The most commonly used kernel with an average filter is a square

kernel of 3×3 , 5×5 , 7×7 , 9×9 , etc. [72, 73]. The average filter (AF) with a square kernel of $Z \times Z$ can be expressed mathematically as in Eq. (2):

$$AF = \frac{ones(Z, Z)}{(ZZ)} \tag{2}$$

To remove small pixel values, including both positive and negative values, the smoothed image is then subtracted from the 2-D grayscale vibration image. Following this treatment, we converted the resulted 2-D grayscale vibration image to binary in order to find the ROI. To achieve this, we employed a global image threshold using Otsu's technique, which automatically selects an optimal threshold from the viewpoint of discriminant analysis [74].

2.1.3 Generation of Colour Images from Vibrations

Intending to produce a 2-D vibration image with useful discriminative characteristics in this step we applied an algorithm to generate connected components-based

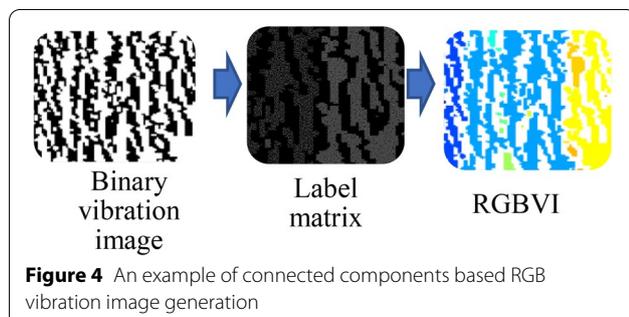
RGB vibration images (RGBVIs) from the binary images generated in the previous step. The connected components labeling process plays a vital role in object extraction in binary image analysis [75]. Therefore, in this step, our algorithm includes three main processes: (1) find connected components in the binary image using 2-D connectivity of 8 where eight pixels are connected if their boundaries or corners touch; (2) create label matrix from the connected components in the binary image with unique values; and (3) convert the created label matrix into RGB color image of labeled regions, which offers a set of colour and texture features. Figure 4 shows an example of connected components based RGB vibration image generated from a binary vibration image.

2.2 Second Stage: CNN Based Fault Diagnosis

A CNN, also known as convNet, is a multi-stage neural network (NN) that is usually composed of an input layer, a convolutional layer, a sub-sampling layer also called a pooling layer, fully connected layers, and an output layer. The CNNs perform two main tasks: (1) feature learning and (2) classification. They learn features from alternating and stacking convolutional layers, activation layer, and pooling process, which apply procedures to learn features from the data. The classification task can then be achieved using the learned features with the fully connected layer and the output layer that is commonly a SoftMax layer. The convolution layers convolve multiple local filters with raw input data and generate invariant local features and the pooling layers extract the most significant features [76, 77]. The convolution computation can be described mathematically as follows [68]:

$$h_j = f\left(\sum_i X_i * W_{ij} + b_j\right). \tag{3}$$

Here, h_j specifies the j th output feature map of the present convolutional layer, X_i denotes the i th output feature of the preceding convolutional layer, $*$ represents the convolution operator, W_{ij} maps the convolution kernel relating to the i th input feature map to the j th output feature



map in the present layer, b_j is the bias of the j th feature kernel, and f is the activation function. The most commonly used activation function is Rectified linear unit (ReLU) that can be described mathematically in Eq. (4):

$$f(x) = \begin{cases} 0, & \text{if } x < 0, \\ x, & \text{if } x \geq 0. \end{cases} \tag{4}$$

The pooling layer performs nonlinear downsampling to reduce the output dimensionality using one of the pooling techniques, such as maximum pooling, averaging pooling, and random pooling. Of these techniques, the maximum pooling, which can be described mathematically in Eq. (5), is commonly utilised

$$X_j = f(\alpha_j \text{down}(X_i) + b_j). \tag{5}$$

Here, X_j is the j th output of the current pooling layer, α_j represents a constant applied to control the change of the data by the pooling layer, $\text{down}(X_i)$ is the down-sampling procedure of the i th output of the previous layer, b_j is the bias of the j th feature kernel of the current pooling layer, and f is the activation function.

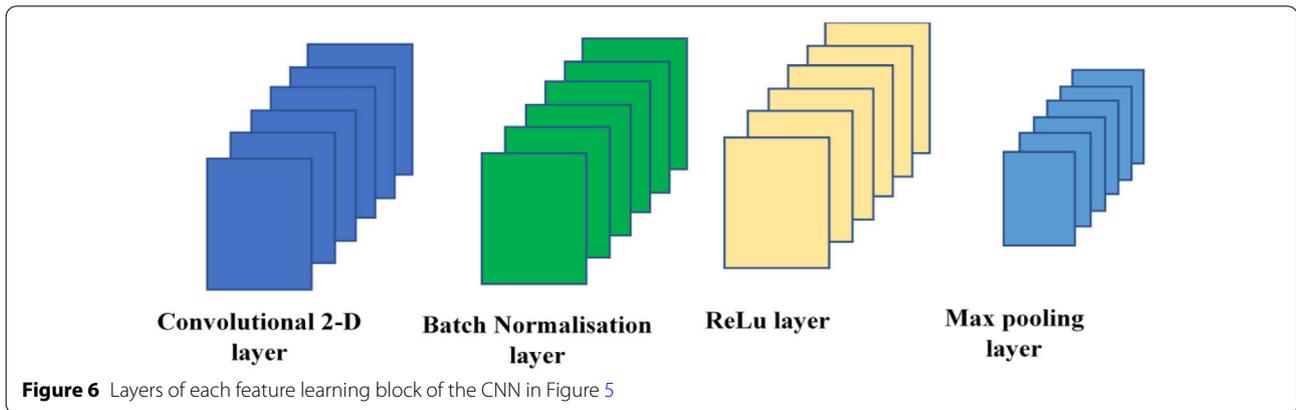
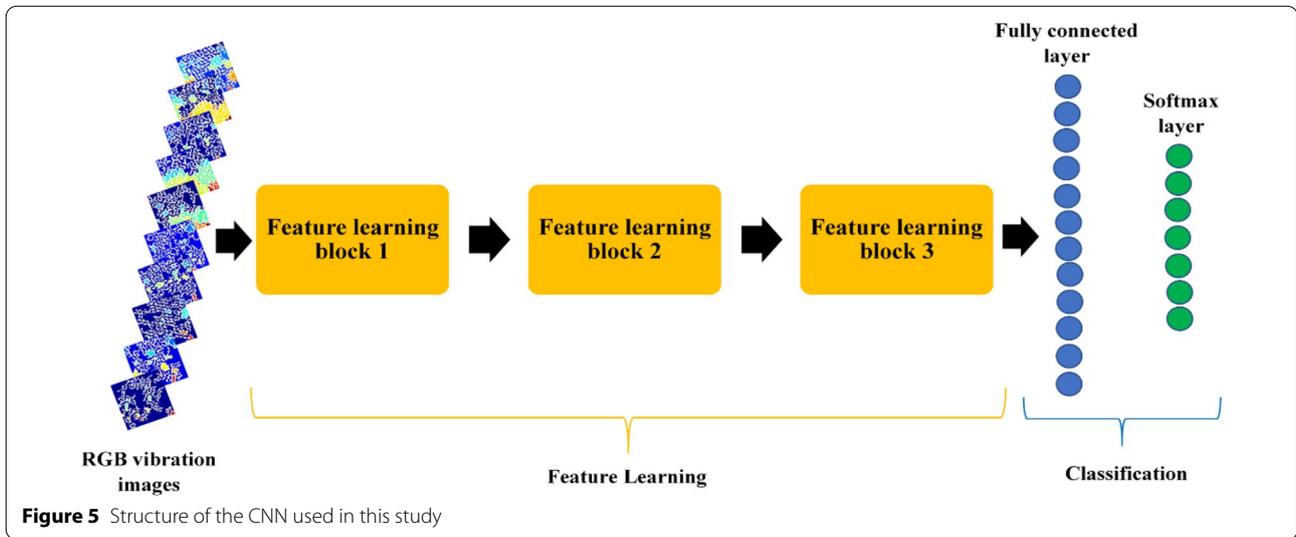
With the learned features, a CNN performs the classification task using the fully connected layer and the output layer that is also called the classifier layer, which is often added as top layers of CNNs to perform predictions. The most used classifier is SoftMax classifier. Briefly, we can describe the SoftMax classifier as follows: Assume we have a training set $\{(x^{(1)}, c^{(1)}), \dots, (x^{(L)}, c^{(L)})\}$ of L labeled examples and input features $x^{(i)} \in R^k$ with multi-labels $c^{(i)} \in \{1, \dots, c\}$. To estimate the probability $P(c = c^{(i)} | x)$ for each value of $c^{(i)} = 1$ to c , such that

$$h_\theta(x) = \begin{bmatrix} P(c = 1 | x; \theta) \\ P(c = 2 | x; \theta) \\ \vdots \\ P(c = K | x; \theta) \end{bmatrix} = \frac{1}{\sum_{j=1}^K \exp(\theta^{(j)T} x)} \begin{bmatrix} \exp(\theta^{(1)T} x) \\ \exp(\theta^{(2)T} x) \\ \vdots \\ \exp(\theta^{(K)T} x) \end{bmatrix}. \tag{6}$$

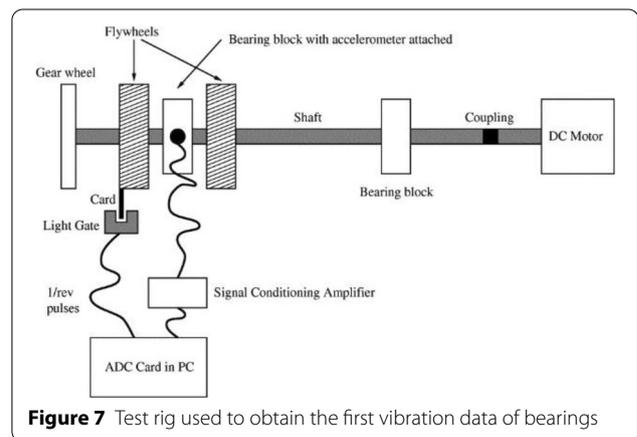
Here, θ contains model parameters that are trained to minimize the cost function $J(\theta)$ defined by

$$J(\theta) = - \left[\sum_{i=1}^L c^{(i)} \log h_\theta(x^i) + (1 - c^{(i)}) \log(1 - h_\theta(x^i)) \right], \tag{7}$$

where $\theta^{(1)}, \theta^{(2)}, \dots, \theta^{(K)} \in R^n$ are the parameters of the SoftMax model.



The CNNs have been demonstrated to be successful in many applications such as medical imaging, object recognition, speech recognition, visual document analysis [78–81]. In the second stage of our proposed method, we employed CNNs to deal with the classification problem of bearing health condition based on the RGB vibration image generated in the first stage described above. The CNNs are chosen for their reliability and validity in image classification. They are mainly advantageous for finding patterns in images for detection and classification purposes. The architecture of the proposed CNN model is presented in Figure 5. Each feature learning layer in Figure 5 consists of four layers: (1) convolutional 2-D layer; (2) Batch Normalisation layer; (3) ReLu layer; and (4) maximum pooling layer as shown in Figure 6.



3 Experimental Study

Two case studies of bearing vibration datasets generated using different health conditions in roller bearings are used to validate the efficacy of RGBVI-CNN in diagnosing bearing faults.

3.1 The First Case Study

The first bearing vibration dataset is obtained from experiments on a test rig that simulates running roller bearings' environment. In these experiments, several substitutable faulty roller bearings are implanted in the test rig to represent the type of faults that can usually happen in roller bearings. The test rig (Figure 7) used to get the first vibration dataset of bearings consists of a 12 V DC electric motor driving the shaft through a flexible coupling. The shaft was supported by two Plummer bearing blocks where a series of damaged bearing were implanted. Two accelerometers were utilised to measure the resultant vibrations in both the horizontal and vertical planes. The outputs from the accelerometers were fed back through a charge amplifier to a Loughborough Sound Images DSP32 ADC card employing a low-pass filter by means of a cut-off of 18 kHz. The sampling rate was 48 kHz, giving slight oversampling. Six health conditions of roller bearings have been recorded with two normal conditions, i.e., brand new condition (NO) and worn but undamaged condition (NW), and four faulty condition including, inner race fault (IR), an outer race fault (OR), rolling element fault (RE), and cage fault (CA). Table 1 shows a description of the matching characteristics of these bearing health conditions.

The data recorded using 16 different speeds in the range 25–75 rev/s. In each speed, ten-time series were recorded for each condition, i.e., 160 examples per condition. This resulted in a total of 960 examples to work with. Figure 8 shows some typical time series plots for the six different conditions.

As shown in Figure 8, each fault modulates the vibration signals by their unique patterns. For example, based on the level of damage to the rolling element and the loading of the bearing, IR and OR fault conditions have a

fairly periodic signal, RE fault condition may or may not be periodic, and CA fault condition generates a random distortion.

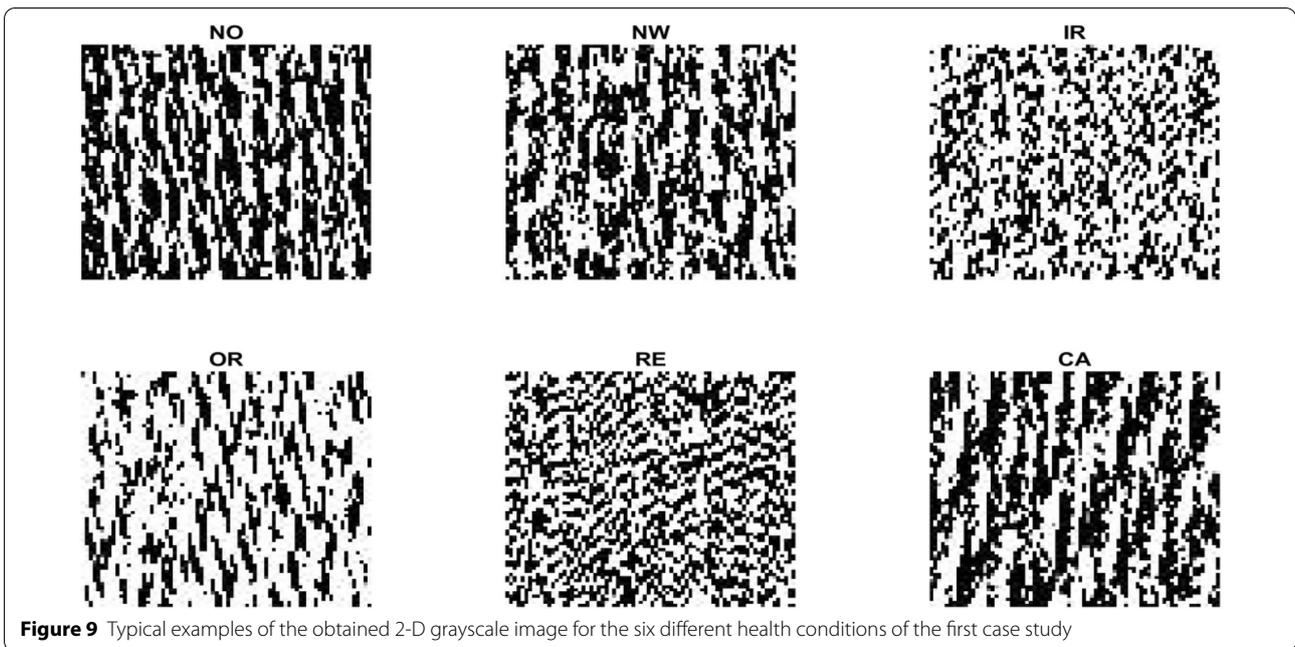
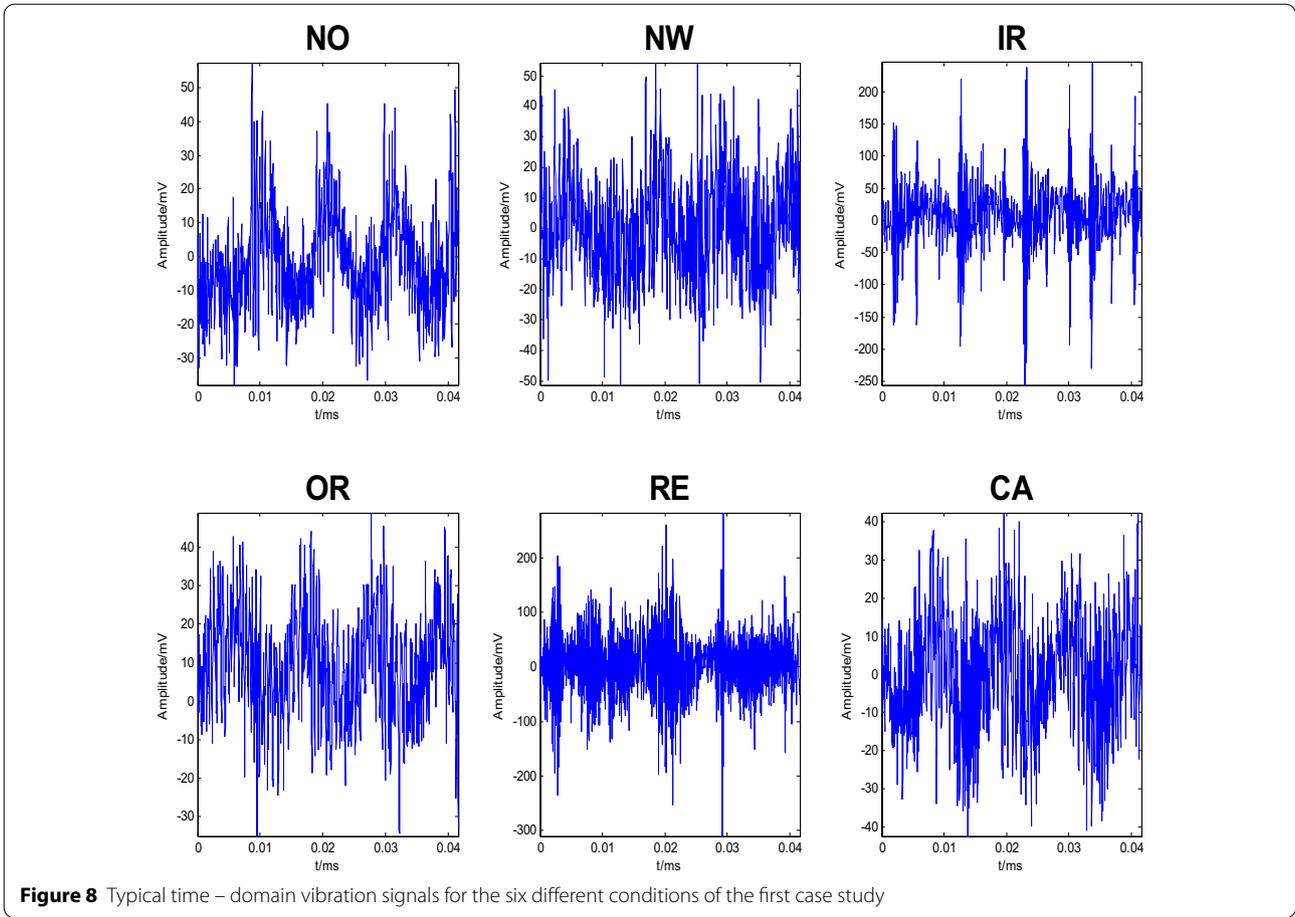
3.1.1 Results

To apply the RGBVI-CNN method in this case study, we began by generating the RGBVIs using the first stage of the RGBVI-CNN described in Section 2.1. First, each raw vibration with 5776 data points is divided into 76 non-overlapping segments of 76 samples and then each segment is aligned as the column of a 76×76 matrix of a 2-D grayscale image. Figure 9 depicts some typical examples of the obtained 2-D grayscale image for the six different health conditions of the first case study. Next, to find the ROI that can offer useful discrimination characteristics, we resized the converted 2-D grayscale image to 128×128 with the intention of adding more pixel information. Then, we employed an average filter with a square kernel of size 9×9 to smooth the image. After, the smoothed image is subtracted from the converted 2-D grayscale image to remove the small positive and negative values of pixels according to the process used in Ref. [82]. Finally, we employed a global image threshold based on Otsu's technique, which automatically chooses an optimal threshold from the viewpoint of discriminant analysis, to convert the resulting 2-D gray image in the previous step to a binary image.

With the obtained binary image, we generated the connected components based RGBVI by first finding the connected components 2-D connectivity of 8 pixels where eight pixels are connected if their edges and corners touch. Here, touching pixels are considered part of the same object if they are both on and are linked along the horizontal, vertical, or diagonal direction. To achieve this, we employed the '*bwconncomp*' function [83]. Based on the connected components we create a label matrix with a unique value for each set of connected components. Finally, we convert the created label matrix into an RGB color image of labeled regions, which offers a set of colour and texture features. Figure 10 shows examples of the generated RGBVIs for the six health conditions of the first case study. It can be clearly seen that the generated

Table 1 Characteristics of bearings health conditions in the first bearing dataset

Condition	Characteristic
NO	The bearing is brand new and in perfect condition
NW	The bearing is in service for some time but in good condition
IR	Inner race fault. This fault is created by cutting a small groove in the raceway of the inner race
OR	Outer race fault. This fault is created by cutting a small groove in the raceway of the outer race
RE	Roller element fault. This fault was created by using electrical etcher to mark the surface of the balls, simulating corrosion
CA	Cage fault. This fault is created by removing the plastic cage from one of the bearings, cutting away a section of the cage so that two of the balls were not held at a regular space and free to move



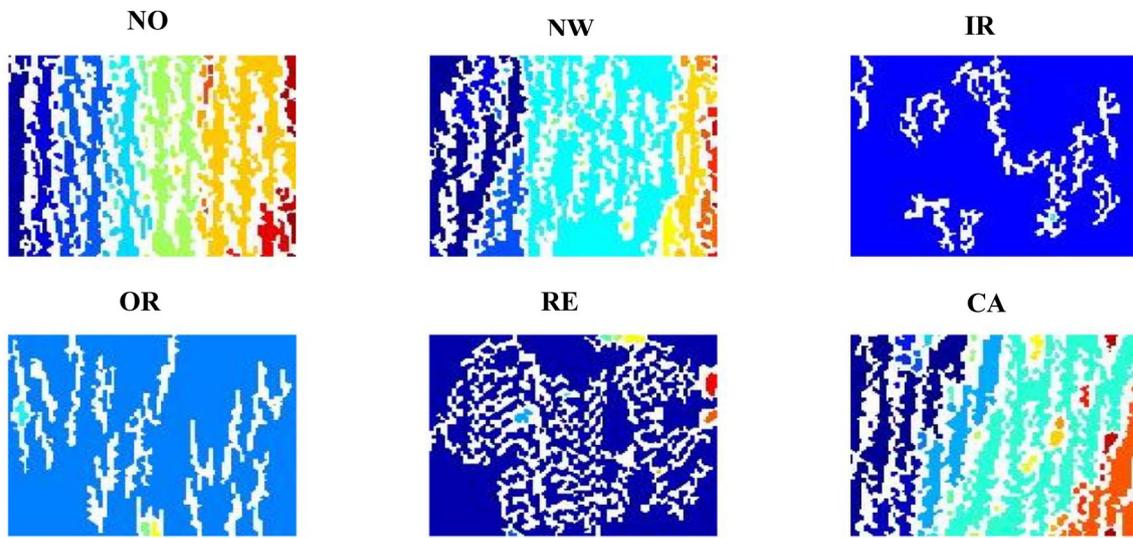


Figure 10 Typical examples of the RGBVI for the six health conditions of the first case study

RGBVIs offers discriminative texture and colour features of the six bearing health conditions that can be recognized visually.

To validate the efficiency of RGBVI-CNN in bearing fault diagnosis, we used the CNN with the generated RGBVIs to classify bearing health conditions of the first case study dataset described above using the CNN architecture described in Figure 5. Experiments were conducted with a training size of 40%, 50%, 60%, and 70% and 10 trials for each experiment. The CNN architecture consists of an image input layer of size $128 \times 128 \times 3$, three blocks of feature learning and each block consists of four layers including the convolutional 2-D layer, Batch Normalisation layer, ReLu layer, and the maximum pooling layer. Based on preliminary observations, we created convolutional layers with 32, 64, and 96 filters in the first, second, and third feature learning blocks respectively. Each filter with a square kernel of 11, i.e., the height and width of each filter, were set to 11, and we set the step size, i.e., the stride, to 2 in the horizontal and vertical directions. To speed up the training process of CNNs we used a batch normalization layer between the convolutional layers and the activation layer, i.e., the ReLU layer. Besides, in each feature learning block, we employed a max-pooling layer of size 2 and stride 2. In the classification stage, we used a fully connected layer with output size 6 (number of classes in the first case study vibration dataset), a SoftMax layer, and a classification layer that computes the cross-entropy loss. For the training procedure of our proposed CNN model, we employed the Adam optimization method [84], which is an efficient computational method with little memory requirements,

Table 2 Overall classification accuracy, precision, recall, F-score results from the first bearing dataset using the RGBVI-CNN method with 40% to 70% training size

Training size (%)	Overall classification accuracy	Overall precision	Overall recall	F-score
40	99.3% \pm 0.7	99.2% \pm 0.8	99.3% \pm 0.6	99.3% \pm 0.7
50	99.8% \pm 0.1	99.8% \pm 0.1	99.8% \pm 0.1	99.8% \pm 0.1
60	99.9% \pm 0.1	99.9% \pm 0.1	99.9% \pm 0.1	99.9% \pm 0.1
70	100% \pm 0.0	100% \pm 0.0	100% \pm 0.0	100% \pm 0.0

with a minimum batch size of 15, an initial learning rate of 0.001, and a maximum number of epochs was set to 20.

Table 2 shows overall testing classification accuracy, precision, recall, F-score results, and their corresponding standard deviations of bearing faults using the first case study vibration dataset. As can be seen from Table 2, the testing results of the four performance measures from the RGBVI-CNN based classification models obtained using training sets of size 40%, 50%, 60%, and 70% are all above 99%. With a training set of 70%, our method achieved a 100% average testing classification accuracy, precision, recall, and F-score. Table 3 shows the sample confusion matrix of the classification results with 30% and 60% testing data. As can be seen from Table 3 that the recognition of all health conditions of bearings is 100% with 30% testing data, i.e., the RGBVI-CNN method misclassified none of the testing examples of all health conditions. But, with 60% the RGBVI-CNN method misclassified 0.2% of NO condition as NW, 0.2% of NW condition as NO

Table 3 Sample confusion matrix of classification results with testing data set of 30% and 60%

Testing set size = 30%							Testing set size = 60%								
NO	48 16.7%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%	NO	95 16.5%	1 0.2%	0 0.0%	0 0.0%	0 0.0%	1 0.2%	97.9% 2.1%
NW	0 0.0%	48 16.7%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%	NW	1 0.2%	95 16.5%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	99.0% 0.1%
IR	0 0.0%	0 0.0%	48 16.7%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%	IR	0 0.0%	0 0.0%	94 16.4%	1 0.2%	0 0.0%	0 0.0%	98.9% 1.1%
OR	0 0.0%	0 0.0%	0 0.0%	48 16.7%	0 0.0%	0 0.0%	100% 0.0%	OR	0 0.0%	0 0.0%	2 0.3%	95 16.5%	0 0.0%	0 0.0%	97.9% 2.1%
RE	0 0.0%	0 0.0%	0 0.0%	0 0.0%	48 16.7%	0 0.0%	100% 0.0%	RE	0 0.0%	0 0.0%	0 0.0%	0 0.0%	96 16.7%	0 0.0%	100% 0.0%
CA	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	48 16.7%	100% 0.0%	CA	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	95 16.5%	100% 0.0%
	100% 0.0%		99.2% 0.8%	99.1% 0.9%	97.9% 2.1%	99.2% 0.8%	100% 0.0%	99.1% 0.9%	99.1% 0.9%						
	NO	NW	IR	OR	RE	CA		NO	NW	IR	OR	RE	CA		

Table 4 Comparison with the results obtained using 2-D grayscale-based CNN

Training size (%)	2-D grayscale vibration images-based CNN classification accuracy	Our proposed RGBVI-CNN classification accuracy
40	86.8% ± 13.1	99.3% ± 0.7
50	88.1% ± 11.6	99.8% ± 0.1
60	91.7% ± 8.2	99.9% ± 0.1
70	92.6% ± 7.4	100% ± 0.0

condition, 0.3% of IR condition as OR condition, 0.2% of OR condition as IR condition, and 0.2% CA condition as NO condition. Overall, these results verify the effectiveness of the RGBVI-CNN method for diagnosing faults in bearings.

3.1.2 Comparison of Results

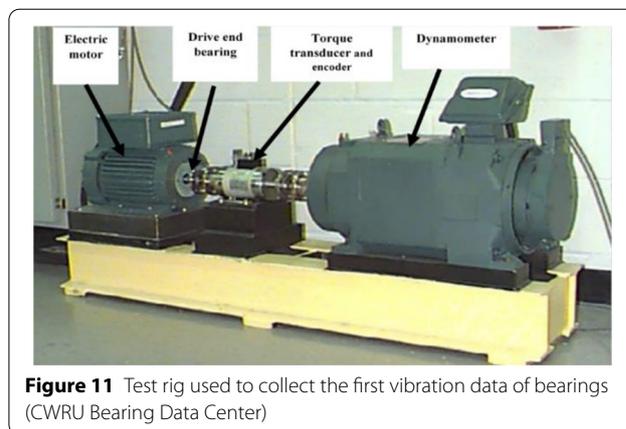
To validate the effectiveness of the RGBVI-CNN method, we conducted two experiments to examine the classification performances in two scenarios. The results are presented in Table 4. The first left column refers to the training sets size used to train the classification models in each scenario. The second column describes the classification results and their corresponding standard deviations obtained using the converted 2-D grayscale vibration images as the inputs to the CNN. The third column presents the classification results and their corresponding standard deviations obtained using the RGBVI images as inputs to the CNN architecture. As Table 4 shows, the classification results from our proposed method, i.e., RGBVI-CNN, are better than those achieved by using the converted 2-D grayscale vibration image-based CNN for all training sets with 40%, 50%, 60%, and 70% of the vibration dataset. Also, it shows that all the classification results from RGBVI-CNN are above 99% for all training sets and it achieved 100% accuracy with 70% training set size. While the highest classification

accuracy achieved using the converted 2-D grayscale vibration images is 92.6% with 70% training set size. Together these results indicate that the RGBVI generated using our proposed method offer greater discriminative characteristics compared to the 2-D grayscale image vibration.

These results may be explained by the fact that resolving a classification problem is enormously helped by appropriate feature representations of the input data. If the features of the input data are carefully devised, and the parameters of the classifiers are carefully tuned, it may be possible to achieve high accuracy in classification performance. In this study, unlike the 2-D grayscale images in Figure 9, the visualized texture and colour features of the RGBVI images in Figure 10 are generated. In the first stage of the proposed method, the colour features of the RGBVI images can visually offer discriminative patterns of bearing health conditions that may sufficiently advance the feature learning process in the second stage of our proposed method. In addition, the 2-D grayscale input image is usually represented as a single matrix, while the RGBVI input image is a 3-D matrix of texture and colour-space features that describe red, green, and blue colour components for each pixel, which offers richer features to the CNN in the second stage of the proposed method. Therefore, the RGBVI images as inputs to the CNN in the second stage of the proposed method are richer and are expected to achieve better classification results than having the 2-D grayscale images as inputs.

To further evaluate the effectiveness of the RGBVI-CNN method, shows comparisons of classification accuracies of the RGBVI-CNN with some recently published results using the same vibration dataset of the first case study. The first left column refers to the reference number in the paper. The second left column presents the method used to obtain the classification results in each reference. The third column describes the testing data size used to validate the classification models in each work while the

fourth column shows the classification results obtained using each method of the compared works. In Ref. [85], classification accuracies from a technique uses entropic features extracted directly from the raw data, and the other two techniques use entropic features extracted from reconstructed signals of compressed measurements. SVM is employed to classify bearing health conditions based on the extracted features. In Ref. [86], with 0.1 sampling rate, compressively sampled measurements (CS), CS followed by principal component analysis (PCA) feature extraction, and CS followed by linear discriminant analysis (LDA) feature extraction. Linear logistic regression (LRC) classifier is used to obtain the classification accuracies results. In Ref. [87], a hybrid model comprising of the Fuzzy Min-Max (FMM) neural network and Random Forest (RF) with Sample Entropy (SampEn) and Power Spectrum (PS) features is employed to classify bearing health conditions. In Ref. [88], a technique based on features extracted using SampEn from the envelope data of the vibration dataset of the first case study is used to classify bearing health condition using SVM and MLP classifiers. In Ref. [89], a Genetic Programming (GP) based approach is studied to extract features from raw vibration data and SVM and ANN are used to classify bearing conditions. As Table 5 shows, the classification results from our proposed method are better than those reported in Refs. [85, 86, 88, 89] using different testing data sizes. Utilising the same vibration dataset of the first case study with 40% testing data size, our proposed method achieved 99.9% classification accuracy, which is



better than the results achieved using CS, CS-PCA, CS-LDA methods in Ref. [86] with 98.6%, 98.5%, and 89.8% respectively. Also, our results are as good as, if not better than results reported in Ref. [87].

3.2 The Second Case Study

The second bearing vibration data is offered by the Case Western Reserve University (CWRU) Bearing Data Center [90]. This data is freely available and widely used in roller bearings fault diagnosis research. Figure 11 displays the test rig used to obtain this vibration data. It is comprised of a 2-horsepower electric motor driving a shaft that encompasses a torque transducer and encoder. A dynamometer and electronic control system are

Table 5 Comparison with the classification results from the literature on the bearing dataset

Reference No.	Method	Testing data size	Classification accuracy (%)
[85]	Raw Vibration data with entropic features.	10-fold cross-validation	98.9 ± 1.2
	Compressed measurements of 0.5 sampling rate followed by signals recovery.		92.4 ± 0.5
	Compressed sampled 0.25 sampling rate followed by signals recovery.		84.6 ± 3.4
[86]	CS	40%	98.6 ± 0.3
	CS-PCA		98.5 ± 0.4
	CS-LDA		89.8 ± 3.5
[87]	FMM-RF (SampEn + PS)	5-fold cross-validation	99.81 ± 0.41
[88]	SVM	10-fold cross-validation	93.5 ± 0.50
	MLP		93.7 ± 0.21
[89]	GP generated feature sets (un-normalised)	50%	
	ANN		96.5
	SVM		97.1
This paper	RGBVI-CNN (Our proposed method)	50%	99.8% ± 0.1
		40%	99.9% ± 0.1
		30%	100% ± 0.0

utilised to apply torque to the shaft. A series of faults with width ranging from 0.18–0.71 mm (0.007–0.028 inch) were planted on the drive end bearing (in this case SKF deep groove ball bearings 6205-2RS JEM were employed) of the electric motor utilising electro-discharger machining. The planted faults include rolling elements, inner race, an outer race faults, and each faulty bearing was run for motor loads 0–3 horsepower at a constant speed in the range of 1720–1797 rev/m. The sampling rates used were 12 kHz for some of the sampled data and 48 kHz for the rest of the sampled data. The bearing vibration signals were acquired under normal NO, IR, OR, and RE conditions for four different speeds. At each speed, 100 time-series were taken for each condition per load.

For IR, OR, and RE conditions, vibration signals for four different fault widths (0.18 mm, 0.36 mm, 0.53 mm, and 0.71 mm) were separately recorded. In this study, from these acquired vibration signals, three datasets were prepared to be used in the assessment of the RGBVI-CNN. The datasets are chosen from the data files of the vibration signals sampled at 48 kHz with fault width (0.18, 0.36, and 0.53), fixed load including 1, 2, and 3 horsepower, and the number of examples chosen are 200 examples per condition. This gave three different datasets A, B, and C with 2000 total number of examples and 2400 data points for each signal. The characteristics of these three datasets are recorded in Table 6.

3.2.1 Results

We applied the same data processing steps of RGBVI-CNN as in the first case study to each dataset, i.e., A, B, and C datasets described above, to obtain the RGBVI (first stage of RGBVI-CNN), which can be used as input to the CNN (second stage of RGBVI-CNN). Figure 12 shows examples of typical time-domain vibration signals for the ten health conditions of dataset A of the second

case study vibration datasets and their corresponding generated RGBVI.

Experiments were conducted with a training size of 30%, 40%, 50%, 60%, 70%, and 80% with 10 trials for each experiment. Figure 13 shows examples of the training progress of the RGBVI-CNN for (a) 50% training data, and (b) 70% training data of dataset A. Figures 14, 15, 16 present the overall classification accuracy, precision, recall, F-score results from datasets A, B, and C of the second case study, respectively, using the RGBVI-CNN method with 30% to 80% training size. As follows in Figures 14, 15, 16, RGBVI-CNN shows high performance in the testing results of the four measures used in this study. In particular, the results for 80% training data achieved 100% average testing classification accuracy, precision, recall, and F-score for datasets A and B as can be seen in Figures 14 and 15, respectively. Results for dataset C, i.e., the dataset with a load of 3 horsepower, achieved 99.9% average testing classification accuracy, precision, recall, and F-score.

Results from the RGBVI-CNN with 30% training data for datasets A, B, and C are less than 98% for average testing classification accuracy, precision, recall, and F-score. Generally, RGBVI-CNN achieved results above 99% for all the performance measures used in this study using classification models trained with training data of size 50% or more of all datasets of the second case study, i.e., datasets A, B, and C.

Taken, together, these results indicate that the RGBVI-CNN method can classify the bearing conditions under different working loads with high accuracy.

3.2.2 Comparison of Results

To further evaluate the effectiveness of the RGBVI-CNN method, Table 7 presents the comparisons with some recently published results with the same bearings vibration datasets A, B, and C. The first left column refers to the reference number in the paper. The second left column presents the method used to obtain the classification results in each reference. The third column describes the testing data size used to validate the classification models in each work while the fourth column shows the classification results obtained using each method of the compared works. In Ref. [34], classification results of deep neural network (DNN) and backpropagation neural network (BPNN) were reported. Also, classification results from the same datasets using a generic multi-layer perceptron (MLP) are reported in Ref. [91]. Moreover, classification results using a technique combining compressive sampling and DNN (CS-DNN) are reported in Ref. [36]. Furthermore, STFT was employed to convert

Table 6 Description of the bearing health conditions for A, B, and C datasets

Fault type	Fault width (mm)	Classification label
NO	0	1
IR1	0.18	2
IR2	0.36	3
IR3	0.53	4
RE1	0.18	5
RE2	0.36	6
RE3	0.53	7
OR1	0.18	8
OR2	0.36	9
OR3	0.53	10

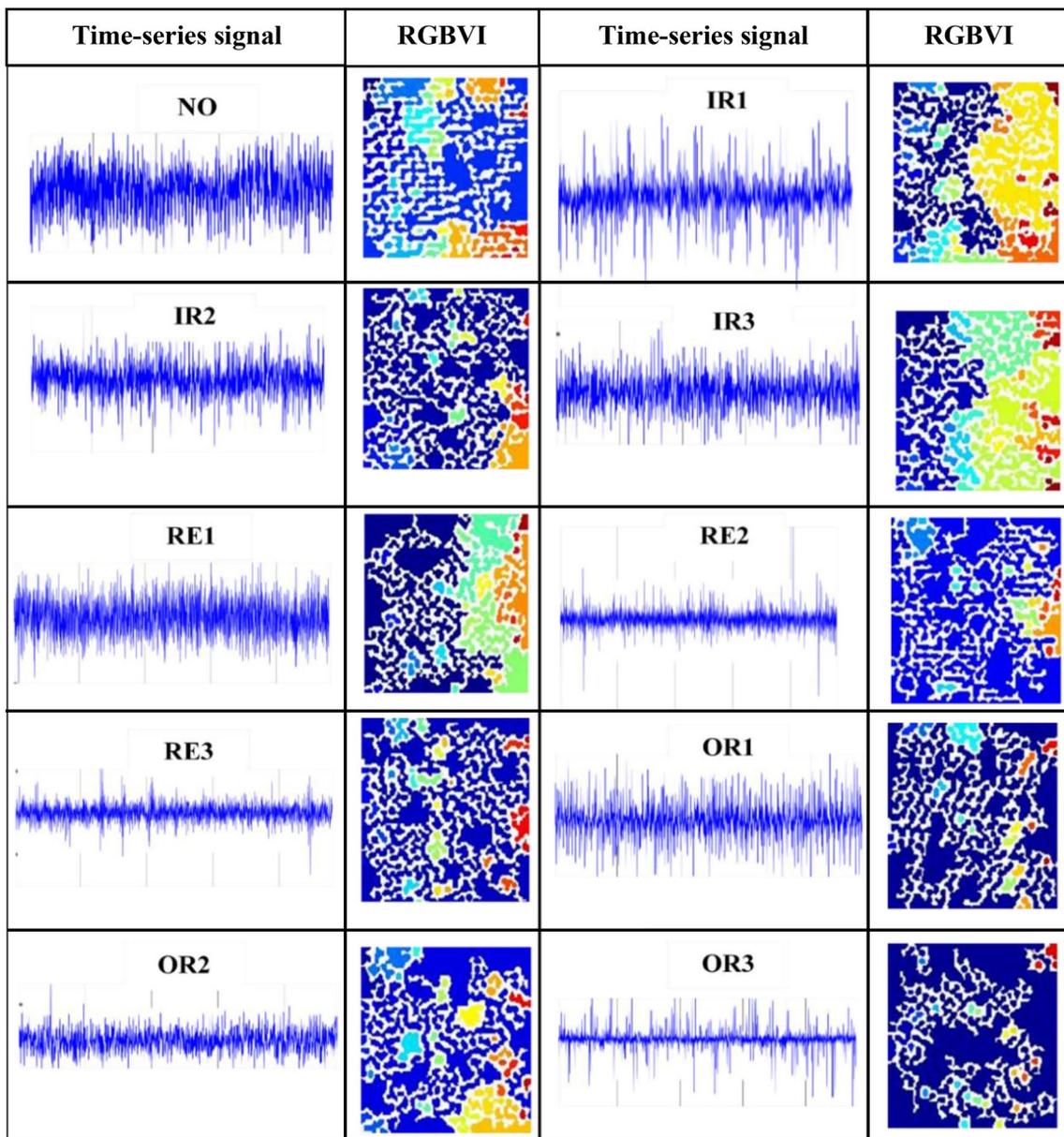


Figure 12 Typical examples of time – domain vibration signals for the ten health conditions of dataset A of the second case study vibration datasets and their corresponding generated RGBVI

1-D vibration vectors of datasets A, B, and C to 2-D time-frequency graphs (TFGs) then a CNN based on a capsule network with an inception block (ICN) was trained on one load of A, B, and C datasets and tested on the other two loads. Also, ICN is compared to the other two CNN based network, namely, AlexNet and ResNet using the TFGs and the results are reported in Ref. [67] where $A \rightarrow B$ means the methods were trained on dataset A and tested on dataset B.

As Table 7 shows, using the same vibration datasets of the second case study, the classification results obtained using the RGBVI-CNN are better than the results reported in Refs. [67, 91] as well as the results of BPNN reported in Ref. [35]. Moreover, using 30%, 40%, and 50% testing data sizes, the results from the RGBVI-CNN remain very competitive if not better than the results reported in [36] and [34] with 50% data size.

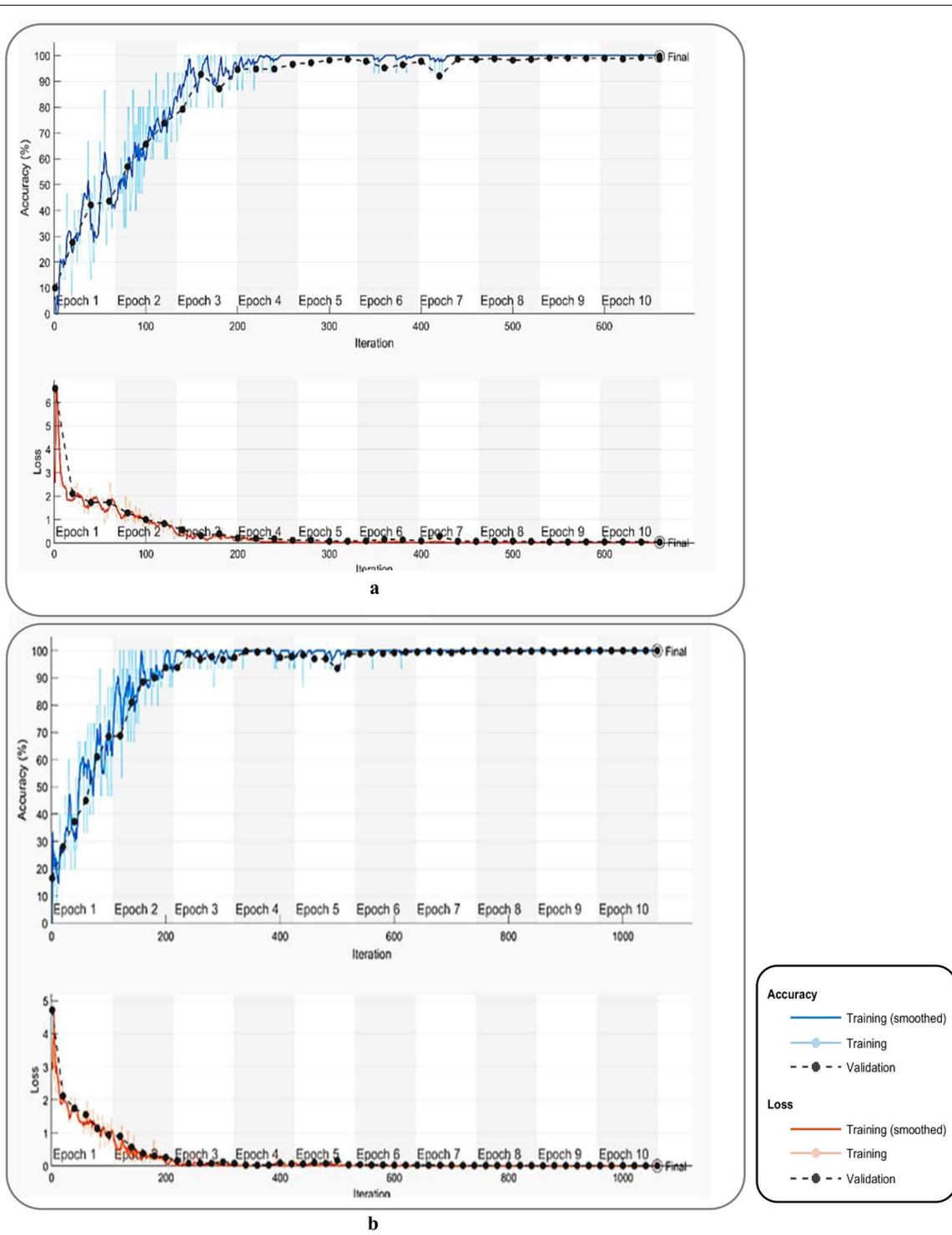
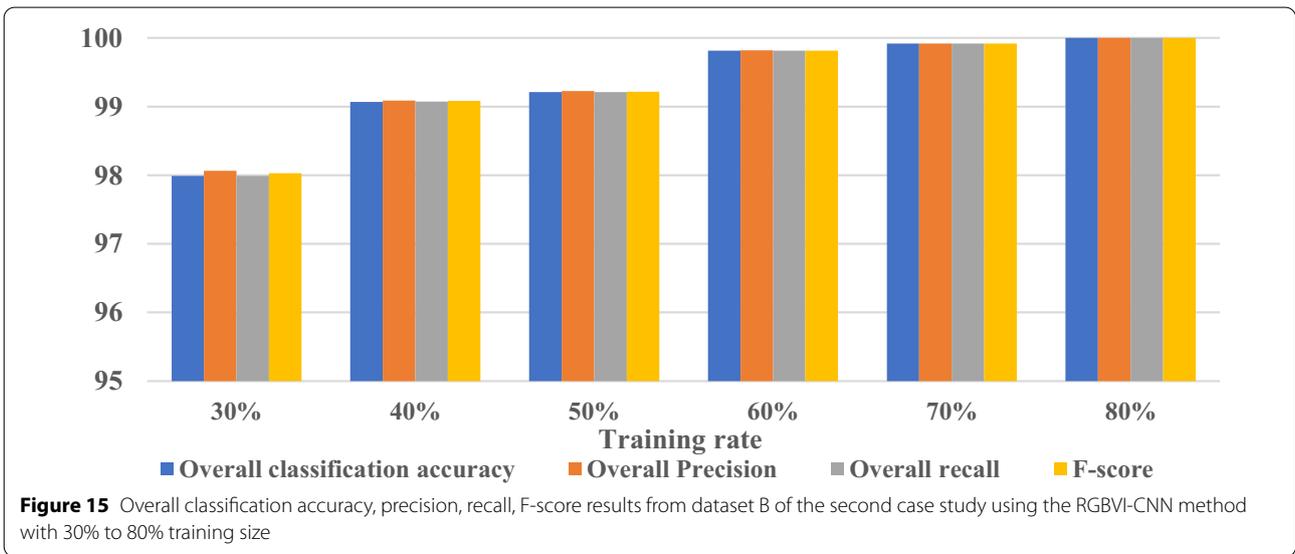
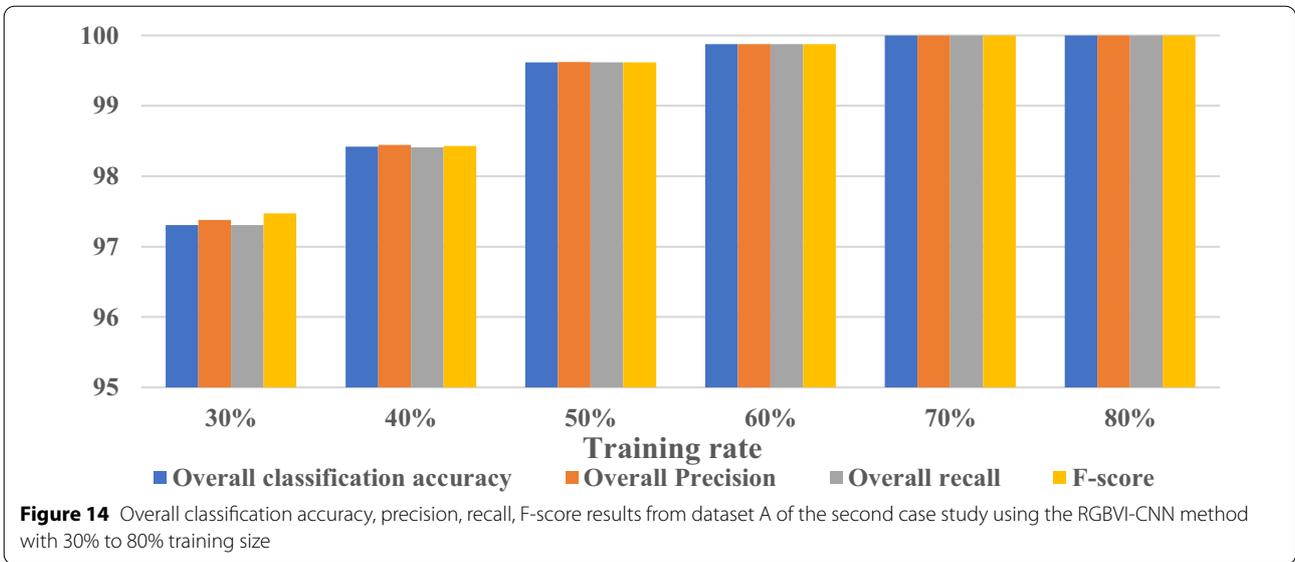


Figure 13 Examples of the training progress of the RGBVI-CNN for **a** 50% training data, and **b** 70% training data



4 Conclusions

The paper developed and evaluated a two-stage method RGBVI-CNN for bearing fault diagnosis. In the RGBVI-CNN, the first stage was employed to generate RGBVIs based on connected components of the ROI of bearing vibration images that demonstrate useful features of bearing vibration signals. These RGBVIs were input to the proposed CNN architecture in the second stage of RGBVI-CNN to train, validate, and classify bearing health conditions. RGBVI-CNN was evaluated experimentally in two case studies of bearing vibration datasets. The results of this investigation showed that RGBVI-CNN can generate advantageous

health condition features from bearing vibration signals and classify the health conditions under different working loads with high accuracy. The classification models trained using RGBVI-CNN showed high performance in the testing results of the overall classification accuracy, precision, recall, and F-score. Besides, a comparison with recently published results indicates that RGBVI-CNN offers better results. Taken together, these findings suggest a significant role for RGBVI-CNN in developing intelligent methods for bearing fault diagnosis, to help improve safety, reliability, and performance.

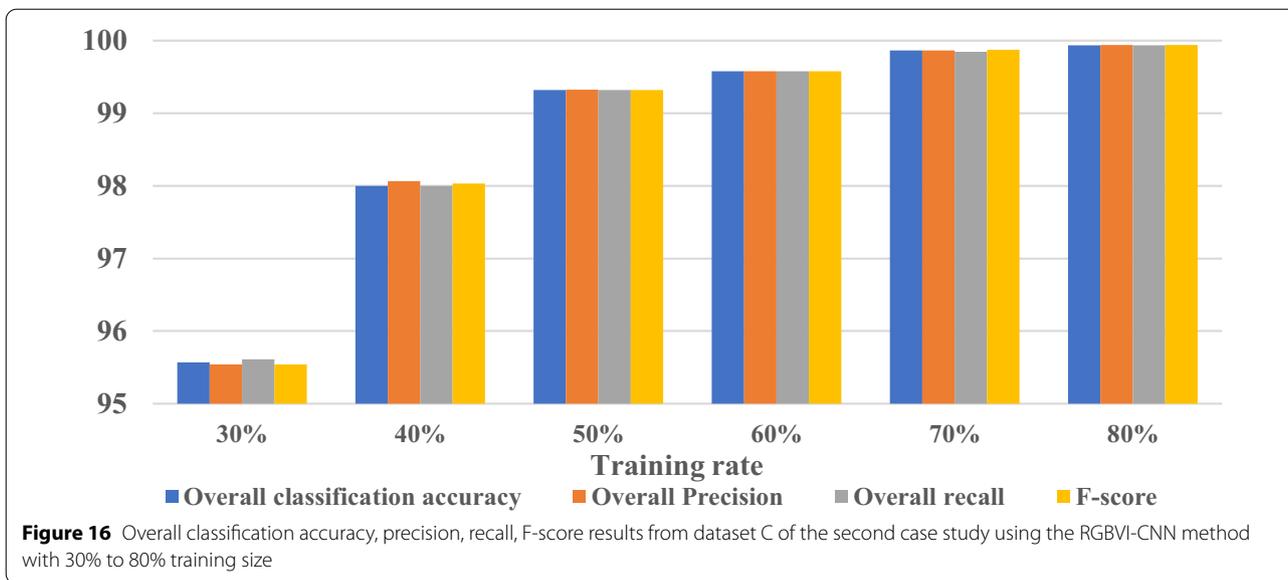


Table 7 A comparison with the classification results from the literature on bearing datasets A, B, and C datasets

Reference No.	Method	Dataset	Testing data size (%)	Classification accuracy (%)
[34]	DNN	A	50	99.95 ± 0.06
		B		99.61 ± 0.21
		C		99.74 ± 0.16
[34]	BPNN	A	50	65.2 ± 18.09
		B		61.95 ± 22.09
		C		69.82 ± 17.67
[36]	CS-DNN with CS sampling rate of 0.1	A	50	99.3 ± 0.6
		B		99.7 ± 0.5
		C		100 ± 0.0
[91]	MLP	A	50	95.7
		B		99.6
		C		99.4
[67]	AlexNet	A → B; A → C; B → A;	-	94.3
	ResNet	B → C; C → A; C → B		94.6
	ICN			97.2
This paper	RGBVI-CNN *the proposed method	A	50	99.6 ± 0.4
		B		99.2 ± 0.7
		C		99.3 ± 0.7
		A	40	99.9 ± 0.1
		B		99.8 ± 0.1
		C		99.6 ± 0.3
		A	30	100 ± 0.0
		B		99.9 ± 0.1
		C		99.9 ± 0.1

In the practice of machine condition monitoring, one would like to use the model developed for a fault classification task of rolling bearings in one machine to work in another machine. Such transfer learning is beyond the

scope of this study. However, we believe that the generated connected components-based colour representations of vibrations using our proposed method can be

used for transfer learning and we will consider this in our future research.

Acknowledgements

Not applicable.

Authors' contributions

HOAA and AKN designed the investigations. HOAA carried out the experiments. HOAA and AKN discussed and analysed the results. HOAA wrote the first draft and both contributed to the final version. All authors read and approved the final manuscript.

Authors' information

Hosameldin O A Ahmed, Ph.D., received the degree of Ph.D. in Electronic and Computer Engineering from *Brunel University London, UK*. His research interests lie in the areas of signal processing, pattern recognition, image processing, feature extraction and generation, predictive modeling, and machine learning with applications to various systems. He has a proven track record of developing and applying condition monitoring systems for rotating machines. He also has developed and applied several techniques in the areas of Human-Computer-interaction, image classification, digital image inpainting to a variety of applications. His work has been published in high-quality journals and international conferences. He published several book chapters and journal articles. He has also co-authored a research monograph entitled *Condition Monitoring with Vibration Signals* (Wiley, 2020).

Asoke K Nandi received the degree of Ph.D. in Physics from the *University of Cambridge (Trinity College), Cambridge, UK*. He held academic positions in several universities, including Oxford, Imperial College London, Strathclyde, and Liverpool as well as Finland Distinguished Professorship in Jyväskylä (Finland). In 2013 he moved to Brunel University London, to become the Chair and Head of Electronic and Computer Engineering. Professor Nandi is a Distinguished Visiting Professor at *Xi'an Jiaotong University (China)* and an Adjunct Professor at the *University of Calgary (Canada)*.

In 1983 Professor Nandi jointly discovered the three fundamental particles known as W^+ , W^- and Z^0 , providing the evidence for the unification of the electromagnetic and weak forces, for which the Nobel Committee for Physics in 1984 awarded the prize to his two team leaders for their decisive contributions. His current research interests lie in the areas of signal processing and machine learning, with applications to communications, gene expression data, functional magnetic resonance data, machine condition monitoring, and biomedical data. He has made many fundamental theoretical and algorithmic contributions to many aspects of signal processing and machine learning. He has much expertise in "Big Data", dealing with heterogeneous data, and extracting information from multiple datasets obtained in different laboratories and different times. Professor Nandi has authored over 600 technical publications, including 250 journal papers as well as five books, entitled *Condition Monitoring with Vibration Signals* (Wiley, 2020), *Automatic Modulation Classification: Principles, Algorithms and Applications* (Wiley, 2015), *Integrative Cluster Analysis in Bioinformatics* (Wiley, 2015), *Blind Estimation Using Higher-Order Statistics* (Springer, 1999), and *Automatic Modulation Recognition of Communications Signals* (Springer, 1996). The H-index of his publications is 80 (Google Scholar) and ERDOS number is 2.

Professor Nandi is a Fellow of the *Royal Academy of Engineering (UK)* and of seven other institutions. Among the many awards he received are the Institute of Electrical and Electronics Engineers (USA) Heinrich Hertz Award in 2012, the Glory of Bengal Award for his outstanding achievements in scientific research in 2010, from the Society for Machinery Failure Prevention Technology, a division of the Vibration Institute (USA) in 2000, the Water Arbitration Prize of the Institution of Mechanical Engineers (UK) in 1999, and the Mountbatten Premium, Division Award of the Electronics and Communications Division, of the Institution of Electrical Engineers (UK) in 1998. Professor Nandi was an IEEE EMBS Distinguished Lecturer (2018-19).

Funding

Supported by the 111 Project (B20038) of China.

Availability of data and materials

The second bearing vibration dataset is publicly available from Case Western Reserve University (CWRU) Bearing Data Center (<http://csegroup.case.edu/bearingdatacenter/home>)

Competing interests

The authors declare there are no competing interests.

Author Details

¹ Department of Electronic and Electrical Engineering, Brunel University London, Uxbridge UB8 3PH, UK. ² School of Mechanical Engineering, Xi'an Jiaotong University, Xi'an, China.

Received: 29 August 2020 Revised: 11 December 2020 Accepted: 11 March 2021

Published online: 12 April 2021

References

- [1] Asoke K Nandi, Ahmed Hosameldin. *Condition monitoring with vibration signals: Compressive sampling and learning algorithms for rotating machines*. John Wiley & Sons, 2020.
- [2] A C McCormick, A K Nandi. A comparison of artificial neural networks and other statistical methods for rotating machine condition classification. In IEE Colloquium on Modeling and Signal Processing for Fault Diagnosis, 1996, Digest No: 1996/260.
- [3] A C McCormick, A K Nandi. Classification of the rotating machine condition using artificial neural networks. *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, 1997, 211(6): 439-450.
- [4] L B Jack, A K Nandi. Genetic algorithms for feature selection in machine condition monitoring with vibration signals. *IEE Proceedings-Vision, Image and Signal Processing*, 2000, 147(3): 205-212.
- [5] A Rojas, A K Nandi. Practical scheme for fast detection and classification of rolling-element bearing faults using support vector machines. *Mechanical Systems and Signal Processing*, 2006, 20(7): 1523-1536.
- [6] V Sugumaran, K I Ramachandran. Automatic rule learning using decision tree for fuzzy classifier in fault diagnosis of roller bearing. *Mechanical Systems and Signal Processing*, 2007, 21(5): 2237-2247.
- [7] M D Prieto, G Cirrincione, A G Espinosa, et al. Bearing fault detection by a novel condition-monitoring scheme based on statistical-time features and neural networks. *IEEE Transactions on Industrial Electronics*, 2013, 60(8): 3398-3407.
- [8] K Christian, N Mureithi, A Lakis, et al. On the use of time synchronous averaging, independent component analysis and support vector machines for bearing fault diagnosis. *Proceedings of the 1st International Conference on Industrial Risk Engineering*, 2007: 610-624.
- [9] A K Nandi, Ed. *Blind estimation using Higher-Order Statistics*. Kluwer Academic Publishers, Amsterdam, Netherlands, and Springer, 1999.
- [10] L Zhang, L B Jack, A K Nandi. Fault detection using genetic programming. *Mechanical Systems and Signal Processing*, 2005, 19(2): 271-289.
- [11] L Zhang, A K Nandi. Fault classification using genetic programming. *Mechanical Systems and Signal Processing*, 2007, 21(3): 1273-1284.
- [12] J Yang, Y Zhang, Y Zhu. Intelligent fault diagnosis of rolling element bearing based on SVMs and fractal dimension. *Mechanical Systems and Signal Processing*, 2007, 21(5): 2012-2024.
- [13] Z Zhang, Y Wang, K Wang. Fault diagnosis and prognosis using wavelet packet decomposition, Fourier transform and artificial neural network. *Journal of Intelligent Manufacturing*, 2013, 24(6): 1213-1227.
- [14] H Ahmed, A K Nandi. Compressive sampling and feature ranking framework for bearing fault classification with vibration signals. *IEEE Access*, 2018, 6: 44731-44746.
- [15] G Betta, C Liguori, A Paolillo, et al. A DSP-based FFT-analyzer for the fault diagnosis of rotating machine based on vibration analysis. *IEEE Transactions on Instrumentation and Measurement*, 2002, 51(6): 1316-1322.
- [16] H O A Ahmed, A K Nandi. Three-stage hybrid fault diagnosis for rolling bearings with compressively sampled data and subspace learning techniques. *IEEE Transactions on Industrial Electronics*, 2019, 66(7): 5516-5524.
- [17] H O A Ahmed, A K Nandi. Multiple measurement vector compressive sampling and fisher score feature selection for fault classification of roller

- bearings. 2017 *22nd International Conference on Digital Signal Processing (DSP)*, London, 2017: 1-5.
- [18] O R Seryasat, M Aliyari shoorehdeli, F Honarvar, et al. Multi-fault diagnosis of ball bearing using FFT, wavelet energy entropy mean and root mean square (RMS). *IEEE International Conference on Systems, Man and Cybernetics*, Istanbul, 2010: 4295-4299.
- [19] A Ghods, H Lee. A frequency-based approach to detect bearing faults in induction motors using discrete wavelet transform. *IEEE International Conference on Industrial Technology (ICIT)*, Busan, 2014: 121-125.
- [20] S Patidar, P K Soni. An overview on vibration analysis techniques for the diagnosis of rolling element bearing faults. *International Journal of Engineering Trends and Technology (IJETT)*, 2013, 4(5): 1804-1809.
- [21] A C McCormick, A K Nandi. Bispectral and trispectral features for machine condition diagnosis. *IEE Proceedings-Vision, Image and Signal Processing*, 1999, 146(5): 229-234.
- [22] J Tian, C Morillo, M H Azarian, et al. Motor bearing fault detection using spectral kurtosis-based feature extraction coupled with K-nearest neighbor distance analysis. *IEEE Transactions on Industrial Electronics*, 2016, 63(3): 1793-1803.
- [23] M He, D He. Deep learning based approach for bearing fault diagnosis. *IEEE Transactions on Industry Applications*, 2017, 53(3): 3057-3065.
- [24] H Tao, P Wang, Y Chen, et al. An unsupervised fault diagnosis method for rolling bearing using STFT and generative neural networks. *Journal of the Franklin Institute*, 2020: 7286-7307.
- [25] X Yu, F Dong, E Ding, et al. Rolling bearing fault diagnosis using modified LFDA and EMD with sensitive feature selection. *IEEE Access*, 2017, 6: 3715-3730.
- [26] P K Kankar, S C Sharma, S P Harsha. Fault diagnosis of ball bearings using continuous wavelet transform. *Applied Soft Computing*, 2011, 11(2): 2300-2312.
- [27] R Kumar, M Singh. Outer race defect width measurement in taper roller bearing using discrete wavelet transform of vibration signal. *Measurement*, 2013, 46(1): 537-545.
- [28] F Al-Badour, M Sunar, L Cheded. Vibration analysis of rotating machinery using time-frequency analysis and wavelet techniques. *Mechanical Systems and Signal Processing*, 2011, 25(6): 2083-2101.
- [29] Y Lei, J Lin, Z He, et al. A review on empirical mode decomposition in fault diagnosis of rotating machinery. *Mechanical Systems and Signal Processing*, 2013, 35(1-2): 108-126.
- [30] A Soualhi, K Medjaher, N Zerhouni. Bearing health monitoring based on Hilbert-Huang transform, support vector machine, and regression. *IEEE Transactions on Instrumentation and Measurement*, 2015, 64(1): 52-62.
- [31] H Li, Y Zhang. Bearing faults diagnosis based on EMD and Wigner-Ville distribution. *Intelligent Control and Automation, 2006, WCICA 2006, The Sixth World Congress, IEEE*, 2006, 2: 5447-5451.
- [32] J Yu, J Lv. Weak fault feature extraction of rolling bearings using local mean decomposition-based multilayer hybrid denoising. *IEEE Transactions on Instrumentation and Measurement*, 2017, 66(12): 3148-3159.
- [33] Y Zhang, R B Randall. Rolling element bearing fault diagnosis based on the combination of genetic algorithms and fast kurtogram. *Mechanical Systems and Signal Processing*, 2009, 23(5): 1509-1517.
- [34] F Jia, Y Lei, J Lin, et al. Deep neural networks: A promising tool for fault characteristic mining and intelligent diagnosis of rotating machinery with massive data. *Mechanical Systems and Signal Processing*, 2016, 72: 303-315.
- [35] F Zhou, Y Gao, C Wen. A novel multimode fault classification method based on deep learning. *Journal of Control Science and Engineering*, 2017: 1-14.
- [36] H O A Ahmed, M L D Wong, A K Nandi. Intelligent condition monitoring method for bearing faults from highly compressed measurements using sparse over-complete features. *Mechanical Systems and Signal Processing*, 2018, 99: 459-477.
- [37] C Li, W Zhang, G Peng, et al. Bearing fault diagnosis using fully-connected winner-take-all autoencoder. *IEEE Access*, 2018, 6: 6103-6115.
- [38] Z Chen, W Li. Multisensor feature fusion for bearing fault diagnosis using sparse autoencoder and deep belief network. *IEEE Transactions on Instrumentation and Measurement*, 2017, 66(7): 1693-1702.
- [39] W Mao, J He, Y Li, et al. Bearing fault diagnosis with auto-encoder extreme learning machine: A comparative study. *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, 2017, 231(8): 1560-1578.
- [40] Z Chen, S Deng, X Chen, et al. Deep neural networks-based rolling bearing fault diagnosis. *Microelectronics Reliability*, 2017, 75: 327-333.
- [41] Z Chen, X Zeng, W Li, et al. Machine fault classification using deep belief network. *Instrumentation and Measurement Technology Conference Proceedings (I2MTC), IEEE International*, 2016: 1-6.
- [42] J Tao, Y Liu, D Yang. Bearing fault diagnosis based on deep belief network and multisensor information fusion. *Shock and Vibration*, 2016: 1-9.
- [43] H Shao, H Jiang, X Li, et al. Rolling bearing fault detection using continuous deep belief network with locally linear embedding. *Computers in Industry*, 2018, 96: 27-39.
- [44] J Guo, P Zheng. A method of rolling bearing fault diagnose based on double sparse dictionary and deep belief network. *IEEE Access*, 2020, 8: 116239-116253.
- [45] C Zhang, Y Zhang, C Hu, et al. A novel intelligent fault diagnosis method based on variational mode decomposition and ensemble deep belief network. *IEEE Access*, 2020, 8: 36293-36312.
- [46] K Lee, J K Kim, J Kim, et al. Stacked convolutional bidirectional LSTM recurrent neural network for bearing anomaly detection in rotating machinery diagnostics. *1st IEEE International Conference on Knowledge Innovation and Invention (ICKII)*, 2018: 98-101.
- [47] H Liu, J Zhou, Y Zheng, et al. Fault diagnosis of rolling bearings with recurrent neural network-based autoencoders. *ISA Transactions*, 2018, 77: 167-178.
- [48] B Zhang, S Zhang, W Li. Bearing performance degradation assessment using long short-term memory recurrent network. *Computers in Industry*, 2019, 106: 14-29.
- [49] M Qiao, S Yan, X Tang, et al. Deep convolutional and LSTM recurrent neural networks for rolling bearing fault diagnosis under strong noises and variable loads. *IEEE Access*, 2020, 8: 66257-66269.
- [50] F Wang, G Deng, L Ma, et al. Convolutional neural network based on spiral arrangement of features and its application in bearing fault diagnosis. *IEEE Access*, 2019: 64092-64100.
- [51] O Janssens, V Slavkovikj, B Vervisch, et al. Convolutional neural network based fault detection for rotating machinery. *Journal of Sound and Vibration*, 2016, 377: 331-345.
- [52] X Guo, L Chen, C Shen. Hierarchical adaptive deep convolution neural network and its application to bearing fault diagnosis. *Measurement*, 2016, 93: 490-502.
- [53] X Ding, Q He. Energy-fluctuated multiscale feature learning with deep convnet for intelligent spindle bearing fault diagnosis. *IEEE Transactions on Instrumentation and Measurement*, 2017, 66(8): 1926-1935.
- [54] L Wen, X Li, L Gao, et al. A new convolutional neural network-based data-driven fault diagnosis method. *IEEE Transactions on Industrial Electronics*, 2018, 65(7): 5990-5998.
- [55] U P Chong. Signal model-based fault detection and diagnosis for induction motors using features of vibration signal in two-dimension domain. *Strojinski vestnik*, 2011, 57(9): 655-666.
- [56] M Amar, I Gondal, C Wilson. Vibration spectrum imaging: A novel bearing fault classification approach. *IEEE Transactions on Industrial Electronics*, 2014, 62(1): 494-502.
- [57] M Kang, J M Kim. Reliable fault diagnosis of multiple induction motor defects using a 2-D representation of Shannon wavelets. *IEEE Transactions on Magnetics*, 2014, 50(10): 1-13.
- [58] W Li, M Qiu, Z Zhu, et al. Bearing fault diagnosis based on spectrum images of vibration signals. *Measurement Science and Technology*, 2016, 27(3): 035005.
- [59] C Lu, Y Wang, M Ragulskis, et al. Fault diagnosis for rotating machinery: A method based on image processing. *PLoS One*, 2016, 11(10): e0164111.
- [60] D Verstraete, A Ferrada, E L Droggett, et al. Deep learning enabled fault diagnosis using time-frequency image analysis of rolling element bearings. *Shock and Vibration*, 2017: 1-17.
- [61] H Oh, J H Jung, B C Jeon, et al. Scalable and unsupervised feature engineering using vibration-imaging and deep learning for rotor system diagnosis. *IEEE Transactions on Industrial Electronics*, 2017, 65(4): 3539-3549.
- [62] W Zhang, G Peng, C Li. Bearings fault diagnosis based on convolutional neural networks with 2-D representation of vibration signals as input. *MATEC Web of Conferences, EDP Sciences*, 2017, 95: 13001.
- [63] M J Hasan, J M Kim. Bearing fault diagnosis under variable rotational speeds using stockwell transform-based vibration imaging and transfer learning. *Applied Sciences*, 2018, 8(12): 2357.

- [64] H Wang, S Li, L Song, et al. A novel convolutional neural network based fault recognition method via image fusion of multi-vibration-signals. *Computers in Industry*, 2019, 105: 182-190.
- [65] X Yan, Z Sun, J Zhao, et al. Fault diagnosis of active magnetic bearing-rotor system via vibration images. *Sensors*, 2019, 19(2): 244.
- [66] D T Hoang, H J Kang. Rolling element bearing fault diagnosis using convolutional neural network and vibration image. *Cognitive Systems Research*, 2019, 53: 42-50.
- [67] Z Zhu, G Peng, Y Chen, et al. A convolutional neural network based on a capsule network with strong generalization for bearing fault diagnosis. *Neurocomputing*, 2019, 323: 62-75.
- [68] P Ma, H Zhang, W Fan, et al. 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.
- [69] X Zhu, D Hou, P Zhou, et al. Rotor fault diagnosis using a convolutional neural network with symmetrized dot pattern images. *Measurement*, 2019, 138: 526-535.
- [70] Y Zhang, K Xing, R Bai, et al. An enhanced convolutional neural network for bearing fault diagnosis based on time-frequency image. *Measurement*, 2020: 107667.
- [71] K Kaplan, Y Kaya, M Kuncan, et al. An improved feature extraction method using texture analysis with LBP for bearing fault diagnosis. *Applied Soft Computing*, 2020, 87: 106019.
- [72] H P Singh, A Nigam, A K Gautam, et al. Noise reduction in images using enhanced average filter. *International Journal of Computer Applications*, 2014: 25-28.
- [73] M Piccardi. Background subtraction techniques: a review. *2004 IEEE International Conference on Systems, Man and Cybernetics (IEEE Cat. No. 04CH37583)*, 2004, 4: 3099-3104.
- [74] N Otsu. A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, 1979, 9(1): 62-66.
- [75] L Di Stefano, A Bulgarelli. A simple and efficient connected components labeling algorithm. *Proceedings 10th International Conference on Image Analysis and Processing*, 1999: 322-327.
- [76] Y LeCun, B E Boser, J S Denker, et al. Handwritten digit recognition with a back-propagation network. *Advances in Neural Information Processing Systems*, 1990: 396-404.
- [77] Y LeCun, L Bottou, Y Bengio, et al. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 1998, 86(11): 2278-2324.
- [78] N Tajbakhsh, J Y Shin, S R. Gurudu, et al. Convolutional neural networks for medical image analysis: Full training or fine tuning. *IEEE Transactions on Medical Imaging*, 2016, 35(5): 1299-1312.
- [79] D Maturana, S Scherer. Voxnet: A 3D convolutional neural network for real-time object recognition. *Intelligent Robots and Systems (IROS), 2015 IEEE/RSJ International Conference on*, 2015: 922-928.
- [80] O Abdel-Hamid, A R Mohamed, H Jiang, et al. Convolutional neural networks for speech recognition. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 2014, 22(10): 1533-1545.
- [81] D Simard, P Y Steinkraus, J C Platt. Best practices for convolutional neural networks. *Proc. Seventh Int'l Conf. Document Analysis and Recognition*, 2003: 958.
- [82] P Bankhead. Analyzing fluorescence microscopy images with ImageJ. *Image*, 2014, 1(195): 10-1109.
- [83] Mathworks.com. (2020). Find connected components in binary image —MATLAB *bwconncomp*. [Online]. Available: <https://uk.mathworks.com/help/images/ref/bwconncomp.html>.
- [84] D P Kingma, J Ba. Adam: A method for stochastic optimization. arXiv preprint. arXiv:1412.6980, 2014.
- [85] M L D Wong, M Zhang, A K Nandi. Effects of compressed sensing on classification of bearing faults with entropic features. *23rd European Signal Processing Conference (EUSIPCO)*, Nice, 2015: 2256-2260.
- [86] H O A Ahmed, M L D Wong, A K Nandi. Compressive sensing strategy for classification of bearing faults. *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, New Orleans, LA, 2017: 2182-2186.
- [87] M Seera, M L D Wong, A K Nandi. Classification of ball bearing faults using a hybrid intelligent model. *Applied Soft Computing*, 2017, 57: 427-435.
- [88] M D Wong, C Liu, A K Nandi. Classification of ball bearing faults using entropic measures. *Proceeding Surveillance*, 2014, 7.
- [89] H Guo, L B Jack, A K Nandi. Feature generation using genetic programming with application to fault classification. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, 2005, 35(1): 89-99.
- [90] Case Western Reserve University Bearing Data Center (<http://csegroup.case.edu/bearingdatacenter/home>).
- [91] L F de Almeida, J W Bizarria, F C Bizarria, et al. Condition-based monitoring system for rolling element bearing using a generic multi-layer perceptron. *Journal of Vibration and Control*, 2015, 21(16): 3456-3464.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)
