Research Article

Fault Diagnosis for Rotating Machinery Based on Convolutional Neural Network and Empirical Mode Decomposition

Yuan Xie and Tao Zhang

Department of Automation, School of Information Science and Technology, Tsinghua University, Beijing, China

Correspondence should be addressed to Yuan Xie; xiey07@sina.com

The analysis of vibration signals has been a very important technique for fault diagnosis and health management of rotating machinery. Classic fault diagnosis methods are mainly based on traditional signal features such as mean value, standard derivation, and kurtosis. Signals still contain abundant information which we did not fully take advantage of. In this paper, a new approach is proposed for rotating machinery fault diagnosis with feature extraction algorithm based on empirical mode decomposition (EMD) and convolutional neural network (CNN) techniques. The fundamental purpose of our newly proposed approach is to extract distinguishing features. Frequency spectrum of the signal obtained through fast Fourier transform process is trained in a designed CNN structure to extract compressed features with spatial information. To solve the nonstationary characteristic, we also apply EMD technique to the original vibration signals. EMD energy entropy is calculated using the first few intrinsic mode functions (IMFs) which contain more energy. With features extracted from both methods combined, classification models are trained for diagnosis. We carried out experiments with vibration data of 52 different categories under different machine conditions to test the validity of the approach, and the results indicate it is more accurate and reliable than previous approaches.

1. Introduction

Rolling-element bearings (REBs) are the most fundamental and important components of rotating machines in industrial manufacture and agricultural production. Therefore, the analysis of REB vibration signals is always considered an important approach in fault diagnosis and condition monitoring. A minor defection of rolling bearings may lead to breakdown of the entire system and cause severe financial losses.

Vibration signals are usually generated from rolling-element bearings, which contain rich information that may assist in the procedure of condition monitoring, fault diagnosis, and machine health management. The research of bearing fault diagnosis has long been receiving extensive attention over years and is becoming more important in modern industry for the need of higher reliability and lower loss possibility.

Essentially, fault diagnosis is a pattern recognition problem, which includes two major steps that are feature extraction and classification. Traditional features of vibration signals are generated from three main kinds of methods as listed below. Time domain analysis and frequency domain analysis are mostly commonly used in feature extraction; also the combination known as time-frequency domain analysis is another significant method.

Time domain features have long been used in the aspect of fault diagnosis for rotating machinery [1]. Most time domain features are statistical features such as mean value, root mean squares, standard deviation, kurtosis, and skewness. They are generally easy to calculate and acquire and therefore are trained in different classifier models for fault diagnosis. Hu et al. [2] and Sreejith et al. [3] combined time domain features with artificial intelligence, namely, artificial neural network (ANN), in bearing fault diagnosis. Another machine learning technique such as support vector machine (SVM) is also

applied in [4]. Chang et al. in [5] summarized other time domain features used in fault diagnosis.

The analysis of the vibration signals' frequency spectrum is the basis of frequency domain analysis. Fundamental frequencies of the signals are calculated through fast Fourier transform. Usually the significant frequencies and the corresponding amplitudes are chosen manually as fault diagnosis features. Frequency domain features are applied with different methods in [6–8]. Time domain features and frequency domain features reflect different characters of the vibration signals, so generally fault diagnosis methods consider them both as classification features. In [9], time domain features and frequency domain features were combined using information fusion and an ANN model was trained for fault diagnosis. Cao et al. in [10] trained a SVM model with feature extraction using PCA method. Other experiments were done trying to take advantage of both domain analyses in [11, 12].

Time-frequency methods are usually effective in extracting the features of the original rotating machinery signals. However, most of the vibration signals may have nonstationary characteristic; other analysis methods are introduced. Wavelet transform is one of the most useful signal analysis methods. Efficient results of applying wavelet transform are shown in [18, 19].

Though traditional analysis methods are mostly effective, however some fundamental mathematic models usually need to be established before applying to the original signals. For instance, the fundamental frequencies need to be selected manually and the bandwidth of filters to preprocess signals is chosen with expert experiences. In real rolling-element bearing systems, signals are more complex and parameters may be hard to extract or determine.

Being a time-frequency analysis technique, empirical mode decomposition (EMD) shows its powerful ability for signal analysis. The analysis process of EMD is not based on predetermined parameters but takes the local time scales of the signals into consideration [20]. In an EMD procedure, the vibration signal of a rotating machine is decomposed into a set of intrinsic mode functions (IMFs). Each IMF may be considered as a basic function of the signal. When the vibration signals are nonlinear and nonstationary, EMD technique may have better performance than traditional techniques. Also, EMD is a self-adaptive processing method, which means less manual work.

Most feature extraction methods mentioned focused on utilizing signal characteristics instead of modeling the signal itself. However, vibration signals still contain rich information. Recently machine learning techniques, especially neural networks, have been widely used in feature engineering. Deep learning technique is a machine learning method proposed in 2006 [21]. The special structure of deep neural network (DNN) makes it possible to extract features for original signals representation [22]. The performance of DNN has been state of the art in many applications, such as computer vision and natural language process [23, 24]. Researchers have applied DNN in fault diagnosis as well [25–28]. Verma et al. [27] purposed a condition monitoring method using sparse autoencoder. In [25], Tagawa et al. built a model based on denoising autoencoder for car fault diagnosis.

Convolutional neural network (CNN) is an important machine learning technique. CNN is a deep neural network structure that mainly focuses on image processing. Like other neural network structures, CNN is formed by a number of neurons, which are organized as the reflection of different overlapping part in the whole field. CNN has been used for image classification and segmentation, and it already has achieved effective results [29, 30].

In this paper, EMD and CNN are both applied as feature extraction method, and a complete structure for fault diagnosis of rolling-element bearing is designed and trained. The following parts of the paper are organized as below. In Section 2, a literature review is given about CNN and EMD applications. Details of CNN and EMD methods and a complete structure of our approach are also described and discussed. In Section 3, the validity of our newly proposed approach for REB fault diagnosis is testified by different experiments which we carried out. In addition, the experiment results are compared with other analysis methods. In the end, the conclusion of this paper is drawn in Section 4.

2. Methodology

First of this section, details of CNN and EMD methods are introduced, after which a complete structure of our approach is described and discussed.

2.1. Convolutional Neural Network. Deep learning methods have outstanding performances in image classification, computer vision, and nature language process. CNN structure is a type of deep neural network. Neurons forming the CNN structure have weights and biases which are changeable and learnable through training.

A number of CNN structures are developed in recent years such as LeNet, GoogleNet, and AlexNet. Figure 1 is a typical structure of LeNet model. Applications in image recognition, video analysis, and nature language process also show the effectiveness of CNN model [31–35].

A CNN structure is made up of three types of layers, which are convolutional layer, subsampling layer, and fully connected layer with a loss function such as SVM or softmax classifier [36]. Typical CNN structure can therefore be divided into two parts. Convolutional layers and subsampling layers work as the feature extractor, while the last layer works as a classifier.

A convolutional layer is the most important and fundamental component of a CNN structure. Each neuron in a convolutional layer receives some inputs of a restricted region in the whole signal. The convolutional layer's weights and biases are considered as a group of convolution kernels (or filter). A kernel only takes a relatively small region of the signal into consideration and projects the whole signal to a brand new feature map, which means dot product is calculated between the signal and each kernel repeatedly. Since the replicated kernel shares the same parameter setup, the number of the network parameters is relatively small.

A $m \times m \times r$ signal vector is input to a convolutional layer as the extractor part of CNN. m is the height and
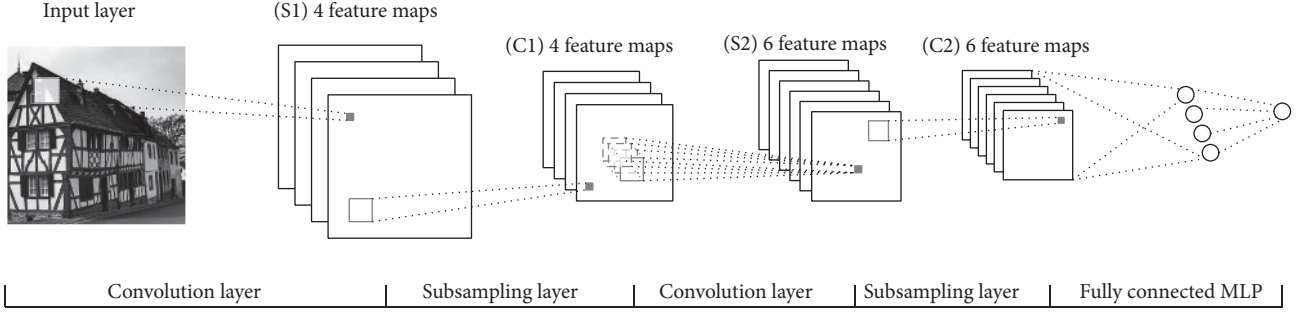
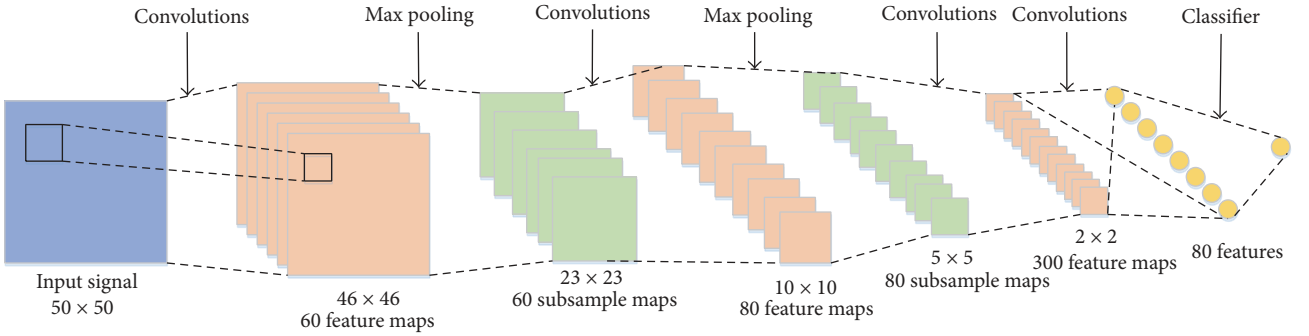FIGURE 1: Typical convolutional neural network structure of LeNet.

FIGURE 2: Structure of convolutional neural network.

width of the input signal, and in general cases the height and width are the same. r is the number of channels of the input. A convolutional layer has k filters (kernels) in the size of $n \times n$, where n is usually less than half the size of the input vector's height m. Each of the filters takes a relatively small local region of the input signal into consideration and projects the whole signal to a brand new feature map, which means dot product is calculated between the signal and each kernel repeatedly. k feature maps are generated with the size of $m - n + 1$. Each feature map is then generally subsampled in contiguous $p \times p$ areas. Types of subsampling techniques include average pooling and maximum pooling depending on the calculation of a restricted area. Also in the pooling process, the pooling areas may be overlapped.

As we know, the convolution layer is used for extracting signal features, and the pooling layer may reduce computation cost. After feature extraction, the extracted features are usually put into a classifier. In this paper, CNN is only used as a feature extractor for fault diagnosis, and the classification part is done after combining other time-frequency domain features.

Figure 2 presents the structure of the CNN structure used in this paper. Consider vibration signals X as the input signals and Y as labels of the signal. In the convolutional layer, a set of feature maps can be acquired by using different filters. Subfeature maps are the result of convoluting multiple input feature maps. The process is calculated as follows:

$$x_j^l = f\left(\sum_{i \in M_j} x_i^{l-1} * w_{ij}^l + b_j^l\right), \tag{1}$$

where M_j represents the selection of input feature maps, l is the lth layer of a network, w is a convolutional filter connecting the $l-1$th layer to the lth layer, f is a nonlinearity active function, and x_j^l represents the feature map generated from the $l-1$th layer. b is the additive bias given to each output feature map.

Traditional nonlinearity active function f used in neural network is sigmoid function ($f(x) = 1/(1 + e^{-x})$), but due to its problem in gradient vanishing, a new active function called Relu (Rectified Linear Units) function is generally used in deep learning methods. The expression of Relu function is $f(x) = \max(0, x)$. Besides solving the gradient vanishing problem in back propagation steps of the neural network training, the amount of calculation would be much less using Relu function. The outputs of some neurons would be zero using Relu function, which leads to the sparsity of the network and avoids the problem of overfitting.

A subsampling layer is calculated as follows:

$$x_j^l = f\left(\beta_j^l \cdot \text{down}\left(x_j^{l-1}\right) + b_j^l\right), \tag{2}$$

where β and b are multiplicative bias and additive bias.

down represents a subsampling function; common subsampling functions are max pooling and average pooling functions. In a max pooling process, the max of the restrict region is chosen as the new feature, while, in an average pooling, a mean value of the same region is calculated as the new feature. Generally speaking, max pooling reflects the most significant characteristic while average pooling smoothens the region and selects the smoothed feature for further use in the following layers.

CNN method has the advantage of extracting feature automatically due to the back propagation (BP) steps. The gradient of the loss function for all the weights in all the layers is calculated by BP algorithm. The mean-squared error (MSE) of the output layer is expressed as follows:

$$E^n = \sum_{k=1}^{m} \left(l_k^n - y_k^n\right)^2. \tag{3}$$

The objective is to minimize the error by reducing the contributions of the network parameters. We calculate the derivative of the MSE to perform gradient descent method on weight w_{ij}^l and bias b_j^l of the neuron. The sensitivities of the error are as follows:

$$\delta = \frac{\partial E}{\partial b} = \frac{\partial E}{\partial z}\frac{\partial z}{\partial b}, \tag{4}$$

where $z = \sum_{i \in M_j} x_i^{l-1} * w_{ij}^l + b_j^l$.

The sensitivities of higher layer are calculated using chain-rule as

$$\delta^l = \left(W^{l+1}\right)^T \delta^{l+1} \circ f'\left(z^l\right). \tag{5}$$

The updating of the weights is then calculated as follows:

$$\Delta W^l = -\eta \frac{\partial E}{\partial W^l} = -\eta x^{l-1}\left(\delta^l\right)^T, \tag{6}$$

where η is the learning rate. The calculations of sensitivities for convolutional layers and subsample layers are different, of which we will not discuss the details in this paper.

In our purposed approach, the CNN structure consists of 4 convolutional layers and 2 subsample layers; detailed parameters are shown in Section 4.

2.2. Empirical Mode Decomposition. The empirical mode decomposition method was first developed by Huang et al. in 1998 [37]. Unlike other signal analysis methods which transform a signal into a certain mode, EMD method focuses on the natural scale and character of the original signal.

In the EMD process, original vibration signal is always decomposed into a certain number of different components which reflect different intrinsic character of the signal. Entropy energy of IMFs contains information of the signal and can be extracted as measurement for fault diagnosis. EMD is superior to traditional signal analysis approach when the signal to be analyzed has nonlinear or nonstationary characters. In addition, EMD technique is self-adaptive analysis processing method which means little manual operation is needed.

After EMD was developed, it has been widely studied in various domains, such as process control [38], voice recognition [39], and system identification [40]. The decomposition result of a simple sample signal is shown in Figure 3.

The fundamental assumption of EMD method is that a sequence of signal is the combination of several different components. In EMD methods, these components are known as intrinsic mode functions. In each of the IMFs, the number
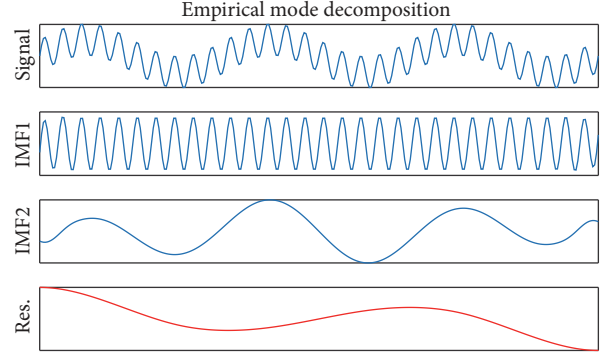

FIGURE 3: Empirical mode decomposition of a sample signal.

of extrema and the number of zero-crossings are the same. Another premise of EMD is that between two contiguous zero-crossings, there is only one extremum [41].

As shown in Figure 2 and mentioned above, the following conditions should be satisfied for IMFs:

(1) In each complete IMF, the difference between the number of extrema and the number of zero-crossings should be less than or equal to one.

(2) In the process of EMD, two envelopes are defined in which the upper envelope is defined by local maxima and the lower envelope by local minima. For each point of an IMF, the mean value of both envelopes should be zero all the time.

The decomposition process of a vibration signal $x(t)$ is described as below:

(1) For a sequence of vibration signal $x(t)$, local extrema are first selected. An envelope is created by connecting the local maxima with cubic spline technique. This envelope is called upper envelope.

(2) Another envelope is created as in (1). All the local minima are connected using the same technique, and the new envelope is called lower envelope. All the points in the signal must be in the range of two envelopes.

(3) The mean value of both envelopes' values is defined as m_1, and we could get h_1 by subtracting the mean value m_1 from the original signal $x(t)$ as follows:

$$x(t) - m_1 = h_1. \tag{7}$$

We validate h_1 to see if both conditions as an IMF are satisfied. If both conditions are satisfied, h_1 is defined as the first composition of $x(t)$.

(4) If either of the conditions is not satisfied, we treat h_1 as the former signal $x(t)$ and then repeat the process from step (1) to step (3), which means a new mean value m_{11} is calculated and then we have

$$h_1 - m_{11} = h_{11}. \tag{8}$$

The process is repeated for k times, until we have h_{1k} which satisfy both premises. We have

$$h_{1(k-1)} - m_{1k} = h_{1k}, \tag{9}$$

and h_{1k} is chosen as the first IMF composition of the signal $x(t)$. c_1 is defined as the first IMF as

$$c_1 = h_{1k}. \tag{10}$$

Normally, c_1 ought to have the most significant feature of the original signal.

(5) Then the IMF is subtracted from signal $x(t)$, and the residue is acquired as

$$r_1 = x(t) - c_1. \tag{11}$$

After that, we consider r_1 as the original signal and repeat the process from step (1) to step (4) until we obtain a new IMF c_2 of $x(t)$.

(6) The whole procedure described above is repeated for n times until we stop the decomposition process. We have

$$r_1 - c_2 = r_2$$
$$\vdots \tag{12}$$
$$r_{n-1} - c_n = r_n.$$

A set of IMFs from c_1 to c_n are acquired. If the residue r_n becomes monotonic, it can reflect the main trend of the original signal. Also no more IMFs could be obtained. In summary, the original signal can be presented as

$$x(t) = \sum_{j=1}^{n} c_j + r_n. \tag{13}$$

Through the EMD process, a combination of n empirical modes is got from decomposing the signals, plus a residue term r_n. Intrinsic mode functions each contain unique frequency bands.

The energy entropy of EMD is calculated and measured as features for fault diagnosis. After decomposing rolling bearing signals into IMFs, energies of the n IMFs are $E_1, E_2, \ldots, E_n$. The energy for one IMF is calculated as

$$E_i = \sum_{j=1}^{m} |c_{ij}|^2, \tag{14}$$

where m is the number of sample data points. And the total energy of all IMFs is calculated as

$$E = \sum_{i=1}^{n} E_i. \tag{15}$$

EMD energy entropy of the signal is calculated as

$$H_{en} = -\sum_{i=1}^{n} p_i \log(p_i), \tag{16}$$

TABLE 1: Time domain features.

Feature	Formula
Mean value	$\bar{x} = \dfrac{1}{N}\sum_{i=1}^{N} x_i$
Standard deviation	$\sigma = \sqrt{\dfrac{1}{N}\sum_{n=1}^{N}(x_i - \bar{x})^2}$
Kurtosis	$K = \dfrac{1}{N}\sum_{n=1}^{N}\dfrac{(x_i - \bar{x})^4}{\sigma^4}$
Skewness	$S = \dfrac{1}{N}\sum_{n=1}^{N}\dfrac{(x_i - \bar{x})^3}{\sigma^3}$
Root mean square	$RMS = \sqrt{\dfrac{1}{N}\sum_{n=1}^{N} x_i^2}$

where $p_i = E_i/E$ is the percentage of the energy entropy of the ith IMF.

In our approach, the energies of the first five IMFs $E_1, E_2, \ldots, E_5$ and the energy entropy H_{en} are chosen as fault features.

2.3. Fault Diagnosis Structure. In this section, the implementation of our proposed fault diagnosis approach is introduced. Figure 4 represents the flowchart of the fault diagnosis process.

In the feature extraction process, five statistical time domain features are selected as fault features, including mean value, standard deviation, skewness, kurtosis, and root mean square (RMS). The formulas of the five features are listed in Table 1.

Fourier transform is applied to vibration signals of rolling-element bearing to obtain the frequency spectrum. A CNN model is designed to extract the spatial information of the frequency spectrum. Eighty features are gained based on CNN methods for classification phase.

Empirical mode decomposition is also applied to vibration signals. Vibration signals in real rolling-element bearing system may be divided into more than 10 IMFs; however the energy of IMF decreases swiftly. In this paper, we only select the first five IMFs. Their energies $E_1, E_2, \ldots, E_5$, as well as the energy entropy H_{en}, are chosen as fault features.

In summary, the vibration signals of rotating machinery are analyzed and a total of 91 features are extracted based on two different methods. In the following classification phase, two effective models, support vector machine (SVM) and softmax classifier, are trained for fault diagnosis of rolling-element bearings.

3. Experiment Results and Analysis

To testify the effectiveness of our approach, experiments were performed on the bearing vibration signal database of Case Western Reserve University (CWRU). CWRU database contains a large amount of data acquired from the experimental setup introduced below.

3.1. Experimental Setup. Figure 5 shows the test platform used in this paper. The experiment apparatus consisted of a motor with horse power of two, a torque transducer, and a
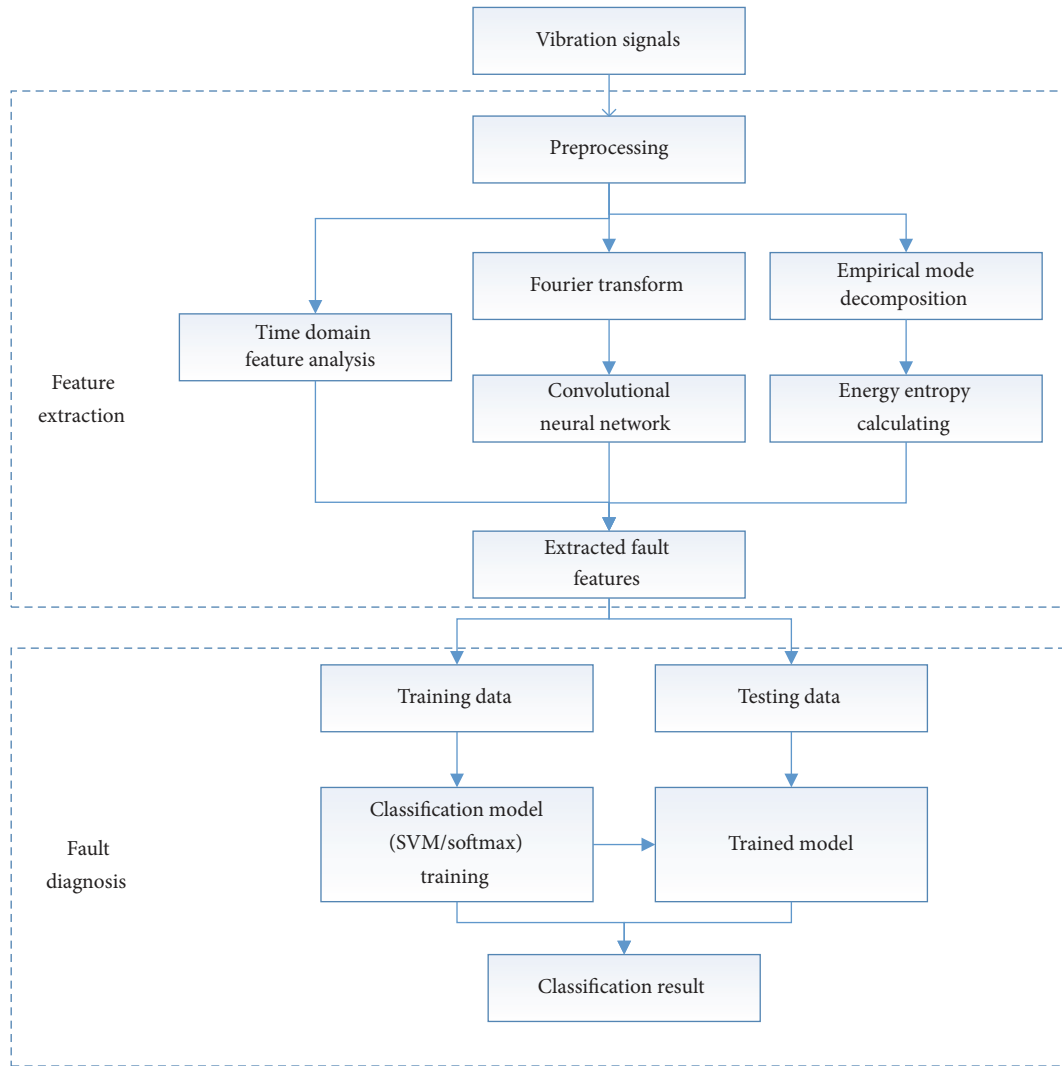
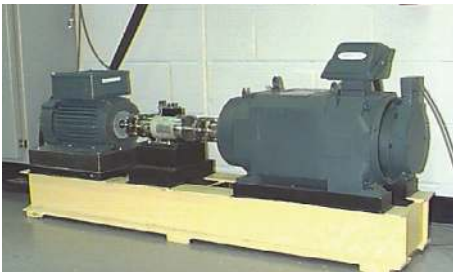FIGURE 4: Representation of proposed fault diagnosis structure.

FIGURE 5: Experiment apparatus for vibration signal acquiring.

dynamometer. Accelerometers are attached to the magnetic bases of the apparatus and vibration signals are acquired under different working conditions which include normal and faulty situations.

3.2. Data Selection and Preprocess. Three bearing components, the inner race (IR), the outer race (OR), and the ball

of rolling bearing (BA), are under study in the database of CWRU. In order to verify this performance of our approach, a set of experiments were conducted. Fault categories of the experiment apparatus include IR faults, BA faults, and OR faults located at three o'clock, six o'clock, and twelve o'clock. In addition, vibration signals under different motor loads and fault diameters are collected for analysis. The sampling frequency of the platform is twelve kHz.

The data set of the bearings used in this paper is arranged in Table 2.

As shown in Table 2, 52 categories of vibration signals are chosen from CWRU database. 1000 samples containing 5000 points each are selected for every category, and 800 samples are randomly selected as training data while 200 samples are left as test data. Two vibration signals and their frequency spectrum are shown in Figures 6 and 7.

3.3. Feature Extraction. As we can see from the vibration signals shown in Figures 6 and 7, original vibration data are disordered and messy, while no recognizable patterns are
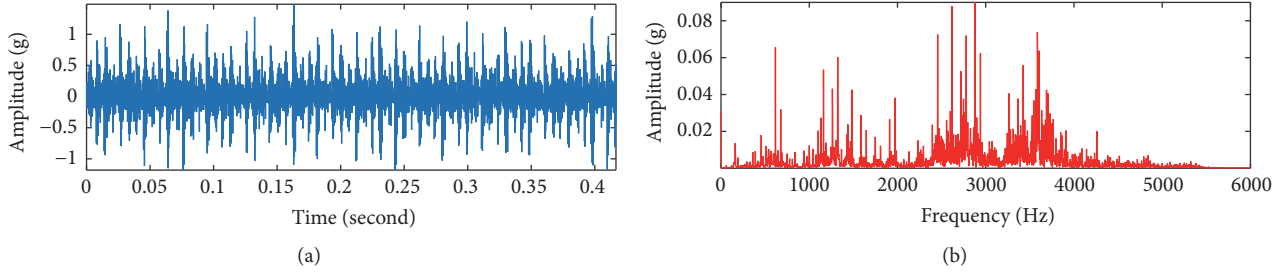
FIGURE 6: Vibration signal and its frequency spectrum under inner race fault with fault diameter of 0.007 inches and motor load of 0.
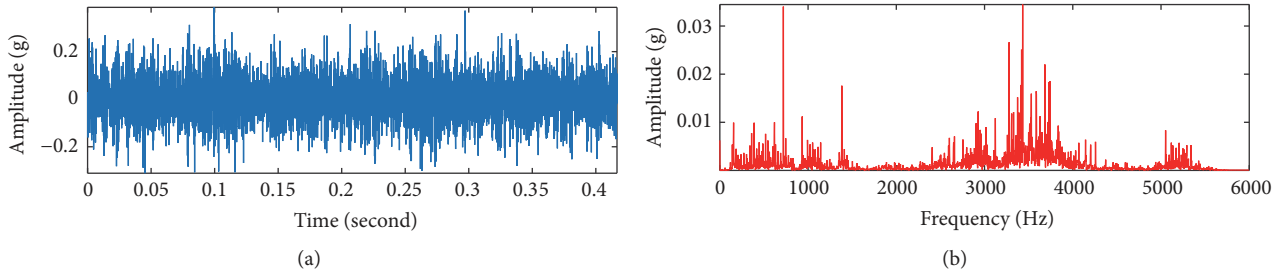

FIGURE 7: Vibration signal and its frequency spectrum under outer race fault at 6 : 00 with fault diameter of 0.014 inches and motor load of 3.

TABLE 2: Bearing fault data arrangement.

Fault position	Motor load /horse power	Fault diameter /inch
IR	0, 1, 2, 3	0.007, 0.014, 0.021
BA	0, 1, 2, 3	0.007, 0.014, 0.021
OR(at 3:00)	0, 1, 2, 3	0.007, 0.021
OR(at 6:00)	0, 1, 2, 3	0.007, 0.014, 0.021
OR(at 12:00)	0, 1, 2, 3	0.007, 0.021

- ○ Training error
- ∗ Test error

FIGURE 8: Training and test error of CNN feature extraction model.

presented. On the other hand, the frequency spectrum may have more notable features, which illustrates that the analysis process using CNN is promising on the side. The original vibration signal contains 5000 points while the frequency spectrum of a signal is a data set of 2500 points. In our approach, the spectrum is reshaped into a 50×50 vector as the input of the CNN model designed above for feature extraction.

In this experiment, mini-batch stochastic gradient descent algorithm was used as approximation method. The batch size was fixed on 100, and the CNN learning rate varied from 0.01 to 0.001. In the training process, we can see the significant ability of CNN in extracting features from the original vibration signals of rotating machinery.

As shown in Figure 8, the training error reduced to almost zero in three epochs, while the test error remained 1.10% after 15 epochs.
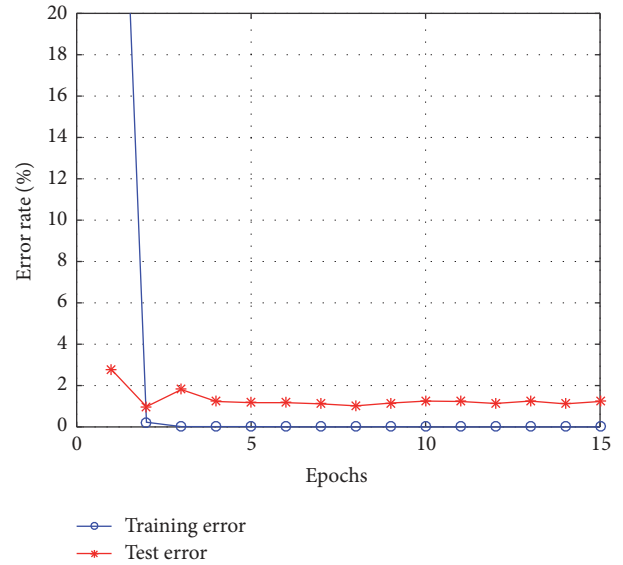
Meanwhile, EMD technique is applied to the original vibration signals as well. Through EMD process, a combination of n empirical modes is got from decomposing the signals, plus a residue term r_n. Intrinsic mode functions each contain unique frequency bands. Figures 9 and 10 show two different vibration signals and their decomposition. Vibration signal from the real platform can be decomposed into about 10 IMFs, and we can get from the functions that the energy decreases rapidly. The sixth IMF usually has an energy level of less than 1, which is less than 1% of the first IMF. So, only the energies of the top five IMFs are chosen as fault features.
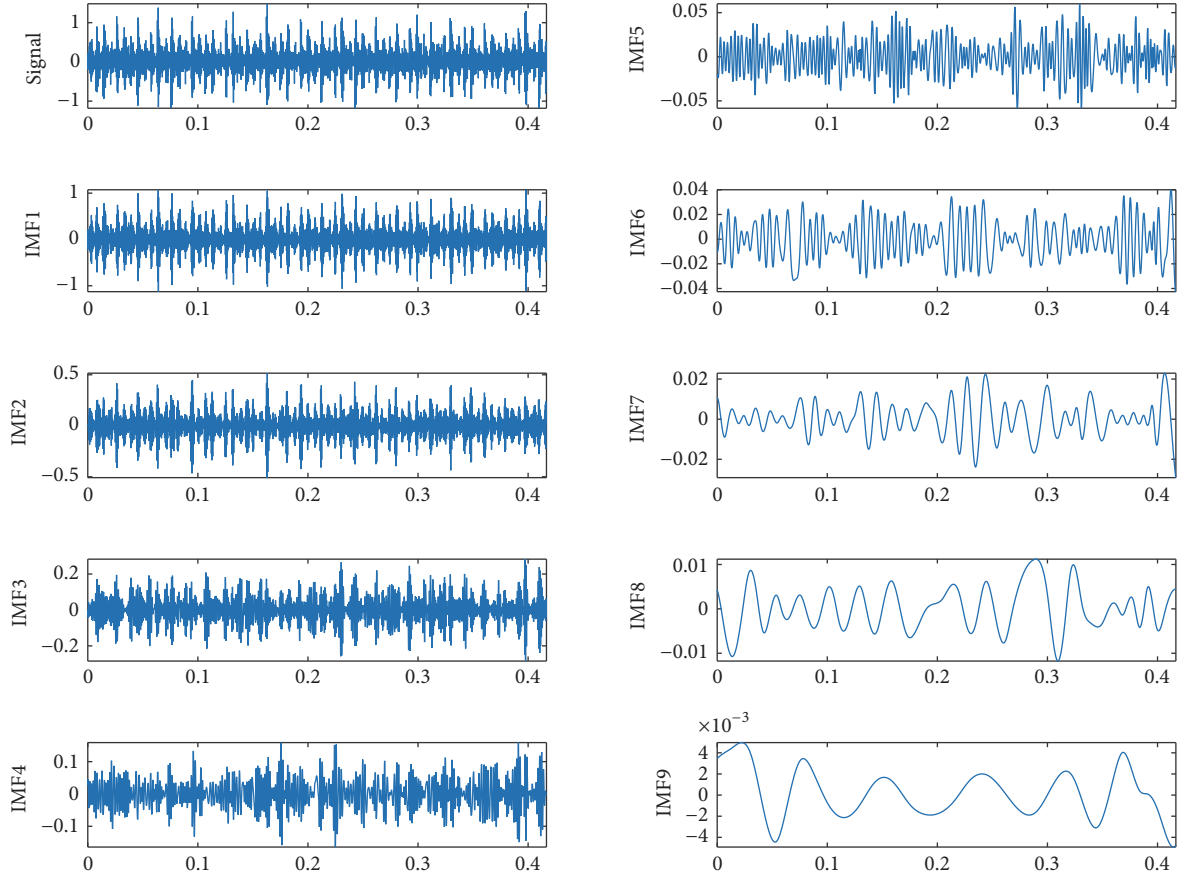
FIGURE 9: Vibration signal and its first 9 IMFs under inner race fault with fault diameter of 0.007 inches and motor load of 0.

TABLE 3: Training accuracy of both classifiers on different features.

Features	SVM	Softmax
11 time-domain and EMD features	88.72% (36908/41600)	89.47% (37220/41600)
80 CNN features	100%	100%
91 combined features	100%	100%

3.4. Result Comparison. After extracting 91 new features of the vibration signal, a classifier model needs to be trained for fault diagnosis. In this paper, both SVM model and softmax classifier are trained to testify the effectiveness of the feature extraction.

We split the 91 features into 2 groups, 80 CNN features and 11 time domain and EMD features, and trained classifiers separately and at last all together. As mentioned in former part, 800 samples of each condition are trained and 200 samples are used as test database, that is, a set of 41600 training data sets and 10400 test data sets. The results are presented in Tables 3 and 4.

The training accuracy of both methods is rather high as shown in Tables 3 and 4 which represented that both classifiers trained on 91 combined features achieved an outstanding

test accuracy. 10374 of 10400 samples are classified correctly using SVM while 10346 samples are correct using softmax classifier. Two classification methods are both competitive and effective, and SVM method shows a slight superiority.

The results also demonstrate the powerful feature extraction ability of CNN. As we can see, features from CNN model alone can reach a relatively high performance; however, features from CNN model have limitation in fault classification. Efforts have been done trying to alter the parameters or even structures of the CNN model, but features extracted can only get a classification accuracy around 99%. Time domain features and EMD features are easier to obtain compared with CNN, and they are also useful in many situations. By combining features from both methods, we can achieve a superior result compared to using them separately.
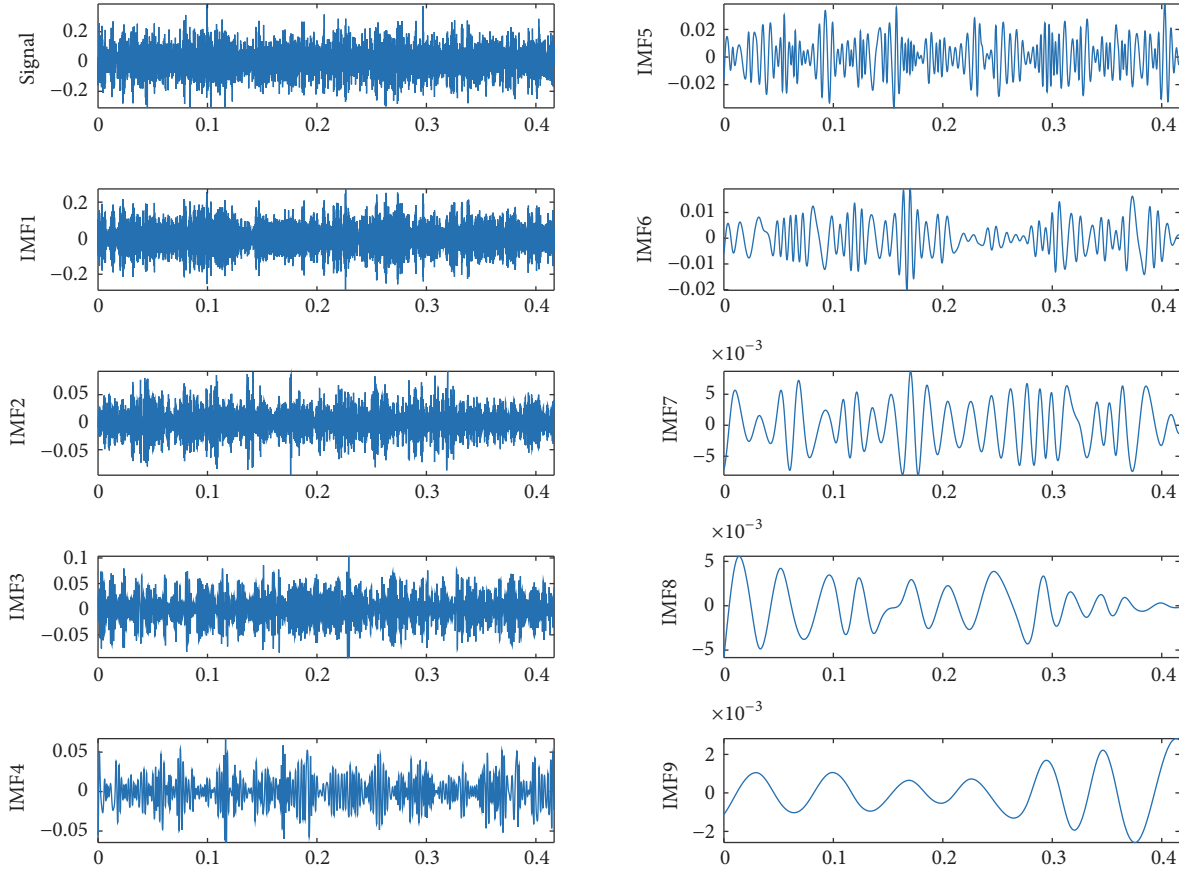
FIGURE 10: Vibration signal and its first 9 IMFs under outer race fault at 6 : 00 with fault diameter of 0.014 inches and motor load of 3.

TABLE 4: Test accuracy of both classifiers on different features.

Features	SVM	Softmax
11 time-domain and EMD features	83.14% (8647/10400)	82.93% (8625/10400)
80 CNN features	99.05% (10301/10400)	98.90% (10286/10400)
91 combined features	99.75% (10374/10400)	99.48% (10346/10400)

The results of our proposed approach are also compared with works in some other papers. Table 5 below shows classification accuracy of some other works.

As shown in Table 5, traditional ANN combined with EMD method already has a high accuracy in [13]. CNN has been applied in fault diagnosis in [14–17]. CNN structures in [15, 16] show great performance in classification. However, with a small number of categories, CNN would not always have better results than traditional methods as shown in [17]. Most works only dealt with a small number of categories, which is not adequate in practical situations, while our approach deals with 52 fault categories. Our proposed approach with 91 features has the best performance in the table.

3.5. Parameter Selection for CNN. In our purposed approach, the CNN structure consists of 4 convolutional layers and 2 subsample layers; detailed parameters are shown in Table 6. In a CNN structure, usually bigger number of filters shows better ability of representation. As there are 52 fault categories, filter numbers should be bigger than 52. Convolutional layers show different kinds of characteristics, and the later convolutional layer represents more delicate details than former layers. Therefore, in layer C3, we select 300 filters for better representation.

The number of features extracted from CNN model is very important. Experiments are implemented with different number of features. The results are shown in Figure 11. As we can see, different numbers of features have different accuracies. 80 features show the best representation ability while more features may lead to the problem of overfitting.

The optimization of parameters of CNN is always important to obtain an effective CNN model. In general, learning

TABLE 5: Classification accuracy of different methods.

Method	Classification accuracy	Number of categories
80 features—softmax	98.90%	52
80 features—SVM	99.05%	52
91 features—softmax	**99.48**%	52
91 features—SVM	**99.75**%	52
EMD-ANN [13]	96.24%	3
Wavelet-ANN [13]	88.54%	3
CNN with 2 pipelines [14]	93.61%	8
CNN with statistical feature [15]	98.02%	12
CNN with statistical feature [15]	98.35%	8
Hierarchical ADCNN [16]	98.13%	3
SVRM [16]	94.17%	3
1D-CNN [17]	97.40%	2
WP-SVM [17]	99.20%	2
FFT-SVM [17]	84.20%	2

TABLE 6: Parameters of the purposed CNN structure.

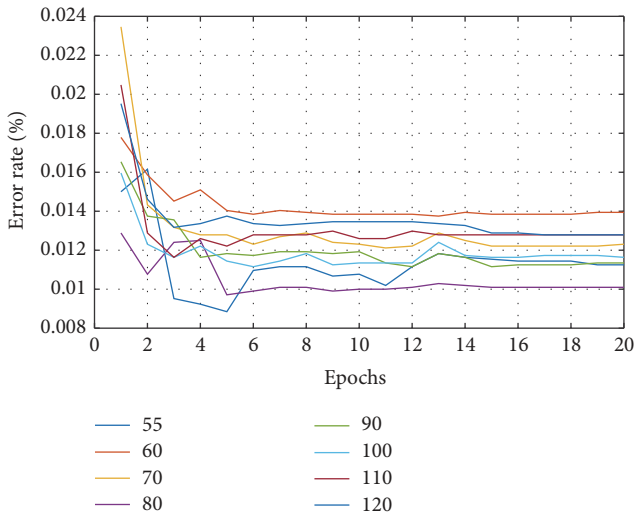| Layer	C1	S1	C2	S2	C3	C4
Length	5	2	5	2	4	2
Strides	1	2	2	2	1	1
Filter numbers	60	/	80	/	300	80
Parameter numbers	1560	/	124800	/	408000	120000

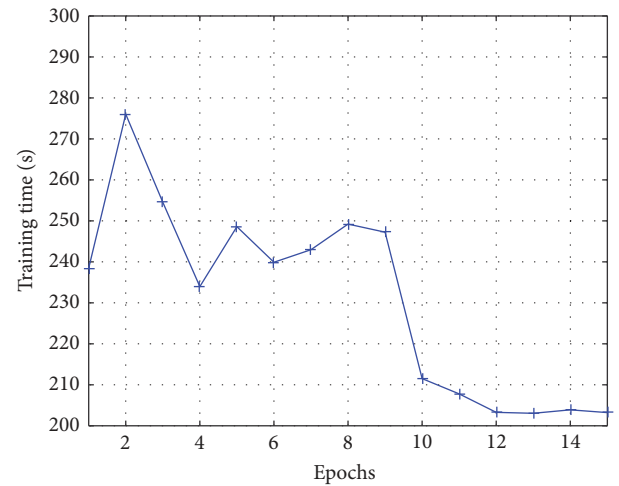FIGURE 11: Error rate with different numbers of CNN features.

FIGURE 12: Training time of CNN model.

rate, number of kernels, number of weights in each layer, and batch size are all parameters to be optimized.

In our purposed CNN model, as shown in Table 6, a total number of 654360 weights and bias parameters need to be calculated in each step, which results in a relatively long training time. Training time of the CNN model in this paper is shown in Figure 12, and the average training time is about 240 seconds.

The selecting of learning rate of the mini-batch SGD algorithm is also considered. An appropriate learning rate is important to the final results. Higher learning rate leads to faster descent, while lower rate may cause the optimization to be local but not global.

A series of experiments were done trying out different learning rate, and some of the results are shown in Figure 13. As shown in the figure, training error collapses to nearly zero in no more than four epochs, except the one with learning rate of 0.001. Due to the small learning rate, the CNN model cannot get a satisfied result. The results in this paper and other CNN parameter-adjusting algorithms indicate that the variational learning rate is the best choice here.
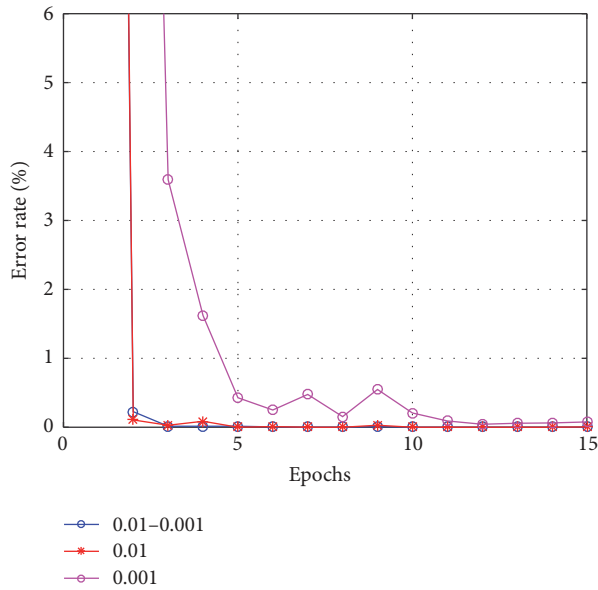
Figure 13: Training error with different learning rate.

Generally, the numbers of weights and filters affect the feature explanation capacity of CNN. Larger number of parameters usually suggest a better representation ability along with a larger computing expense. We conducted experiments with fewer weights and filters, and the performance indicated that effect on the final result is not significant. The parameters in our designed CNN are suitable for application of fault diagnosis.

4. Conclusions

In this paper, a novel approach for rotating machinery fault diagnosis was proposed, in which CNN and EMD were applied to extract features from raw vibration signals. A SVM model and a softmax classification model are trained using combined features. With rolling-element bearing data collected from CWRU experimental setup, experiments are implemented under different situations. Fifty-two thousand samples under 52 working conditions are arranged for the experiment in this paper.

Experiment results also demonstrate the powerful feature extraction ability of CNN. Classification based on features extracted from CNN model alone can reach a relatively high accuracy. However, features from CNN model have a limitation in fault classification due to its generalization ability. To improve the performance of classification, time domain features and EMD features, which are easier to calculate, work as complementary features for CNN model. The proposed approach represents its superior ability of extracting features from original vibration signals self-adaptively, and it is practical and effective in fault diagnosis for rotating machinery.

Deep learning algorithm shows an excellent expression capacity while increasing the expense of computing; on the contrary, traditional signal analysis methods are generally more convenient to calculate. It is important to analyze the ability of feature explanation for both deep learning algorithms and traditional methods. Further exploration about the effectiveness of other deep learning structures will be investigated in future work.
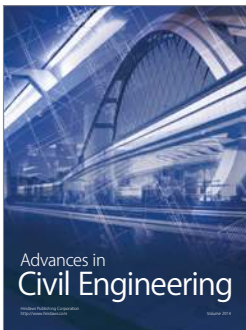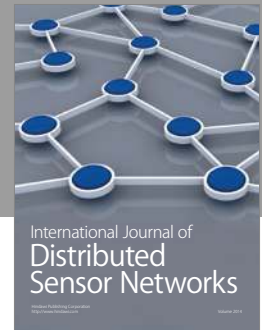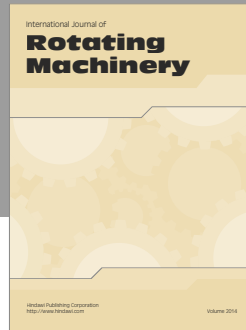
Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] H. Sriyananda and D. R. Towill, "Fault diagnosis using time domain measurements," *Radio and Electronic Engineer*, vol. 43, no. 9, pp. 523–533, 1973.

[2] Q. Hu, S. Zhang, and S. Yang, "Variable condition bearing fault diagnosis based on time-domain and artificial intelligence," *Applied Mechanics and Materials*, vol. 203, pp. 329–333, 2012.

[3] B. Sreejith, A. K. Verma, and A. Srividya, "Fault diagnosis of rolling element bearing using time-domain features and neural networks," in *Proceedings of the IEEE Region 10 Colloquium and 3rd International Conference on Industrial and Information Systems (ICIIS '08)*, pp. 1–6, Kharagpur, India, December 2008.

[4] O. R. Seryasat, M. Aliyari Shoorehdeli, F. Honarvar, and A. Rahmani, "Multi-fault diagnosis of ball bearing based on features extracted from time-domain and multi-class support vector machine (MSVM)," in *Proceedings of the 2010 IEEE International Conference on Systems, Man and Cybernetics, (SMC '10)*, pp. 4300–4303, Istanbul, Turkey, October 2010.

[5] J.-B. Chang, T.-F. Li, and P.-F. Li, "The selection of time domain characteristic parameters of rotating machinery fault diagnosis," in *Proceedings of the 2010 International Conference on Logistics Systems and Intelligent Management (ICLSIM '10)*, pp. 619–623, Harbin, China, January 2010.

[6] X. Zhou and D. Luo, "Research of amplitude-frequency domain parameters analysis for condition detection and fault diagnosis," *Research Journal of Applied Sciences, Engineering and Technology*, vol. 4, no. 19, pp. 3787–3790, 2012.

[7] K. Mao and Y. Wu, "Fault diagnosis of rolling element bearing based on vibration frequency analysis," in *Proceedings of the 3rd International Conference on Measuring Technology and Mechatronics Automation, (CMTMA '11)*, pp. 198–201, Shangshai, China, January 2011.

[8] J. Cao, L. Chen, J. Zhang, and W. Cao, "Fault diagnosis of complex system based on nonlinear frequency spectrum fusion," *Measurement: Journal of the International Measurement Confederation*, vol. 46, no. 1, pp. 125–131, 2013.

[9] Z. Jiang, W. Jiao, and S. Meng, "Fault diagnosis method of time domain and time-frequency domain based on information fusion," *Applied Mechanics and Materials*, vol. 300, pp. 635–639, 2013.

[10] M. Cao, H. Pan, and X. Chang, "Research on automatic fault diagnosis based on time-frequency characteristics and PCA-SVM," in *Proceedings of the 13th International Conference on Ubiquitous Robots and Ambient Intelligence, (URAI '16)*, pp. 593–598, Xi'an, China, August 2016.

[11] Z. T. Yao and H. X. Pan, "The engine fault diagnosis based on time domain and frequency domain," *Advanced Materials Research*, vol. 936, pp. 2243–2246, 2014.

[12] X.-W. Deng, P. Yang, J.-S. Ren, and Y.-W. Yang, "Rolling bearings time and frequency domain fault diagnosis method based on Kurtosis analysis," in *Proceedings of the 6th IEEE PES Asia-Pacific Power and Energy Engineering Conference, (APPEEC '14)*, Hong Kong, China, December 2014.

[13] Y. Yu and C. Junsheng, "A roller bearing fault diagnosis method based on EMD energy entropy and ANN," *Journal of Sound and Vibration*, vol. 294, no. 1, pp. 269–277, 2006.

[14] O. Janssens, V. Slavkovikj, B. Vervisch et al., "Convolutional neural network based fault detection for rotating machinery," *Journal of Sound and Vibration*, vol. 377, pp. 331–345, 2016.

[15] Z. Chen, C. Li, and R.-V. Sanchez, "Gearbox fault identification and classification with convolutional neural networks," *Shock and Vibration*, vol. 2015, Article ID 390134, 10 pages, 2015.

[16] X. Guo, L. Chen, and C. Shen, "Hierarchical adaptive deep convolution neural network and its application to bearing fault diagnosis," *Measurement: Journal of the International Measurement Confederation*, vol. 93, pp. 490–502, 2016.

[17] T. Ince, S. Kiranyaz, L. Eren, M. Askar, and M. Gabbouj, "Real-time motor fault detection by 1-D convolutional neural networks," *IEEE Transactions on Industrial Electronics*, vol. 63, no. 11, pp. 7067–7075, 2016.

[18] L. Han, J. Hong, and D. Wang, "Fault diagnosis of aero-engine bearings based on wavelet package analysis," *Tuijin Jishu/ Journal of Propulsion Technology*, vol. 30, no. 3, pp. 328–341, 2009.

[19] M. Deriche, "Bearing fault diagnosis using wavelet analysis," in *Proceedings of the 2005 1st International Conference on Computers, Communications and Signal Processing with Special Track on Biomedical Engineering, (CCSP '05)*, pp. 197–201, Kuala Lumpur, Malaysia, Malaysia, November 2005.

[20] C. Junsheng, Y. Dejie, and Y. Yu, "A fault diagnosis approach for roller bearings based on EMD method and AR model," *Mechanical Systems and Signal Processing*, vol. 20, no. 2, pp. 350–362, 2006.

[21] G. E. Hinton, S. Osindero, and Y.-W. Teh, "A fast learning algorithm for deep belief nets," *Neural Computation*, vol. 18, no. 7, pp. 1527–1554, 2006.

[22] A. J. M. Timmermans and A. A. Hulzebosch, "Computer vision system for on-line sorting of pot plants using an artificial neural network classifier," *Computers and Electronics in Agriculture*, vol. 15, no. 1, pp. 41–55, 1996.

[23] Y. Yao and Z. Huang, "Bi-directional LSTM Recurrent Neural Network for Chinese Word Segmentation," in *Neural Information Processing*, vol. 9950 of *Lecture Notes in Computer Science*, pp. 345–353, Springer, Cham, 2016.

[24] Li. Deng, "A tutorial survey of architectures, algorithms, and applications for deep learning," in *Transactions on Signal and Information Processing*, 2014.

[25] T. Tagawa, Y. Tadokoro, and T. Yairi, "Structured denoising autoencoder for fault detection and analysis," *ACML*, 2014.

[26] M. Sakurada and T. Yairi, "Anomaly detection using autoencoders with nonlinear dimensionality reduction," in *Proceedings of the 2nd Workshop on Machine Learning for Sensory Data Analysis, (MLSDA '14)*, pp. 4–11, Gold Coast, Australia QLD, Australia.

[27] N. K. Verma, V. K. Gupta, M. Sharma, and R. K. Sevakula, "Intelligent condition based monitoring of rotating machines using sparse auto-encoders," in *Proceedings of the 2013 IEEE International Conference on Prognostics and Health Management, (PHM '13)*, Gaithersburg, MD, USA, June 2013.

[28] B. Yan and Q. Weidong, "Aero-engine sensor fault diagnosis based on stacked denoising autoencoders," in *Proceedings of the 35th Chinese Control Conference, (CCC '16)*, pp. 6542–6546, Chengdu, China, July 2016.

[29] C. Ciresan Dan, "Flexible, high performance convolutional neural networks for image classification," in *Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence (IJCAI '11)*, vol. 22, Barcelona, Catalonia, Spain, July 2011.

[30] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Proceedings of the 26th Annual Conference on Neural Information Processing Systems (NIPS '12)*, pp. 1097–1105, Lake Tahoe, Nev, USA, December 2012.

[31] M. Matsugu, K. Mori, Y. Mitari, and Y. Kaneda, "Subject independent facial expression recognition with robust face detection using a convolutional neural network," *Neural Networks*, vol. 16, no. 5, pp. 555–559, 2003.

[32] C. Szegedy, W. Liu, Y. Jia et al., "Going deeper with convolutions," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR '15)*, pp. 1–9, Boston, Mass, USA, June 2015.

[33] O. Russakovsky and etal., "Imagenet large scale visual recognition challenge," *International Journal of Computer Vision*, vol. 115, no. 3, pp. 211–252, 2015.

[34] A. Karpathy, G. Toderici, S. Shetty, T. Leung, R. Sukthankar, and F.-F. Li, "Large-scale video classification with convolutional neural networks," in *Proceedings of the 27th IEEE Conference on Computer Vision and Pattern Recognition, (CVPR '14)*, pp. 1725–1732, Columbus, OH, USA, June 2014.

[35] E. Grefenstette, P. Blunsom, N. de Freitas, and K. M. Hermann, "A Deep Architecture for Semantic Parsing," in *Proceedings of the ACL 2014 Workshop on Semantic Parsing*, pp. 22–27, Baltimore, MD, USA, June 2014.

[36] Y. LeCun, "Deep learning & convolutional networks," http:// yann.lecun.com/exdb/lenet.

[37] N. E. Huang, Z. Shen, S. R. Long et al., "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," *The Royal Society of London. Proceedings. Series A. Mathematical, Physical and Engineering Sciences*, vol. 454, no. 1971, pp. 903–995, 1998.

[38] R. Srinivasan, R. Rengaswamy, and R. Miller, "A modified empirical mode decomposition (EMD) process for oscillation characterization in control loops," *Control Engineering Practice*, vol. 15, no. 9, pp. 1135–1148, 2007.

[39] E. Ambikairajah, "Emerging features for speaker recognition," in *Proceedings of the 6th International Conference on Information, Communications and Signal Processing, (ICICS '07)*, Singapore, Singapore, December 2007.

[40] Y. B. Yang and K. C. Chang, "Extraction of bridge frequencies from the dynamic response of a passing vehicle enhanced by the EMD technique," *Journal of Sound and Vibration*, vol. 322, no. 4-5, pp. 718–739, 2009.

[41] N. E. Huang, Z. Shen, and S. R. Long, "A new view of nonlinear water waves: the Hilbert spectrum," *Annual Review of Fluid Mechanics*, vol. 31, pp. 417–457, 1999.

Journal of
Engineering

The Scientific
World Journal

International Journal of
Rotating
Machinery

Journal of
Sensors

International Journal of
Distributed
Sensor Networks

Advances in
Civil Engineering

Journal of
Control Science
and Engineering

Journal of
Robotics

Hindawi

Submit your manuscripts at
https://www.hindawi.com

Journal of
Electrical and Computer
Engineering

Advances in
OptoElectronics

VLSI Design

International Journal of
Navigation and
Observation

Modelling &
Simulation
in Engineering

International Journal of
Aerospace
Engineering

International Journal of
Chemical Engineering

International Journal of
Antennas and
Propagation

Active and Passive
Electronic Components

Shock and Vibration

Advances in
Acoustics and Vibration