IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received May 9, 2019, accepted May 20, 2019, date of publication May 27, 2019, date of current version June 10, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2919126

Bearing Fault Classification Based on Convolutional Neural Network in Noise Environment

QINYU JIANG, FALIANG CHANG, AND BOWEN SHENG
School of Control Science and Engineering, Shandong University, Jinan 250061, China

Corresponding author: Faliang Chang (flchang@sdu.edu.cn)

This work was supported in part by the National Key R&D Program of China under Grant 2018YFB1305300, in part by the National Natural Science Foundation of China under Grant 61673244 and Grant 61703240, and in part by the Key R&D Projects of Shandong Province of China under Grant 2017GGX10137.

ABSTRACT Bearing fault diagnosis is an important technique in industrial production as bearings are one of the key components in rotating machines. In bearing fault diagnosis, complex environmental noises will lead to inaccurate results. To address the problem, bearing fault classification methods should be capable of noise resistance and be more robust. In previous studies, researchers mainly focus on noise-free condition, measured signal and signal with simulated noise, many effective approaches have been proposed. But in real-world working condition, strong and complex noises are often leads to inaccurate results. According to the situation, this work focuses on bearing fault classification under the influence of factory noise and the white Gaussian noise. In order to eliminate the noise interference and take the possible connection between signal frames into consideration, this paper presents a new bearing fault classification method based on convolutional neural networks (CNNs). By using the sensitivity to impulse of spectral kurtosis (SK), noises are repressed by the proposed filtering approach based on the SK. Mel-frequency cepstral coefficients (MFCC) and delta cepstrum are extracted as the feature by the reason of satisfactory performance in sound recognition. And in consideration of the connection between frames, a feature arrangement method is presented to transfer feature vectors to feature images, so the advantages of the CNNs in the fields of image processing can be exploited in the proposed method. The proposed method is demonstrated to have strong ability of classification under the interference of factory noise and the Gaussian noise by experiments.

INDEX TERMS Bearing fault, convolutional neural network, fault diagnosis, spectral kurtosis.

I. INTRODUCTION

Rotating machines are indispensable important equipments in industry, such as induction motor and turbine, etc. Rolling element bearings (REB) are one of the most common components in rotating machineries, and they are also one of the most brittle parts [1]. The reasons of failures including heavy loads, inadequate lubrication, friction caused by foreign matters owing to bad sealing, etc. Therefore, effective bearing fault classification methods should be presented to ensure the safe operation of equipments.

Localized faults in REBs result in specific spectral characteristics, called the bearing fault frequencies. The previous bearing diagnosis methods mainly include fault frequency calculation, signal processing, and envelope analysis [2]. Through the processes, fault types can be recognized generally. But it should be mentioned that it is hard for traditional methods to distinguish all the fault types and to evaluate damage degrees. Such as, different radiuses of damages in the inner race will cause similar bearing fault frequencies. In recent years, many advanced approaches about machine learning based fault diagnosis and fault feature extraction have been proposed [3]–[5].

In a real factory environment, noises contain impulses on a wide frequency range which may result in violent spectral changes, and this increases the difficulty of diagnosis. To address this problem, a bearing fault classification method based on convolutional neural networks (CNNs) for heavy noise environment is proposed in this paper.

Spectral kurtosis (SK) is a statistical indicator which can reveal the non-Gaussian components and the corresponding

The associate editor coordinating the review of this manuscript and approving it for publication was Aysegül Ucar.

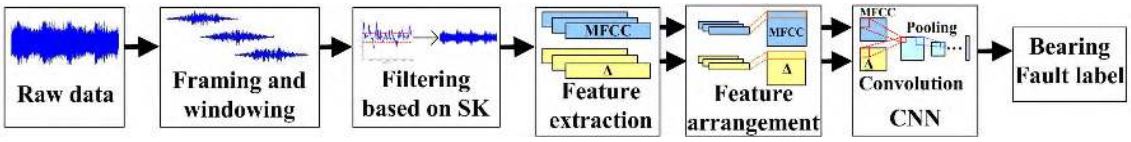

FIGURE 1. Framework of the proposed method.

frequency location in a signal, which was first introduced by Dwyer [6]. If a signal contains series of transients, its SK will be distinctly different from the signal which obeys a Gaussian distribution. Based on the theory above, Jerome Antoni demonstrated the high potential of the SK to detect and characterize non-stationary signals [7], and used the SK as a tool for vibration monitoring and rotating machines fault diagnosis [8]. Recently, SK has been used as feature for fault detection [9], etc.

CNNs was put forward in the late twentieth century, and shows a strong capacity in the field of image processing. Several classical and efficient CNNs models have been proposed, such as LeNet-5 [10] which is identified as the initial CNNs model, AlexNet [11] for images recognition, and so on. Meanwhile, CNNs have been successfully applied in the fields of medical image retrieval and video classification [12], [13], etc. Compared with traditional manually engineered features, CNNs can extract effective features from input data automatically through multilevel convolution and pooling operation, which tends to be more efficient than the features of artificial selection. On account of the satisfactory performance of CNNs, the model was introduced in the area of 1D signal, such as speech and sound recognition, fault detection [14]–[17], etc. In recent years, CNNs have been applied in the field of fault diagnosis. In [18], 1D CNN is built for bearing fault diagnosis, and in [19], authors present a method that transforms bearing signals to explore feature images for CNN to classify bearing data. To take advantage of the strength of CNNs in the fields of 2D data classification, and in consideration of the possible connection between signal frames, a feature map arrangement method is presented to transform 1D feature vectors to 2D feature matrixes for precisely expressing the data and a CNN is built for fault classification. otherwise, in some cases, training data are insufficient for training an accurate fault classifier presented in this work, the proposed feature map arrangement methods can address the problem generally.

Normally, bearing fault results in non-Gaussian components, while other parts obey Gaussian distribution [8], so the SK curve can be seen as the amplitude-frequency response of a filter which can remove Gaussian noises from the raw signal. Meanwhile, the non-Gaussian parts in factory noises have no effect with the result if the noise is quite static. To provide a relatively stable input for the fault classifier, Mel-frequency cepstral coefficients (MFCC), and delta cepstrum are extracted as the feature due to the satisfactory application in the fields of sound and vibration signal processing. At last, a CNN is established as a fault classifier.

In this work, in consideration of the connections between signal frames and reducing the impact of erroneous frames, a novel bearing fault classification method is presented which transforms 1D bearing data to 2D feature matrix and builds a CNN as the fault classifier. Meanwhile, based on the character of outlier sensitively, in response to Gaussian noise effect, a simple filtering method based on SK is presented in this paper. The rest of this paper is organized as follows.

The framework of the proposed bearing fault classification method and the main process are introduced in section II. In section III, this paper briefly recall the definition of SK and presenting the filtering method. A brief introduction of MFCC and delta cepstral is described in section IV. In section V, the feature arrangement method and the applied architecture of CNNs in this paper are introduced. In section VI, several experiments are presented to demonstrate the effectiveness of the proposed method.

II. FRAMEWORK OF BEARING FAULT CLASSIFICATION METHOD

The proposed method mainly includes 5 processes: framing and windowing, filtering, feature extraction, feature arrangement and classification. The frame work is shown in Fig. 1.

In the process of filtering, the SK of the raw data should be calculated first. Then thresholding the SK curve to ensure that the modified SK can be transformed to the amplitude-frequency response of filters, and filtering the input data. To represent the bearing fault signals, MFCC and delta cepstrum are extracted as the two-input-channel features. To maintain the continuity and comparability between samples in the same class, a vector arrangement method is proposed to transfer 1D feature vectors to 2D feature matrixes. At last, on account of the satisfactory performance in 2D signal processing, such as face, license plates detection, and recognition [20], a CNN is established to give the label of input samples. The detailed processes are described in the following sections.

III. FILTERING BASED ON SPECTRAL KURTOSIS

A. SPECTRAL KURTOSIS

Kurtosis has been employed in the signal-processing community to solve 'blind' problems, it is very sensitive to non-Gaussian signals, the kurtosis of a signal x is:

$$k(x) = \frac{E[(x-\mu)^4]}{\delta^4} - 3 \tag{1}$$

where μ and δ are the mean and standard deviation of the signal. The spectral version of kurtosis, SK, is the kurtosis computed at the output of a perfect filter-bank at each frequency band [8], and it is commonly defined as the normalized fourth-order cumulate of the Fourier transform [21]. In order to simplify calculating, Antoni [7] presented to estimate the SK with the short-time Fourier transform (STFT)-based estimator.

Let $Y(n)$ be a sampled signal, then for a N_w wide analysis window $w(n)$ and a given temporal step P, the STFT of process $Y(n)$ over frequency index f is defined as:

$$Y_w(iP, f) = \sum_{n=-\infty}^{+\infty} Y(n)w(n - iP)e^{-j2\pi nf} \tag{2}$$

Then, the definition of the $2m$th order empirical spectral moment of $Y_w(iP, f)$ is:

$$\hat{S}_{2mY}(f) = \left\langle Y_w(iP, f)^{2m} \right\rangle_i \tag{3}$$

where $\langle \cdots \rangle_i$ is the time-averaged operator over index i. At last, the SK can be defined as:

$$\hat{K}_Y(f) = \frac{\hat{S}_{4Y}(f)}{\hat{S}^2_{2Y}(f)} - 2 \tag{4}$$

In the actual situation, the length of STFT analysis window can be set as the length of frames if the signal is framed previously, then the process of STFT can be regarded as discrete Fourier transform (DFT). For instance, a mixed signal x which can be represented as:

$$\begin{aligned} x(t) = e^{-5(t-0.4)} sin(300 \times 2\pi t)\varepsilon(t - 0.4)] + wgn(t) \\ + 0.5[sin(100 \times 2\pi t) + sin(200 \times 2\pi t) \end{aligned} \tag{5}$$

where the 300Hz part is the non-Gaussian part, wgn is white Gaussian noise and ε is step function. The SK curve of x is shown in Fig. 2.

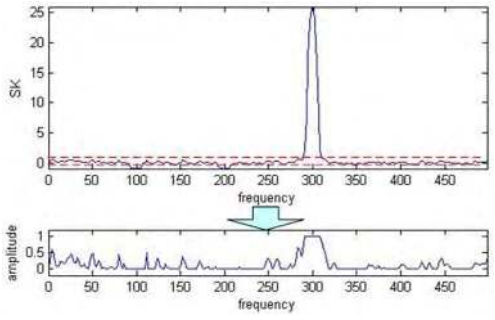

FIGURE 2. The SK curve of the signal x and the amplitude-frequency curve of the filter.

From the SK curve, the non-Gaussian part is distinctly different with other parts. Meanwhile, for vibration or sound signal of bearings, most faults show the non-Gaussian characteristics [8], so it is available to use SK to filter irrelevant information.

B. FILTERING METHOD

To extract the non-Gaussian part of the mixed signal, the SK curve can be used as the basis of filtering. The filtering method based on SK includes:

1) Calculate the SK curve of the noised signal.
2) Thresholding the curve to 0 to 1, as it shows in Fig. 2, consider the SK curve as the magnitude-frequency curve of the filter.

$$\hat{K}_Y(f) = \begin{cases} 0 & K_Y(f) \le 0 \\ K_Y(f) & 0 < K_Y(f) \le 1 \\ 1 & 1 < K_Y(f) \end{cases} \tag{6}$$

3) Filtering the signal by the obtained filter.

For the step 2), the threshold level of the curve from 0 to 1 will prevent over-amplify non-Gaussian parts of the signal in the process of filtering and the meaningless negative value in the magnitude-frequency curve. The filter will keep all the non-Gaussian parts and restrain others according to the magnitude-frequency curve shown in Fig. 2. The temporal waveform and spectrum of the filtered version of $x(t)$ are shown in Fig. 3. The result proves that the filtering method based on SK can extract non-Gaussian components and reduce the effect of Gaussian parts.

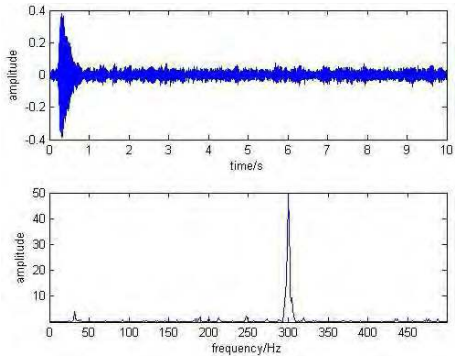

FIGURE 3. The temporal waveform and spectrum of the filtered signal.

IV. MEL-FREQUENCY CEPSTRAL COEFFICIENTS FEATURE

A. MEL-FREQUENCY CEPSTRAL COEFFICIENTS

MFCCs is a commonly used technique in the applications of sound and vibration signal for feature extraction, which could reveal the time-variant features effectively [22], [23]. MFCCs feature was first introduced to simulate the human perception due to the utilized Mel scale and it reveals the short-term power spectrum of the signal [24]. Moreover, the delta spectrum of MFCCs show the dynamic characteristics of signals while MFCCs represent the static characteristics. In consideration of the experimental data in this paper are sound and vibration signal, MFCCs and the delta spectrum are extracted as the feature.

B. FEATURE EXTRACTION

The process of extracting MFCC and delta cepstrum are:

1) FRAMING, WINDOWING, AND DISCRETE FOURIER TRANSFORMATION (DFT)

For processing conveniently, signal should be cut in frames generally. But framing directly will caused frequency leakage and lack of continuity. So overlapping framing and windowing are used. Normally Hanning window is used for framing sound and vibration signal, and the overlap is 10% to 50% window width. A Hanning window can be expressed as:

$$w(t) = 0.5 - 0.5cos[\frac{2\pi(t+1)}{T+1}], \quad 0 \le t \le T-1 \tag{7}$$

where T is the length of the Hanning window, $w(t)$ is the tth point of the window. In addition, the filtering method is based on the experimental data, so data are filtered after framing and windowing.

After frames are got, do DFT to every frame to get the frequency spectrum, the expression of DFT is:

$$X(n) = \sum_{t=0}^{T-1} x(t)w(t)e^{-j\frac{2\pi nt}{T}} \tag{8}$$

where x is a frame of sampled signal and X is the DFT. The magnitude spectrum M of each frame is obtained by:

$$M(n) = |X(n)|, \quad 0 \le n \le T-1 \tag{9}$$

2) MEL-FREQUENCY FILTERING

The main processes of Mel-filtering include mapping the linear frequency magnitude spectrum to Mel-frequency, then filtering by triangle filter bank. Triangle filter bank is a bank of triangle band pass filters. The mapping relation of liner frequency f and Mel-frequency mel is:

$$mel(f) = 2595\log_{10}(1 + \frac{f}{700}) \tag{10}$$

All the filters have the same bandwidth and the cut-off frequency can be expressed as:

$$f_c(j) = f_u(j-1) = f_l(j+1) \tag{11}$$

where f_c, f_u, and f_l represent the center, upper cut off and lower cut off frequency of the jth,$(j-1)$th, and $(j+1)$th filter in the filter bank, respectively. The filter bank B is given by:

$$B(j,n) = \begin{cases} 0 & f_j(n) \le f_c(j-1) \\ \dfrac{f_j(n) - f_c(j-1)}{f_c(j) - f_c(j-1)} & 0 < f_j(n) \le f_c(j) \\ \dfrac{f_j(n) - f_c(j)}{f_c(j) - f_c(j+1)} & f_c(j) < f_j(n) \le f_c(j+1) \\ 0 & f_c(j+1) < f_j(n) \end{cases} \tag{12}$$

where $B(j,n)$ is the nth value of jth filter in the filter bank and $f_j(n)$ is the corresponding Mel-frequency.

The outputs of the filter bank constitute the vector of Mel magnitude spectrum MS:

$$MS(j) = \sum_{n=0}^{T-1} B(j,n)M(n) \tag{13}$$

3) LOGARITHMIC TRANSFORMATION AND DISCRETE COSINE TRANSFORMATION (DCT)

Logarithm MS to the base e to get the logarithmic magnitude s:

$$s(j) = \ln MS(j) \tag{14}$$

At last, do DCT to s to get the MFCC feature vector C:

$$C(n) = \sqrt{\frac{2}{N}}\sum_{j=1}^{J} s(j)cos(\frac{\pi n(2j-1)}{2J}), \quad n = 1, 2, \ldots, N \tag{15}$$

where J is the number of filters in the filter bank, N is the dimension of MFCC feature vector.

4) DELTA CEPSTRUM OF MFCC

Delta cepstrum reflects the variations of MFCC, it shows the dynamic characteristics. The delta cepstrum of a N dimension MFCC feature is:

$$d(n) = \begin{cases} C(n+1) - C(n) & n < K \\ \dfrac{\sum_{k=1}^{K} k[C(n+k) - C(n-k)]}{\sqrt{2\sum_{k=1}^{K} k^2}} & others \\ C(n) - C(n-1) & n \ge N - K \end{cases} \tag{16}$$

where $d(n)$ is the nth value of the delta cepstrum, K is the time difference, normally be 1 or 2.

Based on these steps, the MFCC feature C and the delta cepstrum feature d can be extracted.

V. FEATURE ARRANGEMENT AND CNN ARCHITECTURE

In this paper, in consideration of the possible connection between signal frames and the combination of multiple frame feature can reduce the influence of outliers, feature vectors of frames are arranged into 2D feature matrixes, feature vectors of several contiguous frames constitute the representation of the currant signal. CNNs have been applied in fields of visions and image processing successfully, one major reason is that CNNs can extract distinguishable features automatically and classify accurately. On account of the strong classification ability and mature application of CNNs in the fields of 2D feature maps, CNNs can be considered as a deeper feature extractor of the original cepstrum feature and bearing fault classifier in this work. The keys to apply CNNs to bearing fault classification are: The way of input 1D data into network; Sufficient training samples; Network structure.

A. FEATURE ARRANGEMENT

This paper presents a novel feature arrangement method to generate feature matrixes using 1D feature. It is worth noting that the number of feature vectors may not enough for training a CNN, so data augmentation is required. The proposed feature arrangement method fulfils the two tasks in the same time.

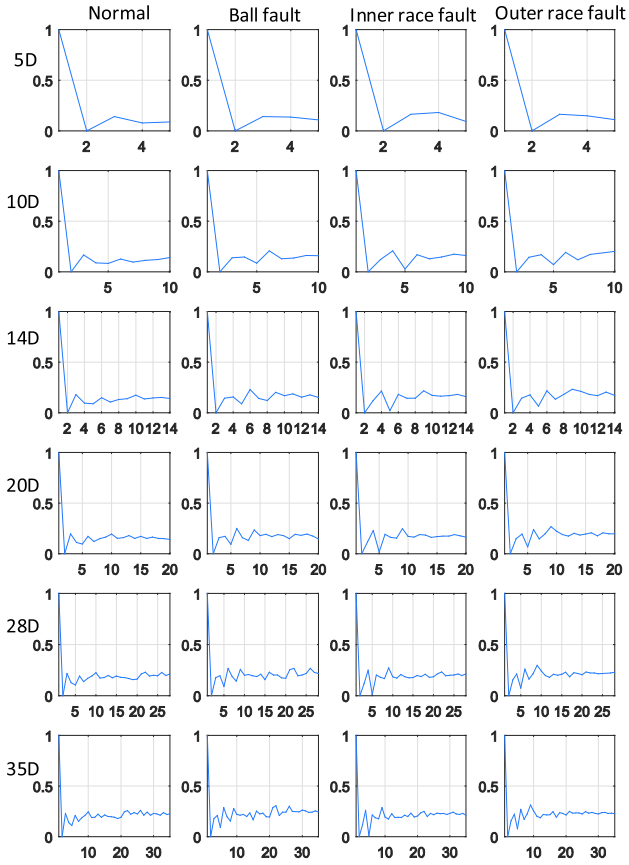

FIGURE 4. Feature vectors of normal, ball, inner race and outer race fault with different dimensions.

To form a feature matrix, first the size of the matrix should be chose properly. For an appropriate feature dimension, the feature vector should accurately represent the bearing fault types generally, and keep lower size to avoid computational burden. The feature vectors of normal, ball, inner race and outer race fault with different dimensions are shown in Fig. 4. It can be seen that the feature vectors getting clearer along with the increase of feature dimensions. Feature vectors of different fault types are similar before 20D, and there is no obvious change when the dimension up to 35. Therefore, 28 is selected as the size of feature vector in this work. Meanwhile, too many or less feature vectors will cause delay or inaccuracy in the real case, and square matrix simplify the calculation for the following steps. Take all into account, the dimension of the feature matrix is 28×28, namely 28 feature vectors(1×28 dimensional) constitute a feature matrix in this paper. After the size is confirmed, pick feature vectors from the feature set randomly and arrange them into the matrix by the order they were picked out. At last, repeat the steps till the number of matrixes is sufficient.

For the general case, the method of feature arrangement can be expressed as: to generate a feature matrix P composed by M MFCC vectors ($1 \times N$ dimensional feature), and the corresponding feature set F which has Z vectors, the basic steps are:

1) Generate an array R consist of a random permutation of the integers from 1 to Z

$$R = randperm(1 : Z) \tag{17}$$

2) Take the first M elements of R as a subset S

$$S(n) = R(n), \quad n = 1, 2, \ldots, M \tag{18}$$

3) Take the $S(n)$th feature vector from the set F and arrange it into nth row of the feature matrix P

$$P(n, :) = F[S(n)], \quad n = 1, 2, \ldots, M \tag{19}$$

4) Repeat 1)-3) until feature matrixes are sufficient for training the network.

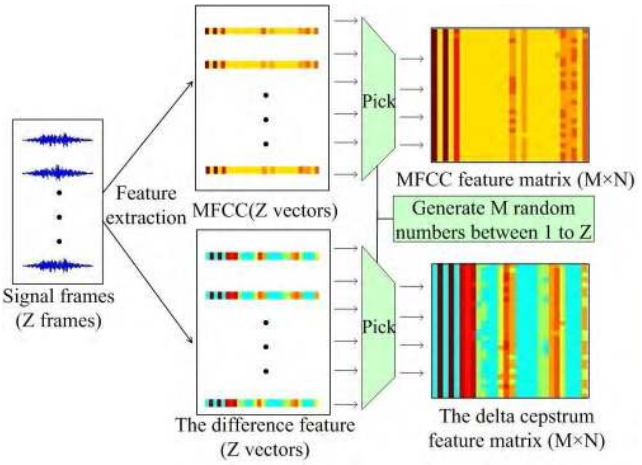

FIGURE 5. The generating process of feature representation matrixes.

Meanwhile, delta cepstrum vectors should be arranged as the same order as the corresponding MFCC feature vectors. The process is depicted as in Fig. 5. To improve the visual effects, the feature vectors and matrixes were expressed by color block from dark blue to bright red. Feature vectors are normalized to 0 to 1 according to the Max-min normalization rule before arrangement operation.

In summary, 28D MFCC feature of filtered data frames are extracted firstly. Then after feature vector sets are built, feature maps are generated by the proposed feature arrangement method, every feature map contains of 28 different feature vectors which are selected by random sampling without replacement. At last, the feature map set is established by repeating the process of Feature arrangement.

B. NETWORK ARCHITECTURE

A fundamental architecture of CNNs is composed of an input layer, hidden layers, which include convolutional and pooling layers, and a fully connected output layer to fulfill the need of classification. To evaluate these architectures of CNNs, normal and 3 types of bearing fault data added with white Gaussian noise at low SNR level, -10dB, are applied to verify several architectures of CNNs, the classification result is shown in Table.1. -10dB SNR level is applied in this

TABLE 1. Classification result of various networks at −10dB SNR.

Networks	C1 (Output maps, Kernel size)	P1 (Pooling size)	C2 (Output maps, Kernel size)	P2 (Pooling size)	Accuracy(%)
1	(1,3)	2× 2	×	×	49.35
2	(2,3)	2× 2	×	×	51.68
3	(2,3)	2× 2	(1,2)	2× 2	49.89
4	(2,3)	2× 2	(2,2)	2× 2	54.88
5	(2,3)	2× 2	(3,2)	2× 2	48.59
6	(2,5)	2× 2	(3,3)	2× 2	59.50
7	(2,5)	2× 2	(3,5)	2× 2	67.75
8	(2,5)	2× 2	(3,7)	2× 2	53.32
LeNet-5 [10]	(6,7)	2× 2	(16,4)	2× 2	67.49

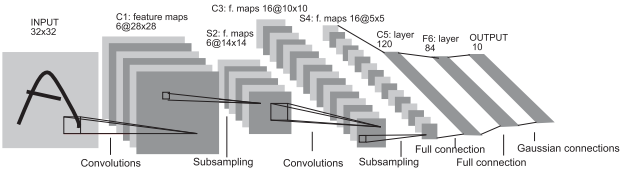

FIGURE 6. Architecture of LeNet-5 [10].

part as a representation of low SNR level condition. Under the condition of higher SNR levels, satisfying fault diagnosis results can be got by applying the mentioned different architectures, it is not a difficult task to classify fault data at higher SNR levels. Meanwhile, too low SNR level condition will lose the practical meaning of fault diagnosis. As a result, −10dB SNR condition is applied for evaluating the different architectures of CNNs in this part. In this part, 8000 training sample, 100 batch size and 150 epochs are applied. In the table, C1 and C2 represent the first and second convolutional layer, P1 and P2 represent the first and second pooling layer. The results of applying different numbers of output maps and convolutional kernel size are included. Meanwhile, a representative existing pre-trained CNN model LeNet-5 [10] is included in the table. The information of convolutional and pooling layers of LeNet-5 are listed in the table, the architecture of LeNet-5 is shown in Fig. 6. On account of the size of input maps, the kernel size of the second convolutional layers of LeNet-5 is set to 4×4.

In the initialization phase, all the bias are initialized to 0 and the weights ω are set as:

$$\omega = random(-1, 1)\sqrt{\frac{6}{Inputpoints + Outputpoints}} \tag{20}$$

Sigmoid function is applied behind each convolutional layer and in the error back propagation training phase, and mean squared error (MSE) is used as the loss function. The table shows that the network No.7 achieves the best result, and it is more accurate than LeNet-5 in this case. Meanwhile, the presented architecture has fewer output maps in C1 and C2 which means fewer parameters need to be trained. According to the result, the convolutional neural network in

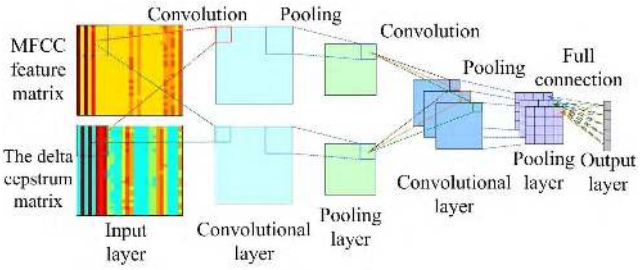

FIGURE 7. The network architecture of the proposed method.

this paper contains a input layer, two convolutional layers, two pooling layers, and a fully connected output layer. The network architecture is depicted in Fig. 7. To improve the performance of feature expression, a two-channel-input layer is applied, one for MFCC feature input and another for delta cepstrum. The detailed comparison is shown in the experiment section.

In [15], Piczak evaluates the potential of CNNs in classifying environmental sound and urban recordings. Unlike the model [15] and LeNet-5, this paper uses less convolution kernels due to bearing fault signals have relatively stable frequency components in a short frame compare to environmental sounds and pictures, which means less parameters need to be trained in the training phase.

VI. EXPERIMENT

In this section, comparison experiments of bearing fault classification of the proposed method with several comparison methods are provided. In experiments, the results of fault classification with white Gaussian noise and real factory noise are presented. Firstly, the experimental bearing data and experimental conditions are introduced. Then, experiments under real factory noise have been presented to verify the effect of SK-based filtering and input features. In comparative experiments, the results of the proposed method and methods in [25]–[32] classifying bearing data with Gaussian white noise at −10dB signal to noise ratio(SNR) are shown first, then comparison with the method in [25] is given. At last, summarized classification results of the proposed method on bearing data with real factory noise are listed. All the bearing data in this section come from Case Western Reserve University (CWRU) bearing data center [33].

A. EXPERIMENTAL SETUP

The CWRU bearing dataset has been widely used for testing fault diagnosis algorithm and new features, and became a standard database [34], [35]. Specification of the experimental bearings in the database is shown in Table 2.

The dataset contains two kinds of fault bearings, drive end bearing fault and fan end bearing fault. Data were collected from the base plate (BA), the vertical direction on the housing of the drive end (DE) and fan end bearing (FE) respectively, with the motor load of 0hp (_0), 1hp (_1), 2hp (_2), and 3hp (_3). The bearing fault mode include no damage, damage

TABLE 2. Specification of the experimental bearings.

Bearing	Drive end bearing	Fan end bearing
Specification (inches)	6205-2RS JEM SKF	6203-2RS JEM SKF
Inside Diameter	0.9843	0.6693
Outside Diameter	2.0472	1.5748
Thickness	0.5906	0.4724
Ball Diameter	0.3126	0.2656
Pitch Diameter	1.537	1.122
Defect frequencies: (multiple of running speed in Hz)		
Inner Ring	5.4152	4.9469
Outer Ring	3.5848	3.0530
Cage Train	0.39828	0.3817
Rolling Element	4.7135	3.9874

on inner race (IR), outer race (OR), and ball (B) with fault diameters of 0.007in (007), 0.014in (014), and 0.021in (021). For outer race fault, the number after '@' means the fault position in a clock.

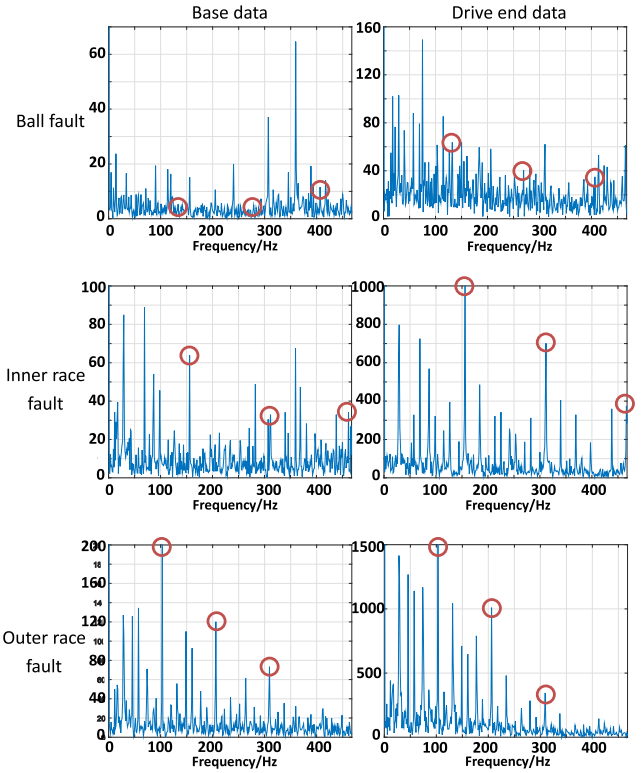

FIGURE 8. The envelop spectrum of drive end fault data.

The envelop spectra of drive end fault data from the equipment base and drive end with 3hp load (rotate speed: 1730rpm) are shown in Fig. 8. The fault character frequencies of each fault type are marked by red circles in the figures. The results show that the fault character frequencies of the base and ball fault data are blur, and the fault type is difficult to be identified by envelop spectrum. Without enough prior knowledge, the envelop spectrum analysis is not reliable for all the fault type in this case.

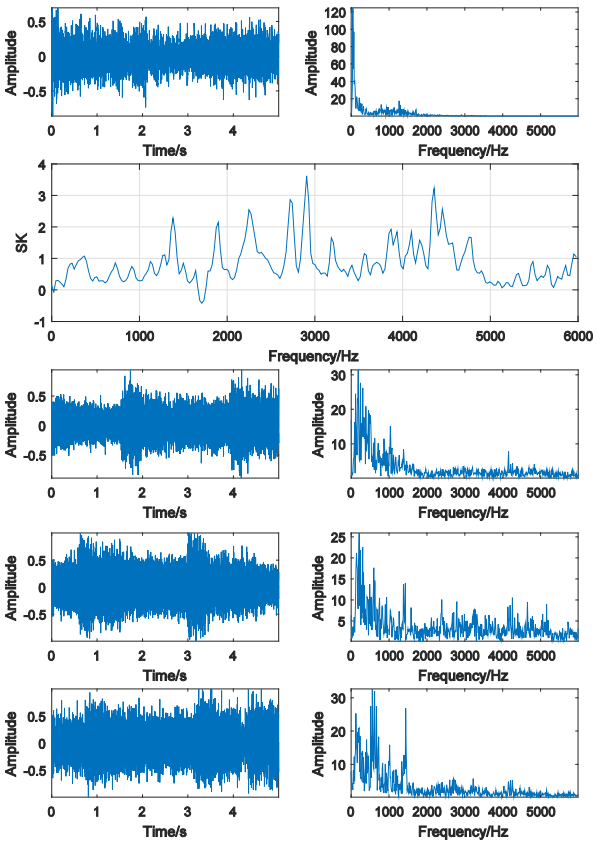

FIGURE 9. (a) The temporal waveform and spectrum of the factory noise. (b) The SK curve of the factory noise. The temporal waveform and spectrum of (c) BA-B021_3, (d) BA-IR021_3, (e) BA-OR021_3 with the factory noise.

To verify the proposed approach, the test dataset used for experiments in this section are the CWRU bearing dataset added with a audible factory noise signal and white Gaussian noise respectively. The factory noise was download from Freesound.org [36], the noise contains continuous and intermittent machine noise which can simulate the real case. The temporal waveform and spectrum of the factory noise, the bearing data, BA-B021_3,BA-IR021_3, BA-OR021_3, with the factory noise, and the SK curve of the factory noise are shown in Fig. 9. From the temporal waveform of the factory noise Fig. 9(a), the shock can be identified clearly, and the SK curve Fig. 9(b) shows that the noise includes several main Gaussian and non-Gaussian components. Moreover, the signal contains weak noises on the entire frequency band as it shown in the spectral waveform. Therefore, the noise signal is regarded as a representative factory noise to verify the proposed method.

In filtering process, the window width (Nw) is the integral power of 2 normally, so let the window be 256 points Hanning window here. Let the points of STFT be $2 \times Nw$, and the step length of window is $0.75 \times Nw$. For framing, to ensure every frame contains the main SK information, let the frame length be 6000 points, the step length of framing window is 0.1 times of the width of the framing window. Let the dimension of a MFCC feature vector be 28 and a feature matrix composed by 28 feature vectors.

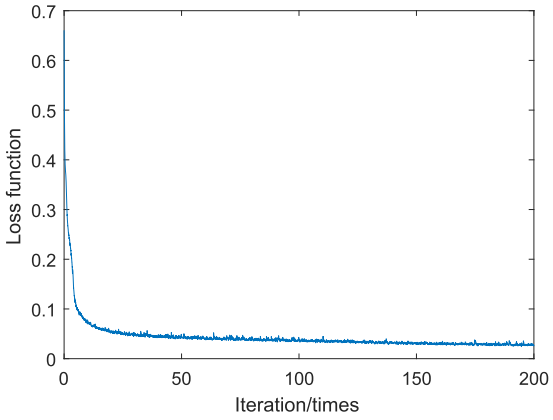

FIGURE 10. Loss function of the proposed method.

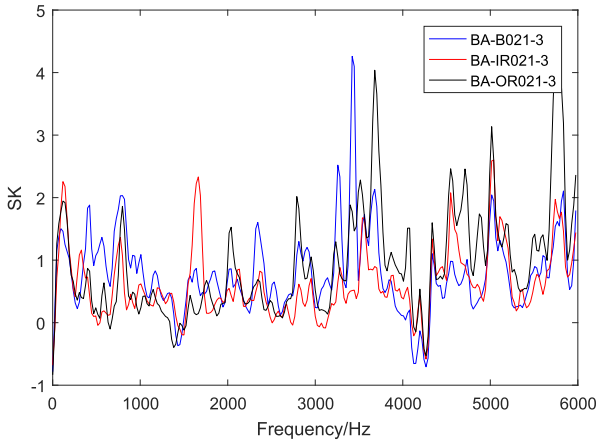

FIGURE 11. SK curves of bearing data with the factory noise.

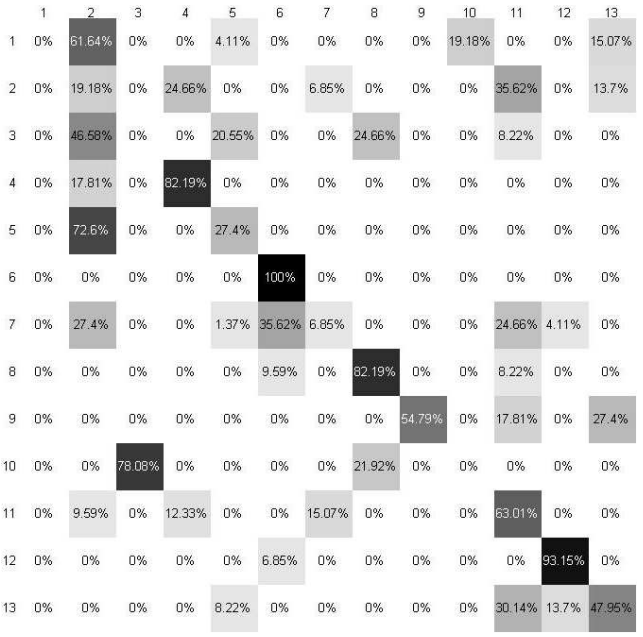

FIGURE 12. Confusion matrix of classifying noised signals.

In training process, sufficient samples should be provided to train the network. In this work, 8000 feature maps were produced for each class according to the length of the original data. In the testing phase, to simulate the real detecting condition, feature vectors should be arranged in time sequence. And 10-fold cross validation is applied in the experiments. The batch size is set to 100. The loss function of the training phase in 10-fold cross validation, MSE, is shown in Fig. 10. In the process of 10-fold cross validation, the proportion of training and validation data samples used for training is 9 to 1 in each time validation. According to the result, 150 times iteration is applied. Since the motor speeds are obviously different under the condition of different motor load, data collected from the same position under one kind of motor load are used in each experiment.

The presented architecture of CNNs is shown in Fig. 7, the Deep-learn Toolbox [37] was applied as the basis to build the CNN in this paper. The first convolutional layer has two convolution kernels, each size is 5×5. The second convolutional layer has 3 convolution kernels with the same size. Mean sampling is adopted in every pooling layer, the size of the sampling kernel is 2×2.

B. RESULTS

1) SK-BASED FILTERING

This experiment tends to verify the effect of filtering based on SK. The testing data are 0 motor load drive end bearing fault BA data with the factory noise, a total of 13 classes fault data. The outer race defect signals with different fault locations may relate to different fault causes and equipment operating conditions in the real case. As a result, the bearing outer race defect signals are used as different types of bearing fault data in this work. Meanwhile, in order to highlight the effect of SK-based filtering and in consideration of delta cepstrum is the derivative of MFCC, only MFCC was used in this part, in other word, the CNN has a one-channel-input layer. The SK curves of base plate data, BA-B021_3, BA-IR021_3, and BA-OR021_3 with the factory noise are shown in Fig. 11.

The confusion matrixes of classifying the noised and filtered data using one-channel-input CNN of each class are shown in Fig. 12 and Fig. 13. The overall accuracy of classifying noised and filtered data are 44.36% and 78.39%, respectively. The result shows that the filtering method increases the classification accuracy.

In the confusion matrixes, label 1 to 13 stand for 13 fault types: BA-B007_0, BA-B014_0, BA-B021_0, BA-IR007_0, BA-IR014_0, BA-IR021_0, BA-OR007@12_0, BA-OR007@3_0, BA-OR007@6_0, BA-OR014@6_0, BA-OR021@12_0, BA-OR021@3_0, and BA-OR021@6_0.

2) COMPARE TWO WITH ONE-CHANNEL-INPUT

To compare with two-channel-input, the result of applying delta cepstrum feature as the single input to one-channel-input CNN is shown in Fig. 14. The overall accuracy is 28.5605%. The delta cepstrum represent the dynamic characteristic of MFCC, without MFCC, delta cepstrum is not very

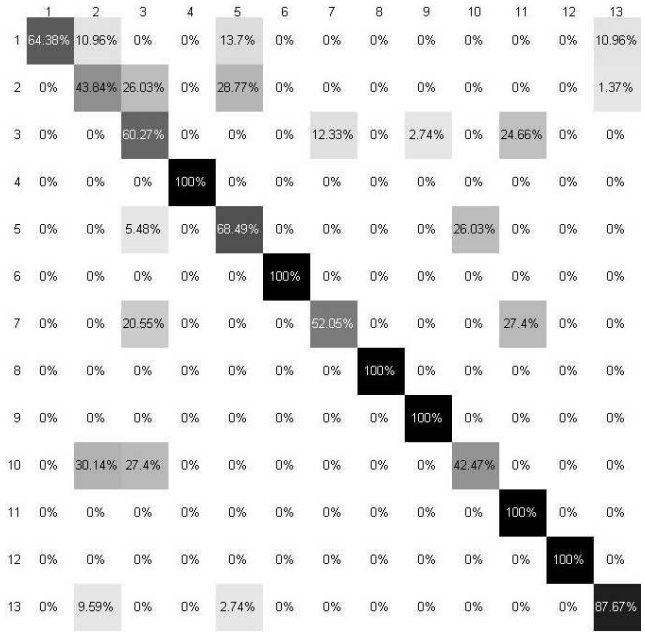

FIGURE 13. Confusion matrix of classifying filtered signals.

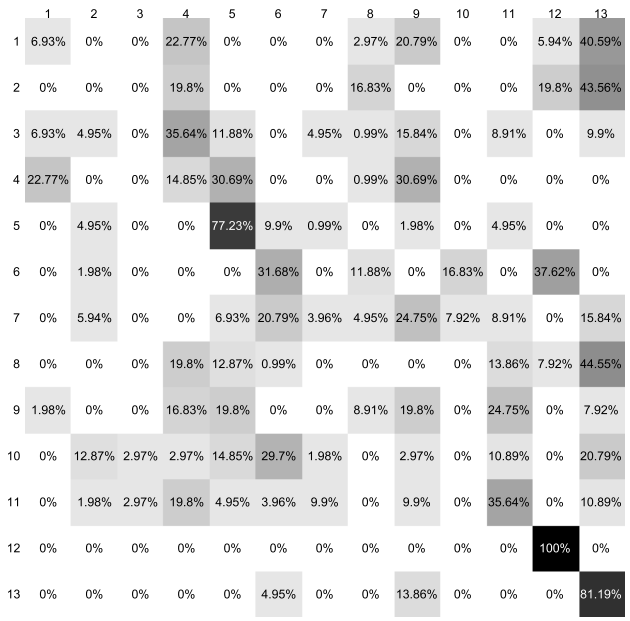

FIGURE 14. Confusion matrix of delta cepstrum feature with one-channel-input CNN.

effective for classification according to the result. In contrast, MFCC feature and delta cepstrum feature are applied as the inputs of a two-channel-input CNN, and the confusion matrix of classifying each class are shown in Fig. 15, the overall accuracy is 86.0906%. Compare to the result of applying MFCC and delta cepstrum feature with one-channel-input respectively, two-channel-input CNN increases the classification accuracy. And basically, most of the data can be classified correctly. The fault type 0.007in, 0.021in ball

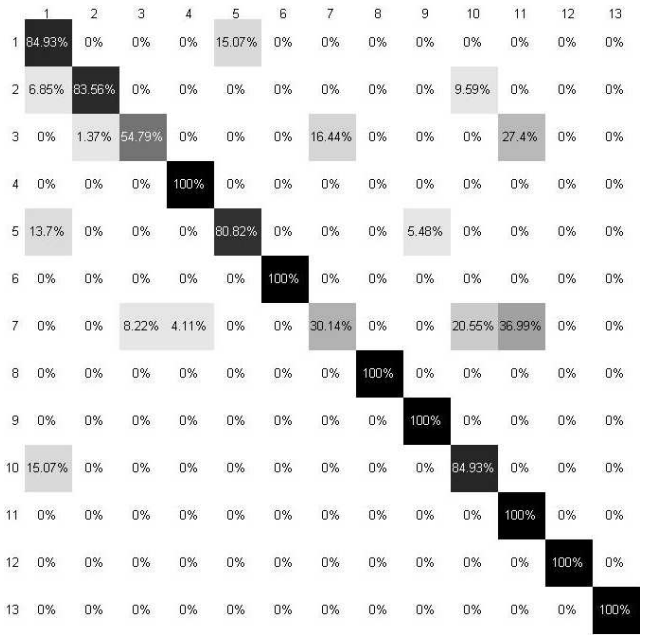

FIGURE 15. Confusion matrix of applying two-channel-input CNN.

fault and 0.014in outer race fault perform poorly by applying no-filtering one-input-channel CNN, and get accurate results with the SK-filtering two-input-channel CNN. The accuracy of 0.007in outer race fault at the 12 o'clock position decreased compare to applying filtering one-channel-input CNN, and it is higher than applying the no-filtering approach. Since the two-channel-input CNN get better results generally, the architecture of the proposed approach is applied in the rest experiments.

3) COMPARISON OF THE PROPOSED METHOD WITH PREVIOUS METHODS

In this part, white Gaussian noise was added in bearing data to evaluate the classification ability of the proposed method and several previous methods [25]–[32]. The experimental data are the normal, B007, IR007, and OR007 data of 2hp load drive end bearing fault DE data. In [25] and [32], the previous methods have been tested all sidedly, and the conclusion indicates that the worst case is with −10dB SNR. Table 3 presents the comparison of the proposed method with the previous methods at −10dB SNR, and the confusion matrix of the proposed method is shown in Fig. 16. In the confusion matrix, number 1 to 4 represent the normal, ball fault, inner race fault, and outer race fault data respectively. The result shows that the better fault classification ability of the proposed method, SK-based CNN, at −10dB SNR level, all the normal samples are classified accurately and few fault samples are misclassified.

In addition, the strong white Gaussian noise changes the SK curve of the bearing data greatly. For instance, the SK curve, average error, and mean of the SK curve of DE-IR007_2 with white Gaussian noise at several SNR levels are

TABLE 3. Comparative performance of the proposed method with previous methods at SNR=−10dB.

Method	Accuracy(%)
S.K.Goumas [26]	33.33
X.Lou [27]	33.33
S.Seker [28]	78.25
F.Li [29]	86.00
B.Samanta [30]	33.33
A.Malhi [31]	38.09
AFHCW [32]	91.23
SK-based CNN	98.76

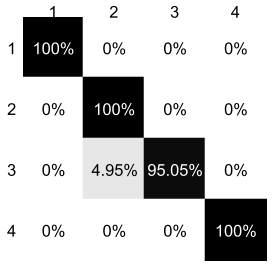

FIGURE 16. Confusion matrix of the proposed method at −10dB SNR.

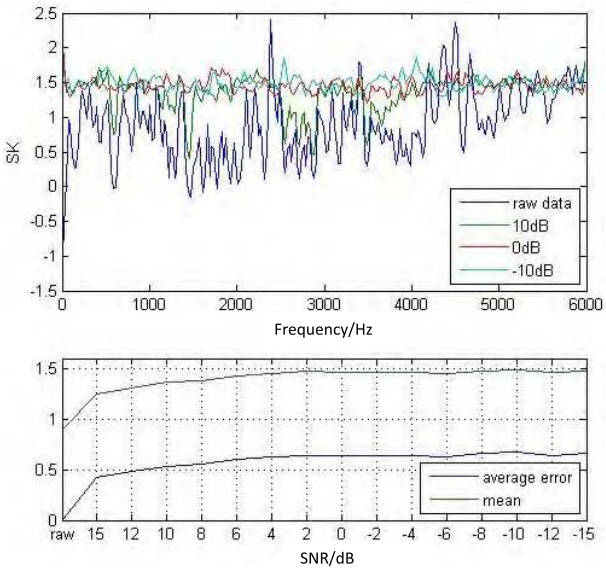

FIGURE 17. The SK curve and the waveforms of the average error and mean of SK of DE-IR007_2 with white Gaussian noise.

shown in Fig. 17. The SK curve of the raw signal and at 10dB level are similar, but changing obviously when the SNR level is lower than 0dB, which indicates that the effectiveness of the proposed approach will decline along with the reduction of SNR.

4) COMPARISON OF THE PROPOSED METHOD WITH VSI-BASED ANN

In [25], Muhammad Amar presents a novel vibration spectrum imaging (VSI) feature enhancement procedure for low SNR conditions, and used an artificial neural network (ANN) as a fault classifier (VSI-based ANN). In the paper, VSI-based ANN method was used to classify the CWRU bearing data that added with white Gaussian noise in various low SNRs. Comparing to other methods mentioned in [25], VSI-based ANN method got better performance in classifying Gaussian white noised bearing data. In this part, comparison of SK-based CNN with VSI-based ANN are shown.

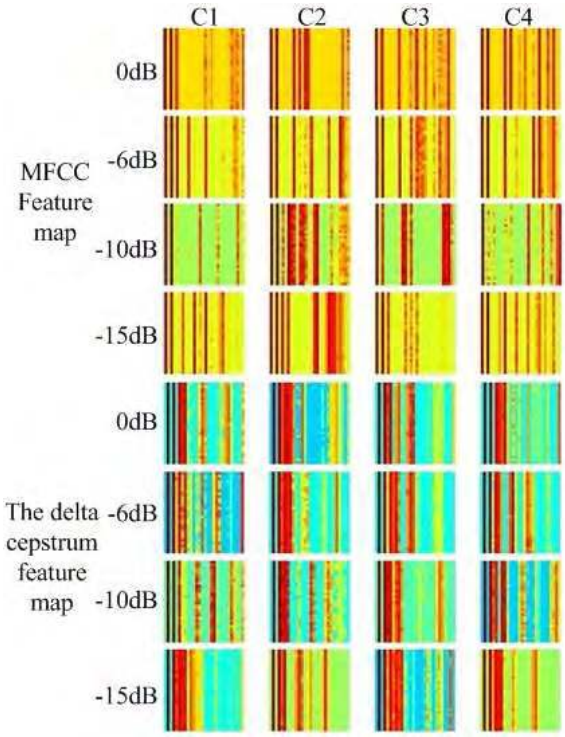

FIGURE 18. The input maps correspond to SNR levels of the CNN.

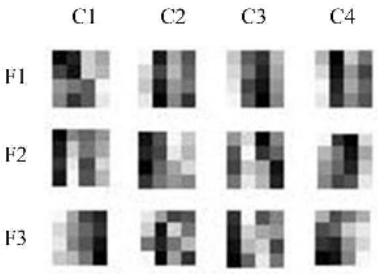

FIGURE 19. The feature maps of −10dB noised data extracted by the CNN.

In the experiment, the bearing data of four classes were added with 0dB to −15dB SNR white Gaussian noise respectively. The corresponding input maps of different SNR levels are shown in Fig. 18. And the feature maps of −10dB noised data extracted by the CNN before the last layer are shown in Fig. 19. The comparison of VSI-based ANN [25] and the proposed approach on low SNR levels are shown in Fig. 20.

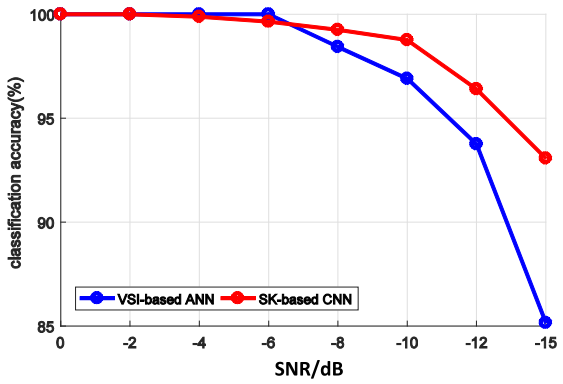

FIGURE 20. The classification accuracy of VSI-based ANN and SK-based CNN.

In Fig. 18 and Fig. 19, C1 to C4 represent the 2hp load drive end bearing DE data of normal, ball fault, inner race fault, and out race fault of 0.007in, F1 to F3 represent the three outputs in the last convolutional layer of the CNN, which can be regarded as the feature extracted by the CNN. As it shown in Fig. 18, there is a huge difference between input maps of same fault type at different SNR level, but several main frequency bands (columns) have been retained, especially for the delta cepstrum feature maps. Meanwhile, input maps of different classes at one SNR level are distinguishable. In Fig. 19, the differences between fault types are even greater among the feature maps extracted by CNN.

According to the results shown in Fig. 20, the VSI-based ANN gets high classification accuracy from 0dB to −6dB, and declined gradually from −6dB to −15dB. The classification ability of the proposed method is weaker than VSI at −4dB to −6dB, but gets better results at lower SNR levels, from −8dB to −15dB, and the downward trend is more slower. The result verifying the Gaussian noise resisting and classifying ability of the proposed method.

5) THE DRIVE AND FAN END BEARING FAULT CLASSIFICATION

In this part, the drive end and fan end bearing fault data with real factory noise are classified by the proposed method, respectively. By contrast, the classification results of one-channel-input CNN without filtering are presented. The classification results are shown in Table 4-Table 7. Table 4 and Table 5 show the results of drive end bearing fault classification, the results of fan end bearing fault classification are shown in Table 6 and Table 7.

TABLE 4. Classification accuracy (%) of no filtering/one-channel-input (drive end fault).

Data recorded	Motor load(hp)			
	0	1	2	3
BA	44.36	36.87	26.87	49.63
DE	89.64	88.52	79.28	71.66
FE	72.21	68.69	81.51	63.50

TABLE 5. Classification accuracy (%) of SK-based filtering/two-channel-input (drive end fault).

Data recorded	Motor load(hp)			
	0	1	2	3
BA	86.10	83.88	77.87	84.83
DE	99.14	99.91	92.89	85.53
FE	95.89	91.26	98.60	93.25

TABLE 6. Classification accuracy (%) of no filtering /one-channel-input (fan end fault).

Data recorded	Motor load(hp)			
	0	1	2	3
BA	26.59	29.76	43.34	37.98
DE	76.92	69.41	78.54	59.36
FE	75.76	90.18	77.16	92.35

TABLE 7. Classification accuracy (%) of SK-based filtering/two-channel-input (fan end fault).

Data recorded	Motor load(hp)			
	0	1	2	3
BA	65.53	63.89	81.07	61.02
DE	96.52	89.27	99.99	85.96
FE	93.05	95.21	98.29	99.88

From the experimental results, the proposed method is more effective than the contrast method. Meanwhile, most of the results of classifying FE data are better than others, which matches the conclusion in paper [2] that the FE data are easier to be diagnosed than other data. Similarly, the results of classifying BA data perform poorly than other locations.

Based on the results above, whether under the influence of Gaussian white noise or real factory noise, the proposed method can classify bearing fault data effectively.

VII. CONCLUSION

This paper proposed a novel and effective bearing fault classification method by combining SK-based filtering method and convolutional neural network. In the experiments, data with real factory noise and white Gaussian noise at low SNR levels have been taking into account. From the experiment results delivered above, whether under the influence of low SNR white Gaussian noise or real factory noise, the proposed method can get high classification accuracies. Based on the accurate classification results, the proposed method can improve the quality of monitoring rolling element bearings in noise environment. Based on the findings in this work, the future direction of the research work includes multi-faults diagnosis and bearing fault anomaly detection in strong noise environment, etc.

ACKNOWLEDGMENT

The authors would like to thank the Case Western Reserve University Bearing Data Center for providing the bearing data in this study.

REFERENCES

[1] X. Jin, M. Zhao, T. W. S. Chow, and M. Pecht, "Motor bearing fault diagnosis using trace ratio linear discriminant analysis," *IEEE Trans. Ind. Electron.*, vol. 61, no. 5, pp. 2441–2451, May 2014. doi: 10.1109/TIE.2013.2273471.

[2] W. A. Smith and R. B. Randall, "Rolling element bearing diagnostics using the Case Western Reserve University data: A benchmark study," *Mech. Syst. Signal Process.*, vols. 64–65, pp. 100–131, Dec. 2015. doi: 10.1016/j.ymssp.2015.04.021.

[3] B. Yao, P. Zhen, L. Wu, and Y. Guan, "Rolling element bearing fault diagnosis using improved manifold learning," *IEEE Access*, vol. 5, pp. 6027–6035, 2017. doi: 10.1109/ACCESS.2017.2693379.

[4] R. Zhang, H. Tao, L. Wu, and Y. Guan, "Transfer learning with neural networks for bearing fault diagnosis in changing working conditions," *IEEE Access*, vol. 5, pp. 14347–14357, 2017. doi: 10.1109/ACCESS.2017.2720965.

[5] S. Singh and N. Kumar, "Detection of bearing faults in mechanical systems using stator current monitoring," *IEEE. Trans. Ind. Informat.*, vol. 13, no. 3, pp. 1341–1349, Jun. 2017. doi: 10.1109/TII.2016.2641470.

[6] R. F. Dwyer, "A technique for improving detection and estimation of signals contaminated by under ice noise," *J. Acoust. Soc. Amer.*, vol. 74, no. 1, pp. 124–130, Mar. 1983. doi: 10.1016/j.ymssp.2015.04.021.

[7] J. Antoni, "The spectral kurtosis: A useful tool for characterising non-stationary signals," *Mech. Syst. Signal Process.*, vol. 20, no. 2, pp. 282–307, Feb. 2006. doi: 10.1016/j.ymssp.2004.09.001.

[8] J. Antoni and R. Randall, "The spectral kurtosis: Application to the vibratory surveillance and diagnostics of rotating machines," *Mech. Syst. Signal Process.*, vol. 20, no. 2, pp. 308–331, 2006. doi: 10.1016/j.ymssp.2004.09.002.

[9] J. Tian, C. Morillo, M. H. Azarian, and M. Pecht, "Motor bearing fault detection using spectral kurtosis-based feature extraction coupled with K-nearest neighbor distance analysis," *IEEE Trans. Ind. Electron.*, vol. 63, no. 3, pp. 1793–1803, Mar. 2016. doi: 10.1109/TIE.2015.2509913.

[10] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proc. IEEE*, vol. 86, no. 11, pp. 2278–2324, Nov. 1998. doi: 10.1109/5.726791.

[11] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Proc. Adv. Neural Inf. Process. Syst. (NIPS)*, vol. 2012, pp. 1097–1105.

[12] J. C. Rangel, J. Martínez-Gómez, C. Romero-González, I. García-Varea, and M. Cazorla, "Semi-supervised 3D object recognition through CNN labeling," *Appl. Soft Comput.*, vol. 65, no. 3, pp. 603–613, 2018. doi: 10.1016/j.asoc.2018.02.005.

[13] Q. Mao, M. Dong, Z. Huang, and Y. Zhan, "Learning salient features for speech emotion recognition using convolutional neural networks," *IEEE Trans. Multimedia*, vol. 16, no. 8, pp. 2203–2213, Dec. 2014. doi: 10.1109/TMM.2014.2360798.

[14] T. Ince, S. Kiranyaz, L. Eren, M. Askar, and M. Gabbouj, "Real-time motor fault detection by 1-D convolutional neural networks," *IEEE Trans. Ind. Electron.*, vol. 63, no. 11, pp. 7067–7075, Nov. 2016. doi: 10.1109/TIE.2016.2582729.

[15] K. J. Piczak, "Environmental sound classification with convolutional neural networks," in *Proc. IEEE 25th Int. Workshop Mach. Learn. Signal Process.*, Sep. 2015, pp. 1–6. doi: 10.1109/MLSP.2015.7324337.

[16] T. Yoshioka, S. Karita, and T. Nakatani, "Far-field speech recognition using CNN-DNN-HMM with convolution in time," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, Apr. 2015, pp. 4360–4364. doi: 10.1109/ICASSP.2015.7178794.

[17] J. Schlüter and S. Böck, "Improved musical onset detection with convolutional neural networks," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, May 2014, pp. 6979–6983. doi: 10.1109/ICASSP.2014.6854953.

[18] L. Wen, L. Gao, X. Li, M. Xie, and G. Li, "A new data-driven intelligent fault diagnosis by using convolutional neural network," in *Proc. IEEE Int. Conf. Ind. Eng. Eng. Manage. (IEEM)*, Singapore, Dec. 2017, pp. 813–817.

[19] D. Peng, Z. Liu, H. Wang, Y. Qin, and L. Jia, "A novel deeper one-dimensional CNN with residual learning for fault diagnosis of wheelset bearings in high-speed trains," *IEEE Access*, vol. 7, pp. 10278–10293, 2019. doi: 10.1109/ACCESS.2018.2888842.

[20] H.-M. Moon, C. H. Seo, and S. B. Pan, "A face recognition system based on convolution neural network using multiple distance face," *Soft Comput.*, vol. 21, no. 17, pp. 4995–5002, 2017. doi: 10.1007/s00500-016-2095-0.

[21] V. C. M. N. Leite, J. G. B. da Silva, G. F. C. Veloso, L. E. B. da Silva, G. Lambert-Torres, E. L. Bonaldi, and L. E. L. de Oliveira, "Detection of localized bearing faults in induction machines by spectral kurtosis and envelope analysis of stator current," *IEEE Trans. Ind. Electron.*, vol. 62, no. 3, pp. 1855–1865, Mar. 2015. doi: 10.1109/ICASSP.2014.6854953.

[22] S. E. Kucukbay and M. Sert, "Audio-based event detection in office live environments using optimized MFCC-SVM approach," in *Proc. IEEE 9th Int. Conf. Semantic Comput.*, Feb. 2015, pp. 6979–6983. doi: 10.1109/ICOSC.2015.7050855.

[23] D. Sharma and I. Ali, "A modified MFCC feature extraction technique for robust speaker recognition," in *Proc. Int. Conf. Adv. Comput., Commun. Inform.*, Aug. 2015, pp. 1052–1057. doi: 10.1109/ICACCI.2015.7275749.

[24] B. Yan, G. Q. Qian, F. H. Wang, and S. Chen, "Noise recognition of power transformers based on improved MFCC and VQ," in *Proc. IEEE/PES Transmiss. Distrib. Conf. Expo.*, May 2016, pp. 1–5. doi: 10.1109/TDC.2016.7519923.

[25] M. Amar, I. Gondal, and C. Wilson, "Vibration spectrum imaging: A novel bearing fault classification approach," *IEEE Trans. Ind. Electron.*, vol. 62, no. 1, pp. 494–502, Jan. 2015. doi: 10.1109/TIE.2014.2327555.

[26] S. K. Goumas, M. E. Zervakis, and G. S. Stavrakakis, "Classification of washing machines vibration signals using discrete wavelet analysis for feature extraction," *IEEE Trans. Instrum. Meas.*, vol. 51, no. 3, pp. 497–508, Jun. 2002. doi: 10.1109/TIM.2002.1017721.

[27] X. Lou and K. A. Loparo, "Bearing fault diagnosis based on wavelet transform and fuzzy inference," *Mech. Syst. Signal Process.*, vol. 18, no. 5, pp. 1077–1095, Sep. 2004. doi: 10.1016/S0888-3270(03)00077-3.

[28] S. Seker and E. Ayaz, "Feature extraction related to bearing damage in electric motors by wavelet analysis," *J. Franklin Inst.*, vol. 340, no. 2, pp. 125–134, 2003. doi: 10.1016/S0016-0032(03)00015-2.

[29] F. Li, G. Meng, L. Ye, and P. Chen, "Wavelet transform-based higher-order statistics for fault diagnosis in rolling element bearings," *J. Vib. Control*, vol. 14, no. 11, pp. 1691–1709, 2008. doi: 10.1177/1077546308091214.

[30] B. Samanta and K. R. Al-Balushi, "Artificial neural network based fault diagnostics of rolling element bearings using time-domain features," *Mech. Syst. Signal Process.*, vol. 17, no. 2, pp. 317–328, 2003. doi: 10.1006/mssp.2001.1462.

[31] A. Malhi and R. X. Gao, "PCA-based feature selection scheme for machine defect classification," *IEEE Trans. Instrum. Meas.*, vol. 53, no. 6, pp. 1517–1525, Dec. 2004. doi: 10.1109/TIM.2004.834070.

[32] M. F. Yaqub, I. Gondal, and J. Kamruzzaman, "Inchoate fault detection framework: Adaptive selection of wavelet nodes and cumulant orders," *IEEE Trans. Instrum. Meas.*, vol. 61, no. 3, pp. 685–695, Mar. 2012. doi: 10.1109/TIM.2004.834070.

[33] Case Western Reserve University Bearing Data Center. Accessed: Feb. 20, 2019. *Seeded Fault Test Data*. [Online]. Available: http://csegroups.case.edu/bearingdatacenter/home

[34] M. Segla, S. Wang, and F. Wang, "Bearing fault diagnosis with an improved high frequency resonance technique," in *Proc. IEEE 10th Int. Conf. Ind. Inform.*, Jul. 2012, pp. 580–585. doi: 10.1109/INDIN.2012.6301378.

[35] T. W. Rauber, F. de A. Boldt, and F. M. Varejão, "Heterogeneous feature models and feature selection applied to bearing fault diagnosis," *IEEE Trans. Ind. Electron.*, vol. 62, no. 1, pp. 637–646, Jan. 2015. doi: 10.1109/TIE.2014.2327589.

[36] *Freesound*. Accessed: Feb. 20, 2019. [Online]. Available: https://freesound.org/s/249636/

[37] R. B. Palm, "Prediction as a candidate for learning deep hierarchical models of data," M.S. thesis, Dept. Inform. Math. Model., Tech. Univ. Denmark, Kongens Lyngby, Denmark, 2012.

QINYU JIANG received the B.S. degree from Shandong University, Weihai, China, in 2014. He is currently pursuing the Ph.D. degree with the School of Control Science and Engineering, Shandong University, Jinan, China. His research interests include fault diagnosis and classification, pattern recognition, and intelligent systems.

FALIANG CHANG received the B.S. and M.S. degrees from Shandong Polytechnic University, Jinan, China, in 1986 and 1989, respectively, and the Ph.D. degree in pattern recognition and intelligence systems from Shandong University, Jinan, in 2003.

He has been a Professor of pattern recognition and machine intelligence with the School of Control Science and Engineering, Shandong University, since 2003. His research interests include computer vision, image processing, intelligent transportation systems, and multi-camera tracking methodology.

BOWEN SHENG received the B.S. degree from the Inner Mongolia University of Technology, Hohhot, China, in 2017. He is currently pursuing the M.S. degree with the School of Control Science and Engineering, Shandong University, Jinan, China. His research interests include fault diagnosis and classification, pattern recognition, and intelligent systems.