

Research Article

A Novel Bearing Fault Diagnosis Methodology Based on SVD and One-Dimensional Convolutional Neural Network

Yangyang Wang ¹, Shuzhan Huang ², Juying Dai ³ and Jian Tang ³

¹Xichang Satellite Launch Center, Xichang 615000, China

²School of Graduate, Army Engineering University of PLA, Nanjing 210000, China

³School of Field Engineering, Army Engineering University of PLA, Nanjing 21000, China

Correspondence should be addressed to Shuzhan Huang; hsz_221618@163.com

Received 12 June 2019; Revised 16 October 2019; Accepted 12 December 2019; Published 30 January 2020

Academic Editor: Salvatore Russo

Copyright © 2020 Yangyang Wang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper constructs a novel network structure (SVD-1DCNN) based on singular value decomposition (SVD) and one-dimensional convolutional neural network (1DCNN), which takes the original signal as input to realize intelligent diagnosis of bearing faults. The output of the first convolution layer was also analyzed from the perspectives of time domain and time-frequency domain in the simulation experiment. Through qualitative analysis and quantitative analysis, it was found that the convolution kernel not only extracted the classification features of signals but also gradually highlighted the learned features in the network training process. Moreover, applying this network in fault diagnosis of bearing data provided by the Case Western Reserve University (CWRU) Bearing Data Center, it was found that the convolution kernel could also achieve the above operation. The novel network of this paper achieved a good classification effect on both the simulated signals and the measured signals.

1. Introduction

A small fault in a mechanical device often affects the stability and safety of the entire system and can even lead to catastrophic consequences [1]. As a key component of mechanical equipment, bearings are widely used in various types of machinery. Failure of the bearing can cause many serious mechanical failures, so the safe and smooth operation of the bearing is critical to the mechanical equipment. Timely detection, positioning, and troubleshooting of bearing faults can effectively improve the safety of industrial production. Therefore, it is of great significance to study the fault diagnosis of bearing.

The traditional fault diagnosis process generally consists of three steps: data acquisition, feature extraction and selection, and fault pattern recognition [2].

The collected data include vibration signal, acoustic signal, and temperature signal, and since the vibration signal can directly characterize the state of the mechanical equipment, the vibration signal is most commonly collected in fault diagnosis [3]. In the fault diagnosis of bearing,

commonly used signal processing methods include Fourier transform (FT) [4], short-time Fourier transform (STFT) [5], wavelet transform (WT) [6], Wigner-Ville distribution (WVD) [7], and empirical mode decomposition (EMD) [8]. The above methods can extract features that are conducive for classification and diagnosis [9, 10] and then pass the extracted features through various classifiers to realize pattern recognition of bearing faults. Among the various pattern recognition methods, the machine learning-based method is the most used. Wang et al. [11] used KPCA to extract features from bearing fault signal and used k-nearest neighbor (KNN) as a classifier to achieve diagnosis; Fei et al. [12] reconstructed the characteristics of bearing vibration signal after singular value decomposition based on wavelet packet transform phase space and established support vector machine (SVM) model of bearing diagnosis; Mahamad and Hiyama [13] performed fast Fourier transform (FFT) and envelope processing on the bearing vibration signal, extracted time domain and frequency domain feature as input, and then used ANN to fulfill the diagnosis. However, the existing intelligent fault diagnosis methods based on the

above feature extraction and classification still have three limitations: first, the feature extraction methods often require the operators to have professional prior knowledge and rich experience. As the research progresses, the form of input signal becomes more diversified, and its objectivity and accuracy may be affected if feature extraction is still based on past experience [14, 15]; second, the feature extraction methods are poor in generality, and often a method only has a good feature extraction result for a certain type of signal; and third, feature extraction and pattern recognition are two independent processes, and the diagnosis model cannot be jointly optimized globally.

In recent years, the successful application of deep learning in the fields of speech recognition [16], face recognition [17], computer vision [18], and image processing [19] has made it a research hotspot. Various deep learning models can extract abstract features directly from the original signal, avoiding manual extraction of feature [20], and they also have better universality [21] and can jointly optimize the two processes of feature extraction and pattern recognition in various classification problems [22]. Thanks to these advantages, researchers have introduced a variety of deep learning models in bearing fault diagnosis; for example, Duong and Kim [23] constructed a DNN structure which is based on the stacked denoising autoencoder (DAE) non-mutually exclusive classifier (NMEC) method for combined modes to realize bearing fault diagnosis, Shao et al. [24] developed a convolutional deep belief network with Gaussian visible units to obtain an excellent accuracy rate of bearing fault diagnosis, Chen and Li [25] utilized the acceleration sensors to collect the vibration signal of the bearing and input the time domain and frequency domain characteristics of the signal into multiple two-layer sparse autoencoder (SAE) neural networks for feature fusion, and then the fused feature was further classified by DBN. Lu et al. [26] established a deep neural network model based on autoencoder (AE) and achieved good results in bearing fault diagnosis. Shao et al. [27] proposed a novel optimization deep belief network (DBN) for bearing fault diagnosis which is verified by the simulation signal and experimental signal of a rolling bearing.

Figure 1 shows the main differences between the traditional fault diagnosis method and the deep learning-based fault diagnosis method.

Convolutional neural network [28] is a typical deep learning model that has also attracted attention. It extracts the characteristics of the signal layer by layer through convolution, pooling, and nonlinear activation function mapping. Compared with the fully connected deep learning model, CNN has stronger robustness and better generalization ability [29]. At the same time, CNN improves network performance and reduces training costs by weight sharing and pooling operation and is less prone to overfitting problem than other deep learning models [30]. From the perspective of input, the existing CNN models include two types: one-dimensional convolutional neural network (1DCNN) and two-dimensional convolutional neural network (2DCNN).

For 2DCNN, its input is actually two-dimensional matrix. In the fault diagnosis of bearing, researchers used a

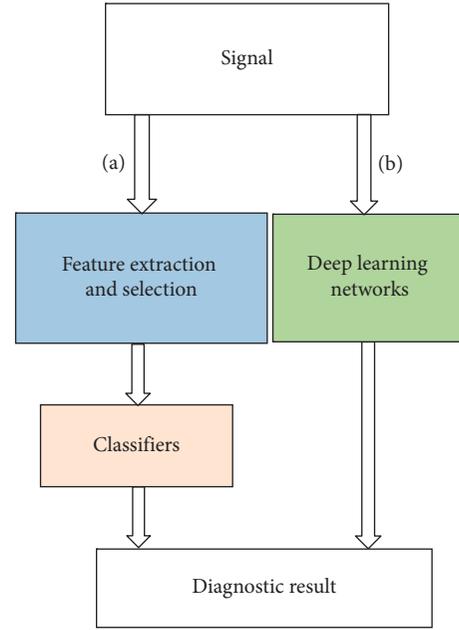


FIGURE 1: Block diagram of diagnosis methods: (a) traditional fault diagnosis method; (b) deep learning-based fault diagnosis method.

variety of methods to convert one-dimensional original signal into two-dimensional matrix and then used it as 2DCNN input. In [20], the one-dimensional signal was converted into two-dimensional gray map as the input of 2DCNN, and the input of 2DCNN in [31] was the root mean square (RMS) map of the characteristics of the vibration signal after Fourier transform (FT). In [32], the continuous wavelet transform scale (CWTS) map was directly classified by 2DCNN.

However, in practice, the bearing vibration signal is a one-dimensional time signal, and the method of converting the original one-dimensional signal to two-dimensional signal also depends on experience. These methods cannot guarantee whether there is torsion, distortion, or even loss of useful information in the conversion process, which may result in insufficient characteristic learning and low accuracy. Therefore, if the original one-dimensional signal is used as input directly, the input of the network will contain all the feature information in the original signal and the above problem can be avoided. In addition, compared with 2DCNN, 1DCNN has better interpretability, and the convolution kernel and its extracted feature are one-dimensional vectors, so that multiple signal processing methods can be used to study the convolution kernel and its extracted feature conveniently, which is conducive to further understand 1DCNN and its feature extraction mechanism.

For 1DCNN, its input is one-dimensional vector. In practice, the actual measured signal often contains a lot of noise, which will greatly increase the difficulty in extracting fault features in a simple shallow CNN model. In the case where the measured noisy signal is input, the diagnosis accuracy can be improved by the following two methods.

$$\mathbf{H}_x = \begin{bmatrix} x_1 & x_2 & \cdots & x_{N-n+1} \\ x_2 & x_3 & \cdots & x_{N-n+2} \\ \vdots & \vdots & \ddots & \vdots \\ x_{N-n+1} & x_{N-n+2} & \cdots & x_N \end{bmatrix}_{m \times n}, \quad (6)$$

where $1 < n < N$, $m = N - n + 1$. Each matrix has multiple Hankel matrices with different column combinations. When constructing Hankel matrix, the product of row number m and column number n of matrix should be maximized as far as possible, and the best way to construct matrix should be square matrix or near square matrix [37]. According to the inequality principle, when m and n are equal or close, the product of the two numbers is the largest, so the structure of the optimal Hankel matrix is determined as follows:

$$m = \begin{cases} \frac{N}{2}, & N \text{ is even,} \\ \frac{(N+1)}{2}, & N \text{ is odd.} \end{cases} \quad (7)$$

The key of denoising noisy signal by SVD is how to determine the singular value of useful signal and the singular value of noise signal. Zhao et al. [38] proposed a method to determine the singular value of useful signal based on singular value difference spectrum. Assuming that the form of Hankel matrix of noisy signal \mathbf{X} is shown in equation (5) and that there are q ($q < k$) singular values of useful signal determined by singular value difference spectrum method, so the Hankel matrix of useful signal separated by this method can be expressed as

$$\mathbf{A}' = \sigma_1 \mathbf{u}_1 \mathbf{v}_1 + \sigma_2 \mathbf{u}_2 \mathbf{v}_2 + \cdots + \sigma_q \mathbf{u}_q \mathbf{v}_q \quad (q < k). \quad (8)$$

Furthermore, the denoised signal can be obtained by reducing \mathbf{A}' to one-dimensional signal. It can be seen that the singular value difference spectrum method is a denoising method based on the characteristics of the data itself.

2.2. Proposal of Diagnostic Model. Figure 2 shows the SVD-1DCNN structure constructed in this paper. The network embeds an SVD layer after C1 to realize further feature transformation. The feature maps in the SVD layer are connected to the corresponding feature maps in C1.

The SVD-1DCNN mainly includes input layer, convolution layers, pooling layers, fully connected layer, output layer, and SVD layer. The convolution layers, the pooling layers, and the SVD layer are the core structures of the SVD-1DCNN, and each of the convolution layers, the pooling layers, and the SVD layer has several feature maps. Each feature map connected to a corresponding feature map on the adjacent layer, and the output of the previous layer feature map is the input of the next layer feature map.

In this structure, SVD layer denoises and reconstructs the output of C1 (primary classification feature) to achieve joint optimization of feature extraction and denoising and reconstruction. The denoised and reconstructed features are used as input to the next layer. In this network structure, the convolution kernels realize adaptive denoising of signals,

and the useful feature components required by the network are highlighted. The SVD layer's denoising and reconstruction process further highlights the features, so it is more conducive to network extraction classification features.

Since SVD-1DCNN includes a SVD layer, in order to ensure that the network can carry out backpropagation, it is necessary to ensure that the error can be backpropagated from S2 layer to C1 layer, and the weights and bias in C1 can be updated. This process of SVD-1DCNN is described below.

Suppose the signal in the input layer is $L^0 = [l_1^0, l_2^0, \dots, l_s^0]$, the output of C1's feature map is $L^1 = [l_1^1, l_2^1, \dots, l_s^1]$, and the output of SVD layer's corresponding feature map is $L^2 = [l_1^2, l_2^2, \dots, l_s^2]$. Therefore, the output of the i -th node of the feature map of C1 is l_i^1 , the output of the corresponding node on the SVD layer's corresponding feature map is l_i^2 , and the difference between the two is $\Delta l_i = l_i^2 - l_i^1$. The singular value difference spectrum method is based on the characteristics of the signal itself to achieve denoising, so when given the input, $\Delta l_i = l_i^2 - l_i^1$ is a constant. In the process of error backpropagation, assuming that the error from S2 layer to the i -th node on SVD layer's feature map is E_i^2 , then the error from this node on SVD layer's feature map to the corresponding i -th node on C1's corresponding feature map is $E_i^1 = E_i^2 + \Delta l_i$. Because Δl_i is a constant, E_i^1 is differentiable to the weights and bias in the backpropagation process. In the process of network training, the weights and bias in C1 can be updated by E_i^1 .

Figure 3 shows the flowchart of the method, and the specific steps are as follows:

Step 1: the original signals are taken as network input. After the feature transformation through the C1 layer, the primary classification features in the original signals are extracted, and the primary classification features are used as the input of the SVD layer.

Step 2: in each training process, the SVD layer denoises and reconstructs the output of C1 to further extract higher-level classification features.

Step 3: the classification features extracted by the SVD layer are used as the input of the next layer to achieve deeper feature transformation.

In each training process, the SVD layer denoises and reconstructs the output of C1, which is conducive to the network to obtain more obvious classification characteristics of signals under the background of noise and enhance the network fault diagnosis ability.

3. Performance Analysis Based on Simulated Signals

3.1. Construction of Simulation Signals. The following are four types of simulation signals with a signal-to-noise ratio of 20 dB. The feasibility of the research is verified by the classification of four types of signals. Table 1 shows the relevant parameters of the signals, where $u(t)$ is the Gaussian white noise and φ_{ij} is the phase of the signals, which is randomly generated between 1 and 100. The time

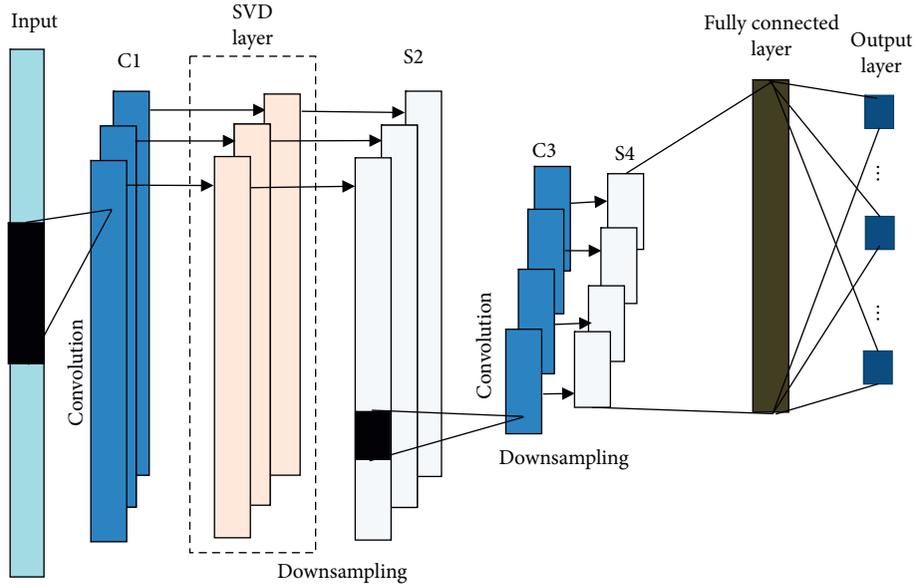


FIGURE 2: Structure of SVD-1DCNN.

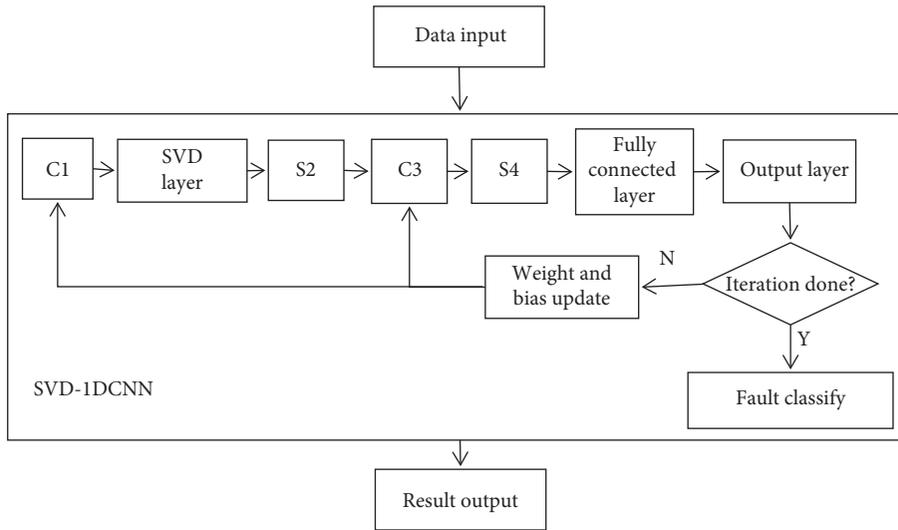


FIGURE 3: SVD-1DCNN diagnosis flowchart.

TABLE 1: Four types of simulated signals.

Signal	Expression	φ_{ij}
Y_1	$\sin(100\pi t + \varphi_{1j}) \cdot \cos(250\pi t + \varphi_{2j}) + u(t)$	$\varphi_{1j} = \text{rand}[1, 100]$ $\varphi_{2j} = \text{rand}[1, 100]$
Y_2	$e^{-t} \sin(400\pi t^2 + \varphi_{3j}) + u(t)$	$\varphi_{3j} = \text{rand}[1, 100]$
Y_3	$2 \cos(200\pi t e^{-t} + \varphi_{4j}) + u(t)$	$\varphi_{4j} = \text{rand}[1, 100]$
Y_4	$t \sin(150\pi t + \varphi_{5j}) + t \cos(550\pi t + \varphi_{6j}) + u(t)$	$\varphi_{5j} = \text{rand}[1, 100]$ $\varphi_{6j} = \text{rand}[1, 100]$

domain and time-frequency diagrams of the four types of signals are shown in Figure 4.

In SVD-1DCNN, since the original signals contain noise, the output of each feature extraction layer also contains

noise. In order to quantify the feature extraction effect of each feature extraction layer, eSNR is defined as an index. Assuming that the output of a layer is x and the denoised signal reconstructed by SVD is s , the noise component can

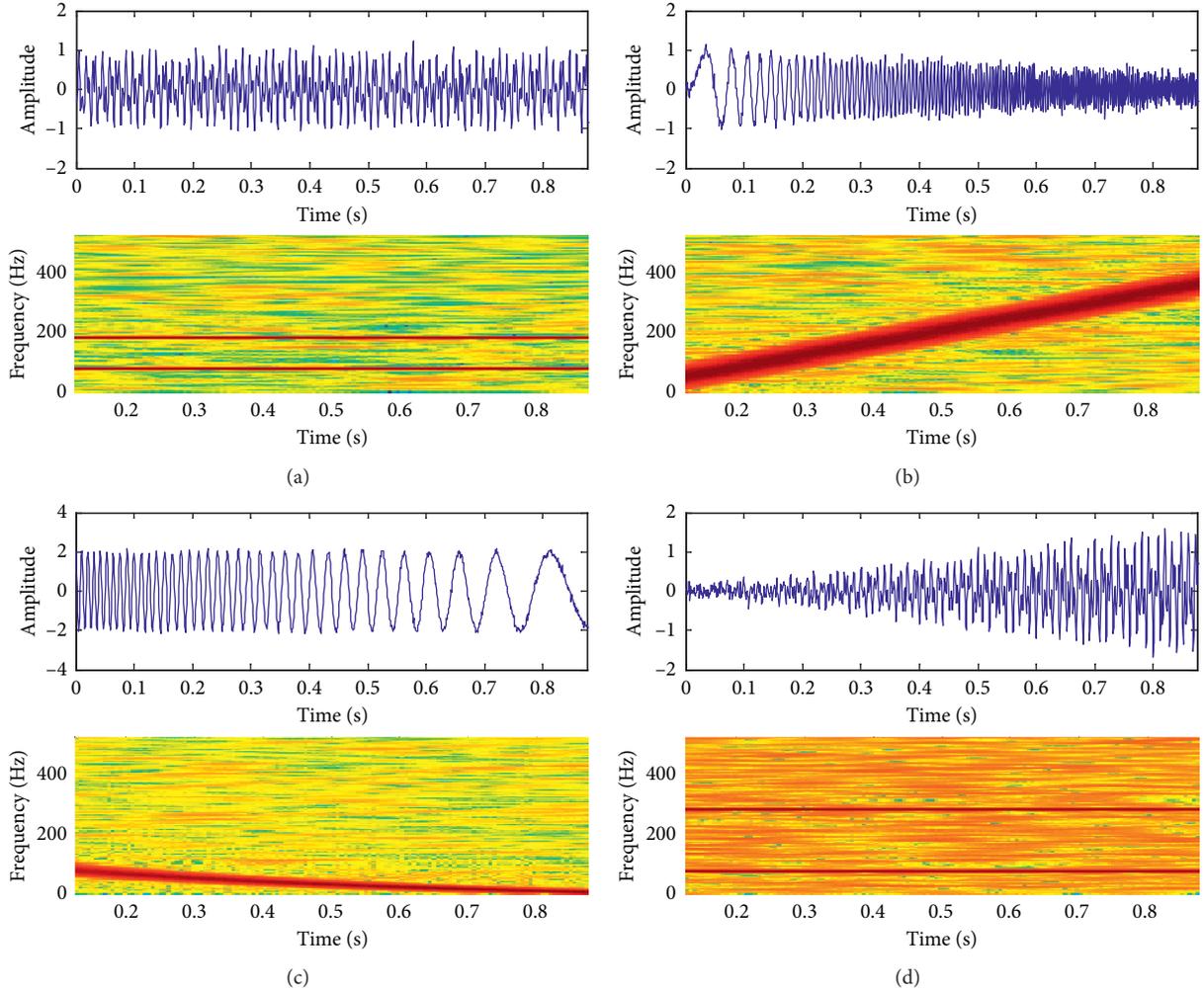


FIGURE 4: Time domain and time-frequency diagrams of the simulated signals: (a) Y_1 , (b) Y_2 , (c) Y_3 , and (d) Y_4 .

be expressed as $x - s$. Then the eSNR of the output of this layer can be expressed by the following equation:

$$\text{eSNR} = 10 \lg\left(\frac{s}{x - s}\right). \quad (9)$$

Further, the eSNR of the simulation signals are calculated by equation (9), and the results are shown in Table 2. As can be seen from Table 2, eSNR can accurately reflect the SNR of the signals.

3.2. Pattern Recognition of Simulated Signal. The network structure of SVD-1DCNN is as shown in Figure 2. Because SVD-1DCNN is an improved network based on 1DCNN, its network structure and network parameters are also based on specific pattern recognition tasks and experience. In this pattern recognition task, there are only four types of simulated signals, so according to the previous experience, the learning rate of the network is set to be 0.1, the training batch is set to be 10, and the maximum number of iterations is set to be 1500, the pooling method of the two pooling layers is average pooling, and the step size is set to be 2. For convenience of

TABLE 2: SNR(dB) and eSNR(dB) of four types of simulated signals.

	SNR (dB)	eSNR (dB)
Y_1	20	19.287
Y_2	20	19.166
Y_3	20	19.109
Y_4	20	19.001

representation, (m, n) - $[p, q]$ is used to represent the relevant parameters in the network, where m and n , respectively, represent the size of the convolution kernels in the two convolution layers, and p and q , represent, the number of convolution kernels in the corresponding convolution layer.

SVD is usually used in the preprocessing in signal processing, that is, the original signals are denoised firstly and then the denoised signals are used in the subsequent analysis. Therefore, in order to compare the classification effects, a network structure (SVD + 1DCNN) is constructed. In the new network, the original signals are denoised firstly, and the denoised signals are used as the input of the network to realize pattern recognition as shown in Figure 5. The other

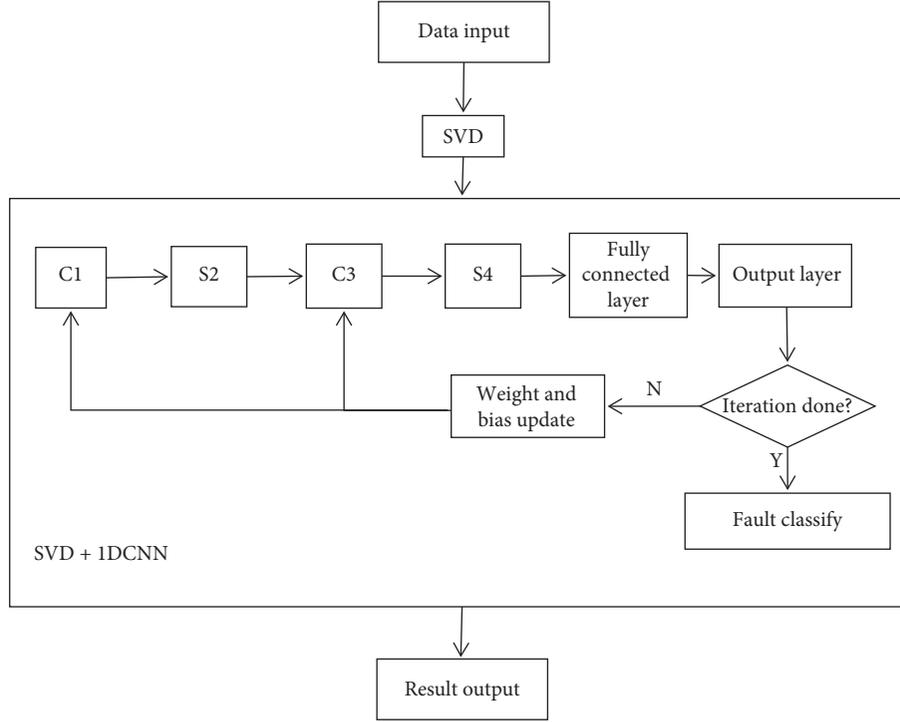


FIGURE 5: SVD + 1DCNN diagnosis flowchart.

parameters setting in the new network are the same as in SVD-1DCNN.

Four types of simulation signals are used for the experiment. Each signal contains 60 samples (50 training samples and 10 test samples). In order to verify the stability of the networks, 10 experiments were conducted for each network structure. In addition, the classification results of each experiment of the two networks were evaluated by confusion matrix and accuracy. Confusion matrix is calculated by the four parts composed of true label and prediction label, which are true positive (TP), false negative (FN), false positive (FP), and true negative (TN), respectively. The confusion matrix is shown in Table 3.

The accuracy is the overall judgment of the classification model and the proportion of correct prediction in the total amount. The calculation method is as follows:

$$A = \frac{TP + TN}{TP + FP + FN + TN}, \quad (10)$$

A_i ($i = 1, 2, \dots, 10$), \bar{A} , and Var were, respectively, used to represent the classification accuracy of each experiment, the average accuracy, and the variance of the accuracy, and then A_i , \bar{A} , and Var satisfy the following equations:

$$\begin{aligned} \bar{A} &= \sum_{i=1}^{10} \frac{A_i}{10}, \\ \text{Var} &= \sum_{i=1}^{10} \frac{(A_i - \bar{A})^2}{10}. \end{aligned} \quad (11)$$

In the classification of simulation signals, different network structures need to be set up for multiple

TABLE 3: Confusion matrix.

Predicted label	True label	
	Positive class	Negative class
Positive class	True positive (TP)	False positive (FP)
Negative class	False negative (FN)	True negative (TN)

experiments to determine the best network structure. Among various network structures, the network whose structure is (351, 80)-[3, 3] is taken as an example to show its confusion matrix in one experiment and the \bar{A} and Var of the network after 10 experiments.

Table 4 is the confusion matrix in one experiment when the structure of SVD-1DCNN is (351, 80)-[3, 3].

Table 5 is the confusion matrix in one experiment when the structure of SVD + 1DCNN is (351, 80)-[3, 3].

As can be seen from Table 4, SVD-1DCNN can classify every type of signal correctly; as can be seen from Table 5, for Y_1 , the classification accuracy of SVD + 1DCNN is 90%, but 10% is classified as Y_4 . For Y_2 , the classification accuracy of SVD + 1DCNN is 80%, but 10% is classified as Y_1 , and the remaining 10% is classified as Y_4 . For Y_3 , the classification accuracy of SVD + 1DCNN is 90%, but 10% is classified as Y_4 . For Y_4 , the classification accuracy of SVD + 1DCNN is 90%, but 10% is classified as Y_1 . In general, SVD-1DCNN has a higher classification accuracy than SVD + 1DCNN.

10 experiments were carried out on both networks. Table 6 shows A_i of each experiment, \bar{A} , and Var of the two networks after 10 experiments.

As can be seen from Table 6, in multiple experiments, the variance of SVD-1DCNN is 0 and the variance of SVD-1DCNN is 1.25×10^{-4} , indicating that both networks have

TABLE 4: Confusion matrix of SVD-1DCNN for classification results of simulated signals.

Predicted label	True label (%)			
	Y_1	Y_2	Y_3	Y_4
Y_1	100	0	0	0
Y_2	0	100	0	0
Y_3	0	0	100	0
Y_4	0	0	0	100

TABLE 5: Confusion matrix of SVD + 1DCNN for classification results of simulated signals.

Predicted label	True label (%)			
	Y_1	Y_2	Y_3	Y_4
Y_1	90	0	0	10
Y_2	10	80	0	10
Y_3	0	0	90	10
Y_4	10	0	0	90

TABLE 6: A_i , \bar{A} , and Var of SVD-1DCNN and SVD + 1DCNN.

	SVD-1DCNN	SVD + 1DCNN
1	100%	85%
2	100%	87.5%
3	100%	90%
4	100%	87.5%
5	100%	87.5%
6	100%	87.5%
7	100%	87.5%
8	100%	87.5%
9	100%	87.5%
10	100%	87.5%
\bar{A}	100%	87.5%
Var	0	1.25×10^{-4}

excellent stability. According to the final experimental results, SVD-1DCNN has a higher \bar{A} than SVD + 1DCNN. In the classification of simulation signals, SVD-1DCNN has a better classification effect.

In addition, \bar{A} and Var of SVD-1DCNN and SVD + 1DCNN were calculated with different network structures, as shown in Table 7.

As shown in Table 7, both networks have excellent stability, and the classification effect of SVD-1DCNN is better than that of SVD + 1DCNN. It can be seen that the number of convolution kernels has a greater impact on the classification results. The number of convolution kernels in the network is too small to make the network fail to achieve high classification accuracy, but the number of convolution kernels is not as good as possible. Excessive convolution kernels may even reduce the training effect of the network.

In the above network structure, the network structure of (351, 80)-[3, 3] has the best classification effect, so it will be taken as the research object in the following part. Therefore, the final parameters of SVD-1DCNN are as follows: the first convolutional layer contains three convolution kernels, each of which has a size of 1×351 ; the second convolutional layer

contains three convolution kernels, each with a size of 1×80 ; the learning rate is 0.1; the training batch is 10; the maximum number of iterations is 1500; the pooling mode of the two pooling layers is average pooling; and the step size is 2. The corresponding parameters in SVD + 1DCNN are the same as those in SVD-1DCNN.

3.3. Analysis of the Role of Convolution Kernel

3.3.1. Qualitative Analysis. In the SVD-1DCNN network structure, each feature map in the convolution layer contains a convolution kernel, and the convolution results of the convolution kernel with the signals are the output of the feature map. In order to analyze the role of the convolution kernels during training, Y_1 is taken as an example. During the training process, the output of the feature map of C1 of SVD-1DCNN is extracted, and its time domain and time-frequency diagrams are as shown in Figure 6.

It can be seen from Figure 6 that, during the training process, the convolution kernel highlights part of the frequency characteristics and suppresses other frequency characteristics. C1 highlights this portion of the frequency characteristics as primary classification features of the input. It is worth noting that the convolution kernel only selects part of the frequency features from the input as the primary classification features, which adaptively realizes the dimensionality reduction of the data and improves the classification efficiency of the network.

Figure 7 shows the time domain and frequency domain diagrams of the output of the C1 feature map for the four types of signals at the end of the network training. It can be seen that the convolution kernel performs different denoising operations on the four types of signals and retains the main frequency components in the original signals.

At the same time, in order to more intuitively analyze the feature extraction effect of the convolution kernel on the signals, the output of C1 is shown in Figure 8. Figure 8 shows the time-frequency diagrams of the four original signals and the time-frequency diagrams of the output of C1 during network training.

It can be seen from Figures 7 and 8 that, in the training process, the convolution kernel realizes the feature extraction of the original signals. As the number of iteration increases, the noise components are gradually eliminated, highlighting the features learned by the network.

3.3.2. Quantitative Analysis. In order to analyze the feature extraction effect of convolution kernels on the original signal, eSNR is used as the index for evaluation. Table 8 is the eSNR of C1's three feature maps' output when the iteration times are 200, 500, and 1000, respectively.

In order to visually explain the change in eSNR, Y_1 is used as an object. Figure 9 shows the eSNR of three feature maps' output according to equation (9).

It can be seen from Figures 8 and 9 that, in the training process, the primary classification features of each convolution kernel extraction have a higher eSNR than the input,

TABLE 7: \bar{A} and Var of different network structures.

	SVD-1DCNN		SVD + 1DCNN	
	\bar{A}	Var	\bar{A}	Var
(351, 80)-[1, 1]	Divergency	Divergency	Divergency	Divergency
(351, 80)-[2, 2]	95%	1.25×10^{-4}	85%	1.25×10^{-4}
(351, 80)-[3, 3]	100%	0	87.5%	1.25×10^{-4}
(351, 80)-[4, 4]	97.5%	0	85%	0
(351, 80)-[6, 6]	90%	1.25×10^{-4}	82.5%	1.25×10^{-4}
(351, 80)-[8, 8]	90%	0	82.5%	0
(351, 80)-[10, 10]	87.5%	0	82.5%	0

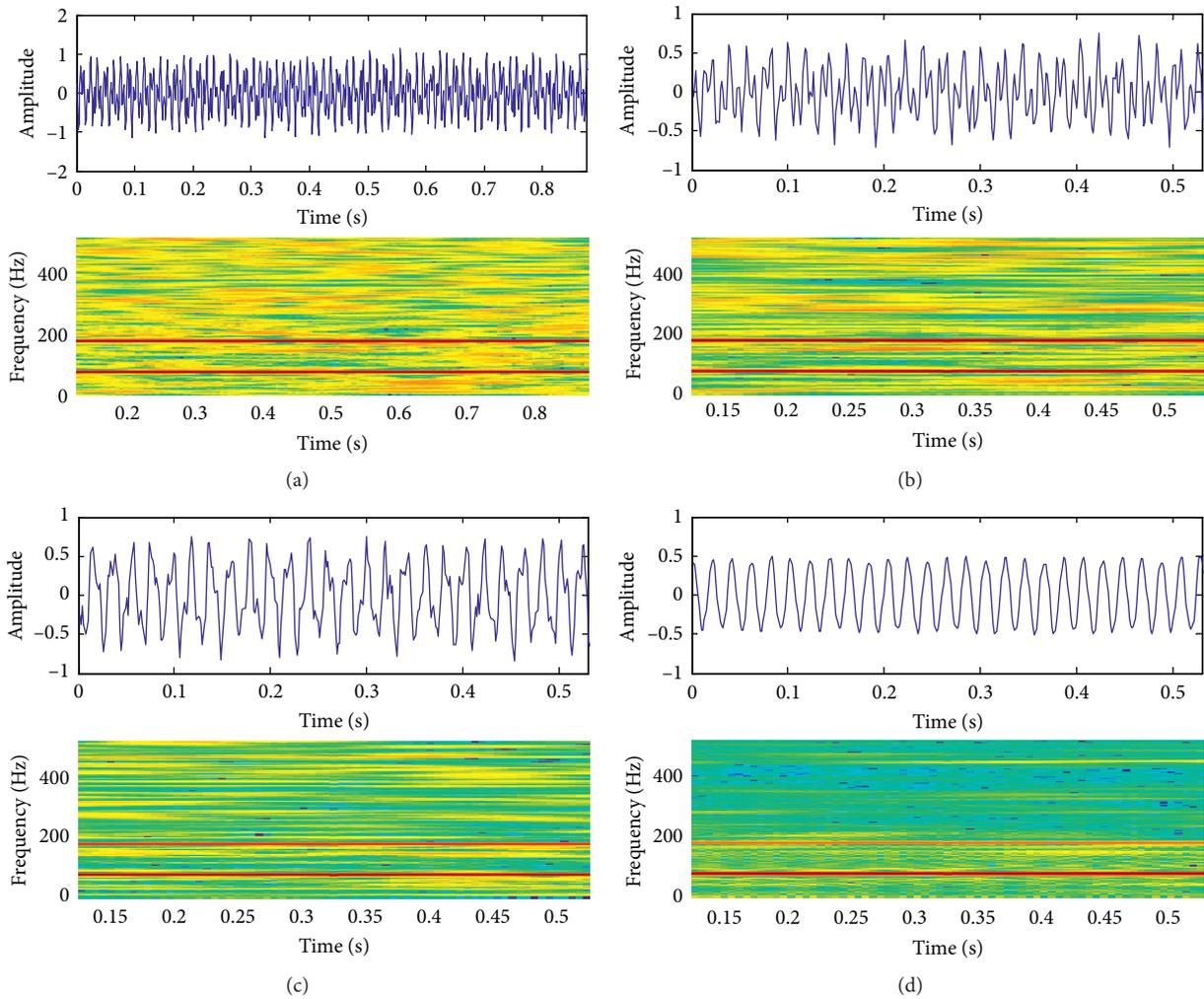


FIGURE 6: Time domain and time-frequency diagrams of the output of C1 in different iteration stages: (a) original signal; (b) iteration is 200; (c) iteration is 500; (d) iteration is 1000.

which highlights the useful feature components in the signals. At the same time, as the number of iterations increases, the eSNR of the primary classification features is higher, and the useful feature components in the signal are more significant.

Through simulation experiment, it can be found that, in the training process, the convolution kernels can adaptively remove the noise components in the signals according to the

characteristics of the original signals and retain the learned features. It can be said that the convolution kernels not only extract the characteristic components in the original signals, but also achieve denoising.

3.4. Analysis of Two Networks' Classification Effects. SVD + 1DCNN and SVD-1DCNN have different classification effects on the same dataset. In the two network

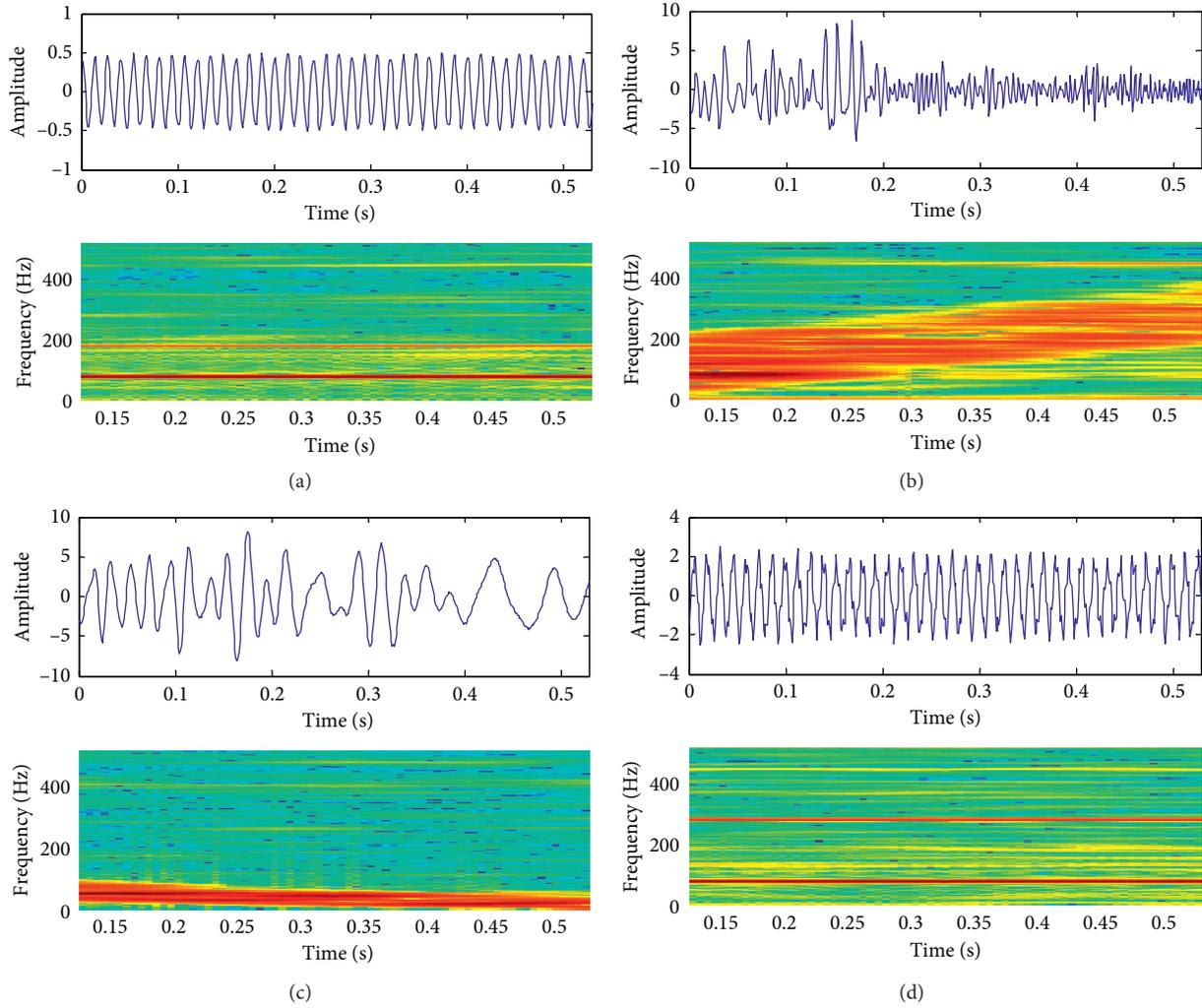


FIGURE 7: Time domain and time-frequency diagrams of the output of C1 of four types of signals at the end of training: (a) Y_1 , (b) Y_2 , (c) Y_3 , and (d) Y_4 .

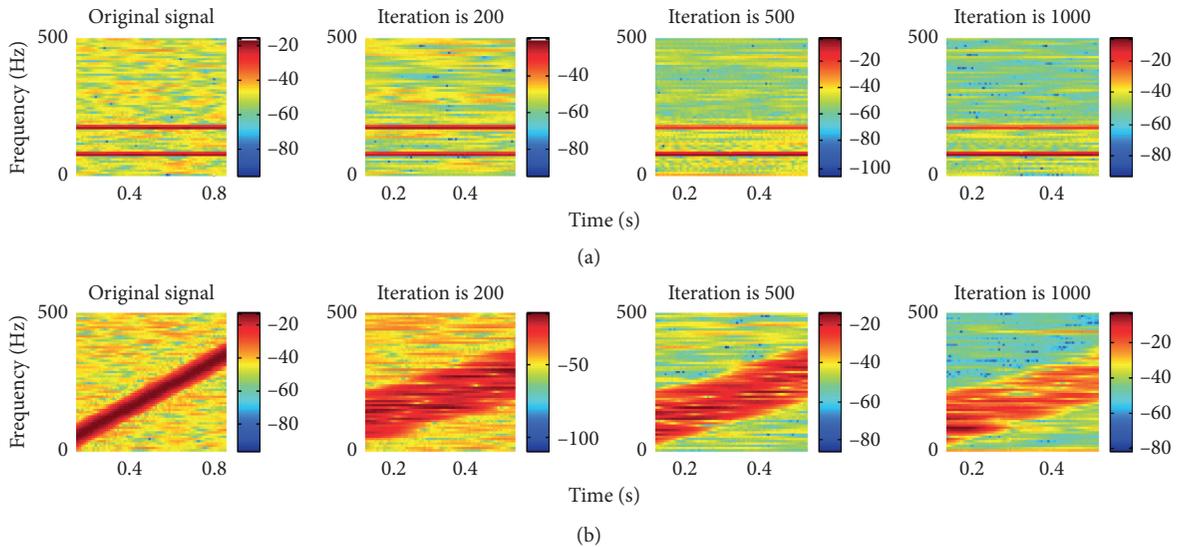


FIGURE 8: Continued.

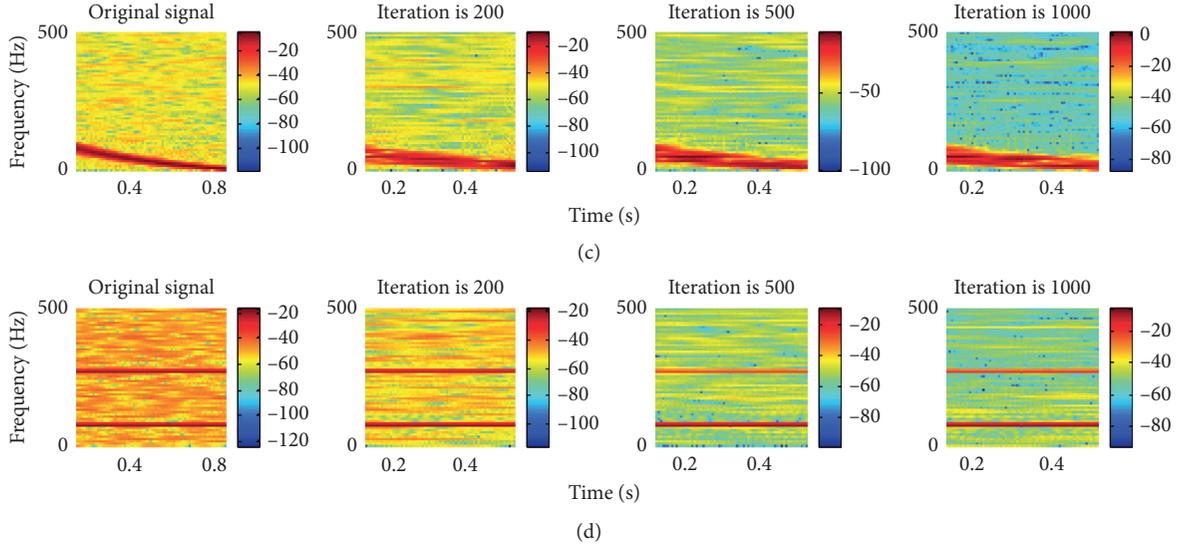


FIGURE 8: Time-frequency diagrams of the output of C1 of four types of signals in different iteration stages: (a) Y_1 , (b) Y_2 , (c) Y_3 , and (d) Y_4 .

TABLE 8: eSNR of C1's three feature maps' output.

Input	C1 output	Iterations		
		200	500	1000
Y_1	Feature map 1	24.090	37.783	41.507
	Feature map 2	27.967	35.906	40.328
	Feature map 3	24.672	37.418	41.003
Y_2	Feature map 1	28.414	33.305	37.819
	Feature map 2	27.756	33.910	38.203
	Feature map 3	29.455	34.477	39.551
Y_3	Feature map 1	35.443	40.511	47.872
	Feature map 2	35.787	41.335	46.914
	Feature map 3	33.706	40.591	45.370
Y_4	Feature map 1	25.801	36.551	41.434
	Feature map 2	25.513	35.857	38.749
	Feature map 3	23.790	35.934	38.023

structures, except for the structure, the other parameters are the same. By comparing the two network structures, it can be seen that in SVD + 1DCNN, the S2's input is the output of C1, and in SVD-1DCNN, that is the output of the SVD layer. Therefore, eSNR of the S2's input is calculated to analyze the feature extraction capabilities of the two networks. For convenience of presentation, map 1, map 2, and map 3 are used to represent the output of C1 of SVD + 1DCNN, and Smap 1, Smap 2, and Smap 3 are used to represent the output of the SVD layer in SVD-1DCNN. Figure 10 shows eSNR of the S2's input in two networks.

It can be seen from Figure 10 that, in the training process of the two networks, the eSNR of S2's input increases with the increase of the number of iteration, and the networks' feature extraction ability is stronger. With the same number of iteration, the eSNR of S2's input in SVD-1DCNN is higher than that in SVD + 1DCNN. This indicates that the input features of S2 in SVD-1DCNN are more obvious. In further feature extraction, features with high eSNR are more conducive to network learning high-level features.

Therefore, it can be said that SVD-1DCNN has stronger feature extraction ability than SVD + 1DCNN, which is conducive to improving the accuracy of pattern recognition.

4. Bearing Fault Diagnosis Based on SVD-1DCNN

4.1. Data Collection and Processing. The experimental data in this paper come from the bearing database of the Case Western Reserve University (CWRU) [38]. The experimental data are the acceleration data of the drive end at a sampling frequency of 12 KHz. The data include four types: data with a fault diameter of 0.007 mils on the rolling element, data with a fault diameter of 0.007 mils on the inner ring, data with a fault diameter of 0.007 mils on the outer ring, and normal data.

The length of each segment of the signal is about 120,000. In order to increase the randomness of the training set and the test set, a window of length 1024 is used to sample the signal in random steps from the first node of each signal, as shown in Figure 11. In the sampling process, 60 samples are obtained from each signal, and among the 60 samples, 10 samples are randomly selected as the test set, and the remaining samples are used as samples of the training set. In this way, a training set containing 200 samples and a test set containing 40 samples are obtained.

4.2. Fault Diagnosis of Measured Signals. Figure 12 shows the signals in the four states in the time domain and time-frequency diagrams. It can be seen that the potential failure modes are masked by noise and the signal characteristics are hidden in strong background noise and unrelated interference.

As can be seen from Figure 12, the four measured signals contain a large amount of noise, which increases the difficulty of pattern recognition. The eSNR of the four types of signals is calculated according to equation (9), as shown in Table 9.

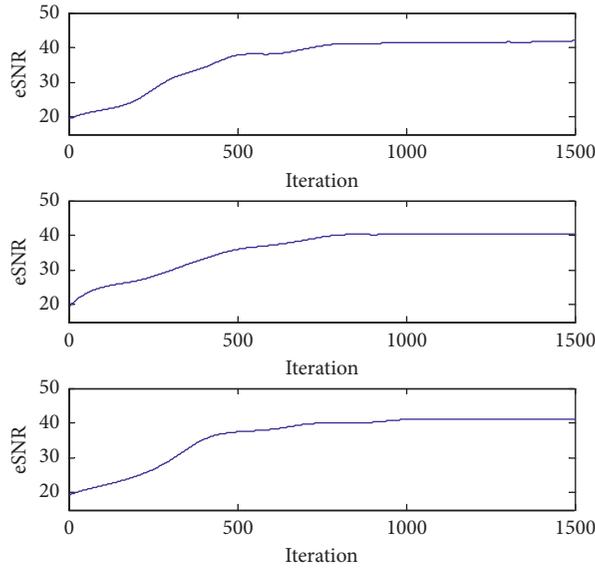


FIGURE 9: eSNR of three feature maps' output.

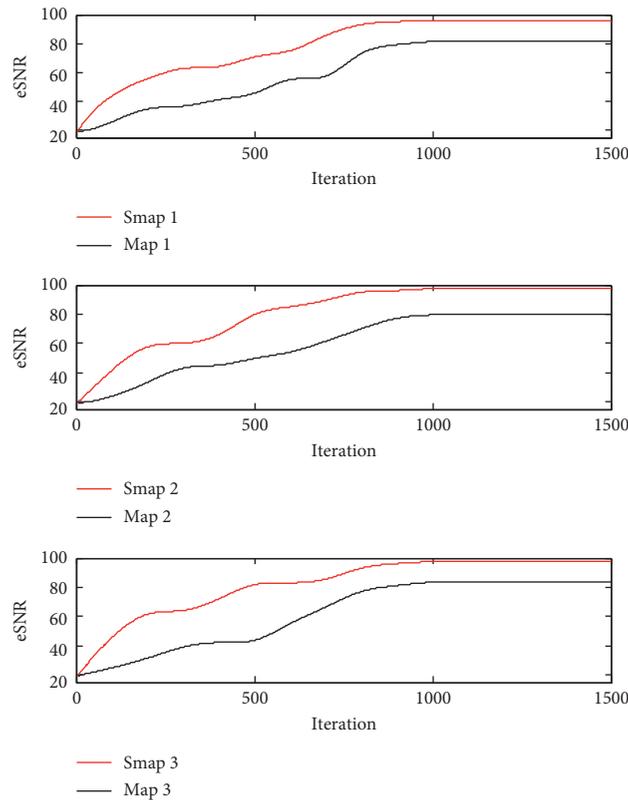


FIGURE 10: eSNR of the S2's input of the two networks.

Figure 13 shows the output of C1 of four types of signals during network training. As can be seen from Figure 13, the characteristics of the original signals are submerged in a large amount of noise, but after C1 convolution, the noise in the original signal is gradually eliminated. As the number of iteration increases, the characteristic components in the original signal gradually become prominent.

In order to visually reflect the change process of eSNR, the rolling element fault signal is selected for explanation. Figure 14 shows the change of eSNR of the output of the three feature maps of C1 during the training.

It can be seen from Figure 14 that the measured signal has a lower eSNR, and the eSNR of the signal is improved after the C1 feature extraction. In the pattern recognition of

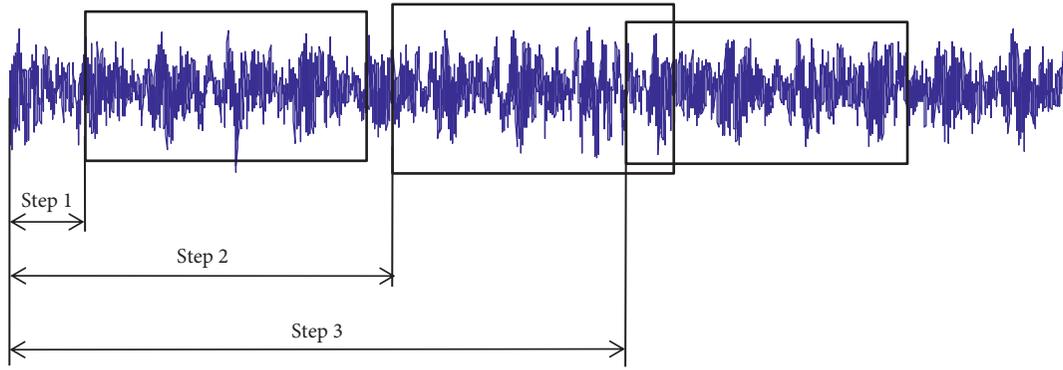


FIGURE 11: Data sampling.

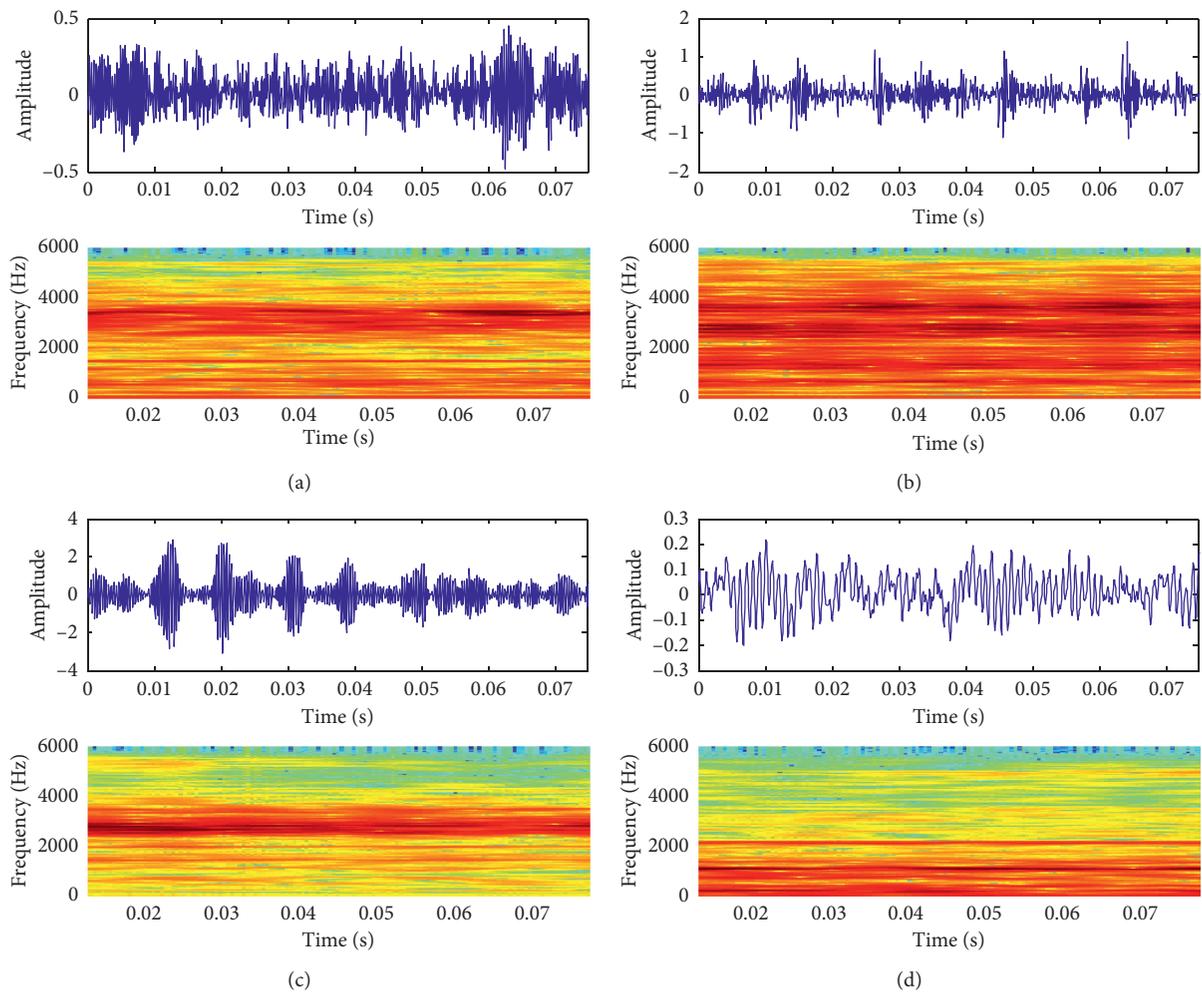


FIGURE 12: Time domain and time-frequency diagrams of measured signals: (a) roll damage; (b) inner ring damage; (c) outer ring damage; (d) normal.

TABLE 9: eSNR of four types of measured signals.

	Roll damage	Inner ring damage	Outer ring damage	Normal
eSNR	5.320 dB	2.429 dB	13.377 dB	20.709 dB

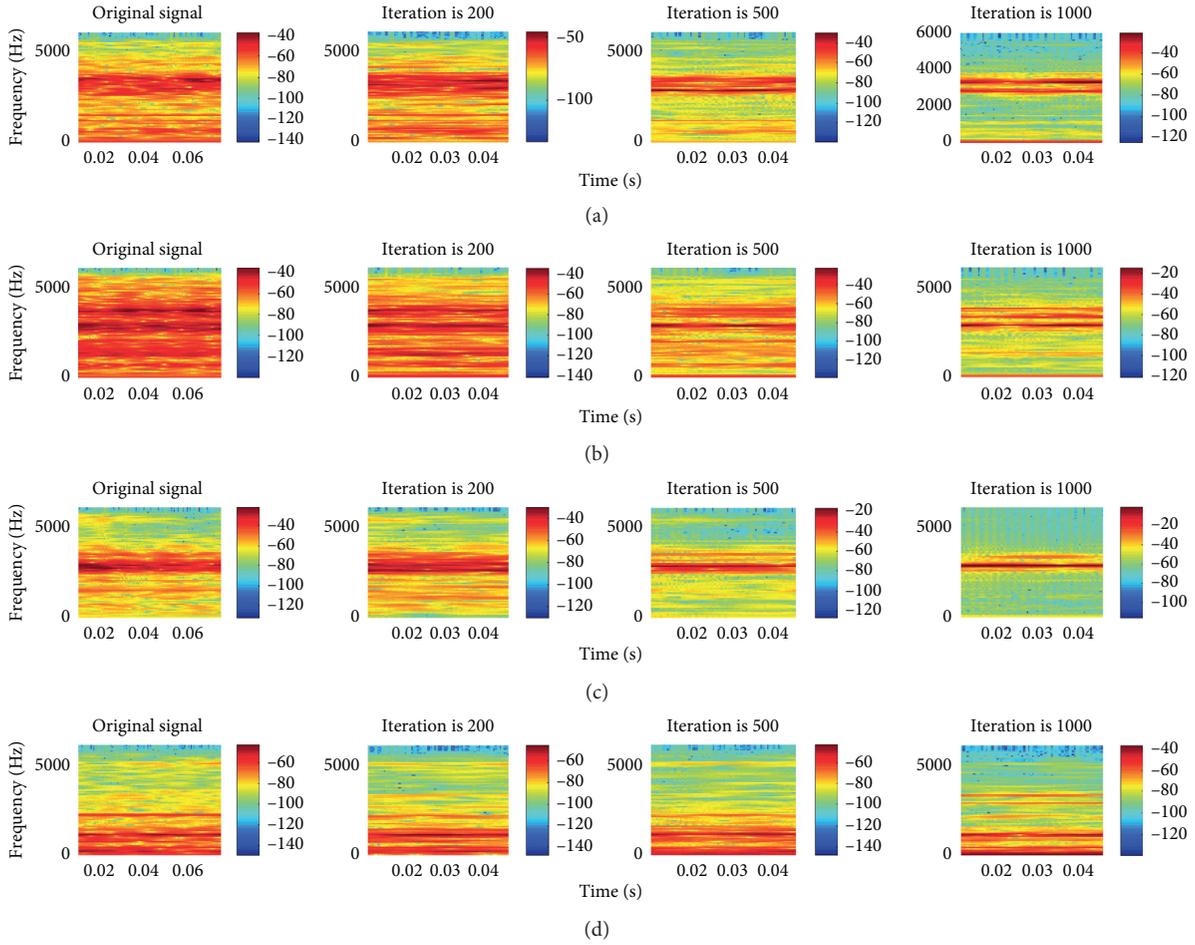


FIGURE 13: Time-frequency diagrams of the output of C1 of four types of measured signals in different iteration stages: (a) roll damage; (b) inner ring damage; (c) outer ring damage; (d) normal.

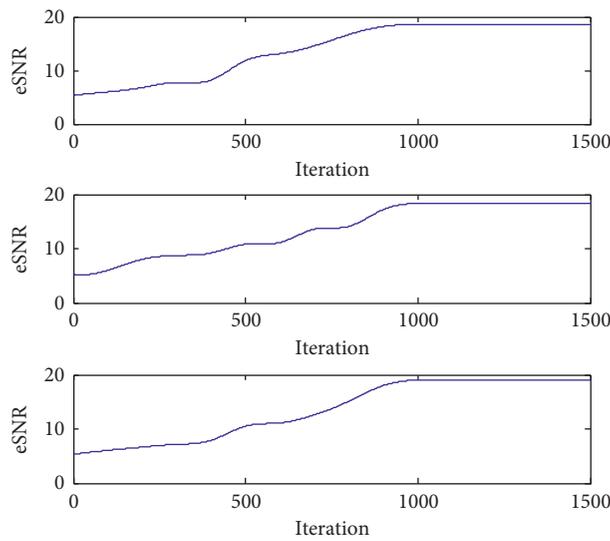


FIGURE 14: eSNR of three feature maps' output.

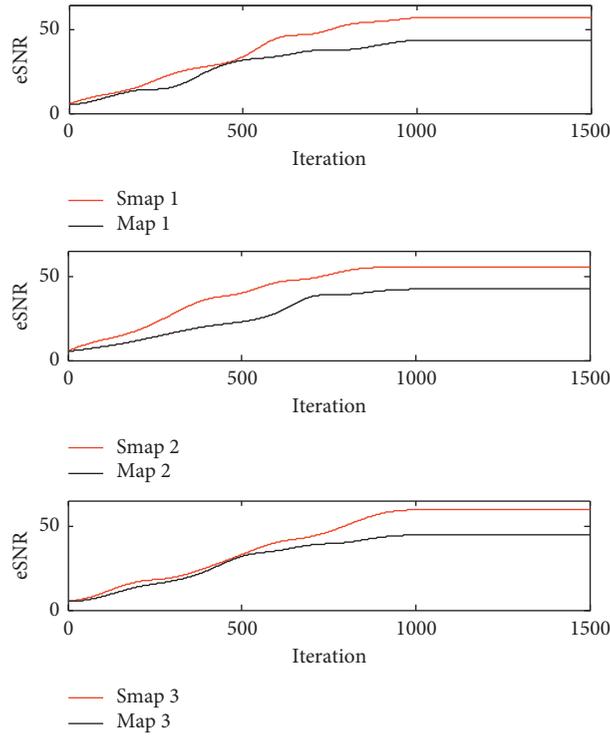


FIGURE 15: eSNR of the S2's input of the two networks.

TABLE 10: Confusion matrix of SVD-1DCNN for classification results of measured signals.

Predicted label	True label (%)			
	Roll damage	Inner ring damage	Outer ring damage	Normal
Roll damage	100	0	0	0
Inner ring damage	0	100	0	0
Outer ring damage	0	0	100	0
Normal	0	0	0	100

TABLE 11: Confusion matrix of SVD + 1DCNN for classification results of measured signals.

Predicted label	True label (%)			
	Roll damage	Inner ring damage	Outer ring damage	Normal
Roll damage	90	0	0	10
Inner ring damage	0	90	10	0
Outer ring damage	0	10	90	0
Normal	10	0	0	90

the measured signals, the convolution kernel also selectively filters out the noise in the original signals, and as the number of iteration increases, the denoising effect is more significant.

In addition, eSNR of the S2' input of SVD + 1DCNN and SVD-1DCNN is calculated as shown in Figure 15.

As can be seen from Figure 15, for the measured signals, similarly, in the training of the two networks, eSNR of the S2's input increases as the number of iteration increases. At the same number of iteration, eSNR of the S2's input in SVD-1DCNN is higher than that in SVD + 1DCNN. Through the experimental analysis of the measured signals,

it can be said that SVD-1DCNN has stronger feature extraction ability than SVD + 1DCNN. The confusion matrices of SVD-1DCNN and SVD + 1DCNN in this classification process are shown in Tables 10 and 11.

As can be seen from Table 10, SVD-1DCNN can correctly classify each type of measured signals. According to Table 11, for roll damage signal, the classification accuracy of SVD + 1DCNN is 90%, but 10% is classified as normal signal. For inner ring damage signal, the classification accuracy of SVD + 1DCNN is 90%, but 10% is classified as outer ring damage signal. For outer ring damage signal, the

TABLE 12: A_i , \bar{A} , and Var of SVD-1DCNN and SVD + 1DCNN.

	SVD-1DCNN	SVD + 1DCNN
1	100%	92.5%
2	100%	87.5%
3	100%	90%
4	100%	90%
5	100%	90%
6	100%	90%
7	100%	90%
8	100%	90%
9	100%	90%
10	100%	90%
\bar{A}	100%	90%
Var	0	1.25×10^{-4}

classification accuracy of SVD + 1DCNN is 90%, but 10% is classified as inner ring damage signal. For normal signal, the classification accuracy of SVD + 1DCNN is 90%, but 10% is classified as roll damage signal. For measured signals, SVD-1DCNN had a higher classification accuracy than SVD + 1DCNN.

10 experiments were carried out on both networks. Table 12 shows A_i of each experiment, \bar{A} and Var of the two networks after 10 experiments.

As can be seen from Table 12, in the classification of measured signals, both networks have excellent stability, and the experimental results of SVD-1DCNN are better than those of SVD + 1DCNN.

5. Conclusions

This paper proposes a fault diagnosis method based on SVD and 1DCNN, which takes the original signals as input and avoids the loss of feature information. The feasibility of the method is verified by experiments of simulated signals and measured signals. In addition, the role of convolution kernels in feature extraction is also analyzed. The main conclusions can be summarized as follows.

A novel network structure, SVD-1DCNN, is constructed by embedding an SVD layer after the first convolution layer of 1DCNN. In the novel network, the SVD layer denoises and reconstructs the output of the first convolution layer (primary classification feature) to achieve joint optimization of feature extraction and denoising and reconstruction, and the output of the SVD layer is used as input to the next pooling layer. Experiments show that the method has higher pattern recognition accuracy, which shows SVD-1DCNN is more conducive to the accurate diagnosis of bearing faults.

By analyzing the output of the first convolution layer, it is found that the convolution kernels in the network extract different frequency components for different signals and filter out other frequency components. In the training process, the convolution kernel plays the role of extracting features and denoising, and as the number of network training increases, the denoising effect of the convolution kernel is better.

Data Availability

The data used to support the findings of this study are available in [39].

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This research was funded by the National Natural Science Foundation of China (grant number: 51705531).

References

- [1] R. Liu, G. Meng, B. Yang, C. Sun, and X. Chen, "Dislocated time series convolutional neural architecture: an intelligent fault diagnosis approach for electric machine," *IEEE Transactions on Industrial Informatics*, vol. 13, no. 3, pp. 1310–1320, 2017.
- [2] W. Sun, B. Yao, N. Zeng et al., "An intelligent gear fault diagnosis methodology using a complex wavelet enhanced convolutional neural network," *Materials*, vol. 10, no. 7, pp. 790–807, 2017.
- [3] G. Jiang, H. He, J. Yan, and X. Ping, "Multiscale convolutional neural networks for fault diagnosis of wind turbine gearbox," *IEEE Transactions on Industrial Electronics*, vol. 66, no. 4, pp. 3196–3207, 2018.
- [4] N. R. Safin, V. A. Prakht, V. A. Dmitrievskii, and A. A. Dmitrievskii, "Stator current fault diagnosis of induction motor bearings based on the fast Fourier transform," *Russian Electrical Engineering*, vol. 87, no. 12, pp. 661–665, 2016.
- [5] J. Burrielvalencia, R. Puchepanadero, J. Martinezroman, and M. Pinedasanchez, "Fault diagnosis of induction machines in a transient regime using current sensors with an optimized slepian window," *Sensors*, vol. 18, no. 2, pp. 146–169, 2018.
- [6] M. Kang, J. Kim, and J.-M. Kim, "An FPGA-based multicore system for real-time bearing fault diagnosis using ultra-sampling rate AE signals," *IEEE Transactions on Industrial Electronics*, vol. 62, no. 4, pp. 2319–2329, 2015.
- [7] F. K. Choy, W. Jia, and R. Wu, "Identification of bearing and gear tooth damage in a transmission system," *Tribology Transactions*, vol. 52, no. 3, pp. 303–309, 2009.
- [8] Z. Zhi, Y. S. Zhu, Y. Y. Zhang, Y. Xing, and H. H. Shi, "Fault diagnosis of rolling bearings based on EMD Interval-Threshold denoising and maximum likelihood estimation," *Journal of Vibration and Shock*, vol. 32, no. 9, pp. 155–159, 2013.
- [9] K. H. Hui, L. M. Hee, M. S. Leong, and A. M. Abdelrhman, "Time-frequency signal analysis in machinery fault diagnosis: review," *Advanced Materials Research*, vol. 845, pp. 41–45, 2014.
- [10] S. P. Mogal and D. I. Lalwani, "A brief review on fault diagnosis of rotating machineries," *Applied Mechanics and Materials*, vol. 541–542, no. 2, pp. 635–640, 2014.
- [11] Q. Wang, Y. B. Liu, X. He, S. Y. Liu, and J. H. Liu, "Fault diagnosis of bearing based on KPCA and KNN method," *Advanced Materials Research*, vol. 986–987, pp. 1491–1496, 2014.
- [12] S.-W. Fei, "Fault diagnosis of bearing based on wavelet packet transform-phase space reconstruction-singular value

- decomposition and SVM classifier," *Arabian Journal for Science and Engineering*, vol. 42, no. 5, pp. 1967–1975, 2017.
- [13] A. K. Mahamad and T. Hiyama, "Development of ANN for diagnosing induction motor bearing failure," *IEEE Transactions on Industry Applications*, vol. 130, no. 7, pp. 838–846, 2010.
- [14] X. Dai and Z. Gao, "From model, signal to knowledge: a data-driven perspective of fault detection and diagnosis," *IEEE Transactions on Industrial Informatics*, vol. 9, no. 4, pp. 2226–2238, 2013.
- [15] Z. W. Gao, C. Cecati, and S. Ding, "A survey of fault diagnosis and fault-tolerant techniques—part II: fault diagnosis with knowledge-based and hybrid/active approaches," *IEEE Transactions on Industrial Electronics*, vol. 62, no. 6, pp. 3768–3774, 2015.
- [16] K. Noda, Y. Yamaguchi, K. Nakadai, H. G. Okuno, and T. Ogata, "Audio-visual speech recognition using deep learning," *Applied Intelligence*, vol. 42, no. 4, pp. 722–737, 2015.
- [17] S. Nagpal, M. Singh, R. Singh, and M. Vatsa, "Regularized deep learning for face recognition with weight variations," *IEEE Access*, vol. 3, pp. 3010–3018, 2015.
- [18] S. Nie, M. Zheng, and Q. Ji, "The deep regression bayesian network and its applications: probabilistic deep learning for computer vision," *IEEE Signal Processing Magazine*, vol. 35, no. 1, pp. 101–111, 2018.
- [19] T.-H. Chan, K. Jia, S. Gao, J. Lu, Z. Zeng, and Y. Ma, "PCANet: a simple deep learning baseline for image classification?" *IEEE Transactions on Image Processing*, vol. 24, no. 12, pp. 5017–5032, 2015.
- [20] W. Long, X. Li, G. Liang, and Y. Zhang, "A new convolutional neural network based data-driven fault diagnosis method," *IEEE Transactions on Industrial Electronics*, vol. 65, no. 7, pp. 5990–5998, 2018.
- [21] E. Özkural, "The Foundations of Deep Learning with a Path towards General Intelligence," in *Proceedings of the International Conference on Artificial General Intelligence*, pp. 162–173, Prague, Czech Republic, August 2018.
- [22] H. Ren, J.-F. Qu, Y. Chai, Q. Tang, and X. Ye, "Deep learning for fault diagnosis: the state of the art and challenge," *Control and Decision*, vol. 32, no. 8, pp. 1345–1358, 2017.
- [23] B. P. Duong and J. M. Kim, "Non-mutually exclusive deep neural network classifier for combined modes of bearing fault diagnosis," *Sensors*, vol. 18, no. 4, pp. 1129–1143, 2018.
- [24] H. Shao, H. Jiang, H. Zhang, W. Duan, T. Liang, and S. Wu, "Rolling bearing fault feature learning using improved convolutional deep belief network with compressed sensing," *Mechanical Systems and Signal Processing*, vol. 100, pp. 743–765, 2018.
- [25] Z. Chen and W. Li, "Multisensor feature fusion for bearing fault diagnosis using sparse autoencoder and deep belief network," *IEEE Transactions on Instrumentation and Measurement*, vol. 66, no. 7, pp. 1693–1702, 2017.
- [26] W. Lu, X. Wang, C. Yang, and T. Zhang, "A novel feature extraction method using deep neural network for rolling bearing fault diagnosis," in *Proceedings of the 27th Chinese Control and Decision Conference (2015 CCDC)*, Qingdao, China, May 2015.
- [27] H. Shao, H. Jiang, X. Zhang, and M. Niu, "Rolling bearing fault diagnosis using an optimization deep belief network," *Measurement Science and Technology*, vol. 26, no. 11, p. 115002, 2015.
- [28] Y. Lecun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.
- [29] 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, 2018.
- [30] X. Min, L. Teng, X. Lin, L. Liu, and C. W. D. Silva, "Fault diagnosis for rotating machinery using multiple sensors and convolutional neural networks," *IEEE/ASME Transactions on Mechatronics*, vol. 23, no. 1, pp. 101–110, 2017.
- [31] S. Li, G. Liu, X. Tang, J. Lu, and J. Hu, "An ensemble deep convolutional neural network model with improved D-S evidence fusion for bearing fault diagnosis," *Sensors*, vol. 17, no. 8, pp. 1729–1747, 2017.
- [32] S. Guo, T. Yang, W. Gao, and C. Zhang, "A novel fault diagnosis method for rotating machinery based on a convolutional neural network," *Sensors*, vol. 18, no. 5, pp. 1429–1444, 2018.
- [33] B. Vidakovic and C. B. Lozoya, "On time-dependent wavelet denoising," *IEEE Transactions on Signal Processing*, vol. 46, no. 9, pp. 2549–2554, 1998.
- [34] X. Zhao and B. Ye, "Selection of effective singular values using difference spectrum and its application to fault diagnosis of headstock," *Mechanical Systems and Signal Processing*, vol. 25, no. 5, pp. 1617–1631, 2011.
- [35] N. Sawalhi, R. B. Randall, and H. Endo, "The enhancement of fault detection and diagnosis in rolling element bearings using minimum entropy deconvolution combined with spectral kurtosis," *Mechanical Systems and Signal Processing*, vol. 21, no. 6, pp. 2616–2633, 2007.
- [36] H. Jiang, J. Chen, G. Dong, T. Liu, and G. Chen, "Study on Hankel matrix-based SVD and its application in rolling element bearing fault diagnosis," *Mechanical Systems and Signal Processing*, vol. 52–53, no. 1, pp. 338–359, 2015.
- [37] Z. H. Meng and C. Wang, "Application of rolling bearing compound fault diagnosis based on combined SVD," *Manufacturing Automation*, vol. 35, no. 21, pp. 90–92, 2013.
- [38] X. Z. Zhao, B. Y. Ye, and T. J. Tong, "Difference spectrum theory of singular value and its application to the fault diagnosis of headstock of lathe," *Journal of Mechanical Engineering*, vol. 46, no. 1, pp. 100–108, 2010.
- [39] CWRU, *Case Western Reserve University Bearing Data Center Website*, CWRU, Cleveland, OH, USA, 2008, <http://csegroups.case.edu/bearingdatacenter/home>.