IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received April 8, 2019, accepted May 6, 2019, date of publication May 10, 2019, date of current version May 29, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2916024

Convolutional Neural Network Based on Spiral Arrangement of Features and Its Application in Bearing Fault Diagnosis

FENGTAO WANG, (Member, IEEE), GANG DENG, LINJIE MA, XIAOFEI LIU, AND HONGKUN LI

School of Mechanical Engineering, Dalian University of Technology, Dalian 116024, China

Corresponding author: Fengtao Wang (wangft@dlut.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 51875075 and Grant 51375067.

ABSTRACT With the coming of artificial intelligence and the era of big data, convolutional neural network (CNN) has become one of the research hotspots in many scientific fields. However, there exist serious edge information loss problems in the information transmission process of CNN. Therefore, in order to suppress the loss of important information during the information transmission process, new methods are proposed which use the singular value decomposition (SVD) algorithm based on the phase space reconstruction to analyze the bearing vibration signal. Singular values are regarded as the features for evaluating the bearing's health condition. Then, the features radiate from the center around to form a spiral matrix as the input of the CNN, which can effectively resist information loss problems during the information delivery process. In order to verify the performance of the proposed methods better than the conventional ones, experiments are carried out. This paper shows that the proposed methods have excellent performances in the field of bearing fault diagnosis.

INDEX TERMS Vibration signal, singular value decomposition, spiral matrix, convolutional neural network, fault diagnosis.

I. INTRODUCTION

A rolling bearing is the core component of rotating machinery whose running state plays a decisive role in economic development and social security [1]. Therefore, it is of great significance to continuously monitor and diagnose the running state of rolling bearings. With the upgrading of industrial demand, the traditional signal processing methods have been increasingly difficult to meet the new requirements for fault diagnosis and the rapid development of artificial intelligence is leading the fault diagnosis to the intelligent direction. In recent years, intelligent diagnosis models such as extreme learning machine (ELM), support vector machine (SVM) and deep neural network (DNN) have been successfully applied to the rolling bearings' fault diagnosis [2]–[7].

Among many intelligent diagnosis models, DNN has attracted much attention due to its outstanding pattern recognition ability. The DNN commonly used in the field of fault diagnosis is mainly composed of stacked auto-encoder (SAE), deep belief network (DBN) and convolutional neural network (CNN). However, compared with the other two models, CNN is the most complex and widely used. Since the AlexNet model of CNN won the championship of the 2012 ImageNet Large Scale Visual Recognition Challenge (ILSVRC) [8], CNN models such as VGGNet, Google Inception and ResNet have also won this award, which has set off a wave of research on the CNN in the field of image recognition. At the same time, with the coming of artificial intelligence and the era of big data, a new data-driven ''industrial revolution'' is quietly emerging and deep learning algorithms presented by the CNN have been widely applied in the field of rotating machinery fault diagnosis. Jiao et al. [9] proposed a multivariate encoder information based convolutional neural network for intelligent fault diagnosis of planetary gearboxes, which offers a promising tool for intelligent diagnosis of rotating machinery. Appana *et al.* [10] proposed an acoustic emission (AE) analysis-based bearing fault diagnosis invariant under fluctuations of rotational speeds

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

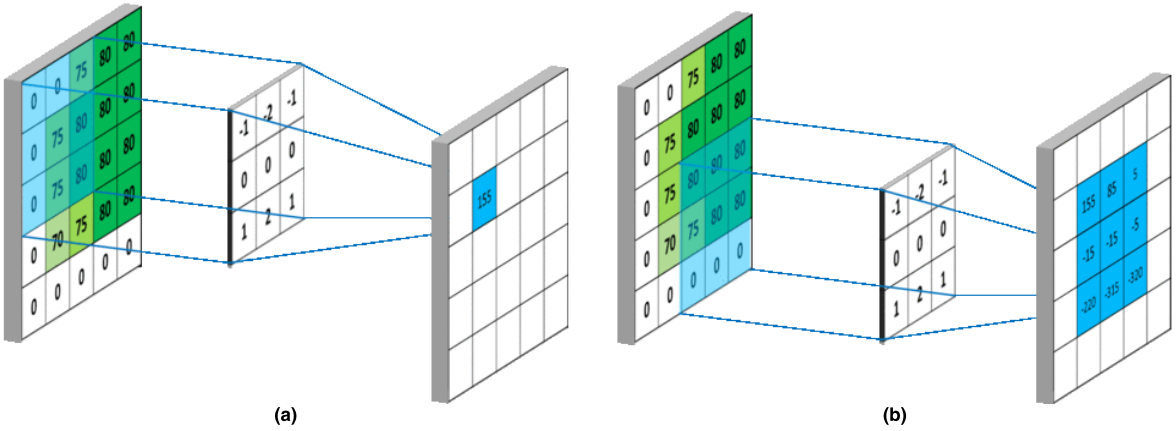

FIGURE 1. The specific process of the convolution operation. (a) Convolution kernel slides to the starting position. (b) Convolution kernel slides to the end position.

using envelope spectrums (ES) and a CNN. They showed that the proposed method leads to highly reliable results. Moreover, Wang *et al.* [11] used the Short-Time Fourier transform (STFT) to pre-process the original signal to obtain the corresponding time-frequency images. Moreover, in order to perform the fault diagnosis of an asynchronous motor, they applied the CNN to adaptively extract time-frequency mapping features.

Although CNN has achieved good results in the fault diagnosis of rotating machinery, it has a problem that cannot be ignored, that is, the loss of edge information. The problem of edge information loss is determined by the operation mechanism of CNN, and the operation mechanism of CNN can be simply described as the following process. Generally, the inputs of the CNN are images and the convolution operation is applied to process pixel values of each image. Taking a single-channel grayscale image as an example, in fact, it is a two-dimensional discrete signal and the convolution operation is to use a convolution kernel to slide on the grayscale image. When the convolution kernel slides to a certain position, pixel values corresponding to this position are multiplied by convolution kernel values and the results are added to obtain the center pixel value of the corresponding position of the output image. The convolution kernel slides in the order from left to right and from top to bottom to complete the calculation of all pixels. Thus, the output of this process is still an image. Figure 1 shows the specific process of the convolution operation.

Figure 1 shows that the input image size and the size of the convolution kernel for the sample image are 5×5 and 3×3, respectively. After completing the convolution operations, the size of the output image is 3×3, which leads to the following problems: First, when the number of CNN layers is too deep, the size difference between input and output images is too large and it is obviously not an ideal result. Second, during the above convolution operations, pixel values of the edge position of the input image are less involved in the operation than those of the center position. For example, pixel values of four vertex positions of the input image participate in only 1 convolution operation, while the pixel value in the middlemost position participates in 9 convolution operations. This causes the loss of information in the edge position of the image and the loss of the edge information accumulatively increases during layer by layer training.

The general solution to the above problems is to perform edge filling processing on the original map (Padding) [12]–[14]. That is, the above input image is expanded to a size of 7×7 to ensure that the final output image size is also 5×5. The commonly used edge filling methods include zero padding, edge copying, edge mirroring and block copying. Although these edge filling processing methods solve the first problem well, the CNN still suffers from loss of the edge information during the information delivery process after padding. It should be indicated that when the edge information is particularly important, the loss is fatal. What's more, no scholars have perfectly solved the problem of edge information loss caused by the operation mechanism of CNN so far.

In the field of rotating machinery fault diagnosis, the inputs of the CNN are usually the time-frequency images of the vibration signals, which are mostly obtained by time-frequency analysis methods such as wavelet transform (WT) and short-time Fourier transform (STFT) [15]–[17]. These time-frequency images not only inevitably suffer from edge information loss in the training process of CNN, but also have excessively large image sizes. The problem is that the time-frequency images with too large size will take up a lot of computing resources, and the image compression processing will also lose some information. In order to solve the above problems, a new image construction method that the extracted features of vibration signals are converted into pixel values to construct the images which are regarded as the inputs of CNN is proposed in this paper. In this way, it is better to extract hundreds of features to form an image of appropriate size while traditional feature extraction methods in time domain, frequency domain

and time-frequency domain are not only tedious, but also difficult to extract so many features. According to the matrix theory, the singular values can reflect the energy distribution of signal and noise, so the singular values obtained by the singular value decomposition (SVD) algorithm can represent the features of the original signal itself. It is worth noting that after the one-dimensional vibration signal is transformed into a two-dimensional matrix by means of phase space reconstruction, the number of obtained singular values is considerable. Moreover, the larger the singular value is, the more important the corresponding feature will be. Therefore, in this paper, the SVD algorithm is first used to extract the low-level features of the vibration signal, and then the grayscale image is constructed from the center to the edge according to the importance of these singular value features which ensures that the pixel values converted by the larger singular values are in the center of the image. As a result, the most important information is concentrated at the center of the image, while the least important features are concentrated at the edge of the image. In this way, the integrity of the important information in the transmission process can be guaranteed, and the impact caused by the loss of edge information can be effectively resisted.

To suppress the loss of important information during the information transfer process, in this paper, a CNN based on spiral arrangement of features is proposed and the phase space reconstruction and singular value decomposition (SVD) algorithms are combined to investigate the accurate diagnosis of the aero-engine inter-shaft bearing.

II. FUNDAMENTAL THEORY

A. SVD BASED ON PHASE SPACE RECONSTRUCTION

The basic idea of the phase space reconstruction is that the evolution of a component in the system is determined by other components that interact with it and the relevant information about these components is implicit in the development of this component. To reconstruct an equivalent state space, we only need to consider one component and treat some of its measurements at fixed-time delay points as new dimensions.

The way how Hankel matrix is constructed is an important phase space reconstruction method [18], which is widely used in the field of signal processing. If there is time-domain signal $x = \{x_1, x_2, \cdots x_N\}$, a p^*q-order Hankel matrix $\mathbf{H}_{p^*q}$ can be constructed as the following:

$$\mathbf{H}_{p^*q} = \begin{pmatrix} x_1 & x_2 & \cdots & x_q \\ x_2 & x_3 & \cdots & x_{q+1} \\ \vdots & \vdots & & \vdots \\ x_p & x_{p+1} & \cdots & x_N \end{pmatrix} \tag{1}$$

Singular value decomposition (SVD) is a commonly used algorithm for mining data features, which is an extension of the eigenvalue decomposition and suitable for arbitrary matrices. Employing the SVD algorithm yields the following expression for $\mathbf{H}_{p^*q}$:

$$\mathbf{H}_{p^*q} = \mathbf{U}_{p^*p}\mathbf{\Sigma}_{p^*q}\mathbf{V}^T_{q^*q} \tag{2}$$

where, $\mathbf{U}$ and $\mathbf{V}$ are orthogonal matrix and $\mathbf{\Sigma}$ is a non-negative diagonal matrix, that is

$$\mathbf{\Sigma}_{p\times q} = \begin{bmatrix} \mathbf{\Delta} & 0 \\ 0 & 0 \end{bmatrix}, \quad \mathbf{\Delta} = diag(\sigma_1, \sigma_2, \cdots \sigma_r),$$
$$\sigma_i = \Sigma(i,i), \quad \sigma_{i-1} \geq \sigma_i, \ i = 1, 2, \cdots r$$

To reach the majority of singular values of the Hankel matrix, it is necessary to maximize the scalar product of the row and column numbers, called p and q, respectively. It should be indicated that the Hankel matrix size varies in accordance with the signal length N:

$$p = \begin{cases} N/2, & N \text{ is an even number} \\ (N+)/2, & N \text{ is an odd number} \end{cases}$$
$$q = N + -p \tag{3}$$

B. CNN

The convolutional neural network (CNN) was originally designed to solve the problem of image recognition and it was proposed by Lecun [19] in 1989. Currently, CNN's applications are not limited to images and video, but can also be used for time-series signals such as audio and vibration signals. The core elements of CNN are local connection, weight sharing, down-sampling and so on. Among them, local connection and weight sharing can greatly reduce the parameter quantity and training complexity, as well as solving the overfitting problem in training. Moreover, the down-sampling of the pooling layer can further reduce the parameter quantity. Furthermore, it is gratifying that down-sampling gives the model a certain degree of deformation tolerance and improves the generalization of the model.

Assuming x_i^l as the input of lth layer in the CNN, the convolution calculation for this layer can be expressed as:

$$x_j^{l+1} = s_f\left(\sum_{i\in M_j} x_i^l * \boldsymbol{k}_{ij}^{l+1} + b_j\right) \tag{4}$$

where $*$, $\boldsymbol{k}_{ij}^{l+1}$, M_j, b_j and s_f are the convolution operator, the weight matrix of the convolution kernel in $l+1th$ layer, a collection of input images, the corresponding bias and the activation function, respectively.

After each layer's convolution operation, it is usually followed by a pooling layer and its main purpose is to compress the images and reduce the parameters without affecting the images' quality by means of down-sampling. In classic CNN models such as AlexNet (see Figure 2) [1], the first few

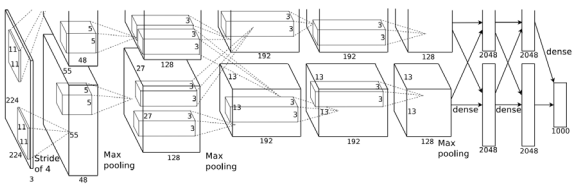

FIGURE 2. AlexNet network structure.

layers of the network are generally designed to be connected alternately between the convolutional layer and the pooling layer. At the end of the network, fully connected layers and a Softmax classifier are used to perform the final classification task. Compared with the conventional machine learning algorithms, CNN does not need to manually extract features. In other words, it can automatically extract features from the input during the training and complete the pattern recognition, which can not only reduce the recognition difficulty but also have higher recognition accuracy.

III. CNN MODEL BASED ON SPIRAL ARRANGEMENT OF SINGULAR VALUES

A. SPIRAL ARRANGEMENT OF SINGULAR VALUES

In this section, it is assumed that the time-domain vibration signal of the inter-shaft bearing is $x = \{x_1, x_2, \cdots x_N\}$, then the p^*q-order Hankel matrix $\mathbf{H}_{p^*q}$ can be constructed by the method of Formula (1). Finally, singular value decomposition of $\mathbf{H}_{p^*q}$ yields a set of singular values of the signal from the largest to the smallest permutations. In order to achieve accurate classification of the bearing's health condition, all the singular values of the signal are arranged as a two-dimensional matrix and converted into a grayscale image, which is the input of the deep convolution neural network. What must be taken into account is that, in the singular value theory of matrices, larger singular values correspond to more important signal features [20]. Therefore, in order to suppress the loss of important information during the information delivery process, it is necessary to avoid large singular values appearing at the edge positions of the matrix when singular values are arranged into a two-dimensional matrix.

The appearance of internal spiral matrix solves the above problems skillfully. The inner spiral matrix refers to a matrix whose elements are arranged in an inner spiral pattern [21]. Equation (5) indicates that elements of the inner spiral matrix increase from the middle to the right, down, left and up directions, respectively.

$$\begin{array}{ccccc} 21 & 22 & 23 & 24 & 25 \\ 20 & 7 & 8 & 9 & 10 \\ 19 & 6 & 1 & 2 & 11 \\ 18 & 5 & 4 & 3 & 12 \\ 17 & 16 & 15 & 14 & 13 \end{array} \tag{5}$$

Except for the inner spiral matrix, the elements of the double spiral matrix also radiate from the center around, which can achieve the same effect. Elements of double spiral matrix are arranged in such a way as the following:

$$\begin{array}{ccccc} 9 & 10 & 11 & 12 & 13 \\ 8 & 3 & 4 & 5 & 6 \\ 7 & 2 & 1 & 2 & 7 \\ 6 & 5 & 4 & 3 & 8 \\ 13 & 12 & 11 & 10 & 9 \end{array} \tag{6}$$

Considering that the singular value sequence obtained by SVD is a one-dimensional array arranged from large to small, if the array elements' indexes are spirally arranged in accordance with Equations (5) or (6), the relatively larger singular values can avoid appearing at the edge of the matrix. This ensures that the pixel values converted by the larger singular values are in the center of the image and the loss of important information during the transmission will be effectively resisted.

After completing the data preprocessing mentioned above, in order to facilitate the subsequent training and testing of CNN model, the matrix composed of singular values is saved as a grayscale image and labeled at the same time.

B. THE PROPOSED CNN MODEL

It is intended to investigate the vibration signal of the aero-engine inter-shaft bearing in the present study. Moreover, fault diagnosis methods of CNN based on spiral arrangement of singular values are proposed. Figure 3 illustrates the proposed fault diagnosis model. It indicates that the model consists of six steps:

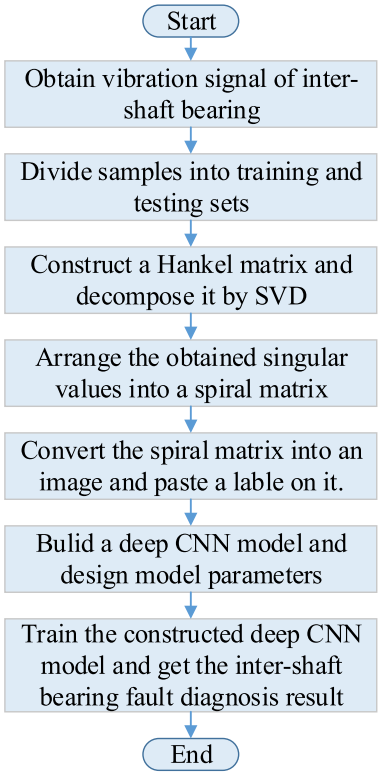

FIGURE 3. Flowchart of the proposed fault diagnosis model.

Step 1: Acquiring original vibration signals of bearings. Then all signals are divided into training and testing signals.

Step 2: Reconstruction of the original signal into a Hankel matrix. Then the matrix size is determined.

Step 3: Applying the SVD algorithm to the constructed Hankel matrix.

Step 4: Arranging obtained singular values into an inner spiral matrix or a double spiral matrix, converting the spiral

matrix into a grayscale image to save and attaching a label at the same time.

Step 5: Building a deep CNN model and designing model parameters.

Step 6: Training the established deep CNN model and testing in batches to obtain diagnostic accuracy.

IV. EXPERIMENTAL VERIFICATION

A. INTRODUCTION OF TEST RIG AND DATA SETS DESCRIPTION

In order to validate the efficiency of the proposed fault diagnosis model, experiments are carried out in the present study. The test setup consists of a double rotor aero-engine with inter-shaft bearing, two motors, three mass disks and four accelerometers. Figure 4 illustrates the configuration of the test setup.

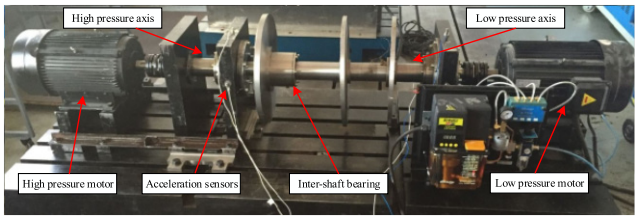

FIGURE 4. The configuration of test setup.

In order to investigate the inter-shaft bearing, four acceleration sensors are installed on the bearing pedestal to gather vibration signals of the bearing, which is the equipment basis for the following 4 groups of parallel experiments carried out in this paper. As we can see from Figure 4 that the bearing is installed between low and high pressure axes to connect motors and it should be indicated that the sampling frequency for all sensors is set to 25.6 KHz.

In this case, data sets contain 10 kinds of operating conditions, which point to the corresponding data of label 1∼10 in Table 1. The faults are introduced to the inter-shaft bearing under the running conditions of high-pressure-motor single rotation (HR), low-pressure-motor single rotation (LR), and high-pressure-motor/low-pressure-motor relative rotation (HLR), respectively. Besides, a normal condition of a two-motor relative rotation is introduced. The rotation speed of the motors is 1200 RPM under each operating condition. The fault grooves in the bearing are machined by an electric spark, as shown in Figure 5. The outer ring fault is displayed by taking one of roller elements because the outer race and the holder cannot be removed.

TABLE 1. Description of operation conditions.

	bearing operating condition	*Motor speed (Hz) HR/LR*	*Size of training/ testing samples*	*Label*
ILR	Inner fault (LR)	0/20	300/100	1
IHR	Inner fault (HR)	20/0	300/100	2
IHLR	Inner fault (HLR)	20/20	300/100	3
OLR	Outer fault (LR)	0/20	300/100	4
OHR	Outer fault (HR)	20/0	300/100	5
OHLR	Outer fault (HLR)	20/20	300/100	6
RLR	Roller fault (LR)	0/20	300/100	7
RHR	Roller fault (HR)	20/0	300/100	8
RHLR	Roller fault (HLR)	20/20	300/100	9
NHLR	Normal (HLR)	20/20	300/100	10

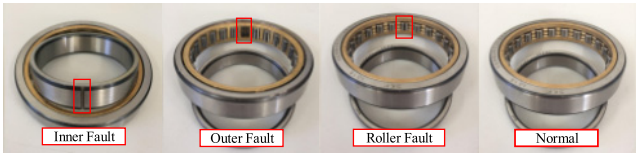

FIGURE 5. Configuration of machined faults on the bearing.

In order to verify the effectiveness of the proposed methods, four groups of parallel experiments were conducted in this paper, which were respectively denoted as Group 1, Group 2, Group 3 and Group 4. In each group, 400 samples are collected for each operating condition, where 300 samples are used for training and 100 samples are remained for testing. Table 1 shows the detailed information of the used data sets. In Group 1, the data sets were collected from the acceleration sensor in the X direction of the high pressure axis or the low pressure axis, and in Group 2, the vibration signals were collected from the acceleration sensor in the Y direction of the high pressure axis or the low pressure axis. In Group 3 and Group 4, the data sets were obtained from the X and Y directions of the bearing housing, respectively.

B. DATA PROCESSING AND COMPARISON EXPERIMENTS

In this study, the sampling length for all operating conditions is set to 2000. Therefore, after each sample is processed by the proposed SVD algorithm, 1000 singular values, arranged in descending order can be obtained. If these singular values are spirally arranged into an inner spiral matrix or a double spiral matrix, the corresponding matrix of $31 \times 31 (31^2 < 1000 < 32^2)$ or 44×44 $((44^2 + 1)/2 < 1000 < (45^2 + 1)/2)$ size is available and the matrix can be converted to a grayscale image by processing each element of the matrix to a pixel value between 0 and 255 and labeled to facilitate subsequent processing. In this study, the image processing method of double-cubic interpolation is applied to reconstruct the saved 31×31 grayscale image, dedicates to get an image size of 32×32 as the input of CNN, which is convenient for data processing.

In order to highlight the advantages of proposed methods over conventional methods, 4 comparison experiments have been carried out in this part. Comparison experiment 1 is performed to arrange the obtained singular values into a two-dimensional matrix whose elements decrease from left to right and from up to down and convert it to a 32×32-size image as the input of the CNN. Comparison experiment 2 is performed to treat the obtained 1000 singular

values as the input of the stacked auto-encoder (SAE) and then identify the inter-shaft bearing vibration signal. In comparison experiment 3, the obtained 1000 singular values were firstly reduced in dimension by principal component analysis (PCA) algorithm, and then the low-dimensional samples were input into the support vector machine (SVM) model to complete the intelligent diagnosis of the inter-shaft bearing. In comparison experiment 4, the traditional artificial neural network (ANN) algorithm is used to classify the singular value features of vibration signals. In order to describe the experiments process more clearly, the proposed fault diagnosis methods are recorded as Method 1 and Method 2, corresponding to the CNN's input constructed by the method of double spiral matrix and inter spiral matrix, respectively. The fault diagnosis methods corresponding to comparison experiments 1∼4 are recorded as Method 3, Method 4, Method 5 and Method 6, respectively. All the above experimental methods used in this paper and corresponding descriptions are shown in Table 2.

TABLE 2. Experimental methods and corresponding description.

Methods	*Methods description*
Method 1 (proposed method)	Arrange the obtained singular values into a double spiral matrix and convert the double spiral matrix into a 44×44-size grayscale image, which is regarded as the input of CNN.
Method 2 (proposed method)	Arrange the obtained singular values into an inner spiral matrix and convert the inner spiral matrix into a 32×32-size grayscale image, which is regarded as the input of CNN.
Method 3 (comparison experiment 1)	Arrange the obtained singular values into a two-dimensional matrix whose elements decrease from left to right and from up to down and convert the matrix into a 32×32-size grayscale image, which is regarded as the input of CNN.
Method 4 (comparison experiment 2)	After the vibration signal is processed by SVD algorithm, treat the obtained 1000 singular values as the input of the stacked auto-encoder (SAE) and then identify the health condition of the inter-shaft bearing.
Method 5 (comparison experiment 3)	The PCA algorithm is firstly used to reduce the dimension of the obtained 1000 singular values, and then the low-dimensional samples were input into the SVM model to complete the intelligent diagnosis of the inter-shaft bearing.
Method 6 (comparison experiment 4)	Use the traditional ANN algorithm to classify the singular value features of vibration signals.

Figure 6 illustrates the network structure and parameters of Method 1, Method 2 and Method 3, which indicates that after passing through the input layer, convolution layers (C), Max-pooling layers (S) and fully connected layers (F), the model automatically learns the grayscale features of singular values and sends them to the Softmax classifier. At the input layer, the image size for Method 1 is $44 \times 44 \times 1$, while the image size for Methods 2 and 3 are $32 \times 32 \times 1$, where 1 denotes the number of color channels, indicating that the input image is a grayscale image. In all convolution operations, the convolution kernel has a common size of 3×3 and all convolution kernels have a same stride of 1×1 with the boundary processing way of 'SAME'. After completing the convolution calculation, the *Relu* activation function is applied to improve the training speed and the accuracy of the model. It should be indicated that a 2×2 filter with a Max-pooling strides of 2×2 is selected in all pooling layers of the present study. After all convolution and Max-pooling layers, all features of Max-pooling layer S6 are taken as inputs to fully connected layer F7 so that the first fully connected layer has 2304 or 1024 nodes. The number of nodes in the second fully connected layer F8 is 200 and the number of nodes in the output layer is 10, corresponding to the health status of inter-shaft bearings for 10 different operating states.

In Method 4, the SAE contains three hidden layers and the network structure parameters are: 1000-400-80-10. Considering the dimension of sample singular values, number of input layer neurons is set to 1000. On the other hand, since there are 10 kinds of fault signals, the number of output layer neurons is set to 10. The learning rate of each AE, Softmax classifier and the fine-tuning process are 0.3, 0.3, 0.3, 2.0 and 1.2, respectively and the momentum term is 0.5.

In Method 5, the low-level 1000-dimensional singular value features are firstly reduced to 50 by PCA algorithm, then the relatively advanced 50-dimensional features are used as the input of the SVM model to complete the training process, and finally the intelligent diagnosis of the inter-shaft bearing is realized.

Method 6 still uses SVD algorithm based on phase space reconstruction for signal preprocessing. The obtained singular values are used as the inputs of the traditional ANN, and the number of hidden layer nodes of ANN is set to 60.

C. EXPERIMENTS ANALYSIS

If the diagnostic accuracy of the bearing is higher, the fault information extracted from the signal will be more comprehensive. In other words, for the CNN methods used in this paper, the network structure corresponding to the higher fault diagnosis accuracy can better inhibit the edge information loss of CNN.

After selecting the gradient descent algorithm to perform 10000 optimization iterations on the model, the diagnosis accuracy in 4 groups of experiments of each method is shown in Figure 7.

Figure 7 shows that the diagnosis accuracy of Method 1 and Method 2 achieved in 4 groups of experiments ranged from 96.1% to 97.9%, indicating that the proposed methods can accurately identify 10 kinds of health status of inter-shaft

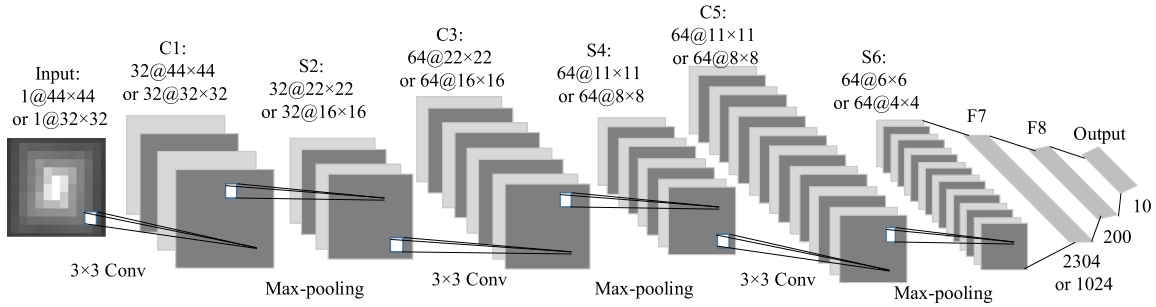

FIGURE 6. CNN structure and parameters of method 1-3.

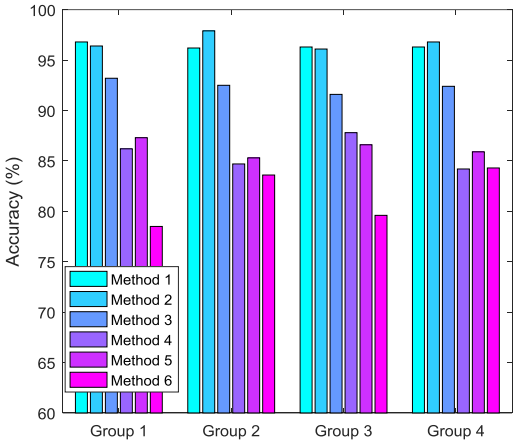

FIGURE 7. Each method's diagnosis accuracy in 4 groups.

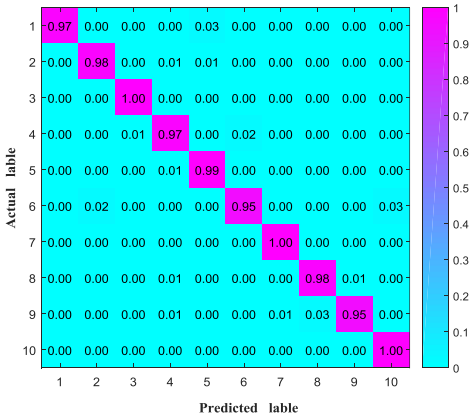

FIGURE 8. The multi-class confusion matrix for the method 2 in group 2.

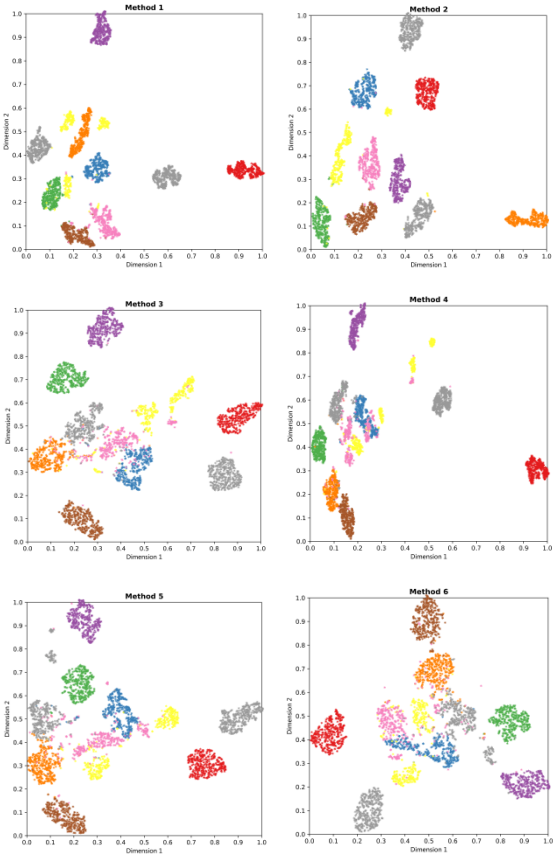

FIGURE 9. Feature space of each method using t-SNE.

bearings for three different working conditions and four different fault positions. However, since the most important singular values are at the top of the image in Method 3 (Comparison experiment 1), this leads to a certain degree of edge information loss in the information transmission process of the CNN, causing the reduction of the average diagnostic accuracy to 92.4%. The average classification accuracy of Method 4, Method 5 and Method 6 are all low than 88%, indicating that the CNN has better classification ability than the SAE, SVM and ANN for the extracted singular value features. Thus, the above results show that the proposed methods can effectively suppress information loss problems during the information delivery process and achieve a very good classification accuracy in classifying faults of vibration signals for the aero-engine inter-shaft bearing.

Experiments carried out in the present study show that the highest accuracy is achieved for Group 2 of Method 2. Among all 1000 test samples, only 21 samples are misjudged and the average diagnosis accuracy reaches to 97.9%. Figure 8 shows the multi-class confusion matrix for the Method 2 in Group 2.

To further verify the ability of the proposed algorithm to resist the edge information loss of CNN, in this section, t-SNE clustering algorithm is used to analyze the last layer features extracted from Method 1-6 in Group 2. The better the clustering effect of a method, the fault information extracted

by the method will be more comprehensive, and if the method is based on CNN, it is also proved that this method can effectively suppress the impact of edge information loss of CNN.

Figure 9 shows the clustering effect of each method using t-SNE algorithm. It can be seen from figures of Method 1 and Method 2 that the bearing signals of the same kind can be clustered together with an obvious clustering center, and the bearing signals of different classes can be separated effectively. In figures of Method 3∼6, because the edge information loss of CNN is not suppressed or the CNN is not used for fault diagnosis at all, the points are aggregated not so well, and the different classes of features are aliased together, which results in an inability to effectively identify.

Considering analysis and discussions, it is concluded that the proposed methods are pretty accurate and effective for fault diagnosis of the aero-engine inter-shaft bearing in all operating conditions. Moreover, it is found that applying proposed methods can effectively resist information loss problems during the information delivery process.

V. CONCLUSION

In the present study, novel methods are proposed for the bearing fault diagnosis by combining SVD and spiral arrangement of features with the deep CNN. The proposed methods can be divided into three major steps: Firstly, the SVD algorithm based on the phase space reconstruction is used to analyze the bearing vibration signal. Subsequently, the obtained singular values are used as features to evaluate the bearing health status and arrange these features into an inner spiral matrix or a double spiral matrix. It converts the spiral matrix into a grayscale image and attaches a label at the same time. Finally, the CNN model is built and the diagnosis accuracy is calculated to verify the efficiency of proposed methods. The present study shows that the proposed methods have the following characteristics:

(1) Highly innovative algorithm. The proposed methods combine the SVD and spiral arrangement of features with the deep CNN, which can effectively resist information loss problems during the information delivery process. This is nearly unprecedented in the field of deep learning and fault diagnosis.

(2) Perfect diagnosis accuracy. Compared with other methods in comparison experiments, the proposed fault diagnosis methods achieve a high recognition accuracy.

(3) Excellent clustering effect. The proposed methods can effectively extract the features from different signals, which have an obvious clustering center after carrying on t-SNE algorithm.

REFERENCES

[1] L. Cui, N. Wu, C. Ma, and H. Wang, "Quantitative fault analysis of roller bearings based on a novel matching pursuit method with a new step-impulse dictionary," *Mech. Syst. Signal Process.*, vols. 68–69, pp. 34–43, Feb. 2016.

[2] M. Luo, C. Li, X. Zhang, R. Li, and X. An, "Compound feature selection and parameter optimization of ELM for fault diagnosis of rolling element bearings," *ISA Trans.*, vol. 65, pp. 556–566, Nov. 2016.

[3] X. Yan and M. Jia, "A novel optimized SVM classification algorithm with multi-domain feature and its application to fault diagnosis of rolling bearing," *Neurocomputing*, vol. 313, pp. 47–64, Nov. 2018.

[4] X. Liu, L. Bo, and H. Luo, "Bearing faults diagnostics based on hybrid LS-SVM and EMD method," *Measurement*, vol. 59, pp. 145–166, Jan. 2015.

[5] J. Xie, G. Du, C. Shen, N. Chen, L. Chen, and Z. Zhu, "An end-to-end model based on improved adaptive deep belief network and its application to bearing fault diagnosis," *IEEE Access*, vol. 6, pp. 63584–63596, 2018.

[6] X. Guo, L. Chen, and C. Shen, "Hierarchical adaptive deep convolution neural network and its application to bearing fault diagnosis," *Measurement*, vol. 93, pp. 490–502, Nov. 2016.

[7] M. Gan, C. Wang, and C. Zhu, "Construction of hierarchical diagnosis network based on deep learning and its application in the fault pattern recognition of rolling element bearings," *Mech. Syst. Signal Process.*, vols. 72–73, pp. 92–104, May 2016.

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

[9] J. Jiao, M. Zhao, J. Lin, and J. Zhao, "A multivariate encoder information based convolutional neural network for intelligent fault diagnosis of planetary gearboxes," *Knowl.-Based Syst.*, vol. 160, pp. 237–250, Nov. 2018.

[10] D. K. Appana, A. Prosvirin, and J.-M. Kim, "Reliable fault diagnosis of bearings with varying rotational speeds using envelope spectrum and convolution neural networks," *Soft Comput.*, vol. 22, no. 20, pp. 6719–6729, 2018.

[11] L. H. Wang, X. P. Zhao, J. X. Wu, Y. Y. Xie, and Y. H. Zhang, "Motor fault diagnosis based on short-time Fourier transform and convolutional neural network," *Chin. J. Mech. Eng.*, vol. 30, pp. 1357–1368, Nov. 2017.

[12] V. K. Gupta and N. Ramani, "A note on convolution and padding for two-dimensional data," *Geophys. Prospecting*, vol. 26, no. 1, pp. 214–217, 2010.

[13] X. Ouyang, X. Ouyang, and P. Zhou, "Spatial pyramid pooling mechanism in 3D convolutional network for sentence-level classification," *IEEE/ACM Trans. Audio, Speech Lang. Process.*, vol. 26, no. 11, pp. 2167–2179, Nov. 2018.

[14] H. Li and H. Zhang, "Research and implementation of convolution boundary expansion," (in Chinese), *Microcomput. Appl.*, vol. 34, no. 10, pp. 47–49, 2018.

[15] D.-T. Hoang and H.-J. Kang, "A survey on Deep Learning based bearing fault diagnosis," *Neurocomputing*, vol. 335, pp. 327–335, Mar. 2019.

[16] O. Janssens *et al.*, "Convolutional neural network based fault detection for rotating machinery," *J. Sound Vib.*, vol. 377, pp. 331–345, Sep. 2016.

[17] X. Ding and Q. He, "Energy-fluctuated multiscale feature learning with deep convnet for intelligent spindle bearing fault diagnosis," *IEEE Trans. Instrum. Meas.*, vol. 66, no. 8, pp. 1926–1935, Aug. 2017.

[18] F. Wang, G. Deng, C. Liu, W. Su, Q. Han, and H. Li, "A deep feature extraction method for bearing fault diagnosis based on empirical mode decomposition and kernel function," *Adv. Mech. Eng.*, vol. 10, no. 9, pp. 1–12, 2018.

[19] Y. LeCun *et al.*, "Backpropagation applied to handwritten zip code recognition," *Neural Comput.*, vol. 1, no. 4, pp. 541–551, 1989.

[20] S. Chatterjee, "Matrix estimation by universal singular value thresholding," *Ann. Statist.*, vol. 43, no. 1, pp. 177–214, 2014.

[21] S. Hamdi, A. B. Abdallah, and M. H. Bedoui, "Scintigraphic image segmentation based on grammatical inference and spiral matrix," in *Proc. 6th Int. Conf. Soft Comput. Pattern Recognit.*, Tunis, Tunisia, Aug. 2014, pp. 186–190.

FENGTAO WANG received the M.S. degree in mechanical engineering from Jilin University of Technology, Jilin, China, in 2000, and the Ph.D. degree in mechanical engineering from the Dalian University of Technology, Dalian, China, in 2003, where he is currently an Associate Professor.

His current research interests include signal processing, machine learning, metal-based additive manufacturing process monitoring and defect identification, and rotating machinery fault diagnosis.

Dr. Wang has received the Second Prize of the National Scientific and Technological Progress Award as the first accomplisher of the Dalian University of Technology.

GANG DENG received the B.S. degree in mechanical design, manufacturing, and automation from the Zhejiang University of Science and Technology, Hangzhou, China, in 2016. He is currently pursuing the M.S. degree with the School of Mechanical Engineering, Dalian University of Technology, China. His current research interests include computer vision, signal processing, metal-based additive manufacturing process monitoring and defect identification, and fault diagnosis using deep learning methods.

LINJIE MA received the B.S. degree in mechanical engineering from the Qilu University of Technology, Jinan, China, in 2017. He is currently pursuing the master's degree with the School of Mechanical Engineering, Dalian University of Technology, China. His current research interests include rotating machinery condition monitoring, signal processing, and fault diagnosis using deep learning methods.

XIAOFEI LIU received the B.S. degree in mechanical engineering from the Nanjing University of Science and Technology, China, in 2016.

His current research interests include mechanical equipment condition monitoring, rotating machinery fault diagnosis, and remaining useful life prediction. He has received the National Scholarship, in 2018.

HONGKUN LI received the M.S. degree in energy and power engineering and the Ph.D. degree in mechanical engineering from the Dalian University of Technology, Dalian, China, in 2000 and 2003, respectively, where he is currently a Professor.

His current research interests include signal processing, dynamic analysis, and the fault diagnosis of mechanical equipments.