

Convolutional neural network intelligent fault diagnosis method for rotating machinery based on discriminant correlation analysis multi-domain feature fusion strategy

Guisheng Lan¹, Haibo Shi²

GuangDong Yue Guang Water Supply Company Limited, Shenzhen, 518002, China

¹Corresponding author

E-mail: ¹3287563626@qq.com, ²349701180@qq.com

Received 5 September 2023; accepted 3 January 2024; published online 18 February 2024

DOI <https://doi.org/10.21595/jve.2024.23612>



Copyright © 2024 Guisheng Lan, 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.

Abstract. Aiming at solving the problems of limited training data, single input information, and limited diagnostic accuracy under the influence of strong background noise in fault diagnosis of rotating machinery, this paper proposes a fault diagnosis method based on the combination of discriminant correlation analysis (DCA) and convolutional neural network (CNN). Firstly, the original vibration signal is divided into several segments in the time domain, and the training data is directly processed by one CNN branch to extract multi-scale time domain features. Simultaneously, the divided data is subjected to discrete wavelet transform (DWT), and processed by another branch of CNN to extract multi-scale time-frequency features. Then, the DCA feature fusion mechanism is adopted to fuse the two-domain features extracted in the parallel branches to improve the model's detection ability. Finally, the fused features are input into the deep CNN for training and learning to extract new features and output the classification results. Through the experimental analysis of two different types of data, the results show that the proposed method can be used for fault diagnosis of rotating machinery effectively. Compared with the single CNN network, the proposed method combines the multi-domain multi-scale feature extraction module with the DCA feature fusion module to enrich the feature information extraction ability. At the same time, the network performance is improved to get higher fault classification accuracy higher.

Keywords: discriminant correlation analysis, convolutional neural network, feature fusion, fault diagnosis.

1. Introduction

With the development and progress of science, the key machines of modern industry are moving towards automation and intelligent application gradually. Bearings and gears are the key supported components of mechanical equipment, and their operating status affects the performance of the entire mechanical equipment. A small fault defect may lead to disastrous consequences. Therefore, fault analysis of the key components of rotating machinery has important practical significance for timely detection of faults to ensure the normal and healthy operation of mechanical equipment [1].

The traditional fault diagnosis method includes three parts: signal acquisition, feature extraction and pattern recognition, among which the most critical part is feature extraction and pattern recognition [2]. The collected vibration signals is first processed by signal processing technology to extract features. Then, pattern recognition algorithms are applied on the extracted features for fault diagnosis. The most commonly used pattern recognition algorithms are some shallow learning models such as BP neural network [3], k-nearest neighbor (KNN) [4] and support vector machine (SVM) [5]. The diagnostic performance of these shallow learning models largely depends on the extracted feature information. At present, the most popular feature extraction method is to construct features manually [6], which requires people to select useful features

containing fault information based on prior knowledge. Most of them rely on advanced signal processing technology, which not only consumes a lot of time, but also may not be adaptable to the change of working conditions and environments. Furthermore, the shallow structure of these learned models limits their ability to learn the complex nonlinear relationships between fault features and patterns [7].

In response to this situation, deep learning methods have been introduced into the field of fault diagnosis gradually in recent years. Deep learning algorithms have powerful feature extraction capabilities. By building a multi-layer network structure and multiple nonlinear transformations, they can extract deep features from raw data directly and adaptively. Therefore, they do not require much prior knowledge or signal processing techniques [8]. Currently, various deep learning models such as Deep Belief Network (DBN) [9], Stacked Autoencoder (SAE) [10] and CNN [11] have been applied on mechanical fault diagnosis successfully. Compared with the shallow learning model, deep learning shows better performance. As the typical representative method of deep learning, CNN has achieved excellent performance in fault diagnosis. Liu et al. proposed a fault diagnosis method based on variational mode decomposition and CNN [12]. A fault diagnosis model based on LeNet5 is proposed [13], and high-precision fault classification is accomplished by transforming the one-dimensional fault signal of a rotating electrical machine into a two-dimensional image. Shao et al. [14] improved the accuracy of CNN in diagnosing bearing and gearbox systems with multiple sensors. By integrating support vector machine (SVM) with deep convolutional neural network (DCNN), a DCNN-SVM network model was proposed [15], which improved the accuracy of fault classification and recognition, and the convergence speed and generalization ability of the model were also improved significantly.

The above studies have achieved good results due to their excellent feature extraction network. However, CNN still has some shortcomings. The shallow learning model does not require a large amount of data for classification. In contrast, the high classification accuracy of CNN relies on a large number of training samples, because the mathematical model of its network is complex, and more samples are needed to increase generalization capabilities and prevent model overfitting. Unfortunately, most of the current fault diagnosis methods only use single-sensor measurement data. The actual working environment of rotating machinery is complex usually, and the amount of data measured by a single sensor is usually small, and the contained fault information may be lost due to external interference, which may lead to the problem of low diagnosis accuracy. Therefore, some methods are needed to realize the fusion of multiple features to obtain the global information of the original signal and reduce the possibility of loss of fault information. In fact, vibration signals can be represented in different domains such as time domain, frequency domain and wavelet domain, and different domains lead to different sensitivities to failure modes. So feature fusion based on multi-domain information can be performed to enhance diagnostic performance, which may be more effective than only depending on network alone in some cases [16]. Wang et al. [17] proposed a new method for fault identification of rotating machinery based on multi-vibration signal fusion and bottleneck layer optimized convolutional neural network (MBCNN). Ding et al. [18] proposed a multi-scale feature mining method for spindle bearing energy fluctuations based on wavelet packet energy images and CNN. Chen et al. [19] proposed a technique to fuse time-domain and frequency-domain features of multi-sensor data through multiple two-layer sparse autoencoder neural networks, and to identify machine operations through deep belief networks. Demetgul et al. [20] proposed a multi-purpose fault detection method, which uses DM, LLE and AE dimensionality reduction method together. Xue et al. [21] used 1D-CNN and 2D-CNN parallel multi-channel structure to extract deep features, and then used feature fusion strategy to realize the fault diagnosis of rolling bearings. Although the above various methods have certain effects, there are still some problems. On the one hand, for multiscale feature extraction from a single domain, it may happen that the global information of the original signal cannot be obtained accurately due to different fault types. In fact, the time domain, frequency domain and time-frequency domain contain features of the vibration signal in different aspects, which are helpful to improve the diagnostic accuracy. On the other hand, the mapping

between different signals and fault types is complex, and common feature fusion strategies may lead to loss of fault information in practical application.

This paper proposes a fault diagnosis method based on the combination of DCA with CNN. The proposed model mainly consists of three parts: multi-domain feature extraction module, multiscale feature extraction module and DCA feature fusion module. First, a branch of CNN is used to extract the multiscale time-domain feature information from the original data directly. At the same time, the original signal is processed by discrete wavelet transform (DWT) and another branch CNN to extract multiscale time-frequency feature information. Then, the DCA feature fusion strategy is used to fuse the time domain and frequency domain features to improve the detection ability of the model. Finally, the fused training features are further learned and trained by CNN to extract new advanced features and output the classification results finally. The effectiveness of the method is verified by two different sets of data. In addition, noise interference experiment is also carried out to verify the superiority of the proposed method.

2. Discrete wavelet transform

The essence of wavelet transform is to obtain a series of wavelet functions by performing multiple different stretching and translation transformations on a mother wavelet function, so as to decompose the original signal into the superposition of these wavelet functions. It can also obtain the signal' time and frequency information while obtaining the local characteristics of the signal. Wavelet transform is divided into continuous wavelet transform (CWT) [22], discrete wavelet transform (DWT) [23] and so on.

A function cluster $\phi_{(a,b)}(t)$ could be obtained by stretching and translating the mother wavelet function $\phi(t)$, which is called wavelet basis function, and can be expressed by the following formula:

$$\phi_{(a,b)}(t) = \frac{1}{\sqrt{|a|}} \phi\left(\frac{t-b}{a}\right), \quad (1)$$

where, a is the scaling factor and b is the translation factor.

The expression of continuous wavelet transform is:

$$WT_{x(t)}(a, b) = \langle x(t), \phi_{(a,b)}(t) \rangle = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{+\infty} x(t) \phi^*\left(\frac{t-b}{a}\right) dt, \quad (2)$$

where $x(t)$ represents the input signal, $\langle \cdot \rangle$ represents the inner product operation, and ϕ^* represents the conjugate function of ϕ .

As an important time-frequency analysis method, DWT is obtained by discretization on the basis of CWT, and binary discretization is usually selected. Its essence is the binary discretization scale factor a and translation factor b , in which $a = 2^j$, $b = 2^j k$ ($j, k \in Z$). The expression of DWT is as following:

$$WT_{x(t)}(a, b) = \frac{1}{\sqrt{|2^j|}} \int_{-\infty}^{+\infty} x(t) \phi^*\left(\frac{t-2^j k}{2^j}\right) dt. \quad (3)$$

Compared with CWT, the signal after DWT not only has the characteristics of no redundant decomposition and accurate reconstruction, but also can show the time-frequency characteristics of the fault fully. At the same time, the calculation time is also reduced greatly [24].

Multi-resolution analysis (MRA) is a stepwise analysis method proposed by Mallat and Meyer in 1986, also known as multiscale analysis. Wavelet decomposition is a multi-resolution analysis

process. By studying the multi-resolution representation of signals from the perspective of function space, MRA not only provides a simple method to construct orthogonal wavelet basis, but also provides a theoretical basis for the fast algorithm of orthogonal wavelet transform. The core of MRA is to decompose the signal hierarchically [25]. The wavelet function can form a multi-resolution function space through stretching and translation, and then project on this function space to form a multi-scale analysis of the signal. This process can be expressed as Eq. (4):

$$f(t) = \sum_b a_0(b)\phi(t-b) + \sum_b \sum_{j=0}^{j-1} d_j(b)\frac{1}{2^j}\varphi(2^j t - b), \quad (4)$$

where b is the displacement coefficient, a_0 is the scale coefficient at 0 scale, d_j is the wavelet coefficient at scale j , ϕ and φ are scale function and wavelet function.

Through multi-resolution analysis, DWT can decompose the acquired original signal into low-frequency approximation signal (CA) and high-frequency detail signal (CD) by using scaling function and wavelet function. The low-frequency approximation signal is also called scale coefficient, and the high-frequency detail signal is the wavelet coefficient. The decomposed approximation signal is then decomposed on the scale, and the decomposed signal is decomposed into a finite layer by iterating the decomposing process.

3. Discriminant correlation analysis

Information fusion can usually be divided into data layer fusion, feature layer fusion and decision layer fusion. Among them, the feature layer fusion is mainly to calculate and process the multi class feature vectors extracted from the original data, so as to realize information fusion. At present, the most classic feature fusion methods based on deep learning models are point-by-point addition (ADD) and vector concatenation (Con-cat) algorithms. ADD could reflect some characteristics of the original features through reducing the parameters and the amount of calculation, but this operation will lead to the loss of some useful information of the original features. The latter concatenates the feature vectors extracted by the network model through the con-cat operation to generate new features directly, and let the network learn without losing information during this process. Although this approach is relatively simple, the two eigenvectors generate redundant information due to their weak correlation, which brings unnecessary increasing in parameters, thus exerting invisible pressure on the network. Therefore, it is not only necessary to analyze the connection and difference between categories while performing feature fusion, but also to consider saving time to improve the performance of the algorithm. To this end, the feature fusion strategy of DCA [26] is introduced. This strategy is improved based on the basis of CCA [27]. While reducing the correlation between features, the redundant correlation between different categories of features is also reduced, so that the feature information extracted in different modes has a better fusion effect to enrich the feature information.

The schematic diagram of the DCA-based feature fusion method is shown in Fig. 1. It is assumed that the extracted two sets of feature matrices are X and Y respectively, and the high level fused feature is Z .

First, according to the Eqs. (5) and (6), the average value of the feature vectors within the class \bar{x}_i and the average value of the feature vectors between the classes \bar{x} are calculated:

$$\bar{x}_i = \frac{1}{n} \sum_{j=1}^n x_{ij}, \quad (5)$$

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n n_i \bar{x}_i, \tag{6}$$

where $x_{ij} \in X$ is the j sample of the i class, $n = \sum_{i=1}^N n_i$, n_i is the number of samples of the i class, and N is the number of categories.

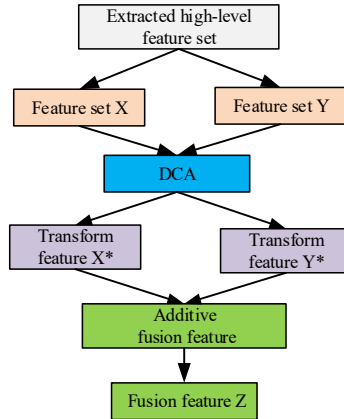


Fig. 1. Schematic diagram of feature fusion based on DCA

The scatter matrix S_{bx} measuring the relationship between different feature classes can be calculated through \bar{x}_i and \bar{x} , which is shown in Eq. (7):

$$S_{bx} = \sum_{i=1}^N n_i (\bar{x}_i - \bar{x})(\bar{x}_i - \bar{x})^T = \Phi_{bx} \Phi_{bx}^T, \tag{7}$$

$$\Phi_{bx} = [\sqrt{n_1}(\bar{x}_1 - \bar{x}), \sqrt{n_2}(\bar{x}_2 - \bar{x}), \dots, \sqrt{n_N}(\bar{x}_N - \bar{x})], \tag{8}$$

where, the matrix Φ_{bx} is the degree of difference between different categories of features in X , and $\Phi_{bx}^T \Phi_{bx}$ is the covariance matrix, which is a symmetrical diagonal matrix. Diagonalizing $\Phi_{bx}^T \Phi_{bx}$ can make the different categories better separated, that is, $Q^T (\Phi_{bx}^T \Phi_{bx}) Q = \Lambda$, where Q is an orthogonal eigenvector matrix, and Λ is a diagonal matrix of eigenvalues in descending order. The largest non-zero r features can be obtained as following:

$$(\Phi_{bx} \varphi)^T S_{bx} (\Phi_{bx} \varphi) = \Lambda_{r \times r}. \tag{9}$$

Define the transformation matrix $W_{bx} = \Phi_{bx} \varphi \Lambda^{-1/2}$ to unitize S_{bx} , that is, $W_{bx}^T S_{bx} W_{bx} = I$, where I is the inter-class scatter matrix after transformation and dimensionality reduction. The dimensionality of the feature matrix X can be reduced from p to r by W_{bx} , as shown in Eq. (10), and this process can reduce the connection between different categories greatly in high-level features:

$$X'_{r \times n} = W_{bx(r \times p)}^T X_{(p \times r)}. \tag{10}$$

Similarly, the transformation matrix Y' of the feature set Y is obtained. In order to enhance the correlation between the corresponding features of the same type in X and Y , singular value decomposition (SVD) can be used to diagonalize the inter-class covariance matrix $S'_{xy} = X'Y'^T$ of X' and Y' , that is:

$$S'_{xy(r \times r)} = U \Sigma V^T, \quad (11)$$

where U and V are left and right singular matrices respectively, and Σ is a singular value matrix, which only has a non-zero singular value on the principal diagonal, $(U \Sigma^{-1/2})^T S'_{xy(r \times r)} (V \Sigma^{-1/2}) = I$. The final transform feature sets X^* and Y^* can be obtained through the transformation matrix $W_{cx} = U \Sigma^{-1/2}$, $W_{cy} = V \Sigma^{-1/2}$, as shown in Eqs. (12) and (13):

$$X^* = W_{cx}^T X' = W_{cx}^T W_{bx}^T X = W_x X, \quad (12)$$

$$Y^* = W_{cy}^T Y' = W_{cy}^T W_{by}^T Y = W_y Y, \quad (13)$$

where W_x , W_y are the transformation matrices of high-level feature X and Y respectively. Finally, this paper adopts the addition method for feature fusion as shown in Eq. (14) in order to keep the dimension of the feature vector unchanged:

$$Z = X^* + Y^* = \begin{bmatrix} W_x \\ W_y \end{bmatrix}^T \begin{bmatrix} X \\ Y \end{bmatrix}. \quad (14)$$

4. Convolution neural network

CNN is a multi-level feedforward neural network with strong feature extraction ability, and its network parameters are updated by back propagation algorithm. Typical CNN consists of input layer, convolution layer, pooling layer, FC layer and output layer, in which convolution layer and pooling layer are the core of feature extraction. The convolution layer is connected to the previous layer through local connection and weight sharing, and the convolution operation is carried out to generate features, which reduces the number of required training parameters greatly. The pooling layer extracts deep local features by down-sampling the data dimensions, and reduces the complexity of the network by reducing the required parameters, which not only improves the robustness of the model, but also avoids the overfitting phenomenon effectively. At the same time, it makes easier for CNN to use back propagation algorithms for training. The fully connected layer is mainly used to complete the task of classification or regression, and it has the same structure and calculation method as the traditional feedforward neural network. The basic structure of CNN is shown in Fig. 2.

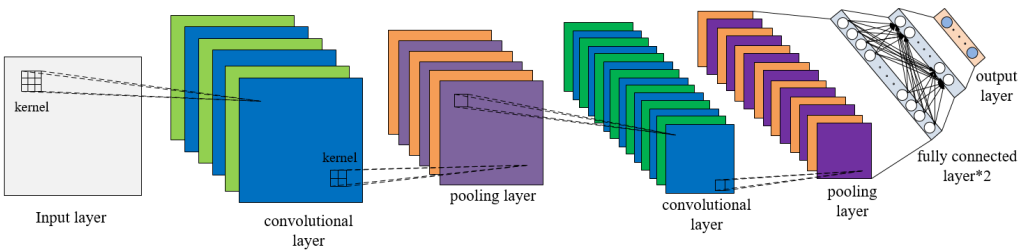


Fig. 2. Typical CNN structure

4.1. Convolution layer

The convolution layer contains multiple convolution cores (also known as filters), that is, weight matrices. Each neuron of each feature graph is connected to the local region of the previous feature graph through a set of weights. This local area is called the receptive field of neurons, and this set of weights is called convolution nucleus. Different convolution kernels have different weights, which are calculated and updated by the error back propagation algorithm. The convolution layer convolutes the input local region with the convolution kernel, and then passes

the results to the nonlinear activation function. Different convolution kernels can generate different feature graphs and generate several new feature graphs as the input of the next layer. Convolution layer has the characteristics of weight sharing and local connection, and the convolution operation is defined as:

$$x_j^l = f \left(\sum_{i \in M_j} x_i^{l-1} * w_{ij}^l + b_j^l \right), \quad (15)$$

where x_j^l is the output feature map of the l layer. $f()$ is the activation function, M_j is the input feature map. x_i^{l-1} is the input information of the l layer; w_{ij}^l is the convolution kernel weight and b_j^l is the bias.

4.2. Activation layer

After the convolution operation, the nonlinear transformation activation function is applied on the output of the convolution calculation, the purpose of which is to improve the representation ability of the network and make the learned features more sufficient. Different activation functions can obtain different nonlinear transformations. The ReLU function is the most commonly used activation function. It can accelerate the convergence of CNN and improve computational efficiency. Its calculation formula is as following:

$$y = f(x) = \begin{cases} x, & x > 0, \\ 0, & x \leq 0, \end{cases} \quad (16)$$

where x is the output value of the convolutional layer, and y is the activation value of x .

4.3. Pooling layer

In the CNN structure, the pooling layer is usually located after the convolutional layer. It mainly reduces the dimensionality of the output of the convolutional layer through down-sampling operations, thereby reducing network parameters and suppressing overfitting to obtain more representative features. Maximum pooling is the most commonly used pooling operation. Maximum pooling is to perform the local maximum operation on the input features to reduce parameters and obtain position-invariant features. It is defined as follows:

$$p_i^{l+1} = \max_{(j-1)W+1 < t < jW} \{q_i^l(t)\}, \quad j = 1, 2, \dots, \quad (17)$$

where the value corresponding to the t neuron of the i feature map of the l layer of $q_i^l(t)$; p_i^{l+1} is the value corresponding to the $l + 1$ layer neuron. W is the pooling width and j is the moving step size.

4.4. Full connection layer and output layer

At the end of CNN, a fully-connected layer is added basically. After the network extracting the deep-level features of the input data through the convolution pooling operation, the feature information is first flattened into a one-dimensional feature vector, which is further used as the input of the fully connected layer. The definition of the fully connected layer is shown in Eq. (18):

$$x^l = f(w^l x^{l-1} + b^l), \quad (18)$$

where l is the serial number of the network layer, x^l is the output of the fully connected layer,

x^{l-1} is the one-dimensional feature vector, w^l is the weight coefficient, b^l is the bias item, and $f()$ is the activation function.

In the fully connected layer, all neuron nodes are connected to all neuron nodes in the previous layer to fully extract input features, and the hidden layer in the middle uses the ReLU activation function. Finally, the Softmax activation function is used in the output layer to complete classification recognition. The Softmax function can normalize the probability distribution of different types of fault characteristics, and any obtained real-valued vector could be compressed into the value range from 0 to 1. The closer the value is to 1, the more likely the output is the actual fault type, which has conducive to the establishment of multi-classification objective function.

4.5. Loss function

The loss function is an integral part of CNN, which can reflect the difference between the predicted value output by the model forward propagation process and the real value. Its main function is to supervise the learning process of the network model, and help the network model to adjust the weight automatically, so that the model can get the best fitting data to minimize the loss value. The smaller the loss value, the closer the actual value of the model is to the expected value. Commonly used loss functions include mean square error and cross-entropy loss function. In multi-classification problems, the cross-entropy loss function is used by CNN usually, which measures the output of the output layer Softmax function by calculating the value of cross-entropy to judge the training effect of the model. The calculation formula of the cross-entropy loss function is as following:

$$Loss = -\frac{1}{m} \sum_{k=1}^m \sum_j p_k^j \log q_k^j, \quad (19)$$

where m represents the number of samples input in batches, p_k^j represents the real value of the input layer samples, and q_k^j represents the predicted value of the output layer Softmax.

5. Construction of DCA-CNN network model

5.1. Framework for diagnostic methods

A CNN based intelligent fault diagnosis model by combing with DCA is proposed to solving the problems that limited training data and single input information make feature learning insufficient and low classification accuracy. The model consists of two parallel branches, and each branch is a convolutional pooling extraction process with specific parameters. First, multi-domain feature information of the original signal is extracted by increasing the network's feature extraction scale. In one branch, the divided training samples in time-domain are processed through the multi-scale CNN branch directly to extract more abundant time-domain feature information. In the other branch, the original training data is transformed by DWT, then further extracts the time-frequency domain feature information of the original data through multi-scale CNN. Subsequently, feature fusion is performed on the extracted multi-domain feature information by DCA feature fusion strategy to enhance the diagnostic performance of the network model. Finally, new features are further extracted from the fused features through a deep multiscale CNN, and the extracted new features are used as the input of the classifier to complete the classification result.

As shown in Fig. 3, the proposed method can be divided into the following four steps:

1) Preprocess the data: the acquired original vibration signal is divided into non-overlapping samples by a sliding window to generate training, validation and test samples.

2) Construct the DCA-CNN network model and set relevant network parameters: number of

iteration steps, training batches, number of convolutional layer and so on.

3) Train the Network model: the original training samples are input into the model, and the multi-scale time-domain feature information of the original data training samples is extracted directly through one branch of the multi-scale CNN. At the same time, the training samples handled by DWT are processed by another branch of the multi-scale CNN to extract the multiscale time-frequency feature information. Then, the extracted time-domain and time-frequency domain feature information are fused through the DCA feature fusion strategy to generate fused training features. Finally, the fused training features are further learned and trained through the deep-level of CNN, and the training of network parameters of each layer is completed through iterative training, and finally the trained model is obtained.

4) Input the test sample data into the trained model, and output the classification diagnosis result through the classifier finally.

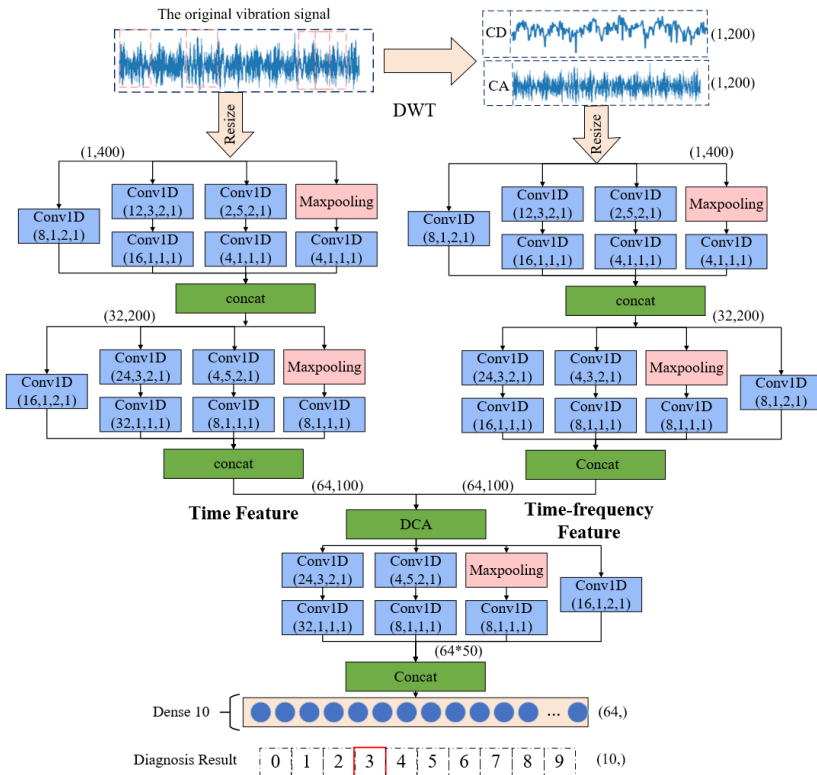


Fig. 3. DCA-CNN network model structure

5.2. DCA-CNN model parameter selection

In the process of feature extraction, the DCA-CNN network model proposed in this paper analyzes from the perspective of multiscale feature extraction in order to maximize the feature information of the input data. By using convolution kernels with different sizes, a four-layer structure is designed for multiscale feature extraction, and two layers of multiscale feature extraction layers are designed in the parallel two-branch CNN respectively. The used kernel size and step size are same. The number of convolution kernels is different to maximize the extraction of multiscale temporal features and multiscale temporal features of the original data. Through a branched multi-scale CNN, the time-domain features of different scales extracted directly from the original data training samples are fused through Con-cat to output the extracted time-domain feature information. The original training sample is decomposed into the first-order low-frequency

approximation signal and the first-order high-frequency detail signal by DWT, and the data fusion is carried out by serial operation. Subsequently, the fusion data is passed through another branch of multiscale CNN processes to extract the multiscale time-frequency feature information of the training data, and Con-cat is also used to fuse the extracted multiscale time-frequency feature information to output the extracted time-frequency domain features.

The extracted feature information in time domain and time-frequency domain is fused by DCA feature fusion strategy. The fused features are further extracted with multiscale features through deep multi-scale CNN, and Con-cat is also used to fuse the extracted multiscale features to generate new features. Then the extracted new features are flattened into one-dimensional feature vectors and used as the input of the fully connected layer, and the classification and diagnosis of faults are completed according to the output of the classifier at last. Table 1 lists the specific parameters of the network. The convolution layer Conv1D(8,2,2) means that the number of convolution kernels is 8, the size is 2, and the step size is 2. After each convolution layer, there is a batch normalization layer to normalize the features into a suitable data distribution. The size of the maximum pooling layer is set to 3, the step size is 2, and the convolutional layer and the maximum pooling layer are both zero-filled. Using the Adam optimizer, the learning rate is set to 0.005, and finally the Softmax classifier is used for classification, in which the accuracy rate is selected as the evaluation standard.

Table 1. Detailed parameters of the DCA-CNN network model

Operation	Sample data partition	
feature domain	Time Domain	Time-frequency domain (DWT)
Branch multi-scale CNN feature extraction	1) Conv1d (8,2,2) 2) Conv1d(12,3,2) →Conv1d(16,1,1) 3) Conv1d (2,5,2)→Conv1d(4,1,1) 4) Maxpooling→Conv1d (4,1,1)	1) Conv1d (8,2,2) 2) Conv1d(12,3,2) →Conv1d(16,1,1) 3) Conv1d (2,5,2)→Conv1d(4,1,1) 4) Maxpooling→Conv1d (4,1,1)
Branch multi-scale feature fusion method	Con-cat	Con-cat
Multi-domain feature fusion method	DCA	
Multi-scale CNN feature extraction	1) Conv1d (16,2,2) 2) Conv1d(24,3,2) →Conv1d(32,1,1) 3) Conv1d (4,5,2)→Conv1d(8,1,1) 4) Maxpooling→Conv1d (8,1,1)	
Multi-scale CNN feature fusion method	Con-cat	
Fully connected layer 1	64	
Fully connected layer 2	10	

6. Experiment verification and analysis

In this section, effectiveness of the method is verified through two experiments. In addition, noise interference experiments are carried out and compared with other models to verify the superiority of the proposed model.

6.1. Experiment 1: rolling bearing failure data of case western reserve university (CWRU)

6.1.1. Data description

The experimental data use the rolling bearing fault data set of CWRU [28], and its experimental platform is shown in Fig. 4. The type of bearing selected in the experiment is SKF6205, which is installed on the driving end of the induction motor. The signal is recorded by the accelerometer installed at the drive end of the induction motor, and the data is collected by the

16-channel acquisition instrument, and the sampling frequency is set to 12 kHz. In this experiment, the vibration data set of the first 10 seconds collected by the sensor is selected. In this data set, three types of bearing defects are machined by EDM technology: inner ring, outer ring and rolling body defects. Each type of defect has three different sizes: 0.1778 mm 0.3556 mm and 0.5334 mm. The bearing data of different fault locations and different defect degrees are regarded as one class separately, which includes a total of 10 types. Based on the frequency and rotational speed of the collected data, it can be inferred that the data points collected in each circle are as follows: sampling points (circle) = sampling frequency \times 60 / rotational speed = $12000 \times 60 / 1797 = 400$. Therefore, the sample data is divided by the way of non-overlap. The sample length is set to 400 and the displacement is set to the length of a sample. A total of 3000 samples corresponding to 10 types of bearing states are constructed. The health status is marked with label 0, and the other 9 faults are marked by label 1-9. 70 % of the experimental data set is randomly selected as the training set, 10 % as the verification set, and 20 % as the test set. The details of the sample division of the data set are shown in Table 2.

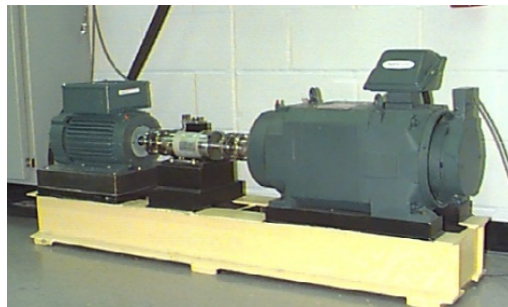


Fig. 4. CWRU bearing fault test rig

Table 2. Sample division of the rolling bearing data set

Fault type	Damage diameter	Sample length	Number of samples	Label
Normal	0	400	300	0
OR007	0.1778	400	300	1
OR014	0.3556	400	300	2
OR021	0.5334	400	300	3
IR007	0.1778	400	300	4
IR007	0.3556	400	300	5
IR007	0.5334	400	300	6
B007	0.1778	400	300	7
B014	0.3556	400	300	8
B021	0.5334	400	300	9

In order to verify the outstanding performance of the proposed network under different load conditions, the original vibration data of the drive bearing under four loads of 0-3hp are selected and defined as data sets A, B, C and D respectively. Each data set contains 10 states, and the description information of the data set is shown in Table 3.

Table 3. Dataset description

Data set	Load (hp)	Number of samples /categories	Sample division /categories	Number of categories
A	0	300	7:1:2	10
B	1	300	7:1:2	10
C	2	300	7:1:2	10
D	3	300	7:1:2	10

Fig. 5 shows the time domain waveforms of 10 kinds of vibration signals with sample length

of 400 under 0 hp load. Based on the proposed DCA-CNN network, the training sample data from the original vibration signal is input to a CNN branch for feature extraction firstly. At the same time, the training data set is transformed by DWT, and then input to another CNN branch for time-frequency domain feature extraction. Fig. 6 shows the low-frequency and high-frequency signals decomposed by DWT. Generally speaking, the low-frequency components usually contain the important characteristics of the signal, while the high-frequency components can show the details or differences of the signal.

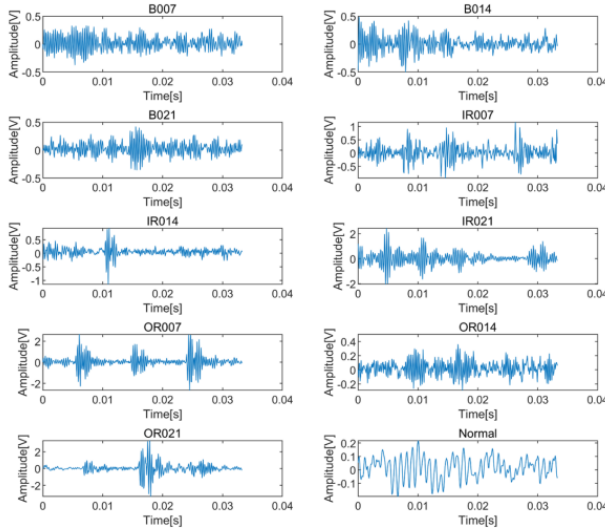


Fig. 5. Original waveform diagram of bearing vibration signals in different states

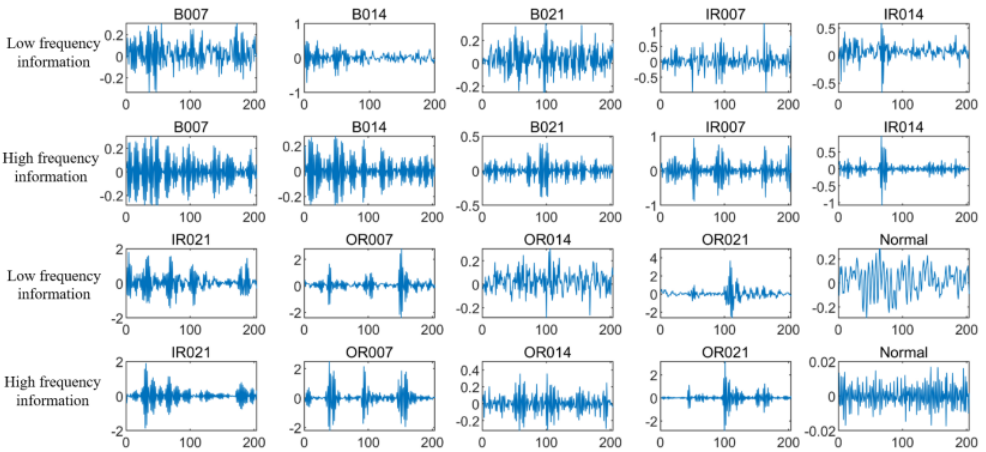


Fig. 6. Low-frequency information and high-frequency information of DWT decomposition of bearing vibration signals in different states

6.1.2. Analysis of experimental results

Fig. 7 is one of the confusion matrix diagrams of 10 experiments performed on 4 data sets using the method model in this chapter. The classification accuracy of the diagnostic model on 4 data sets is 97.5 %-99.33 %, among which the classification effect on data set B is the best. The multi-class confusion matrix is able to record the classification results of all conditions in detail, including the classification accuracy and the number of mis-classifications. The dark area on the

diagonal of the confusion matrix represents the accuracy rate corresponding to each type of failure, 60 is the number of test sets for each type of failure, and the values in the rest of the area represents the number of misclassifications. The vertical axis is the actual label of the classification, and the horizontal axis is the predicted label of classification. The confusion matrix presented in Fig. 7(b) shows the classification and recognition of the 10 states, and it can be seen from the figure that the model has achieved good classification results under the four types of data sets, indicating that the proposed DCA-CNN diagnosis model has better recognition ability and higher diagnosis accuracy.

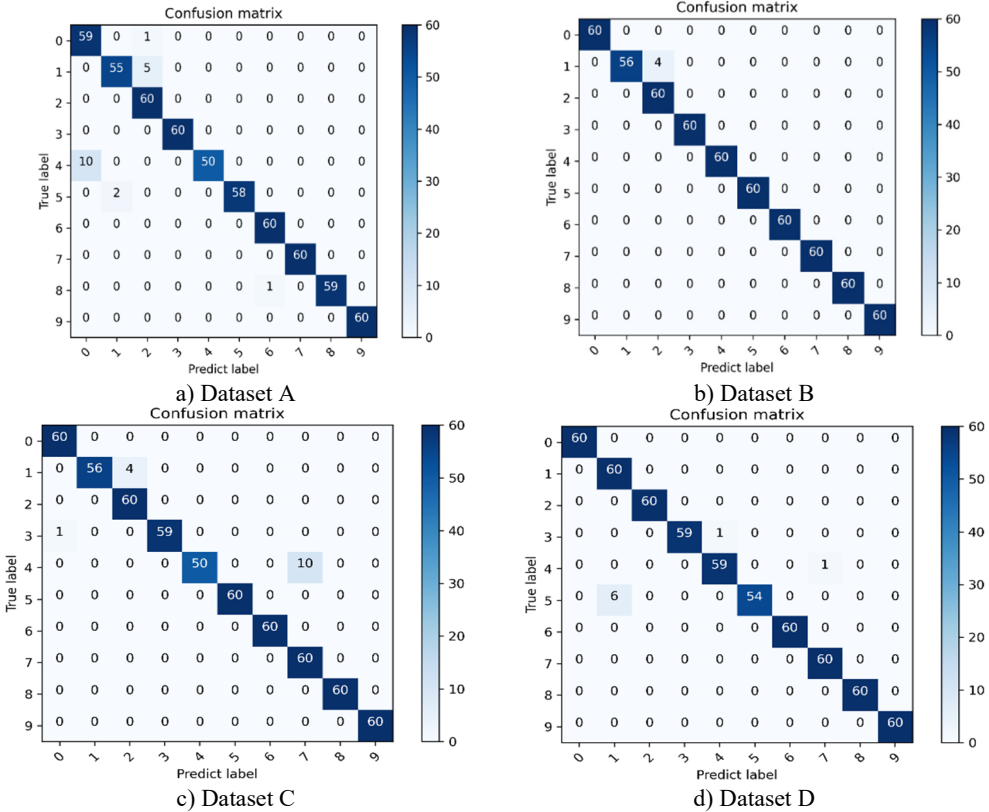


Fig. 7. The multi-class confusion matrix visualization of the proposed method under 4 data sets

In order to further illustrate the impact of the proposed multi-domain feature fusion strategy on the classification results, the most common t-SNE method in manifold learning is introduced. By mapping high-dimensional feature vectors to three-dimensional space, the t-SNE method is used for dimensionality reduction and feature visualization. Here, it is selected as the feature visualization result graph of the last fully connected layer of the DCA-CNN network model based on data set B, and the corresponding result is shown in Fig. 8(b). Besides, the t-SNE visualization of CNN network model based on data set B is also given in Fig. 8(a) for comparison.

The CNN model compared in this chapter uses the same parameters as the DCA-CNN. It also extracts multiscale time-domain and time-frequency domain features from the original data through two parallel branch CNNs. The difference is that there is no DCA feature fusion in the CNN model. In the process of fusion of CNN, Con-cat is used to fuse the extracted time domain and time-frequency domain features directly. In Fig. 8, the classification accuracy of the test set of the CNN network model is 93.67 %, and the classification accuracy of the DCA-CNN network model is 98.83 %. Compared Fig. 8(b) with Fig. 8(a), the feature classes of the former are more

concentrated, and the distance between classes is larger, which proves that the DCA-CNN network model can not only be well clustered into categories, but also easier to be identified.

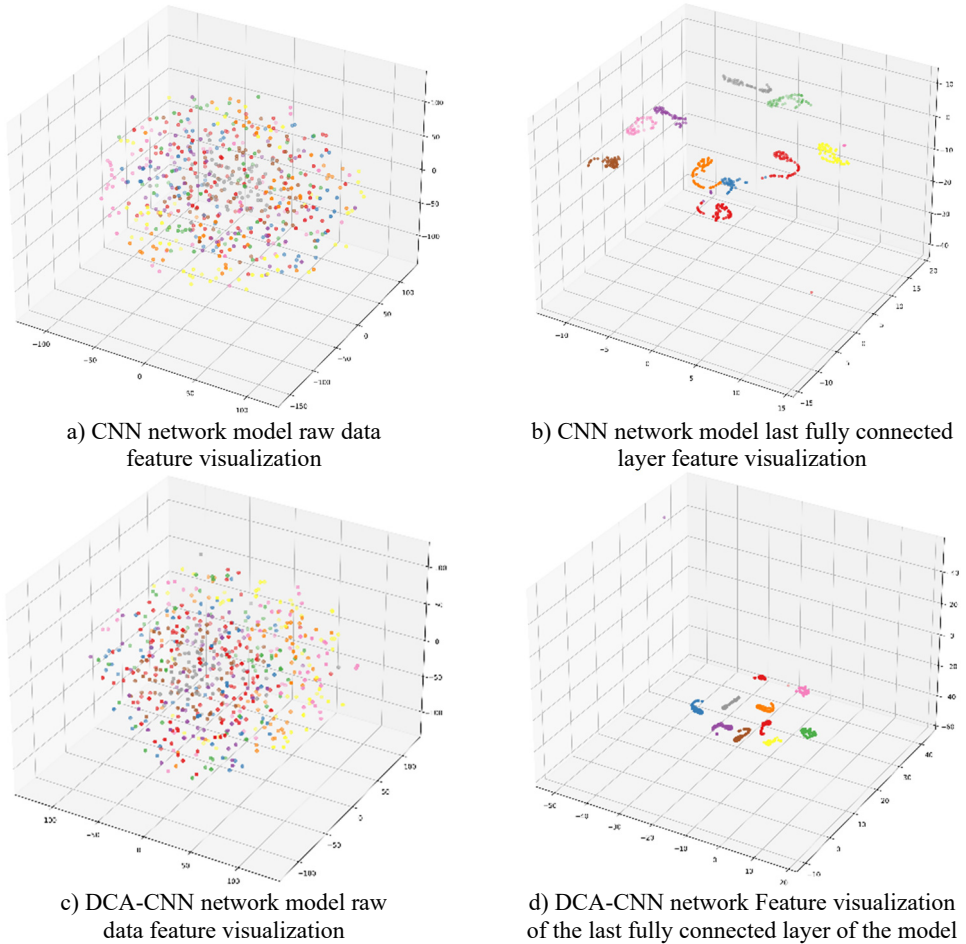


Fig. 8. t-SNE feature visualization

In the actual operation scene of rotating machinery, environmental noise exists inevitably, and the measured vibration signal is often interfered by noise, which reduces the effectiveness of the fault diagnosis method. In this regard, Gaussian white noise is added into the original data of the test set to simulate the impact of noise on the classification and diagnosis results, and the trained model is used for testing to evaluate the performance of the proposed method in noisy environment. The definition of SNR is shown in Eq. (6):

$$SNR_{dB} = 10 \log_{10} \left(\frac{P_{signal}}{P_{noise}} \right), \quad (20)$$

where P_{signal} represents the power of the original signal and P_{noise} represents the power of the noise signal. The smaller the SNR, the greater the noise interference.

Fig. 9 shows the test accuracy of the proposed method at different levels of noise compared with the CNN method. It can be observed that adding extra noise usually reduces the diagnostic performance in different scenarios, and stronger noise generally leads to lower test accuracy. Specifically, the model is less accurate when $SNR < 2$ dB, and is higher than 94 % when

SNR > 4 dB. Besides, the proposed method is better than the CNN method in most cases, and high test accuracy can still be obtained in the case of additional environmental noise. It can be seen that the method proposed in this chapter has strong anti-noise stability and generalization ability.

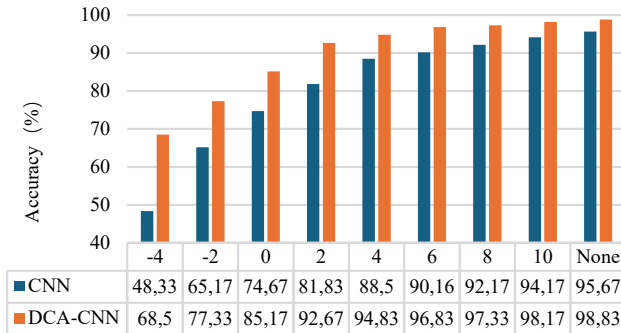


Fig. 9. Accuracy rate of rolling bearing test set under different SNR

6.1.3. Comparison

For the method proposed in this study, several conventional bearing fault recognition models are used for comparison: 1D-CNN [29], EMD+SVM [30], DNN network [31]. Here, 1D-CNN directly trains and tests the original signal data set in time domain. EMD+SVM decomposes the input vibration signal by using EMD, and then uses fuzzy entropy to extract the characteristics of the vibration signal effectively. The DNN uses nine time-domain statistical features of the original as input. This test is divided into 4 groups to verify the diagnostic performance of the proposed network under different load conditions. At the same time, the above conventional fault identification methods were tested on various data sets for 10 times in order to reduce the interference of uncertain factors on the experimental results and ensure the reliability of the proposed method, and the average value was taken as the final diagnostic accuracy. Fig. 10 shows the average recognition accuracy trend graph of the different methods under different data sets.

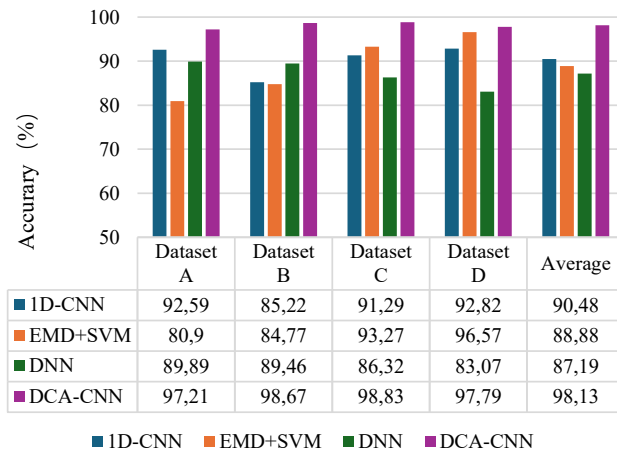


Fig. 10. The average recognition accuracy of different methods under different data sets

To verify the superiority of the method, we compared it with the advanced methods: (1) SR-DEEP [32], (2) BFD-2DCNN [33], (3) LSSA-VMD-GRU [34], (4) WPD-CSSOA-DBN [35], (5) ACP SO-BP [36]. The experimental results are shown in Fig. 11.

From the results as shown in Fig.10, it can be seen that the 1D-CNN model and the DNN diagnostic model have relatively stable diagnostic performance under the 4 datasets, but their

diagnostic accuracies are low. EMD+SVM has a better diagnostic performance on the data set D, but has the worst diagnostic accuracy on dataset A. From the results as shown in Fig. 11, our proposed method achieves optimal results. Furthermore, the DCA-CNN diagnostic model proposed in this chapter has achieved relatively good diagnostic accuracy in the four types of data sets, and its average recognition accuracy and recognition accuracy of each data set are basically higher than other methods.

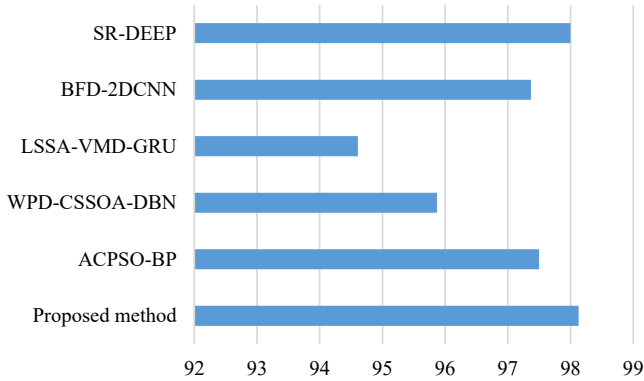


Fig. 11. The average recognition accuracy of different methods

6.2. Experiment 2: gearbox failure data of QPZZ-II test bench

6.2.1. Description of experimental data

The gear fault vibration signals collected from the QPZZ-II rotating machinery vibration test bench [37] is used to further verify the proposed method, and Fig. 12 shows the experiment rig. Ten different kinds of gear failures are simulated on the test bench, and the teeth numbers of the test gears are 75 and 55 respectively. The modulus of the test gears is 2. In the experiment, the wire EDM process was used to create faults in the large gear. By replacing the faulty gear in the gear box, a total of 10 different gear states were simulated, and their corresponding faulty parts are shown in Fig. 13. Vibration data in different states are collected through an acceleration sensor installed on the gearbox. The motor speed is 1500 r/min, the sampling frequency is set to 12,800 Hz, and the sampling time is set as 10 s. A total of 128,000 data points are obtained for each state. In the experiment, 400 data points were selected as the data of one sample, and the number of samples for each type of fault was 320 using non-overlapping division. The sample was divided into training set, verification set and test set according to the ratio of 7:1:2. Table 4 shows the details of the ten running states. The original waveforms of the ten kinds running states and their corresponding decomposed low-frequency signal components and high-frequency signal components using DWT are presented in Fig. 14 and Fig. 15 respectively.

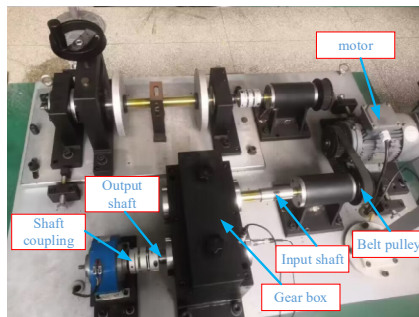


Fig. 12. QPZZ-II rotating machinery vibration test bench

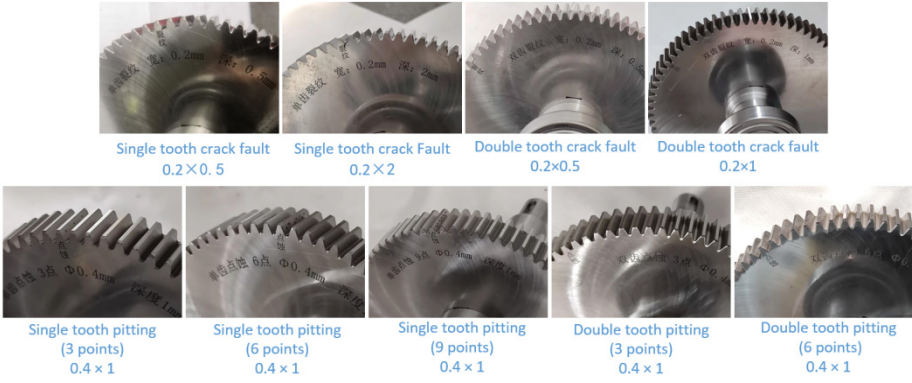


Fig. 13. Gear failure parts

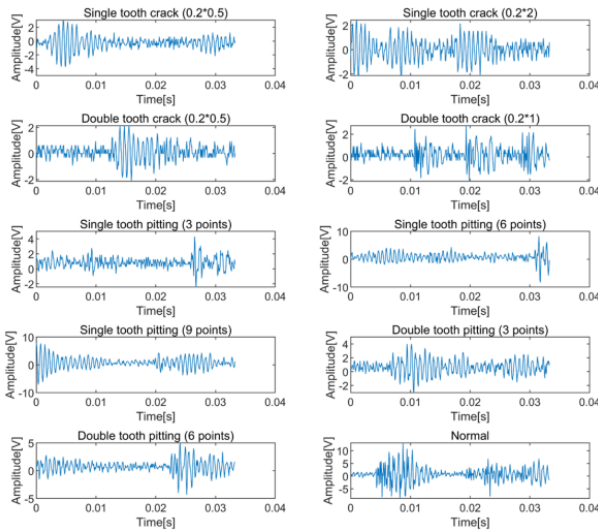


Fig. 14. Original waveform diagram of gear vibration signal in different states

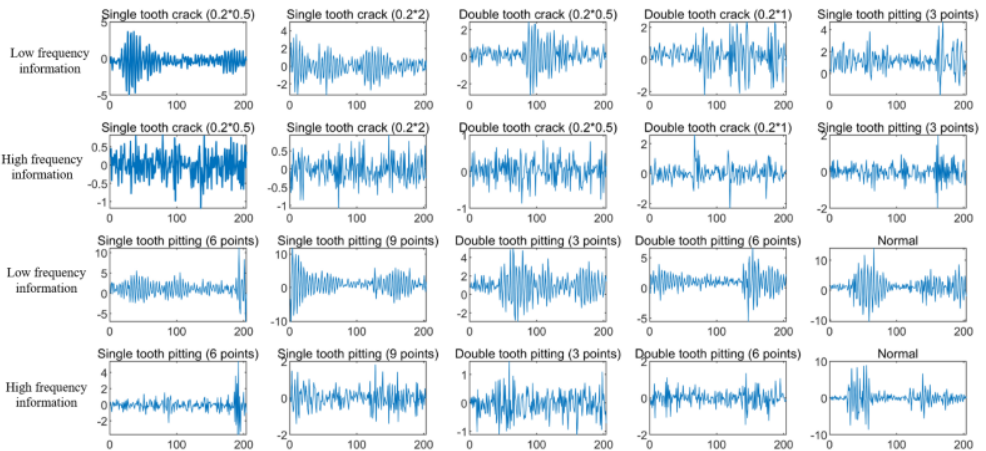


Fig. 15. Low-frequency information and high-frequency information of DWT decomposition of gear vibration signals in different states

Table 4. Gear failure dataset description

Fault type	Failure points	Damage diameter	Sample length	Number of samples	Sample division	Label
Normal	–	–	400	320	224/32/64	0
Single tooth crack fault	–	0.2×0.5	400	320	224/32/64	1
	–	0.2×2	400	320	224/32/64	2
Double tooth crack fault	–	0.2×0.5	400	320	224/32/64	3
	–	0.2×1	400	320	224/32/64	4
Single tooth pitting fault	3 Point	Ø0.4×1	400	320	224/32/64	5
	6 Point	Ø0.4×1	400	320	224/32/64	6
	9 Point	Ø0.4×1	400	320	224/32/64	7
Double tooth pitting fault	3 Point	Ø0.4×1	400	320	224/32/64	8
	6 Point	Ø0.4×1	400	320	224/32/64	9

6.2.2. Analysis of experimental results

The accuracy curves using the proposed method and CNN network directly are shown in Fig. 16 after conducting 10 experiments, and their average accuracy ratios are 98.96 % and 95.41 % respectively. The advantage of the proposed method over the CNN network is verified. The multi-class confusion matrix and t-SNE feature visualization of the results obtained in the fourth of 10 experiments using the proposed method and CNN model are given in Fig. 17 and Fig. 18 respectively, and their recognition accuracy rates are 99.69 % and 94.84 % respectively, which further verify the effectiveness of the proposed method intuitively.

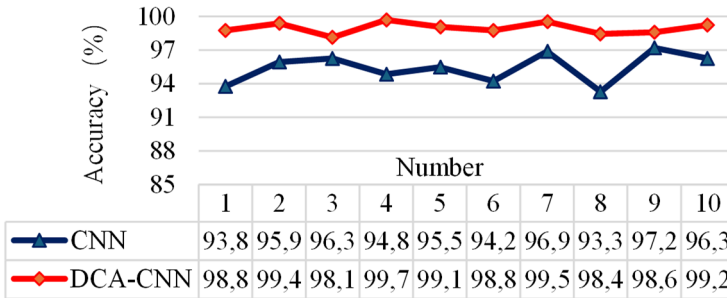
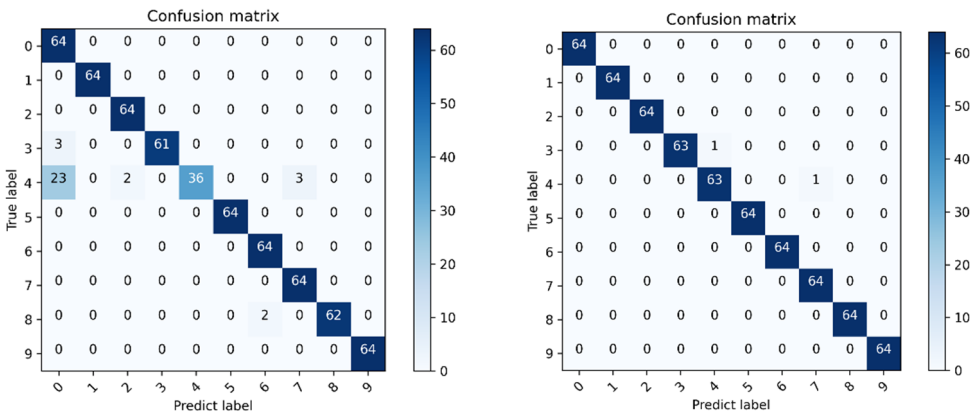


Fig. 16. Test set accuracy for ten trials

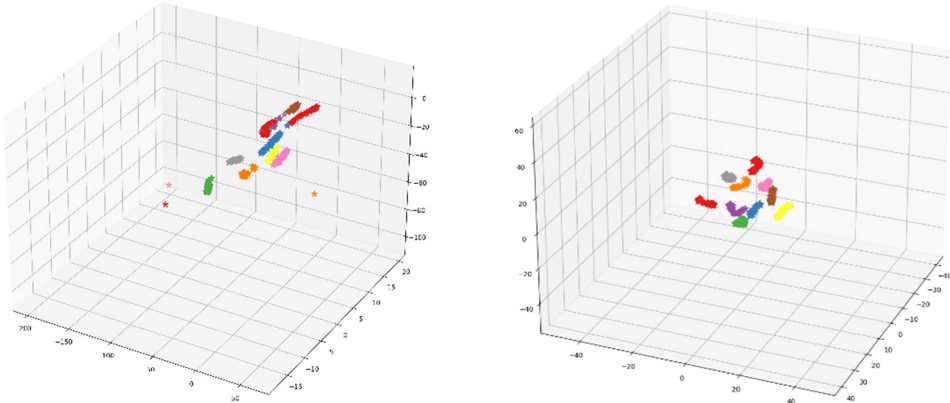


a) CNN network model multi-class confusion matrix visualization

b) DCA-CNN network model multi-class confusion matrix visualization

Fig. 17. The proposed method is used to visualize the multi-class confusion matrix under the gear dataset

By adding noise interference with different degree of SNRs to the gear test set data, the diagnostic effect of the method on the gear fault data set under the noise interference is verified. Accordingly, the noise interference on the CNN model method is compared, and the results are shown in Fig. 19. It can be seen from the figure that when the SNR is -4 dB, the classification accuracy of the CNN model is 50.78 %, while the classification accuracy of the DCA-CNN model is 66.41 %. When the SNR becomes 6dB, the classification accuracy of the DCA-CNN model can reach more than 96 %, which shows the effectiveness of the model for fault diagnosis of gearboxes in noisy environments.



a) Feature visualization of the last fully connected layer of the CNN network model b) Feature visualization of the last fully connected layer of the DCA-CNN network model

Fig. 18. The proposed method is used for t-SNE feature visualization under the gear dataset

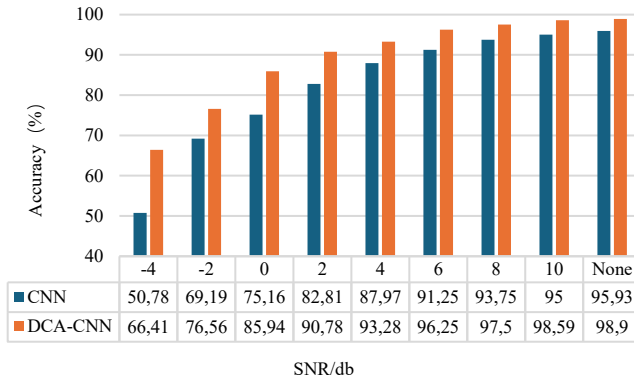


Fig. 19. Accuracy of gear test set under different SNR

6.2.3. Comparison

Similarly, 1DCNN, P1-CNN, CNN, and DCA-CNN were used to conduct ten experiments on the gearbox fault data set for comparison, and their corresponding average values are used as the final accuracy. Among them, the 1DCNN model adopts the five-layer convolution and pooling layer structure built in Chapter 3 in this paper. The P1-CNN model is constructed by combining the time-domain multi-scale feature extraction with deep multi-scale feature extraction part of the DCA-CNN model. The CNN model is constructed by combining the multi-domain and multi-scale feature extraction part of the DCA-CNN model with the deep multi-scale feature extraction part, in which the multi-domain feature is not fused by DCA, but a simple concept operation is used.

As can be seen from Fig. 20, P1-CNN model does not need a deeper network structure to achieve higher diagnostic accuracy than 1DCNN model due to the addition of multi-scale feature

extraction module. CNN model has higher diagnosis accuracy than IDCNN and P1-CNN model due to the adding multi-domain feature extraction module and multi-scale feature extraction module. The diagnostic accuracy of DCA-CNN model is the highest, which is 98.96 %. These show that the method proposed in this chapter can improve the diagnosis performance of CNN network effectively through multi-scale feature extraction module, multi-domain feature extraction module and feature fusion module based on DCA. Besides, the proposed method also could achieve higher classification diagnosis accuracy when the input data is relatively single and the actual data set is limited.

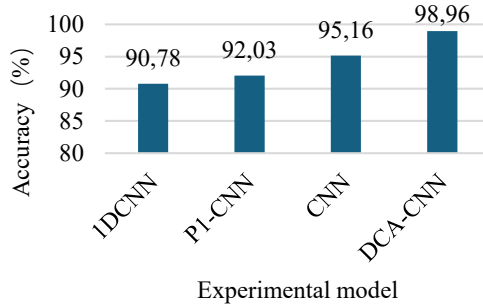


Fig. 20. The average accuracy of the gear failure test set under different methods

7. Conclusions

In this paper, a network model based on DCA and CNN is proposed for fault diagnosis of rotating machinery. The following conclusions could be obtained according to the analysis of two groups of different types of experimental data by using the proposed method:

1) Through the combination original signal processing handling ability of DWT with the strong nonlinear feature learning ability of CNN, the proposed method can get rid of the dependence on signal processing technology and manual feature extraction and can extract effective deep fault features adaptively and improve the diagnosis accuracy of the model.

2) The most representative feature information could be obtained by adding the DCA feature fusion module into the multi-domain CNN network model to fuse the extracted multi-domain feature information. Meanwhile, the correlation between similar fused features is improved, and the redundant correlation between different types of features is reduced.

3) Compared with series of other fault diagnosis methods, the results show that the proposed model can still achieve higher diagnosis in the case of limited data set and single input information and has a wide range of application. Besides, the anti-noise performance of the proposed model is verified.

In the future work, we will verify the scalability of the proposed DCA-CNN diagnosis model on a large-scale equipment. In addition, we will further study the fault diagnosis of rotating machinery under variable working conditions based on this method and study the applicability of fault diagnosis based on two-dimensional image data, so as to further improve the performance of the proposed model.

As for the engineering application of the proposed method, it could be embedded as an algorithm package into the device online monitoring software, and the algorithm package only needs to define the input and output. The proposed method can also be embedded as algorithm package into offline monitoring instruments to service the low-cost offline monitoring instrument users. Of course, as the proposed method is developed based on Matlab and Python languages, both software engineering and hardware engineering applications cannot be separated from the mixing programming of Matlab, Python with other development languages such as C++ and Java.

Acknowledgements

The authors have not disclosed any funding.

Data availability

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Author contributions

Guisheng Lan: paper conception and writing; Haibo Shi: theoretical research and algorithm implementation.

Conflict of interest

The authors declare that they have no conflict of interest.

References

- [1] I. Attoui, B. Oudjani, N. Boutasseta, N. Fergani, M.-S. Bouakkaz, and A. Bouraiou, "Novel predictive features using a wrapper model for rolling bearing fault diagnosis based on vibration signal analysis," *The International Journal of Advanced Manufacturing Technology*, Vol. 106, No. 7-8, pp. 3409–3435, Jan. 2020, <https://doi.org/10.1007/s00170-019-04729-4>
- [2] Y. Xu, Z. Li, S. Wang, W. Li, T. Sarkodie-Gyan, and S. Feng, "A hybrid deep-learning model for fault diagnosis of rolling bearings," *Measurement*, Vol. 169, p. 108502, Feb. 2021, <https://doi.org/10.1016/j.measurement.2020.108502>
- [3] J. Ben Ali, N. Fnaiech, L. Saidi, B. Chebel-Morello, and F. Fnaiech, "Application of empirical mode decomposition and artificial neural network for automatic bearing fault diagnosis based on vibration signals," *Applied Acoustics*, Vol. 89, pp. 16–27, Mar. 2015, <https://doi.org/10.1016/j.apacoust.2014.08.016>
- [4] D. H. Pandya, S. H. Upadhyay, and S. P. Harsha, "Fault diagnosis of rolling element bearing with intrinsic mode function of acoustic emission data using APF-KNN," *Expert Systems with Applications*, Vol. 40, No. 10, pp. 4137–4145, Aug. 2013, <https://doi.org/10.1016/j.eswa.2013.01.033>
- [5] L. Saidi, J. Ben Ali, and F. Fnaiech, "Application of higher order spectral features and support vector machines for bearing faults classification," *ISA Transactions*, Vol. 54, pp. 193–206, Jan. 2015, <https://doi.org/10.1016/j.isatra.2014.08.007>
- [6] F. Jia, Y. Lei, J. Lin, X. Zhou, and N. Lu, "Deep neural networks: A promising tool for fault characteristic mining and intelligent diagnosis of rotating machinery with massive data," *Mechanical Systems and Signal Processing*, Vol. 72-73, pp. 303–315, May 2016, <https://doi.org/10.1016/j.ymsp.2015.10.025>
- [7] G. Jiang, H. He, P. Xie, and Y. Tang, "Stacked multilevel-denoising autoencoders: A new representation learning approach for wind turbine gearbox fault diagnosis," *IEEE Transactions on Instrumentation and Measurement*, Vol. 66, No. 9, pp. 2391–2402, Sep. 2017, <https://doi.org/10.1109/tim.2017.2698738>
- [8] D.-T. Hoang and H.-J. Kang, "Rolling element bearing fault diagnosis using convolutional neural network and vibration image," *Cognitive Systems Research*, Vol. 53, pp. 42–50, Jan. 2019, <https://doi.org/10.1016/j.cogsys.2018.03.002>
- [9] H. Shao, H. Jiang, F. Wang, and Y. Wang, "Rolling bearing fault diagnosis using adaptive deep belief network with dual-tree complex wavelet packet," *ISA Transactions*, Vol. 69, pp. 187–201, Jul. 2017, <https://doi.org/10.1016/j.isatra.2017.03.017>
- [10] Y. Qi, C. Shen, D. Wang, J. Shi, X. Jiang, and Z. Zhu, "Stacked sparse autoencoder-based deep network for fault diagnosis of rotating machinery," *IEEE Access*, Vol. 5, pp. 15066–15079, Jan. 2017, <https://doi.org/10.1109/access.2017.2728010>

- [11] C. Wang, H. Li, K. Zhang, S. Hu, and B. Sun, "Intelligent fault diagnosis of planetary gearbox based on adaptive normalized CNN under complex variable working conditions and data imbalance," *Measurement*, Vol. 180, p. 109565, Aug. 2021, <https://doi.org/10.1016/j.measurement.2021.109565>
- [12] C. Liu, G. Cheng, X. Chen, and Y. Pang, "Planetary gears feature extraction and fault diagnosis method based on VMD and CNN," *Sensors*, Vol. 18, No. 5, p. 1523, May 2018, <https://doi.org/10.3390/s18051523>
- [13] L. Wen, X. Li, L. Gao, 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, Jul. 2018, <https://doi.org/10.1109/tie.2017.2774777>
- [14] S. Shao, R. Yan, Y. Lu, P. Wang, and R. X. Gao, "DCNN-based multi-signal induction motor fault diagnosis," *IEEE Transactions on Instrumentation and Measurement*, Vol. 69, No. 6, pp. 2658–2669, Jun. 2020, <https://doi.org/10.1109/tim.2019.2925247>
- [15] Y. Xue, D. Dou, and J. Yang, "Multi-fault diagnosis of rotating machinery based on deep convolution neural network and support vector machine," *Measurement*, Vol. 156, p. 107571, May 2020, <https://doi.org/10.1016/j.measurement.2020.107571>
- [16] A. Khan, J. E. Sung, and J.-W. Kang, "Multi-channel fusion convolutional neural network to classify syntactic anomaly from language-related ERP components," *Information Fusion*, Vol. 52, pp. 53–61, Dec. 2019, <https://doi.org/10.1016/j.inffus.2018.10.008>
- [17] H. Wang, S. Li, L. Song, and L. Cui, "A novel convolutional neural network based fault recognition method via image fusion of multi-vibration-signals," *Computers in Industry*, Vol. 105, pp. 182–190, Feb. 2019, <https://doi.org/10.1016/j.compind.2018.12.013>
- [18] X. Ding and Q. He, "Energy-fluctuated multiscale feature learning with deep ConvNet for intelligent spindle bearing fault diagnosis," *IEEE Transactions on Instrumentation and Measurement*, Vol. 66, No. 8, pp. 1926–1935, Aug. 2017, <https://doi.org/10.1109/tim.2017.2674738>
- [19] 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, Jul. 2017, <https://doi.org/10.1109/tim.2017.2669947>
- [20] M. Demetgul, K. Yildiz, S. Taskin, I. N. Tansel, and O. Yazicioglu, "Fault diagnosis on material handling system using feature selection and data mining techniques," *Measurement*, Vol. 55, pp. 15–24, Sep. 2014, <https://doi.org/10.1016/j.measurement.2014.04.037>
- [21] F. Xue, W. Zhang, F. Xue, D. Li, S. Xie, and J. Fleischer, "A novel intelligent fault diagnosis method of rolling bearing based on two-stream feature fusion convolutional neural network," *Measurement*, Vol. 176, p. 109226, May 2021, <https://doi.org/10.1016/j.measurement.2021.109226>
- [22] Y. Cheng, M. Lin, J. Wu, H. Zhu, and X. Shao, "Intelligent fault diagnosis of rotating machinery based on continuous wavelet transform-local binary convolutional neural network," *Knowledge-Based Systems*, Vol. 216, No. 1, p. 106796, Mar. 2021, <https://doi.org/10.1016/j.knosys.2021.106796>
- [23] S. H. Syed and V. Muralidharan, "Feature extraction using discrete wavelet transform for fault classification of planetary gearbox – a comparative study," *Applied Acoustics*, Vol. 188, p. 108572, Jan. 2022, <https://doi.org/10.1016/j.apacoust.2021.108572>
- [24] S. Sharma, S. K. Tiwari, and S. Singh, "Diagnosis of gear tooth fault in a bevel gearbox using discrete wavelet transform and autoregressive modeling," *Life Cycle Reliability and Safety Engineering*, Vol. 8, No. 1, pp. 21–32, Jul. 2018, <https://doi.org/10.1007/s41872-018-0061-9>
- [25] S. G. Mallat, "A theory for multiresolution signal decomposition: the wavelet representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 11, No. 7, pp. 674–693, Jul. 1989, <https://doi.org/10.1109/34.192463>
- [26] M. Haghighat, M. Abdel-Mottaleb, and W. Alhalabi, "Discriminant correlation analysis: real-time feature level fusion for multimodal biometric recognition," *IEEE Transactions on Information Forensics and Security*, Vol. 11, No. 9, pp. 1984–1996, Sep. 2016, <https://doi.org/10.1109/tifs.2016.2569061>
- [27] Q. Sun, "Combined feature extraction based on canonical correlation analysis and face recognition," *Journal of Computer Research and Development*, Vol. 42, No. 4, p. 614, Jan. 2005, <https://doi.org/10.1360/crad20050413>
- [28] Case Western Reserve University Bearing Data Center.
- [29] X. Liu, Q. Zhou, and J. Zhao, "Real-time and anti-noise fault diagnosis algorithm based on 1-D convolutional neural network," *Journal of Harbin Institute of Technology*, Vol. 51, pp. 89–95, 2019.

- [30] W. Deng, R. Yao, M. Sun, H. Zhao, Y. Luo, and C. Dong, "Study on a novel fault diagnosis method based on integrating EMD, fuzzy entropy, improved PSO and SVM," *Journal of Vibroengineering*, Vol. 19, No. 4, pp. 2562–2577, Jun. 2017, <https://doi.org/10.21595/jve.2017.18052>
- [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, p. 1729, Jul. 2017, <https://doi.org/10.3390/s17081729>
- [32] Y. Niu et al., "A sparse learning method with regularization parameter as a self-adaptation strategy for rolling bearing fault diagnosis," *Electronics*, Vol. 12, No. 20, p. 4282, Oct. 2023, <https://doi.org/10.3390/electronics12204282>
- [33] D. Gilbert Chandra, U. Srinivasulu Reddy, G. Uma, and M. Umapathy, "Group normalization-based 2D-convolutional neural network for intelligent bearing fault diagnosis," *Journal of the Brazilian Society of Mechanical Sciences and Engineering*, Vol. 45, No. 11, pp. 1–13, Oct. 2023, <https://doi.org/10.1007/s40430-023-04491-5>
- [34] G. Ma, X. Yue, J. Zhu, Z. Liu, and S. Lu, "Deep learning network based on improved sparrow search algorithm optimization for rolling bearing fault diagnosis," *Mathematics*, Vol. 11, No. 22, p. 4634, Nov. 2023, <https://doi.org/10.3390/math11224634>
- [35] F. Zhao, Y. Jiang, C. Cheng, and S. Wang, "An improved fault diagnosis method for rolling bearings based on wavelet packet decomposition and network parameter optimization," *Measurement Science and Technology*, Vol. 35, No. 2, p. 025004, Feb. 2024, <https://doi.org/10.1088/1361-6501/ad0691>
- [36] S. Song, S. Zhang, W. Dong, X. Zhang, and W. Ma, "A new hybrid method for bearing fault diagnosis based on CEEMDAN and ACPSO-BP neural network," *Journal of Mechanical Science and Technology*, Vol. 37, No. 11, pp. 5597–5606, Nov. 2023, <https://doi.org/10.1007/s12206-023-1003-7>
- [37] F. Wei, G. Wang, and B. Ren, "Multisensor fused fault diagnosis for rotation machinery based on supervised second-order tensor locality preserving projection and weighted-nearest neighbor classifier under assembled matrix distance metric," *Shock and Vibration*, pp. 1–14, 2016, <https://doi.org/10.1088/1361-6501/ac2b72/meta>



Guisheng Lan received bachelor's degree in Zhejiang university, Hangzhou, China, in 1991. Now he works at GuangDong Yue Guang Water Supply Co., Ltd. His current research interests include transmission technology and fault diagnosis.



Haibo Shi received master's degree in Wuhan university, Wuhan, China, in 2019. Now he works at GuangDong Yue Guang Water Supply Co., Ltd. His current research interests include transmission technology and fault diagnosis.