Zhao, K, Xiao, J, Li, C, Xu, Z and Yue, M

Fault diagnosis of rolling bearing using CNN and PCA fractal based feature extraction

Fault Diagnosis of Rolling Bearing using CNN and PCA Fractal Based on Feature Extraction

Kaicheng Zhao[a],Junqing Xiao[a] Chun Li[a, b], Zifei Xu[a,c], Minnan Yue[a, *]

[a] *School of Energy and Power Engineering, University of Shanghai for Science and Technology, Shanghai, 200093. China*

[b] *Shanghai Key Laboratory of Multiphase Flow and Heat Transfer in Power Engineering, Shanghai 200093, China*

[c] *Department of Maritime and Mechanical Engineering, Liverpool John Moores University, Liverpool, Byrom Street, L3 3AF, UK*

Abstract

A novel adaptive decomposition algorithm based on CEEMDAN and fractal dimension is proposed in this study to overcome limitations like redundancy and mode confusion in traditional EMD-based algorithms. An intelligent fault diagnosis model is developed using CNN and the proposed CEEMDAN to enhance rolling bearing state recognition. Sub-signals generated by CEEMDAN are selected and reconstructed using PCA and fractal dimension. In feature extraction and pattern recognition, the proposed Improve Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (ICEEMDAN)., coupled with CNN, extracts advanced features from the reconstructed signal for intelligent diagnosis. The methodology is validated through empirical experiments involving rolling bearings, where its superiority and reliability are compared with approaches based on CNN. The accuracy of this method reaches 99.79%

Key Words: complete ensemble empirical mode decomposition with adaptive noise; convolutional neural network; rolling bearing; box dimension; fault diagnosis

1. Introduction

During the explosive development of computer and sensor technology over the years, rotating equipment malfunction monitoring and diagnosis has facilitated the emergence of enormous and high-dimensional big data [1-2]. Energy security and environmental conservation are becoming increasingly essential worldwide, while the quick development of new energy sources such as wind, water, and nuclear energy needs the utilization of effective and reliable rotating equipment [3]. However, rolling bearing of rotating equipment is the most fundamental and susceptible fundamental component, and its steady and efficient operation is closely associated to the overall performance of energy extraction equipment [4-6]. Meanwhile, its vibration signals are not only solely handy to collect, but also rich in a giant quantity of superb fault information. It has become as the preferred signal to analyze the fault traits of rolling bearings, which can also effectively minimize the downtime and useless maintenance triggered caused by equipment failure and lengthen the service life of equipment [7-9]. Therefore, the predominant focal point have to be on rolling bearing situation monitoring and fault diagnosis. The irregular vibration and implications of the device in the actual work setting can easily cause the typical nonlinear and nonstationary characteristics that appear when collecting signal by acceleration sensors, and it is also difficult to guarantee the purity of the signal, posing a serious challenge to effective the features extraction and timely fault warning [10]. The early fault diagnosis process is mainly divided signal acquisition, extraction of fault features, identification and classification, among which the feature extraction is the most critical process in preprocessing [11].

In order to effectively accomplish fault diagnosis tasks in the presence of noise interference, vibration signals must be preprocessed to decrease noise in order to extract fault features of signal more effectively. Time-frequency analysis has recently been popular in the field of fault detection, as it can successfully identify fault features from nonlinear signals under noise, allowing for improved mechanical equipment preventive and maintenance [12-13].A large number of studies have been carried out by many scholars, Gabor [14] proposed the short-time Fourier transform (STFT), which uses a time-frequency representation to obtain the power spectrum at different times by moving the window function to realize fault analysis. Considering the relation of frequency and time resolution, Morlet [15] proposed the wavelet transform, a time-frequency local transform algorithm that receives and expands the localization idea of the STFT.

In order to address the issue of limited adaptability arising from the manual configuration requirement for wavelet transform metrics such as wavelet basis and decomposition layers in the aforementioned methods .Huang et al [16] was the one to propose the empirical mode decomposition (EMD) approach, which can deconstruct an extremely complex noisy signal into a series of intrinsic modal functions (IMFs) without requiring human interaction. Despite its widespread application in the fields of electrocardiogram (ECG) image processing [17], signal filtering [18] and rotating machinery fault diagnosis [19] etc. The EMD was later discovered to have mode aliasing, which has a direct impact on fault diagnostic precision [20].

In order to address the issue of mode aliasing inherent in EMD, Wu and Huang [21] proposed the ensemble empirical mode decomposition (EEMD), which may retain data continuity and suppress mode aliasing by adding white noise to the input signal for EMD. Furthermore, the consequences of a couple of decomposition are averaged to attain the last IMF. Despite the fact that this method has been applied to diagnose rotating machinery faults by the domestic and international scholars[22-24], adding white noise to the raw signal repeatedly via the EEMD method will result in reconstruction errors. YEH [25] presented the complete EEMD (CEEMD) as a key upgrade to EEMD. The CEEMD algorithm can makes sure decomposition accuracy and successfully eliminates residual white noise in the rebuilt signal decomposing the time domain signal by adding the two opposite white noise. Both EEMD and CEEMD, however, generally require a significant amount of computing, and the decomposition is overly reliant on the amplitude of adding white noise and the times of ensemble average [26].

In order to address the computational intensity and the excessive reliance on the amplitude of added white noise and the number of ensemble averages in both EEMD and CEEMD decomposition processes. Improved versions of the CEEMD analysis have been developed to overcome the problem of inefficiency, culminating in the Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (CEEMDAN) [27]. At each level of EMD decomposition, the CEEMDAN adds adaptive white noise and calculates the residual signal to obtain IMF. Furthermore, the decomposition process is comprehensive, and the reconstruction error is relatively tiny regardless of the number of integration times [28]. Moreover, Smith [29] devised the local mean decomposition (LMD) approach, which used the slider mean instead of cubic spline interpolated in EMD but had the similar terminal effect. Eventually, the CEEMDAN method is utilized to filter and extract defect feature information from rolling bearings in this work, outperforming the above traditional modal decomposition strategy.

Precise and reliable extraction of faults signal data is an important factor of fault detection and diagnosis, and it has a direct impact on fault diagnostic recognition rate. Although the time-frequency decomposition method is an effective modal decomposition, precisely selecting the modal components with fault features is problematic. Furthermore, the extracted initial feature components have a high dimension and may contain redundant or insensitive information, adding to the calculation's complexity.

In order to effectively reduce data dimensionality and computational complexity while avoiding dimension disasters. In 1901, Karl Pearson created PCA algorithm which is one of the most extensively used data dimension reduction algorithms today [30]. WANG et al. [31] used PCA algorithm to recognize bearing faults. It can reduce the number of variables in regression and clustering techniques by extracting the largest individual differences from the principal components and reduce the dimension of feature vectors drawn from raw vibration signals, improving real-time performance and fault diagnosis accuracy. This technique has been utilized by several researchers to diagnose rotating machine faults [32-36]. Although the studies above can effectively minimize the complexity of data dimension and calculation and avoid dimensional disasters, they all overlook a critical issue: fault information is not included in a single IMF component, but rather in a number of them.

In order to address the issue of fault information being distributed across multiple IMF components. The fractal dimension proposed by Mandelbrot in 1975 not only better describes the complexity and nonlinear characteristics of vibration signals, but it also has good anti-noise and relatively simple calculation, making fault information pleasant to showcase and improving fault identification performance and generalization ability [37].

In order to address the issue of spurious, redundant, and pseudo components encountered during the CEEMDAN algorithm's processing of nonlinear vibration signals, The application of PCA for the purpose of filtering IMF components is an effective method for mitigating redundancy within the context of processing nonlinear vibration signals[38-39]. We propose a pre-processing filter approach that fractal box dimension combines PCA algorithm.

After selecting low-dimensional sensitive fault feature components for reconstruction, a suitable fault diagnosis approach is used to identify and classify the different types of bearing states. [40]. Traditional defect diagnostic approaches are primarily data-driven and relied on mechanism models. The mechanism model diagnosis approach necessitates the development of a comprehensive mathematical model, therefore its application breadth is limited. The data driver does not need to create a mathematical model; instead, it depends on an expert system or a fault library to perform problem diagnostics [41-42]. Although the approaches listed above have a high fault detection rate, they all require a shallow learning algorithm to recognize fault diagnosis since they cannot learn their own characteristics, self-adaptive fault feature extraction, and weak model generalization. Convolutional neural networks are one of the most prominent deep learning models, capable of combining feature extraction with state categorization. Furthermore, its convolutional kernel can adaptively train and extract fault features in signals, avoiding the error of artificial feature extraction and selection, and improving fault diagnostic accuracy. It's also been commonly used in the industrial industry [43-46]. The results show that CNN is not only capable of digesting large amounts of high-dimensional data, but also of self-learning. However, it is still not possible to exclude external noise interfering with defect diagnostics. A new rolling bearing fault diagnosis approach based on ICEEMDAN fusion deep learning is presented for these reasons. The following are the article's primary contributions and superiorities:

(i) We proposed ICEEMDAN method to solve the limitations in EMD、EEMD or CEEMDAN algorithm for dealing with unstable signals.

(ii) We proposed a new signal fusion method for IMF modes reconstruction based on PCA algorithm and fractal box dimension.

(iii) We developed an intelligent ICEEMDAN-CNN model for fault diagnosis, which combining with the advantages of filter.

The rest of the paper is laid out in the following states. Section 2 discusses the strategies for recognizing states and preprocessing. Section 3 describes the new structure of the ICEEMDAN-CNN model framework. Section 4 presents the test experimental data as well as the validation of the ICEEMDAN-CNN approach. Section 5 wraps off with the findings.

2. The relevant theory of state recognition and preprocessing

2.1 CEEMDAN algorithm

CEEMDAN is the ultimate improved algorithm based on Torres' EMD [47]. Using adaptive white noise, it is possible to successfully eliminate modal aliasing, redundant and false components. It also avoids the need to compute the various order components. The specific procedural steps of CEEMDAN encompass the following sequence: initially, introducing adaptive white noise into the original signal; subsequently, decomposing the signal. For each IMF component, employing a collective averaging method, iteratively repeating the process of decomposition and noise introduction to further enhance the precision and stability of the decomposition. Lastly, amalgamating all the processed IMF components culminates in the ultimate CEEMDAN decomposition result. The CEEMDAN decomposition steps are shown in Fig.1.

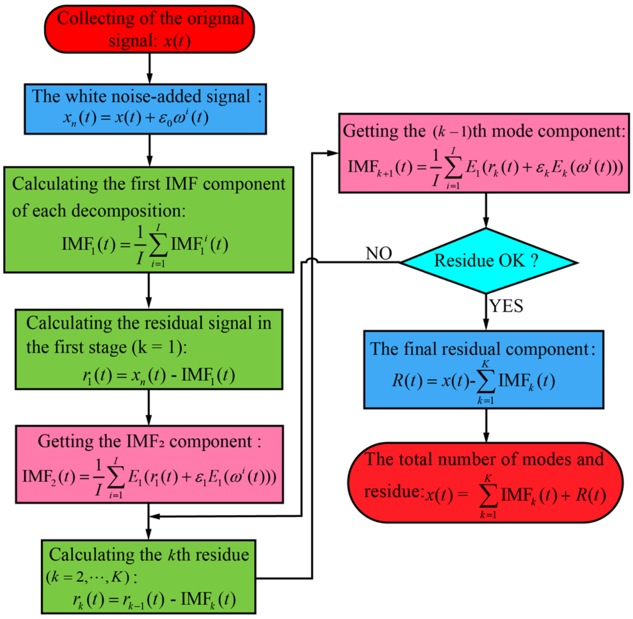


Fig.1 The flow chart of CEEMDAN

where $E_i(\cdot)$ is the k th IMF decomposed by the EMD, $\omega^i(t),(i=1,\cdots,I)$ is the added white noise, I is the number of times to add white noise.

2.2 Principal component analysis

Principal component analysis is a method for preparing high-dimensional feature data. It can keep the most critical elements of high-dimensional data while removing complex noise and unexpected characteristics in order to improve data processing performance. As a result, PCA has been widely employed in exploratory data analysis and the development of prediction models. It is widely used to reduce dimensionality through trying to project each data point onto only the first few principal components to get lower-dimensional data while retaining as much of the data's variation as feasible [48]. The steps of PCA [49] as follows:

Step 1: Normalize the original feature space data $X=\{x_1,x_2,x_3,\cdots,x_n\}$ to get standardized X^*.

Step 2: Calculate the covariance matrix $\mathrm{cov}(X^*)$.

Step 3: Using eigenvalue decomposition to calculate the eigenvalues and eigenvectors of covariance matrix $\mathrm{cov}(X^*)$.

Step 4: Ordering eigenvalue and Choosing components and forming a feature vector.

Step 5: Transform the data into a new space constructed by feature vectors.

2.3 Fractal theory

The original meaning of fractal is irregular, fractional, and fragmented things, which can be regarded as the similarity between the part and the whole in some ways [50]. Fractal dimension is a useful metric for measuring fractals, as it accurately describes the complexity and nonlinearity of vibration signals [51]. Although there are a variety of fractal dimensions that can be used to describe the complexity of signals, such as Hausdorf dimension, box dimension, capacity dimension, information dimension, correlation dimension, etc. The box dimension is now the most widely used because of its simple and efficient calculation [52]. As a result, it is preferred among nonlinear field researchers [53-55]. The definition of fractal box dimension is Eq. (1) [35].

$$D_B = \lim_{\varepsilon \to 0} \frac{\ln N(\varepsilon)}{\ln(1/\varepsilon)} \tag{1}$$

where $Y \subset R^n, Y \neq \Phi$, if there $N(\varepsilon)$ hypercube can cover Y.

2.4 CNN Classification Model

Convolutional neural network is derived from the neurons of primate visual nervous system. It not only has neural network with deep structure, but also has powerful data mining and feature extraction capabilities [56]. Local receptive field weight sharing and down-sampling, as its unique features, can not only realize the deep mining of data features, but also enhance the self-learning ability of data features, eliminating algorithm overfitting well [57]. The CNN model structure is shown in Fig.2.

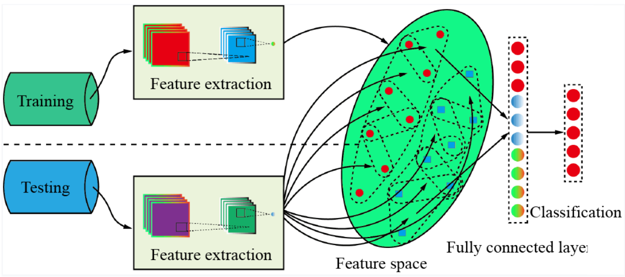


Fig.2 The model structure of CNN

2.4.1. Convolutional and Pooling layer

The convolutional layer is the CNN core, which uses the product and reconstruction of corresponding regions overlapped by the convolutional kernel and the input features, and achieves feature information extraction by adding bias to obtain the feature values. Although the feature information extraction ability is improved through the convolutional layer, the dimension of data is increased, resulting dimension disaster. However, the pooling layer can reduce the number of parameters while retaining its key features to achieve the purpose of reducing and screening the main features [58].

2.4.2. Fully connected layer and Dropout

The full connection layer can classify feature information effectively, and the hidden layer of multi-layer perceptron can better integrate the data information after convolutional pooling [59]. The regularization technique of Dropout can omit some elements of the hidden layer to prevent overfitting and improve the generalization performance of the model [60]. The effect schematic diagram of Dropout operation is shown in Fig.3.

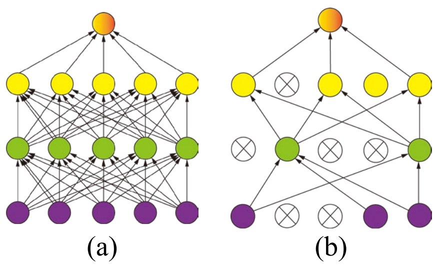


Fig.3 The effect schematic diagram of Dropout operation: (a) Fully connected network of standard; (b)The net after Dropout

3. The proposed ICEEMDAN-CNN model architecture structure

3.1 ICEEMDAN algorithm

Although the CEEMDAN algorithm with adaptive adding white noise can effectively reduce the mode aliasing problem, it still cannot completely eliminate the influence of redundant components and false components, which interferes with the selection of principal components. However, Pearson correlation coefficient, kurtosis and grey correlation have been used to screen the optimal IMF components by a large number of scholars [61-62-63-64]. However, they all ignore the fault information often exists in some IMF components, which makes it difficult to extract all effective information completely. In this paper, the combine method of PCA and fractal dimension is used to improve CEEMDAN. PCA can extract effective fault information by dimensionality reduction of data, and use cross analysis method to calculate the fractal dimension before and after dimensionality reduction to reconstruct the optimal component group. According to the fractal theory, the fractal dimension has a positive correlation with the signal stability. The flow chart of the improved CEEMDAN algorithm is shown in Fig.4.

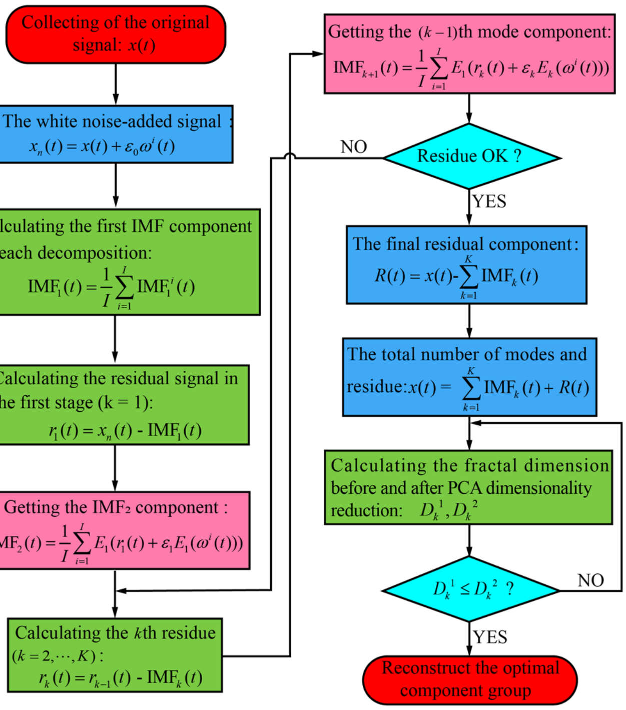


Fig.4 The improved CEEMDAN flow chart

3.2 ICEEMDAN model architecture

The ICEEMDAN-CNN model architecture proposed in this paper consists of CEEMMDAN decomposition denoising, PCA dimensionality reduction processing, fractal dimension screening fault feature component reconstruction group, feature learning layer and classification layer of one-dimensional CNN. Figure 5 shows the proposed ICEEMDAN-CNN model framework structure.

As shown in Fig. 5, the proposed method is divided into two steps altogether: model training and verification. The collected vibration signals need to be decomposed by CEEMDAN, avoiding effectively endpoint effect and reducing redundant and residual noise in IMF component. Furthermore, PCA method is used to reduce the dimension of the decomposed high-dimensional components, which can map the faults that are difficult to identify to another subspace for dimensionality reduction to extract key feature information and improve the ability of fault feature extraction. According to the dimension properties of fractal box, the best reconstruction components can be screened out, eliminating redundant components and false components. Finally, it is input into CNN with powerful data processing advantages to realize accurate identification and classification of faults through the full connection layer and SOFTMAX.

The structural parameters of one-dimensional CNN are shown in detail in Table 1. Compared with the traditional CNN, Adam gradient descent optimization algorithm is adopted in the proposed model. The CNN is composed of five convolutional layers and five pooling layers in an alternating fashion, followed by dropout and a fully connected layer. The learning rate has been maintained at 0.001. The length of one-dimensional signal under random screening conditions is 2048. Its initial weights are randomly initialized, and the activation function is ReLU. The cross entropy is the loss function. In addition, multi-layer network structures such as small-size convolution and pooling, Batch Normalization, and Dropout are used continuously. It can not only simplify the calculation, avoid gradient explosion, over-fitting and gradient disappearance, but also improve the performance of fault diagnosis identification and classification [65].

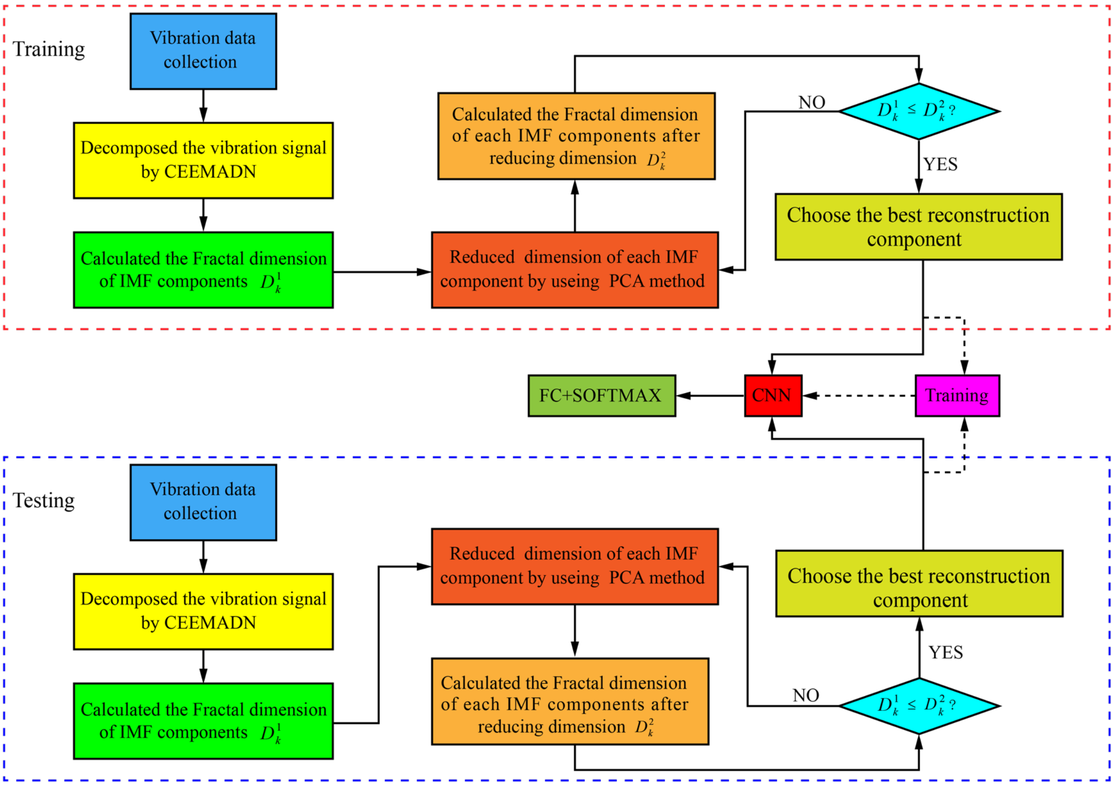


Fig.5 The ICEEMDAN and CNN model architecture

Table.1 The parameter of CNN network structure

Layer (Type)	Convolution kernels	Step length	Size
Input layer	---	---	1@2048×1
Layer 1	16@11×1	[4 1]	16@510×1
Batch Normalization (BN)	---	---	---
Pooling layer 1	---	[2 1]	16@255×1
Layer 2	32@5×1	[2 1]	32@126×1
Batch Normalization (BN)	---	---	---
Pooling layer 2	---	[2 1]	32@63×1
Layer 3	32@3×1	[1 1]	32@61×1
Batch Normalization (BN)	---	---	---
Pooling layer 3	---	[2 1]	32@30×1

Layer 4	64@2×1	[1 1]	64@29×1
Batch Normalization (BN)	---	---	---
Pooling layer 4	---	[2 1]	64@14×1
Layer 5	128@2×1	[1 1]	128@13×1
Batch Normalization (BN)	---	---	---
Pooling layer 5	---	[1 2]	128@6×1
Dropout (DR)	---	---	128@6×1
Full connection layer	---	---	4@1×1
Softmax	---	---	---

4. Experiments and verification method

4.1 Sample dataset

For better evaluation of reality and integration with reality, the Xi'an Jiao-tong University experimental data [66], as the standard bearing vibration data set, is used to examine the performances of three kinds of decomposition method that include EMD, EEMD and CEEMDAN and to compare the performance with CNN-based models for proving the superiority of the ICEEMDAN-CNN model. The bearing experimental test platform is shown in Fig.6. The motor speed of the test platform was 2100 r/min and 2250 r/min, The sensor utilized is an accelerometer sensor. and the sampling frequency was 25.6 kHz. The sampling duration for each instance is 1.28 seconds, with a sampling interval of one minute. Consequently, each dataset comprises 32,768 data points, and a total of 100 datasets were collected for each operational condition. The test conditions [67] are shown in Table 2.

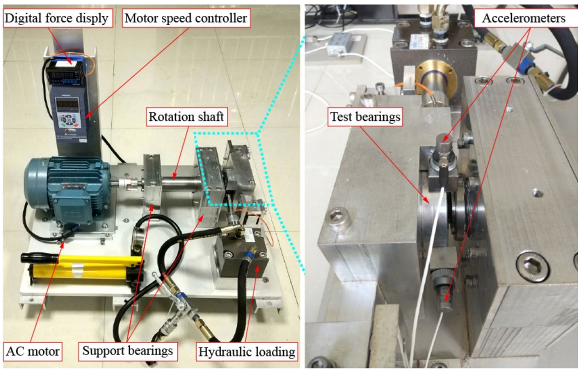


Fig.6 The test platform of bearing

Table.2 The test conditions of bearing

Fault types	Mixed damage	Inner ring wear	Cage wear	Outer ring wear
Speed / (r/min)	2100	2250	2250	2250
Radial force / kN	12	11	11	11

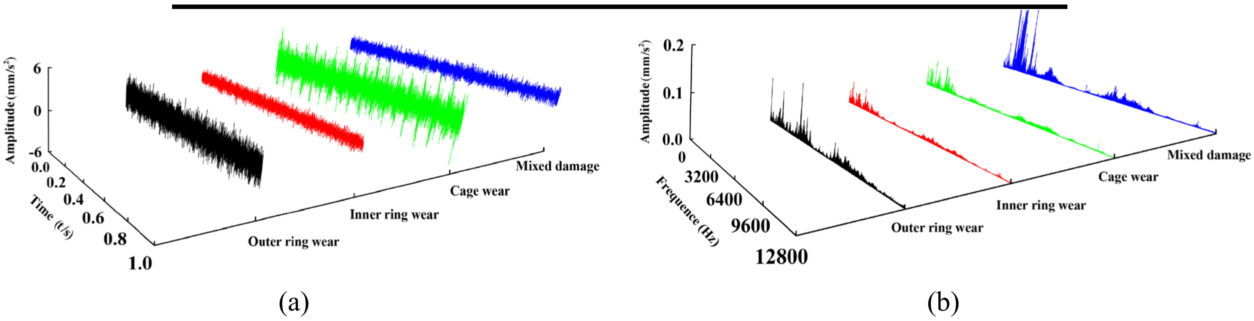


Fig.7 Time and frequency domain waveform of bearing vibration signals: (a) The bearing vibration signal time

domain waveform; (b) The bearing vibration signal frequency domain waveform

As can be seen from Table 2, there are four faults of mixed damage, inner ring wear, cage wear and outer ring wear. Their time domain diagram is shown in Fig.7. Although there are some differences in time domain and frequency domain of vibration signals of different faults of rolling bearings, fault identification and classification cannot be carried out directly, so it is more difficult to ensure the diagnosis accuracy. However, the extraction of pure and effective fault features is the basis of diagnosis, so the original signal should be de-noised to highlight the fault information and enhance the practicability of the signal.

4.2 Data preprocessing and method analysis

In this paper, the original signal was processed by EMD, EEMD and CEEMDAN algorithms respectively to achieve noise reduction. However, there are too much data to display all of them directly. Therefore, only the decomposition results of inner ring wear are shown in Fig. 8. The time domain and frequency domain plots of the three kinds of decomposition results cannot accurately select the best reconstructed component groups for fault analysis. Compared with EMD EEMD algorithm, CEEMDAN algorithm adaptively adds white noise to improve the effect of mode aliasing, but it still can not accurately remove redundant components and false components.

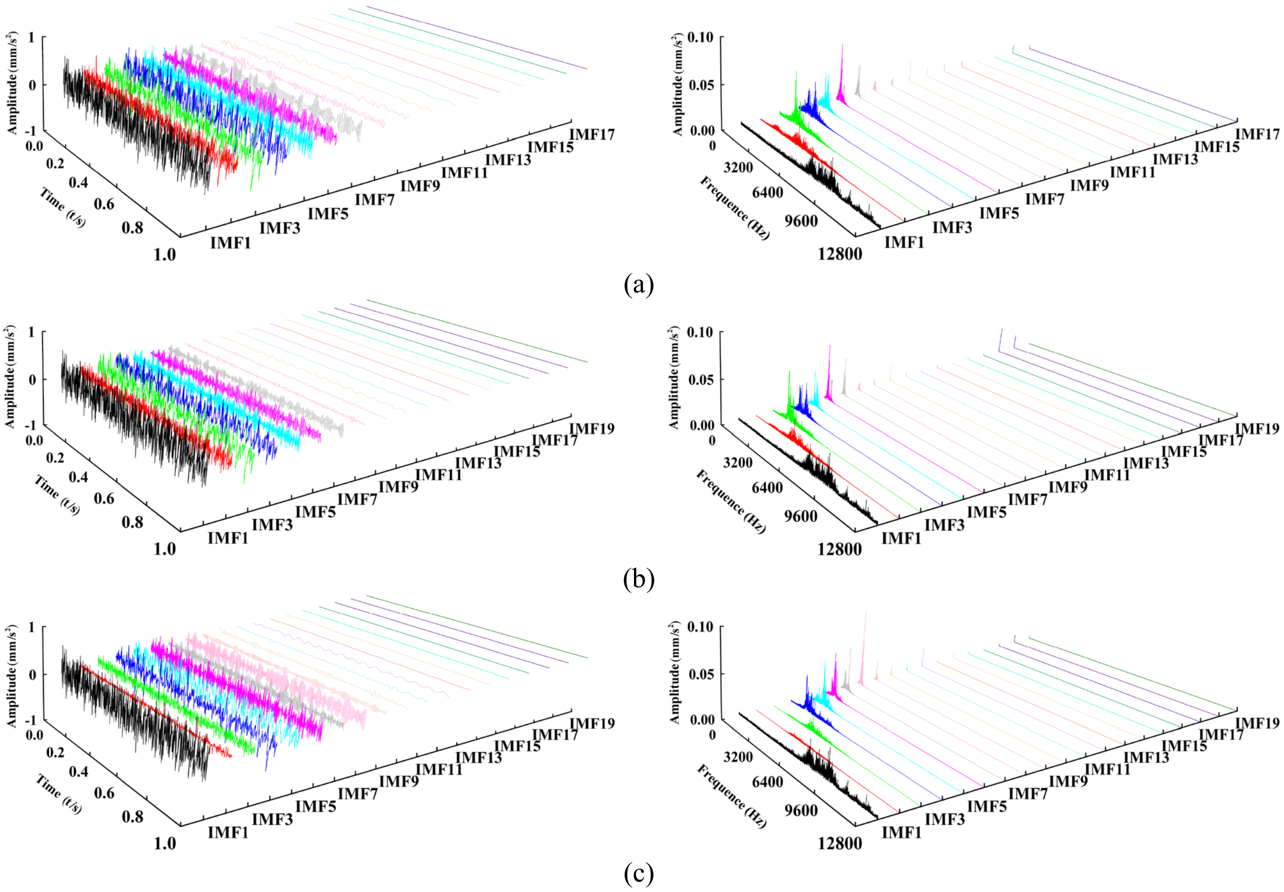


Fig.8 The different decompose of inner ring wear fault signals: (a) The result of EMD decomposed; (b) The result of EEMD decomposed; (c) The result of CEEMDAN decomposed

Based on the above problems, this paper proposes a method combining PCA dimension reduction and fractal dimension to screen the best reconstruction components, which can effectively avoid the loss of fault information, eliminate redundant components and improve the identification accuracy of fault diagnosis. Fig.9 shows the comparison results of fractal dimension before and after dimension reduction of four kinds of rolling bearing faults.

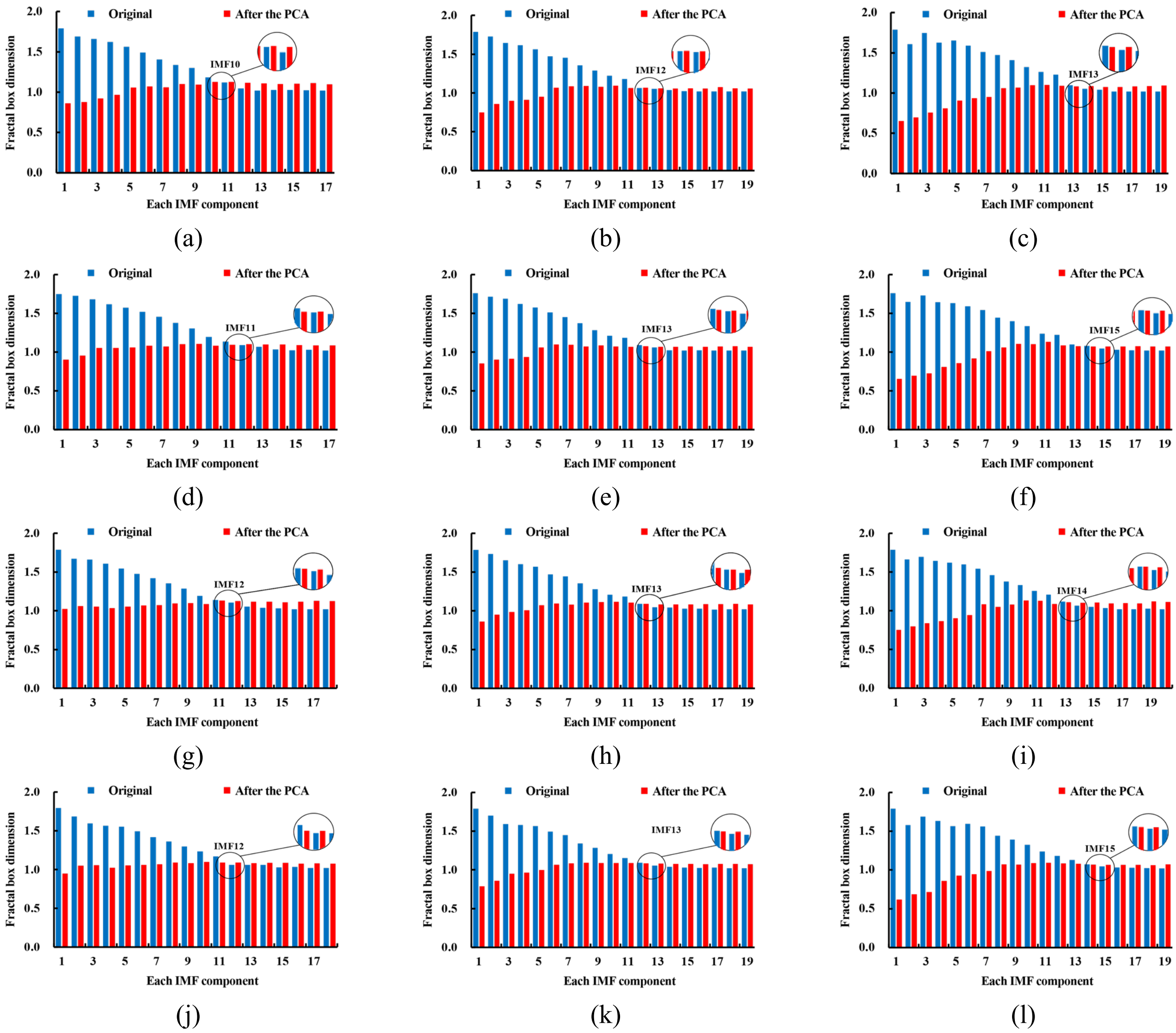


Fig.9 The different methods are used to screen the optimal reconstruction components of each fault: (a) Inner ring wear (EMD); (b) Inner ring wear (EEMD); (c) Inner ring wear (CEEMDAN); (d) Outer ring wear (EMD); (e) Outer ring wear (EEMD); (f) Outer ring wear (CEEMDAN); (g) Mixed damage (EMD); (h) Mixed damage (EEMD); (i) Mixed damage (CEEMDAN); (j) Cage wear (EMD); (k) Cage wear (EEMD); (l) Cage wear (CEEMDAN)

Fig.9 shows that fractal dimension is negatively correlated before and after PCA dimension reduction, and CEEMDAN algorithm is significantly superior to EMD and EEMD. Based on the properties of fractal box dimension, the size of box dimension can indirectly judge the stability of signal. Before using PCA to reduce dimension, the residual noise in IMF component directly affects the overall stability, which leads to the box dimension decreasing gradually with the component. After PCA dimensionality reduction, main fault information is further extracted and purified to enhance the stability of data. The cross method can not only avoid the influence of residual noise, but also eliminate redundant components and false components. The time domain diagram of the optimal reconstruction component group of each fault is shown in Fig.10.

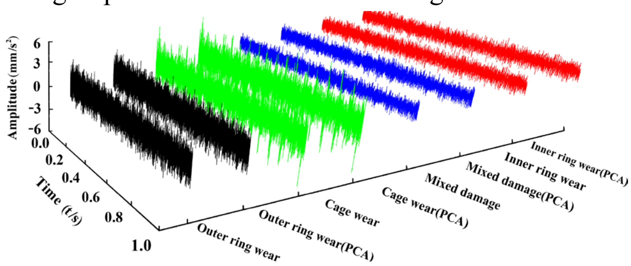


Fig.10 The Time domain diagrams of the optimal reconstructed component group and the original signal component

4.3 Model performance verification

In order to better validate the proposed improved method and its performance, this paper compares ICEEMDAN with EMD and EEMD algorithms: and inputs the optimized reconstructed filter component group into CNN for fault diagnosis. In addition, The length of each segment of the original signal is 2048, and they have been divided into training, testing, and validation sets in a ratio of 8:1:1. The training process consists of 10 epochs, with 10 iterations per epoch. The accuracy and loss of the validation set of the best reconstructed component in the absence of noise are shown in Fig.11.

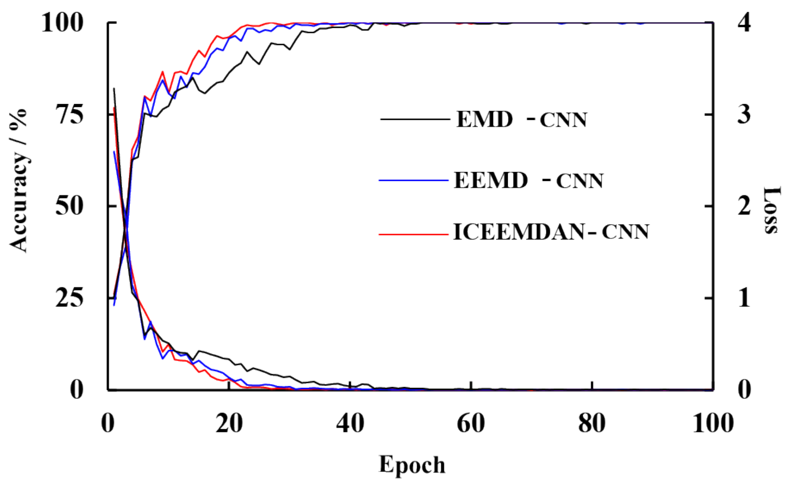
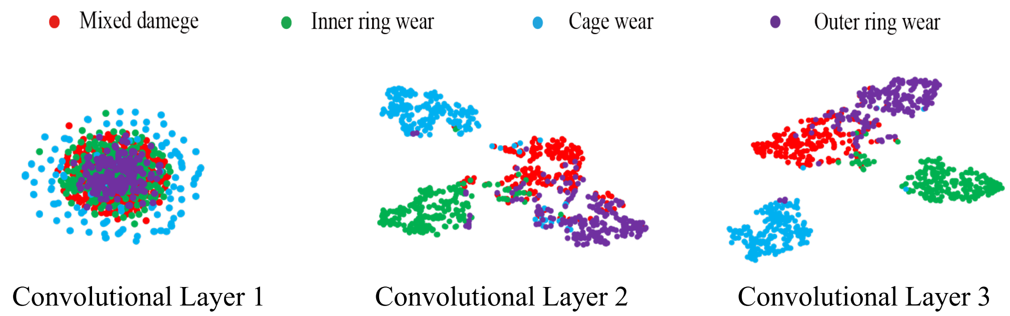


Fig.11 The accuracy and loss of validation for optimal reconstruction component

As can be seen from Fig.11, the ICEEMDAN-CNN method proposed in this paper is superior to other algorithms in both accuracy and loss. However, there are mixed noises in the actual environment. In order to further verify the practicability and generalization of the method, it is necessary to add noises with different SNR to the signal to restore the actual operating environment as much as possible.

4.4 Visualization and Generalization comparison validation

Deep network learning is mostly based on the analysis of data attributes, which is difficult to restore the actual complex operating conditions. Therefore, this paper not only adds noise-assisted comparison verification, but also uses a comparison method to screen the optimal component groups with different entropy. It is compared with the sample entropy and fuzzy entropy fault diagnosis methods which are widely used in many fields [68-69]. The t-SNE clustering visualization analysis results under different SNR are shown in Fig.12.



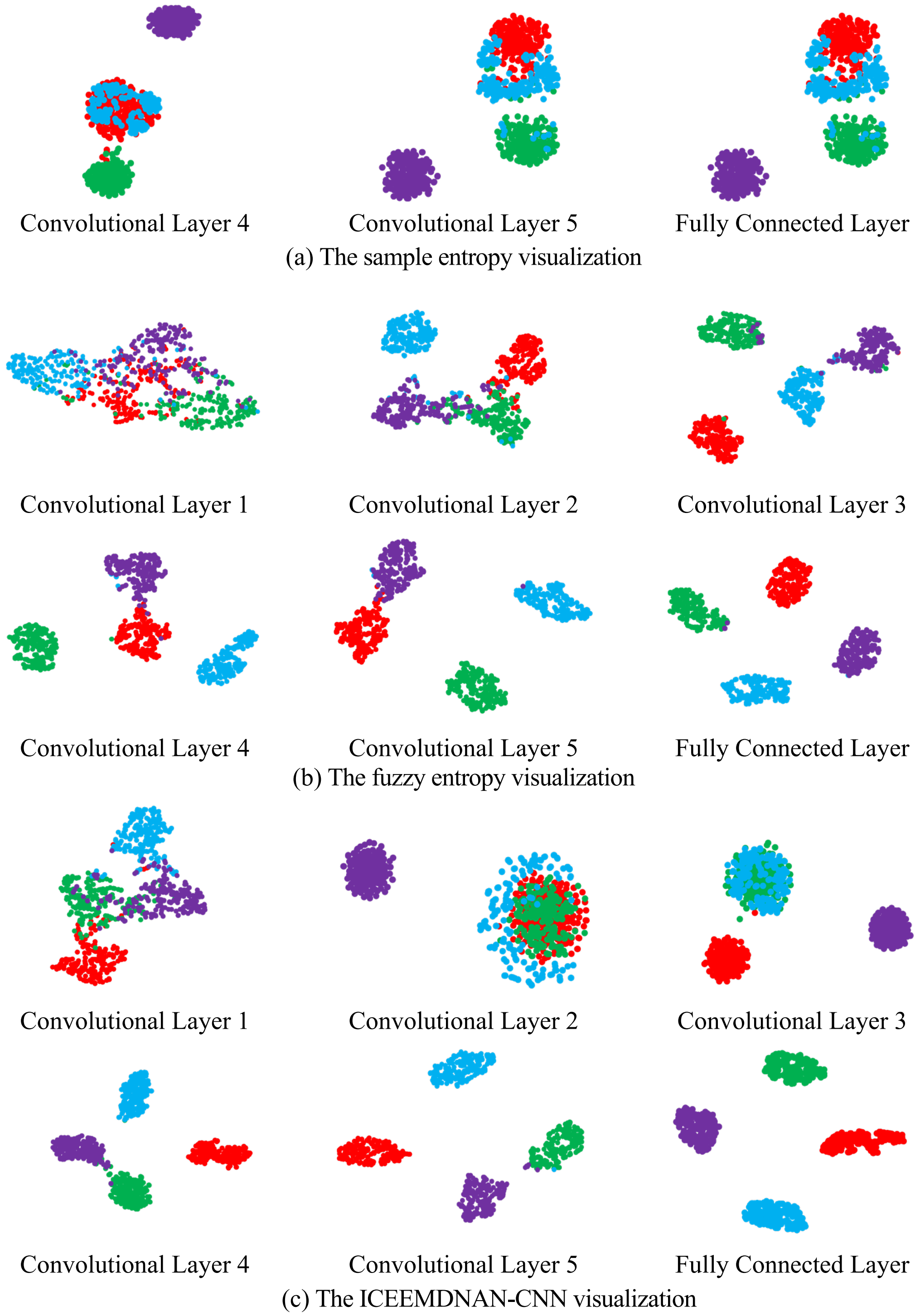


Fig.12 Visual analysis results of t-SNE clustering

Fig.12 shows that the data processed by deep network can be well classified by t-SNE clustering analysis method for the four kinds of rolling bearing faults. The proposed ICEEMDAN-CNN method can completely separate the four faults, while the fuzzy entropy method is better than the sample entropy screening method, but they still cannot achieve the best clustering effect. The effect of noise on fault diagnosis can be better observed from the visualization results. It is evident that the choice of distinct methodologies for the selection and reconstruction of IMFs during the CEEMDAN procedure can exert a discernible influence on the ultimate performance of the entire model, The method substitutes PCA for sample entropy or fuzzy entropy to select and reconstruct IMF components during the CEEMDAN process. in this paper can better perform data mining for massive high-dimensional big data, analyze the laws hidden behind the data, realizing fault classification and visual analysis. Therefore, the ICEEMDAN-CNN method proposed in this paper can better conduct data mining for massive and high-dimensional big data, exploring the laws hidden behind the data and realizing fault classification and visual analysis. In order to further highlight the good generalization performance of ICEEMDAN-CNN algorithm proposed in this paper, noises with different SNR were added to original signals and compared with existing methods respectively, and the results are shown in Fig.13.

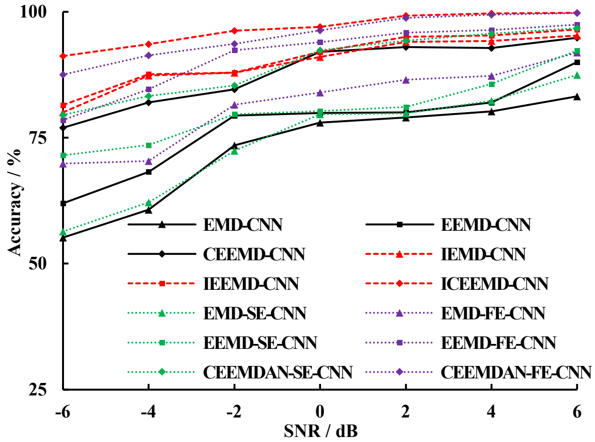


Fig.13 The recognition accuracy of each method under different SNR

As shown in Fig.13, The proposed ICEEMDAN-CNN algorithm has obvious advantages and good generalization performance compared with other existing algorithms. The actual operating environment can be restored well under different SNR, and the correlation between SNR and accuracy is positive, which indirectly reflects the influence of noise on diagnosis accuracy. The improved algorithm has higher recognition and classification accuracy than the original algorithm. The accuracy of the proposed method is up to 99.79%, and the recognition accuracy is still 87.13% at the lowest SNR of -6dB, which is 0.54 - 10.33% higher than other algorithms. However, there is still room for improvement in this method at low (SNR)." In addition, it can eliminate redundant signals and false components well and realize noise reduction, enhance the accuracy of extracting effective fault signs, and further improve the accuracy of fault diagnosis recognition and classification.

5. Conclusions

A novel fault diagnosis method of rolling bearing is proposed using CNN and PCA fractal based feature extraction in this paper. The method can effectively solve the problems of redundant components and false components in the decomposition process of existing methods, screening also accurately the optimal component group. CEEMDAN algorithm is used to process raw signals to achieve noise reduction and decomposition. PCA can efficiently extract effective fault features by reducing the dimension of high-dimensional data, and fractal box dimension filters the best reconstruction component groups to eliminate irrelevant components. Finally, CNN further excavates the optimal component group to realize fault diagnosis recognition and classification. In addition, the effectiveness and feasibility of this method are verified by a variety of data verification and comparison with existing methods. The specific conclusions are as follows:

(i) The proposed model framework of ICEEMDAN-CNN fault diagnosis, testing by experiment, can effectively filter out the noise disturbance and accurately extract the effective fault features, achieving better classification effect of four kinds of rolling bearing faults and reducing the diagnosis error.

(ii) The PCA and fractal box dimension combine method are used to select the best reconstructed component groups, which can effectively eliminate redundant components and false components. The reconstructed component group is input into CNN with strong nonlinear fitting ability, which can adaptively extract features to eliminate the interference caused by human factors and improve the accuracy of CNN fault identification and classification. The robustness and feasibility of the proposed method are verified by rolling bearing fault analysis under different working conditions.

(iii) Compared with the existing fault diagnosis models, the proposed ICEEMDAN-CNN model has the highest recognition accuracy by 99.79% at different SNR. Meanwhile, the generalization of the proposed model method is superior to EMD-CNN, EEMD-CNN, CEEMDAN-CNN, CEEMDAN-SE-CNN, CEEMDAN-FE-CNN, IEMD-CNN, IEEMD-CNN etc.

The following future the optimization model analyses are need to investigate multifractal and multiscale convolutional neural networks. In addition, the effect of adding different forms of noise on generalization can also be considered.

Acknowledgement

The authors would like to acknowledge the financial support of National Natural Science Foundation (NSFC) of China (grant numbers: 51976131 and 52006148). The National Science Fund for Distinguished Young Scholars (grant numbers: 52106262). Science and Technology Commission of Shanghai Municipality (grant number: 1906052200), the Xi'an Jiao-tong University for providing the benchmarking datasets of rolling bearing vibration condition monitoring.

References

[1] Chen Z, Wang Y, Wu J, et al. Sensor data-driven structural damage detection based on deep convolutional neural networks and continuous wavelet transform[J]. Applied Intelligence, 2021, 51(3): 5598-5609.

[2] Abhishek Dhananjay Patange,Jegadeeshwaran R..A machine learning approach for vibration-based multipoint tool insert health prediction on vertical machining centre (VMC)[J]. Measurement, 2021, 108649.

[3] Xu Z, Xuan M, Wang X, Yue M, et al. Fault diagnosis of wind turbine bearing using a multi-scale convolutional neural network with bidirectional long short term memory and weighted majority voting for multi-sensors. Renewable Energy, 2021: 615-626.

[4] Shan P, Lv H, Yu L, et al. A Multisensor Data Fusion Method for Ball Screw Fault Diagnosis Based on Convolutional Neural Network with Selected Channels. IEEE Sensors Journal, 2020, 20(14): 7896-7905.

[5] Wei Q, Lu D. A Survey on Wind Turbine Condition Monitoring and Fault Diagnosis - Part II: Signals and Signal Processing Methods. IEEE Transactions on Industrial Electronics, 2015, 62(10):1-1.

[6] Lu O, Yu D, Yang H. A new rolling bearing fault diagnosis method based on GFT impulse component extraction. Mechanical Systems & Signal Processing, 2016, 81: 162-182.

[7] Siddiqui O, Dincer I. Comparative Assessment of the Environmental Impacts of Nuclear, Wind and Hydro-Electric Power Plants in Ontario: A Life Cycle Assessment. Journal of Cleaner Production, 2017, 164(15): 848-860.

[8] Li C, Sanchez V, Zurita G, et al. Rolling element bearing defect detection using the generalized synchosqueezing trasform guided by time-frequency ridge enhacement. ISA Transactions, 2015(60): 274-284.

[9] Hu Z, Wang Y, Ge M, et al. Data-driven Fault Diagnosis Method based on Compressed Sensing and Improved Multi-scale Network. IEEE Transactions on Industrial Electronics, 2020, 67(4): 3216-3225

[10] Du Y, Zhang W, Zhang Y, et al. Fault diagnosis of rotating machines for rail vehicles based on local mean decomposition energy moment directed a cyclic graph support vector machine. Advances in Mechanical Engineering, 2016, 8(1): 1-6.

[11] He W, Zi Y, Chen B, et al. Automatic fault feature extraction of mechanical anomaly on induction motor bearing using ensemble super-wavelet transform. Mechanical Systems and Signal Processing, 2015, 54: 457-480.

[12] Shi J, Liang M, Necsulescu D S, et al. Generalized stepwise demodulation transform and synchrosqueezing for time﹣frequency analysis and bearing fault diagnosis. Journal of Sound & Vibration, 2016, 368: 202-222.

[13] Wang L, Liu Z, Miao Q, et al. Time-frequency analysis based on ensemble local mean decomposition and fast kurtogram for rotating machinery fault diagnosis. Mechanical Systems & Signal Processing, 2018, 103: 60-75.

[14] Gabor D. Theory of communication. Journal of the Institution of Electrical Engineers - Part I: General, 1946, 94: 58-58.

[15] Polikar R. The story of wavelets. Physics and Modern Topics in Mechanical and Electrical Engineering. World Scientific and Engineering Academy and Society. 1999: 5481-5486.

[16] Huang N E, Shen Z, Long S R, et al. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Proceedings Mathematical Physical & Engineering Sciences, 1998, 454(1971): 903-995.

[17] Suchetha M, Kumaravel N, Jagannath M, et al. A Comparative Analysis of EMD based Filtering Methods for 50Hz Noise Cancellation in ECG Signal. Informatics in Medicine Unlocked, 2017: S2352914817300072.

[18] Xiong H, Zheng C, Liu J, et al. ECG Signal In-Band Noise De-Noising Base on EMD. Journal of Circuits, Systems and Computers, 2018, 28(1): 1950017.1-1950017.13.

[19] Li S, Sun Y, Wang X. Fault diagnosis of rolling bearing in multi-dimensional entropy space// 2020 Chinese Control And Decision Conference (CCDC). IEEE, 2020: 4689-4694.

[20] Sagar M, Vivekkumar G, Reddy M, et al. Research on intelligent fault diagnosis of gears using EMD, spectral features and data mining techniques. IOP Conference Series: Materials Science and Engineering, 2017, 263: 062047-.

[21] Wu Z, Huang N E. Ensemble empirical mode decomposition a noise-assisted data analysis method. Advances Adaptive Data Analysis, 2011, 1(01): 1-41

[22] Wang H, Chen J, Dong G. Feature extraction of rolling bearing's early weak fault based on EEMD and tunable Q-factor wavelet transform. Mechanical Systems and Signal Processing, 2014, 48: 103-119.

[23] Amarouayache I, Saadi M N, Guersi N, et al. Bearing fault diagnostics using EEMD processing and convolutional neural network methods. International Journal of Advanced Manufacturing Technology, 2020, 107(4): 1-19.

[24] Xu F, Song X, Tsui K L, et al. Bearing Performance Degradation Assessment Based on Ensemble Empirical Mode Decomposition and Affinity Propagation Clustering. IEEE Access, 2019, 7(99): 54623-54637.

[25] Yeh J R, Shieh J S, Huang N E. Complementary ensemble empirical mode decomposition:a novel nos enhanced data analysis method. Advances in Adaptive data analysis, 210, 2(2) 135-156.

[26] Chen J, Zhou D, Lyu C, et al. An integrated method based on CEEMD-SampEn and the correlation analysis algorithm for the fault diagnosis of a gearbox under different working conditions. Mechanical Systems and Signal Processing, 2017: 102-111.

[27] Torres M E, Colominas M A, Schlotthauer G, et al. A complete ensemble empirical mode decomposition with adaptive noise//Proceedings of 2011 IEEE International Conference on Acoustics, Speech and Signal Processing. Praque. Czech Republic: IEEE, 2011: 4144-4147.

[28] Wang L, Shao Y. Fault feature extraction of rotating machinery using a reweighted complete ensemble empirical mode decomposition with adaptive noise and demodulation analysis. Mechanical systems and signal processing, 2020, 138: 106545.1-106545.20.

[29] Smith, Jonathan S. The local mean decomposition and its application to EEG perception data. Journal of The Royal Society Interface, 2005, 2(5):443-454.

[30] Pearson K. On lines and planes of closest fit to systems of points in space. Philosophical Magazine, 1901, 2(11): 559-572.

[31] Wang F, Sun J, Yan D, et al. A Feature Extraction Method for Fault Classification of Rolling Bearing based on PCA. Journal of Physics Conference, 2015, 628: 012079.

[32] Zhang D, Chen Y, Guo F, et al. A New Interpretable Learning Method for Fault Diagnosis of Rolling Bearings. IEEE

Transactions on Instrumentation and Measurement, 2021, 70 (99): 1-10.
[33] Ahmed H, Nandi A K. Three-Stage Hybrid Fault Diagnosis for Rolling Bearings with Compressively Sampled Data and Subspace Learning Techniques. IEEE Transactions on Industrial Electronics, 2019, 66(7): 5516-5524.
[34] Moshen K , Gang C , Pang Y , et al. Research of Planetary Gear Fault Diagnosis Based on Permutation Entropy of CEEMDAN and ANFIS. Sensors, 2018, 18(3): 782.
[35] Hu K, Jiang M, Zhang H, et al. Design of fault diagnosis algorithm for electric fan based on LSSVM and Kd-Tree[J]. Applied Intelligence, 2021, 51(2): 804-818.
[36] Wu J, Wu C, Cao S, et al. Degradation Data-Driven Time-To-Failure Prognostics Approach for Rolling Element Bearings in Electrical Machines. IEEE Transactions on Industrial Electronics, 2019, 66(1): 529-539.
[37] Albers D, Alexanderson G L. Mathematical people. 2008.
[38] Wang J, Zhang S, Li C, et al. A data-driven method with mode decomposition mechanism for remaining useful life prediction of lithium-ion batteries. IEEE Transactions on Power Electronics, 2022, 37(11): 13684-13695.
[39] Li X,Tang G,Ren Y, et al. Combined CEEMDAN-CNN-BiLSTM-ATT Model for Forex Forecasting[J]. Academic Journal of Computing & Information Science,2023,6(7).
[40] Gao P, Wu W, Li J. Multi-source fast transfer learning algorithm based on support vector machine[J]. Applied Intelligence, 2021: 1-15.
[41] Aljemely A H, Xuan J, Xu L, et al. Wise-local response convolutional neural network based on Nave Bayes theorem for rotating machinery fault classification[J]. Applied Intelligence, 2020, 51: 6932-6950.
[42] Kale Archana P.,Wahul Revati M.,Patange Abhishek D.,Soman Rohan,Ostachowicz Wieslaw.Development of Deep Belief Network for Tool Faults Recognition[J].Sensors,2023,23(4).
[43] Goodfellow I, Bengio Y, Courville A. Deep learning. Cambridge, MA, US: MIT Press, 2016: 326-366.
[44] Liu X, Tian Y, Lei X, et al. Deep forest based intelligent fault diagnosis of hydraulic turbine. Journal of Mechanical Science and Technology, 2019, 33(3): 2049-2058.
[45] Song X, Cong Y, Song Y, et al. A bearing fault diagnosis model based on CNN with wide convolution kernels. Journal of Ambient Intelligence and Humanized Computing, 2021, :1-16.
[46] Li Y T, Jiang W B, Zhang G Y, et al. Wind turbine fault diagnosis based on transfer learning and convolutional autoencoder with small-scale data. Renewable Energy, 2021, 171: 103-115.
[47] Torres M E, Colominas M A, Schlotthauer G, et al. A complete ensemble empirical mode decomposition with adaptive noise[C]//Proceedings of 2011 IEEE International Conference on Acoustics, Speech and Signal Processing. Praque. Czech Republic: IEEE, 2011: 4144-4147.
[48] Wang T, Xu H, Han J, et al. Cascaded H-Bridge Multilevel Inverter System Fault Diagnosis Using a PCA and Multiclass Relevance Vector Machine Approach. IEEE Transactions on Power Electronics, 2015, 30(12): 7006-7018.
[49] Gu Y K, Zhou X Q, Yu D P, et al. Fault diagnosis method of rolling bearing using principal component analysis and support vector machine. Journal of Mechanical Science and Technology, 2018, 32(11): 5079-5088.
[50] Mandelbrot B B. The Fractal Geometry of Nature. San Francisco: Freeman, 1982.
[51] Zhang Z Y, Wu J D, Ma J, et al. Fault diagnosis for rolling bearing based on lifting wavelet and morphological fractal dimension, [C]//The 27th Chinese Control and Decision Conference (2015 CCDC), Praque. Czech Republic: IEEE, 2015: 6351-6354.
[52] James T S. The geometry of fractal sets. London: Cambridge University Press,1985.
[53] Gu H, Liu X X, Zhao B L, et al. A Compensation Method for Long-term Zero Bias Drift of MEMS Gyroscope Based on Improved CEEMD and ELM// 2018: 13-14.
[54] Zhao J, Li H, Teng H, et al. A novel method for detecting bearing defects based on EMD and fractal dimension. Vibroengineering

Procedia, 2016,10: 120-125.

[55] Li X, Ma J, Wang X, et al. An improved local mean decomposition method based on improved composite interpolation envelope and its application in bearing fault feature extraction. ISA Transactions, 2020, 97: 365-383.

[56] Goodfellow I, Bengio Y, Courville A. Deep Learning. Cambridge, MA: The MIT Press, 2016.

[57] Shao H, Jiang H, Lin Y, et al. A novel method for intelligent fault diagnosis of rolling bearings using ensemble deep auto-encoders. Mechanical Systems and Signal Processing, 2018, 102: 278-297.

[58] Luo Y, Cheng Y, Uzuner A, et al. Segment convolutional neural networks (Seg-CNNs) for classifying relations in clinical notes. J Am Med Inform Assoc. 2017, 25(1): 93.

[59] Kuo C C J. Understanding Convolutional Neural Networks with A Mathematical Model. Journal of Visual Communication & Image Representation, 2016, 41: S104732031 6302267

Associates Inc, 2017, 60(6), 84-90.

[61] Cerrada M, Sanchez R V, Li C, et al. A review on data-driven fault severity assessment in rolling bearings. Mechanical Systems and Signal Processing, 2018, 99(15): 169-196.

[62] Rai A, Upadhyay S H. A review on signal processing techniques utilized in the fault diagnosis of rolling element bearings. Tribology International, 2016: 289-306.

[63] Zhang X, Wan S, He Y, et al. Bearing Fault Diagnosis Based on Iterative 1.5-Dimensional Spectral Kurtosis[J]. IEEE Access, 2020, 8: 174233-174243.

[64] Yan X, Liu Y, Jia M. A Feature Selection Framework-Based Multiscale Morphological Analysis Algorithm for Fault Diagnosis of Rolling Element Bearing. IEEE Access, 2019, 7: 123436-123452.

[65] Kensert A, Harrison P J, Spjuth O. Transfer Learning with Deep Convolutional Neural Networks for Classifying Cellular Morphological Changes. Slas Discovery, 2019, 24(4): 466-475.

[66] Wang B, Lei Y G, Li N P, et al. A hybrid prognostics approach for estimating remaining useful life of rolling element bearings. IEEE Transactions on Reliability, 2018: 1-12.

[67] Wang B, Lei Y G, Li N P, et al. A hybrid prognostics approach for estimating remaining useful life of rolling element bearings. IEEE Transactions on Reliability, 2018: 1-12.

[68] Liu X, Zhang X, Luan Z, et al. Rolling bearing fault diagnosis based on EEMD sample entropy and PNN. The Journal of Engineering, 2019(23): 8696-8700.

[69] Malhotra A, Minhas A S, Singh S, et al. Bearing fault diagnosis based on flexible analytical wavelet transform and fuzzy entropy approach. Materials Today: Proceedings, 2021, 43(1): 629-635.