KSME

Springer

Journal of Mechanical Science and Technology 34 (6) 2020

Original Article
DOI 10.1007/s12206-020-0506-8

Keywords:
· Aero-engine
· Rolling bearing
· Casing
· Convolutional neural network
· Wavelet scale spectrum
· Fault diagnosis

Correspondence to:
Guo Chen
cgzyx@263.net

Citation:
Zhang, X., Chen, G., Hao, T., He, Z. (2020). Rolling bearing fault convolutional neural network diagnosis method based on casing signal. Journal of Mechanical Science and Technology 34 (6) (2020) ?~?.
http://doi.org/10.1007/s12206-020-0506-8

Received June 17th, 2019
Revised March 25th, 2020
Accepted April 1st, 2020

† Recommended by Editor Chongdu Cho

Rolling bearing fault convolutional neural network diagnosis method based on casing signal

Xiangyang Zhang[1], Guo Chen[1], Tengfei Hao[2] and Zhiyuan He[1]

[1]College of Civil Aviation, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China, [2]School of Automotive and Rail Transit, Nanjing Institute of Technology, School of Automotive and Rail Transit, Nanjing 211167, China

Abstract Affected by the transmission path, it is very difficult to diagnose the vibration signal of the rolling bearing on the aircraft engine casing. A fault diagnosis method based on convolutional neural network is proposed for the weak vibration signal of the casing under the excitation of rolling bearing fault. Firstly, the processing method of vibration signal is studied. Through comparison and analysis, it is found that the fault characteristics of rolling bearing are more easily expressed by continuous wavelet scale spectrum, and a better recognition rate is obtained. Finally, the experiment was carried out with an aero-engine rotor tester with a casing, and the method based on wavelet scale spectrum and convolutional neural network was used for diagnosis. The results were compared with the support vector machine method. The results show that the method has a high recognition rate for the weak fault signals of different fault types collected on the aero engine case, and its fault recognition rate reaches 95.82 %, which verifies the superiority and potential of the method for rolling bearing fault diagnosis.

1. Introduction

As a key component in aero-engines, rolling bearing is of a high failure rate due to their high temperature, high speed and large load change range. Once a fault occurs, it will cause abrasion of the rotor, transmission failure, and even cause the engine to stop in the serious case. Therefore, the condition monitoring and fault diagnosis of aircraft engine rolling bearing is of great significance [1, 2].

Machine learning is an effective method for fault diagnosis of rolling bearings. Chen et al. [3] extracted the characteristics of time domain and frequency domain, reduced the feature by principal component analysis (PCA), and then used Gaussian mixture model to diagnose bearing faults. Zhang et al. [4] proposed to use self-organizing neural network for rolling bearing state. Chen [5] used BP neural network for early fault intelligent diagnosis of rolling bearings. Saidi et al. [6] inputed the artificially acquired high-order spectral fault features into the SVM for bearing fault identification.

In recent years, deep learning has achieved great success in the fields of natural language processing, computer vision, and image recognition [7, 8]. Because of its powerful feature learning and representation ability, and the ability to adaptively learn features, it is widely used in the field of fault diagnosis. For example, Chen et al. [9] used the deep neural network to evaluate the damage degree of rolling bearings. Lei et al. [10] used deep learning to achieve health monitoring of mechanical equipment. Shao et al. [11] proposed a rolling bearing fault feature learning method based on compressed sensing and an improved convolutional deep confidence network, compared with other standard deep learning methods and manually extracted features, it achieved a higher recognition rate. Zeng et al. [12] obtained a time-frequency diagram of the signal into S transform, and then used the convolutional neural network to identify the fault of the gearbox. Li et al. [13] processed the signal by short-time Fourier

transform to obtain the time-frequency diagram, and then used the convolutional neural network to diagnose the bearing fault. Janssens et al. [14] studied the application of convolutional neural networks in rolling bearing fault diagnosis, the frequency spectrum of the vibration signal is used as the input of the CNN, the CNN is used to directly learn fault features and use them for intelligent diagnosis, compared with the diagnosis of the features, higher fault diagnosis accuracy is achieved, indicating that CNN can be effectively used for fault feature learning. Wu et al. [15] used a one-dimensional convolutional neural network to realize the fault diagnosis of the gearbox. Zhang et al. [16] directly used the original vibration signal of the rolling bearing as the input of the CNN to realize the fault diagnosis of the rolling bearing. Jia et al. [17] used the weighted Softmax loss function to study the imbalance between normal and fault samples, and proposed a neuron activation maximization algorithm to understand the CNN fault feature learning process.

For the fault diagnosis of the rolling bearing vibration signal collected on the aircraft engine casing, due to the complicated transmission path, the structural aerodynamic noise, combustion noise and vibration noise are usually coupled together, thus that the fault characteristics of the casing signal are very weak. In the case of bearing fault diagnosis, it is difficult to obtain the desired effect.

Due to the huge advantage of convolutional neural network in image recognition, this paper proposes a fault diagnosis method based on convolutional neural network [18, 19]. Using the characteristic learning advantage of convolutional neural network, the fault diagnosis of rolling bearing based on casing signal is realized by using the characteristic learning advantage of convolutional neural network. At the end of the paper, the method is verified, analyzed and compared by experiments.

2. Convolutional neural network

Convolution neural network (CNN) [20] is a feedforward neural network, which is mainly composed of a convolutional layer, a pooling layer, and a fully connected layer. These layers are used to complete the task of feature learning and classification.

2.1 Convolutional layer

The convolution layer uses multiple convolution kernels and the input image to perform convolution. After adding the bias term, the activation function can obtain a series of feature maps. The mathematical expression of the convolution is

$$y = \beta_0 + \sum_{i=1}^{n_d} \beta_i x_i + \sum_{i=1}^{n_d} \sum_{j \geq i}^{n_d} \beta_{ij} x_i x_j \tag{1}$$

where X_J^l is the lth element of the j -th layer; M_j is the $l-1$ -th convolutional region of the $y = \beta_0 + \sum_{i=1}^{n_d} \beta_i x_i + \sum_{i=1}^{n_d} \sum_{j \geq i}^{n_d} \beta_{ij} x_i x_j$ -th layer feature map; X_i^{l-1} is the element therein; w_{ij}^l is the corresponding weight matrix; b_j^i is the bias term. $y = \beta_0 + \sum_{i=1}^{n_d} \beta_i x_i + \sum_{i=1}^{n_d} \sum_{j \geq i}^{n_d} \beta_{ij} x_i x_j$ is the activation function whose mathematical expression is:

$$f(x) = \max(0, \lg(1 + e^x)) . \tag{2}$$

2.2 Pooling layer

The pooling layer is usually behind the convolution layer. Its main function is to reduce the dimension of the feature map and ensure the translation invariance of its features. Pooling methods used commonly include: Max pooling, mean pooling, stochastic pooling, etc.

Generally, only the dimensionality reduction is performed on the pooling layer, and there is no need to update the weights. It performs a pooling operation on the feature map output by the convolution layer at each non-overlapping size region. Both dimensions have been reduced.

2.3 Full connection layer

After the input image has been propagated through the convolutional layer and the pooling layer for several times, it is classified using the fully connected layer. The input of the fully connected layer is the one-dimensional feature vector expanded by all feature maps, which can be obtained after weighted summing and activation function:

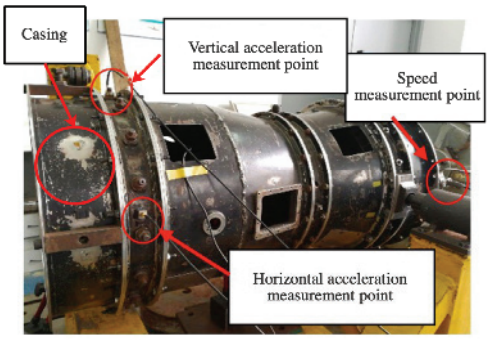
$$y^k = f\left(w^k x^{k-1} + b^k\right) \tag{3}$$

where $y = \beta_0 + \sum_{i=1}^{n_d} \beta_i x_i + \sum_{i=1}^{n_d} \sum_{j \geq i}^{n_d} \beta_{ij} x_i x_j$ is the serial number of the network layer; $y = \beta_0 + \sum_{i=1}^{n_d} \beta_i x_i + \sum_{i=1}^{n_d} \sum_{j \geq i}^{n_d} \beta_{ij} x_i x_j$ is the output of the fully connected layer; $y = \beta_0 + \sum_{i=1}^{n_d} \beta_i x_i + \sum_{i=1}^{n_d} \sum_{j \geq i}^{n_d} \beta_{ij} x_i x_j$ is the one-dimensional feature vector; $y = \beta_0 + \sum_{i=1}^{n_d} \beta_i x_i + \sum_{i=1}^{n_d} \sum_{j \geq i}^{n_d} \beta_{ij} x_i x_j$ is the weight coefficient; $y = \beta_0 + \sum_{i=1}^{n_d} \beta_i x_i + \sum_{i=1}^{n_d} \sum_{j \geq i}^{n_d} \beta_{ij} x_i x_j$ is the bias term. Multi-classification tasks generally use the Softmax activation function.

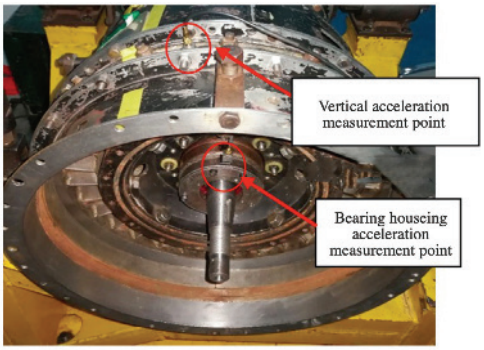
3. Simulation experiment of rolling bearing failure of aero engine rotor experimenter with casing

3.1 Aero-engine rotor experimental

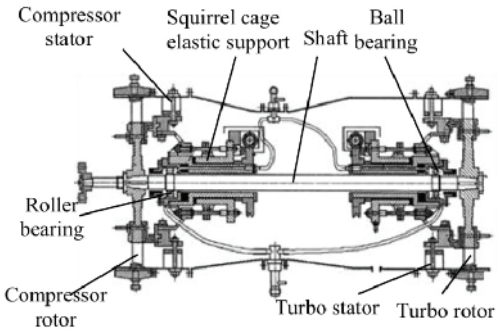
The experimental platform is the aero-engine rotor tester with



(a) External



(b) Innel



(c) Section drawing

Fig. 1. Aero-engine rotor experimental.

casing, which is designed and manufactured in a 1:3 size based on a real aero-engine, as shown in Fig. 1. Firstly, the casing of the tester is consistent with the shape of the core engine of the aero engine. Secondly, its internal structure is simplified. Finally, the structure of the multi-stage compressor is simplified into a single-stage roulette. The structure of the tester can be expressed as "rotor-abut-van-casing system", which has the structure of a real aeroengine and can effectively simulate the attenuation process of rolling bearing vibration signal through the transmission path. Chen et al. [21] have made relevant verifications and discussions on the above.

3.2 Rolling bearing fault simulation experiment

In this experiment, a 6206 single-row deep groove ball bearing was used, and a 6 mm wide crack was processed on the inner and outer ring surfaces by electric discharge wire cutting. and a cylindrical pit with a radius of 0.5 mm and a depth of 2 mm was machined on the surface of the rolling element, as shown in Fig. 2. The bearing parameters are shown in Table 1.

Table 1. Ball bearing dimensions.

Model	Thickness	Rolling body diameter/mm	Section diameter/mm	Number of balls
6206	16	9.5	46	9

Table 2. Rolling bearing fault simulation experiment scheme.

Serial number	Rotating speed (r/min)	Bearing status	Measuring point position
1	1500	N IF OF BF	Vertical measuring point; horizontal measuring point; housing position
2	1800	N IF OF BF	Vertical measuring point; horizontal measuring point; housing position
3	2000	N IF OF BF	Vertical measuring point; horizontal measuring point; housing position

(a) Outer ring fault (b) Inner ring fault (c) Rolling element failure

Fig. 2. Bearing 6206 after fault processing.

The rolling bearing failure simulation experiment is to install normal (N), inner ring fault (IF), outer ring fault (OF), and ball fault bearings (BF) into a rotor experimenter with a casing. In the experiment, two vibration acceleration sensors are arranged vertically above the casing and horizontally in the casing. As shown in Fig. 1, a vibration acceleration sensor (B&K 4805) and a data collector (NI USB9234) are used to collect vibration acceleration signals with a sampling frequency of 10240 Hz. Each data sample contains 8192 points. This experiment was performed at three different speeds, and the experimental scheme is shown in Table 2.

3.3 Weakness of rolling bearing fault characteristics in casing signals

The artificial fault bearing tests were performed at three dif-

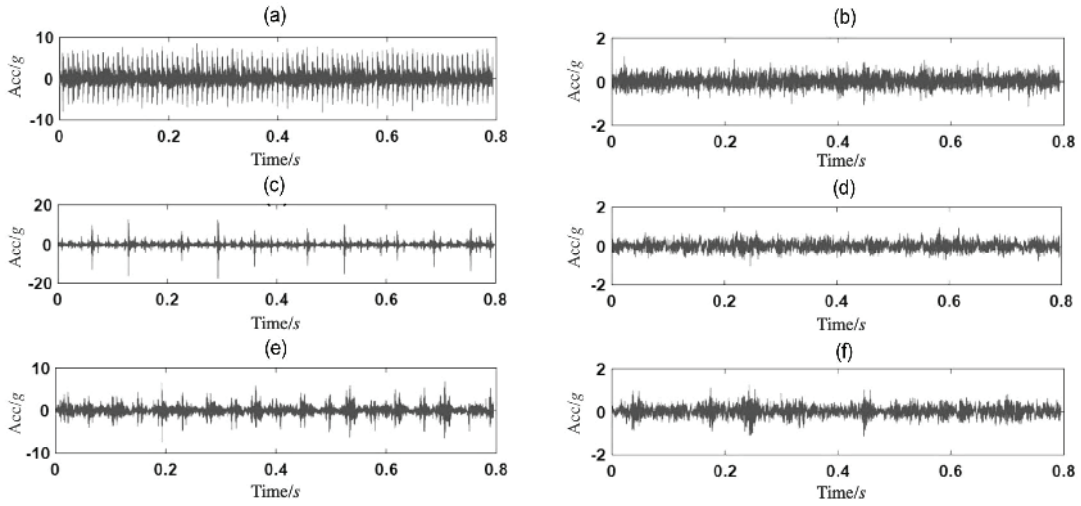


Fig. 3. Time-domain waveforms of different faults: (a) Outer race fault from bearing housing; (b) inner race fault from bearing housing; (c) rolling element fault from bearing housing; (d) outer race fault from casing; (e) inner race fault from casing; (f) rolling element fault from casing.

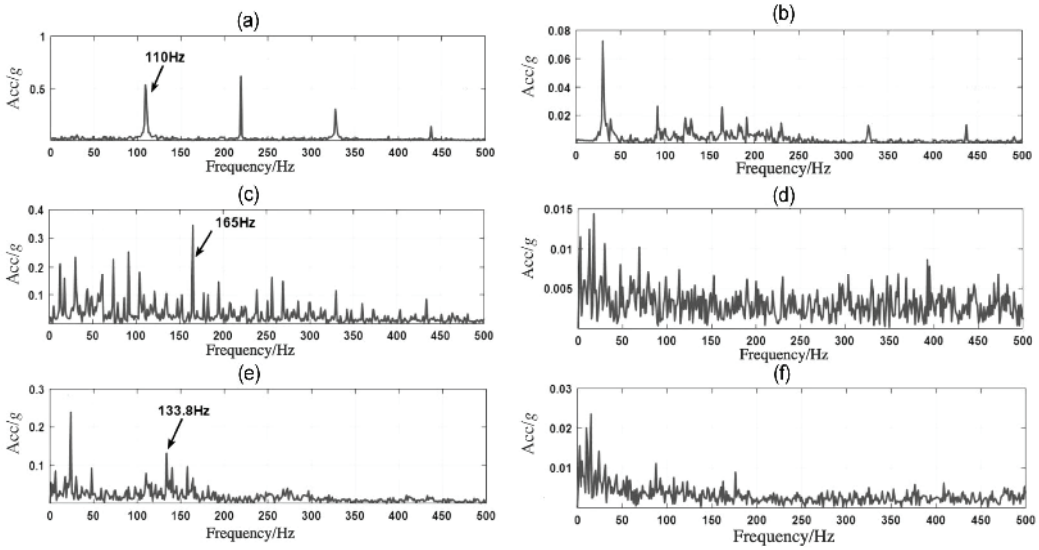


Fig. 4. Envelope spectrum of vibration signals: (a) Outer race fault from bearing housing; (b) inner race fault from bearing housing; (c) rolling element fault from bearing housing; (d) outer race fault from casing; (e) inner race fault from casing; (f) rolling element fault from casing.

ferent speeds, and three types of fault time domain waveforms of bearing fault vibration signals at 1800 rpm were taken as examples, as shown in Fig. 3. Figs. 3(a)-(c) are the time domain waveforms of the outer ring fault, inner ring fault and ball fault collected on the bearing house, respectively; Figs. 3(d)-(f) are collected on the horizontal measurement points of the casing time domain waveform diagram of outer ring fault, inner ring fault and ball fault. It can be known from Fig. 3 that the fault impact of the bearing point measurement signal is obvious, and the impact amplitude is large. However, the shock characteristics in the box's measuring point signal are masked by a lot of noise, and the amplitude of the signal transmitted to the box becomes very weak.

The envelope spectrum is shown in Fig. 4 obtained by using a Hilbert transform to analyze the envelope spectrum of the signal in Fig. 3. It can be seen from Fig. 4(a) that the passing frequency of the rolling element on the outer ring raceway is 110 Hz, and its harmonics are particularly obvious, but in its corresponding casing signal, as shown in Fig. 4(b), only some weak harmonics appear weight. Similarly, in Fig. 4(c), the passing frequency of the rolling element on the inner ring raceway shows 165 Hz, but the passing frequency of the rolling element on the inner ring raceway is not found in its casing signal, and only some noise frequency. The same is true in the rolling body failure analysis. It can be known that due to the influence of the transmission path, the fault characteristics in the receiver signal are seriously attenuated, so it is very difficult to extract the rolling bearing fault characteristics from the receiver signal.

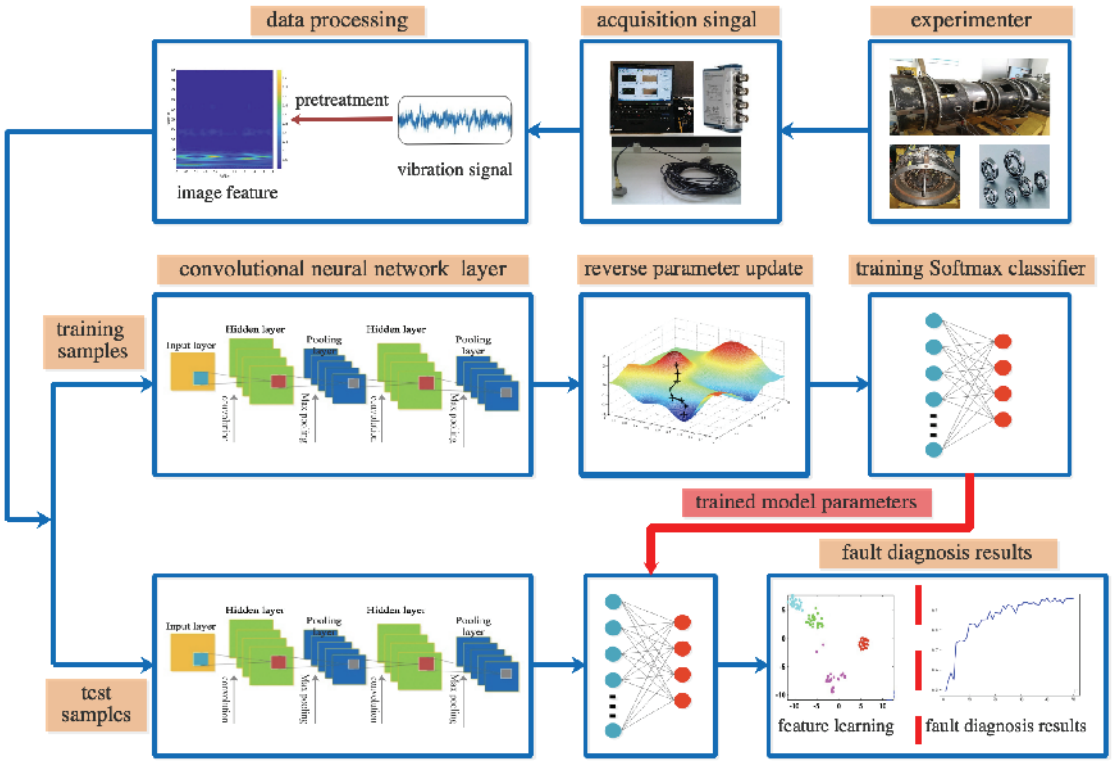


Fig. 5. CNN fault diagnosis method flow chart.

4. Rolling bearing fault diagnosis method based on casing signal

4.1 Intelligent fault diagnosis process

This paper proposes a convolutional neural network method that can adaptively extract fault features and implement intelligent diagnosis. The specific flowchart is shown in Fig. 5, which is summarized as follows:

Step 1: Vibration signals of rolling bearings are collected by sensor measurement and data acquisition system;

Step 2: Processing the vibration signal to obtain a two-dimensional image signal;

Step 3: Divide the data set into a training sample and a test sample;

Step 4: Establish a CNN network model;

Step 5: The gradient descent method is used to update the parameters in reverse, and then used to train the unsupervised feature learning of the sample. Extracting the depth features for rolling bearing fault diagnosis;

Step 6: Verify the test sample using the proposed method and output the diagnosis result.

The algorithm used in this paper is shown in Fig. 6.

Algorithm :Pesudo code of CNN

Input: Image-labeling pairs$\{(x^{(i)}, y^{(i)})\}$**, in the data set.**

Output: Classification results of CNN

Parameters:

Alpha: the rate of learning for highway networks

Batchsize: the size of samples for highway networks

Numepochs:the number of TSF in training algorithm

Kernelsize:the size of convolutional kernel in architechture

1. read data, and standardize the data.
2. set cnn layers with input layer, max sampling layer and fully connected layer.
3. employ train samples and start to train

   CNN= cnntrain(cnn, train_x, train_y, opts);
4. calculate the loss, and use BP algorithm to update parameters in reverse.
5. employ test samples and start to test

   (lab, acc)=cnntest (cnn, test_x, test_y);

Tips:

Opts.alpha =1

Opts.batchsize=128

Opts.numepochs=50

Fig. 6. Pseudocode for CNN.

4.2 Convolutional neural network structure

The network model structure of CNN mentioned in this paper is shown in Fig. 7. The CNN is composed of a convoluted layer, a pooled layer, and a fully connected layer. C, P, and F to represent sepeartely the convolutional layer, the pooled layer, and the fully connected layer. The size of the convolution kernel is N*D*H, where N represents the number of convolution kernels, D represents the depth of the convolution kernel, and H represents the height of the convolution kernel. When inputting the wavelet scale spectrum, CNN first uses the convolution layer C1 to adaptively learn the features, then reduces the dimensions of the convolution layer through the pooling layer, repeats the above process, and finally features in the fully connected

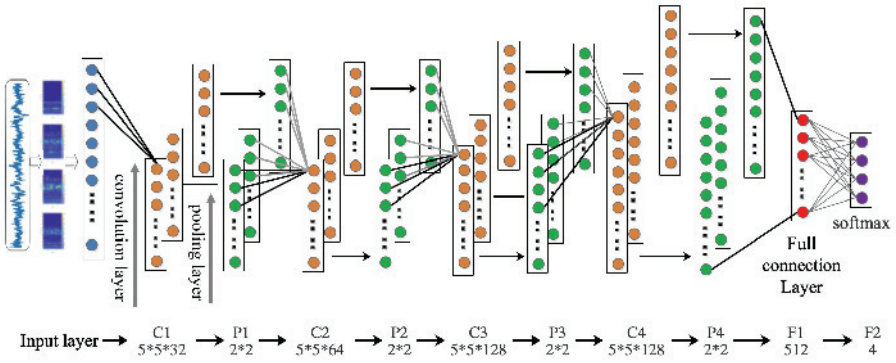


Fig. 7. Convolutional neural network model.

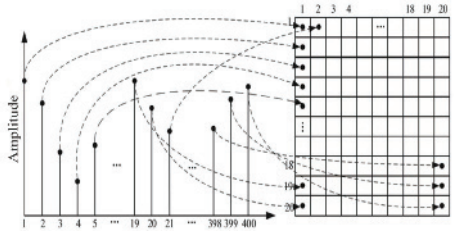
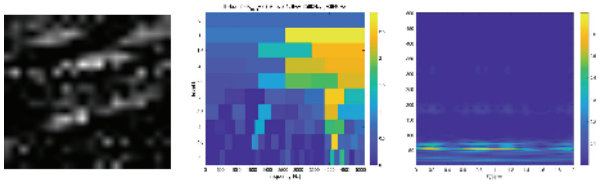


Fig. 8. Method of building a matrix diagram.



(a) Matrix diagram (b) Kurtosis map (c) Wavelet scale spectrum

Fig. 9. Data sample.

layer F9 are flatten to a one-dimensional vector and input into the softmax classifier to identify the fault category of the bearing. The detailed parameters of the CNN model are shown in Fig. 5.

4.3 Data processing methods

In this paper, three different data processing methods are used to process the collected vibration signals. The specific methods are: Matrix diagram method (MDM), kurtosis diagram method (KDM) [22], wavelet scale spectrum (WSS) [23].

4.3.1 Matrix diagram method

First, the measurement data with standard time scale and amplitude scale is obtained; then the data set of each state is divided into a series of time subsequences, and these time subsequences are sorted by successive interleaving sampling to generate data matrices; finally, the matrix graph is established by using the method shown in Fig. 8, and the matrix graph sample is shown in Fig. 9(a).

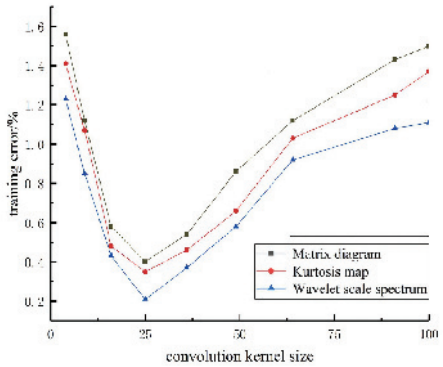


Fig. 10. Error rate trends for different convolution kernel sizes.

4.3.2 Fast spectral kurtosis method

The fast spectrum is decomposed by the band alternating three-decomposition method, and the spectrum kurtosis is expressed in the plane region to obtain the spectral kurtosis value under the combination of the optimal frequency and frequency resolution, the signal instability can be detected and characterized, as shown in Fig. 9(b).

4.3.3 Wavelet scale spectrum

A wavelet scale spectrum is a two-dimensional depiction of a time and frequency of a signal. The principle is to give a corresponding frequency description on each time scale on the time axis, which constructs a two-dimensional time-frequency distribution diagram with the vertical axis as the frequency and the horizontal axis as the time, as shown in Fig. 9(c).

4.4 Analysis of the influence of convolution kernel size on test results

In the training process, the size of the convolution kernel is very important for achieving high performance; therefore, the data sets obtained by the three data processing methods are respectively used for experiments to observe the variation trend of the classification error based on the convolution kernels of different sizes, as shown in the Fig. 10. It can be seen from Fig. 10 that the three methods have the lowest training error rate when the convolution kernel size is 5*5, and the sub-

Table 3. Data sample introduction.

Sample type	Number of samples			
	N	OF	IF	BF
MDM	2000	2000	2000	2000
KDM	2000	2000	2000	2000
WSS	2000	2000	2000	2000

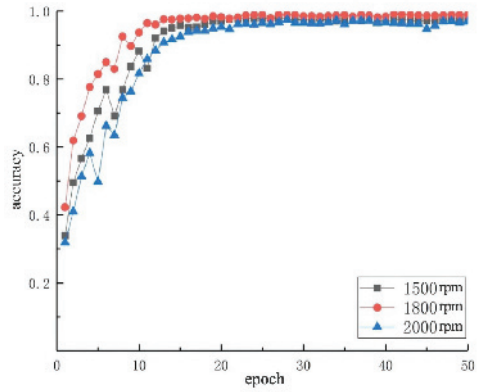


Fig. 11. Fault classification results at different speeds.

sequent error rate gradually increases; when the convolution kernel continues to grow, it means that the super-large convolution kernel may bring unnecessary negative effects in the training process.

4.5 Analysis of test results

4.5.1 Fault classification results at different speeds

The data samples used in this paper are shown in the table. Each data sample has 8000 pictures, and each fault type has 2000 pictures.

First, the different fault signals at different speeds are regarded as one kind of fault data. Here, the signals collected at the vertical measurement points of the case are used as data samples. Three preprocessing methods are used to process the vibration signals. The obtained data set is based on the ratio of 2:8, which is randomly divided into a training set and a test set, and the sample set is used as the input of the CNN. The results of fault classification at different speeds are shown in the Fig. 11. As can be seen from the Table 3, the fault classification results of the proposed method at different speeds have reached more than 98 %.

4.5.2 Classification results of bearing house signals

In order to further verify the effectiveness of this method, the signal samples collected by the same kind of fault bearing at different speeds are taken as a kind of fault data, and then the vibration signal is processed by three different data preprocessing methods. The obtained picture data set is in the ratio of 2:8. It is randomly divided into training set and test set, and the

Table 4. Bearing house signal classification result.

Measuring point	Accuracy/%			
	CNN+MDM	CNN+KDM	CNN+WSS	SVM
Bearing housing	99.43	99.75	100	99.26

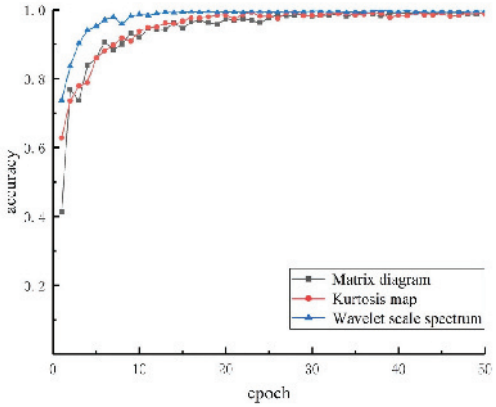
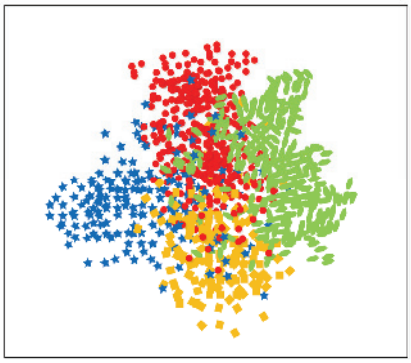
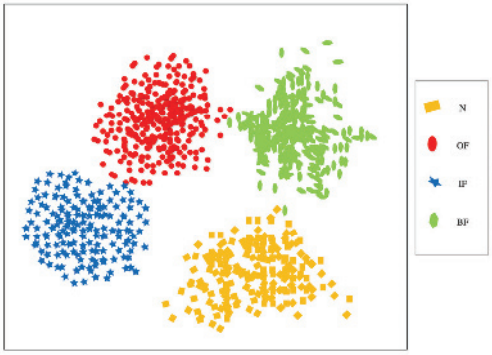


Fig. 12. Fault classification result of bearing house signal.



(a) Feature distribution map extracted from the first convolution layer



(b) Feature classification result graph

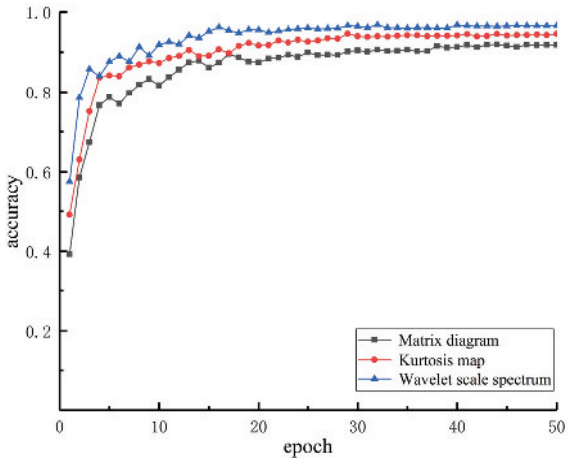
Fig. 13. Classification of bearing house signal characteristics.

sample set is used as the input of CNN. The CNN network model used in this paper is shown in Fig. 6 and finally the fault classification results of different data processing methods combined with CNN are shown in Fig. 12 and Table 4.

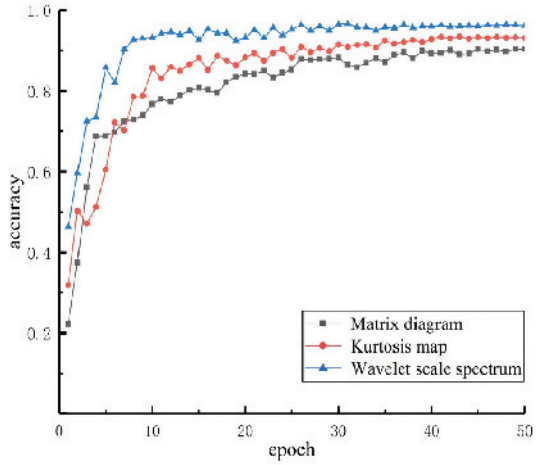
In addition, a total of 13 time-domain and frequency domain

Table 5. Fault classification result of casing signal.

Measuring point	Accuracy/ %			
	CNN+MDM	CNN+KDM	CNN+WSS	SVM
Vertical measuring point	92.44	93.58	96.32	86.16
Horizontal measuring point	90.03	92.17	95.82	88.73



(a) Vertical measuring point



(b) Horizontal measuring point

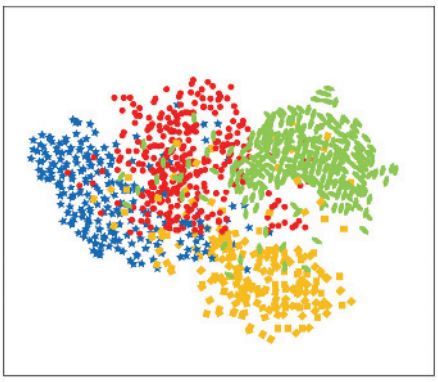
Fig. 14. Fault classification result of casing signal.

features are extracted for the vibration signal, including:

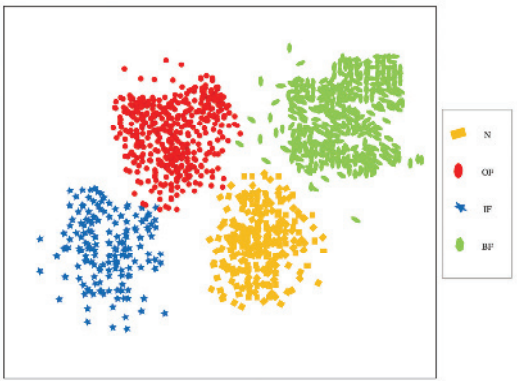
(1) Time-domain dimensionless features: Absolute average amplitude, square root amplitude, effective value, peak value, form factor, peak index, impact index, twist, kurtosis, margin index;

(2) Frequency domain dimensionless features: Center of gravity frequency, mean square frequency, frequency variance.

The extracted 13 time-frequency features are input into the parameter-optimized support vector machine (SVM) model, and the fault classification results are obtained. As shown in Table 4, the fault classification accuracy of CNN and SVM on the bearing seat measurement points. Both reached more than



(a) Feature distribution map extracted from the first convolution layer



(b) Feature classification result graph

Fig. 15. Classification of bearing casing signal characteristics.

99 %. It can be seen from Fig. 12 that with the three methods of fault classification using CNN, the classification accuracy rate increases with the increase of the number of iteration steps, but the convergence speed of the proposed method is faster.

From Fig. 13, the feature distribution extracted by the first layer of the convolutional neural network is relatively chaotic, as shown in Fig. 13(a). After multiple convolutional layers and fully connected layers, the final classifier can extract the convolutional neural network well. The features are effectively classified.

4.5.3 Fault classification results on machine points

As shown in Table 5, the fault classification accuracy of SVM on the horizontal and vertical measuring points is 86.16 % and 88.73 %, respectively, and the fault classification accuracy based on the convolutional neural network algorithm is significantly higher than that of the support vector machine. As a result, the lowest fault classification accuracy rates of the horizontal and vertical points are: 92.44 % and 90.03 %, respectively.

It can be seen from Fig. 14 and Table 5 that the average accuracy of the proposed method at the horizontal and vertical measuring points reaches 96.32 % and 95.82 %, respectively, and is higher than the other two combinations. The results

show that the method described in this paper has better bearing fault diagnosis results. In particular, the experimental data is the bearing vibration signal collected on a casing with a real aero-engine structure. The fault characteristics of the signal are very weak due to the influence of the transmission path and other noise. The fault identification accuracy of the proposed method further proves that the proposed method has good generalization ability and fault recognition rate.

From Fig. 15, the method proposed in this paper can effectively classify the fault of the receiver signal. As the network layer of the convolutional neural network is closer to the front end, the features extracted are more general, and the features extracted from the neural network layer closer to the end are more specialized. After multiple layers of convolutional layers and pooling layers, more abstract deep features can be extracted for fault classification. It further proves the effectiveness of the method proposed in this chapter for the diagnosis of weak signals in the receiver.

5. Conclusion

The rolling bearing fault diagnosis method based on wavelet scale spectrum and convolutional neural network combines the depth feature and pattern recognition of automatic learning, effectively solves the problem of insufficient shallow feature representation of vibration signals, and avoids artificial features extraction to rely on experience and expertise. The main features of the method proposed in this paper are:

(1) Converting the vibration signal into a two-dimensional image; giving full play to the advantages of CNN in the field of image recognition;

(2) After research, it is found that using the time-frequency characteristics of the wavelet scale spectrum, the fault features can be adaptively learned through CNN, and the dependence of artificial extraction features on expert knowledge is avoided. The results show that the proposed method has better fault recognition rate and generalization ability.

Acknowledgments

This research is sponsored by the National Natural Science Foundation of China (No. 51675263), and National Science and Technology Major Project (2017-IV-0008-0045).

References

[1] H. B. Mei, *Vibration Monitoring and Diagnosis of Rolling Bearings*, Mechanical Industry Press (1995).

[2] R. B. Randall and J. Antoni, Rolling element bearing diagnostics - A tutorial, *Mechanical Systems and Signal Processing*, 25 (2) (2011) 485-520.

[3] Z. Q. Chen, C. Li and R. V. Sanchez, Gearbox fault identification and classification with convolutional neural networks, *Shock and Vibration*, 2 (2015) 1-10.

[4] Q. D. Zhang et al., State evaluation method of rolling bearing based on self-organizing neural network, *China Mechanical Engineering*, 5 (2017).

[5] G. Chen, Feature extraction and intelligent diagnosis of early faults in rolling bearings, *Acta Aeronautica Sinica*, 30 (2) (2009) 362-367.

[6] L. Saidi, A. J. Ben and F. Fnaiech, Application of higher order spectral features and support vector machines for bearing faults classification, *ISA Transactions*, 54 (2015) 193-206.

[7] R. Socher, B. Huval, B. P. Bath, C. D. Manning and A. Y. Ng, Convolutional-recursive deep learning for 3D object classification, *NIPS*, 3 (7) (2012) 8.

[8] Y. LeCun, Y. Bengio and G. Hinton, Deep learning, *Nature*, 521 (7553) (2015) 436-444.

[9] R. X. Chen et al., Diagnosis of rolling bearing damage degree by stack sparse noise-added self-coded deep neural network, *Journal of Vibration and Shock*, 36 (21) (2017) 132-138, 144.

[10] Y. G. Lei et al., The big data health monitoring method for mechanical equipment based on deep learning theory, *Journal of Mechanical Engineering*, 51 (21) (2015) 49-56.

[11] H. D. Shao et al., Rolling bearing fault feature learning using improved convolutional deep belief network with compressed sensing, *Mechanical Systems and Signal Processing*, 100 (2018) 743-765.

[12] X. Zeng, Y. Liao and W. Li, Gearbox fault classification using S-transform and convolutional neural network, *International Conference on Sensing Technology* (2016).

[13] H. Li et al., Fault diagnosis method for rolling bearings based on short-time Fourier transform and convolution neural network, *Journal of Vibration and Shock*, 37 (19) (2018) 124-131.

[14] O. Janssens et al., Convolutional neural network based fault detection for rotating machinery, *Journal of Sound and Vibration*, 377 (2016) 331-345.

[15] C. Z. Wu et al., Gearbox fault diagnosis based on one-dimensional convolutional neural network, *Journal of Vibration and Shock*, 37 (22) (2018) 56-61.

[16] W. Zhang et al., A deep convolutional neural network with new training methods for bearing fault diagnosis under noisy environment and different working load, *Mechanical Systems and Signal Processing*, 100 (2018) 439-453.

[17] J. Feng et al., Deep normalized convolutional neural network for imbalanced fault classification of machinery and its understanding via visualization, *Mechanical Systems and Signal Processing*, 110 (2018) 349-367.

[18] X. Guo, L. Chen and C. Shen, Hierarchical adaptive deep convolution neural network and its application to bearing fault diagnosis, *Measurement*, 93 (2016) 490-502.

[19] A. Krizhevsky, I. Sutskever and G. Hinton, ImageNet classification with deep convolutional neural networks, *NIPS* (2012).

[20] L. C. Yang et al., Gradient-based learning applied to document recognition, *Proceedings of the IEEE*, 86 (11) (1998) 2278-2324.

[21] G. Chen et al., Sensitivity analysis and experimental research on ball bearing early fault diagnosis based on testing signal from casing, *Journal of Dynamic Systems, Measurement, and Control*, 136 (6) (2014) 061009-061019.

[22] J. Antoni, Fast computation of the kurtogram for the detection of transient faults, *Mechanical Systems and Signal Processing*, 21 (1) (2007) 108-124.
[23] S. J. Cheng et al., Application of scale-wavelet energy spectrum in fault diagnosis of rolling bearings, *Journal of Vibration Engineering*, 17 (1) (2004) 82-85.

X. Y. Zhang is a Master's student at the College of Civil Aviation, Nanjing University of Aeronautics and Astronautics, Nanjing, China. His current research interests include deep learning and pattern recognition, and their applications in bearing fault diagnosis.

G. Chen received a Ph.D. degree in the School of Mechanical Engineering from the Southwest Jiaotong University, Chengdu, P. R. China, in 2000. Now he works at the College of Civil Aviation, Nanjing University of Aeronautics and Astronautics, Nanjing, P. R. China. His current research interests include the whole aero-engine vibration, rotor-bearing dynamics, rotating-machine fault diagnosis, pattern recognition and machine learning, signal analysis and processing.