Fault diagnosis of satellite flywheel bearing based on convolutional neural network

Ying Liu
School of Transportation Science and Engineering,
Beihang University
Beijing, China
xiaoyingzi@buaa.edu.cn

Qiang Pan
School of Transportation Science and Engineering,
Beihang University
Beijing, China
panqiang@buaa.edu.cn

Hong Wang
Beijing Key Laboratory of Long-life Technology of
Precise Rotation and Transmission Mechanisms, Beijing
Institute of Control Engineering
Beijing, China

Tian He
School of Transportation Science and Engineering,
Beihang University
Beijing, China
hetian@buaa.edu.cn

Abstract—**The bearing is one of the core components of the flywheel, providing a stable slewing support for the flywheel, and its operating state often directly affects the flywheel and even the entire spacecraft's normal operation. In view of the problem of automatic and accurate identification of the bearing faults, this paper uses convolutional neural network (CNN) to develop a satellite flywheel bearing fault intelligent diagnosis method. First, the vibration signal characteristics of satellite flywheel bearing under different faults are studied. Second, the time-domain signal graphs are constructed by combining vibration signals under multiple rotational speeds and used as feature input maps. Finally, the bearing fault intelligent diagnosis method is presented based on the excellent image recognition characteristics of CNN and the constructed feature maps. The experimental verification shows that the proposed method can achieve better diagnostic results.**

Keywords- satellite flywheel bearing; CNN; fault diagnosis; Multi- information fusion

I. INTRODUCTION

The flywheel is a typical space inertial actuator for attitude control and accuracy maintenance of spacecraft. Among the various components and factors that determine the health and longevity of satellites, the flywheel is considered one of the most critical components [1]. The bearing assembly is one of the core components of the flywheel, providing a solid slewing support for the flywheel. In the actual experiment, there are many failure modes of rolling bearings [2]. It is very important to monitor and identify the types of these faults. The traditional fault diagnosis methods of bearings are mostly based on vibration characteristics, and the resonance demodulation method is used to eliminate the disturbance effect to extract the fault features, thereby performing fault diagnosis [3,4,5]. Common methods include Hilbert transform demodulation method, generalized detection filtering method, empirical mode decomposition method and wavelet transform method [6]. By using the above various demodulation methods to filter the interference signal to highlight the fault characteristics, the neural network, support vector machine, fuzzy clustering, expert system and other methods can be used for fault identification [7]. These methods have made significant advances and widespread applications in the diagnosis and prediction of bearings. However, there are some limitations. The extraction of fault features relies on a wealth of engineering experience. It is difficult to find fault characteristics for certain faults. The automated processing capability is poor [8]. Feature extraction and fault classification process are related to each other, so it is urgent to adapt the general algorithm with strong adaptability to design multiple feature extraction and classification methods for different signals [9].

With the improvement of sensor technology, computer technology, big data analysis technology and the popularity of intelligent algorithms, artificial intelligence-based diagnostic methods, especially deep learning methods, have unique advantages in bearing fault diagnosis and bright application prospects. It has been a current research hotspots.

Deep learning is a kind of neural network [10]. Compared with the traditional shallow artificial neural network, it has more network layers, is not easy to produce over-fitting, has better ability to approximate complex functions, and can realize complex high-dimensional functions [11]. The representation has good feature extraction and classification capabilities [12]. Deep learning provides new ideas and new methods for mechanical system fault diagnosis with its unique advantages. Among the various methods of deep learning, CNN has the characteristics of sparse connection, weight sharing, and subsampling, and has received more attention in fault diagnosis [13,14] . In recent years, Wei Zhang et al. used the time domain signal as input to use the deep convolutional neural network for fault diagnosis of rolling bearings under noisy environments and changing workloads, simplifying the manual extraction process of raw data [15]. Xia M et al. combine the multi-sensor data to form a two-dimensional

978-1-7281-0860-5/19/$31.00 ©2019 IEEE

matrix as input, which can achieve higher diagnostic accuracy than traditional rotary mechanical fault diagnosis methods [16]. Janssens et al. used the time-domain map after the mean value processing as the CNN input, which solved the fault diagnosis that is difficult to achieve by the traditional methods of bearing outer ring raceway failure and lubrication performance degradation [17]. Zeng Xueqiong and others used the vibration signal time-frequency image as input to realize the fault classification and identification of the automobile transmission [18]. In short, in terms of fault diagnosis, CNN has excellent performance in data integration, feature extraction and feature classification. In this paper, based on the difficulty and demand of space bearing fault diagnosis, CNN was introduced and an intelligent diagnosis method for satellite flywheel bearing fault was developed.

II. THEORETICAL BACKGROUND

A convolutional neural network is a multi-layered neural network consisting of an input layer, alternating convolutional and pooling layers, a fully connected layer, and an output layer. The input form of the input layer is an image. Both the convolutional layer and the pooling layer are composed of a plurality of two-dimensional planes, each of which is a feature map outputted by each layer after processing, and the number of convolutional layers and pooling layers is determined according to actual needs. The image information is calculated by convolution of each layer, which is equivalent to passing through a multi-layer digital filter, and finally obtains the most significant feature of the original input.

A. Convolutional layer

Assuming that the input layer input image pixel size is M×N, and the input layer two-dimensional image is represented as p = f(x, y), then f(x, y) represents the gray of the xth row and the yth column of pixels on the image p’s degree value. The convolution kernel size is a × b, then any point on the convolution kernel can be expressed as k(x, y). The weight of a point on the convolution kernel indicates the ability of the corresponding neighborhood point to contribute to the final result. C(s, t) is a matrix obtained by convolution of p and k. The operation expression is as follows:

$$C(s,t) = f(x,y) * k(x,y) = \sum_{x=1}^{a}\sum_{y=1}^{b} k(x,y) f(s+x-1, t+y-1) \quad (1)$$

Wherein, $1 \leq s \leq M - a + 1$, $1 \leq t \leq N - b + 1$.

In the convolutional layer of the convolutional neural network, the image of the previous layer is used as an input, convoluted with the convolution kernel, and then subjected to nonlinear transformation to obtain a feature map of this layer, and each convolution kernel corresponds to a feature map.

B. Pooling layer

After the features are obtained by the operation of the convolutional layer, if these features are directly classified, they will face a huge amount of calculation and are prone to over-fitting. Therefore, it is necessary to use the pooling layer to sample the features obtained by the previous convolution layer to reduce the dimension and reduce the computational complexity. The input feature image is divided into non-overlapping rectangular regions, and each rectangular region is subjected to a corresponding operation. Such an operation is called pooling.

The essence of pooling is the expression of local features. After the pooling area is scale × scale, the output feature edge length becomes the original 1/scale, but the number of feature maps is unchanged. Considering the loss of information, the pooling matrix should not be too large.

In addition to reducing the dimension, the role of the pooling layer also achieves quadratic feature extraction. In the convolutional neural network, the convolutional layer is followed by the structure of the pooling layer, which is equivalent to two feature extractions.

C. The fully connected layer

Each pixel of all the feature maps obtained after the pooling is sequentially expanded and arranged in a column to form a feature vector. The feature vector is fully connected to the output layer as a fully connected layer. The fully connected layer is located at the end of the convolutional neural network and is used to calculate the output of the entire network. When a convolutional neural network is used for classification tasks, a classifier is trained at the fully connected layer. The classifier typically chooses a classifier with a weight-of-weight that allows the network to train using gradient-based learning methods. The output layer outputs a real number vector whose number of nodes is consistent with the number of categories, and the output value of each node indicates the probability that the sample belongs to the corresponding category. The classifier chosen in this paper is Softmax function, a classifier extended by Logistic Regression Models that is suitable for dealing with multi-classification problems.

D. Convolutional neural network structure

The convolutional neural network can select various structural parameters of the network according to actual conditions. Fig.1 shows the structure of a typical convolutional neural network.

The network includes an input layer, two convolution layers, two pooling layers, one feature vector layer, and one output layer. The input layer is an image with a pixel size of p ×p, and is convolved with m convolution kernels of size k1×k1 respectively, with a step size of 1, and after the activation function, m× (p-k1+1) ×(p-k1+1) feature maps are obtained, which is the convolution layer C1. After the pooling area is c × c, the pooling layer S2 is obtained. At this time, the feature map is still m and the side length becomes (p-k1+1)/c. The S2 layer is convolved with n convolution kernels of size k2×k2, and then the activation function is used to obtain the convolutional layer C3. After pooling, the pooling layer S4 is obtained. The n feature maps of S4 are connected into a one-dimensional vector as the finally extracted feature vector layer V5. V5 is fully connected to the output layer to get the result [19].

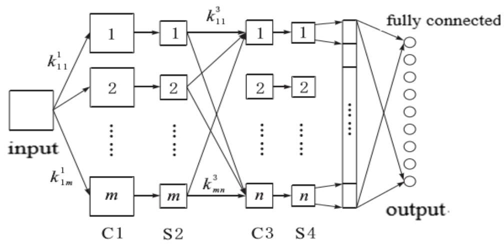


Fig. 1 Typical structure of convolutional neural network

III. PROPOSED INTELLIGENT DIAGNOSIS METHOD

The realization of the satellite flywheel bearing fault diagnosis method based on CNN can be divided into four steps. The fault diagnosis process is shown in Fig. 3 below:

i. Determine the structure of the CNN: set the number of layers in the CNN, the convolution kernel size and the classifier, wherein the convolution layer and the pooling layer extract fault features, and the classifier is used for classification identification of fault types;

ii. Bearing data processing: processing the bearing data, selecting samples, transforming the vibration signal into a time domain image as a kind of sample, and then obtaining a new training sample integrating multi-speed information through image stitching.

iii. Training CNN: training CNN with two training samples in ii, assigning the trained weight parameter matrix and offset to each layer of the network, and then it has classification and feature extraction functions;

iv. Testing CNN: After obtaining the CNN and the classifier after training, test the CNN with the test sample.

The convolutional neural network used herein includes an input layer, two convolutional layers, two pooling layers, and one output layer. The picture size of the input layer is 40×40, and the convolution operation is performed separately with the three 3×3 convolution kernels to obtain three 38×38 feature maps, which form the convolution layer C1, and then pass the size 2×2 pooling operation to obtain three 19×19 feature maps to form the pooling layer S2. S2 layer and three 2×2 convolution kernels are convoluted separately to obtain three 18×18 convolutional layers C3, and then pooled by size 2×2 to obtain three 9×9 pools to form layer S4, so the number of parameters of the finally obtained S4 layer is 243. The three feature maps of S4 are converted into one-dimensional vectors as the extracted feature vectors, and then fully connected with the output layer to obtain an output result. The network structure diagram and the number of parameters are shown in Fig. 2.

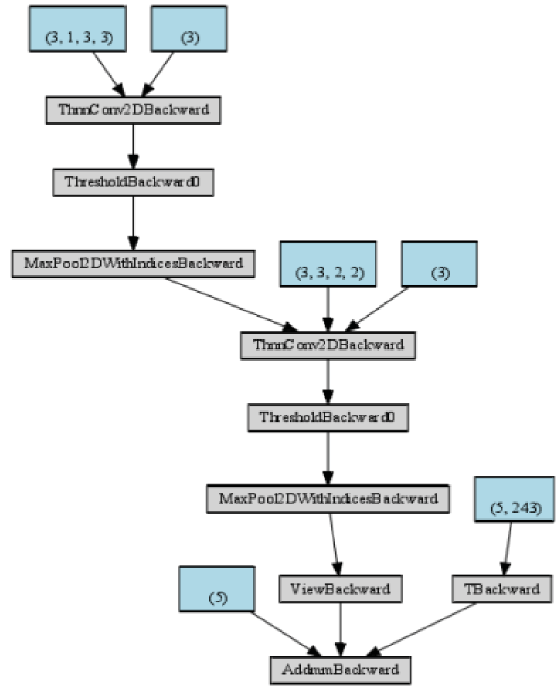


Fig. 2 Network structure diagram and number of parameters

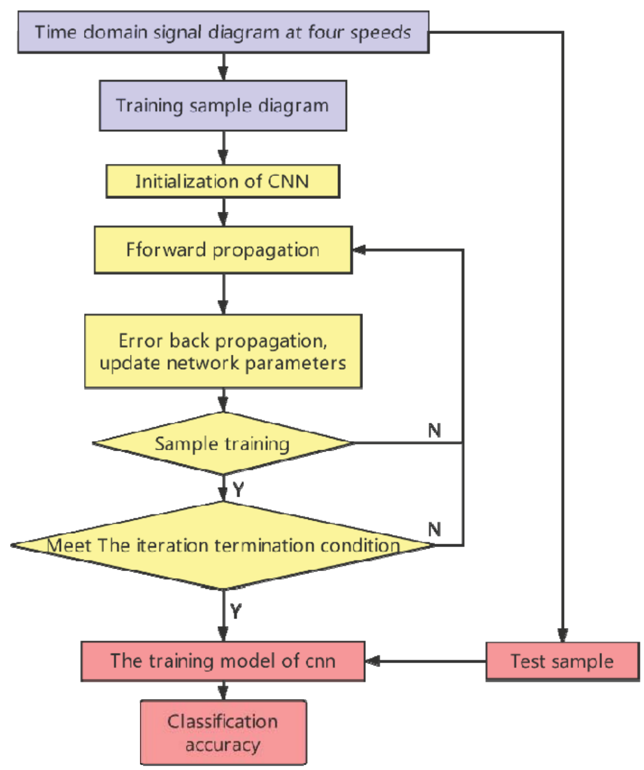


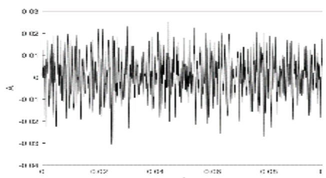
Fig. 3 Fault diagnosis flow chart

IV. EXPERIMENTAL VERIFICATION

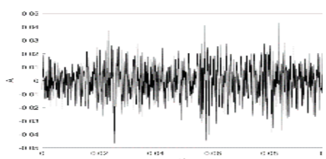
This paper uses the laboratory console to conduct experiments and collect data of different working conditions.

Due to the complexity of multi-fault diagnostic systems, applying machine learning to fault diagnosis is an organizational and learning issue for integrated networks. In order to minimize the complexity of the network, reduce the network organization and training time, use a single point radial vertical vibration signal to draw the time domain waveform map as input for further analysis.

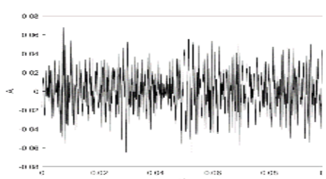
As can be seen from the figure below, the original vibration time domain signals are more cluttered under different working conditions, and there are many background noises that are difficult to distinguish. The original fault diagnosis often requires pre-processing of the signal. However, this paper takes the original time domain map as input, retains all fault information, and uses the powerful nonlinear mapping capability of the convolutional neural network for feature extraction.



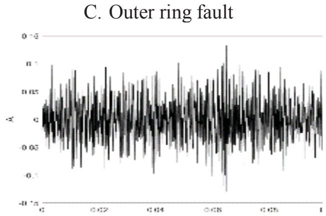
A. normal



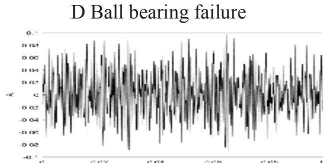
B. The structure is loose



C. Outer ring fault



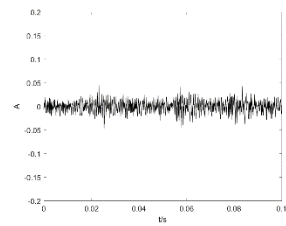
D Ball bearing failure



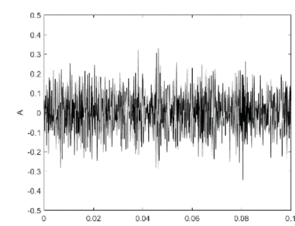
E. Cage failure

Fig. 4 Time domain diagram and working condition code of vibration signals collected at a speed of 300r/min

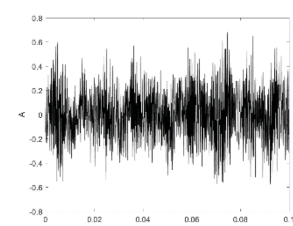
Training a good convolutional neural network requires a large amount of sample data input. The data needs to be initially normalized before training the network. According to the working principle of the convolutional neural network, the vibration signals collected at different speeds in the same position are first divided into multiple pieces of data and drawn into multiple time domain maps. The resulting time domain map is then converted to a grayscale image and the pixels are reduced as much as possible while retaining their features. Finally, the multi-speed time domain map is directly merged, and the image pixel compression is performed again after the combination. The adjusted composite map constitutes a large number of input samples. A part of each type of data is randomly selected to form a training sample, and the rest constitutes a test sample. The ratio of the training sample to the test sample is 9:1. The training samples are input into the initialized convolutional neural network and the operations are performed until the iterative operation is completed. The test samples are input into the trained convolutional neural network model, and the fault diagnosis results are obtained by using the softmax classifier.



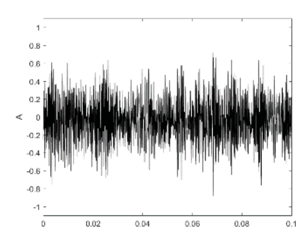
a.300r/min



b.1500r/min



c.3000r/min



d.4600r/min

Fig. 5 Time-domain signal diagram under the same working condition and different rotating speeds

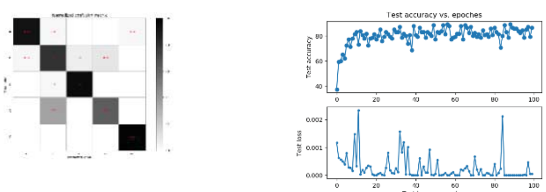
It can be seen from the figure that under the same working condition, the vibration signals collected at the same position at different speeds are very different, the characteristic frequencies caused by different faults are different, and the degree of fault characteristics reflected at different speeds is different. The vibration time domain signal contains a wealth of information, and the feature extraction and classification recognition of the fault can be realized to a certain extent by directly extracting and amplifying the feature.

The analysis object of this paper consists of five flywheels. The time domain signal map with total time of 30s is selected as input to train the above network. The data of each flywheel training group is 270, the test group data is 30, the time length of each time domain graph is 0.01s, and when training network, set 2 trainings per batch, the number of iterations is 100, parameter learning rate is 0.001. Meanwhile, a group of BP neural networks with the same parameter setting were trained for comparison.

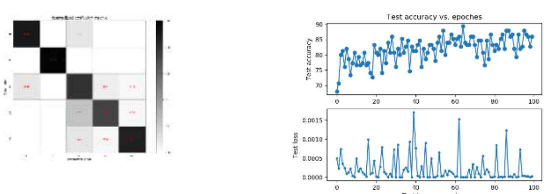
TABLE I. THE TRAINING RESULTS

revolving speed （r/min）	300	1500	3000	4600	all
BPNN accuracy	60%	50%	66%	65%	81%
CNN accuracy	87%	86%	82%	95%	98%

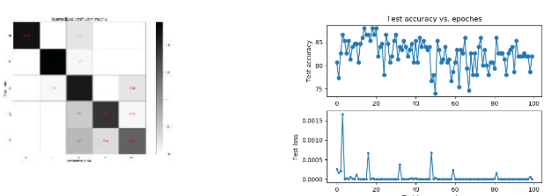
As can be seen from the above table, under similar network parameter Settings and the same input and output conditions, the accuracy of CNN is obviously better than that of BPNN.



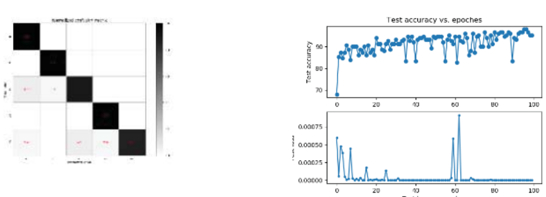
a.300r/min



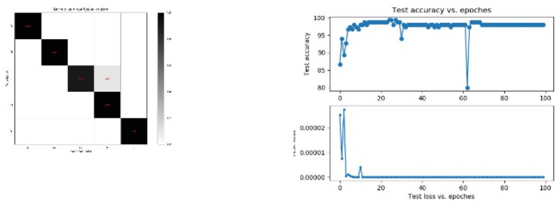
b.1500r/min



c.3000r/min



d.4600r/min



e. Four speed information fusion

Fig. 6 Confusion matrix and training process loss rate and accuracy change chart

Comparing the above confusion matrix diagrams, it can be clearly seen that when a single channel with a single speed is used as an input, the fault classification result is poor. At different speeds, the types of faults in the misclassification set are also different. Taking the speed of 1500r/min as an example, the number of iterations is far from insufficient from the error-reduced image. From a comprehensive comparison, the classification accuracy rate is greatly improved when multi-data fusion is performed, and the number of iterations required is small, and the efficiency and reliability of fault diagnosis are greatly improved. In short, the multi-information fusion has a higher accuracy rate, which is 3% to 19.5% higher than the single rotation speed.

V. CONCLUSION

This paper presents an intelligent diagnosis method of satellite flywheel bearing fault based on convolution neural network and multi-information fusion. Compared to the traditional fault diagnosis mode, this method does not require complex, professional artificial feature extraction process and professional prior reserve knowledge. Compared to traditional artificial intelligent methods, the accuracy of fault diagnosis is greatly improved. The diagnostic accuracy rate of the five types of working conditions reached 98%. Based on this research, more fault types and different severity fault data can be introduced for network training. It is expected that a more accurate and efficient flywheel bearing fault diagnosis network with better generalization performance can be obtained. In addition, this method can be further applied to other rotating machinery fault diagnosis processes.

ACKNOWLEDGMENT

This work was finically supported by the Beijing Key Laboratory of Long-life Technology of Precise Rotation and Transmission Mechanisms under Grant BZ0388201703 the Beijing Key Laboratory of Long-life Technology of Precise Rotation and Transmission Mechanisms.

REFERENCES

[1] Li B , Chow MY , Tipsuwan Y , et al. Neural-network-based motor rolling bearing fault diagnosis [J]. IEEE Transactions on Industrial Electronics, 2000, 47(5): 1060-1069.

[2] Tan SC , Lim CP . Evolutionary fuzzy ARTMAP neural networks and their applications to fault detection and diagnosis. [J]. Neural Processing Letters, 2010, 31(3): 219-242.

[3] Xiang Ling, Tang Guiji, Hu Aijun. Time-frequency analysis of non-stationary vibration signals of rotating machinery[J]. Journal of Vibration and Shock, 2010, 29(2): 42-45.

[4] Yang Jiangtian, Chen Jiayu. Extraction of Fault Symptoms of Rotating Machinery Based on Higher Order Spectrum[J]. Journal of Vibration Engineering, 2001(1): 13-18.

[5] Abbasion S , Rafsanjani A , Farshidianfar A , et al. Rolling element bearings multi-fault classification based on the wavelet denoising and support vector machine [J]. Mechanical Systems and Signal Processing, 2007, 21(7): 2933-2945.

[6] Bin GF , Gao JJ , Li XJ , et al. Early fault diagnosis of rotating machinery based on wavelet packets—Empirical mode decomposition feature extraction and neural network[J]. Mechanical Systems and Signal Processing, 2012, 27(none):696- 711.

[7] Li Yuhua, Weng Shenglong, Zhang Shaohui. A Firefly Neural Network and Its Application in Bearing Fault Diagnosis[J]. Journal of Mechanical Engineering, 2015(7): 99-106.

[8] Chen Dingyue. Composite vibration fault diagnosis based on information fusion and neural network[J]. Journal of Vibration, Testing and Diagnosis, 2004, 24(4): 290-293.

[9] Zhang Zhongkui. Application of RBF Neural Network in Fault Diagnosis of Rotating Machinery[J]. ChinaMachinery, 2014(6): 260-261.

[10] Yanjie D , Yisheng L V , Jie Z , et al. Research status and prospects of deep learning in the field of control[J]. Acta Automatica Sinica, 42(5).

[11] Ren Hao, Qu Jianfeng, Chai Yi, et al. Research status and challenges of deep learning in the field of fault diagnosis[J]. Control and Decision, 2017, 32(8): 1345-1358.

[12] Jia F , Lei Y , Lin J , et al. Deep neural networks: A promising tool for fault characteristic mining and intelligent diagnosis of rotating machinery with massive data [J]. Mechanical Systems and Signal Processing, 2016, 72-73:303- 315.

[13] Wei Dong, Gong Qingwu, Lai Wenqing, et al. Research on fault diagnosis and fault phase selection method in transmission line based on convolutional neural network[J]. Proceedings of the CSEE, 2016, 36(s1): 21 -28.

[14] Olivier Janssens, Viktor Slavkovikj, Bram Vervisch, et al. Convolutional neural network based fault detection forrotating machinery [J]. J of Sound and Vibration, 2016, 337: 331-345.

[15] Zhang W , Peng G , Li C , et al. A New Deep Learning Model for Fault Diagnosis with Good Anti-Noise and Domain Adaptation Ability on Raw Vibration Signals [J]. Sensors, 2017, 17(2): 425.

[16] Xia M , Li T , Xu L , et al. Fault Diagnosis for Rotating Machinery Using Multiple Sensors and Convolutional Neural Networks [J]. IEEE/ASME Transactions on Mechatronics, 2017, PP(99):1-1.

[17] Janssens O, Slavkovikj V, Vervisch B, et al. Convolutional Neural Network Based Fault Detection for Rotating Machinery [J]. Journal of Sound and Vibration, 2016: S0022460X16301638.

[18] Zeng Xueqiong. Research on classification and identification of transmission fault based on convolutional neural network [D]. 2016.

[19] Hu Zhi, Duan Lixiang, Zhang Laibin. Application of Optimized KNNC Algorithm in Fault Pattern Recognition of Rolling Bearings[J]. Journal of Vibration and Shock, 2013, 32(22): 84-87.