# A Robust Fault Diagnosis Method for Rolling Bearings Based on Deep Convolutional Neural Network

Zhenxiang Li
Key Laboratory of Space Utilization,
Technology and Engineering Center for Space Utilization,
Chinese Academy of Sciences
University of Chinese Academy of Sciences,
Beijing, China
lizhenxiang17@mails.ucas.ac.cn

Taisheng Zheng
Key Laboratory of Space Utilization,
Technology and Engineering Center for Space Utilization,
Chinese Academy of Sciences
University of Chinese Academy of Sciences,
Beijing, China
zhengtaisheng17@csu.ac.cn

Wang Yang
Key laboratory of Space utilization
Technology and Engineering Center for Space Utilization,
Chinese Academy of Sciences
Beijing, China
wangyangphm@ csu.ac.cn

Hongyong Fu
Key laboratory of Space utilization
Technology and Engineering Center for Space Utilization,
Chinese Academy of Sciences
Beijing, China
fuhongyong@csu.ac.cn

Wenbo Wu
Key laboratory of Space utilization
Technology and Engineering Center for Space Utilization,
Chinese Academy of Sciences
Beijing, China
wuwenbo@csu.ac.cn

***Abstract*—Fault diagnosis of rolling bearings has been an important and challenging research issue. The existing conventional algorithms for diagnosing rolling bearings rely on artificial feature extraction requiring a wealth of knowledge of signal handling, expertise and human efforts, lacking adaptability. With the capacity for learning features from original signals automatically, deep learning methods can solve the shortcomings of conventional diagnosis methods. This paper puts forward a deep learning method for rolling bearing fault diagnosis based on two-dimensional convolutional neural network, called 2dCNN-FD. The method has the capability of automatic feature extraction and robust noise tolerance. Experimental results demonstrate that comparing with other traditional diagnosis algorithms, the 2dCNN-FD model achieves high diagnosis accuracy and better robustness under different noise conditions.**

***Keywords- rolling bearing; fault diagnosis; deep convolutional neural network; robustness***

## I. INTRODUCTION

Being widely used in industrial equipment, rolling bearings play a crucial role in modern rotating machinery and their health conditions has a great impact on the machinery health and its remaining lifetime [1-3]. With an amount of effort putted into the topic, fault diagnosis of rolling bearings has been a one of research hotspots for decades [4]. The most popular way for bearing fault diagnosis is intelligent methods combined shallow machine learning (ML) with artificial feature extraction. An intelligent fault diagnosis algorithm usually contains three main steps: feature extraction, feature selection and feature classification. In general, the original vibration signal collected by the sensors contains a lot of useless noise, so the purpose of feature extraction is to obtain helpful information from original vibration data for fault classification. The widely-used methods of feature extraction include fast Fourier transform (FFT) [5], discrete wavelet transform (DWT) [6], empirical mode decomposition (EMD) [7] and so on. Feature selection is carried out after feature extraction to discard the insensitive and useless features and further reduce the number of effective features. Principle component analysis (PCA) [8], independent component analysis (ICA) [9] and manifold learning [10] are widely applied to select features. The last step, the selected features are inputted to classifiers such as multi-layer perceptron (MLP) [11], hidden Markov model (HMM) [12], support vector machine (SVM) [13] and so on for training to realize fault classification. The conventional methods mentioned above have been widely used to diagnosing faults of rolling bearings. Lin et al. [14] introduced a novel feature extraction approach based on Morlet wavelet and applied it for mechanical fault diagnosis. Cai et al. [15] combined EMD with genetic neural network adaptive boosting (GNN-AdaBoost) to realize fault feature extraction and classification. Kankar et al. [16] extracted features of the bearing original data by statistical methods and then made use of artificial neural network and SVM respectively for feature classification to complete diagnosis.

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

Although the traditional methods listed above achieve good diagnostic performance, they have two inevitable drawbacks. (1) Their feature extraction dependents on manual extraction and expertise within the domain. In engineering practice, the features suitable for one kind of mechanical equipment may fail in other equipment, so it is necessary to select other appropriate features for fault diagnosis, which increases the difficulty and workload of diagnosis. (2) The shallow architectures adopted in these conventional machine learning methods have limited capacity of mining the useful information from raw data [1].

As a major progress of machine learning, deep learning provides a feasible solution to the above problems. Deep learning utilizes the deep structure to realize the adaptive extraction of features through the nonlinear activation function of each layer, which overcomes the flaws of conventional methods, therefore it provides an efficient and novel method for fault diagnosis. Convolutional neural network (CNN), as an efficient deep learning method, has been widely used in image processing, speech recognition and other fields in recent years [17-19]. However, the research on using two-dimensional CNN for fault diagnosis is still in the early stage. Being inspired by the above facts, this paper establishes a two-dimensional CNN model to realize fault diagnosis using original data, which is named 2dCNN-FD. In the first place, original data are converted into two-dimensional data. Then the CNN model learns features of these two-dimensional samples automatically and classify them in the output layer. The algorithm requires neither any expert knowledge for signal processing nor additional feature extraction. A series of experiments are carried out to check the performance of the method.

The following structure of this paper is as follows. Section II briefly presents the theoretical knowledge of CNN. Section III introduces the 2dCNN-FD model at length. Experimental verification is given in Section IV. And Section V is the conclusion of this paper.

## II. CONVOLUTIONAL NEURAL NETWORK

CNN is a kind of neural networks with typical feed-forward structure, whose essence is extracting the required features from input by establishing multiple filters. A typical CNN usually consists of input layer, convolutional layer, pooling layer, fully connected (FC) layer and output layer, shown as Fig. 1.

### *A. Convolutional layer*

The function of convolutional layer is to convolve the output of the previous layer through its filters to realize features

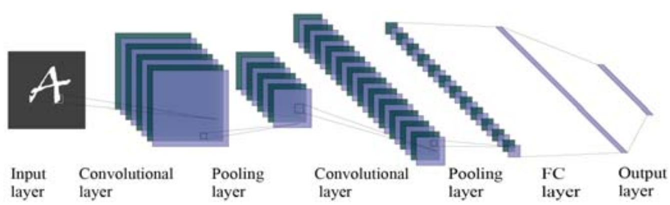


Figure 1. Schematic diagram of convolutional neural network

extraction. The convolutional layer can be expressed mathematically as follows.

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

where the symbol "*" represents convolution operation; $l$ represents the $l_{th}$ layer; $x_j^l$ is the output of the $j_{th}$ channel of the convolutional layer $l$; $M_j$ is a input feature vector; k is a kernel matrix and $b_j^l$ is a bias; $f(\cdot)$ represents an activation function.

### *B. Pooling layer*

A convolutional layer is usually followed by a pooling layer, which is also called subsampling layer and used to remove the redundant information of the output of the previous convolutional layer. The pooling layer can be expressed mathematically as follows.

$$x_j^l = f\left(\beta_j^l down(x_j^{l-1}) + b_j^l\right) \tag{2}$$

where $\beta$ is a weighting coefficient of pooling layer, $b_j^l$ is a bias and the symbol $down(\cdot)$ represents the downsampling function.

### *C. Fully connected layer*

In the FC layer, all neuron nodes are interconnected with all of those in the feature maps from the previous layer. Output of the FC layer can be expressed mathematically by the following equation:

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

where $w^l$ is a weighting coefficient and $b^l$ is a bias of the FC layer. $f(\cdot)$ represents an activation function.

### *D. Training*

The objective of training CNN is minimizing the loss function of the network $L(W,b)$. The widely-used loss functions are the mean square error (MSE) and the negative log-likelihood (NLL) function.

$$MSE(W,b) = \frac{1}{|Y|}\sum_{k=1}^{|Y|}\left(Y(k) - \overline{Y}(k)\right)^2 \tag{4}$$

$$NLL(W,b) = -\sum_{k=1}^{|Y|}\log Y(k) \tag{5}$$

where $Y(k)$ is the classification result of the $k_{th}$ class; $W$ is a weight matrixc and $b$ is a bias matrix.

To relieve overfitting, a L2 norm is usually added into the loss function. And the intensity of overfitting can be controlled by changing parameter $\lambda$. The final loss function is shown as (6).

$$R(W,b) = L(W,b) + \frac{\lambda}{2} W^T W \tag{6}$$

In training, the residual error propagates backward through gradient descent, which updates the trainable parameters of CNN layer by layer, shown as (7) and (8).

$$W_i = W_i - \eta \frac{\partial R(W,b)}{\partial W_i} \tag{7}$$

$$b_i = b_i - \eta \frac{\partial R(W,b)}{\partial b_i} \tag{8}$$

where $\eta$ is learning rate, whose role is to restrain the strength of the residual error backpropagation

## III. 2dCNN-FD Based Intelligent Diagnosis Method

The traditional fault diagnosis methods are lack of adaptability since they are based on artificial feature extraction. Besides, almost all of fault diagnosis models based on CNN just use CNN to classify fault types after artificial feature extraction at present, which does not give full play to the strong feature learning ability of convolutional neural network. For these reasons, an intelligent approach based on two-dimensional CNN, called 2dCNN-FD, is put forward to diagnose faults of rolling element bearings in the paper, which can adaptively mine the fault features from one-dimensional original signals of rolling bearings and automatically classifies bearing fault types and severity with these fault features.

### *A. Architecture of the Proposed 2dCNN-FD Model*

The 2dCNN-FD model is shown in Fig. 2, which contains three main parts: input part, feature extraction part and classification part.

In the input part, the original data is partitioned to many fragments by sliding window technique, and each fragment is taken as an input sample. Then, these 1D signals are transformed into 2D samples before being transmitted into the feature extraction part. The feature extraction part contains two convolutional layers and two pooling layers. The first convolutional layer has 64 kernels with a size of 4×4, which carries out feature extraction on the 2D signals to obtain 64 feature maps with size of 32×32. Then, the pooling layer compress these feature maps into a size of 16×16, which can not only reduce the dimension of these features to simplify the computational complexity of the network, but also extract main features from lots of fault information. After two convolutional layers and pooling layers, the 2dCNN-FD model completes feature extraction and input the extracted features into a multi-layer perception (MLP) for classification.

### *B. Model Training Procedure*

The training procedure of the proposed method is as follows.

(1). Augment the original dataset to obtain enough training samples. Input the one-dimensional data into the model, and convert it into a two-dimensional format.

(2). Initialize the 2dCNN-FD model by initializing the parameters such as the bias of the convolution kernels and the weight of the full connection layers to random numbers close to 0, and setting the number of iterations of the network, learning rate, etc.

(3). Input training samples to CNN network in fixed batch size. Obtain features by layer-by-layer forward propagation and get the diagnosis results in the output layer, then calculate the error between the expected output and the actual output.

(4). Utilize back-propagation (BP) algorithm to update weights layer by layer.

(5). Repeat step (3) and step (4) until all training samples have been trained and the desired accuracy or iteration times have been achieved.

## IV. Experiment and Result Analysis

To demonstrate the proposed method, two experiments are designed and conducted in this paper. In Experiment 1, the 2dCNN-FD model is trained and used to classify fault types and severity of rolling bearings. Experiment 2 is a contrast experiment that compares the 2dCNN-FD model with several widely-used fault diagnosis methods. The results indicate that, the proposed method has stronger capability of feature self-learning and robustness.

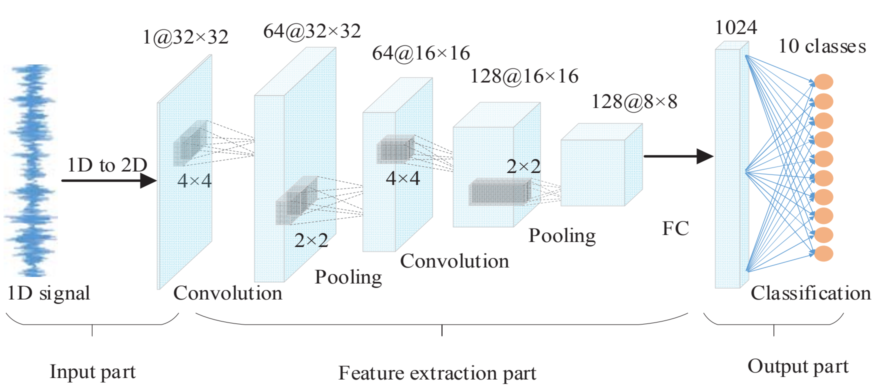


Figure 2. Architecture of the Proposed 2dCNN-FD Model

## A. *Data introduction*

The datasets used in this paper are acquired from the Case Western Reserve University Bearing Data Center Website [20]. The test-bed shown in Fig. 3 consists of two motors, a torque sensor, a dynamometer and other devices. Single point faults are created by using electro-discharge machining with damage diameters of 0.007, 0.014, 0.021 and 0.028 inches. Under loads from 0 to 3hp, vibration signals are gathered at a sampling frequency of 12 kHz and 48 kHz.

In this paper, fault data of 10 health states collected at a sampling frequency of 12000Hz and under a load of 2hp is selected as the experimental samples. 600 samples of each class is selected randomly for multi-state fault classification, which contains 400 training samples, 100 validation samples and 100 test samples. In this way, the total size of training set is 4000, and the total size of validation set and test set is 1000 respectively. The details of chosen samples are listed in Table I.

## B. *Experiment 1: 2dCNN-FD based Fault Classification*

In the training process, appropriate model parameters can accelerate the rate of convergence of the model and make the model obtain high fault diagnosis accuracy. Cross validation is adopted to select the optimal parameters for 2dCNN-FD models in this paper, shown in Table II. With the parameter setting, the model is trained to classify the input samples, and model performances are recorded in the training progress, shown in Fig. 4 and Fig. 5. For better illustration, only the first 5000 epochs of the model loss curve and the first 1000 epochs of the classification accuracy curve are shown.

As for the classification accuracy, as shown in Fig. 5, the curves of the training accuracy and validation accuracy both rise rapidly and almost reach 100% after about 600 epochs. The final testing accuracy is 99.9%, which indicates that the proposed algorithm performs outstanding in fault diagnosis.

## C. *Experiment 2: Comparison with Other Algorithms*

To further check the performance of the 2dCNN-FD model, a series of tests are performed to compare it with four widely-used diagnosis algorithms, namely one-dimensional CNN (1dCNN), MLP, SVM and random forest (RF). It can be demonstrated that the proposed model has a stronger ability in feature self-learning and the robustness.

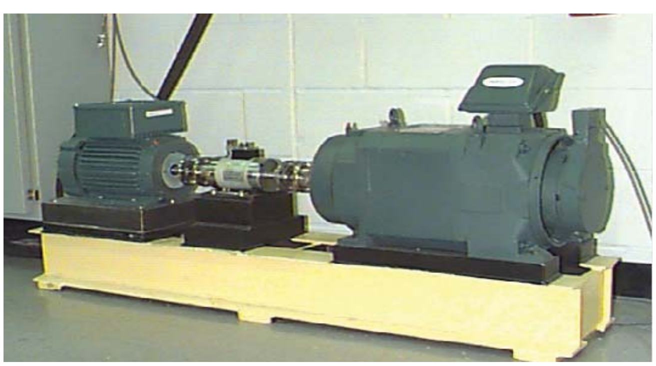

Figure 3. Schematic diagram of sampling

TABLE I. THE EXPERIMENTAL SAMPLES

| **Sample Types** | **The Details of Samples** | | | | | |
|---|---|---|---|---|---|---|
| | ***Damage diameters (inch)*** | ***Sample length*** | ***Training set*** | ***Validation set*** | ***Testing set*** | ***Label*** |
| Normal | —— | 1024 | 400 | 100 | 100 | 0 |
| Ball-1 | 0.007 | 1024 | 400 | 100 | 100 | 1 |
| Ball-2 | 0.014 | 1024 | 400 | 100 | 100 | 2 |
| Ball-3 | 0.021 | 1024 | 400 | 100 | 100 | 3 |
| Inner-1 | 0.007 | 1024 | 400 | 100 | 100 | 4 |
| Inner-2 | 0.014 | 1024 | 400 | 100 | 100 | 5 |
| Inner-3 | 0.021 | 1024 | 400 | 100 | 100 | 6 |
| Outer-1 | 0.007 | 1024 | 400 | 100 | 100 | 7 |
| Outer-2 | 0.014 | 1024 | 400 | 100 | 100 | 8 |
| Outer-3 | 0.021 | 1024 | 400 | 100 | 100 | 9 |

TABLE II. TRAINING PARAMETERS OF THE 2DCNN-FD MODEL

| ***Batch size*** | ***Kernel size*** | | ***Epochs*** | ***Learning rate*** |
|---|---|---|---|---|
| 50 | 4×4 | 2×2 | 10000 | $10^{-6}$ |

*1) Feature Self-learning Ability:* For conventional fault diagnosis methods, artificial feature extraction is an indispensable work before classification. However, with feature self-learning ability, the 2dCNN-FD model can achieve high diagnosis accuracy even if the original vibration signal is taken as input directly. To verify that, the five methods mentioned above are used to classify the original data and the data after artificial feature extraction respectively. The feature extraction methods used here are FFT and DWT, therefore the corresponding data can be called *Data after FFT* and *Data after DWT* respectively. The test accuracy of the five diagnosis models is shown in Fig. 6. From Fig. 6, we can find that the test accuracy of 2dCNN-FD can reach a very high value close to 100%, no matter using the original data, the Data after FFT or the Data after DWT, while the diagnosis accuracy of the other methods for raw data is obviously lower than that for the data after feature extraction. It has a strong ability in feature self-

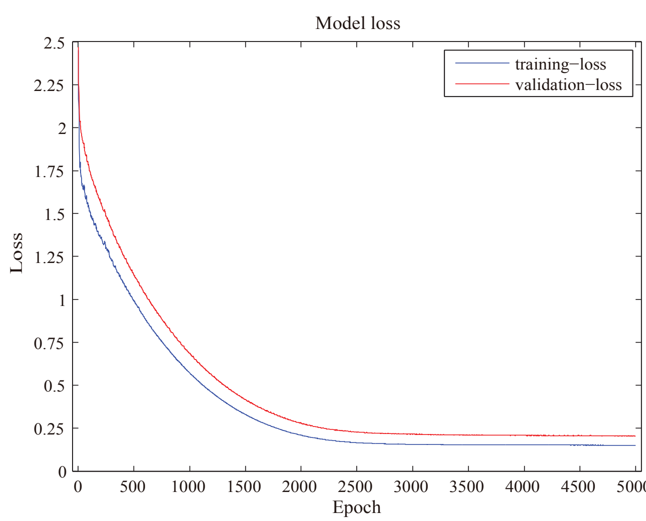


Figure 4. Loss of the 2dCNN-FD model

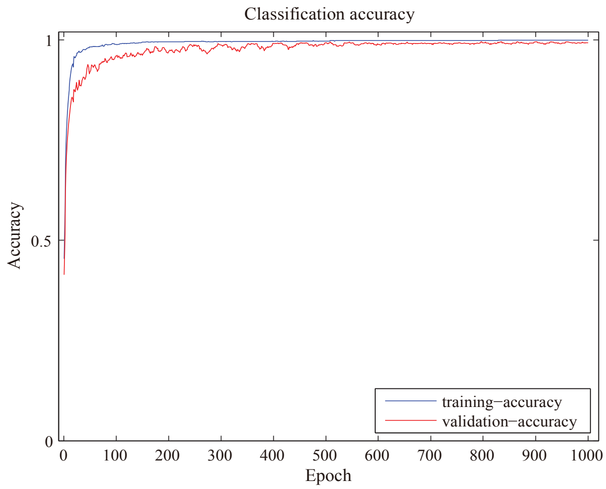


Figure 5. Accuracy of the 2dCNN-FD model

learning for original fault data, so that it can achieve a high diagnosis accuracy even without artificial feature extraction.

*2) Robustness:* In real application, rolling bearings generally work in a noisy environment, which would increase the difficulty in fault diagnosis. Therefore, in this section we test the robustness of the 2dCNN-FD model under the interference of noise samples.

The test is conducted as follows. Random noise is used to produce a certain number of noise samples. Then these noise samples are added to the original training set to simulate a contaminated input while the test set remains the same. The diagnosis model will be trained with the contaminated training set and be test with the original test set. Table III shows the diagnosis accuracy of the 2dCNN-FD model with noise samples varying from 0 to 2000. Obviously, with the increase of noise samples, the accuracy of the 2dCNN-FD model fluctuates little and remains a high level. Even when the ratio of noise sample to original sample is as high as 1 to 2, the diagnosis accuracy can still reach 99.1%.

As comparison, the other four methods are used for the same test and the result is record in Table IV. For a more intuitive comparison, the results of the five methods are plotted in Fig. 7. As can be seen from Fig. 7, with the increase of noise samples, the test accuracy of 2dCNN-FD remains close to 1, while the test accuracy of MLP, SVM and RF all descend, especially that of SVM. This indicates that the 2dCNN-FD model is more robust than other methods. However, the test accuracy of 1dCNN increases slowly with the number of noise samples. With a simple analysis, the reason may be that the added noise samples help alleviate the overfitting problem of the model to some extent.

To further verify the robustness of the model, the five methods are also used for tests on the *Data after FFT* and *Data after DWT* under different noise environments. Table V and Table VI respectively give the results of the test on the *Data after FFT* and the test on *Data after DWT*. From Table V, we find that the test accuracy of 2dCNN-FD and RF both remain 100% no matter how many noise samoles are added into the training set, and MLP also achives a high test accuracy close to 100%. Meanwhile, the test accuracy of CNN increases slowly

TABLE III. PERFORMANCE OF 2DCNN-FD MODEL UNDER DIFFERENT NOISE CONDITIONS

| ***Number of training samples*** | ***Number of noise samples*** | ***The ratio of noise sample to fault sample*** | ***Diagnosis accuracy*** |
|---|---|---|---|
| 4000 | 0 | —— | 0.999 |
| 4000 | 500 | 1:8 | 0.998 |
| 4000 | 1000 | 2:8 | 0.998 |
| 4000 | 1500 | 3:8 | 1 |
| 4000 | 2000 | 4:8 | 0.991 |

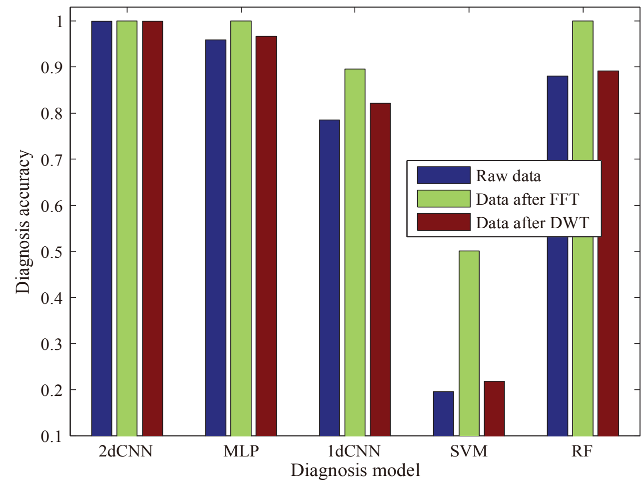


Figure 6. Diagnosis accuracy before and after artificial feature extraction

TABLE IV. PERFORMANCE OF THE OTHER FOUR MODELS WITH NOISE SAMPLES ADDED TO RAW DATA

| **Model** | **Number of Noise Samples Added to Raw Data** | | | | |
|---|---|---|---|---|---|
| | ***0*** | ***500*** | ***1000*** | ***1500*** | ***2000*** |
| 1dCNN | 0.785 | 0.794 | 0.8 | 0.805 | 0.827 |
| MLP | 0.959 | 0.911 | 0.886 | 0.867 | 0.863 |
| SVM | 0.196 | 0.137 | 0.138 | 0.097 | 0.092 |
| RF | 0.88 | 0.844 | 0.862 | 0.835 | 0.823 |

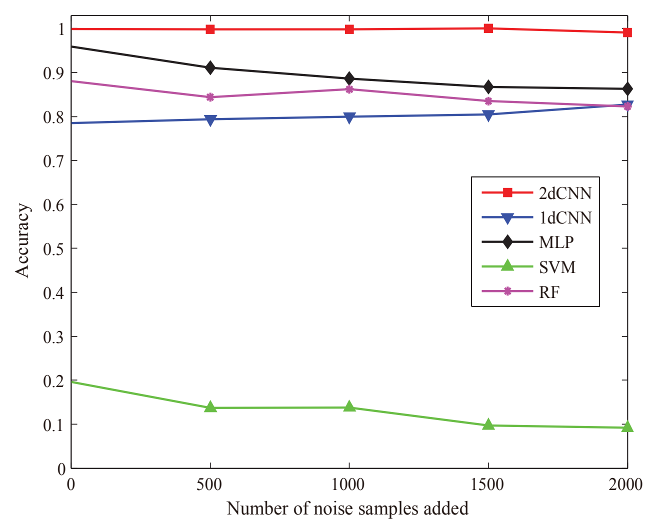


Figure 7. Accuracy of methods with noise samples added to raw data

TABLE V. PERFORMANCE OF ALL FIVE MODELS WITH NOISE SAMPLES ADDED TO DATA AFTER FFT

| Model | Number of Noise Samples Added to Data after FFT | | | | |
|---|---|---|---|---|---|
| | *0* | *500* | *1000* | *1500* | *2000* |
| 2dCNN-FD | 1 | 1 | 1 | 1 | 1 |
| 1dCNN | 0.896 | 0.913 | 0.91 | 0.919 | 0.901 |
| MLP | 1 | 1 | 1 | 0.996 | 0.997 |
| SVM | 0.501 | 0.101 | 0.101 | 0.101 | 0.101 |
| RF | 1 | 1 | 1 | 1 | 1 |

TABLE VI. PERFORMANCE OF ALL FIVE MODELS WITH NOISE SAMPLES ADDED TO DATA AFTER DWT

| Model | Number of Noise Samples Added to Data after DWT | | | | |
|---|---|---|---|---|---|
| | *0* | *500* | *1000* | *1500* | *2000* |
| 2dCNN-FD | 0.999 | 0.994 | 0.996 | 0.994 | 0.987 |
| 1dCNN | 0.821 | 0.799 | 0.82 | 0.823 | 0.832 |
| MLP | 0.966 | 0.906 | 0.885 | 0.87 | 0.867 |
| SVM | 0.218 | 0.114 | 0.11 | 0.08 | 0.088 |
| RF | 0.891 | 0.888 | 0.862 | 0.866 | 0.853 |

with the added noise samples. However, the accuracy of SVM descends dramatically after adding noise samples to training set and then remains stable at a low value. As for Table VI, it is easy to see that the result is similar with that of test using raw data, therefore it will not be repeated here.

The comparison results of these tests indicate that the 2dCNN-FD method has strong robustness and can achieve high diagnosis accuracy even with dataset contaminated by noise.

## V. CONCLUSION

In this paper, a new method based on 2D CNN is proposed for fault diagnosis of rolling bearings. By transforming 1D vibration signals into 2D data and giving full play to the advantages of CNN in processing 2D data, the proposed algorithm can achieve an accuracy of 99.9% even 100% in CWRU bearing data set. Compared to other fault diagnosis methods based on traditional machine learning, the main advantage of the method can be generalized as follows:

(1) For the strong feature self-learning ability, it can achieve high diagnosis accuracy in classifying multi-state fault types and severity with the raw vibration signals without any artificial feature extraction.

(2) Furthermore, the proposed method has strong robustness and tolerance for noise. It does not require any denoising steps, but still achieve satisfactory diagnosis results under different noise conditions.

In the following work, testing the CNN model on more datasets is a work to look forward to. Moreover, the exploration of the self-learning ability and the robustness of CNN is another interesting but challenging problem.

## REFERENCES


[1] F. Jia, Y. G. Lei, J. Lin, X. Zhou, and N. Lu, “Deep neural networks: A promising tool for fault characteristic mining and intelligent diagnosis of rotating machinery with massive data,” Mechanical Systems and Signal Processing, vol. 72-73, pp. 303-315, May, 2016.

[2] I. El-Thalji, and E. Jantunen, “A summary of fault modelling and predictive health monitoring of rolling element bearings,” Mechanical Systems and Signal Processing, vol. 60-61, pp. 252-272, Aug, 2015.

[3] W. A. Smith, and R. B. Randall, “Rolling element bearing diagnostics using the Case Western Reserve University data: A benchmark study,” Mechanical Systems and Signal Processing, vol. 64-65, pp. 100-131, 2015.

[4] H. Qiu, J. Lee, J. Lin, and G. Yu, “Robust performance degradation assessment methods for enhanced rolling element bearing prognostics,” Advanced Engineering Informatics, vol. 17, no. 3-4, pp. 127-140, Jul-Oct, 2003.

[5] Y. Y. He, J. Huang, and B. Zhang, “Approximate entropy as a nonlinear feature parameter for fault diagnosis in rotating machinery,” Measurement Science and Technology, vol. 23, no. 4, pp. 14, Apr, 2012.

[6] L. Ying, X. Yuejun, H. Yigang, U. Zhen Zhang %+ Changsha, T. Hefei University of, U. Changsha, T. Hefei University of, and U. Changsha, “A new switched current circuit fault diagnosis approach based on pseudorandom test and preprocess by using entropy and Haar wavelet transform,” Analog Integrated Circuits and Signal Processing, vol. 91, no. 3, 2017.

[7] D. J. Yu, J. S. Cheng, and Y. Yang, “Application of EMD method and Hilbert spectrum to the fault diagnosis of roller bearings,” Mechanical Systems and Signal Processing, vol. 19, no. 2, pp. 259-270, Mar, 2005.

[8] B. B. Hu, and B. Li, “A new multiscale noise tuning stochastic resonance for enhanced fault diagnosis in wind turbine drivetrains,” Measurement Science and Technology, vol. 27, no. 2, pp. 14, Feb, 2016.

[9] W. P. He, Y. Y. Zi, B. Q. Chen, F. Wu, and Z. J. He, “Automatic fault feature extraction of mechanical anomaly on induction motor bearing using ensemble super-wavelet transform,” Mechanical Systems and Signal Processing, vol. 54-55, pp. 457-480, Mar, 2015.

[10] Q. S. Jiang, M. P. Jia, J. Z. Hu, and F. Y. Xu, “Machinery fault diagnosis using supervised manifold learning,” Mechanical Systems and Signal Processing, vol. 23, no. 7, pp. 2301-2311, Oct, 2009.

[11] H. Q. Wang, and P. Chen, “Intelligent diagnosis method for rolling element bearing faults using possibility theory and neural network,” Computers & Industrial Engineering, vol. 60, no. 4, pp. 511-518, May, 2011.

[12] A. Soualhi, G. Clerc, H. Razik, A. Lebaroud, and Ieee, Fault Detection and Diagnosis of Induction Motors Based on Hidden Markov Model, New York: Ieee, 2012.

[13] A. Widodo, and B. S. Yang, “Support vector machine in machine condition monitoring and fault diagnosis,” Mechanical Systems and Signal Processing, vol. 21, no. 6, pp. 2560-2574, Aug, 2007.

[14] J. Lin, and L. S. Qu, “Feature extraction based on Morlet wavelet and its application for mechanical fault diagnosis,” Journal of Sound and Vibration, vol. 234, no. 1, pp. 135-148, Jun, 2000.

[15] G. Q. Cai, C. Yang, Y. Pan, and J. J. Lv, “EMD and GNN-Adaboost Fault Diagnosis For Urban Rail Train Rolling Bearings,” Discrete and Continuous Dynamical Systems-Series S, vol. 12, no. 4-5, pp. 1471-1487, Aug-Sep, 2019.

[16] P. K. Kankar, S. C. Sharma, and S. P. Harsha, “Fault diagnosis of ball bearings using machine learning methods,” Expert Systems with Applications, vol. 38, no. 3, pp. 1876-1886, Mar, 2011.

[17] R. Girshick, J. Donahue, T. Darrell, J. Malik, and Ieee, "Rich feature hierarchies for accurate object detection and semantic segmentation," 2014 Ieee Conference on Computer Vision and Pattern Recognition, IEEE Conference on Computer Vision and Pattern Recognition, pp. 580-587, New York: Ieee, 2014.

[18] A. Karpathy, G. Toderici, S. Shetty, T. Leung, R. Sukthankar, L. Fei-Fei, and Ieee, "Large-scale Video Classification with Convolutional Neural Networks," 2014 Ieee Conference on Computer Vision and Pattern Recognition, IEEE Conference on Computer Vision and Pattern Recognition, pp. 1725-1732, New York: Ieee, 2014.

[19] S. Q. Ren, K. M. He, R. Girshick, and J. Sun, “Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks,” Ieee Transactions on Pattern Analysis and Machine Intelligence, vol. 39, no. 6, pp. 1137-1149, Jun, 2017.

[20] "The Case Western Reserve University Bearing Data Center Website," http://csegroups.case.edu/bearingdatacenter/home.