



One-Dimensional Convolutional Neural Network Based Bearing Fault Diagnosis

Jiaxue Chen, Xiaoqi Yin*, Chenxue Li, Hang Yang, Li Hong

Faculty of Electronic Information Engineering, Huaiyin Institute of Technology, Huai'an, China

Email: *hy_xuebao2009@126.com

How to cite this paper: Chen, J.X., Yin, X.Q., Li, C.X., Yang, H. and Hong, L. (2022) One-Dimensional Convolutional Neural Network Based Bearing Fault Diagnosis. *Open Access Library Journal*, 9: e8644. <https://doi.org/10.4236/oalib.1108644>

Received: March 23, 2022

Accepted: April 10, 2022

Published: April 13, 2022

Copyright © 2022 by author(s) and Open Access Library Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

Fault diagnosis based on time-domain signals has become mainstream in recent years. Traditional methods require signal processing to extract fault features before feeding them into a neural network for diagnostic classification, which is a cumbersome process. This paper proposes an adaptive fault diagnosis model based on a one-dimensional convolutional neural network, and the structure and parameters of the model are analyzed and designed in detail. The segmented pre-processed vibration signal is fed directly into a convolutional neural network, where fault features can be extracted adaptively, and finally classify the diagnostic results using a Softmax classifier. This method directly processes the vibration signals in an end-to-end way, which improves the timeliness of diagnosis. The effectiveness of the method is verified through bearing experiments and compared with KNN, SVM, LSTM and AlexNet models. The results show that the model is accurate for fault diagnosis of bearings.

Subject Areas

Electric Engineering

Keywords

Convolutional Neural Network, Fault Diagnosis, Rolling Bearing, Intelligent Diagnosis

1. Introduction

With the rapid development of science and technology, machinery and equipment in modern industry is developing in the direction of precision and automation. If a part of the equipment fails, it will have a series of serious consequences. Bearings are an important part of rotating machinery, in which they

play a role in supporting the spindle and transmitting torque. In rotating machinery with rolling bearings, about 30% of mechanical failures are caused by rolling bearings [1]. This is why it is of great significance to carry out fault diagnosis on rolling bearings.

Fault diagnosis of bearings is achieved by analysing and processing the vibration signal on the bearing and extracting the signal characteristics. Traditional fault diagnosis methods are mainly time domain analysis, frequency domain analysis and time-frequency domain analysis. The time domain analysis method is to display all the information contained in the vibration signal on the time axis for analysis, the main method is time domain statistical analysis [2]. Frequency domain analysis is the Fourier transform of the vibration signal to obtain a spectrum, commonly used methods are power spectrum analysis [3], cepstrum analysis [4] and envelope spectrum analysis [5]. The analysis of time-frequency domain is not only to analyse the frequency domain characteristics of the signal, but also to analyse the transient information of the signal. The main methods include Fourier transform [6], Wigner-Ville distribution [7], wavelet transform [8] and empirical modal decomposition [9].

Nowadays more and more machine learning algorithms are widely used in the fault diagnosis, such as K-Nearest Neighbour algorithm (KNN) and Support Vector Machine (SVM), both of which possess strong data processing and automatic fault identification capabilities. Lu Dunli *et al.* [10] combined KNN with Naïve Bayes making and inputting the results of KNN preliminary classification to Naïve Bayes for further classification, which effectively improved the fault diagnosis rate. Zhao Chunhua *et al.* [11] automatically extracted fault features through deep learning, and then diagnosed them through a fault diagnosis model combining the whale optimization algorithm and SVM. Chen Renxiang *et al.* constructed the time-frequency matrix of the signal through wavelet transform and then used it as the input of convolutional neural network for classification, which has high generalisation performance [12]. Jian Di establishes a CAPSO-DAEN fault diagnosis model based on Deep Auto-encoder Network optimized by Cloud Adaptive Particle Swarm Optimization. The model uses the randomness and stability of CAPSO algorithm to optimize the connection weight of DAEN, reduce the constraints on the weights and extract fault features adaptively [13].

Although many effective fault diagnosis methods have been proposed, most of them need to extract the fault features of the vibration signal first, which greatly reduces the timeliness. Based on previous research, a one-dimensional convolutional neural network based method is proposed for fault diagnosis in this paper. The method does not require human extraction of fault features and directly uses the original vibration signal as the input of the convolutional neural network to complete the classification of faults.

2. One-Dimensional Convolutional Neural Networks

The process of convolutional neural networks takes raw data as the input to the

algorithm and gradually abstract the raw data into the desired feature representation through operations such as convolution, pooling and mapping of non-linear activation functions. Ends with a feature-to-target task mapping [14]. The structure of a convolutional neural network consists of an input layer, a convolutional layer, a pooling layer, a fully connected layer and an output layer.

As an indispensable part of convolutional neural network, convolutional layer serves to enhance the original features of signals and reduce the noises [15]. Each element of the convolution kernel has its corresponding weight coefficient and bias [16]. During the training process, the weight coefficients and biases are constantly modified by back propagation to obtain the best parameters. The calculation function of convolution layer is as follows:

$$g(i) = \sum_{x=1}^m \sum_{y=1}^n \sum_{z=1}^p a_{x,y,z} \times w_{x,y,z}^i + b^i \quad (1)$$

where $g(i)$ is the mapping graph obtained from the i th convolution kernel and a is the input data. w is the weight factor of the convolution kernel. b is the bias of the convolution kernel. If the input data is an image, x and y represent the pixel size of the image in the length and width directions respectively, and z is the number of channels. If the input data is a one-dimensional time-domain signal, y and z are simplified.

Pooling layers between convolutional layers aim to strengthen the translational invariance and reduce the dimension of the feature maps [17]. The pooling mainly includes the average pooling and the maximum pooling. Average pooling is to extract the local average value, and the maximum pooling is to extract the local maximum value. The function representation of the maximum pool is as follows:

$$P_i^{l+1}(j) = \max_{(j-1)W+1 \leq t \leq jW} \{q_t^l(t)\} \quad (2)$$

where $t \in [(j-1)W+1, jW]$ and W is the width of the region to be pooled. $q_t^l(t)$ represents the value of the t th neuron in the l th feature vector on the l th layer. $P_i^{l+1}(j)$ represents the value of the neuron on the $l+1$ layer.

The fully-connected layer connects each neuron to all the neurons in the previous layer. This means that the extracted features are stitched together and the Softmax classifier is used to output the multi-classification result. The function expression is as follows:

$$O = f(b_o + w_o f_v) \quad (3)$$

where f_v represents the eigenvector, and the b_o and w_o subscales represent the deviation vector and the weight matrix.

The one-dimensional convolutional neural network is also composed of three main parts: the convolutional layer, the pooling layer and the fully connected layer. A one-dimensional signal is fed into the input layer of the one-dimensional convolutional neural network. A one-dimensional signal is convolved with a one-dimensional convolution kernel to generate an input feature map. Then the

input feature maps are passed through the activation function to generate the output feature maps of the convolution layer [18].

3. 1D-FDCNN Fault Diagnosis Algorithms

This paper proposes a fault diagnosis model based on a one-dimensional convolutional neural network (1D-FDCNN), which is divided into three parts, namely the input layer, the fault feature extraction layer and the classification layer (**Figure 1**). The input layer mainly accomplishes the pre-processing of the input data and cuts the signal to a certain length. The feature extraction layer consists of a convolutional layer and a pooling layer. Fault features of the input data are first extracted using a convolutional kernel, and then a dimensionality reduction operation is performed in the pooling layer. Multiple alternating convolutional and pooling layers can extract the non-linear features of the input signal [19]. The classification layer then uses the Softmax function to classify the target task. The model structure is an input layer, a group of 3 convolutional layers (consisting of alternating convolutional and pooling layers), 2 fully connected layers, and an output layer. Since vibration signals are periodic, the final pooling layer of the proposed model in this paper should have a sensory field size greater than the size of a periodic signal in the input signal so that more comprehensive fault information can be obtained.

The relationship between signal period and sampling frequency and rotational speed is as follows:

$$T = f \times \frac{60}{r} \quad (4)$$

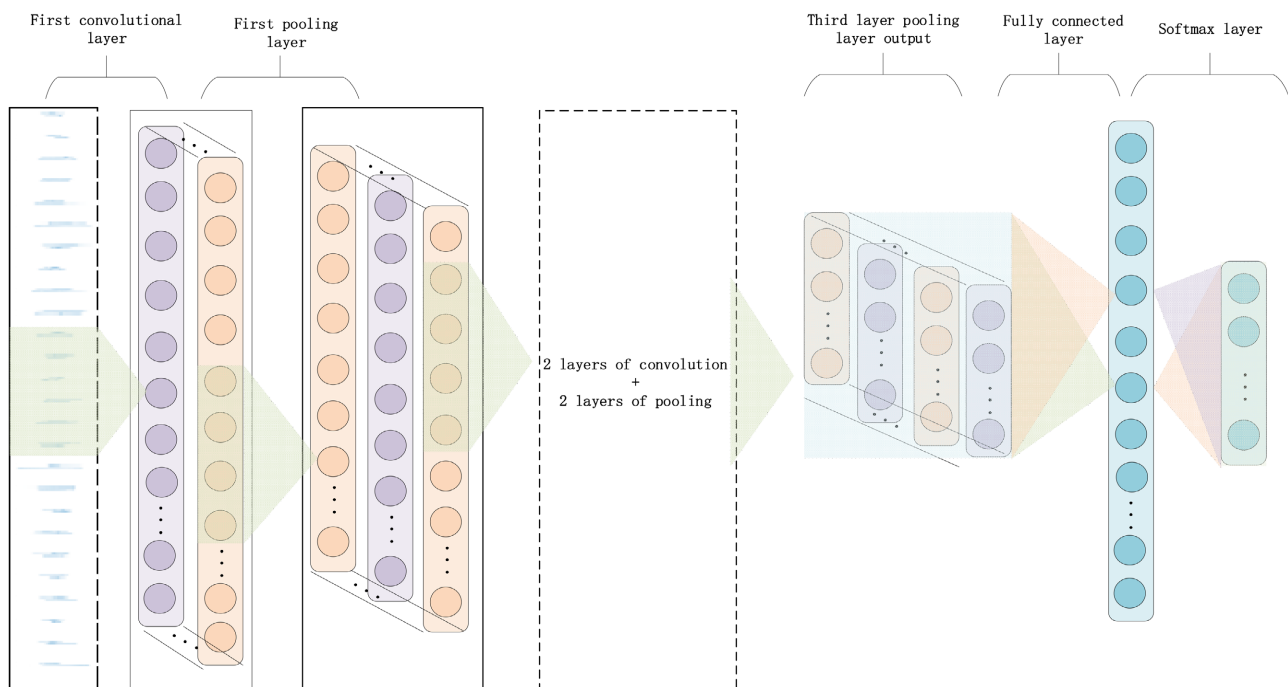


Figure 1. Architecture of 1D-FDCNN.

where T is the period, f is the sampling frequency and r is the speed. A period of 400 is obtained when f is 12 kHz and r is 1797 r/min, which indicates that the last pooling layer cannot have a perceptual field of less than 400 samples in the input signal. The perceptual field of the last pooling layer in the input signal cannot be less than 400 samples. In a one-dimensional convolutional neural network, the relationship between the receptive field $F^{(i)}$ on the i th pooling layer and the receptive field $F^{(i-1)}$ on the $(i-1)$ th pooling layer is as follows:

$$F^{(i-1)} = S^{(i)} \left(P^{(i)} F^{(i)} - 1 \right) + W^{(i)} \quad (5)$$

where W is the width of the convolution kernel on the convolution layer, S is the step length on the convolution layer, and P is the size of the pooling layer. In 1D-FDCNN, when $i > 1$, $W^{(i)} = 3$, $S^{(i)} = 1$ and $P^{(i)} = 2$, the above function (5) can be simplified to function (6):

$$F^{(i)} = \frac{F^{(i-1)}}{2} - 1 \quad (6)$$

When both i and the number of convolution kernels are 3, $F^{(3)} = 1$. So the final pooling layer has a perceptual field of 10 above the first pooling layer, that is, $F^{(1)} = 10$. Substituting this into function (4) gives the magnitude of the perceptual field $F^{(0)}$ of the final pooling layer on the input signal, as shown in the following function (7):

$$F^{(0)} = S^{(1)} \left(P^{(1)} F^{(1)} - 1 \right) + W^{(1)} = 10P^{(1)}S^{(1)} - S^{(1)} + W^{(1)} \approx 10P^{(1)}S^{(1)} \quad (7)$$

When $P^{(1)}$ is 2, $F^{(0)}$ is $20S^{(1)}$. According to the principle that $T \leq F^{(0)} \leq L$ and $S^{(1)}$ should be able to divide L , where L is the length of the input signal. The following function (8) can be obtained by substitution (8):

$$\begin{aligned} T &\leq 20S^{(1)} \leq L \\ S^{(1)} &| L \end{aligned} \quad (8)$$

According to the above equation, when L is taken as 2048, $S^{(1)}$ can be taken as 32 or 64, and the step size of the first convolutional layer of 1D-FDCNN in this paper is taken as 32. In 1D-CNN, a large convolutional kernel can bring a larger perceptual field and does not increase too many parameters and computation [20], so a convolutional kernel of size 64×1 is used in the first convolutional layer of 1D-FDCNN to obtain a more comprehensive fault feature of the input signal sequence. And considering that the network framework should not be too deep, the step size of the second and third convolution layers is set to 1, and the size of the pooling layer is set to 2.

The specific parameter table is shown in **Table 1** below.

4. Experiments

4.1. Experimental Data Sets

The data for this experiment was taken from the Case Western Reserve University (CWRU) open bearing dataset [21], where the bearing type is SKF6205 and

the sampling frequency in the system is 12 kHz.

The data set for this experimental platform was acquired under four loads, denoted by 0, 1, 2 and 3, where the different loads correspond to different rotational speeds. The speed of load 0 is 1797 r/min, load 1 is 1772 r/min, load 2 is 1750 r/min and load 3 is 1730 r/min. There are three types of defective locations for bearings with faults, namely, inner and outer rings and rolling elements, with diameter sizes of 0.007 inch, 0.014 inch and 0.021 inch. There are nine defective states, plus one healthy state, for a total of ten states. In this experiment, the length of each sample is 2048, each state contains 1000 samples, and is divided into training set, test set and validation set in the ratio of 0.7:0.2:0.1, that is, a data set contains 10,000 samples. The experimental data set is shown in **Table 2** below.

Table 1. 1D-FDCNN network parameters

Layer	Kernel size	Stride	Kernels numbers	Outputs size	Padding
Convolutional layer 1	64×1	32	16	64×16	yes
Pooling layer 1	2×1	2	16	32×16	no
Convolutional layer 2	3×1	1	32	32×32	yes
Pooling layer 2	2×1	2	32	16×32	no
Convolutional layer 3	3×1	1	64	16×64	yes
Pooling layer 2	2×1	2	64	8×64	no
Fully-connected layer	200	1	1	200×1	
Softmax	10	1	1	10×1	

Table 2. Experimental data set.

Fault area		Normal		Inner ring		Outer ring			Rolling body			Load
Diameter of damage		0	0.007	0.014	0.021	0.007	0.014	0.021	0.007	0.014	0.021	
A	Training	700	700	700	700	700	700	700	700	700	700	0
	Testing	200	200	200	200	200	200	200	200	200	200	
	Validation	100	100	100	100	100	100	100	100	100	100	
B	Training	700	700	700	700	700	700	700	700	700	700	1
	Testing	200	200	200	200	200	200	200	200	200	200	
	Validation	100	100	100	100	100	100	100	100	100	100	
C	Training	700	700	700	700	700	700	700	700	700	700	2
	Testing	200	200	200	200	200	200	200	200	200	200	
	Validation	100	100	100	100	100	100	100	100	100	100	
D	Training	700	700	700	700	700	700	700	700	700	700	3
	Testing	200	200	200	200	200	200	200	200	200	200	
	Validation	100	100	100	100	100	100	100	100	100	100	

4.2. Optimisation of Model Parameters

Table 2 above shows that the number of experimental datasets is relatively large. If only one sample is input to the model for training at each time, it will not only take a lot of time, but also over-fitting will occur, which will greatly reduce the robustness of the model. Therefore, it is necessary to find a suitable number of batch samples so that the model can achieve a high diagnostic accuracy without falling into a local optimum. In this paper, 6 batch sample numbers were set. The experiments were conducted in sequence and the diagnostic results obtained are shown in **Table 3** below. It can be seen from the table that the accuracy of the validation set reaches the highest when the number of batch samples is 32, and the average training time is close to the best result. Therefore, the number of batch samples in this model is 32.

4.3. Bearing Fault Diagnosis Experiments

To validate the effectiveness of the 1D-FDCNN model proposed in Section 2, the model is trained and evaluated using the datasets A, B, C and D in Section 3.1, respectively, where the accuracy of the validation set and loss are used to evaluate the model.

The 1D-FDCNN model training process is shown in **Figure 2** below.

The epoch was set to 100 times for training, the number of batch samples was 32, and the loss function was cross-entropy. The model was run on a software environment of Python version 3.8 and a hardware environment of AMD Ryzen 7 4800U with Radeon Graphics 1.80 GHz with 16.0 GB of memory. The training results are shown in **Figure 3** and **Figure 4** below.

In **Figure 3** and **Figure 4**, it can be seen that after 30 epochs, the model basically stabilises, with the training loss of the four data sets being basically smooth and tending to 0. The accuracy curve shows that the accuracy of data set A is higher than that of data sets B, C and D, but the overall accuracy is at 95% and above. The experimental results show that the model can achieve high accuracy diagnosis of bearing faults and has good robustness.

In order to verify the performance of the algorithm proposed in this paper

Table 3. The influence of batch sample number on accuracy.

Batch sample size	Average diagnostic accuracy $\times 100\%$		Average training time/s
	Training	Validation	
8	97.27	96.65	3.21
16	98.53	98.16	2.18
32	98.39	98.39	1.68
64	98.35	97.97	1.04
128	97.39	96.73	0.72
256	96.26	92.03	0.59

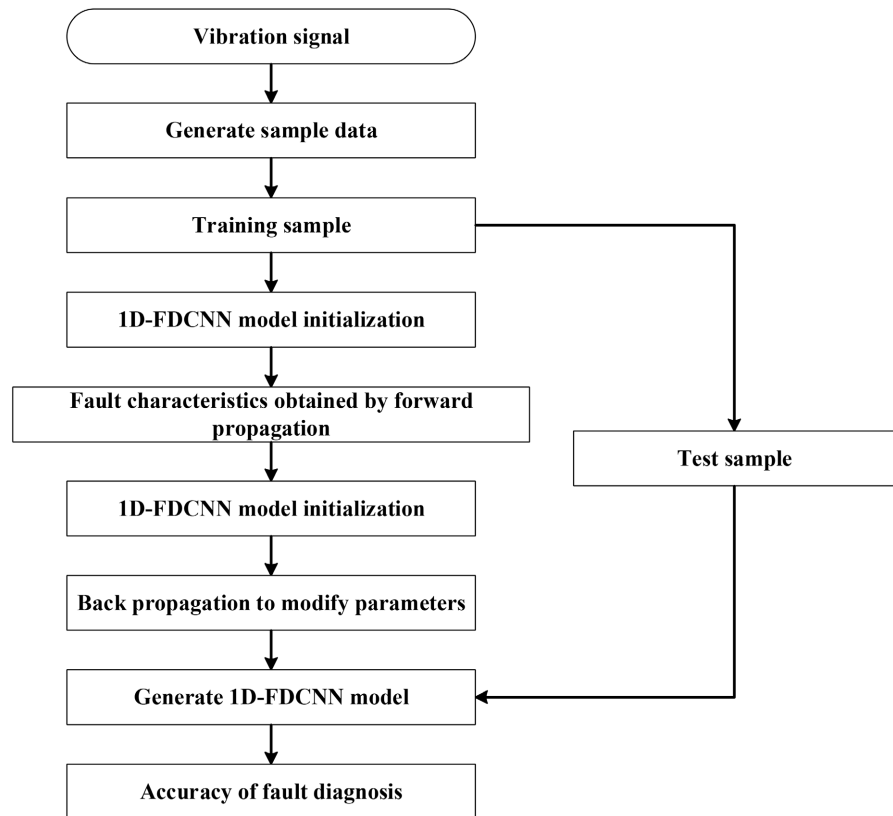


Figure 2. 1D-FDCNN model training process.

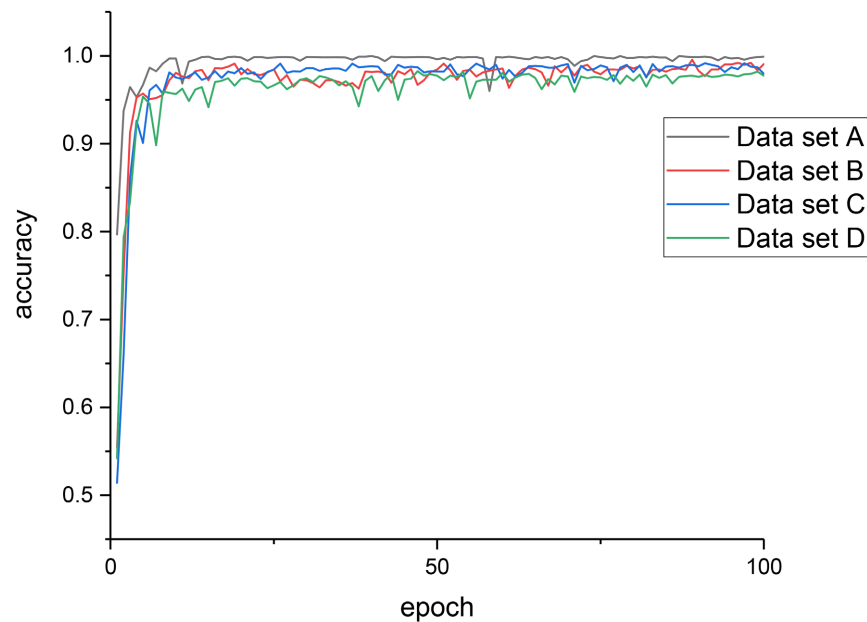


Figure 3. 1D-FDCNN model accuracy curve.

with other algorithms, experiments were conducted on four algorithms, KNN, SVM, LSTM and AlexNet, respectively, and the experimental results are shown in **Table 4** below. From the diagnostic results, the proposed model 1D-FDCNN in this paper has improved diagnostic accuracy compared to the classical AlexNet

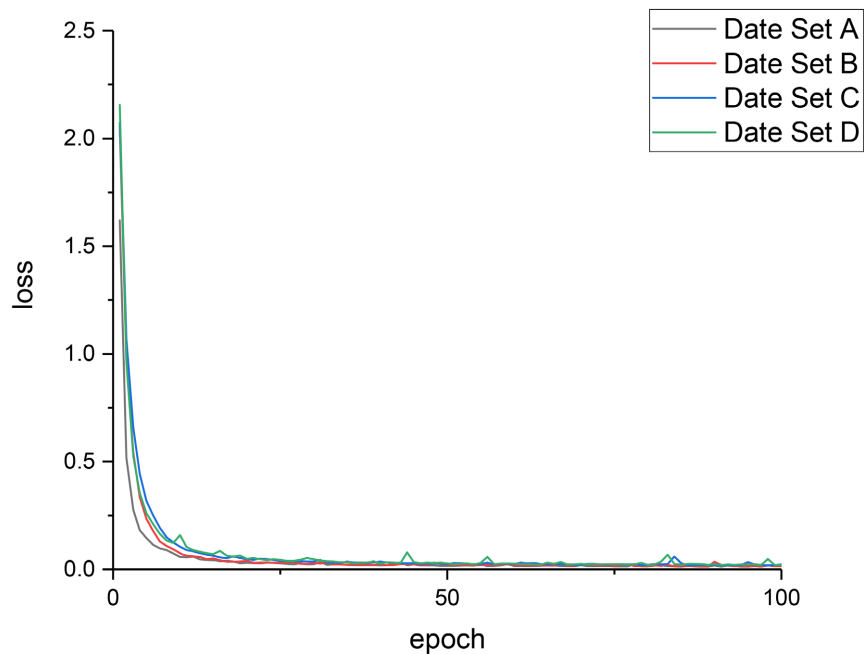


Figure 4. 1D-FDCNN model training loss curve.

Table 4. Experimental results of different algorithms.

Algorithm	Accuracy/%
KNN	74.45
SVM	91.85
LSTM	90.70
AlexNet	96.53
1D-FDCNN	99.81

algorithm, indicating that the improved 1D convolutional neural network model is suitable for time-domain signal processing. Compared with the traditional models KNN, SVM and LSTM algorithms, the improvement in diagnostic accuracy is greater, mainly because these three traditional algorithms have a more limited learning ability for features, and their learned features do not have strong classification characteristics.

5. Conclusion

In this paper, 1D-FDCNN is proposed for bearing fault diagnosis, and the model is used to diagnose vibration signals in various fault states, with experimental results achieving an accuracy rate of up to 99%. The model proposed in this paper has the following advantages: firstly, it can achieve an accurate diagnosis of bearing faults with high robustness; secondly, fault features can be extracted directly from the original vibration signal to realise the end-to-end fault diagnosis process; finally, the training time is relatively short and has good timeliness, which can be applied in engineering.

Conflicts of Interest

The authors declare no conflicts of interest.

References

- [1] Li, W., Chen, J., Li, J. and Xia, K. (2021) Derivative and Enhanced Discrete Analytic Wavelet Algorithm for Rolling Bearing Fault Diagnosis. *Microprocessors and Microsystems*, **82**, Article ID: 103872. <https://doi.org/10.1016/j.micpro.2021.103872>
- [2] Cheng, X.F. and Wang, P. (2020) Fault Diagnosis of Rolling Bearing Based on Time Domain and Frequency Domain Analysis. *Journal of North China University of Science and Technology (Natural Science Edition)*, **42**, 58-64.
- [3] Ma, H.Z., Zhang, Z.D., Shi, W.J., Chen, J.N. and Chen, T.T. (2014) Doubly-Fed Induction Generate Stator Fault Diagnosis Based on Rotor Instantaneous Power Spectrum. *Automation of Electric Power System*, **38**, 30-35.
- [4] Jiang, Z.N., Zhang, Y.S., Feng, K., Hu, M.H. and He, Y. (2019) Gear Fault Diagnosis Method Based on Feature-Enhanced Cepstrum Analysis. *Journal of Mechanical Transmission*, **43**, 13-17.
- [5] Feng, Z.P., Zhao, L.L. and Zhu, F.L. (2013) Amplitude Demodulation Analysis for Fault Diagnosis of Planetary Gearboxes. *Proceedings of the CSEE*, **33**, 107-111.
- [6] Wang, Y.R., Wang, J. and Huang, H.A. (2018) Fault Diagnosis of Planetary Gearboxes Based on NLSTFT Order Tracking under Variable Speed Conditions. *China Mechanical Engineering*, **29**, 1688-1695.
- [7] Liu, Z., Wu, K., Ma, Z. and Ding, Q. (2020) Vibration Analysis of a Rotating Flywheel/Flexible Coupling System with Angular Misalignment and Rubbing Using Smoothed Pseudo Wigner-Ville Distributions. *Journal of Vibration Engineering & Technologies*, **8**, 761-772. <https://doi.org/10.1007/s42417-019-00189-y>
- [8] Zhu, W.Y. and Feng, Z.P. (2016) Fault Diagnosis of Planetary Gearbox Based on Improved Empirical Wavelet Transform. *Chinese Journal of Scientific Instrument*, **37**, 2193-2201.
- [9] Singh, S. and Kumar, N. (2014). Combined Rotor Fault Diagnosis in Rotating Machinery Using Empirical Mode Decomposition. *Journal of Mechanical Science and Technology*, **28**, 4869-4876. <https://doi.org/10.1007/s12206-014-1107-1>
- [10] Lu, D.L., Ning, Q. and Yang X.M. (2018) Fault Diagnosis of Rolling Bearing Based on KNN-Naïve Bayesian Algorithm. *Computer Measurement & Control*, **26**, 21-23.
- [11] Zhao, C.H., Hu, H.X., Chen, B.J., Zhang, Y.N. and Xiao, J.W. (2019) Bearing Fault Diagnosis Based on the Deep Learning Feature Extraction and WOA SVM State Recognition. *Journal of Vibration and Shock*, **38**, 31-37.
- [12] Chen, R.X., Huang, X., Yang, L.X., Tang, B.P. and Yu, T.W. (2018) Rolling Bearing Fault Identification Based on Convolution Neural Network and Discrete Wavelet Transform. *Journal of Vibration Engineering*, **31**, 883-891.
- [13] Di, J. and Wang, L. (2018) Application of Improved Deep Auto-Encoder Network in Rolling Bearing Fault Diagnosis. *Journal of Computer and Communications*, **6**, 41. <https://doi.org/10.4236/jcc.2018.67005>
- [14] Chen, C. and Qi, F. (2019) Review on Development of Convolutional Neural Network and Its Application in Computer Vision. *Computer Science*, **46**, 63-73.
- [15] Zhang, Q., Zhuo, L. and Li, J. (2018) Vehicle Color Recognition Using Multiple-Layer Feature Representations of Lightweight Convolutional Neural Network. *Signal Processing*, **147**, 146-153. <https://doi.org/10.1016/j.sigpro.2018.01.021>

- [16] Yang, H.B. and Gong, W.Y. (2019) Improvement of Back Propagation Algorithm based on Convolution Neural Network. *Computer Engineering and Design*, **40**, 126-130.
- [17] Zhang, B., Zhao, Q. and Feng, W. (2018) AlphaMEX: A Smarter Global Pooling Method for Convolutional Neural Networks. *Neurocomputing*, **321**, 36-48.
<https://doi.org/10.1016/j.neucom.2018.07.079>
- [18] Lu, G., Wang, Y. and Yang, H. (2020) One-Dimensional Convolutional Neural Networks for Acoustic Waste Sorting. *Journal of Cleaner Production*, **271**, Article ID: 122393. <https://doi.org/10.1016/j.jclepro.2020.122393>
- [19] Qu, J.L., Yu, L., Yuan, T., Tian, Y.P. and Gao, F. (2018) Adaptive Fault Diagnosis Algorithm for Rolling Bearings Based on One-Dimensional Convolutional Neural Network. *Chinese Journal of Scientific Instrument*, **39**, 134-143.
- [20] Jin, L.J., Zhan, J.M., Chen, J.H. and Wang, T. (2020) Drill Pipe Fault Diagnosis Method Based on One-Dimensional Convolutional Neural Network. *Journal of Zhejiang University (Engineering Science)*, **54**, 467-474.
- [21] Guo, L., Lei, Y.G., Xing, S.B., Yan, T. and Li, N.P. (2018) Deep Convolutional Transfer Learning Network: A New Method for Intelligent Fault Diagnosis of Machines with Unlabeled Data. *IEEE Transactions on Industrial Electronics*, **66**, 7316-7325.
<https://doi.org/10.1109/TIE.2018.2877090>