



# Gear Tooth Fault Detection Based on Designed Convolutional Neural Networks

Xiaoqiang Du<sup>1</sup>, Yongbo Li<sup>2</sup>(✉), and Shubin Si<sup>1</sup>

<sup>1</sup> School of Mechanical Engineering, Northwestern Polytechnical University, Xian, China

<sup>2</sup> School of Aeronautics, Northwestern Polytechnical University, Xian, China  
yongbo@nwpu.edu.cn

**Abstract.** Gearbox is one of the most important parts of the rotating machinery, so health monitoring of the gearbox is essential. The accurate positioning of tooth failure of gear is an important function of the fault diagnosis system. This paper proposes a detection strategy based on designed convolutional neural networks to detect and locate gear tooth failure. The detection strategy aims to compare the characteristic gap between the normal gear and the faulty gear in the same period extracted by the convolutional neural network, and assign weights to the faulty gear vibration signal to obtain the weight sequence of the faulty vibration signal, so as to obtain the faulty tooth weight. Finally, the health condition of the gear can be evaluated by comparing the weight between all teeth of the gear. The proposed detection strategy is tested through simulation vibration signal and experiment vibration signal. The result shows that the proposed method can successfully identify gear failure and effectively detect single tooth failure on gear.

**Keywords:** Gear tooth · Detection strategy · Designed convolutional neural networks

## 1 Introduction

Gearbox is a vital transmission element, which has widely applied in small, medium and large machines. Excessive load, overspeed and improper operating conditions may cause defects on its surface, which can cause abnormal vibration for the entire machine structure. This may eventually lead to accidental casualties and property damage. Therefore, the fault diagnosis of the gearbox is of great significance to the reliability and safety of the mechanical system [1–4].

The accurate positioning of tooth failure of gear is an important function of the fault diagnosis system. In the past researches on gearbox, many researchers have devoted themselves to the detection of gear tooth failure. Liang et al. [5] proposed a windowing and mapping strategy to interpret the vibration signal of a planetary gear at the tooth level and detect the gear tooth failure by comparing the differences among the signals of all teeth. Xue et al. [6] developed a 20 degree of freedom planetary gear lumped-parameter model to obtain the gear dynamic response and detect the gear failure further. Shahin et al. [7] presented a statistical analysis on a fault index computed based on the stator

current space vector instantaneous frequency on a real-time platform, which can evaluate gear tooth surface damage fault detection using a threshold. Liu et al. [8] proposed a new time-domain fault detection method combining fast dynamic time warping (DTW) and correlated kurtosis techniques to detect and identify damaged planetary gear and its position. Yang et al. [9] proposed a method using an autoregressive model with exogenous variables (ARX) to detect and localize the gear failure for varying load conditions.

All of them use the artificial design features to evaluate the health condition of the gearbox. As we know, feature extraction is the core of the entire fault diagnosis work and determines the effectiveness of the entire diagnosis system. Traditional feature extraction methods use artificial design features that experts design features based on experience. However, due to limited expert experience, artificial design features will inevitably bring many errors.

Therefore, in order to overcome the shortcomings of artificial design features, it's a better choice to use deep learning methods for feature extraction. In recent years, deep learning has been used widely in fault diagnosis [10–12], performance evaluation [13, 14], remaining life prediction [15, 16] and many other aspects. Compared with traditional fault diagnosis methods, deep learning can directly process the original data to achieve adaptive feature learning, which reduces manual participation and helps to improve the adaptability and intelligence of prediction methods. Layer-by-layer feature learning is more likely to learn the basic features hidden in the monitoring data, which improves prediction accuracy. Convolutional neural networks (CNN), as an effective neural network of deep learning, has been widely used in the field of fault diagnosis [17, 18]. CNN can automatically extract the original features of the input signal, which greatly avoids the diagnostic errors caused by human factors. The design operation of convolution and pooling has a non-negligible effect on data feature extraction and dimensionality reduction.

In this paper, this designed CNN has 3 stages including data preprocessing stage, feature extraction stage, and classification stage [19]. The data preprocessing stage is a key part of the implementation of the detection strategy, which mainly includes the symbolization process [20] and the one-hot encoding process [21]. The symbolization process aims to denoise the vibration signal and recombine the sequence in an integer manner. The one-hot encoding process aims to convert the recombination sequence from a 1-dimensional signal to a 2-dimensional signal. The feature extraction stage is through convolutional layers and pooling layers. The final classification stage uses binary classification logistic regression to score the extracted features.

The detection strategy is mainly based on the idea of probability. For a binary classification problem, suppose that the normal signal label is 0 and the abnormal signal label is 1. The higher score of the test signal, the more fault information characteristics the signal has. And for high-score test signals, they must have some special properties that are different from normal signals in certain areas. Therefore, when there are enough high-score signals, by calculating the probability value of the symbol value of each point of the high-score signals, the weight sequence can be characterized by the maximum probability value.

Therefore, by combining the designed CNN and the detection strategy, we proposed a method of gearbox fault region location based on convolutional neural network (CNN).

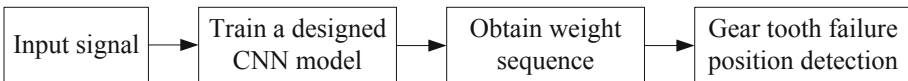
The method mainly consists of three steps. First, train a designed CNN model with dataset. Then, use the detection strategy to obtain the weight sequence of the fault signal. Finally, detect the gear tooth failure by processing the weight sequence. The main contributions of this paper can be summarized as follows.

- (1) A designed CNN structure is proposed by introducing symbolization and one-hot encoding. The designed CNN model can automatically learn to discriminate features directly from the original vibration signal without much prior knowledge about vibration signal feature extraction. And the designed CNN can effectively identify gear fault.
- (2) A detection strategy is proposed to locate the gear tooth failure. The effectiveness is validated by the simulation signal and the experiment signal. The detection result can give certain instructions for failure mechanism.

The focus of this article is gearbox fault area detection based on the designed CNN. The rest comment is as follows. Section 2 gives the designed CNN model and the explanation of fault area detection of gearbox. Section 3 is the validation using the simulation signal. Section 4 is the validation using experiment signal. Finally, conclusions are given in Section 5.

## 2 Methodology

Figure 1 gives the flowchart of the gear tooth failure detection. As we can see from that, firstly, it's essential to train a designed CNN model. Then, obtain the weight sequence of the fault signal using the detection strategy. Finally, detect the position of the gear tooth failure by calculating the weights of all teeth.



**Fig. 1.** Implementation process of proposed methodology

### 2.1 The Designed CNN Model

Figure 2 gives the schematic diagram of the designed CNN model. For training, the scoring CNN model uses a set of sequences. For each sequence, binary class label is adopted. As we can see from Fig. 2, the designed CNN model consists of three stages: data preprocessing stage, feature learning stage, and classification stage. The data preprocessing stage plays a vital role for the tooth failure detection.

For data preprocessing, it contains symbolization step and one-hot encoding step. In this paper, equal probability symbolization is used to symbolize the vibration data.

The idea of equal probability symbolization was first introduced by Lin et al. [22] in 2007. It is widely used in the field of data mining. The equal probability symbolization

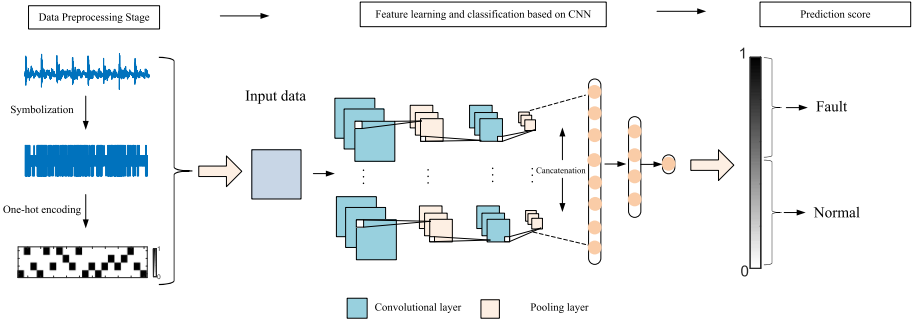


Fig. 2. The designed CNN model structure.

process is briefly described as follows: For time sequence  $\{x_i : 1 \leq i \leq N\}$ , the sequence is sorted by amplitude first. When the number  $n$  of symbols is given,  $n-1$  bisectors (denoted as  $t_1, t_2, \dots, t_{n-1}$ ) can be found as the threshold of symbol division. After that, a new symbolic sequence is obtained according to the rule as follows:

$$s_i = \begin{cases} 0 & (x_i \leq t_1), \\ 1 & (t_1 < x_i \leq t_2), \\ \dots & \\ n - 2 & (t_{n-2} < x_i \leq t_{n-1}), \\ n - 1 & (t_{n-1} < x_i), \end{cases}$$

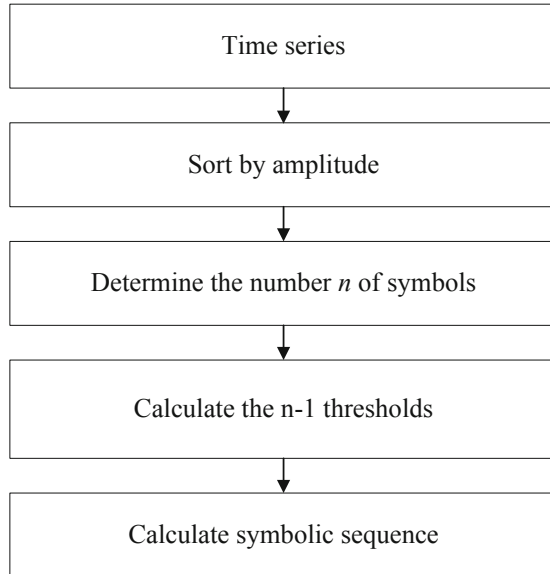


Fig. 3. The flowchart of the equal probability symbolization.

The detailed flowchart of the equal probability symbolization is illustrated in Fig. 3.

For one-hot encoding step, one-hot encoding will assign different values (m types of values) of variables to an m-dimensional vector with only 1 bit of 0. One-hot encoding also plays a role in expanding features to a certain extent. Its values are only 0 and 1, and different types are stored in vertical space.

### 2.2 The Gear Tooth Failure Detection

Section 2.1 has described the process of training a designed CNN model in detail. Then we can detect the gear tooth failure. Figure 4 illustrates the detailed steps about fault area detection of vibration signal. From Fig. 4, we can see that the detection of gear tooth failure can be divided into two parts: the detection strategy and tooth failure detection.

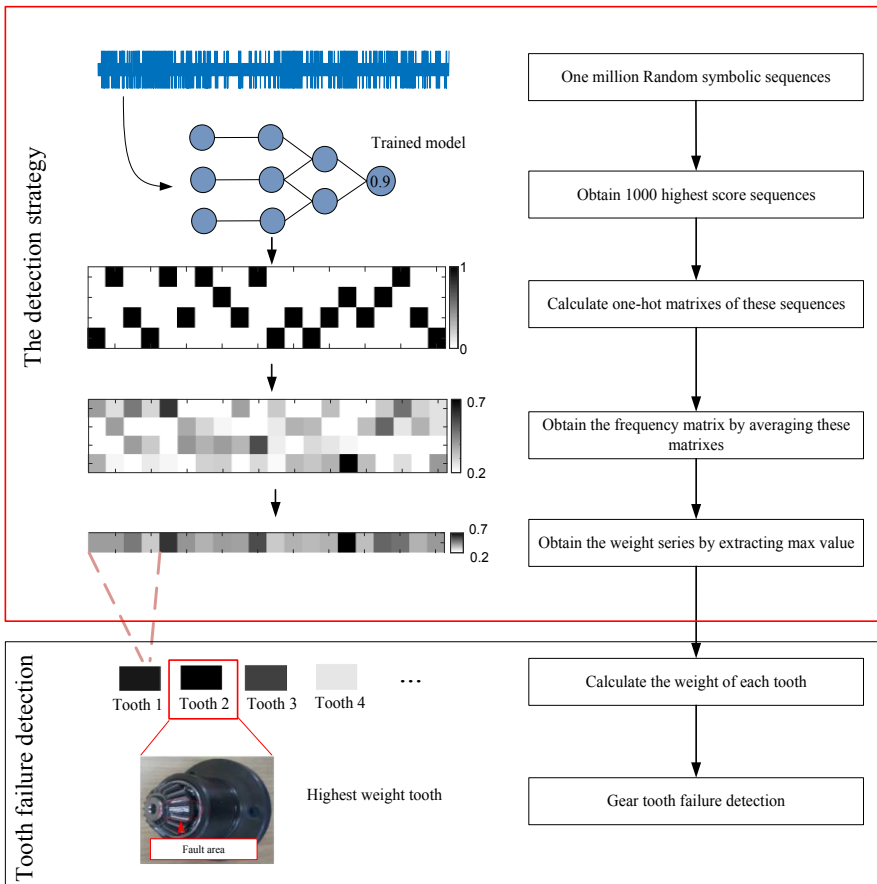


Fig. 4. The flowchart of gear tooth failure detection

### 2.2.1 The Detection Strategy

As we can see from Fig. 4, five steps are required for high score sequences acquisition as follows:

Step 1: Randomly generate 1,000,000 sets of symbol sequences.

Step 2: Using the trained CNN model to test these symbol sequences. Each symbol sequences would generate a score between 0 and 1. Then rank these sequences according to the output score from high to low. Finally, select the top 1000 highest scoring sequences. These high score sequences acquisition aims to provide the sequences which have more fault information.

Step 3: Obtain the one-hot matrixes of these highest scoring sequences by one-hot encoding.

Step 4: For these top 1000 highest-scoring sequences, they have the same fault features in certain areas. This means they must have some same values in certain areas. By adding the 1000 corresponding one-hot matrixes and averaging the final result, each position shows the frequency. When a position of a column have higher value compared with the other values of this column, this indicates this value is more effective and this position can be important. When the value is near to the other values of this column, this demonstrates this point of the vibration signal is not important. Therefore, the average matrix is the weight matrix.

Step 5: Due to the importance for each position of the vibration signal can be reflected by the max value of each position. Therefore, for the weight matrix, we can extract the max weight for each column to obtain the weight series. The weight series shows the importance of each position of the vibration signal for the classification result.

### 2.2.2 Tooth Failure Detection

When we have obtained the weight series for the signal of fault condition. It's easy to calculate the weight of each tooth. To get the weight of each tooth, it's essential to calculate how many points a tooth has according to the related parameters of the experiment data. Then we can get the weight of each tooth by adding these weights for each tooth and each period.

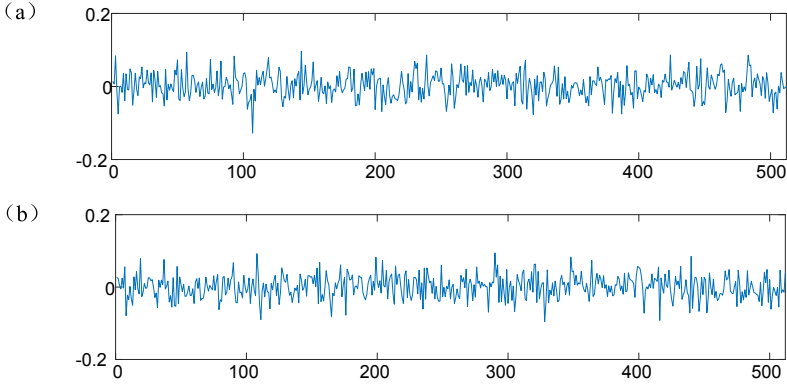
In the end, by normalizing the final result, we can get relative importance of each tooth. By observing the tooth of highest weight, the fault area of gearbox can be found.

## 3 Simulation Validation

In this section, we use the simulation signal to validate the effectiveness of the proposed method.

For dataset, we use a normal gearbox vibration signal and a fault vibration signal. Each fault type number is 2000, and the length of each sample is 512. The training rate is 0.8, which means each condition has 1600 samples to train and 400 samples to test. The simulation vibration signal is shown in Fig. 5.

By symbolizing and one-hot encoding, we can use the dataset to train the designed CNN model. The parameter setting of CNN model is given in Table 1. The learning



**Fig. 5.** Simulation signal: (a) normal (b) fault

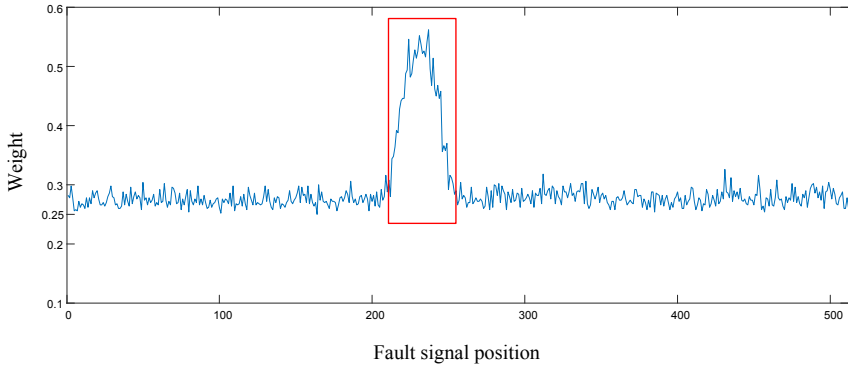
rate is 0.0001, and the epoch is 100. The Adam optimization algorithm [23] is used to minimize the loss function.

**Table 1.** The adopted CNN structure

Layer Type	Number of Filter	Size of the feature map	Size of kernel	Number of Stride	Activation function
Input layer	–	$4 \times 512 \times 1$	–	–	–
C1	12	$1 \times 488 \times 12$	$4 \times 25$	$1 \times 1$	ReLU
P1	–	$1 \times 244 \times 12$	$1 \times 2$	$1 \times 2$	–
C2	24	$1 \times 244 \times 12$	$1 \times 2$	$1 \times 1$	ReLU
P2	–	$1 \times 244 \times 12$	$1 \times 2$	$1 \times 2$	–
F1	–	200	–	–	ReLU
F2	–	40	–	–	ReLU
Output layer	–	1	–	–	Sigmoid

For the testing result, the testing accuracy of 100% after 20 epochs almost, which means we get a very effective model to distinguish these two fault types. Then, we can randomly generate 1,000,000 symbolic sequences to test the trained model. Each sequence’s length is the same as the sample of the dataset and each vibration point’s value is between 1 and 4. Finally, we get the average one-hot matrix by averaging the top 1000 highest score samples’ one-hot matrixes.

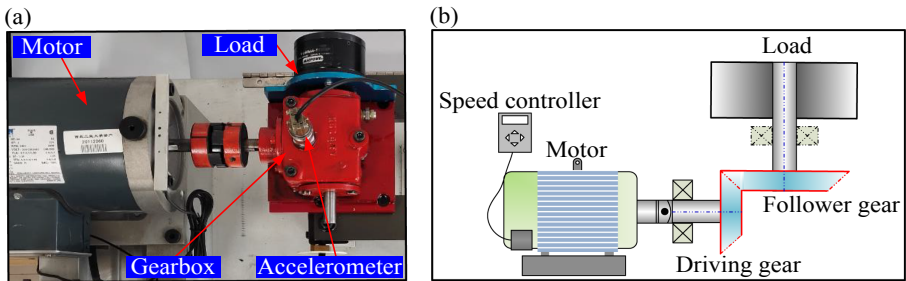
According to the average matrix, we can get the weight series by extracting the biggest value for each column. The weight series means the weight of each position in the diagnosis process, which is shown in Fig. 6. The high weight of the fault signal position shows which part of the fault signal is more important than others for the diagnosis result.



**Fig. 6.** The weight series for fault condition

## 4 Experiment Validation

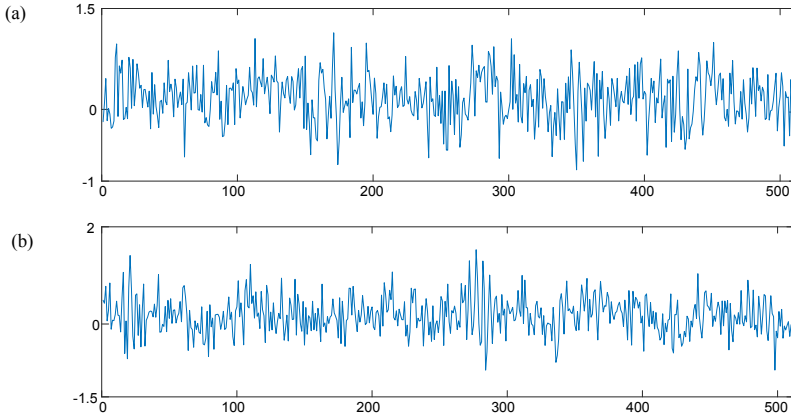
In this section, the gear fault experiments were conducted on rotating machinery called SpectraQuest Machinery Fault Simulator (MFS). The test rig is shown in Fig. 7. An accelerometer was installed on the top of the gearbox to collect the vibration signals. The sampling frequency was 12800 Hz and the rotating speed was 3000 rpm.



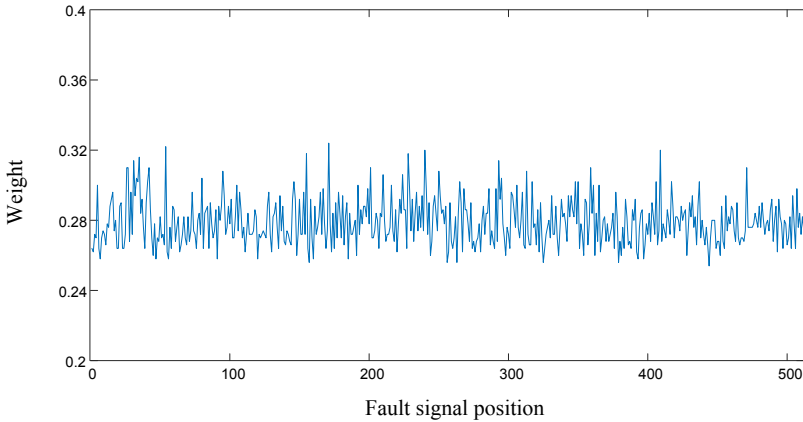
**Fig. 7.** (a) The machinery fault simulator system; (b) the layout of the test rig.

The proposed method is used for binary problem. In the experiment, 2 conditions are introduced including normal condition and missing tooth condition (fault condition). For the dataset, each condition has 480 samples and the length of each sample is 512. The length contains 2 period for the vibration signal. The training rate is 0.8. So, the total number of training samples is 768 and the testing number is 192. An example of dataset is given in Fig. 8.

By symbolizing and one-hot encoding, we can use the dataset to train a designed CNN model. The testing accuracy keeps 100% after almost epoch 20, which means the model is effective to distinguish the fault. Then, by testing 1,000,000 random sequence, we can get the wight matrix. After that, a weight series is obtained by extracting the biggest value for each column, which is shown in Fig. 9.



**Fig. 8.** The experiment vibration signal: (a) normal (b) fault



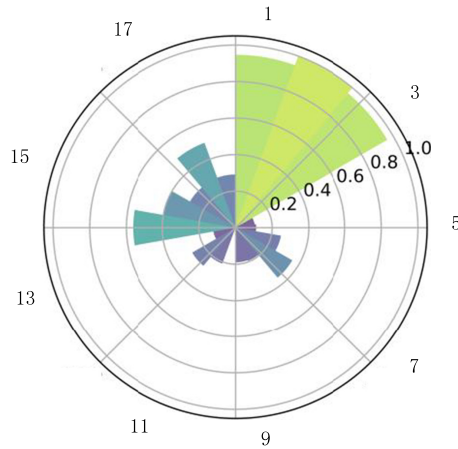
**Fig. 9.** The weight series of the fault condition

As we can see from Fig. 9, it's hard to see which part is more important for the fault condition directly. Because experiment data has more noise impact. Therefore, it's necessary to add the weights corresponding to each tooth. The experiment gear has 18 teeth. The points for each period can be calculated according to the sampling frequency and the rotating speed, the equation is given in Eq. (2).

$$Points\ for\ one\ circle\ of\ the\ gear = \frac{sampling\ frequency \times 60}{rotating\ speed} \tag{2}$$

It's 256 points by calculate the equation. Then we can calculate each gear has 14 points almost. Therefore, each tooth's weight can be gotten by accumulating the weight sum of each tooth and each period, which is shown in Fig. 10. To avoid the randomness of experiment, the experiment is implemented 20 times to get the final result.

From the result, we can find the tooth 1, tooth 2 and tooth 3 obviously have higher weights than other teeth. And the tooth 2 has the biggest weight. The result shows that



**Fig. 10.** The weights of each tooth

the proposed method can detect the fault tooth. And the fault tooth and two teeth around are obviously more important than other teeth for the diagnosis result.

## 5 Conclusion

In this paper, we proposed a method for gear tooth failure detection based on the designed CNN model. And the effectiveness of the proposed method is validated by the experiment signal. The validation result shows that the proposed method can effectively detect the fault area of the vibration signal and the fault tooth. The fault area detection can give instructions on failure mechanism and subsequent maintenance, which deserves research in future.

## References

1. Lin, J., Zuo, M.J.: Gearbox fault diagnosis using adaptive wavelet filter. *Mech. Syst. Signal Process.* **17**(6), 1259–1269 (2003)
2. Liu, B., Riemenschneider, S., Xu, Y.: Gearbox fault diagnosis using empirical mode decomposition and Hilbert spectrum. *Mech. Syst. Signal Process.* **20**(3), 718–734 (2006)
3. Lei, Y., Han, D., Lin, J., et al.: Planetary gearbox fault diagnosis using an adaptive stochastic resonance method. *Mech. Syst. Signal Process.* **38**(1), 113–124 (2013)
4. Li, C., Sanchez, R.V., Zurita, G., et al.: Gearbox fault diagnosis based on deep random forest fusion of acoustic and vibratory signals. *Mech. Syst. Signal Process.* **76**, 283–293 (2016)
5. Liang, X., Zuo, M.J., Liu, L.: A windowing and mapping strategy for gear tooth fault detection of a planetary gearbox. *Mech. Syst. Signal Process.* **80**, 445–459 (2016)
6. Xue, S., Howard, I.: Torsional vibration signal analysis as a diagnostic tool for planetary gear fault detection. *Mech. Syst. Signal Process.* **100**, 706–728 (2018)
7. Kia, S.H., Henao, H., Capolino, G.: Fault index statistical study for gear fault detection using stator current space vector analysis. *IEEE Trans. Ind. Appl.* **52**(6), 4781–4788 (2016). <https://doi.org/10.1109/TIA.2016.2600596>

8. Hong, L., Dhupia, J.S.: A time-domain fault detection method based on an electrical machine stator current measurement for planetary gear-sets. In: 2013 IEEE/ASME International Conference on Advanced Intelligent Mechatronics, Wollongong, NSW, pp. 1631–1636 (2013). <https://doi.org/10.1109/AIM.2013.6584330>
9. Yang, M., Makis, V.: ARX model-based gearbox fault detection and localization under varying load conditions. *J. Sound Vib.* **329**(24), 5209–5221 (2010)
10. 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. *Sensors* **17**(2), 425 (2017)
11. He, M., He, D.: Deep learning based approach for bearing fault diagnosis. *IEEE Trans. Ind. Appl.* **53**(3), 3057–3065 (2017)
12. Jia, F., Lei, Y., Guo, L., et al.: A neural network constructed by deep learning technique and its application to intelligent fault diagnosis of machines. *Neurocomputing* **272**, 619–628 (2018)
13. Lyu, W., Zhang, Z., Jiao, C., et al.: Performance evaluation of channel decoding with deep neural networks. In: 2018 IEEE International Conference on Communications (ICC), pp. 1–6. IEEE (2018)
14. Duan, Y., Lv, Y., Wang, F.Y.: Performance evaluation of the deep learning approach for traffic flow prediction at different times. In: 2016 IEEE International Conference on Service Operations and Logistics, and Informatics (SOLI), pp. 223–227. IEEE (2016)
15. Zhang, J., Wang, P., Yan, R., et al.: Deep learning for improved system remaining life prediction. *Procedia CIRP* **72**, 1033–1038 (2018)
16. Ren, L., Zhao, L., Hong, S., et al.: Remaining useful life prediction for lithium-ion battery: a deep learning approach. *IEEE Access* **6**, 50587–50598 (2018)
17. Xia, M., Li, T., Xu, L., et al.: Fault diagnosis for rotating machinery using multiple sensors and convolutional neural networks. *IEEE/ASME Trans. Mechatron.* **23**(1), 101–110 (2017)
18. Chen, Z.Q., Li, C., Sanchez, R.V.: Gearbox fault identification and classification with convolutional neural networks. *Shock Vibr.* **2015**, 1–11 (2015)
19. Jiang, G., He, H., Yan, J., et al.: Multiscale convolutional neural networks for fault diagnosis of wind turbine gearbox. *IEEE Trans. Industr. Electron.* **66**(4), 3196–3207 (2018)
20. Hu, R.F., Wang, L., Mei, X.Q., et al.: Fault diagnosis based on sequential pattern mining. *Comput. Integr. Manuf. Syst.* **16**(7), 1412–1418 (2010)
21. Cassel, M., Lima, F.: Evaluating one-hot encoding finite state machines for SEU reliability in SRAM-based FPGAs. In: 12th IEEE International On-Line Testing Symposium (IOLTS 2006), pp. 6. IEEE (2006)
22. Lin, J., Keogh, E., Lonardi, S., et al.: A symbolic representation of time series, with implications for streaming algorithms. In: Proceedings of the 8th ACM SIGMOD Workshop on Research Issues in Data Mining and Knowledge Discovery, pp. 2–11 (2003)
23. Kingma, D., Ba, J.: Adam: a method for stochastic optimization. In: Proceedings of 3rd International Conference on Learning Representations (2015)