Multi-view Bearing Fault Diagnosis Method Based on Deep Learning

Wang Ruihan¹, Meng Xinyu¹, Xiong Bangru¹, Wang Zhengxia¹,²,*

¹ College of Information Science and Engineering, Chongqing Jiaotong University, No.66, Xuefu Avenue, Nan'an District, Chongqing, 400074, China

² College of Computer Science and Cyberspace Security, Hainan University, No.58, Renmin Avenue, Meilan District, Haikou, Hainan, 570228, China

*zxiawang@163.com

Abstract: In recent years, the use of Time-frequency analysis to generate Time-frequency diagrams of vibration signals, and then the use of deep learning methods to classify them has become one of the mainstream methods for bearing fault diagnosis. However, a single Time-frequency analysis method usually cannot extract a complete vibration signal, so the accuracy of the model will be affected to a certain extent. Therefore, this paper proposes a multi-view bearing fault diagnosis method based on deep learning. The same segment of vibration signal is generated by different Time-frequency analysis methods, and then the pre-trained network is used to train the model. The Flatten layer was replaced by the Global Max Pooling (GMP) layer before the layer. The experimental results show that compared with the traditional feature fusion method, the method in this paper can not only achieve better accuracy, but also has stronger generalization.

Keywords: Deep learning, Bearing fault diagnosis, Time-frequency analysis, Global Max Pooling

1. Introduction

Bearings are one of the important rotating machinery parts in the industry. When failures occur and go undetected, the stability and life of the equipment will be greatly affected. Therefore, the identification of bearing failures has great research significance. Time-frequency analysis is currently one of the most common bearing fault identification methods.
In recent years, Deep learning has become a common method for bearing fault classification. The original vibration signal is transformed into a Time-frequency diagram by Time-frequency analysis method, and then fault on the generated image is identified through a Convolutional Neural Network [1]. However, a single Time-frequency diagram generated from a single sensor often cannot obtain complete information of vibration signals in different frequency bands, which affects the accuracy of fault diagnosis. The method of using feature fusion can extract more information from different sensors [2], but the increase in sensors and the high number of features will inevitably increase the cost and time of fault diagnosis. Therefore, under the premise of the original single sensor signal, improving the accuracy of fault recognition has become one of the goals of current research.

In response to the above problems, this paper proposes a multi-view bearing fault diagnosis method based on deep learning. This paper has the following innovations.

1) Use different Time-frequency analysis methods to perform Time-frequency analysis on the same sensor, and then perform feature fusion on the obtained Time-frequency diagram. The method of different Time-frequency analysis can more accurately extract the Time-frequency analysis characteristics of the signal in different frequency bands, so as to obtain higher accuracy.

2) Fine-tune the pre-trained weighted VGG16 model to process the Time-frequency graph separately. Using the transfer learning method can reduce the time of model retraining and reduce the number of model iterations.

3) The Global Max Pooling (GMP) layer is used to replace the traditional Flatten layer, which greatly reduces the number of features, reduces the training time and reduces the model's overfitting problem.

2. Theory Introduction

2.1. VGG16 network
Convolutional Neural Network (CNN) is the most common method in the field of deep learning, which can extract useful corresponding features from image information, so it is widely used in image recognition, detection and other fields. The VGG16 [3] network is one of the most commonly used CNN networks. The network contains a total of 16 layers. The VGG16 network connects multiple convolution cores in series and has more than a single convolution layer. The nonlinear transformation, so it can have a higher accuracy.

2.2. Global Max Pooling
Traditional CNN usually use the Flatten layer to convert multi-dimensional data into one-dimensional data, so it is usually used for the transition from convolutional layer and pooling layer to fully connected layer. However, this method involves too many features, which leads to the characteristics of excessive model scale, slow running speed, and easy overfitting. Therefore, using a global pooling layer to replace the traditional Flatten layer has become one of the current main methods. The global pooling layer calculates all pixel values of a feature channel to obtain the corresponding average or maximum value, and uses this value to represent this feature channel. Using the global pooling layer [4] can reduce the number of parameters, the amount of calculation and the ability to overfit.
The global pooling layer includes a Global Average Pooling (GAP) layer and a Global Max Pooling (GMP) layer. The GAP layer is more suitable for extracting global features in an image, and the GMP layer is more suitable for extracting local features in an image. In the Time-frequency diagram of bearing faults, those fault amplitudes are more likely to be reflected in the local picture features. Therefore, the accuracy of handling bearing faults with the GMP layer is higher.

3. Model introduction

In this section, we mainly introduce the multi-view bearing fault diagnosis model based on deep learning. The detailed model is shown in Fig.1.

3.1. Data preprocessing and Time-frequency analysis

The signal of the sensor is divided into equal length parts, and then these signals are respectively subjected to Short-Time Fourier Transform (STFT) and Continuous Wavelet Transform (CWT) to obtain their corresponding Time-frequency diagrams. Both STFT and CWT are common Time-frequency analysis methods. STFT can extract high-frequency information of vibration signals, while CWT can extract high-frequency information. Feature extraction methods from multiple perspectives can improve the complementarity of feature extraction and improve model performance accuracy.

3.2. Multi-view fusion network based on VGG16

Input the Time-frequency map obtained by Time-frequency analysis as different features into the VGG16 migration network layer. This layer uses the weight value of the VGG16 network pre-trained on the ImageNet data set [5] as the initial network weight, and the pre-trained weight value can significantly improve the convergence of the model and reduce the time of model retraining.

The results obtained by migrating the network layer of the two Time-frequency graphs through VGG16 are respectively input into the two GMP layers to replace the Flatten layer. Using GMP can not only reduce the number of features in the model, but also extract local fault features in the Time-frequency diagram. After passing through the two GMP layers, the Concatenate layer is used to splice the obtained feature data to superimpose the obtained information. Because the key information obtained by different Time-frequency analysis is different and has stronger complementarity, the accuracy of the model is higher. After Concatenate, it passes through a fully connected layer thickness and performs fault classification through the SoftMax layer to obtain the final result.

4. Experimental
In this section, experiments are conducted using public bearing vibration data sets, and the superiority of the model is verified by comparison with other commonly used methods.

4.1. Data set introduction

The experiment uses the bearing vibration data set published by the Universität Paderborn [6] for model training and testing. This data set simulates the vibration signal data of the bearing under health and various failures. It contains four operating states, namely Healthy, Electric Discharge Machining (EDM) fault, Drilling fault and Electric Engraver fault, and its sampling frequency is 64000 Hz. The data set contains the vibration signal data of the same bearing under different working conditions. These different working conditions are marked as A, B and C. The details of the different working conditions are shown in Tab.1.

| Working condition | Rotating speed /rpm | Load Torque /Nm | Radial force /N | Sampling frequency /Hz |
|-------------------|---------------------|-----------------|-----------------|------------------------|
| A                 | 1500                | 0.1             | 1000            | 64000                  |
| B                 | 1500                | 0.7             | 400             | 64000                  |
| C                 | 1500                | 0.7             | 1000            | 64000                  |

Tab.1 Detailed information of different working conditions

Experimental results

In this section, the superiority of this model is verified by comparing with several common deep learning networks. Tab.2 shows the accuracy of the model under different working conditions, where "A+B" represents the result obtained by generating the corresponding Time-frequency diagram from the original data through the A method, and then classifying it through the B deep learning network. The results in the table are the average values of ten cross-validation experiments. It can be seen from the table that compared with other models, the accuracy of this model is close to 100%, and it has a good accuracy.

| Method                | Average accuracy rate under different working conditions /% | A   | B   | C   | average |
|-----------------------|----------------------------------------------------------|-----|-----|-----|---------|
| FFT-DNN               | 76.66 90.06 89.12                                        | 85.28 |
| WDCNN                 | 89.17 95.01 96.35                                        | 93.51 |
| TICNN                 | 96.19 96.31 98.81                                        | 97.10 |
| CWT+VGG16             | 99.94 99.85 100.00                                       | 99.93 |
| Feature Fusion+VGG16  | 99.94 99.94 100.00                                       | 99.96 |
| Proposed              | 99.96 99.96 100.00                                       | 99.97 |

Tab.2 Results of different methods

In order to further verify the superiority of this method, the experiment uses several common bearing fault diagnosis methods, and puts the model trained in one working condition and tested under other working conditions to compare its generality under different working conditions. Tab.3 shows
the results of the comparative experiment, in which "A→B" in the table represents the results obtained after training under A working condition, testing under B working condition.

It can be seen from Table 3 that using the feature fusion method and the GMP layer can further improve the average accuracy of the model. Therefore, this model can have stronger generalization.

| Method                      | A→B | A→C | B→A | B→C | C→A | C→B | average |
|-----------------------------|-----|-----|-----|-----|-----|-----|--------|
| STFT+VGG16                  | 78.56 | 69.35 | 72.79 | 99.93 | 80.89 | 83.59 |
| CWT+VGG16                  | 77.67 | 99.97 | 80.28 | 89.47 | 99.86 | 75.53 | 87.13 |
| Feature fusion (FC)         | 85.31 | 100.00 | 86.15 | 85.78 | 98.72 | 73.06 | 88.17 |
| Feature fusion (GAP)        | 79.19 | 100.00 | 94.42 | 99.86 | 99.86 | 81.97 | 90.53 |
| Proposed                   | 88.90 | 100.00 | 91.44 | 94.55 | 98.90 | 100.00 | 93.57 |

**Tab.3** Result under different working conditions

5. Conclusion

This paper proposes a multi-view bearing fault diagnosis method based on deep learning, which improves the complementarity of Time-frequency analysis in different frequency domains. We first extract features from different perspectives of data, and then fuse these features. The proposed method extracts the local features of the Time-frequency map through the GMP layer to reduce the training time of the model, and shows the good accuracy and excellent generalization through experiment results.

Acknowledgements

This work is partially supported by the Science Project of Hainan University (KYQD(ZR)20021, KYQD(ZR)20022), Natural Science Foundation Project of CQCSTC (No. cstc2018jcyj AX0398)

References

[1] Zhiyu Z, Gaoliang P, Yuanhang C and Huijun G 2019 A convolutional neural network based on a capsule network with strong generalization for bearing fault diagnosis J. Neu. **323** 62-75.

[2] Xingwei X, Zhengrui T, Weivei M, Qinglong A and Ming C 2020 Intelligent monitoring and diagnostics using a novel integrated model based on deep learning and multi-sensor feature fusion. *D. Mea.* **165** 108086

[3] Karen S and Zisserman A 2014 Very Deep Convolutional Networks for Large-Scale Image Recognition. Preprint arxiv.org/abs/1409.1556

[4] Min L, Qiang C and Shuicheng Y 2013 Network In Network. Preprint arxiv.org/abs/1312.4400

[5] Jia D, Wei D, Richard S, Li-Jia L, Kai L and Fei-Fei L 2009 ImageNet: A large-scale hierarchical image database. 2009 IEEE conference on computer vision and pattern recognition pp 248-255

[6] Christian L, James Kuria K, Detmar Z and Walter S 2016 Condition Monitoring of Bearing Damage in Electromechanical Drive Systems by Using Motor Current Signals of Electric Motors: A Benchmark Data Set for Data-Driven Classification *Proceedings of the European conference of the prognostics and health management society* pp 05-08