Gearbox fault diagnosis based on transfer learning with RseNet50 model

Chen He¹, Yiqiang Jiang¹, Li Sun¹, Bin Wu*¹
¹School of Business, Shanghai Dianji University, Shanghai, China
Corresponding author’s email: wub@sdju.edu.cn,

Abstract: The fault diagnosis method based on ResNet50 convolutional neural network and migration learning is proposed as a way of improving the gearbox fault diagnosis. Firstly, exporting the normal and faulty data of the gearbox. Then converting the exported data into one-dimensional images to and generate the corresponding training and test sets, train in the model to get accuracy of the test set and training set. After fine adjustment, it is used for gearbox fault diagnosis, compared with the VGG16, ResNet101, and GoogleNet model, the accuracy of ResNet50 is above 86.6%. It has a good prospect of application, and its effect is obviously better than that of other models.

1. Introduction

As the core component of various wind power equipment, gearbox is prone to failure[1]. Especially, wind turbines are basically installed outdoors, and sometimes especially bad weather occurs. Once the gearbox fails[2], if it is not eliminated in time, after the power system is subjected to such a large disturbance, there may also be dramatic changes in parameters and oscillations, which will affect the stability of the power system[3]. The cost of fault maintenance in wind power generation[4] is very high, so it is necessary to do a good job of fault diagnosis and prediction. However, the traditional fault diagnosis of gearbox cannot meet our safety, stability and reliability requirements for wind turbines. With the development of this sensor and the way of transmitting data[5], the data of various operation modes of the gearbox can be saved, which facilitates our judgment and maintenance of faults. With the development of machine learning[5], various machine learning has been used for gearbox fault diagnosis[7] and other fault diagnosis, such as convolutional neural network, VGG16[8], GoogleNet[9], ResNet101[10] models in transfer learning[11]. Support vector machine[11] and logistics regression are the problems existing in traditional machine learning methods: 1. Need a lot of pictures or data input for training, and generally the training set is too large and takes a long time, and the hardware requirements are too high. 2. The accuracy of the trained model for fault diagnosis is too low. Aiming at the above two problems, Chen et al. proposed a planetary gear fault diagnosis method based on Deep Belief Network(DBN) migration learning[13]. Tao et al.[14] proposed an improved diagnostic method for transfer learning: the two-step transfer learning method, which uses DCGAN to make auxiliary data sets. First, transfer learning is carried out on the auxiliary data set, and then the network is placed on the target data set for transfer learning training.

This paper proposes a transfer learning method based on ResNet50 convolutional neural network for gearbox fault diagnosis. The RseNet50 network pre-trained in the source domain is applied to the gearbox
fault data set to train its own model. Firstly, the original one-dimensional vibration signal data is normalized and transformed into (224*224) gray image by signal conversion processing method. After processing, the one-dimensional gray image is transformed into a three-channel image. The data set and the training set are the divided in a ratio of 9:1. After fine-tuning, the fault classification accuracy of the model is verified through the test set.

2. Materials and Methods

2.1 Signal data to image method

It is difficult to directly train the signal data directly in the convolutional neural network, so we deal with the original number of signals Transformed into gray image pictures \[15\]for processing. Usually the signal data collected by the sensor is a time domain signal, which is stored in a one-dimensional matrix format. According to the size of the data set, the image is transformed into a N*N pixel format. Firstly, the signal data are grouped and 1966860 data are divided into a group of 50176 data to form a gray image. Normalization of data by range change due to negative data\[16\]:

\[
Y = \frac{X - \text{min}(X)}{\text{max}(X) - \text{min}(X)},
\]

(1)

Round the normalized data:

\[
L = \text{round}(Y),
\]

(2)

Figure 1. Data signal to gray image

2.2. Convolutional neural network and transfer learning model

2.2.1 Convolutional Neural Network

In 1996, Professor Yann Lecun of New York University proposed a convolutional neural network\[17\], which is essentially a multilayer perceptron. The basic structure of convolutional neural network mainly includes input layer, convolutional layer, pooling layer, standard layer(Batch Normalization) \[18\], fully connected layer, Softmax layer \[18\]. The input layer is the input of the entire neural network. In the convolutional neural network for processes images, it generally represents the pixel matrix of a picture. The convolutional layer performs a more detailed analysis of the features of each small block of the picture in the neural network, so that it can get more abstract and simple features, when training, the general convolution layer is set to 3*3 or 5*5. Pooling layer can effectively reduce the size of the parameter matrix, thereby reducing the number of parameters in the final connected layer, so adding pooling layer can speed up the calculation speed and prevent overfitting. The BN layer can accelerate the convergence speed of the network, and the principle is generally added behind the convolution layer as
follows:

\[
\mu = \frac{1}{m} \sum_{i=1}^{m} x_i \tag{3}
\]

\[
\sigma^2 = \frac{1}{m} \sum_{i=1}^{m} (x_i - \mu)^2 \tag{4}
\]

\[
\hat{x}_i = \frac{x_i - \mu}{\sqrt{\sigma^2 + \varepsilon}} \tag{5}
\]

\[
y_i = \gamma \hat{x}_i + \beta = BN_{\gamma, \beta}(x_i) \tag{6}
\]

The fully connection layer mainly acts as a classifier in the convolutional neural network. The local connection and weight sharing method adopted by convolutional neural network: on the one hand, it reduces the number of weights to make the network easy to optimize, on the other hand, it reduces the complexity of the model, that is, it reduces the risk of overfitting. This advantage is more obvious when the input of the network is the image, which makes the image directly as the input of the network and avoids the complex process of feature extraction and data reconstruction in the traditional recognition algorithm. It has good robustness and computational efficiency in dealing with the problem of one-dimensional image, especially in the application of recognition displacement, scaling and other forms of distortion invariance.

2.2.2 ResNet50 Convolutional Neural Network Model

ResNet50 is a residual model proposed by He Kaiming\(^{[10]}\) and others in 2015. Among many ResNet models, ResNet50 is selected to perform well in image processing speed and accuracy. Even if ResNet101 has higher accuracy, it sacrifices the training speed, and the accuracy is lower than ResNet50 for this training. Therefore, this paper finally selects ResNet50, and the most important is that ResNet uses the residual structure\(^{[11]}\) model as shown in Figure 2. Firstly, input a convolution, and then after \(3 + 4 + 6 + 3 = 16\) building blocks, each block has 3 layers, so there are \(16 \times 3 = 48\) layers, and finally there is a fc layer (for classification), so \(1 + 48 + 1 = 50\) layers, so it is a 50 layer network. The ResNet50 convolutional neural network structure is shown in Figure 3.

Output size calculation formula:

\[
L = \frac{a + 2d - b}{c} + 1 \tag{7}
\]

In the formula, \(L\) represents output size, \(a\) represents input size, \(d\) represents padding, \(b\) represents
kernel_size, and c represents stride.

2.2.3 ResNet50 transfer learning model
One of the many machine learning methods of transfer learning is to train a type or a model at first, and then reuse it in the process of developing another task model. In the traditional classification learning, a large number of experimental samples and a good classification model such as support vector machine and logistics regression are needed, but these two conditions are often unable to meet at the same time. It takes a lot of time to collect and label and classify large amounts of data. Transfer learning can well avoid the problem of insufficient data and a lot of time for training.

As hardware continues to upgrade, we can make the original very shallow network deepening, but this approach followed by such a problem, the deep training effect is not as good as the shallow network, this obstacle is mainly due to the gradient explosion, which hinders convergence from the beginning. When deeper networks begin to converge, the degradation problem is exposed: as the depth of the network deepens, accuracy becomes saturated for a long time, and then decreases rapidly. The deep residual learning framework can well avoid the degradation problem. This paper introduces a ResNet50 deep learning framework. The structure is shown in Figure 3. Firstly, the input image size is 3 × 244 × 224, that is, 3 channels, and the image size is 244 × 224. Entering the first convolution layer, the convolution kernel size is 7 × 7, the number of convolution kernels is 64, the step size is 2, and padding is 3. Using the (7) formula, the input is calculated to be 64 × 112 × 112 at the maxpool layer. After a series of residual structures such as conv2, conv3, conv4 and conv5, 11 probabilities are output, and the most possible fault type is obtained by softmax layer.
3. Results and Discussion

3.1 Data collection and processing
In order to verify the effectiveness of ResNet50, the normal data and the data of gearbox in various faults are collected by the gearboxes experimental platform for fault diagnosis. The experiment platform includes gearbox, drive motors, and brakes as shown in Figure 3. When a fault occurs, the acceleration sensor collects the signal installed on the gearbox. The health status of the gearbox is divided into normal, dented teeth, varying degrees of tooth wear, tooth cracks and other three cases of bearing defects. The time domain signal of the data is converted to a gray image as shown in Figure 4. The rotational speed and the training and testing set of various faults are shown in Table 1.
Table 1 Data set description and distribution

| gearbox status           | speed (r/min) | Training set | Validation set |
|-------------------------|---------------|--------------|----------------|
| normal                  | 294,588,882   | 49           | 4              |
| tooth wear 1            | 294,588,882   | 49           | 4              |
| tooth wear 2            | 294,588,882   | 49           | 4              |
| tooth wear 3            | 294,588,882   | 49           | 4              |
| tooth wear 4            | 294,588,882   | 49           | 4              |
| tooth wear 5            | 294,588,882   | 49           | 4              |
| tooth pitting           | 294,588,882   | 49           | 4              |
| tooth fracture          | 294,588,882   | 49           | 4              |
| bearing inner ring defect | 294,588,882 | 49           | 4              |
| ball roller defect      | 294,588,882   | 49           | 4              |
| cage defect             | 294,588,882   | 49           | 4              |

3.2 Data preprocessing
As shown in Figure 5, the healthy state of the gearbox is divided into normal, dented teeth, tooth wear of different degrees, and a large number of data sets and processing procedures of bearing defects in tooth cracks and other three cases. After transforming the vibration signal into a gray image, ResNet50 is input into an image with a format of 224*224*3. The gray image on the right side of Figure 5 is the gray image of partial faults, and then the gray image is processed into a three-channel image for input.
3.3 Experiments and results

In the application process of the model, the image is first input into it, and the epochs are adjusted to make the success rate of the training set achieve the best preservation of its model. Finally, epochs=20 and batch_size=16, and the accuracy of the test set is determined after each training. The accuracy comparison of ResNet101, VGG16, GoogleNet and ResNet50 is shown in Table 2.

Figure 5 Data preprocessing
Table 2 Gearbox fault diagnosis accuracy

| fault type | 1th | 2th | 3th | 4th | 5th |
|------------|-----|-----|-----|-----|-----|
| C1         | 92.86 | 43.75 | 75.00 | 73.9 | 22.79 |
| C2         | 85.71 | 40.62 | 67.85 | 75.6 | 25.24 |
| C3         | 82.14 | 28.12 | 82.14 | 70.6 | 24.69 |
| C4         | 84.38 | 41.25 | 81.25 | 75.0 | 19.60 |
| C5         | 85.73 | 28.12 | 78.57 | 65.6 | 18.66 |
| C6         | 96.43 | 33.48 | 82.16 | 72.6 | 25.09 |
| C7         | 85.28 | 42.79 | 71.88 | 66.7 | 18.92 |
| C8         | 84.64 | 46.87 | 97.56 | 74.7 | 25.44 |
| C9         | 86.25 | 31.25 | 78.58 | 59.4 | 19.98 |
| C10        | 84.56 | 34.38 | 85.71 | 59.0 | 26.14 |
| C11        | 85.01 | 40.69 | 71.96 | 71.4 | 18.58 |

average value 86.64 37.39 79.33 69.5 22.28

This experiment is divided into 11 groups for testing, C1-C11 in turn is normal, a various of tooth wear 1, tooth wear 2, tooth wear 3, tooth wear 4, tooth wear 5, tooth pitting, tooth fracture and inner ring defects, ball roller defects, cage defects such as 11 kinds of fault diagnosis, from ResNet50 and VGG16, ResNet101, GoogleNet, the average accuracy were 49.25, 7.31 and 17.4 percentage points higher. Figure 6 shows the accuracy of various transfer learning training sets. It can be seen from the figure that ResNet50 has the fastest convergence speed and higher accuracy than other models. Figure 7 shows the loss reduction rate of various transfer learning validation sets. Figure 8 shows the accuracy of various transfer learning validation sets. The accuracy of ResNet50 and ResNet101 is almost the same, and sometimes even exceeds ResNet50, but the convergence speed of ResNet50 is much faster, so it can also reflect the superiority of ResNet50 for the fault. Figure 9 is the specific training accuracy of ResNet50, and Figure 10 is the loss rate of ResNet50.

Figure 6 Training set accuracy
Figure 7 Transfer learning training process loss
4. Conclusions
Aiming at the fault diagnosis task of wind turbine gearbox, this paper proposes a diagnosis method based on ResNet50 transfer learning. The main innovations are as follows: Transfer learning is mainly aimed at training color images. One-dimensional data cannot be directly introduced into training, and the original collected one-dimensional data are transformed into three-channel images for input, which reduces the complexity of fault signal extraction. A ResNet50 transfer learning model for gearbox fault diagnosis is designed. In the pre-trained ResNet50 convolutional neural network model, some layers are frozen to accelerate the training speed. After that, the global pooling layer is added, the dropout layer is added to prevent overfitting, and the softmax layer is added to increase the accuracy.

Acknowledgments
Humanity and Social Science Youth foundation of Ministry of Education of China (Project No.20YJCZH027).

References
[1] Jiao J, Zhao M, Lin J, et al. A multivariate encoder information based convolutional neural network
for intelligent fault diagnosis of planetary gearboxes [J]. Knowledge-Based Systems, 2018, 160(Nov.15):237-250.

[2] Infield D, Wang Y. Supervisory control and data acquisition data-based non-linear state estimation technique for wind turbine gearbox condition monitoring [J]. Renewable Power Generation Iet, 2013, 7(4):350-358.

[3] Zhang R, Zhang X, Yan Z. Study on Impact of Wind Turbines on Static Voltage Stability of Power System [J]. Electrical (English version), 2011, 22(1):43-47.

[4] Wang MH, Chen HC. Application of ENN-1 for Fault Diagnosis of Wind Power Systems [J]. Mathematical Problems in Engineering, 2012, 2012(pt.7):720-750.

[5] Chang, Robert W. Synthesis of band-limited orthogonal signals for multichannel data transmission [J]. Bell System Technical Journal, 1966, 45(10):1775-1796.

[6] Gershon RC, Karon C. Use of Computer Adaptive Testing in the Development of Machine Learning Algorithms [J]. Pain Medicine, 2011(10):1450-1452.

[7] Fang W, Subcompany GT . Analysis for Gearbox Failure Cause and Preventive Measures. Modern Industrial Economy and Informationization, 2014, 4(10):53-54.

[8] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition [J]. Computer Science, 2014, arXiv:1409.1556.

[9] Szegedy C, Liu W.Jia Y, et al. Going Deeper with Convolutions [C]. IEEE Conference on Computer Vision and Pattern Recognition, 2015: 1-9.

[10] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition [C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV: IEEE, 2016: 770-778.

[11] Pan SJ, Qiang Y. A Survey on Transfer Learning [J]. IEEE Transactions on Knowledge and Data Engineering, 2010, 22(10):1345-1359.

[12] Saunders C, Stitson MO, Weston J, et al. Support Vector Machine [J]. Computer Science, 2002, 1(4):1-28.

[13] CHEN RX, YANG X, HU XL, LI J, CHEN C, TANG LL. Planetary gearbox fault diagnosis method based on deep belief network transfer learning [J]. Journal of Vibration and Shock, 2021, 40(01):127-133, 150.

[14] TAO QS, PENG C, Man JF, Liu Y. Two-step transfer learning method for bearing fault diagnosis [J/OL]. Computer Engineering and Applications: 1-12.

[15] ZHAO YK, XU GW, LIU M. Bearing fault diagnosis method based on VGG16 migration learning [J]. Spacecraft Environment Engineering, 2020, 37(05): 446-451.

[16] Liu XT. Research on the normalization of input layer data of BP neural network [J]. Journal of Mechanics Engineering and Automation, 2010(03):122-123, 126.

[17] LECUN Y, BOTTOU L, BENGIO Y, et al. Gradient-based learning applied to document recognition [J]. Proceedings of the IEEE, 1998, 86(11): 2278-2324.

[18] Ba JL, Kiros JR, Hinton GE. Layer normalization [EB/OL]. [2019-01-10]. https://arxiv.xilesou.top/pdf/1607.06450.pdf.

[19] Qi X, Wang T, Liu J. Comparison of Support Vector Machine and Softmax Classifiers in Computer Vision [C] // International Conference on Mechanical. IEEE Computer Society, 2017: 151-155.

[20] Iskol'Dskii AM, Nesterikhin YE, Patashinskii AZ, et al. On the instability of a gradient explosion [J]. Doklady Akademii nauk SSSR, 1977, 236:589-591.