Published by :
http://www.ijert.org

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 12 Issue 07, July-2023

Rolling Bearing Fault Detection in CWRU Data Based on Markov Transition Field and Convolutional Neural Network

Cong Chen
Tianjin University of Technology and Education
School of Mechanical Engineering
Tianjin, China

Sen Zhang
Tianjin University of Technology and Education
School of Mechanical Engineering
Tianjin, China

***Abstract*—As one of the most important components of rotating machinery, rolling bearings are widely used in various fields such as aerospace, railway traffic metallurgy, etc. Whether rolling bearings can operate normally plays a decisive role in the safety work of a system, so accurate bearing fault diagnosis method plays a very important role. Traditional bearing fault diagnosis based on signal processing method has been developed very mature, but it relies on manual feature extraction efficiency is low, in the era of industrial intelligence development, its effect is limited. With the rise of deep learning, convolutional neural network with its automatic learning characteristics is widely used in various fields, especially in the visual field. Thus, a method based on Markov transition field-Convolutional neural network (MTF-CNN) is proposed to convert one-dimensional bearing vibration signals into two-dimensional images and then recognize and classify them automatically by convolutional neural network. Experiments show that this method has a high accuracy.**

***Keywords*—convolutional neural network; bearing; Markov transition field**

I. INTRODUCTION

Rolling bearing fault diagnosis is mainly divided into bearing fault diagnosis based on signal processing and bearing fault diagnosis based on pattern recognition.

In practical engineering, due to the impact of environmental noise, transmission path of signal acquisition and other factors, the periodic transient impact signal representing bearing faults is submerged by severe background noise, power frequency vibration, and other strong interference signals. At this time, direct demodulation analysis of fault time domain signals cannot find weak fault characteristics, and it is easy to cause missed diagnosis of faults [1]. Therefore, looking for an effective impact feature extraction method has always been the focus of rolling bearing fault diagnosis research. At present, there are four main methods to extract impact features from the original bearing fault vibration signals: time domain feature extraction, frequency domain feature extraction, time frequency domain feature extraction and sparse representation. Tang et al. [2] introduced MCKD method into the impact feature extraction of rolling bearings, and combined with envelope demodulation analysis, effectively realized bearing fault diagnosis. Mcdonald etc. [3] proposed correlation kurtosis to balance the impact and continuous periodicity of signals.

The essence of mechanical equipment fault diagnosis is to carry out feature extraction and pattern recognition according to the running state of equipment [4]. Usually, when the equipment is faulty, the characteristics of the signal in the time domain, frequency domain and frequency domain will be changed. On this basis, fault information characterization can be realized by designing feature sets in different domains in combination with signal processing methods [5]. Based the extracted features, the fault pattern recognition can be carried out. The purpose of fault pattern recognition is to mine the effective feature information contained in the data, to predict or identify the equipment fault. Many scholars at home and abroad have carried out a lot of work on fault pattern recognition based on extracted features.

II. CORRELATION THEORY

A. *Markov Transition Field*

Markov transfer field is a time series image coding method based on Markov transfer matrix. During this method, the time progression of time series is regarded as a Markov process, that is, its future evolution does not depend on its past evolution under the condition of known current state. Therefore, the Markov transfer matrix is constructed, and then the Markov transfer field is extended to realize image coding [6].

For time series $X = (x_t, t = 1, 2 \cdots T)$,The image coding steps are as follows:

- Divide the time $X(t)$ series into Q compartments (labeled $1, 2 \cdots Q$, each with the same number of compartments).
- Change each piece of data in the time series to the serial number of its corresponding bit box.
- Construct the transition matrix W (w_{ij} represents the frequency at which cell *i* is transferred to cell *j*):

$$W = \begin{bmatrix} w_{11} & \cdots & w_{1Q} \\ w_{21} & \cdots & w_{2Q} \\ \vdots & \ddots & \vdots \\ w_{Q1} & \cdots & w_{QQ} \end{bmatrix} \tag{1}$$

- Construct Markov transfer field M :

$$W = \begin{bmatrix} w_{ij} \mid x_1 \in q_i, x_1 \in q_j & \cdots & w_{ij} \mid x_1 \in q_i, x_N \in q_j \\ w_{ij} \mid x_2 \in q_i, x_1 \in q_j & \cdots & w_{ij} \mid x_2 \in q_i, x_N \in q_j \\ \vdots & \ddots & \vdots \\ w_{ij} \mid x_N \in q_i, x_1 \in q_j & \cdots & w_{ij} \mid x_N \in q_i, x_N \in q_j \end{bmatrix} \tag{2}$$

B. Convolutional Neural Network

Convolutional neural network is a kind of feedforward neural network, which is mainly composed of input layer, convolutional layer, pooling layer, fully connected and output layer.

- For this paper, the input is the two-dimensional picture obtained by the original signal through the calculation process of Markov transition field.
- Convolution layer is mainly used to extract image features. The local and global features of 2D images obtained by Markov transition field calculation for different fault types are different.
- The pooling layer is mainly a function of downsampling, because the computation amount of the neural network is very large, in order to reduce the computation amount and improve the computational efficiency of the convolution.
- The fully connected layer is mainly composed of neurons, which is used to integrate differentiated local feature information between different categories, and finally sent to the output layer for classification and recognition.

Fig. 1 shows the most typical LeNet-5 convolutional neural network structure diagram, which was proposed by LeCun Yann in 1998[7].

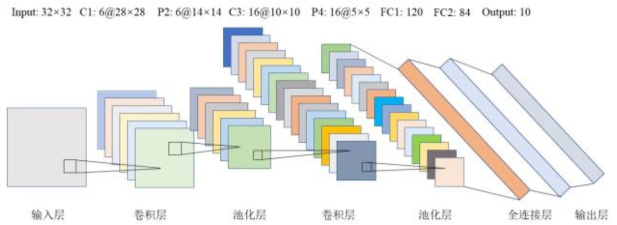


Fig. 1. The structure of LeNet-5

III. EXPERIMENTAL SETUP

A. EXPERIMENTAL DATA

The experimental data of this article uses the motor rolling bearing data set of Case Western Reserve University. The data set acquisition system is shown in Fig. 2, mainly including the drive motor, torque sensor and encoder, ergograph and data acquisition card. The acceleration sensor is installed on the motor shell at the fan end, drive end and base to collect vibration acceleration signals. The data sampling frequency is divided into 12kHz and 48kHz.Each sampling frequency has 1hp(1772rpm), 2hp(1750rpm) and 3hp(1730rpm) four different loads, each load has inner ring, outer ring, ball damage in three different positions, each damage corresponds to three different damage diameters of 0.007inch, 0.014inch and 0.021inch. Together with the normal state data, each load can be divided into 9 fault states and normal state vibration signal data, a total of 10 fault categories.

In this paper, vibration signal data with sampling frequency of 48kHz is used. The diagnosed object is the driving end deep groove ball bearing, whose model is SKF6205, and its geometric parameters are shown in Table. Ⅰ.

TABLE I. BEARING GEOMETRIC PARAMETER

Inner ring diameter	*Outer ring diameter*	*Rolling diameter*	*Bearing pitch diameter*	*Number of balls*
0.9843 inch	*2.0472 inch*	*0.3136 inch*	*1.537 inch*	*9*

The total number of samples divided into this experimental data is 10,000, among which the ratio of training set, verification set and test set is 7:2:1. The data partitioning method is shown in Table. Ⅱ. The three damage diameters (0.007inch, 0.014inch and 0.021inch) corresponding to the inner ring, outer ring and ball bearing are marked with C0、C1、C2、C3、C4、C5、C6、C7、C8、C9 respectively, and the normal state data is marked with C9, totaling 10 fault types.

TABLE II. DATA SET PARTITIONING

Damage location	*Outer*			*Inner*			*Ball*			*NC*
Mark	*C1*	*C2*	*C3*	*C4*	*C5*	*C6*	*C7*	*C8*	*C9*	*C0*
Train set	*700*	*700*	*700*	*700*	*700*	*700*	*700*	*700*	*700*	*700*
Val set	*200*	*200*	*200*	*200*	*200*	*200*	*200*	*200*	*200*	*200*
Test set	*100*	*100*	*100*	*100*	*100*	*100*	*100*	*100*	*100*	*100*

B. CONVOLUTIONAL NEURAL NETWORK PARAMETERS

This experiment uses a neural network model with 2 convolutional layers, 2 pooled layers, and 2 fully connected layers. Using the Dropout function at the fully connected layer results in half of the neurons losing activity and less data overfitting. Loss function was selected as Cross Entropy loss function, which had a good effect in the classification experiment. The optimization method was selected as Adam optimizer, and the specific hyperparameter Settings were shown in Table. Ⅲ:

TABLE III. DATA SET PARTITIONING

Number	*Name*	*Parameter settings*
1	*Training times*	*100*
2	*Batch size*	*256*
3	*Learning rate*	*0.01*
4	*Weight decay*	*0.0005*
5	*gamma*	*0.9*

It is worth mentioning that the hardware configuration of this experiment is mainly intel i7-6700@3.40GHz CPU and NVIDIA GEFORCE GT 730 graphics card. The programming environment is Python3.9, and the construction of convolutional neural network is mainly based on Pytorch deep learning framework.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 12 Issue 07, July-2023

C. FAULT DIAGNOSIS PROCESS

The rolling bearing fault diagnosis process based on Markov transition field and convolutional neural network is shown in Fig. 2. First, the vibration data of Case Western Reserve University motor rolling bearing data set is read, and then classified according to the method in Table. Ⅱ. Each data sample after classification is divided into a bit box, and then each data in the time series is changed to the serial number of its corresponding bit box, and then the transfer matrix W is constructed. Finally, the Markov transfer field is calculated, and M is used as the input of the neural network for training, testing and visualization.

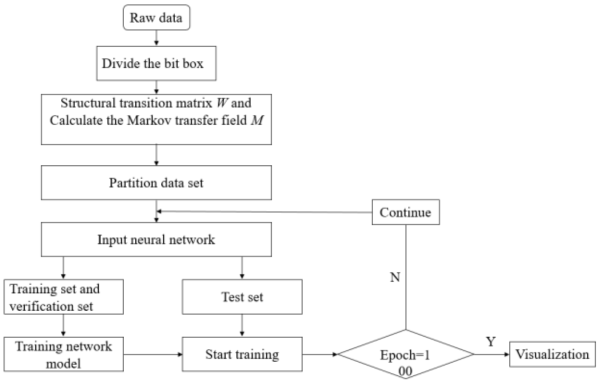


Fig. 2. Fault Diagnosis Process

IV. ANALYSIS OF EXPERIMENTAL RESULTS

Vibration signal data with a sampling frequency of 48kHZ and bearing load of 1hp(1772rpm) were selected as input, and the sampling length was 4096. The data were divided according to the method in Table 2. Fig. 3 and Fig. 4 shows the original one-dimensional signal visualization of C0、C3、C6 and C9 data and the two-dimensional visualization of the calculated Markov transition field M .

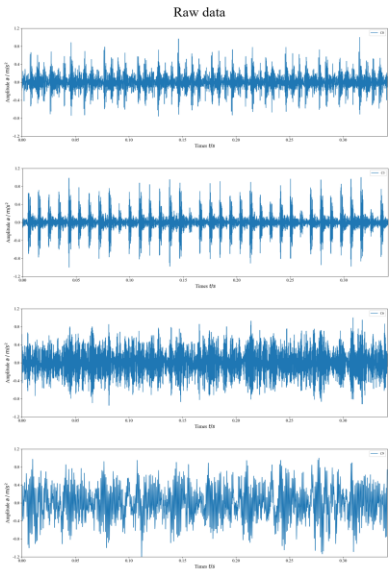


Fig. 3. Raw data

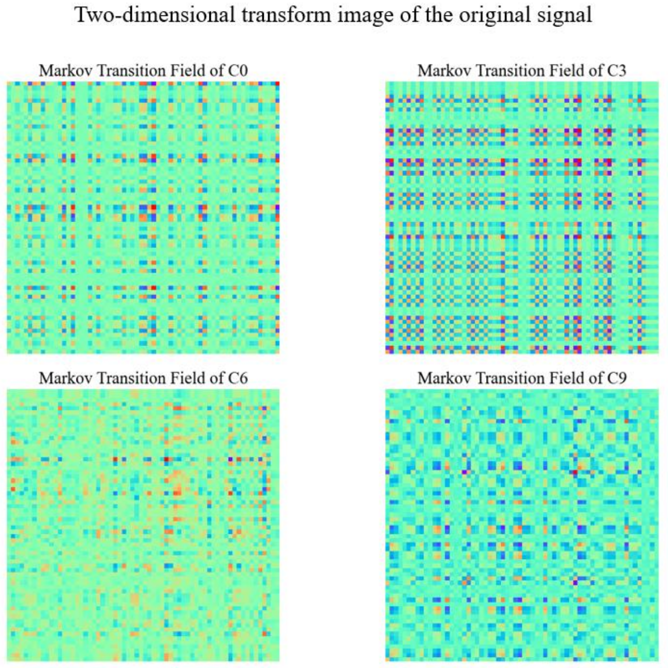
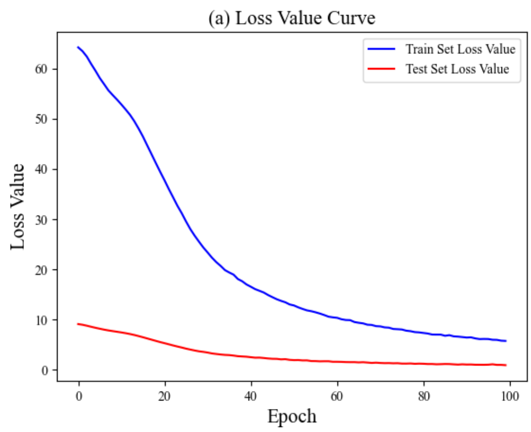


Fig. 4. Two-dimensional transform image of the original signal

It can be seen from Fig. 3 that C0、C3、C6 and C9 faults have different periodic impact components in the time domain, but it is difficult to distinguish the differences between C0 and C3 and between C6 and C9 in the time domain. Figure 4 is the two-dimensional image of Markov transition domain after calculation of the four signals, and the difference of different signals can be intuitively distinguished from the figure.

Convolutional neural networks are characterized by poor interpretability, etc. so we use visualization to analyze experimental results, including loss function curve, accuracy curve, etc. The strengths and weaknesses of the model can be analyzed through different visualizations.

Fig. 5 shows the visualization of the loss function and accuracy curve of the training set and the test set. It can be seen from the Fig. 5(a) that when the training is close to 100 times, the loss function approaches 0, indicating that the training of the network model is close to the saturation state. It can be seen from the Fig. 5(b) that the accuracy of the final training set is about 93%, and the accuracy of the test set is 90%.



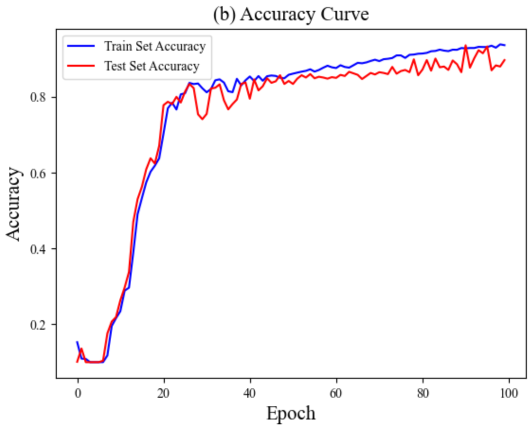


Fig. 5. Loss function and accuracy curve

Fig. 6 shows the confusion matrix of the training set and the test set. The confusion matrix can show the prediction effect of different fault categories. It can be seen from the Fig. 6 that the prediction effect of faults C1, C2 and C3 is not ideal. The results show that the model can distinguish the damage at different locations well, but it is not sensitive to the three different kinds of damage in the outer ring. The causes of this phenomenon will be further explored in the future research.

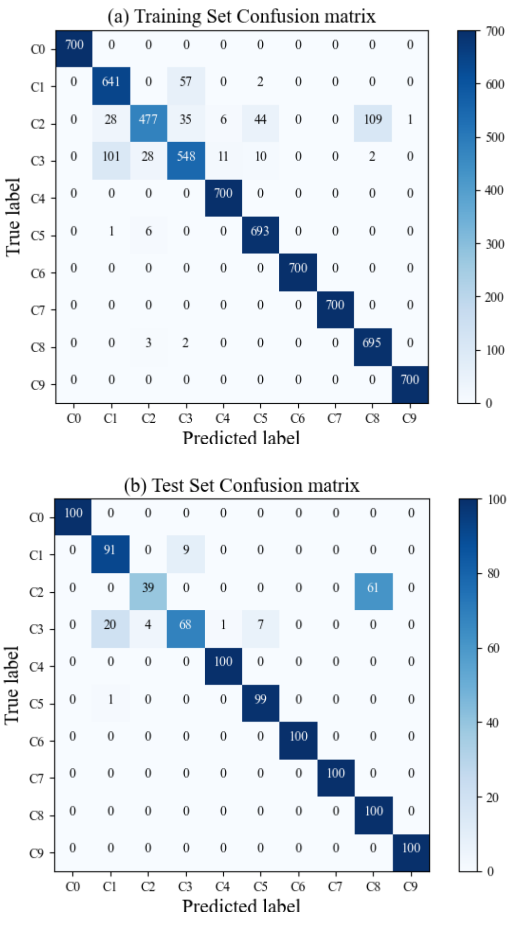


Fig. 6. Confusion matrix of training set and test set

Fig. 7 shows the learning degradation curve set in this experiment. With the initial learning rate set at 0.001, the model has a high learning speed and becomes stable after 60 training times. Therefore, it is necessary to train at a smaller learning rate to obtain more stable parameter weights. The learning rate degradation parameters are shown in Table 2, and the learning rate update formula is $lr\_new = lr\_old * gamma$.

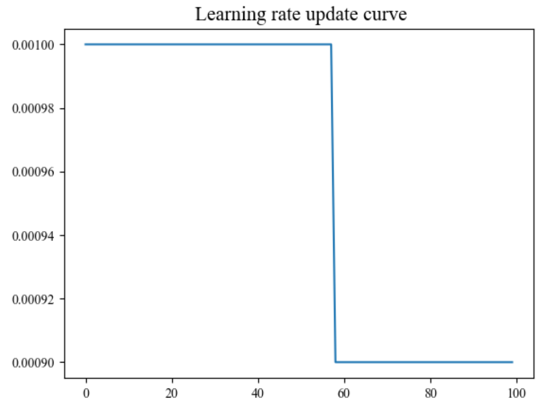


Fig. 7. Loss function and accuracy curve

V. CONCLUSION

The fault diagnosis of rolling bearing based on MTF-CNN is proposed by using one-dimensional signal as input of convolutional neural network after Markov transition field calculation, and good results are obtained on CWRU data set. It provides a new method and idea for bearing fault diagnosis.

REFERENCES

[1] Lei Ya Guo, Han Dong, Lin Jing. New Adaptive Stochastic Resonance Method and its Application in Fault Diagnosis [J]. Journal of Mechanical Engineering, 2012,48(7): 62-67.

[2] Tang G, Wang X, He Y. Diagnosis of compound faults of rolling bearings through adaptive maximum correlated kurtosis deconvolution[J]. Journal of Mechanical Science and Technology, 2016, 30(1): 43-54.

[3] Mcdonald G L, Zhao Q, Zuo M J. Maximum correlated Kurtosis deconcolution and application on gear tooth chip fault detection[J]. Mechanical Systems and Signal Processing, 2012, 33:237-255.

[4] Diallo D, Benbouzid M E H, Hamad D. Fault detection and diagnosis in an induction machine drive: A pattern recognition approach based on concordia stator mean current vector[J]. IEEE Transactions on Energy Conversion, 2005, 20(3):512-519.

[5] Pan Haiyang, Yang Yu, Li Yongguo. Rolling Bearing Fault Diagnosis Method Based on Manifold Learning and improved VPMCD [J]. Chinese Journal of Mechanical Engineering, 2014, 27(06): 934-941.

[6] Liu, Wuping, and F. Wei. "USING LOCAL TRANSITION PROBABILITY MODELS IN MARKOV RANDOM FIELD FOR MULTI-TEMPORAL IMAGE CLASSIFICATION." International Conference on Manufacturing Technology and Electronics Applications 2014.

[7] Lecun Y. LeNet-5, convolutional neural networks[J]. [2023-07-05] .