Early Gear Pitting Fault Diagnosis Based on Bi-directional LSTM

Xueyi Li, Jialin Li, and Chengying Zhao
School of Mechanical Engineering and Automation
Northeastern University
Shenyang, China
lixueyineu@gmail.com

Yongzhi Qu
School of Mechanical and Electronic Engineering
Wuhan University of Technology
Wuhan, China
quwong@whut.edu.cn

David He*
Department of Mechanical and Industrial Engineering
University of Illinois at Chicago
Chicago, USA
*Corresponding author: davidhe@uic.edu

***Abstract*—The early gear pitting fault diagnosis has received much attention in the industry. In recent decades, with the popularity growth of artificial neural network, researchers have applied deep learning methods to figure out early gear pitting faults. However, the classical fault diagnosis methods usually use deep neural networks according to the time sequence of the collected signals. In this case, the feature extraction in the direction of the inverse time-domain signals is usually ignored. Aimed at overcoming this shortage, ground on a traditional Long Short Term Memory (LSTM) network, this paper proposes a Bi-directional LSTM (Bi-LSTM) to construct a fault diagnosis model of early gear pitting using raw vibration signals. Using the Bi-LSTM network, feature extraction of the vibrational signals in both directions is simultaneously carried out to evaluate the degree of the early gear pitting faults to better extract the gear pitting characteristics from the raw vibration signals of the gear. Through the analysis of the experimental data, compared with the traditional LSTM model, the Bi-directional LSTM has a classification accuracy of over 96% for early gear pitting fault diagnosis, which is an increase of 4.1%.**

Keywords- artificial neural network; gear; vibration signal; long short term memory; gear pitting diagnosis

I. Introduction

Gears are important fundamental parts of mechanical equipment and industrial production. Gears play an essential role. The early diagnosis of pitting of gears has been the focus of research by researchers. In the past decade, traditional machine learning methods have made some progress in the diagnosis of gear pitting faults. Researchers use support vector machines in gearboxes or helicopters [1]. there are some studies on the use of feature sequencing methods and support vector machines based classifiers. However, many studies are limited to summarizing the rules of training sets with a small order of magnitude. Due to the limitation of formula derivation, it is difficult for many methods to find rules from big data. Fortunately, thanks to the rise of artificial neural networks, many laws that were previously difficult to summarize can be realized through big data derivation. Thanks to this great discovery, we can derive laws beyond human comprehension.

With the development of neural networks, some ideas of backpropagation(BP) are also reflected in the fault diagnosis of gears. There has also been some BP neural network-based gearbox fault diagnosis [2]. And some changes based on this. For example, BP neural network based on genetic algorithms for the intelligent diagnosis of gear faults [3]. Gear fault diagnosis based on the harmonic wavelet packet and BP neural network [4].

Deep learning is a hot research topic recently. Many methods of deep learning are applied in gear fault diagnosis, such as combining the convolutional neural network Gated Recurrent Unit (GRU) to detect the fault of the gear [5], and the pitting of the gear is detected by an autoencoder [6]. Feature learning and fault diagnosis method for gearbox condition monitoring based on convolutional neural network(CNN) [7]. Although CNN has strong local feature extraction ability. However, it is limited by the size of the receptive field. For a relatively large input window length, the shallower CNN cannot effectively extract the information far away from the raw vibration signal of the gear. For time series problems with long-term dependencies, Long Short Term Memory (LSTM), GRU and other Recurrent Neural Network(RNN) with gating mechanism show good processing results. For example, using a GRU network based on local features for machine health monitoring [8]. But the above writing method only extracts the fault characteristics from one dimension of time, in order to make full use of the information in the raw gear raw vibration signal. In this paper, the Bi-directional LSTM (Bi-LSTM) is used to correlate the local characteristics of the vibration signals extracted from different directions to build up the accuracy of early gear pitting diagnosis. The advantages of Bi-LSTM also include excellent stability.

This work was funded by NSFC (Grant No. 51675089).

978-1-7281-0860-5/19/$31.00 ©2019 IEEE

The other chapters of this paper are arranged as follows. The second part introduces the theory of gear fault diagnosis. The third part puts forward the method of test verification. The fourth part analyzes and discusses the experimental results. The fifth part is to summarize the full paper.

II. The Methodology

The basic idea of the Bi-LSTM neural network is to use two LSTM neural networks simultaneously for each training sequence, one forward and one backward. And both are connected to the same output layer. The characteristic of this structure is that the nodes of each output layer can make full use of the complete history and further raw vibration signal information of the gears. Based on this idea, the raw vibration signal is trained to diagnose the early pitting fault of the gear. The overall framework is shown in Fig. 1. The raw gear vibration signal is extracted by the LSTM network in both forward and reverse directions. The signal is connected to obtain the diagnosis result. Bi-LSTM is an extension of traditional LSTM, which can improve the model performance of sequence classification problems.

A. Long Short Term Memory

LSTM effectively mitigates the gradient disappearance problem in RNN by constructing multiple gates. A typical LSTM unit is shown in Fig. 2. The LSTM neural network refers to the RNN of the hidden layer replaced by the long short term memory. Compared with the hidden layer of the RNN, the biggest improvement of the LSTM unit is the introduction of memory cells to preserve historical information. In the LSTM unit, it is called the cell state, or simply referred to as the state, and the state update depends on the interaction of the three gates, that is, the input gate, the output gate, and the forgetting gate. The internal relationship [9] is:

$$f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t] + b_f\right) \tag{1}$$

$$i_t = \sigma\left(W_i \cdot [h_{t-1}, x_t] + b_i\right) \tag{2}$$

$$o_t = \sigma\left(W_o \cdot [h_{t-1}, x_t] + b_o\right) \tag{3}$$

$$\tilde{C}_t = tanh\left(W_C \cdot [h_{t-1}, x_t] + b_C\right) \tag{4}$$

$$C_t = f_t \times C_{t-1} + i_t \times \tilde{C}_t \tag{5}$$

$$h_t = o_t \times tanh\left(C_t\right) \tag{6}$$

where, x_t is the input signal vector, o is the Logistic Sigmoid function, and *tanh* () is the hyperbolic tangent map, which provides a nonlinear activation function. i_t, f_t, o_t and C_t are the input gate, the forget gate, the output gate, and the state vectors in the LSTM unit, respectively, which have the same data dimensions as the hidden layer vector h_t. W refers to the weight matrix, and b is the corresponding bias term. σ represents the corresponding function, which is obtained by the neural network through iteration.

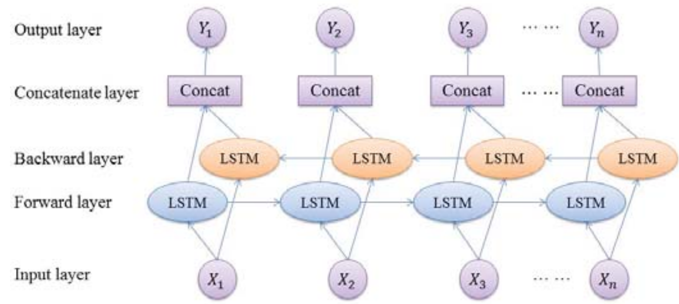

Figure 1. The overall framework of the method proposed in this paper.

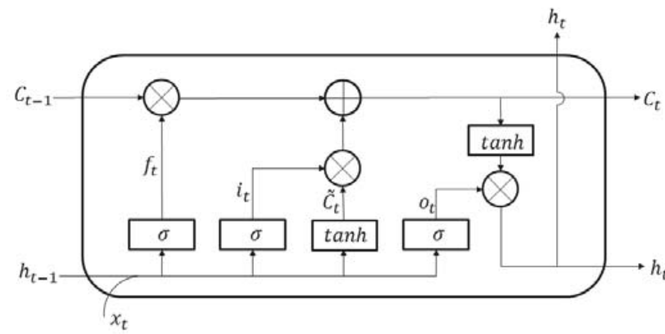

Figure 2. Long and short memory unit structure.

B. Bi-directional LSTM

The single direction transmission mode of the LSTM unit information is such that the cell state in the LSTM unit at each moment contains only the current time and the previous time information. In other words, the LSTM unit only integrates the previous raw vibration signal information of the gear with that at the current moment. If the raw gear vibration signal is processed simultaneously in both forward and reverse chronological order, features that the LSTM can extract but are ignored by the unidirectional LSTM can be obtained. In order to enable the LSTM unit to fuse the gear raw vibration signal information of the current moment with all of its gear raw vibration signal information and output the final result, we use a Bi-LSTM [10]. The structure of the Bi-LSTM layer is shown in Fig 3. It is known from Fig. 3. The number of parameters of the bidirectional LSTM layer is twice that of the normal timing LSTM layer. The Bi-LSTM method used in this paper mainly reverses the gear raw vibration signal sequence along the time dimension by a variant of the data generator.

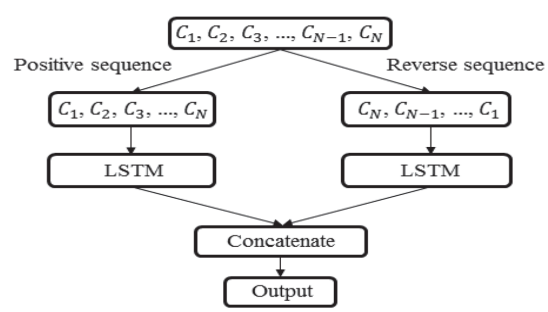

Figure 3. The principle of the Bi-LSTM layer.

Because the judgment of the gear pitting condition does not depend on the specific position of the vibration signal in the collected sample sequence. In fact, the representations learned in the reverse sequence are not the same as those learned in normal timing signals. Bi-LSTM is an integrated idea to improve the performance of traditional positive-sequence LSTM. The Bi-LSTM can extract the characteristics of the raw vibration signal from two directions, resulting in a richer feature representation, especially the features that may be neglected by using only positive raw vibration signal of the gear.

III. Gear Test Experimental Verification

On behalf of check on the efficacy of the arranged Bi-LSTM method for early gear pitting fault diagnosis, a set of experiments was designed to verify it. The test rig is shown in Fig. 4 and constitutes of two 45kw motors. The sensor used is a three-axis acceleration sensor. This experiment was carried out under the conditions of a rotational speed was 1500 RPM and a load was 50 Nm.

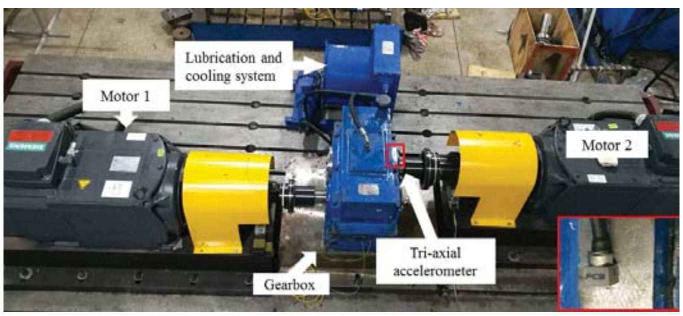

Figure 4. Photo of the experimental equipment.

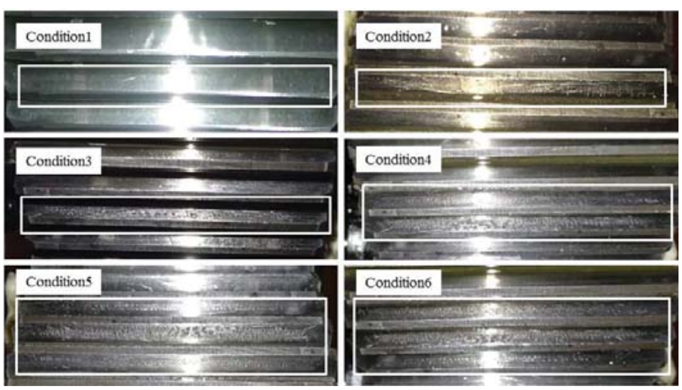

Figure 5. Pitting degree of driven gears.

TABLE I. The centage of gear Pitting region

Condition	Upper	Middle	Lower
C1	healthy	healthy	healthy
C2	healthy	10%	healthy
C3	healthy	30%	healthy
C4	10%	50%	healthy
C5	10%	50%	10%
C6	30%	50%	10%

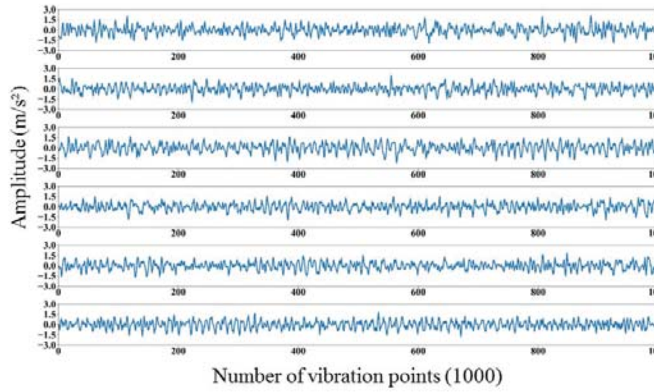

Figure 6. Raw gear vibration signals (From top to bottom: condition1, 2, 3, 4,5, and 6).

The test was carried out using six different pitting gears. C1 is a healthy gear, and the intermediate gears of C2 and C3 have 10% and 30% tooth surface pitting, respectively. In order to simulate a more realistic situation, C4 to C6 has approximately 50% pitting in accretion to the intermediate teeth. The specific pitting ratio is shown in Table 1. The tooth surface pitting centage of the six pitting gears is given in Fig. 5.

The initial vibration signal of the six conditions is shown in Fig. 6 under the operating conditions of 1500 RPM and 50 Nm. It can be seen from the value that the raw vibration signal is almost the case of different kinds of gear pitting. In the same way, it is difficult for the naked eye to distinguish the raw vibration signal corresponding to the gear with different pitting degrees.

In this experiment, a total of 5 sets of samples were collected, the sampling rate was 10.24 Hz, and the number of each set of samples was 62,464. The first four sets of samples were used to train the Bi-LSTM, and the remaining set was used for verification and testing. There are no data reused between the training set, the validation set, and the test set. The network of this experiment is superimposed with four sets of Bi-LSTM neural networks. The cell size is 256, 128, 64, 64 respectively. The first two Bi-LSTM layers are followed by a dropout layer with a parameter of 0.3 to avoid over-fitting. A dense link layer is added behind the Bi-LSTM and softmax classifies the severity of gear pitting. In the case of no increase in the verification accuracy for 50 times, the training was stopped automatically and the experimental results were saved. This test experiment was performed using the NVIDIA 1080Ti graphics card to increase the speed of calculation.

IV. Results and Discussions

The test set is tested on the trained Bi-LSTM network. The confusion matrix is shown in Fig. 7. The Bi-LSTM method for the early pitting diagnosis of gears is ideal. The overall accuracy of pitting in six different cases exceeds 96%, of which 5 types of gear pitting diagnostic accuracy exceeds 95%. In contrast, the gear pitting of the condition2 has only 10% light pitting, and the result is relatively low. The results of the recognition error are mostly close to the actual working conditions. This may be due to a fault that is not obvious enough.

The Bi-LSTM proposed in this paper is compared with the traditional LSTM method for gear pitting fault diagnosis, as shown in Table 2. The diagnostic effect of Bi-LSTM on early gear pitting is better than the traditional LSTM in the test set, verification set, and training set. It is indicated that the feature extraction in two directions can extract the neglected features from the signal than the conventional one-way feature. The accuracy criteria are based on the correctness of the different gear fault types.

The Bi-LSTM method and the traditional LSTM method would be used to compare the diagnosis results of gear pitting in different kinds of working conditions, as shown in Fig. 8. The gear pitting for the first five conditions Bi-LSTM has better accuracy for gear pitting than the traditional LSTM method for gear pitting. In a classic LSTM network, state transmission is unidirectional. However, in some problems, the output of the current time is not only related to the previous state, but also related to the subsequent state. At this time, Bi-LSTM is needed to solve such problems. For the last condition, the two methods of pitting gears have similar diagnostic results. Because the pitting of the last gear is more severe than the pitting of the previous gears, this may indicate that the Bi-LSTM has a better diagnostic effect on earlier pitting gears. It is undeniable that LSTM is difficult to train on very long data, and this problem is generally not likely to be encountered. Similarly, replacing the LSTM in Bi-LSTM with a GRU structure can constitute a Bi-GRU structure. In the future, the author will conduct further research. In the future, the author will conduct further research.

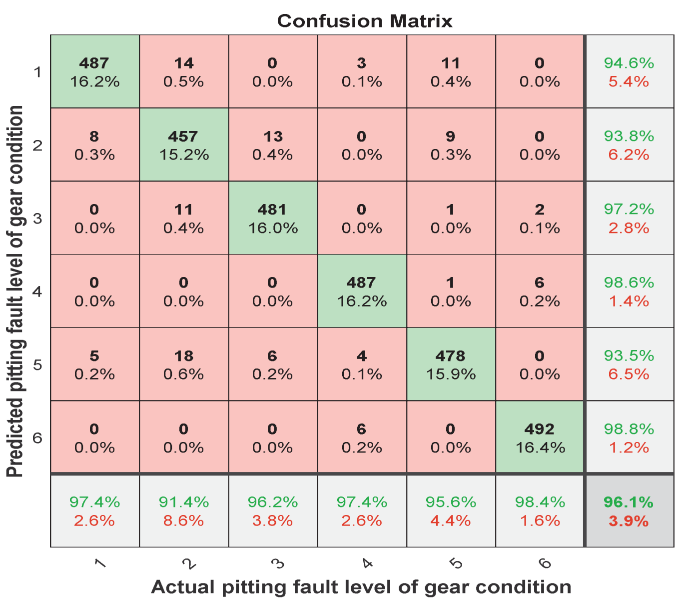

Figure 7. The confusion matrix by the Bi-LSTM.

TABLE II. THE COMPARISON OF ACCURACY BETWEEN THE BI-LSTM AND THE LSTM FOR TEXT SETS

Accuracy	Training set	Validation set	Testing set
Bi-LSTM	0.999	0.965	0.961
LSTM	0.987	0.941	0.923

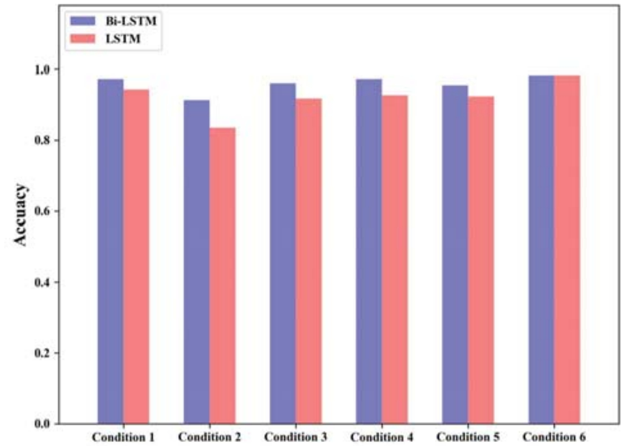

Figure 8. Accuracy comparison of gear pitting in different states.

V. CONCLUSIONS

Early gear pitting fault diagnosis plays a crucial role in maintaining good operational conditions of industrial equipment. A Bi-LSTM model early gear pitting fault diagnosis was proposed in this paper. It can effectively extract the characteristic features in two directions of the raw vibration signals to diagnose the gear pitting faults more effectively. The experimental results show that the prediction performance of the Bi-LSTM model is better than the traditional LSTM network model.

ACKNOWLEDGEMENT

This work was funded by National Natural Science Foundation of China (Grant No. 51675089). This work was supported by the China Scholarship Council.

REFERENCES

[1] V. Camerini, G. Coppotelli, and S. Bendisch, “Fault detection in operating helicopter drivetrain components based on support vector data description,” Aerosp Sci Technol, 2018, 73, pp.48-60.

[2] C. Sreepradha, A. K. Kumari, A. E. Perumal, R. C. Panda, K. Harshabardhan, and M. Aribalagan, “Neural network model for condition monitoring of wear and film thickness in a gearbox,” Neural Comput Appl, 2014, 24, pp.1943-1952.

[3] J. Fang, “Gear fault diagnosis method of intelligence based on genetic algorithm to optimize the BP neural network,” Proceedings of the 2nd International Conference on Computer and Information Applications (Iccia 2012), 2012, pp.815-818.

[4] C. Jiang, X. T. Weng, and J. J. Lou, “Gear Fault Diagnosis Based on Harmonic Wavelet Packet and BP Neural Network,” Advanced Materials and Process Technology, Pts 1-3, 2012, 217-219, pp.2683-2687.

[5] X. Y. Li, J. L. Li, Y. Z. Qu, and D. He, “Gear Pitting Fault Diagnosis Using Integrated CNN and GRU Network with Both Vibration and Acoustic Emission Signals,” Appl Sci-Basel, 2019, 9.

[6] Y. Z. Qu, M. He, J. Deutsch, and D. He, “Detection of Pitting in Gears Using a Deep Sparse Autoencoder,” Appl Sci-Basel, 2017, 7.

[7] L. Y. Jing, M. Zhao, P. Li, and X. Q. Xu, “A convolutional neural network based feature learning and fault diagnosis method for the condition monitoring of gearbox,” Measurement, 2017, 111, pp.1-10.

[8] R. Zhao, D. Z. Wang, R. Q. Yan, K. Z. Mao, F. Shen, and J. J. Wang, "Machine Health Monitoring Using Local Feature-Based Gated Recurrent Unit Networks," Ieee T Ind Electron, 2018, 65, 1539-1548.

[9] L. Yu, J. L. Qu, F. Gao, and Y. P. Tian, "A Novel Hierarchical Algorithm for Bearing Fault Diagnosis Based on Stacked LSTM," Shock Vib, 2019.

[10] B. Kim, K. Chung, J. Lee, J. Seo, and M. W. Koo, "A Bi-LSTM memory network for end-to-end goal-oriented dialog learning," Comput Speech Lang, 2019, 53, pp.217-230.