Fault Diagnosis of Rolling Bearing Using Wireless Sensor Networks and Convolutional Neural Network

Liqun Hou (✉), Zijing Li, Huaihong Qu
North China Electric Power University, Baoding, China
houliquin@ncepu.edu.cn

https://doi.org/10.3991/ijoe.v16i11.15959

Abstract—Rolling bearings are widely used in modern production equipment. Effective bearing fault diagnosis method will improve the reliability of the machinery and increase its operating efficiency. In this paper, a novel fault diagnosis method based on WSN and CNN has been proposed to fully utilize the strong fault classification capability of CNN and the inherent merits of WSNs, such as relatively low cost, convenience of installation, and ease of relocation. The feasibility and effectiveness of proposed system are evaluated using the vibration data sets of seven motor operating conditions released by the Case Western Reserve University Bearing Data Center. The experimental results show the fault diagnosis accuracy of the proposed approach can reach 97.6%.

Keywords—Fault diagnosis, wireless sensor networks, convolutional neural network, rolling bearing.

1 Introduction

Rolling bearings are widely used in modern production equipment and nearly 50 percent mechanical faults are occurred on bearing and related components [1]. For this reason, much attention has been focused on effective bearing fault diagnosis methods [2-4]. It is well known that the structural damage of rolling bearings usually leads to high-frequency and periodic shock on the vibration signal [5-7]. Therefore, identifying such shock signals from the original vibration signal is the key step for fault diagnosis [8]. At present, vibration signals of rolling bearings are generally collected by wired communication. However, the traditional wired network requires a great deal of man-power and material input in wiring, assembling and later maintenance, which may even exceed the price of the sensor itself [9-11].

The emergence of wireless sensor networks (WSNs) solves the above problems in wired networks. In recent years, Internet of things technology gradually attracts people's attention. WSNs, as the carrier of the underlying IoT, develop rapidly as well [12]. WSNs are communications networks composed of a large number of inexpensive wireless microcontrollers that can transmit wireless data in a given area [13]. WSNs are widely used in most fields because of its advantages such as flexible installation position, economical network maintenance, and convenient node replacement [14]. In terms of bearing monitoring, J. Sun et al. [15] proposed a data acquisition and reconstruction
scheme based on compressed sensing and sparse Bayesian learning algorithm to solve the problem of insufficient power supply of WSN in the industrial environment. This method can reduce the transmission pressure of WSN nodes and improve the life of nodes without affecting the monitoring effect. G. Feng et al. [16] achieve signal acquisition, data processing using fast Fourier transform and Hilbert transform, and feature extraction using envelope spectrum analysis on the end node of WSNs, and then only transmit the fault characteristics to the coordinator node and the PC. Experiments show that this method can reduce data transmission by 95% compared with direct transmission of raw data. Besides machine condition monitoring and fault diagnosis, WSNs also plays an irreplaceable role in the fields of structure monitoring [17-19], home automation [20], mass transportation [21-22] and security prevention [23].

Deep learning was formed in the 1980s and booming by 2012 with the development of convolutional neural network (CNN) [24-25]. Deep learning has been widely used in object image recognition [26], natural language processing [27], medical diagnosis and treatment [28], infrastructure engineering [29], and other fields due to its excellent feature extraction capability. CNN as one of the classic deep learning methods has been successfully applied in fault diagnosis. D. Verstraete et al. [30] built a fault diagnosis model based on time-frequency images of vibration signals by utilizing the powerful image recognition capability of CNN. Experiment results show the presented method can get better results with fewer parameters. M. Sadoughi et al. [31] divided feature extraction and CNN fault diagnosis into blocks, first selecting specific data features, and then taking these features as CNN input. This method can effectively solve the problem of simultaneous state monitoring and fault diagnosis of multiple rolling bearings. S. Udmale et al. [32] presented and evaluated a bearing fault diagnosis approach based on kurtogram and deep learning sequence models. D. Belmiloud [33] presented and tested a novel bearing failure prediction method based on wavelet packet decomposition and deep convolutional neural networks. T. Pan et al. [34] proposed and verified a fault diagnosis method for shipborne antenna using 1D-CNN based on multiscale inner product and locally connection feature extraction.

In summary, compared with traditional wired systems WSNs have many inherent merits, while CNN is a promising deep learning method for bearing fault diagnosis. However, bearing fault diagnosis by combining WSNs and CNN is a relatively unexplored area.

Compared with the above-mentioned bearing fault diagnosis methods, this paper proposes a novel bearing fault diagnosis method based on WSNs and CNN, in which WSNs are used to measure and transmit the bearing vibration signal, while CNN algorithm on a laptop is used for bearing fault diagnosis.

The reminder of this paper is organized as follows. Section II briefly introduces the principle of CNN, while Section III describes the system architecture and implementation methodology. Section IV gives the experimental evaluation of the proposed method. Finally, Section V presents the overall conclusions.
2 CNN Principle

CNN is generally composed of an input layer, several convolutional layers and pooling layers, a fully connected layer, and an output layer. The convolutional layer and pooling layer can appear alternately more than one times. The convolutional layer is responsible for extracting the feature information from the input signal, while the pooling layer is used to reduce the number of parameters and the computation load. In addition, the fully connected layer can be replaced by the convolutional layer. The structure of the convolutional neural network is shown as Fig. 1.

![Convolutional neural network structure diagram](http://www.i-joe.org)

**Fig. 1.** Convolutional neural network structure diagram

### 2.1 Convolutional layer

The convolutional layer is the core layer of a CNN. It convolves the input signal with different receptive fields to get a new signal. This procedure is named as feature mapping. Different receptive fields can extract different feature maps, which represent various feature information in the original signal. This feature information is the main basis for classification or regression in the fully connected layer. The convolution layer function expression is

\[
H_j = f \left( \sum_{i=1}^{p} H'_i \times W_{ij} + b_j \right)
\]

where \(H_j\) is the \(j\)-th output feature map of the convolutional layer, \(H'_i\) is the \(i\)-th input signal of the previous layer, \(p\) is the number of receptive fields in the previous layer, \(W_{ij}\) represents the weight matrix connecting \(H'_i\) and \(H_j\), \(b_j\) represents the additive bias corresponding to the \(j\)-th receptive field, \(f(\cdot)\) is the excitation function.
2.2 Pooling layer

The main function of the pooling layer is to reduce the size of its input signal, and then reduce the total number of parameters and calculation load of the CNN. By combining the signal in a certain range into a representative value reflecting the most essential characteristics in the range, the pooling layer can effectively reduce the signal size. The pooling area needs to be set according to the actual situation. The pooling layer function is

\[ H_j = f(\beta_j \cdot \text{down}(H_j') + b_j) \]  

Where, \( H_j \) is the \( j \)-th output feature map of the current layer, \( H_j' \) is the output feature map of the previous layer, \( \text{down}(\cdot) \) represents the pooling rules applied to the pooling layer, \( \beta_j \) is the \( j \)-th multiplicative offset of the layer, \( b_j \) is the \( j \)-th additive offset of the layer.

2.3 Fully connected layer

The fully connected layer, the last layer before the output layer, is generally used to conduct classification or regression. In the modeling of the migration process, the fully connected layer can also play the role of "firewall", especially in the greater difference between the source domain and target domain cases. The connection layer function is given by

\[ y = f(W^T x + b) \]  

Where \( y \) is the constant output of the layer, \( W \) is the receptive field of the layer, \( x \) represents the input signal of the layer, while \( b \) represents the additive bias of the layer. In this article, the role of the fully connected layer is to classify bearing faults, so a Softmax function is employed as the activation function for this layer.

3 System Architecture and Implementation

This architecture of the proposed bearing fault diagnosis system is illustrated in Fig. 2. The system consists of several WSN end nodes, one WSN coordinator, and a centralized computer. Star topology, IEEE 802.15.4 and ZigBee protocols are used for the WSN. The end nodes and coordinator are used to collect and transmit bearing vibration signal to the centralized computer, while the centralized computer is employed to achieve bearing fault diagnosis based on CNN. This architecture fully utilizes the advantages of WSNs and the computing power of the centralized computer to obtain accurate fault diagnosis result.
4 Experimental Validation

4.1 Experimental setup and procedure

A series of experiments was conducted to evaluate the proposed bearing fault diagnosis method. The experimental setup used in this research is shown in Fig. 3. One WSN end node transmits the vibration data to the coordinator through ZigBee3.0 protocol, which sends the data to the host computer via USB. Then, the developed fault diagnosis APP on the host computer receives and displays vibration data, completes bearing fault diagnosis based on CNN, and gives the final results.

In this project, two NXP JN5169 carrier boards are used as the basic hardware platform for the end node and coordinator. The carrier board includes an NXP JN5169 microcontroller and its peripherals. JN5169 microcontroller supports IEEE 802.15.4
and ZigBee3.0 protocols, and it integrates a 32-bit RISC architecture with 512 kB Flash, 4 kB EEPROM, and 32 kB RAM, also a fully compliant 2.4 GHz IEEE 802.15.4 transceiver [35]. In addition, JN5169 contains a 128-bit AES security processor to ensure the security during data transmission. The JN5169 is chosen, representing a state-of-the-art small WSN node suitable for industrial application.

As this paper mainly explores the feasibility of bearing fault diagnosis using WSN and CNN, instead of building up a testbed, this research uses the data set released by the Case Western Reserve University Bearing Data Center [36] and widely used in bearings fault diagnosis fields [4, 37–40]. This data set contains vibration data under fan end fault (FE), driver end fault (FE), and normal state (NOR). Both DE and FE include outer ring fault (OR), inner ring fault (IR) and rotor fault (B). Each fault has three different fault depths, namely 7, 14, and 21 mils.

In this experiment, the bearing vibration data under a rotating speed of 1730 r•min$^{-1}$ and a load of 1hp are used to evaluate the feasibility of the proposed approaches. The vibration data is from seven bearing operating conditions, namely normal (NOR), driving-end rotor fault (DE-B), driving-end outer ring fault (DE-OR), driving-end outer ring fault (DE-OR), fan-end rotor fault (FE-B), fan-end outer ring fault (FE-IR) and fan-end outer ring fault (FE-OR). 700 sets of vibration data, 100 sets for every operating condition, are employed. Every data set includes 1024*3 sample points, namely the vibrations at base plate, the drive end and the fan end of motor.

The implementation of the proposed bearing fault diagnosis algorithm based on CNN is shown as Fig. 4. It is code using MATLAB on a laptop with Intel i3-3240 3.4GHz CPU and 8GB RAM. Part of the vibration data sets are used to train and optimize the CNN. After obtained the optimized parameters of the CNN, these parameters are then embedded in the fault diagnosis APP. Finally, the remained data sets are employed for verifying the effectiveness of the proposed method.
The interface of designed fault diagnosis APP is shown as Fig. 5. As an example, Fig. 5 (a) is for normal condition and Fig. 5(b) is for DE-B condition. The upper portion of the panel is used to display vibration signal curves. The serial port setting including the transmission rate, data bits, input buffer size, and output buffer size is on the lower left corner. The control buttons for “Input Data” and “Fault Diagnosis” and result indication for data transmission and bearing working status are located on the lower right corner.

4.2 Experiment results

Vibration data transmission: In this experiment, the selected vibration data sets are loaded to the end node in advance to verify the data transmission capability of the proposed system. These data sets are then transmitted to the coordinator wirelessly. When the coordinator receives the data, it sends the data to the host computer by USB immediately. Experimental results show that the proposed bearing fault diagnosis is able to transmit the vibration data wirelessly. Typical vibration signal waveforms for the seven bearing working conditions are given in Fig. 6. The red, blue, and green curves are the vibration signal measured at the base plate, the drive end and the fan end of motor, respectively.
Fault diagnosis accuracy: Considering the accuracy of fault diagnosis and calculation load, the data set used in this experiment consists of 1024*3 sample points, namely three 1024-points vibrations signals that measured at base plate, the drive end and the fan end of motor. Therefore, this paper constructs a basic 2D-CNN consisting of an input layer, an output layer, a convolutional layer, a pooling layer, and a fully connected layer. In order to prevent over-fitting, a dropout layer with a probability of 0.5 is added following the pooling layer during the training process of the CNN and closed in the fault diagnosis process. The main parameters of convolution layer and pooling layer used in this experiment are given in Table 1. The learning rate, maximum training node, and the mini-batch size of proposed CNN model is set as 0.01, 20, and 5, respectively.

Table 1. Parameters of convolutional layer and pooling layer

<table>
<thead>
<tr>
<th>Number of layers</th>
<th>Receptive width</th>
<th>Receptive height</th>
<th>Step size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Convolutional layer 1</td>
<td>5</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Pooling layer 1</td>
<td>10</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

Shown as Table 2, the output of the CNN is number 0 to 6 representing the seven bearing working conditions. To achieve accurate fault classification, the Softmax function is used as the output function of the fully connected layer.

Fault type and its corresponding label

<table>
<thead>
<tr>
<th>Fault type</th>
<th>NOR</th>
<th>DE-B</th>
<th>DE-IR</th>
<th>DE-OR</th>
<th>FE-B</th>
<th>FE-IR</th>
<th>FE-OR</th>
</tr>
</thead>
<tbody>
<tr>
<td>CNN output</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
</tbody>
</table>

Like most neural networks, the fault diagnosis accuracy of CNN depends on the number and quality of the training data set. Therefore, various training-testing data ratio, namely 7:3, 5:5, 3:7, and 1:9, are adopted in this experiment to evaluate the fault diagnosis accuracy of the proposed method for the different training data sets size.
The Classification results of the proposed method with various training-testing data ratio are illustrated by confusion matrix in Fig. 7. It can be seen that when the ratio is 7:3 and 5:5 the fault diagnosis accuracy of the proposed method reaches 100%. When the ratio is 3:7 and 1:9 the fault diagnosis accuracy drops to 99.80% and 97.6%. For the ratio of 7:3, one FE-B data set is incorrectly classified as FE-IR. However, its location, fan end, is still correct. For the ratio of 1:9, eight DE-OR data set is misdiagnosed as DE-IR, one FE-B data set is misdiagnosed as FE-IR, while six FE-OR data set is misdiagnosed as FE-IR. Although fault diagnosis accuracy significantly decreases for the ratio of 1:9, the classification of fault location is still right.

Fig. 7. Confusion matrix with various training-testing ratios
5 Conclusion

In this paper, a novel fault diagnosis method based on WSN, CNN has been proposed, and evaluated using the vibration data sets of seven motor operating conditions that released by the Case Western Reserve University Bearing Data Center. The experimental results show: (1) the proposed bearing fault diagnosis based on WSNs is able to transmit and display the vibration data smoothly; (2) the fault diagnosis accuracies of the presented method using CNN reach 100% when the ratios of training data against testing data are 7:3 and 5:5, while the accuracies reach 99.8% and 97.6% for the ratios of 3:7 and 1:9.

6 Acknowledgement

The Natural Science Foundation of Hebei Province, China (Grant No. supported this work F2016502104), the Scientific Research Foundation for the Returned Overseas Chinese Scholars by The Ministry of Education of the People's Republic of China, and the Fundamental Research Funds for the Central Universities of China.

7 References


8 Authors

Liqun Hou (corresponding author) is an Associate Professor in the School of Control and Computer Engineering, North China Electric Power University, Baoding 071003, China. E-mail: houliqun@ncepu.edu.cn

Zijing Li is currently working toward the M. Eng. degree at North China Electric Power University, Baoding 071003, China. E-mail: lizijing9602@163.com

Huaisheng Qu is currently working toward the M. Eng. degree at North China Electric Power University, Baoding 071003, China. E-mail: quhuaisheng@163.com