A fault diagnosis equipment of motor bearing based on sound signal and CNN

Dong Li¹, Binbin Li¹*, Chaoqun Wang¹, Pengyu Cheng¹ and Bin Jiao¹
¹Shanghai DianJi University, Shanghai,200000,China
*Corresponding author's e-mail: libb@sdju.edu.cn

Abstract. As an important part of mechanical equipment, the motor bearing damage rate is very high. In order to realize the fast and accurate diagnosis of motor bearing faults, this paper designs a fault diagnosis equipment based on sound signals. First, perform wavelet transform on the collected sound signal, then use the spectrogram generated by the fast Fourier transform to preliminarily determine whether the motor bearing is faulty, and finally use the convolutional neural network model that has been imported into the processor to diagnose the faulty parts of the motor bearing, The accuracy rate is above 98.41%.

1. Introduction.

Motor bearings are prone to damage when the motor is running at high speed, which is of great significance to its health monitoring and fault diagnosis. Xu Aihua et al. [1] used a motor bearing fault diagnosis algorithm based on convolutional neural network and gated recurrent unit (C-GRU). Chen Zhiheng [2] proposed using EMD (Empirical Mode Decomposition) as a method of vibration signal feature extraction, using an improved particle swarm optimization (Particle Swarm Optimization, PSO) algorithm to optimize the BP neural network. Zhou Yongqiang et al. [3] proposed a pattern recognition method of compound multi-scale permutation entropy partial mean value feature extraction and GK clustering. Zhang Yi et al. [4] diagnosed and analyzed bearing faults based on the envelope demodulation method, combined with the bandpass filtering optimization algorithm of spectral kurtosis, and applied it to the fault diagnosis of escalator motor bearings. Yang Xiaozhen et al. [5] proposed a fault diagnosis method based on improved Bayesian classification. Downs et al. [6] proposed a multi-scale incentive attention convolutional neural network, which effectively overcomes the limitations of traditional CNN models for feature extraction. Ge Xinglai et al. [7] studied a new method of bearing fault diagnosis based on singular value energy spectrum and improved ELM. The above methods have achieved good results, but they need to take up a lot of computing resources. At the same time, the acquisition of vibration signals and current signals in industrial production is very inconvenient. Therefore, this article combines the embedded system and the deep learning strategy, can collect and analyze the sound signal on the spot, diagnose the fault of the motor bearing, and has a high accuracy rate.

2. System framework design

This article mainly designs and implements a motor bearing fault diagnosis system based on sound signals. The composition of the system is shown figure1.
System consists of hardware and software components, wherein the hardware includes a noise sensor, arm processor and a display screen, the application software is mainly used convolutional neural network.

3. Hardware part

3.1 Hardware structure

3.1.1 Main processor. NanoPi is a low-power ARM main control board. It contains UART, SPI, I2C, IO and other pin resources. It also has a parallel camera interface and a full-color LCD interface. The CPU runs at 400Mhz and the internal storage is 64M. 5v DC power supply, the software supports u-boot, Linux-4.1, Debian8 jessie, Rabbit linux and other systems. It is very suitable for application development such as the Internet of Things and man-machine interface.

3.1.2 Display. The display uses the LCD screen P43 designed and developed by Friendly Arm. It is a 4.3-inch resistive touch display module with a resolution of 480 x 272, supports one-line touch, adjustable backlight, and supports embedded systems such as Linux, WinCE, Android, and is driven Open source. The display screen is used to display the results of data processing, which is convenient for personnel to check and repair.

3.1.3 Noise sensor. The noise sensor uses the noise decibel meter BYZ-08, which uses imported components to calibrate the output signal TTL and IIC. The measurement range is 30~150dB, the resolution is 0.1dB, and the frequency range is 20Hz~12.5KHz. It has the characteristics of high detection accuracy and high sensitivity. Suitable for portable/hand-held testing equipment. According to the processor resources, the noise sensor selects IIC as the output signal.

3.1.4 microSD card. The choice of memory card is generally larger than 16GB. This system uses a 64GB micro memory card to make system files. Insert the made microSD card into the NanoPi, and use the microUSB cable to connect to the computer. The NanoPi will power on automatically and you will see the on-board The blue LED flashes, which indicates that the system has started to operate normally.

4. Software part

4.1 Sound signal preprocessing
Principle and application of wavelet transform. The wavelet transform decomposes the noisy signal layer by layer, and then adds it back layer by layer. Since noise cannot be decomposed, it is generally decomposed to the end, and a large part of it is noise. Therefore, the useful signals are generally already
in the upper layer, and the noise is concentrated in the lower layer, so as to filter the noise. The main steps: (1) Select the appropriate wavelet basis function, and perform the wavelet decomposition of the signal at the specified level (2) Process and reconstruct each decomposition level.

Figure 1 is the time domain diagram of the sound signal when the bearing is normal, and Figure 2 is the time domain signal after wavelet noise reduction. Through comparison, it can be clearly seen that the filtered sound signal has much less clutter, and the filtering effect is very good. The processed sound signal is used for Fourier transform and convolutional neural network.

4.2 Fast Fourier and preliminary fault diagnosis

Principle of Fast Fourier Transform. Fast Fourier Transform is a fast algorithm of Discrete Fourier Transform, abbreviated as fft. It is obtained by improving the discrete Fourier transform algorithm based on the odd, even, imaginary, and real characteristics of the discrete Fourier transform. Fourier analysis can transform the signal from the original domain (usually time or space) to the frequency domain for representation.

For the sequence $x(n) =\{x_0, x_1, x_2, \ldots, x_{N-1}\}$, $0 \leq n \leq N$ The expression of discrete Fourier transform for is

$$x^\wedge(k) = \sum_{n=0}^{N-1} x(n)e^{-\frac{2\pi i nk}{N}} \quad (k = 0, 1, \ldots, N - 1) \quad (1)$$

Among them, $e$ is the base of the natural logarithm, and $i$ is the imaginary unit. FFT quickly calculates this type of transformation by decomposing the DFT matrix into the product of sparse factors.

In the laboratory, we collected a total of five sets of data, namely the sound signals collected when the bearing is normal, the small cut in the bearing, the outer ring of the bearing is penetrated, the outer ring of the bearing is damaged, and the outer ring of the bearing is worn. The acquisition time of each group of data is about 100s, and we convert the analog signal of sound into digital signal through code. In order to reduce the impact of environmental noise on the data, it is necessary to perform wavelet noise reduction on the data, and perform the noise reduction on the data after noise reduction. The fast Fourier transform can get the frequency spectrum under normal conditions and various bearing faults, showed in Figure 4 to Figure 8.
By comparing the five frequency spectrograms, it can be found that the amplitude corresponding to the frequency of 635Hz under normal conditions of the bearing is quite different. Therefore, it can be preliminarily judged whether the bearing is faulty through the spectrogram, and the threshold value can be set on the human-computer interaction interface. When a threshold value different from this occurs, a fault alarm signal is sent.

4.3 Convolutional Neural Network and Fault Feature Recognition

4.3.1 Building a convolutional neural network model. Convolutional Neural Network (CNN) is a feedforward neural network, which is mainly composed of an input layer, multiple convolutional layers, multiple pooling layers, a fully connected layer, and an output layer.

The composition of CNN: Input layer. The quality of the data read directly affects the performance of the entire CNN model. If the input is a picture that passes through the input layer, one or more matrices will be obtained. The size of the matrix is proportional to the pixel size of the picture. Since the sound signal is one-dimensional and cannot be input directly, the pre-processed sound signal data is reconstructed into a 32*32*1 black and white picture, and the number of signals in each data matrix is 1024 (the number of data includes the amount of data collected by one rotation of the motor bearing). There are five types of bearing failures, so the data type needs to be labeled, as shown in Table 1.

| label | 0  | 1              | 2              | 3              | 4              |
|-------|----|----------------|----------------|----------------|----------------|
| Fault type | normal | Internal small incision | Outer ring runs through | Outer ring wear | Outer ring one word wound |

Table 1. Fault label table

Convolution Layer and Pooling Layer. The convolutional layer is the core part of the network. The size of the convolution kernel in the convolutional layer is 3*3 or 5*5. Each node in the convolutional
layer passes through a specific convolution kernel and the input matrix. A small piece of convolution is obtained, and the detailed features of the input picture can be obtained through the convolution kernel. The principle is as follows

\[ x^l_j = ReLU \left( \sum_{i \in \mathcal{S}_j} x^{l-1}_i \ast \omega^l_{ij} + b^l_j \right) \]  

(2)

\( \mathcal{S}_j \) represents the set of input feature maps, \( x^{l-1}_i \) represents the i-th input feature map of the l-1th layer, \( \omega^l_{ij} \) represents the convolution kernel from the i-th input feature map of the l-1th layer to the j-th output feature map of the l-th layer, \( b^l_j \) represents the offset value corresponding to the j-th output feature map of the l-th layer, \(*\) represents the convolution operation, and uses the ReLU nonlinear activation function to output the j-th feature map of the l-th layer \( x^l_j \).

Through the pooling layer, the number of nodes in the final fully connected layer can be further reduced, so as to achieve the purpose of reducing the parameters of the entire neural network. The principle is as follows,

\[ x^l_j = f \left( \beta^l_j down \left( x^{l-1}_i \right) + b^l_j \right) \]  

(3)

Where down() is the down-sampling function max or ave, and \( \beta \) is its coefficient.

Fully connected layer and output layer. After multiple rounds of convolutional layer and pooling layer processing, at the end of the CNN, 1 to 2 fully connected layers are generally used to give the final classification result. After several rounds of convolutional layer and pooling layer processing, it can be considered that the information in the image has been abstracted into features with higher information content. We can regard the convolutional layer and the pooling layer as the process of automatic image feature extraction. After the extraction is completed, a fully connected layer is still needed to complete the classification task. Through the Softmax function, the current sample belongs to different types of probability distribution problems, and the output formula is as follows,

\[ \text{Softmax}(y)_i = \frac{\exp(y_i)}{\sum_{j=1}^{n} \exp(y_j)} \]  

(4)

\( y \) is the output of the convolutional neural network.

In summary, CNN uses convolution and pooling to extract the abstract features of the processed sound signal, and then uses the fully connected layer to classify the learned effective features. The model parameters of CNN are shown in Table 2.

|structure| parameter | value | structure | parameter | value |
|---|---|---|---|---|---|
|Conv1| Conv1_input | 32*32*1 | Conv1_input | 14*14*1 |
|Kernel_size| 5*5 | Conv2| Kernel_size | 3*3 |
|strides| 1 | strides | 1 |
|Conv1_output| 28*28*1 | Conv1_output | 12*12*1 |
|Activation_fn| Relu | Activation_fn | Relu |
|Pool2| Kernel_size | 2*2 | Pool2| Kernel_size | 2*2 |
|strides| 2 | strides | 2 |
|Pool2_out| 14*14*1 | Pool2_out | 6*6*1 |
|Fully connected| Dense | 512 | Softmax | Output | 5 |
|Activation_fn| Relu | Activation_fn | Softmax | |

4.3.2 Result analysis

In 4.2, comparing the four fault frequency spectrums with the normal frequency spectrum, it can be determined that the bearing is faulty, but the fault type cannot be distinguished. Therefore, the filtered
sound signal is divided into normal, inner small cuts, and marks outer ring penetration, outer ring wear, and outer ring penetration. Each set of label data is 1.8 million, of which 70% is used as the training set and 30% is used as the test set. The training model is used as the processing system, and the test set is used to enter the training model for testing. The five features are successfully classified through calculations and Figure 9 and Figure 10 are obtained.

![Figure 9. Failure feature visualization](image)

![Figure 10. Confusion matrix](image)

The five colors in the figure9 represent the five fault features. It can be seen that the convolutional neural network has a very obvious effect on various fault classifications, and the recognition rate of various fault features can be obtained through the confusion matrix diagram.

| Bearing failure | Internal small incision | Outer ring runs through | Outer ring wear | Outer ring one word wound |
|----------------|-------------------------|-------------------------|----------------|--------------------------|
| Accuracy(%)     | 100                     | 100                     | 98.41          | 100                      |

Through the deep learning method, the accuracy of the fault recognition rate is as high as 98.41% to 100%. In practical applications, the fault type of the motor bearing can be effectively judged according to the sound signal.

5. Conclusion
In this paper, after collecting sound signals through experiments, fast Fourier transform and convolutional neural network are used to process the sound signals, and the fault type of the motor bearing can be judged successfully and accurately. The system can be customized as a specific portable fault identification device. In the future development, the combination of embedded system and deep learning will be more used in the maintenance of industrial equipment, which greatly reduces the difficulty of maintenance.

References
[1] Xuai Hua, Du Yang, Duan wave, Xuhan Feng motor bearing fault diagnosis based on C-GRU [J] in combination with the automatic machine processing technology, 2021 (02): 47-51.
[2] Chen Zhiheng, Song Dongli, Zhang Weihua, Deng Cong, Dong Jianxiong. Motor bearing fault diagnosis based on EMD and improved PSO_BP [J]. Measurement and control technology, 2020,39 (11): 33-38 + 125.
[3] Zhou Yongqiang, Bu Wenshao. Multiscale arrangement entropy feature extraction of motor bearing fault with GK identification [J]. Combined machine tool and Automation Processing Technology, 2021 (04): 70-74.
[4] Zhang Yi. Application of envelope demodulation method in escalator bearing diagnosis [J]. Equipment Management and Maintenance, 2021 (05): 152-154.
[5] Yang Xiaozen, Zhou continued, Deng Juming. Study on Motor Bearing Fault Diagnosis System Based on Improving Bayesian Classification [J]. Machine Tool & Hydraulic, 2020,48 (20):
172-175.

[6] Downs, Chen Xinchu, Zheng Song. Motor bearing fault diagnosis based on attention and multiscale convolutional neural networks [J]. Electrical Technology, 2020, 21 (11): 32-38.

[7] Ge Xinglai, Zhang Xin. Using ular energy spectrum and improved ELM [J]. Motor and Control Journal, 2021, 25 (05): 80-87.