

## Research Article

# Motor Fault Diagnosis Algorithm Based on Wavelet and Attention Mechanism

Yong Yan <sup>1</sup>, Qiang Liu,<sup>2</sup> and Xiao qin Gao<sup>3</sup>

<sup>1</sup>The College of Electronics Information and Automation, Aba Teachers University, Aba 623002, China

<sup>2</sup>The College of Electronics and Information Engineering, Sichuan University, Chengdu 610065, China

<sup>3</sup>The Department of Information Engineering, Sichuan Technology Business College, Dujiangyan 611830, China

Correspondence should be addressed to Yong Yan; 20069623@abtu.edu.cn

Received 18 May 2021; Accepted 22 June 2021; Published 8 July 2021

Academic Editor: Mu Zhou

Copyright © 2021 Yong Yan et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In order to improve the maintenance efficiency of the motor and realize the real-time fault diagnosis function of the motor, a motor fault diagnosis algorithm based on wavelet and attention mechanism is proposed. Firstly, the motor vibration signal is decomposed by wavelet transform, and the high-frequency signal is denoised to improve the signal-to-noise ratio. Secondly, the frequency band and time dimension after wavelet decomposition are taken as input data, the convolution neural network is used to fuse the frequency band features of data, and the bidirectional gated loop unit is used to fuse the time series features. Then, the attention mechanism is used to adaptively integrate the features of different time points. Finally, motor fault diagnosis and prediction are realized by classifier recognition. Experimental results show that, compared with the existing deep learning fault diagnosis model, this method has higher diagnosis accuracy and can accurately diagnose the running state of the motor.

## 1. Introduction

Motor is the most commonly used device to drive all kinds of machinery and industrial equipment. In order to ensure the safe and stable operation of the system, staff often needs to carry out regular overhaul and maintenance of the system [1]. However, this manual method not only requires a lot of manpower, material, and financial resources but also ensures that the fault can be eliminated in time. With the development of computer technology, large- and medium-sized motors are equipped with computer-centered condition monitoring and fault diagnosis system, in order to prevent or repair the functional failure or local failure before the failure, minimize the loss, and prevent the occurrence of catastrophic accidents [2].

The motor system is composed of multiple functional modules, and its faults are often uncertain and nonlinear. In the motor system, any small deviation of function module may lead to the collapse of the whole system. Therefore, it is particularly important to find a reasonable and effective fault diagnosis method. The common faults of large- and

medium-sized motors include mass imbalance of moving parts, bearing wear, improper matching of moving parts, oil film oscillation, shaft misalignment, and cracks. The characteristics of these faults are most obvious in the frequency of vibration signal. Therefore, correlation analysis and spectrum analysis are usually used to monitor and diagnose the state of motor vibration signal. These methods usually deal with the data first, extract the characteristic parameters, and then refer to some fault diagnosis specifications [3].

At present, motor fault diagnosis methods mainly include quantitative analysis and qualitative analysis. Quantitative analysis methods include analytical model-based method and data-driven method. At present, the fault characteristic signal processing of rotor and bearing in the motor system usually uses the data driven by quantitative analysis method. In reference [4], the Prony method is combined with band selection and thinning technology to amplify and analyze the small frequency band near the characteristic frequency of broken rotor bars, which improves the calculation accuracy. However, in the low noise background, the diagnosis effect is not ideal. In reference [5], a rotor fault

diagnosis method based on improved matrix beam filtering and detection is proposed. The improved matrix beam algorithm can effectively overcome the influence of noise and has better resolution for short-term data. In reference [6], the Park vector mode transformation technique and rotation invariant signal parameter estimation technique are proposed. The Park vector mode transformation technology is used to eliminate the influence of power frequency components and reduce the amount of calculation on the premise of ensuring the accuracy. In reference [7], a fault diagnosis technology of motor bearing based on wavelet transform and Hilbert transform is proposed. Hilbert transform is used to demodulate the modulation signal, and the effective analysis of fault characteristic signal is realized. At the same time, this method introduces the thselect function into the wavelet transform and uses the adaptive ability of this function to adjust the threshold under different noise conditions. It can eliminate the high-frequency signal and improve the adaptive ability of the algorithm. However, this method ignores the noise interference of low-frequency signal, which will still lead to misjudgment and omission. Although the quantitative analysis method based on data can carry out fault analysis according to characteristic components, the accuracy of fault detection is often affected by interference signals. Therefore, the quantitative analysis method is difficult to meet the requirements of motor system fault diagnosis.

Qualitative analysis is a method of building system model based on internal knowledge of the system [8], mainly including graph theory method, expert system, and qualitative simulation. Petri net is a graph theory method, which uses the relationship between the components in the target system to establish a directed graph combination model and accurately deals with the relationship among the sequence, concurrency, and conflict of discrete events [9]. In recent years, with the research of scholars at home and abroad, the application of Petri net in motor system fault detection is gradually increasing and has achieved good results. Aiming at the uncertainty in the process of fault propagation, a probabilistic transition method of fault Petri nets is defined in reference [10], which solves the shortcoming that the fault diagnosis process of Petri nets only focuses on the state of the repository. But the way of probability change is not adaptive. In reference [11], a colored Petri net model is established by using colored Petri nets and modeling tools cpntools, which improves the efficiency of fault diagnosis and the intuitiveness of the model. In reference [12], Petri net and fuzzy reasoning are combined to solve some stochastic and uncertain fault problems effectively. It greatly improves the ability to locate the fault source in fault diagnosis. In reference [13], fuzzy Petri net (FPN) carries out strict mathematical reasoning and gives the reasoning process of matrix. However, the acquisition of its weight still depends on the experience of experts, and its adaptability is poor.

In the practical use of the motor system, the components of motor vibration signal are more complex. At the same time, due to the fluctuation of working conditions, the vibration signal has strong time varying, and the signal characteristics at different times are very different. Therefore, aiming

at the characteristics of complex frequency information and strong time varying of motor vibration signal, a motor fault diagnosis algorithm based on wavelet and attention mechanism is proposed. The wavelet denoising method is used to remove the original signal noise and improve the signal-to-noise ratio. Convolution neural network and bidirectional gating loop unit are used to fuse time-frequency information. The attention mechanism module is used to dynamically weigh the features of different time points in the time dimension, so as to more accurately predict and diagnose motor faults.

## 2. Method

*2.1. Principle of Motor Bearing Vibration.* The fault frequency of motor bearing detection is divided into high-frequency detection and low-frequency detection. When a part of motor bearing fails, pulse signal is generated due to periodic motion. Most of the low-frequency signals are generated by the rotation of the fault parts, while the high-frequency signals are generated by the natural frequencies of other parts when the bearing fails. The natural frequency range of bearing components is 1 kHz~10 kHz, and its level is mainly affected by material, structure, processing technology, assembly, and other factors.

The natural frequency of bearing rolling element is

$$f_1 = 0.212 \frac{Hg}{r\gamma}. \quad (1)$$

The natural frequency of the inner and outer ring of the bearing is

$$f_2 = \frac{m(m^2 - 1)}{2\pi(d_a/2)^2 \sqrt{m^2 + 1}} \sqrt{\frac{HIg}{\gamma A}}, \quad (2)$$

where  $H$  is the modulus of elasticity.  $g$  is the acceleration of gravity.  $r$  is the radius of the steel ball.  $\gamma$  is the material density.  $I$  is the moment of inertia of the cross section of the bearing ring.  $m$  is the vibration order.  $A$  is the cross-sectional area of the bearing ring.  $d_a$  is the neutral shaft diameter of bearing cross section.

When the outer ring of the bearing fails, its characteristic frequency is

$$f_3 = \frac{s}{2 \times 60} \left( 1 - \frac{d}{D} \cos \alpha \right) N. \quad (3)$$

When there is bearing inner ring failure, its characteristic frequency is

$$f_4 = \frac{s}{2 \times 60} \left( 1 + \frac{d}{D} \cos \alpha \right) N, \quad (4)$$

where  $s$  is the inner ring speed.  $d$  is the diameter of roller.  $N$  is the number of rollers.  $D$  is the diameter of bearing pitch circle.  $\alpha$  is the contact angle.

**2.2. Wavelet Transform and Noise Reduction.** Wavelet transform is a relatively new mathematical theory, which has outstanding characteristics of time-frequency localization. The wavelet transform is an integral transform, while the inner product of the continuous wavelet transform  $f(t)$  and the stretching coefficient of the wavelet function is

$$W_f(x, y) = \int_{-\infty}^{+\infty} f(t)p(x, y, t)dt, \quad (5)$$

where

$$p(x, y, t) = |x|^{-1/2}h\left(\frac{t-y}{x}\right), \quad (6)$$

where  $p(t)$  is the parent wavelet, which is the normalized coefficient.  $y$  and  $x$  are the translation and stretching factors of the wavelet function, respectively. The resolution of the local structure can be changed by adjusting the translation factor and the stretching factor. The excitation function of the hidden layer of the wavelet neural network [14] is the wavelet basis function, which is obtained by translating and stretching the parent wavelet function.

The principle of wavelet transform denoising is to decompose the collected signal to obtain low-frequency signal and high-frequency signal. The noise is generally contained in the obtained high-frequency signal so that the threshold value can be set to filter out the noise and the remaining signal can be reconstructed to achieve noise reduction.

**2.3. Fusion of Time-Frequency Features.** According to the characteristics of complex and time-varying frequency information of motor vibration signal, convolution neural network and bidirectional gating loop unit are used to fuse the time-frequency information. Then, the attention mechanism is used to dynamically weigh the time-varying characteristics, and the fault diagnosis of planetary gearbox is realized.

Wavelet packet is another important development of wavelet theory in signal processing. It has the characteristics of multidimensional and multiresolution analysis and can provide more complex signal analysis methods [15]. The signal can be decomposed into a series of frequency bands layer by layer. It further decomposes the undivided high-frequency part of the wavelet analysis and selects the corresponding frequency band according to the characteristics of the analyzed signal to match the signal spectrum. Therefore, the wavelet packet decomposition method is used to decompose the original vibration signal into two-dimensional signals in time domain and frequency domain. For nonstationary signals such as motor vibration signals, wavelet packet decomposition can effectively represent the energy information in different frequency bands and different times and can reflect the rapid change of information in each frequency band in a short time. The process of wavelet packet decomposition is shown in Figure 1. The two-dimensional wavelet coefficient matrix obtained from the original signal decomposition is used as the input of the deep learning model.

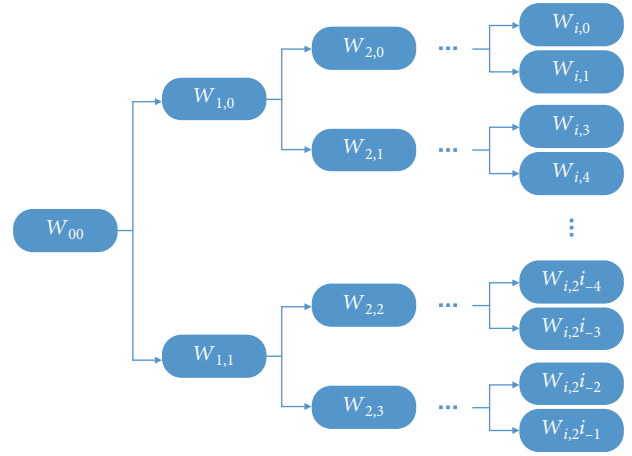


FIGURE 1: Schematic diagram of wavelet packet decomposition.

**2.3.1. Frequency Band Feature Fusion.** The vibration information of the motor contains multiple vibration frequencies, which are expressed in the wavelet packet decomposition coefficient matrix of the vibration signal. The amplitude in each frequency band reflects the working state of different structures, and the related information between different frequency bands can reveal the deep health status. Therefore, it is necessary to extract and fuse the band features in the signal and use these features to solve the problem of complex frequency components.

CNN is a neural network structure with local sensing ability. It scans the input data through the convolution check formed by locally connected network nodes and extracts the detailed features of the data. CNN has been widely used in image recognition and other fields because of its powerful feature learning ability. The calculation process of CNN can be expressed as follows:

$$x_j^l = p\left(\sum_i x_i^{l-1} * d_{ij}^l + v_j^l\right), \quad (7)$$

where  $x_j^l$  is the  $j$ th feature graph output at the  $l$  layer and  $x_i^{l-1}$  is the  $i$ th feature graph of layer  $(l-1)$ .  $d_{ij}^l$  is the convolution kernel of the  $i$ th input feature graph connected with the  $j$ th feature graph.  $v_j^l$  is the offset term.  $*$  is the convolution operation.  $p(\bullet)$  is the activation function. In this paper, the Rectified Linear Unit (ReLU) is adopted, and its function expression is as follows.

$$p(x) = \max\{0, x\}. \quad (8)$$

In this paper, one-dimensional CNN is used to fuse the wavelet packet characteristic matrix of planetary gearbox vibration signal in frequency domain. Through multiple one-dimensional convolution kernels, the feature information of signals in different frequency bands is fused into new features. Therefore, the model can adaptively learn the complex frequency information in the vibration signal of planetary gearbox and reorganize the original frequency

information into new features. Now, the one-dimensional convolution kernel will scan along the time axis of the wavelet packet characteristic matrix, so the output data still contains the characteristics of the time dimension.

**2.3.2. Time Series Feature Fusion.** In the vibration signal, in addition to the correlation characteristics between different frequency bands, the amplitude variation characteristics of each frequency band with time also reflect the health status of the planetary gearbox. However, CNN is mainly used to extract the local features of signals, and it is difficult to perceive the overall change features in long sequences. For the vibration signal of planetary gearbox, it is necessary to integrate the time dimension features to improve the recognition ability of its health state.

Strobe recursive unit (GRU) is a kind of improved recurrent neural network, which is specially used to process time series data and can maintain the change characteristics before and after the series, so it can be used to integrate the characteristics of data in time dimension. The internal structure of GRU is shown in Figure 2.

Its calculation process can be expressed as

$$h_t = z_t \otimes \tilde{h}_t + (1 - z_t) \otimes h_{t-1}, \quad (9)$$

$$r_t = (W_r x_t + U_r h_{t-1} + v_r) \sigma_g, \quad (10)$$

$$z_t = (W_z x_t + U_z h_{t-1} + v_z) \sigma_g, \quad (11)$$

$$\tilde{h}_t = (W_h x_t + U_t (r_t \otimes h_{t-1}) + v_h) \sigma_h, \quad (12)$$

where  $W \in R^{d \times k}$ ,  $V \in R^{d \times d}$ , and  $b \in R^d$  are trainable parameters.  $\otimes$  is the cross product operation.  $x_t$  and  $h_t$  are input and output vectors, respectively.  $z_t$  and  $r_t$  are update and reset gate vectors, respectively.  $\sigma_g$  and  $\sigma_h$  are activation functions.

In this paper, Hard Sigmoid and Tanh are used as activation functions, respectively. The Hard Sigmoid function is expressed as follows.

$$p(x) = \begin{cases} 0, & x < -2.5, \\ 0.2x + 0.5, & -2.5 \leq x \leq 2.5, \\ 1, & x > 2.5. \end{cases} \quad (13)$$

The Tanh function is expressed as follows.

$$p(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}. \quad (14)$$

GRU unit receives a group of data each time, feeds back the state of the previous time, and merges it with the current data as input. After the output is calculated, the cycle is used for the next calculation until the last time scale, so as to obtain the overall change characteristics of the input data in the time dimension. Compared with the standard RNN structure, GRU can selectively learn and forget the features in time series by controlling the state of update gate and reset gate and has the ability of long-term memory. In order to fur-

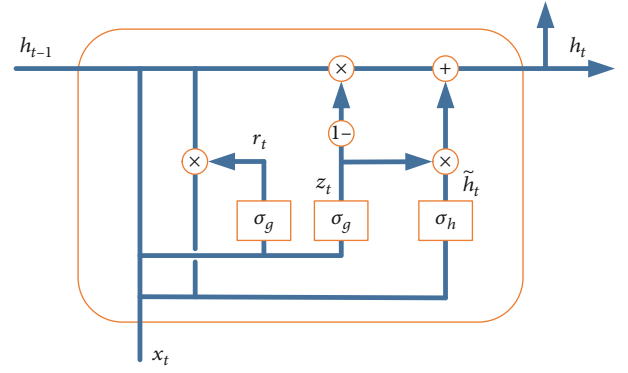


FIGURE 2: The internal structure of GRU.

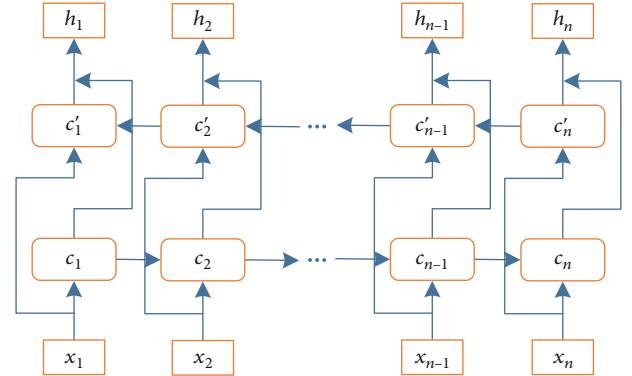


FIGURE 3: Schematic diagram of time sequence expansion of bidirectional GRU.

ther enhance the time feature fusion ability of the model, this paper adopts the bidirectional GRU structure, and its time series expansion is shown in Figure 3. As can be seen from Figure 3, bidirectional GRU can obtain the time series features before the current time through forward scanning and obtain the time series features after the current time through backward scanning. Compared with one-way GRU, the efficiency of feature fusion is greatly improved. We can get the time series characteristics after the current time. Compared with one-way GRU, the efficiency of feature fusion is greatly improved.

**2.4. Dynamic Weighted Fusion Based on Attention Mechanism.** Under the influence of alternating working conditions, the vibration signal of planetary gearbox on display has strong time-varying characteristics. Therefore, the signals collected at different times have different characteristics, and there are great differences between different samples, which will greatly reduce the generalization ability of deep learning model.

The attention mechanism in deep learning is a bionic concept, which simulates the process of human focusing on the specific part of the observed object and improves the ability of information acquisition by focusing on different parts of the object information. Based on the deep learning model, attention mechanism is added to dynamically fuse the output features of bidirectional GRU layer at different times so that



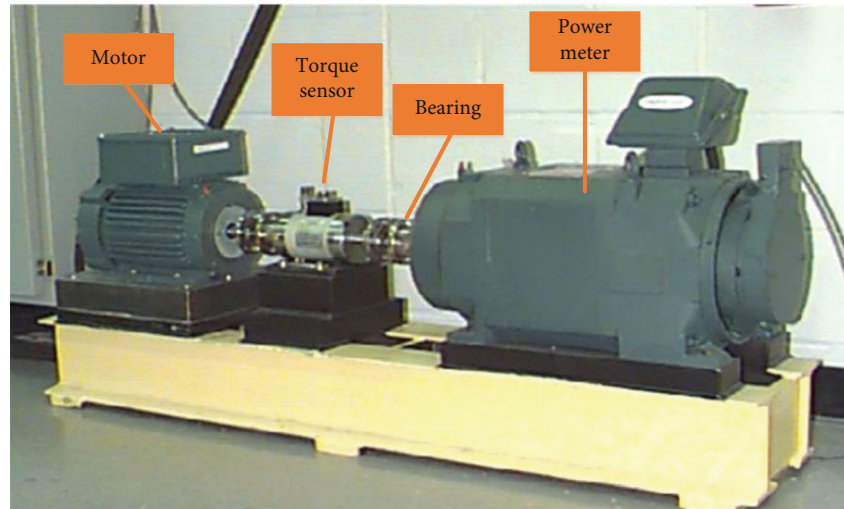


FIGURE 4: The bearing test system.

the model can adaptively focus attention at different times when facing signals with obvious differences in time features and solve the nonstationary vibration signal caused by sample difference.

The structure of this paper is as follows:

$$e_t = p \left( \sum_i W_t^i h_t^i + v_t^i \right), \quad (15)$$

$$\alpha_t = \frac{\exp(e_t)}{\sum_{k=1}^T \exp(e_k)}, \quad (16)$$

$$s_t = \alpha_t h_t, \quad (17)$$

where  $p(\bullet)$  is the activation function. The hidden state sequence of bidirectional GRU is entered into the full connection layer, resulting in a dynamic weight  $\alpha_t$ .  $s_t$  is the feature vector after dynamic weighted fusion.

The neural network outputs a weight  $\alpha_t$  for the feature at every moment by learning the information of the input data itself, which represents the “attention” of the network. The size of the weight represents the degree of focus of attention. Since the vibration signals of the planetary gearbox have different characteristics at different times, the attention machine can help the neural network to dynamically give weight to the characteristics at different times. Thus, it can more effectively improve the utilization degree of useful information and reduce the sensitivity of the mode to the input data.

**2.5. The Process of the Proposed Algorithm.** The motor fault diagnosis algorithm based on wavelet and attention mechanism can be divided into four steps: signal acquisition and truncation, sample wavelet packet decomposition, deep learning model design and training, and fault diagnosis testing.

- (1) Signal acquisition and truncation: the vibration sensor is used to collect the vibration signal of the plan-

etary gearbox, and the collected signal is truncated to the same length to get the sample

- (2) Sample wavelet packet decomposition: one-dimensional vibration sample signal is decomposed by wavelet packet to obtain two-dimensional time-frequency diagram samples, which are divided into the training set and the test set
- (3) Design and training of deep learning model: use training set data to train model parameters, and constantly adjust super parameters in the training process to achieve better performance
- (4) Fault diagnosis test: after the training of the model, the data of the training set is used for fault diagnosis test to verify the diagnosis performance of the model

According to the above methods, the deep learning model is constructed. Among them, batch normalization is to normalize the output data of the convolution layer. It can make the activation of neurons tend to normal distribution and avoid oversaturation. Therefore, it can alleviate the problem of gradient vanishing to a certain extent, accelerate the convergence speed, and improve the accuracy of the model.

The global average pool is used to compress the dimension of the fused feature data to maintain the significance of the feature. It can effectively reduce the number of network parameters, reduce the complexity of the network, and prevent overfitting. In this paper, we use global average pool operation after structure fusion. Statistics and dimension reduction of the fused features are carried out to facilitate the subsequent classification and recognition.

Shedding is an effective way to reduce over assembly. In this process, some weights are discarded randomly when the network parameters are updated in each training. Only retained weights are updated for backpropagation. This can reduce the excessive dependence on some neurons, reduce the complex adaptive relationship between neurons, and

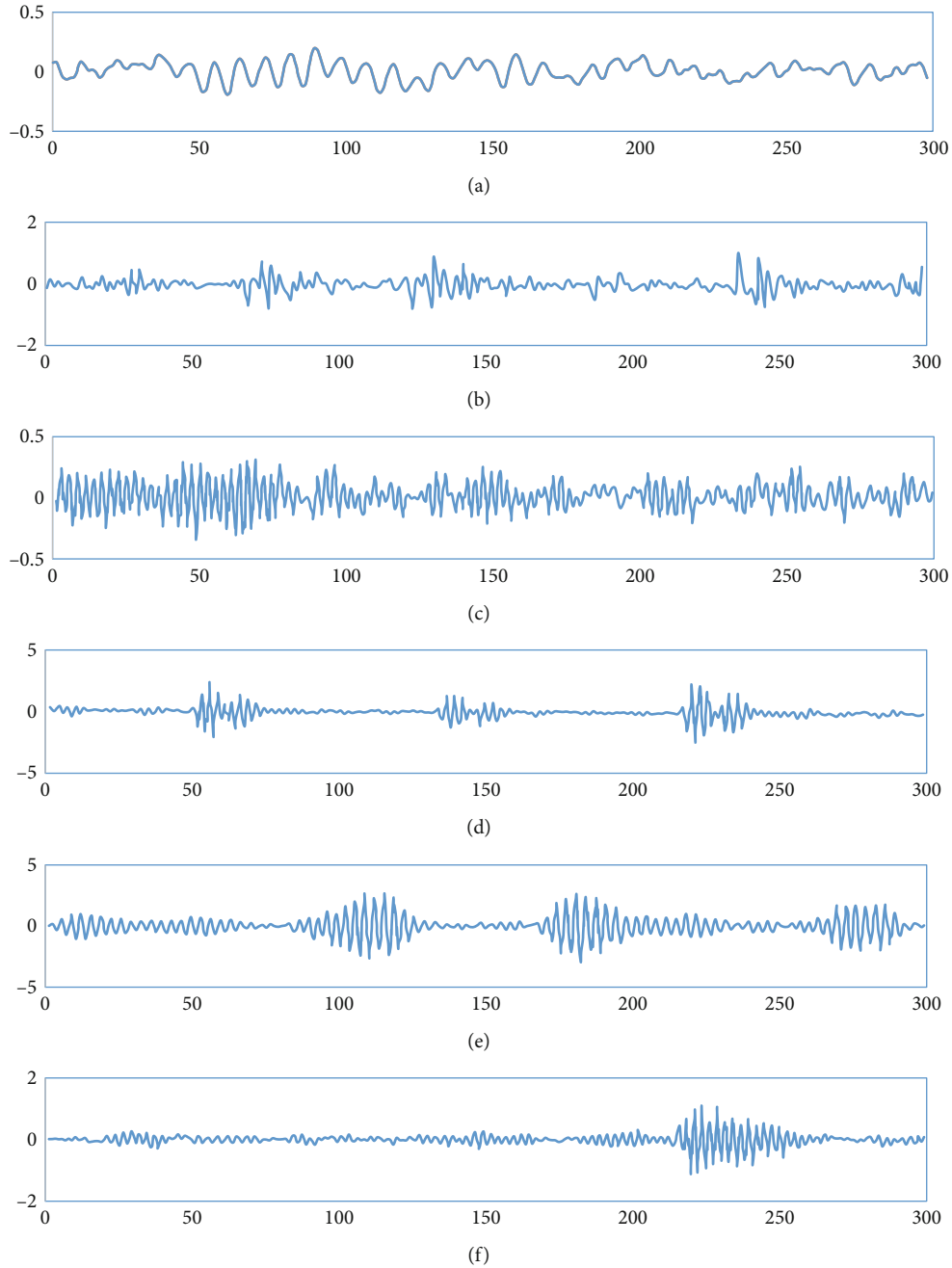


FIGURE 5: Examples of six vibration data.

improve the generalization ability of the network. In this paper, exit operation is set between all connection layers, which only takes effect during training, but not during testing.

The final output of the deep learning model generates probability distribution through Softmax classifier to judge the fault types of the input data. Its mathematical representation is as follows.

$$y_j = \frac{e^{x_j}}{\sum_{i=1}^M e^{x_i}}. \quad (18)$$

In this paper, cross entropy is selected as the loss function to calculate the error between the calculated value and the real value of the deep learning model and promote its back-propagation to update the parameters. The loss function is expressed as follows.

$$L = - \sum_{j=1}^M t_j \log y_j, \quad (19)$$

where  $M$  is the number of fault types,  $t_j$  is the true sample label, and  $y_j$  is an estimate of the output of Softmax.

The RMSProp optimization algorithm [16] was used as the optimizer for model training, and the parameters were iterated and updated in the process of backpropagation. It brings the model's predictions closer to the real value. RMSProp limits the oscillations of gradient descent during backpropagation, allowing the network to obtain more accurate results.

### 3. Experimental Results and Analysis

**3.1. Feasibility Verification of the Proposed Algorithm.** In this paper, experimental data from the Bearing Data Center of Case Western Reserve University [17] were used to test the effectiveness of the proposed method. Vibration signals were measured by the bearing test system, and the bearing test system is shown in Figure 4.

The bearing model is skf6203, the sampling frequency is 1.2 kHz, the rotating speed is 1797 r/min, and there are four kinds of bearings: normal bearing, inner ring fault bearing, in which the damage diameter is 0.53 mm, rolling element fault bearing, in which the damage diameter is 0.53 mm, and outer ring fault bearing, in which the damage diameter is 0.53 mm (including center @ 6:00, orthogonal @ 3:00, and opposite @ 12:00). For each type of data, 120000 sample points are obtained, which means that the length of time series data is at least 120000. If it is directly used as the input of deep neural network, the training will be too large. Here, a group of vibration signals is grouped according to every 300 sampling points, and a total of 2400 datasets are obtained. Figure 5 shows an example of six vibration data.

In order to verify the performance of the model proposed in this paper, based on the same dataset, it is compared with the AlexNet [18], GoogleNet [19], and ResNet [20] deep learning methods that have been proposed at present.

The 400 hundred groups were selected as the test data, and 500, 1000, 1500, and 2000 samples were selected for training. The algorithm proposed in this paper was used to carry out diagnostic tests, and the comparison of diagnosis rate is shown in Figure 6.

As can be seen from Figure 6, with the increase of the number of training samples, the accuracy rate of the proposed algorithm is gradually increased.

**3.2. Actual Experimental Results and Analysis.** Field diagnosis tests are carried out on motor bearings, in order to further verify the effectiveness of the improved random forest algorithm and verify whether the diagnosis method proposed in this paper is accurate when the bearing fault signal is weak. The motor model is SEW-DV132M4, the speed is 1430 r/min, the sampling frequency is 1 kHz, the bearing model is SKF6209-2Z, and the vibration data collector is COMMTEST-VB7. There are 4 kinds of bearings: normal, inner ring fault (damage diameter of 10  $\mu\text{m}$ ), rolling body fault (damage diameter of 18  $\mu\text{m}$ ), and outer ring fault (damage diameter of 9  $\mu\text{m}$ ). The number of diagnostic samples was 1000, and the number of collected samples was 2000.

In order to reduce the influence of unexpected factors, the training set and the test set are divided randomly for 10 times. Each group of data was used for independent training

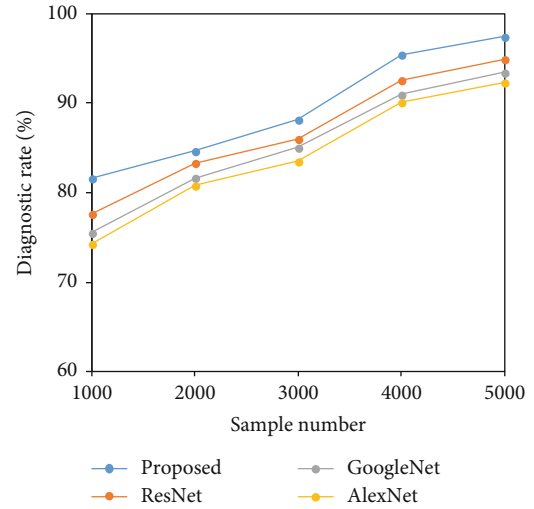


FIGURE 6: Results of diagnostic rate.

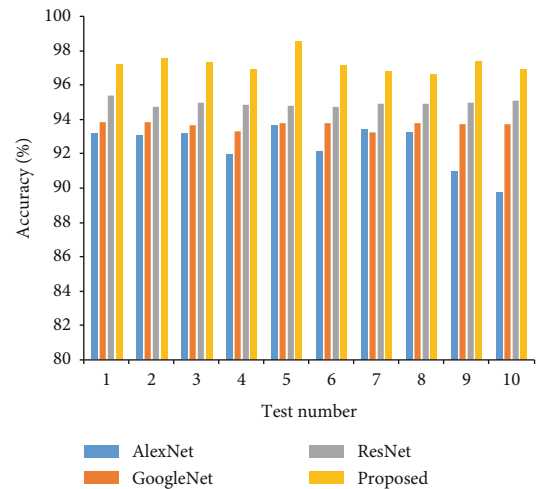


FIGURE 7: The results of 10 experiments.

TABLE 1: Experimental results.

Accuracy	Train	Test
AlexNet [18]	96.18	92.47
GoogleNet [19]	97.69	93.68
ResNet [20]	98.23	94.95
Proposed	99.51	97.27

and testing, and the experimental results were recorded, respectively. The results of 10 experiments are used as performance evaluation indexes for comparative verification. The specific results are shown in Figure 7. The average accuracy and standard deviation are shown in Table 1. The accuracy of the AlexNet model is poor. Because the model simply overlaps multiple convolution layers, it is difficult to effectively extract complex data features. The GoogleNet model adds convolution kernels of different sizes in each convolution layer, which has certain frequency feature extraction ability, and its accuracy is 1.21% higher than that of AlexNet.

ResNet adds residual structure on the basis of GoogleNet and realizes cross layer connection of convolution layer, and its accuracy is 1.27% higher than that of GoogleNet. However, due to the single network structure for feature extraction, the above models cannot make full use of the information contained in the signal, so the overall accuracy is low. The method proposed in this paper can fully combine the time-frequency characteristics of the signal and introduce the attention mechanism of dynamic weighting. Therefore, in each test, its accuracy is higher than other deep learning models. The above results show that, compared with other deep learning fault diagnosis methods, the proposed fault diagnosis method has obvious advantages [21–24].

#### 4. Conclusion

In order to improve the accuracy of motor fault diagnosis, a motor fault diagnosis algorithm based on wavelet and attention mechanism is proposed. The wavelet denoising method is used to remove the original signal noise and improve the signal-to-noise ratio. Convolution neural network and bidirectional gating loop unit are used to fuse time-frequency information. The attention mechanism module is used to dynamically weigh the features of different time points in time dimension. In the experimental part, the feasibility of the algorithm is verified by the dataset of Case Western Reserve University Bearing Data Center. The field diagnostic test was used for test comparison. Experimental results show that, compared with other deep learning methods, the proposed fault diagnosis method has obvious advantages.

#### Data Availability

The labeled dataset used to support the findings of this study are available from the corresponding author upon request.

#### Conflicts of Interest

The authors declare no competing interests.

#### Acknowledgments

The study was supported by the ABA Science and Technology Bureau (19YYJSYJ0091) and the 13th five-year plan of educational informatization construction in Sichuan Province (Chuan Jiao Guan 2019-142).

#### References

- [1] I. H. Kao, W. J. Wang, Y. H. Lai, and J. W. Perng, "Analysis of permanent magnet synchronous motor fault diagnosis based on learning," *IEEE Transactions on Instrumentation and Measurement*, vol. 68, no. 2, pp. 310–324, 2019.
- [2] O. AlShorman, M. Irfan, N. Saad et al., "A review of artificial intelligence methods for condition monitoring and fault diagnosis of rolling element bearings for induction motor," *Shock and Vibration*, vol. 2020, Article ID 8843759, 20 pages, 2020.
- [3] Y. M. Hsueh, V. R. Ittangihal, W. B. Wu, H. C. Chang, and C. C. Kuo, "Fault diagnosis system for induction motors by CNN using empirical wavelet transform," *Symmetry*, vol. 11, no. 10, p. 1212, 2019.
- [4] W. Dehina, M. Boumechraz, and F. Kratz, "On-line detection and estimation of harmonics components in induction motors rotor fault through a modified Prony's method," *International Transactions on Electrical Energy Systems*, vol. 31, no. 2, article e12737, 2021.
- [5] K. C. D. Kompella and G. V. Madhav, "An improved matrix pencil method based bearing fault detection in three phase induction motor," in *2020 IEEE International Conference on Computing, Power and Communication Technologies (GUCON)*, pp. 51–56, NCR New Delhi, India, 2020.
- [6] F. Husari and J. Seshadrinath, "Inter-turn fault diagnosis of induction motor fed by PCC-VSI using Park vector approach," in *2020 IEEE International Conference on Power Electronics, Drives and Energy Systems (PEDES)*, pp. 1–6, Jaipur, Rajasthan, India, 2020.
- [7] F. Ding and F. W. Qin, "Application of wavelet denoising and Hilbert transform in fault diagnosis of motor bearing," *Electric Machines and Control*, vol. 21, no. 6, pp. 89–95, 2017.
- [8] D. Kiel, C. Arnold, and K. I. Voigt, "The influence of the Industrial Internet of Things on business models of established manufacturing companies - a business level perspective," *Technovation*, vol. 68, pp. 4–19, 2017.
- [9] D. Angeli and S. Manfredi, "A Petri Net approach to consensus in networks with joint-agent interactions," *Automatica*, vol. 110, article 108466, 2019.
- [10] Z. Jiang, Z. Li, N. Wu, and M. Zhou, "A Petri net approach to fault diagnosis and restoration for power transmission systems to avoid the output interruption of substations," *IEEE Systems Journal*, vol. 12, no. 3, pp. 2566–2576, 2018.
- [11] G. P. Bhandari, R. Gupta, and S. K. Upadhyay, "Colored Petri nets based fault diagnosis in service oriented architecture," *International Journal of Web Services Research*, vol. 15, no. 4, pp. 1–28, 2018.
- [12] K. Q. Zhou, W. H. Gui, L. P. Mo, and A. Zain, "A bidirectional diagnosis algorithm of fuzzy Petri net using inner-reasoning-path," *Symmetry*, vol. 10, no. 6, p. 192, 2018.
- [13] Y. Zhang, Y. Zhang, F. Wen et al., "A fuzzy Petri net based approach for fault diagnosis in power systems considering temporal constraints," *International Journal of Electrical Power & Energy Systems*, vol. 78, pp. 215–224, 2016.
- [14] Ying Chen, P. B. Luh, Che Guan et al., "Short-term load forecasting: similar day-based wavelet neural networks," *IEEE Transactions on Power Systems*, vol. 25, no. 1, pp. 322–330, 2010.
- [15] X. Li, X. Zhang, P. Zhang, and G. Zhu, "Fault data detection of traffic detector based on wavelet packet in the residual subspace associated with PCA," *Applied Sciences*, vol. 9, no. 17, p. 3491, 2019.
- [16] D. Xu, S. Zhang, H. Zhang, and D. P. Mandic, "Convergence of the RMSProp deep learning method with penalty for nonconvex optimization," *Neural Networks*, vol. 139, pp. 17–23, 2021.
- [17] Y. Li, X. Wang, S. Si, and S. Huang, "Entropy based fault classification using the Case Western Reserve University data: a benchmark study," *IEEE Transactions on Reliability*, vol. 69, no. 2, pp. 754–767, 2020.
- [18] C. Alippi, S. Disabato, and M. Roveri, "Moving convolutional neural networks to embedded systems: the alexnet and VGG-16 case," in *2018 17th ACM/IEEE International Conference*



- on Information Processing in Sensor Networks (IPSN)*, pp. 212–223, Porto, Portugal, 2018.
- [19] G. Cao, K. Zhang, K. Zhou, H. Pan, Y. Xu, and J. Liu, “A feature transferring fault diagnosis based on WPDR, FSWT and GoogLeNet,” in *2020 IEEE International Instrumentation and Measurement Technology Conference (I2MTC)*, pp. 1–6, Dubrovnik, Croatia, 2020.
- [20] L. Wen, X. Li, and L. Gao, “A transfer convolutional neural network for fault diagnosis based on ResNet-50,” *Neural Computing and Applications*, vol. 32, no. 10, article 4097, pp. 6111–6124, 2020.
- [21] S. Deb, Z. Tian, S. Fong, R. Wong, R. Millham, and K. K. L. Wong, “Elephant search algorithm applied to data clustering,” *Soft Computing*, vol. 22, article 3076, no. 18, pp. 6035–6046, 2018.
- [22] Z. Tang, G. Zhao, and T. Ouyang, “Two-phase deep learning model for short-term wind direction forecasting,” *Renewable Energy*, vol. 173, pp. 1005–1016, 2021.
- [23] K. K. L. Wong, G. Fortino, and D. Abbott, “Deep learning-based cardiovascular image diagnosis: a promising challenge,” *Future Generation Computer Systems*, vol. 110, pp. 802–811, 2020.
- [24] D. Yao, H. Liu, J. Yang, and J. Zhang, “Implementation of a novel algorithm of wheelset and axle box concurrent fault identification based on an efficient neural network with the attention mechanism,” *Journal of Intelligent Manufacturing*, vol. 32, article 1701, no. 3, pp. 729–743, 2021.