

Research Article

A Fault Identification Method of Mechanical Element Action Unit Based on CWT-2DCNN

Hongyu Ge, Yujiao Guo , Tianyu Luo, Manzhi Yang, and Chuanwei Zhang

School of Mechanical Engineering, Xi'an University of Science and Technology, Xi'an 710054, China

Correspondence should be addressed to Yujiao Guo; 787234506@qq.com

Received 14 July 2022; Revised 7 October 2022; Accepted 4 November 2022; Published 12 November 2022

Academic Editor: Andrea Spaggiari

Copyright © 2022 Hongyu Ge 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.

Aiming at the problems of low recognition rate and human intervention in the traditional fault diagnosis of mechanical equipment, a fault identification method based on continuous wavelet transform (CWT) and two-dimensional convolutional neural network (2DCNN) is proposed. By collecting the vibration signals of four kinds of fault states and normal states of the worm rotation unit of the CNC machine tool, the data are preprocessed and identified. Firstly, the vibration signals of each fault state of the element action unit are CWT transformed into the corresponding two-dimensional time-frequency diagram; then, the 2DCNN fault identification model is established, and the time-frequency diagrams of various faults are input to the network as characteristic diagrams for training and testing. Through the adjustment of network parameters, the network performance is gradually optimized; finally, the hybrid domain attention module CBAM is added to further improve the network structure, and the recognition effect is compared with the initial 2DCNN. The results show that the CWT-2DCNN meta-action unit fault recognition model with an attention module can recognize the different states of meta-action units more effectively, and the fault recognition effect is better. By using this method, the different fault types of mechanical element action units can be accurately identified, which has a certain application in the field of mechanical fault identification and diagnosis.

1. Introduction

The metal action unit is the key movement unit to ensure the normal operation of mechanical equipment, and the state of mechanical components in the operation process directly affects the operation efficiency and life of the system [1]. Therefore, the identification of the fault state of the metal action unit is of great significance to ensure the safe operation of mechanical equipment.

Vibration signal analysis is a commonly used mechanical fault diagnosis method, which extracts valuable features that provide internal machine information from the collected chaotic signals [2]. The vibration signal of the CNC machine tool has typical nonlinear and nonstationary characteristics. A large number of scholars have studied the extraction of its signal characteristics. Chen et al. [3] and Keshtan and Nouri Khajavi [4] extracted the fault characteristics of vibration signals through EMD and EEMD modal decomposition methods, respectively and diagnosed the fault of rolling

bearings; Zhao et al. [5] proposed a new method of fast icaeemd for feature extraction of vibration signals; Zhou et al. [6] proposed a fault diagnosis method of marine propulsion shafting based on partial integration of empirical mode decomposition and support vector machine; Zhao et al. [7] proposed a fault diagnosis method of VMD, Hilbert transform, and deep belief network for rolling bearing fault diagnosis under complex working conditions; He et al. [8] proposed a fault feature extraction method based on empirical wavelet transform and spectral kurtosis (ewt-sk), which effectively suppressed the noise and applied it to the fault diagnosis of ship shafting. Many signal recognition methods studied by many scholars [8–10] also show their advantages in signal feature extraction, but the traditional recognition methods for fault signal extraction still have certain limitations, and the universality is not high. The time-frequency analysis results of continuous wavelet transform (CWT) signal can reflect the two-dimensional diagram of the energy intensity of the signal at different

times and frequencies and can display the detailed changes of the signal from multiple angles, to more effectively describe the subtle fault characteristics of the signal. Therefore, this paper selects the method of continuous wavelet transform to extract the characteristics of the fault signal of the element action unit.

The convolutional neural network is a multilayer perceptron designed to recognize two-dimensional feature maps. As a popular scheme to study mechanical fault diagnosis, the convolutional neural network has been studied by scholars at home and abroad in recent years. An et al. [11] proposed an intelligent fault diagnosis framework for bearing under time-varying working conditions based on recurrent neural networks; Li et al. [12] proposed a fault diagnosis method based on multiscale permutation entropy and multichannel fused convolutional neural network; Chen et al. [13] realized the diagnosis and identification of gearbox faults through CNN; Jing et al. [14] used the deep convolution model to adaptively extract the original fault signal and determine the motor fault; Babu et al. [15] and Li [16] both used convolutional neural networks to predict the remaining life of bearings and verified their effectiveness. Much literature has proved that convolutional neural network is superior to other mechanical fault recognition methods.

To sum up, aiming at the problems of low efficiency and human intervention in the fault identification of mechanical element action units of CNC machine tools, this paper proposes research on the fault identification method of mechanical element action units based on continuous wavelet transform and two-dimensional convolution neural network model. CWT is used to generate a two-dimensional time-frequency diagram of fault vibration signals collected during the operation of the mechanical element action unit, and then input it as a characteristic diagram into the built 2DCNN for identification. The built 2DCNN adds the mixed domain attention mechanism module CBAM to better extract and recognize the input image features. In the convolution neural network, the time-frequency diagram of the element action unit is trained, and the network parameter structure is continuously adjusted to optimize the network model and to realize the fault identification of the mechanical element action unit.

2. Construction of Meta-Action Unit Test-Bed and Signal Acquisition

2.1. Construction of Meta-Action Unit Test Bed. The assembly and working process of the CNC machine tool are completed by multiple action units. The “function motion action” (FMEA) structural decomposition method [17] is used to decompose the functional motion of the CNC machine tool until the minimum motion unit-meta action unit. In this paper, the worm rotation element action unit of the CNC machine tool is selected as the research object, and the worm rotation element action unit test platform is built, as shown in Figure 1.

According to the assembly relationship of worm rotation element action unit and the failure forms of various parts,

the four most common failure modes are determined through many tests. They are coupling looseness fault, flat key wear fault, poor bearing assembly fault, and worm axis offset fault. According to the test bench shown in Figure 1, the speed and vibration signals generated by the worm rotation element action unit during operation are collected by the multifunctional data acquisition card through the sensor. The speed and vibration fault data generated during the operation of the element action unit are uploaded to the data acquisition terminal by the multifunctional data acquisition card. The vibration signal corresponding to each fault type is determined by analyzing the variation of vibration and speed with the running time. Finally, the one-dimensional vibration signal data are input into the built convolutional neural network fault identification model through the two-dimensional image after continuous wavelet transform for training and testing.

2.2. Fault Identification Process of Element Action Unit. The research idea of intelligent fault identification method of mechanical element action unit based on CWT-2DCNN proposed in this paper is as follows:

- (1) The signal acquisition system collects the vibration signals in five different states in the test bed built by the worm rotation element action unit and constructs the signal sample set through data segmentation;
- (2) The vibration signals in the sample set are pre-processed, that is, the time-frequency map is generated by CWT, and then the time-frequency map is compressed to an appropriate size to construct the feature map sample set. The purpose of compression is to reduce the size of each dimension of the CNN input characteristic map, improve the training speed of the network, and ensure that the useful information in the time-frequency map is not submerged;
- (3) Establish the 2DCNN model, select a certain number of wavelet time-frequency characteristic images as training samples, train CNN, and optimize the parameters and network in the training process;
- (4) The remaining samples in the time-frequency feature map sample set are used as test samples to test the trained CNN and identify the different states of meta-action units.

The research idea framework is shown in Figure 2.

2.3. Fault Signal Acquisition of Element Action Unit. In this experiment, the CT1005LC vibration acceleration sensor is used to test five states of the worm rotation element action unit, including the normal operation state, coupling looseness fault state, flat key wear fault state, poor bearing assembly fault state, and the worm axis offset fault state.

The coupling looseness fault of the worm rotation element action unit is usually caused by the wear of the internal contact surface of the coupling and the deformation of the plum blossom star elastic key connecting the coupling.

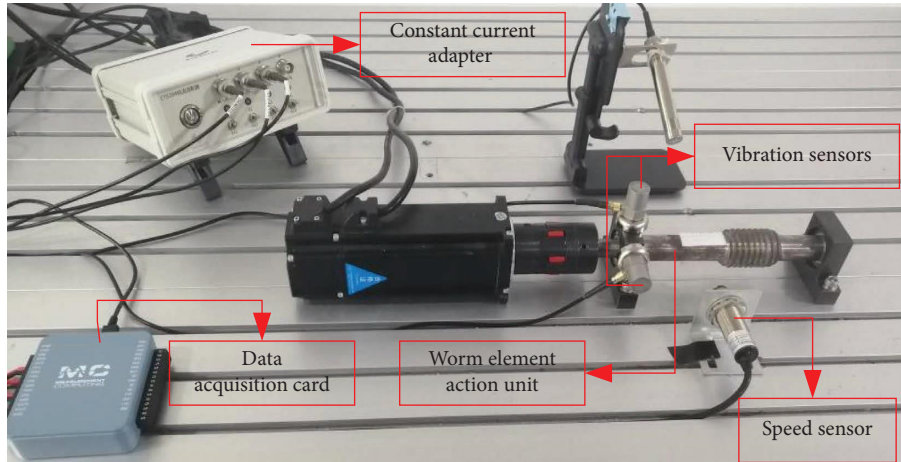


FIGURE 1: Worm rotation element action unit test bench.

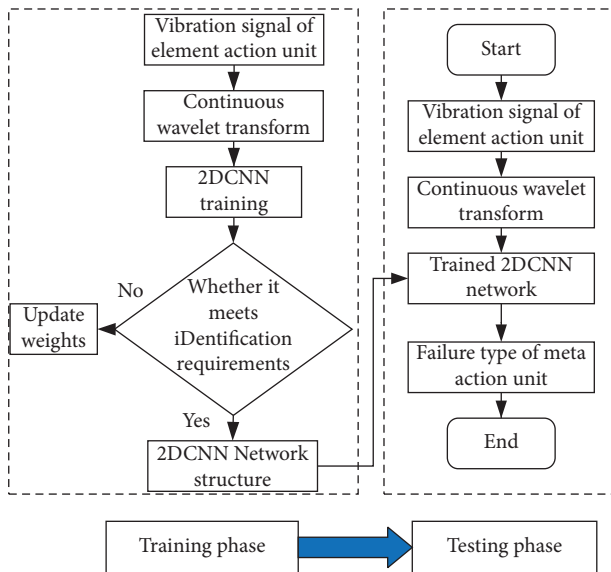


FIGURE 2: Schematic diagram of cwt-2dcnn mechanical element action unit fault identification process.

Therefore, on the premise that the functions and assembly of other parts of the worm rotation element action unit are in good condition, the coupling looseness fault in these two cases is set in the test, as shown in Figures 3(a) and 3(b); The wear failure of the flat key is usually caused by the wear of the flat key surface and the wear of the flat key locating hole. Therefore, the flat key wear fault in these two cases is shown in Figures 3(c) and 3(d); bearing assembly failure is usually caused by excessive bearing assembly clearance and poor bearing lubrication. Therefore, the bearing assembly failure in these two cases is shown in Figures 3(e) and 3(f); The worm axis offset fault is usually caused by the axis offset of the worm rotation element action unit due to the instability of the base or bearing support during operation. The fault setting is shown in Figure 3(g).

Various types of fault division and signal acquisition make the data set more diverse and universal, which can effectively improve the generalization ability of network

identification. The settings of various fault status modes are shown in Figure 3. In the test, the sampling frequency is set at 1 kHz, and three different motor speeds ($n = 1800, 2400,$ and 3000 r/min) are used for operating conditions. 1024 data sampling points are used as a sample to divide the sample data. Under different fault types, there are 140 samples under different working conditions. Finally, 420 samples under the normal state and worm axis offset fault state are obtained and 840 samples under the other three fault states, a total of 3360 samples. The samples collected for each fault of the worm rotation element action unit are shown in Table 1.

3. Preprocessing of Fault Signals of Meta-Action Unit

3.1. Continuous Wavelet Transform of Signals. In the identification and diagnosis technology of mechanical vibration faults, feature extraction needs to rely on signal processing. The analysis methods of signal processing mainly include time-domain analysis, frequency-domain analysis, and time-frequency domain analysis [18]. Because the working environment of the mechanical element action unit is complex and changeable and is often affected by the background noise, the features contained in the collected vibration signal will be impacted by the noise components, so it is difficult to extract the signal features through the depth learning network. In addition, the vibration signals of mechanical faults actually measured are nonstationary and nonlinear in most cases, and their frequency components contain time-varying characteristics. Although time-frequency analysis is simple and intuitive, it cannot provide effective fault feature information, and frequency domain analysis alone cannot obtain the changing relationship between time and frequency of signals.

Fourier transform, as the most basic time-frequency transform method, cannot effectively depict the local characteristics of signals in the time domain. To solve this problem, windowed Fourier transform is introduced. However, the size of the window is difficult to select, and this method still cannot meet the requirements for changing the frequency of nonstationary signals [19]. Continuous wavelet

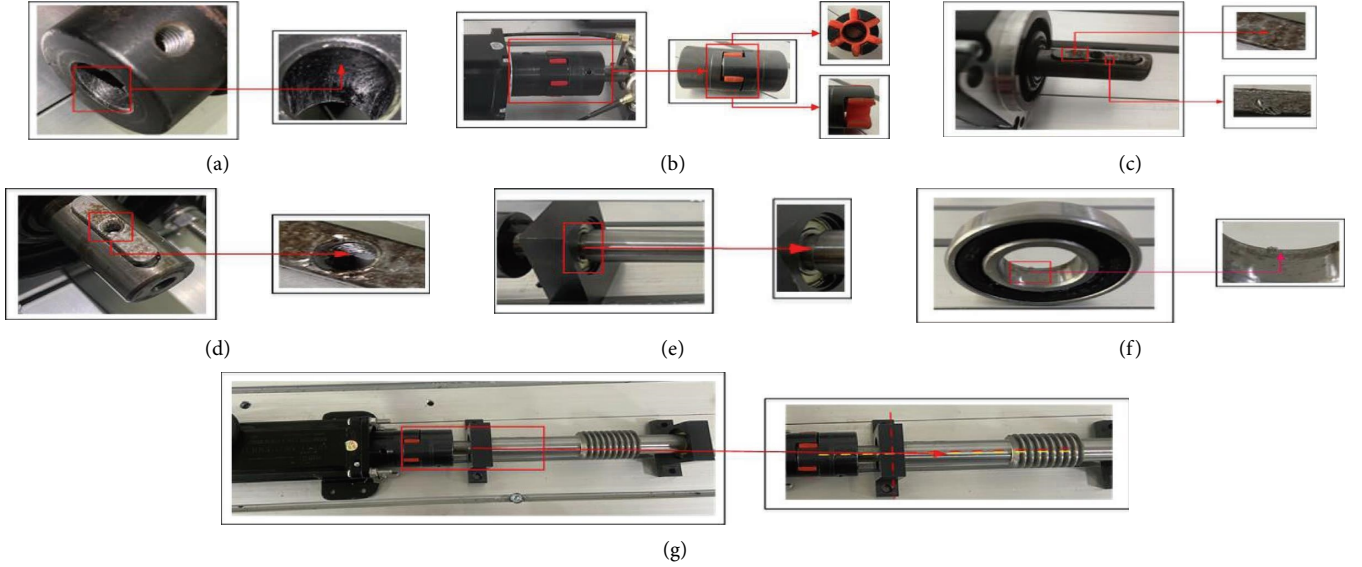


FIGURE 3: Schematic diagram of each fault setting. (a) The contact surface in the coupling is worn. (b) Deformation of plum star elastic bond. (c) Wear of flat key surface. (d) The flat key locating hole is worn. (e) Bearing assembly clearance is too large. (f) Poor bearing lubrication. (g) Worm axis offset.

TABLE 1: Worm fault sample table.

Working condition	Normal state	State type						
		Worm axis offset fault status	Coupling looseness fault status	Flat key wear fault status	Poor bearing assembly fault status			
			The contact surface in the coupling is worn	Deformation of quincunx star elastic bond	The flat key locating hole is worn	Wear of flat key surface	Bearing assembly clearance is too large	Poor bearing lubrication
$N = 1800$	140	140	140	140	140	140	140	140
$N = 2400$	140	140	140	140	140	140	140	140
$N = 3000$	140	140	140	140	140	140	140	140
Total	420	420		840		840		840

transform replaces infinite triangular function basis with finite attenuated wavelet basis. This method can intuitively observe signals in the time domain and frequency domain and is widely used in signal denoising, image compression, mechanical equipment fault detection, and other fields [20]. Therefore, continuous wavelet analysis, an adjustable time-frequency window analysis method, is selected in this paper, which can better extract the fault features in the vibration signal and improve the accuracy of subsequent mechanical fault identification.

The basic idea of the wavelet transform is similar to that of the Fourier transform. It also uses a family of functions to represent signals. This family of functions is called the wavelet function system. Set a function " $f(t) \in L^2(R)$," then

$$\begin{aligned} CWT_f(a, t) &= [f(T), \varphi_{a,\tau}(t)], \\ &= \frac{1}{\sqrt{a}} \int f(T) \varphi^*\left(\frac{t-\tau}{a}\right) dt. \end{aligned} \quad (1)$$

In the above formula " $\varphi_{a,\tau}(t)$ " is obtained from the mother wavelet through scaling and translation;

$$\varphi_{a,\tau}(t) = \frac{1}{\sqrt{a}} \varphi\left(\frac{t-\tau}{a}\right), a, \tau \in R, a > 0, \quad (2)$$

where " a " is the size factor, which represents the frequency-dependent expansion; " τ " is the translation factor; $\varphi_{a,\tau}(t)$ is the wavelet basis function; both " a " and " τ " are continuous variables, so they are called continuous wavelet transforms. The continuous wavelet transform method proposed in this paper can automatically adjust factors a and τ . It makes the vibration signals of the transformation pair action unit with different time intervals have the characteristics of adaptability and multiresolution. The algorithm steps for transforming a vibration signal into a time-frequency diagram are as follows:

Step 1: Let " a " be the scaling factor, " f_s " be the sampling frequency, " F_C " be the wavelet center frequency, then the actual frequency " Fa " corresponding to " a " is as follows:

$$F_a = \frac{F_c \times f_s}{a}. \quad (3)$$

Step 2: According to the above formula, to make the converted frequency sequence an equal difference sequence, the scale sequence takes the following form:

$$\frac{c}{\text{totalscal}}, \dots, \frac{c}{(\text{totalscal} - 1)}, \frac{c}{4}, \frac{c}{2}, c, \quad (4)$$

Where “total scale” is the length of the scale sequence used in the continuous wavelet transform of the signal (256 is present in this paper); “ c ” is a constant.

Step 3: According to the above formula, the actual frequency corresponding to the scale “ $c/\text{total scale}$ ” is “ $f_s/2$,” so we can get the following:

$$c = 2 \times F_c \times \text{totalscal}. \quad (5)$$

By substituting the above formula, the required scale sequence t can be obtained.

Step 4: After the wavelet base and scale are determined, the wavelet coefficients “ $W_f(a, b)$ ” are obtained by using the principle of continuous wavelet transform. Then, the scale sequence is transformed into the actual frequency sequence “ f .” Finally, combined with the time series “ t ,” the wavelet time-frequency map is drawn to obtain the characteristic information.

The key of CWT is the selection of wavelet basis function, and the waveform of the selected wavelet basis function should be similar to the fault characteristics of the signal. Because the waveform of the Morlet wavelet (morl wavelet) is similar to the impact characteristics caused by mechanical faults of CNC machine tools, and the Complex Morlet wavelet (cmor wavelet) is the complex form of Morlet wavelet, and its adaptive performance is better. Therefore, the CMOR wavelet is selected as the wavelet basis function of CWT. Using continuous wavelet transform, we can clearly identify the frequency components contained in the original signal and their corresponding time windows. The time-frequency power map effectively contains the relevant features of the original vibration signal [21].

3.2. Acquisition of Time-Frequency Diagram of Fault Signals of Meta-Action Unit. Collect the vibration signals of the worm rotation element action unit under five state modes on the built element action unit test bench. Under the working condition of $n = 1800$ r/min, the vibration signal diagrams and their corresponding time-frequency diagrams obtained by continuous wavelet transform are shown in Figure 4.

By CWT processing the original vibration signal collected by the element action unit, the one-dimensional characteristics of the signal can be projected into the two-dimensional space. While reflecting the frequency characteristics of the vibration signal of the element action unit, it can also show the transformation relationship of the frequency of the signal in the time domain and space. The obtained two-dimensional time-frequency image can more

clearly reflect its state characteristics and the information contained in each fault so as to facilitate the identification of various faults.

4. Fault State Identification of Meta-Action Unit

After the CWT continuous wavelet transform of the signals of the element action units, to further establish the mapping relationship between the strain signals and various fault types, this paper selects the convolution neural network model commonly used in deep learning to map the vibration signals of the element action units one by one with each state mode and realizes the identification of different fault states of the element action units through the learning and training of the network.

Compared with ordinary one-dimensional convolutional neural networks, two-dimensional convolutional neural networks (2DCNN) can generally reflect the fault status of mechanical equipment through fewer data samples and shorter network training time. Therefore, this paper adopts the fault identification method of meta-action units based on the 2DCNN model (convolutional neural network with RGB three-channel two-dimensional images as input samples). This method can make full use of the filter level and classification level structure of CNN and integrate signal feature extraction and pattern recognition, to realize the “end-to-end” fault diagnosis of electromechanical equipment and improve the recognition rate of mechanical vibration faults. The following is the process of building the 2DCNN meta-action unit fault identification model described in this paper and the function realization and function explanation of each part of the network.

4.1. Construction and Training of the Initial 2DCNN Recognition Model

- (1) Input layer: The input data of the CNN network built in this paper are a two-dimensional time-frequency diagram of the vibration signal of the worm rotation element action unit processed by CWT. The input image is normalized and the image size is 64×64 . To improve the accuracy of the algorithm and avoid overfitting, data enhancement processing are performed on the training data set, that is, the training data are processed through random center clipping, horizontal flipping and image brightness, saturation, and contrast, and new data are generated from the existing data set to expand the training data.
- (2) Convolution layer: Each convolution layer in the convolution neural network has several convolution cores. The parameters in each convolution core are optimized by the backpropagation algorithm, and each convolution core undertakes the task of identifying different features. To obtain the best recognition effect, after repeated debugging, 8 3×3 convolution cores are set in the first convolution layer, 16 3×3 convolution cores are set in the second convolution layer, 32 3×3 convolution cores are set in the third and fourth convolution layers, and 64

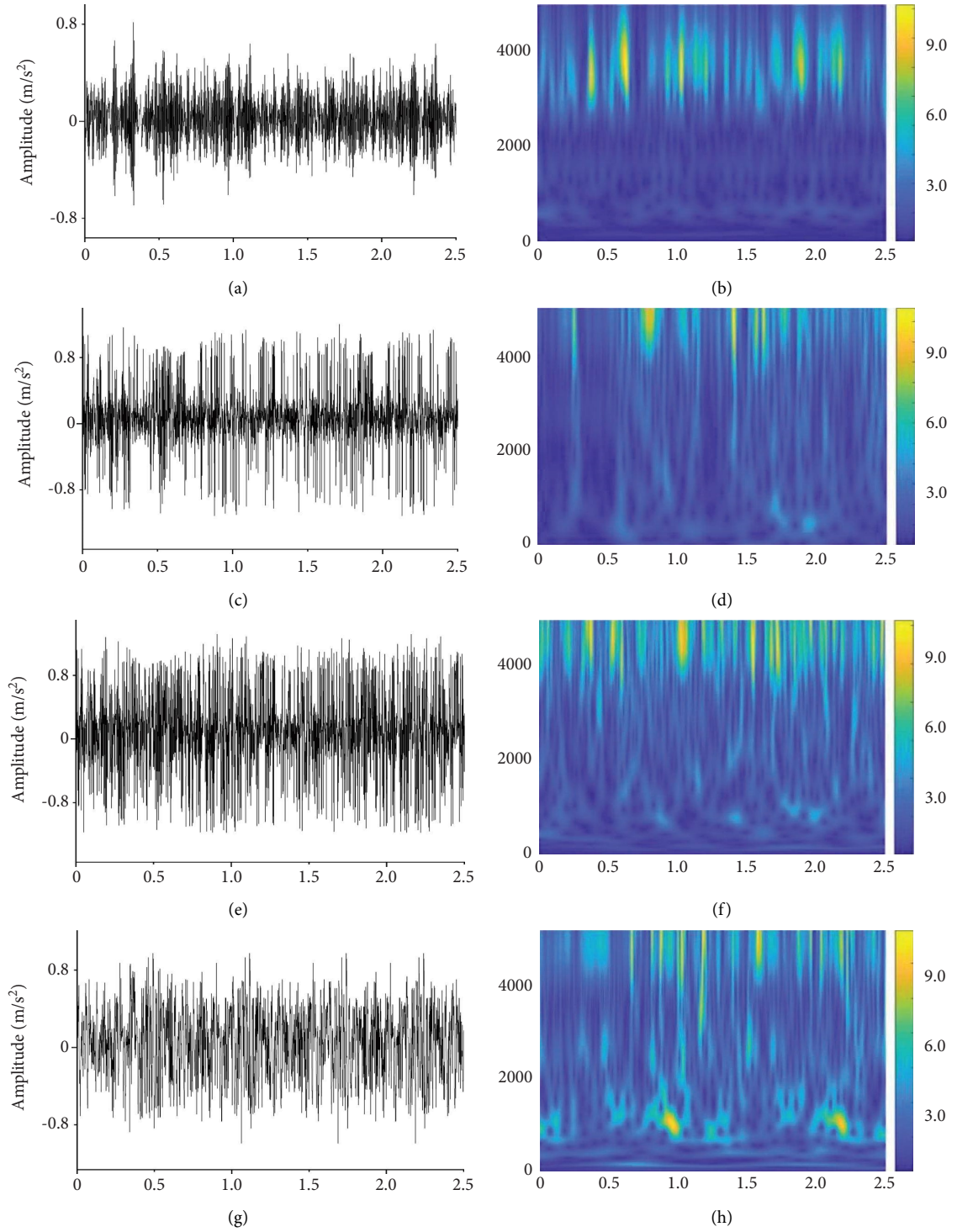


FIGURE 4: Continued.

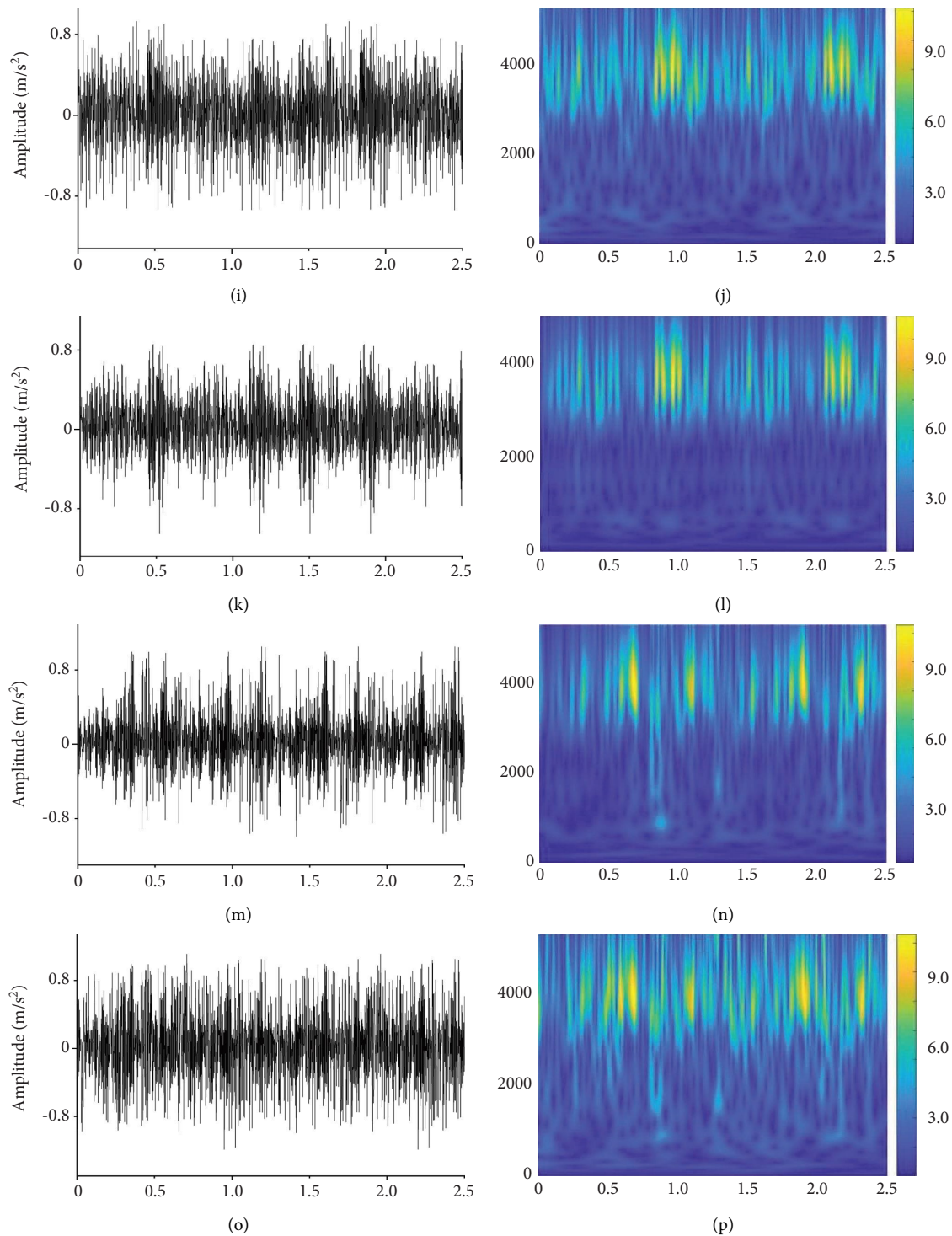


FIGURE 4: Vibration signals and corresponding time-frequency diagrams. (a) Normal state vibration image. (b) Normal time-frequency image. (c) The contact surface in the coupling is worn. (d) The contact surface in the coupling is worn. (e) Deformation of plum star elastic bond. (f) Deformation of plum star elastic bond. (g) Worm axis offset vibration image. (h) Worm axis offset time frequency image. (i) Flat key surface wear vibration image. (j) Time frequency image of flat key surface wear. (k) The flat key locating hole is worn. (l) The flat key locating hole is worn. (m) Bearing assembly clearance is too large. (n) Bearing assembly clearance is too large. (o) Poor bearing lubrication. (p) Poor bearing lubrication.

3×3 convolution cores are set in the fifth and sixth convolution layers. The convolution operation is as follows:

$$\chi_j^l = \sum_{i \in M_j} x_i^{l-1} k_{ij}^l + b_j^l, \quad (6)$$

Where “ k ” represents the convolution kernel; “ M_j ” represents the j th characteristic diagram; “ l ” stands for layer “ l ”; and “ b ” stands for offset.

- (3) Pooling layer: The pooling layer is mainly used to compress the input characteristic map. On the one hand, it makes the characteristic graph smaller and simplifies the computational complexity of the network; on the other hand, feature compression is used to extract the main features. Pooling operations are generally divided into average pooling and maximum pooling. In this paper, “Maximum Pooling” is selected, that is, the maximum value of the image area is selected as the pooled value of the area.
- (4) Activation function: The activation layer is responsible for activating the features extracted from the convolution layer. Since convolution is a linear transformation of the input image and convolution kernel, it is necessary to introduce a nonlinear function for nonlinear mapping. Relu function will make the output of some neurons 0, reduce the interdependence of parameters, and alleviate the occurrence of fitting problems. Therefore, this paper adds the relu activation function after each convolution layer.
- (5) Full connection layer: The function of the full connection layer is to connect all features and send the output value to the classifier. In this paper, two fully connected layers are designed to output features after the attention module.
- (6) Output layer: The output layer is mainly used to prepare for the output of the final target results. The common CNN models have two learning functions: regression and classification. The network designed in this paper is a regression model. In the regression analysis, the error and loss functions are calculated as follows:

$$\text{Loss} = \frac{1}{2} (y_i - y_i'). \quad (7)$$

In the backpropagation process of network training, to obtain the optimal weight, it is necessary to create a cost loss function for the model, and then select an appropriate optimization algorithm to obtain the minimum function loss value. In this paper, the most commonly used cross-entropy loss function for multiclassification problems is selected as the error cost function to evaluate the difference between the currently trained probability distribution and the real distribution. It describes the distance between the actual output (probability) and the expected output (probability), that is, the smaller the value of cross-entropy, the closer the two probability distributions are. The mathematical expression is as follows:

$$L = -\frac{1}{N} \sum_{i=1}^n [Y_i * \log(y_i') \log(1 - y_i) \log(1 - y_i)], \quad (8)$$

where “ N ” is the number of Samples, “ y_i ” is the expected value of the model, and “ y_i' ” is the predicted value of the model.

Gradient descent algorithm can accelerate the learning speed of depth network faster and solve the minimum error of loss function. In this paper, the adaptive moment estimation (Adam) algorithm of the gradient descent algorithm is used for optimization. Adam algorithm can dynamically adopt different learning rates for different parameters to make the objective function converge faster, increase the training speed and avoid falling into the local optimum.

The training method of 2DCNN designed in this paper adopts the method of batch sample input, and the batch size is set to 32. The training of the network includes two parts: the forward propagation of data and the backpropagation of error. In order to prevent overfitting of the model in the propagation process and strengthen the generalization ability of the neural network, a “Dropout Layer” is added after the full connection layer. The dropout regularization is adopted and its parameter size is set to 0.2 so as to strengthen the robustness of the network nodes.

After collecting the vibration signals of the five states of worm rotation on the built worm rotation element action unit test bed, through continuous wavelet transform and other pretreatments, a total of 3360 samples of two-dimensional time-frequency diagrams of various states under different working conditions were obtained, 80% of which were selected as training samples and the rest as test samples. Input data set into the 2DCNN built in the previous chapter for training and optimization of parameters. To improve the identification stability of the network and eliminate the interference of uncertain factors to the greatest extent, after many training iterations, the parameter values are constantly changed, and finally, a group of super parameters with the best network performance, the best iteration efficiency, and the highest accuracy are determined as the parameter settings of the network. The value settings of each super parameter are shown in Table 2.

4.2. Improvement and Optimization of 2DCNN Identification Model. To improve the fault identification effect of the network model with the meta-action unit, the attention mechanism module is added to the model structure. The main function of the attention mechanism is to effectively fuse the fault information collected by sensors by suppressing the information irrelevant to the fault and highlighting the information closely related to the fault information.

4.2.1. Selection of Attention Module. The attention mechanism is mainly divided into channel attention mechanism and spatial attention mechanism. The channel attention mechanism pays more attention to the characteristic

TABLE 2: Value table of network super parameters.

Super parameter name	Numerical value
Epoch	200
Learning_rate	0.001
Dropout_rate	0.2
Batch_size	32
Conv1	$64 \times 64 \times 32$
Max Pooling1	$32 \times 32 \times 64$
Conv2	$32 \times 32 \times 64$
Max Pooling2	$16 \times 16 \times 128$
Conv3	$16 \times 16 \times 128$
Conv4	$16 \times 16 \times 128$
Max Pooling3	$8 \times 8 \times 256$
Conv5	$8 \times 8 \times 256$
Conv6	$8 \times 8 \times 256$
FC1	$1 \times 1 \times 512$
FC2	$1 \times 1 \times 100$
Output_units	5

information of the input channel of the image, and its formula is as follows:

$$\begin{aligned} M_C(F) &= \sigma(MLP(\text{Avg}(F))) + (MLP(\text{Max}(F))) \\ &= \sigma(W_1(W_0(F_{\text{avg}}^c))) + (W_1(W_0(F_{\text{Max}}^c))), \end{aligned} \quad (9)$$

where “ F ” is the characteristic diagram of the input network; “ W_0 ” and “ W_1 ” are full connection layers; and “ σ ” represents the “Sigmoid function.”

The spatial attention mechanism pays more attention to the feature information of the input image in the spatial dimension, and its calculation formula is as follows:

$$\begin{aligned} M_S(F) &= \sigma(f^{7 \times 7}[\text{Avg}(F); \text{Max}(F)]) \\ &= \sigma(f^{7 \times 7}[F_{\text{avg}}^S; F_{\text{Max}}^S]). \end{aligned} \quad (10)$$

Considering the limitations of channel attention and spatial attention, this paper chooses to add a CBAM hybrid domain attention module that combines space and channels to the model [22]. Its principle is shown in Figure 5. The specific operations of adding the CBAM module are as follows: the input feature maps are, respectively, pooled by global maximum and global average based on width and height, and then the shared MLP. Add the MLP output features based on “elementwise,” and then generate the channel attention feature map through “sigmoid” activation. The attention feature map and input features of the channel are “elementwise” multiplied to generate the input features required by the spatial attention module. Then, take the characteristic graph output by the channel attention mechanism module as the input of the spatial attention mechanism module. First, do a global maximum pool and a global average pool of the input based on the channel, and then do a “concat” operation based on the channel. After a convolution operation, the dimension is reduced to one channel. Then, the spatial attention feature map was generated by “sigmoid.” Finally, the

feature and the input feature of the module are multiplied to get the final feature.

4.2.2. Construction and Training of the Improved 2DCNN.

In order not to change the network structure of the convolutional neural network, this paper adds the channel spatial attention mechanism module to the convolution of the last layer of the 2DCNN network, which can ensure the normal operation of the pretraining parameters. That is, the 2DCNN network is composed of the input layer, convolution layer, pooling layer, full connection layer, attention module layer, and output layer. Among them, there are 6 convolution layers and 3 pooling layers. In order not to change the original network structure, the CBAM module is placed after the convolution of the last layer, and finally, the classifier is composed of two full connection layers and an output layer. The classifier uses “softmax” to output the status labels of the meta-action units. In this paper, the 2DCNN network regression learning method will be used to establish the mapping relationship between each state signal and its label after training the wavelet time-frequency image data corresponding to five types of vibration signals in four types of fault states and normal states of the unit action unit, to have the ability to identify and output continuous unknown vibration signals. The fault identification model of the 2DCNN element action unit built in this paper is shown in Figure 6.

After the data set is input into the improved meta-action unit fault identification network model with the CBAM attention module for training, the parameters are consistent with the initial network model, and then the network identification effect is verified on the test set.

4.3. Analysis of Fault Identification Results of Meta-Action Unit Based on CWT-2DCNN.

After constructing the network, the recognition and prediction results are obtained after 200 epochs. To describe the improved network performance more clearly, this paper will take the same parameters and data sets to train and test the initial network model without the channel spatial attention mechanism CBAM module and compare it with it. The training convergence curve of the recognition network without CBAM module for recognizing the five states of the worm rotation element action unit is shown in the following figure. Figure 7(a) shows the accuracy curve of the training set and the test set, and Figure 7(b) shows the loss curve.

It can be seen from Figure 7 that the loss value of the recognition network model without attention mechanism reaches 0.12, and the accuracy rate after convergence is 89.5%. Although the network model can accurately identify the five state types of worm rotation element action units, the recognition effect and accuracy are general. Therefore, the 2DCNN with the mixed domain mechanism module is used to identify the five states of the worm rotation element action unit again.

Figures 8(a) and 8(b) are the accuracy and loss curves of the network on the training set and the test set after adding the CBAM module. As can be seen from Figure 8, the

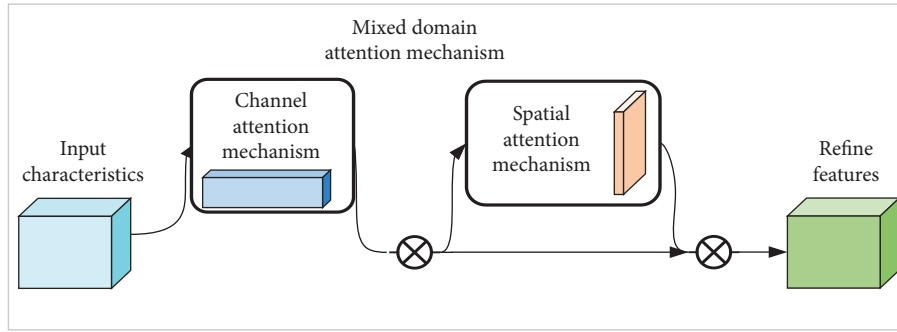


FIGURE 5: Schematic diagram of CBAM module of hybrid domain attention mechanism.

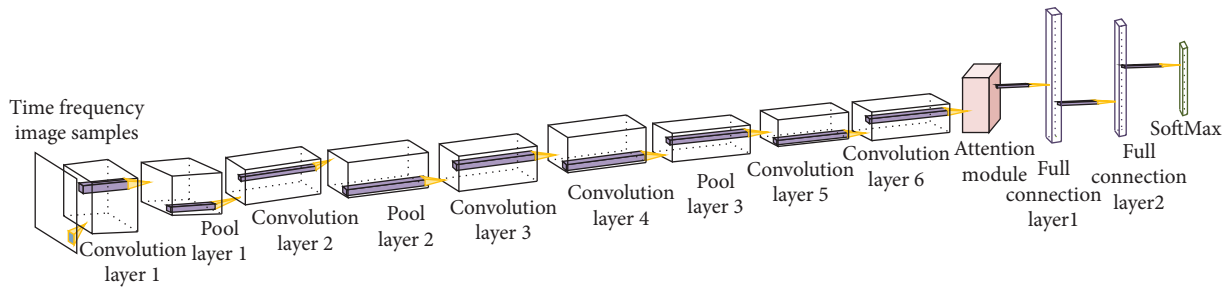


FIGURE 6: Improved 2DCNN model structure.

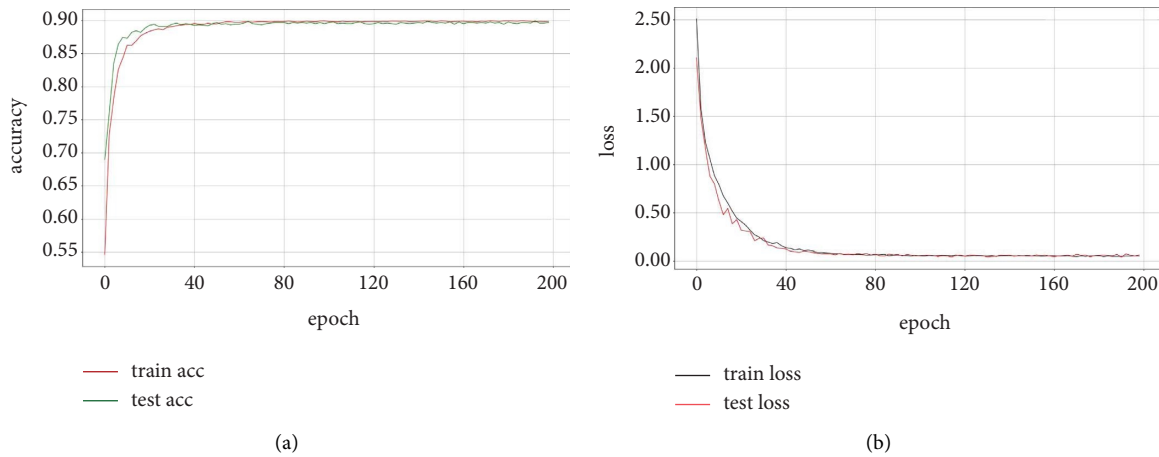


FIGURE 7: Change curve of the initial network training. (a) Change curve of the network training accuracy before improvement. (b) Change curve of network training loss value before improvement.

network error decreases with the increase in epochs. After about 120 epochs, the network has fully converged. At this time, the recognition accuracy of the meta-action unit recognition network for various states has reached the best accuracy of 97.6%, and the loss value is 0.063.

It can be seen from the above figure that the 2DCNN with the mixed domain attention mechanism module has a certain improvement in convergence speed and accuracy compared with the initial network. Compared with the improved network with an average iteration time of

40.25 s, the average iteration time of the improved network is 38.62 s. This shows that the addition of the CBAM module does not cause too much burden on the network, and the recognition accuracy has been significantly improved, which also reflects the effectiveness and superiority of the improved network model with the addition of the CBAM module.

To further analyze the recognition effect and the error of the two models, the confusion matrix is used to carry out error classification statistics for each type of state data.

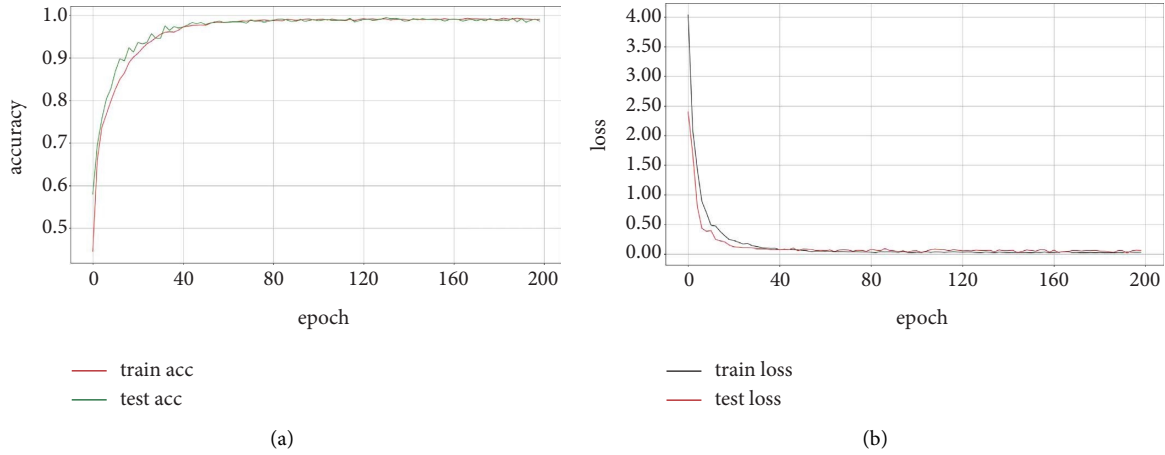


FIGURE 8: Change curve of network training after improvement. (a) Change curve of network training accuracy after improvement. (b) Change curve of network training loss value after improvement.

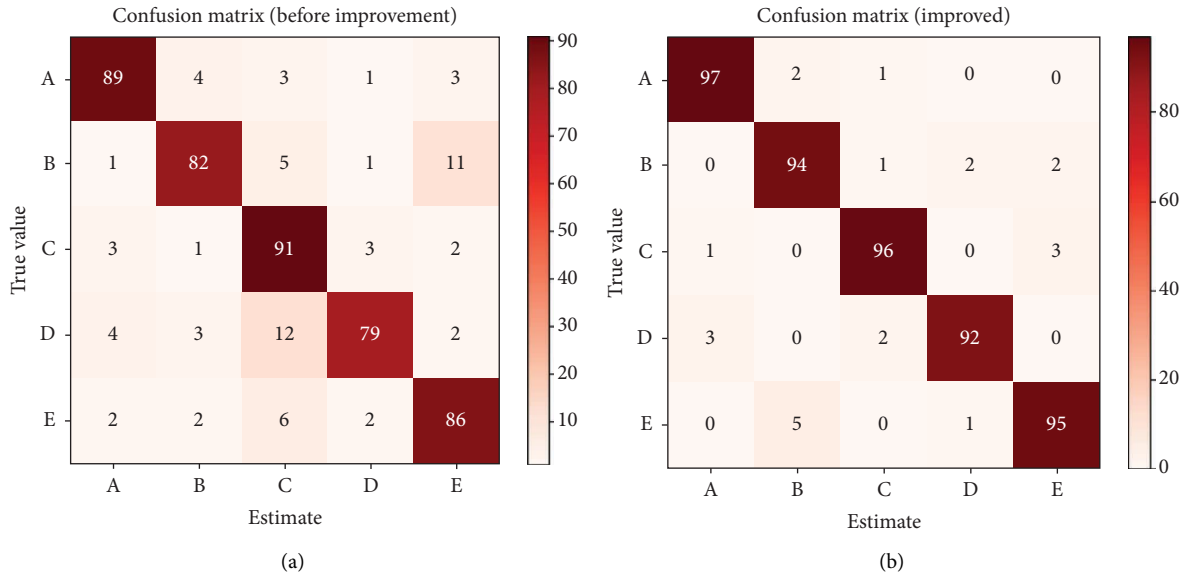


FIGURE 9: Network classification confusion matrix. (a) Confusion matrix of network classification before improvement. (b) Improved network classification confusion matrix.

Taking the state classification data of the worm rotation unit operating at 1800 r/min as an example, the result statistics are shown in Figure 9.

Figures 9(a) and 9(b) are the confusion matrix obtained by identifying and classifying the five states of the worm rotation element action unit with the improved 2DCNN network. The abscissa represents the predicted classification results, and the ordinate represents the real fault status categories. Among them, A, B, C, D, and E respectively represent the normal working state of the worm rotation element action unit, the loose coupling fault state, the flat key wear fault state, the poor bearing assembly fault state, and the worm axis offset fault state. The color depth represents the recognition accuracy, and the lighter the color, the closer the predicted value is to the real value. It can be calculated from the confusion matrix that in the network recognition

before and after the improvement, the precision rate “P”, recall rate “R”, and “F1 score” of the above five states of the worm rotation element action unit are shown in Table 3.

The improvement of each evaluation index shows that the improved 2DCNN network has better recognition and classification effect than the network without an attention module. The confusion matrix shows that the error between the predicted value and the actual value of the improved 2DCNN with the mixed domain attention mechanism CBAM module is significantly lower than the recognition error of the improved 2DCNN, and the accuracy is also relatively improved.

To verify the ability of the improved network to extract features, “t-sne” is used to reduce the dimension of the improved network feature layer data added to the CBAM module and output its visualization results. The output results are shown in Figure 10. The abscissa and ordinate,

TABLE 3: Network evaluation index values before and after improvement.

Evaluating indicator	Network before improvement					Improved network				
	A	B	C	D	E	A	B	C	D	E
Status type										
Precision rate (P)	0.9	0.89	0.78	0.92	0.83	0.96	0.92	0.97	0.97	0.95
Recall (R)	0.89	0.82	0.91	0.79	0.86	0.97	0.96	0.96	0.95	0.94
F1 fraction	0.89	0.85	0.84	0.85	0.84	0.96	0.94	0.96	0.96	0.94

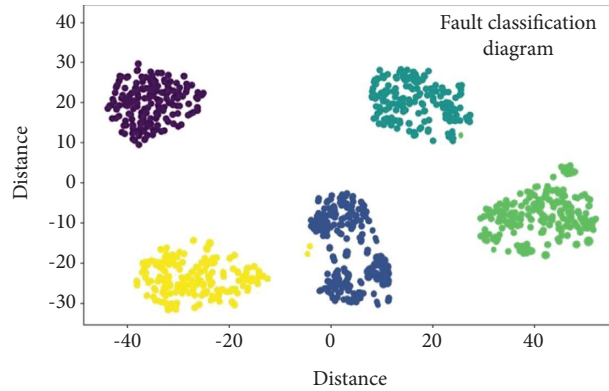


FIGURE 10: Schematic diagram of t-sne dimension reduction visualization results.

respectively, represent the horizontal and vertical distance of data features in the feature space. Under the improved 2DCNN recognition method, each state type of worm rotation element action unit presents a state of clustering by class, and different states have obvious differentiation. This shows that the fault identification method of the meta-action unit designed in this paper, which combines 2DCNN and CWT with the hybrid domain attention mechanism, has greatly improved the accuracy and classification efficiency of the worm rotation meta-action unit and can play a certain reference role in the fault identification and diagnosis of mechanical equipment.

5. Conclusions

Aiming at the problems of low recognition rate and too much human intervention in the traditional mechanical fault diagnosis methods, this paper presents a CWT-2DCNN fault identification method of mechanical element action unit based on the worm rotation element action unit of CNC machine tool. The conclusions of this paper are as follows:

- (1) In this paper, a mechanical fault identification method based on CWT and 2DCNN is used to identify the five state types of worm rotation element action units. Through CWT, the vibration signals of each state of the element action unit are converted into two-dimensional time-frequency diagrams as the input characteristic diagram of the two-dimensional convolutional neural network. Compared with the one-dimensional original vibration signal, the transformed time-frequency diagram of each state can better reflect the time-frequency characteristics of the signal. Combined with the 2DCNN, each state type can be better identified;

- (2) The 2DCNN designed in this paper adds the hybrid domain attention module CBAM to the network structure for improvement, and the recognition accuracy of each state type of worm rotation element action unit reaches 97.6%. Compared with the original network without the attention module with the same structural parameters, this network can extract the internal features of different states more efficiently, and the recognition effect is better.

Data Availability

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

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Project supported by the National Natural Science Foundation of China (No. 51705417, 51805428), the Shaanxi Provincial Natural Science Fund (No. 2019JQ-086), and the National Key Research and Development Program of China (2018YFB1703402).

References

- [1] W. Xu, L. Jing, J. Tan, and L. Dou, "A multimodel decision fusion method based on DCNN-IDST for fault diagnosis of rolling bearing," *Shock and Vibration*, vol. 2020, pp. 1–12, Article ID 8856818, 2020.

- [2] Z. Qiao, Y. Q. Liu, and Y. Y. Liao, "An improved method of EWT and its application in rolling bearings fault diagnosis," *Shock and Vibration*, vol. 2020, pp. 1–13, Article ID 4973941, 2020.
- [3] T. Chen, S. Dai, and X. Bi, "Fault diagnosis of rolling bearing based on EEMD," *Computer Simulation*, vol. 38, no. 02, pp. 361–364, 2021.
- [4] M. N. Keshtan and M. Nouri Khajavi, "Bearings fault diagnosis using vibrational signal analysis by EMD method," *Research in Nondestructive Evaluation*, vol. 27, no. 3, pp. 155–174, 2016.
- [5] K. Zhao, R. Jia, H. Wu, K. Dong, and J. Dang, "Extraction of vibration signal features based on Fast ICA-EEMD," *Journal of Hydroelectric Engineering*, vol. 36, no. 3, pp. 63–70, 2017.
- [6] S. T. Zhou, X. M. Zhu, C. J. Wu, and W. C. Peng, "Marine propulsion shaft system fault diagnosis method based on partly ensemble empirical mode decomposition and SVM," *Journal of Vibroengineering*, vol. 17, no. 4, pp. 1783–1795, 2015.
- [7] H. Zhao, H. Liu, J. Xu, C. Guo, and W. Deng, "Research on a fault diagnosis method of rolling bearings using variation mode decomposition and deep belief network," *Journal of Mechanical Science and Technology*, vol. 33, no. 9, pp. 4165–4172, 2019.
- [8] Y. He, X. Wang, and J. Dong, "A method for feature extraction of ship shafting faults based on empirical wavelet transform and spectral kurtosis," *Chinese Ship Research*, vol. 15, no. 1, pp. 98–106, 2020.
- [9] Y. S. Ahmed, A. F. M. Arif, and S. C. Veldhuis, "Application of the wavelet transform to acoustic emission signals or built-up edge monitoring in stainless steel machining," *Measurement*, vol. 154, Article ID 107478, 2020.
- [10] R. Yuan, Y. Lv, H. Li, and G. Song, "Robust fault diagnosis of rolling bearings using multivariate intrinsic multiscale entropy analysis and neural network under varying operating conditions," *IEEE Access*, vol. 7, pp. 130804–130819, Article ID 130804, 2019.
- [11] Z. An, S. Li, J. Wang, and X. Jiang, "A novel bearing intelligent fault diagnosis framework under time-varying working conditions using recurrent neural network," *ISA Transactions*, vol. 100, pp. 155–170, 2020.
- [12] H. Li, J. Huang, X. Yang, J. Luo, L. Zhang, and Y. Pang, "Fault diagnosis for rotating machinery using multiscale permutation entropy and convolutional neural networks," *Entropy*, vol. 22, no. 8, p. 851, 2020.
- [13] Z. Q. Chen, C. Li, and R. V. Sanchez, "Gearbox fault identification and classification with convolutional neural networks," *Shock and Vibration*, vol. 2015, no. 8, pp. 1–10, Article ID 390134, 2015.
- [14] L. Jing, T. Wang, M. Zhao, and P. Wang, "An adaptive multisensor data fusion method based on deep convolutional neural networks for fault diagnosis of planetary gearbox," *Sensors*, vol. 17, no. 2, pp. 414–429, 2017.
- [15] G. S. Babu, P. Zhao, and X. L. Li, *Deep Convolutional Neural Network Based Regression Approach for Estimation of Remaining Useful Life*, "International Conference on Database Systems for Advanced Applications, pp. 214–228, Dallas, TX, USA, 2016.
- [16] S. P. Li, *Research on Remaining Useful Life Prediction Method of a Rolling Bearing Combining CNN and LSTM*, Harbin University of Science and Technology, Harbin, Heilongjiang, China, 2019.
- [17] Y. Li, G. Zhang, Y. Wang, and G. Zhang, "Research on FMEA analysis technology of NC machine tools based on meta action," *Journal of Hunan University*, vol. 46, no. 10, 2019.
- [18] A. Jawadekar, S. Paraskar, S. Jadhav, and G. Dhole, "Artificial neural network based induction motor fault classifier using continuous wavelet transform," *Systems Science & Control Engineering*, vol. 2, no. 1, pp. 684–690, 2014.
- [19] Y. Yoo and J. G. Baek, "A novel image feature for the remaining useful lifetime prediction of bearings based on continuous wavelet transform and convolutional neural network," *Applied Sciences*, vol. 8, no. 7, p. 1102, 2018.
- [20] J. Ma, S. Zhang, X. Yan, and H. Chen, "Research on feature extraction and life prediction method of mechanical vibration signal based on image recognition," *Aerospace General Technology*, vol. 5, no. 04, pp. 33–42, 2021.
- [21] R. Liang, W. Ran, C. Yu, W. Chen, and De Ni, "Fault status identification of gearbox operation based on CWT-CNN," *Journal of Aerodynamics*, vol. 36, no. 12, pp. 2465–2473, 2021.
- [22] F. Zhou, Li Yang, and X. Fan, "Improvement of feedback loss calculation method of image classification convolution neural network," *Small microcomputer system*, vol. 40, no. 7, pp. 1532–1537, 2019.