(Received 11 March 2024, Accepted 21 Nay 2024, Scheduled 1 June 2024)

Fault Diagnosis Output of Motor Bearings Based on ReliefF Feature Selection

Ming Tang*, Aiyuan Wang, and Zhentian Zhu

School of Electrical Engineering, Shanghai Dianji University, Shanghai 201306, China

ABSTRACT: The problem of unstable vibration signal and accurate fault feature extraction of motor bearing fault causes the low accuracy of motor bearing fault diagnosis. In order to improve the accuracy of motor bearing fault diagnosis, the variational mode decomposition (VMD) is used to decompose the vibration signal and combine with the convolutional neural network (CNN). The bearing faults are categorized into inner ring wear, outer ring wear, and cage fracture; then each category of faults is further subdivided into the degree of loading, which is categorized into 0, 25%, and 50%, with a total of 9 cases. In order to select sensitive fault features, the vibration signals of motor bearings in three dimensions are collected, decomposed into multiple endowment modal function (IMF) components by VMD. The energy entropy of each IMF in each dimension is extracted, and the sensitive fault features are selected by feature selection (ReliefF), and then input into CNN for fault diagnosis. At the same time, the fault diagnosis of transverse vibration signal and three-dimensional vibration signal is also carried out respectively. The experimental results show that the accuracy of the method is greatly improved, and the fault diagnosis can be realized.

1. INTRODUCTION

With the rise of industry 5.0, the demand for motors is increasing. An important part of the motor is motor bearing [1, 2]. Bearings are susceptible to corrosion, foreign bodies, etc. Once damaged, it will have a serious impact on the operation of the entire motor [3]. Bearing failures have the highest incidence, accounting for 40% of overall motor failures. The fault diagnosis of motor bearing can be divided into two parts. The first part is to extract its fault characteristics, and the second part is fault identification. If the motor bearing is abnormal or faulty, its fault characteristics can be reflected in the vibration signal spectrum. By collecting the vibration signal and decomposing the characteristics in the signal, it is one of the commonly used methods for the first half of the motor bearing fault [4]. Ref. [5] analyses motor bearings and then uses orthogonal wavelet method to diagnose the location where the fault occurs. Ref. [6] is an early fault diagnosis method for rolling bearings based on VMD. VMD is used to decompose the bearing fault signal, and then the extracted fault features are input into the envelope spectrum for analysis to diagnose the bearing fault. Ref. [7] uses a combination of adaptive variational modal decomposition and least squares support vector machine to decompose multi-frequency complex signals for fault diagnosis of rolling bearings. A bearing fault diagnosis method based on CNN and Support Vector Machine (SVM) is proposed in [8]. The proposed method is verified on the dataset from the Case Western Reserve University (CWRU) and the Machine Failure Prevention Technology (MFPT). Ref. [9] uses short-time Fourier transform to convert multi-channel original bearing signals into two-dimensional time-frequency images and sends them to the residual network for fault classification. Ref. [10] proposes a bearing fault diagnosis method based on information entropy. The integrated empirical mode decomposition (EEMD) was used to extract the information entropy of each order component, and the SVM after parameter optimization was used for bearing fault diagnosis. In order to improve the accuracy of bearing fault diagnosis. Ref. [11] optimizes the parameters of VMD, extracts the kurtosis coefficient and energy equivalent of the reconstructed signal, and quantifies the fault signal. In [12], a method combining wavelet packet energy entropy and deep belief network (DBN) is proposed to diagnose bearing faults. Firstly, the wavelet packet is used to decompose the fault vibration signal of the bearing, and then the feature vector is composed in the form of energy entropy. The depth model of DBN is used to identify the fault of the energy entropy feature vector. Ref. [13] uses a one-dimensional convolutional neural network for fault feature extraction and fault diagnosis output. Although artificial fault feature extraction is omitted, it will cause the problem of feature loss.

In this paper, a fault diagnosis mode combining variational mode decomposition (VMD) and convolutional neural network (CNN) is adopted. The vibration signals of the transverse, radial and axial dimensions of the bearing are collected and decomposed by VMD. The energy entropy fault feature vector is extracted from the decomposed IMF components, and then the feature weight (ReliefF) is used to select the sensitive fault feature vector. Finally, it is input into the convolutional neural network (CNN) for bearing fault diagnosis.

* Corresponding author: Ming Tang (1979705361@qq.com).

2. FAULT SIGNAL DECOMPOSITION AND FEATURE EXTRACTION

2.1. Principle of Variational Modal Decomposition

VMD is an analysis method applied to signal analysis. The original composite signal is decomposed into several frequency modulation sub-signals in different frequency bands, which effectively avoids the phenomenon of modal aliasing and is conducive to the analysis of fault information and the feature extraction of fault signals.

VMD algorithm is used to decompose the fault signal into several intrinsic mode functions, which are expressed as:

$$u_k(t) = w_k(t)\cos(\varphi_k(t)) \tag{1}$$

In formula (1), $u_k(t)$ is the kth signal component after VMD decomposition; $w_k(t)$ is the instantaneous amplitude of the signal; $\varphi_k(t)$ is the phase of the signal.

In order to obtain the center frequency and bandwidth of each mode $U_k(t)$, Hilbert transform is performed on each mode to construct an analytical function to obtain the unilateral spectrum, and then the spectrum is shifted to the baseband of the estimated large center frequency. The modal bandwidth is determined by the calculated value obtained from the L^2 norm. The model is the following formula

$$\min\left\{\sum_k \left\| dt\left[\left(\delta(t)+\frac{j}{\pi t}\right)*U_k(t)\right]e^{-j\omega t}\right\|_2^2\right\}$$
$$s.t.\sum_{k=1}^{k} U_k(t) = f(t) \tag{2}$$

In formula (1), $\delta(t)$ is the unit impulse function; * is the convolution; dt is the derivation of function; $s.t.$ is the constraint condition.

In order to solve the constraint problem, the quadratic penalty factor α and Lagrange multiplier λ are introduced to obtain the augmented Lagrange equation, which is transformed into an unconstrained problem. Among them, the value of α can ensure the reconstruction accuracy of the signal in the presence of Gaussian noise, and the Lagrange multiplier makes the constraint conditions strict. The equation is solved using the alternating direction multiplier method to achieve alternating optimisation for U_k, ω_k, and λ.

$$\hat{U}_k^{n+1}(\omega) = \frac{\hat{f}(\omega) - \sum_{i\neq k}\hat{U}_i(\omega) + \frac{\hat{\lambda}(\omega)}{2}}{1+2\alpha(\omega-\omega_k)^2} \tag{3}$$

$$\omega_k^{n+1} = \frac{\int_0^\infty \omega\left|\hat{U}_k(\omega)\right|^2 d\omega}{\int_0^\infty \left|\hat{U}_k(\omega)\right|^2 d\omega} \tag{4}$$

$$\hat{\lambda}^{n+1}(\omega) = \hat{\lambda}^n(\omega) + \tau\left(\hat{f}(\omega) - \sum_k \hat{u}_k^{n+1}(\omega)\right) \tag{5}$$

According to the given solution accuracy ε, the iteration is stopped by satisfying Equation (6).

$$\sum_k \frac{\left\|\hat{U}_k^{n+1} - \hat{U}_k^n\right\|_2^2}{\left\|\hat{U}_k^n\right\|_2^2} < \varepsilon \tag{6}$$

The iterative solution is initialised with $\left\{\hat{U}_k^1\right\}$, $\{\hat{\omega}_k^1\}$, $\{\lambda^1\}$ with the maximum number of iterations. According to formulas (3), (4), and (5), U_k, ω_k, λ are updated in turn. Once the convergence accuracy is satisfied, or the maximum number of iterations is reached, the iteration is stopped, and the final k modal components and center frequencies are obtained.

The advantage of VMD is that it can deal with nonlinear and non-stationary signals, and can adaptively select the number of decomposed frequency bands. It is widely used in signal processing, image processing, machine learning, and other fields.

2.2. Energy Entropy Characteristics

Energy entropy can reflect the energy change between the normal signal and fault signal of the bearing [14]. Different IMF components can be obtained by VMD decomposition of motor bearing fault signals. Due to the different fault signals, the energy is also different. The energy is distributed in each component of the decomposition, which can be used as a feature vector to reflect the fault signal. The expression for energy entropy is:

$$H_i = -\sum_{r=1}^{n} p_r \ln p_r \tag{7}$$

p_r denotes the share of the energy of the rth IMF component in the total energy $E_{sum}\cdot p_r = E_r/E_{sum}$, and H_i denotes the ith IMF energy entropy, $i = (1, 2, \ldots n)$.

2.3. ReliefF Feature Selection

ReliefF algorithm is a feature weighting algorithm that gives different differentiated weights to the features according to the degree of association between each feature and type. If the weight is less than a certain set value, the feature values will be removed, and the remaining will form a new feature subset to complete the feature selection process, thereby improving the accuracy of fault feature extraction. However, it cannot handle more than two dimensional data well [15], so the ReliefF algorithm is proposed. The steps of the ReliefF feature selection algorithm to calculate the weight are as follows:

(1) A training set (R) is drawn from the overall data.

(2) Then from this training set, K similar sets are extracted, and each feature set proposes k such sets.

(3) Finally, the weight value of the most feature set is iterated again:

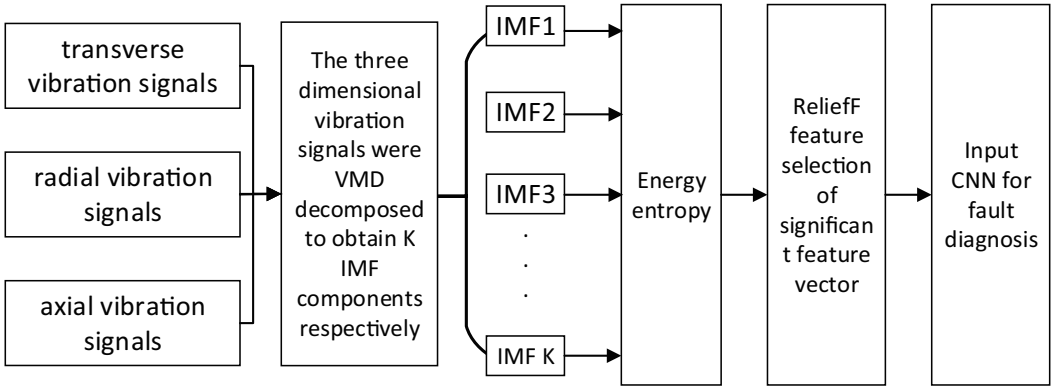

FIGURE 1. Fault diagnosis flow chart.

$$W(A) = W(A) - \frac{\sum_{j=1}^{K} diff(A,R,H_j)}{mk} + \sum_{C\neq clsss(R)} \left[\frac{p(C)\sum_{j=1}^{k} diff(A,R,M_j(C))}{1-p(Class(R))} \right] /(mk) \quad (8)$$

In the above formula, $W(A)$ is the weight value, and sample R_1 and sample R_2 are subtracted above feature A to find the difference that is $diff(AR_1R_2)$. $M_j(C)$ is the jth similarity set in class C.

$$diff(A,R_1,R_2) = \begin{cases} 0, & A \text{ is discrete and } R_1[A] = R_2[A] \\ 1, & A \text{ is discrete and } R_1[A] \neq R_2[A] \\ \dfrac{|R_1[A] - R_2[A]|}{\max(A) - \min(A)}, & A \text{ is continuous} \end{cases} \quad (9)$$

3. CONVOLUTIONAL NEURAL NETWORK

Convolutional neural network (CNN) belongs to a feed-forward neural network, which is a common deep learning network model. CNN can automatically extract features. Compared with traditional feature methods, it can greatly improve the accuracy of recognition [16]. CNN consists of an input layer, a convolutional layer, a pooling layer, a fully connected layer, and an output layer.

The main role of the convolutional layer is to convolve the convolutional kernel with part of the receptive domain of the input signal, to extract the features into part of the receptive domain of the input signal, and to use the result of the convolution of multiple input features under the activation function as the output feature quantity.

$$a_i^{l+1} = f(y_i^{l+1}(j)) \quad (10)$$

$$y_i^{l+1}(j) = K_i^l * x^l(j) + b_i^l \quad (11)$$

In the above formula, a_i^{l+1} is the final output of the jth neuron in the ith feature plane of the lth layer; $y_i^{l+1}(j)$ is the result of convolution operation; $x^l(j)$ is the jth part of the layer l receptive field; K_i^l is the weight vector of the ith convolution kernel in the lth layer; b_i^l is the bias of the ith convolution kernel in the lth layer; $f()$ is the activation function.

The gradient saturation effect is not avoided. The ReLU piecewise function is applied in the deep convolutional network, and its expression is as follows:

$$a_i^{l+1}(j) = f\left(y_i^{l+1}(j)\right) = \max\left\{0, y_i^{l+1}(j)\right\} = \begin{cases} y_i^{l+1}(j), & y_i^{l+1}(j) \geq 0 \\ 0, & y_i^{l+1}(j) < 0 \end{cases} \quad (12)$$

The pooling layer is to reduce the dimension of the feature map and filter out the residual parameters in the data, which can prevent feature degeneration and over-fitting of the network. The general pooling methods are mean pooling and maximum pooling. In this paper, the maximum pooling layer is used to output the maximum value of the receptive field of each feature surface, ignoring the secondary factors and extracting the important features, and its expression is as follows:

$$p_i^{l+1}(j) = \max_{(j-1)w+1\leq t\leq jw} q_i^l(t) \quad (13)$$

where $p_i^{l+1}(j)$ is the pooling result for neurons in the $(l+1)$th layer; w is the width of the pooling area; $q_i^l(t)$ is the value of the tth neuron inside the ith feature in the lth layer.

When the input signal is processed by the convolution layer and pooling layer, it will be transformed into high-level information features, and one or more fully connected layers are connected to classify the information features.

4. FAULT DIAGNOSIS PROCESS

Firstly, the vibration signals of the transverse, radial and axial dimensions of the bearing fault are collected and decomposed by the VMD method. Each dimension is decomposed into K IMFs. Then, K energy entropy features are extracted from K IMF components respectively, and then the ReliefF algorithm is used to select the features with large weight values to form a sensitive feature set. Finally, it is input into the CNN fault

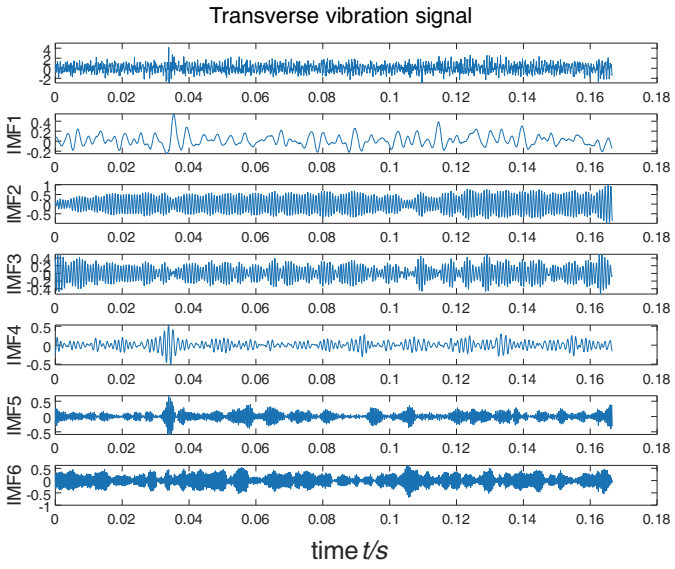

FIGURE 2. Transverse decomposition diagram of inner ring wear.

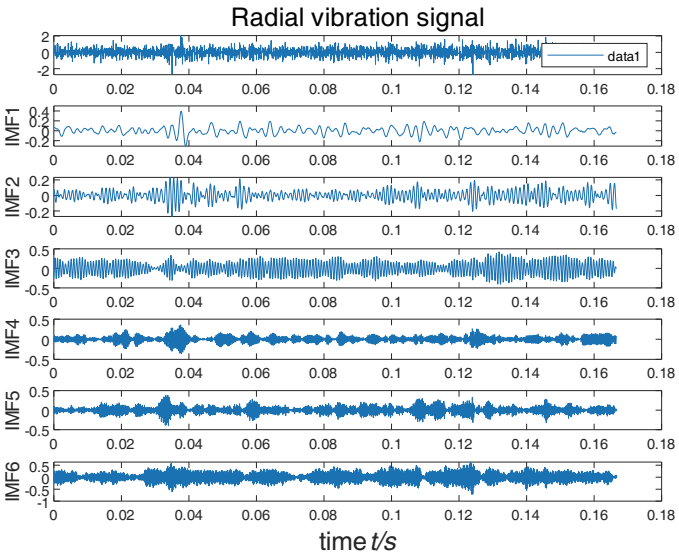

FIGURE 3. Radial decomposition diagram of inner ring wear.

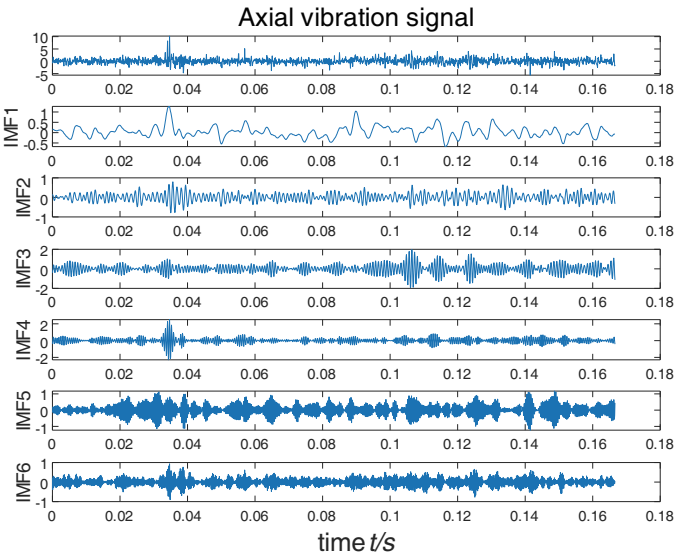

FIGURE 4. Decomposition diagram of wear axial of inner ring.

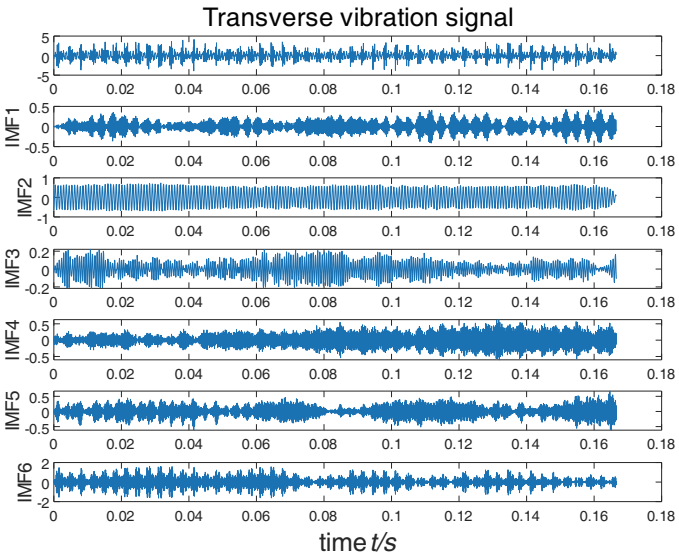

FIGURE 5. Transverse decomposition diagram of outer ring wear.

diagnosis model for testing. The flowchart of bearing fault diagnosis is shown in Figure 1.

5. EXPERIMENTAL RESULTS AND ANALYSIS

5.1. Experimental Data

Bearing faults are divided into three types: inner ring wear, outer ring wear, and cage fracture. Then, each fault is subdivided into 0 load, 25% load, and 50% load, a total of 9 bearing fault states. For each fault, 50 samples are taken, and the training set and test set are divided according to 70%, that is, 35 training samples and 15 test samples. Before entering CNN for fault diagnosis, each fault condition is labeled for easy output. For example, the case where the inner ring wear load is 0 is marked as 1, and so on. The bearing fault table is shown in Table 1.

TABLE 1. Bearing fault table.

Type of fault	Loadinglevel	Label
Inner ring wear	0	1
Inner ring wear	25%	2
Inner ring wear	50%	3
Outer ring wear	0	4
Outer ring wear	25%	5
Outer ring wear	50%	6
Cage fracture	0	7
Cage fracture	25%	8
Cage fracture	50%	9

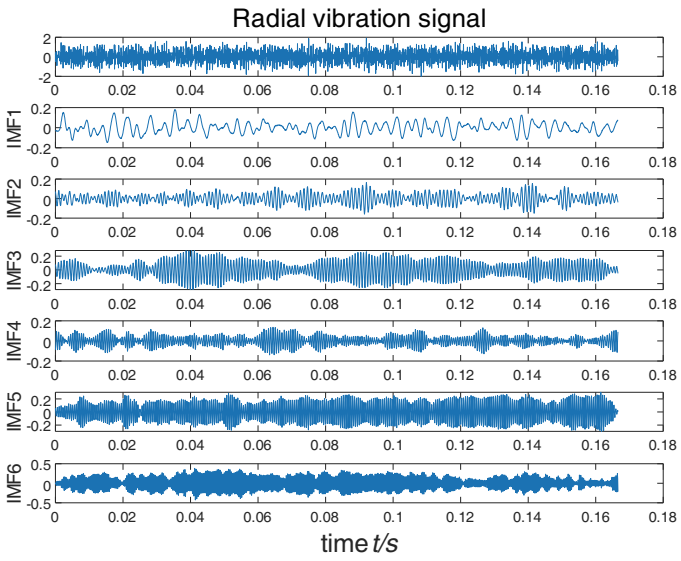

FIGURE 6. Radial decomposition diagram of outer ring wear.

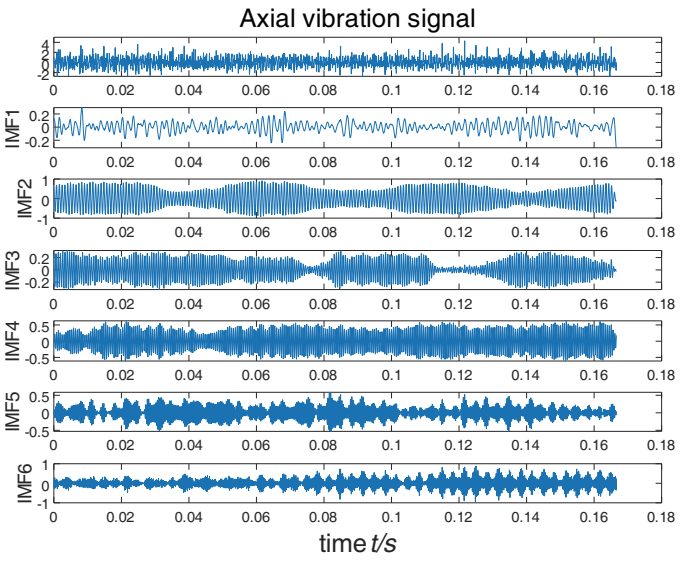

FIGURE 7. Decomposition diagram of wear axial of outer ring.

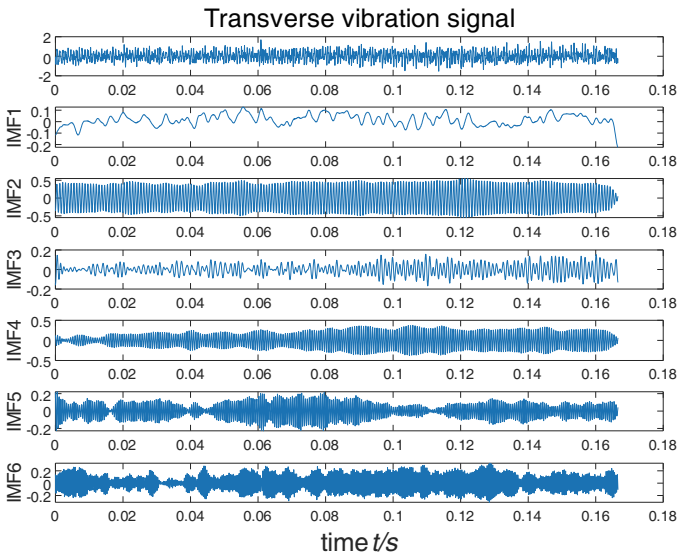

FIGURE 8. Transverse decomposition diagram of cage fracture.

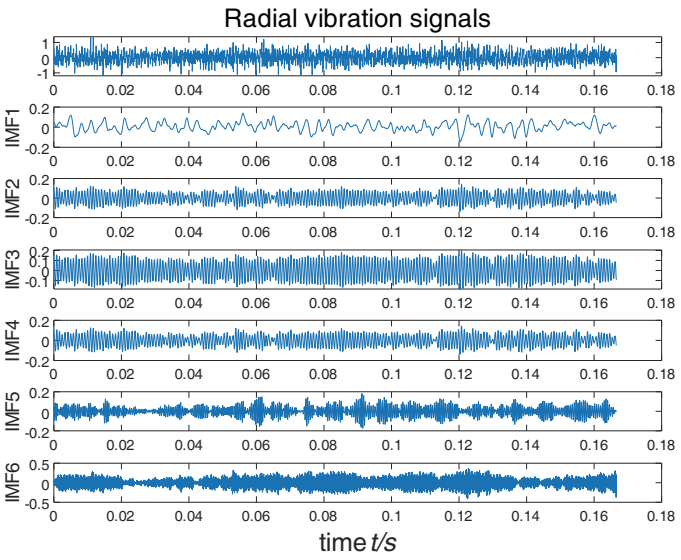

FIGURE 9. Radial decomposition diagram of cage fracture.

5.2. VMD Energy Entropy

VMD can effectively separate the similar components in the signal, and has good robustness, which can effectively solve the interference problem at the end of the signal. The two main parameters of the penalty factor α and the number of modes K in the VMD are set to 2500 and 6, respectively, and the remaining parameters are set to default values. Due to the space relationship, only the decomposition graph with the load of 0 is displayed. The transverse, radial and axial decomposition diagrams of the wear of the inner ring are shown in Figures 2, 3, 4, respectively; the transverse, radial and axial decomposition diagrams of the wear of the outer ring are shown in Figures 5, 6, 7, respectively; the transverse, radial and axial decomposition diagrams of the cage fracture are shown in Figures 8, 9, 10, respectively.

Comparing the above 9 decomposition diagrams, the energy of different frequency bands in each dimension of each bearing fault is different. In feature extraction, VMD is used for decomposition processing, and energy entropy is used as feature vector to distinguish bearing fault types in turn. The fault feature vector of energy entropy is shown in Figures 11, 12, and 13.

5.3. Feature Selection

Six energy entropies decomposed from the three dimensions of each bearing fault state are extracted, with a total of 18 features. The ReliefF algorithm is used to calculate the feature weight, and the proportion of feature weight is shown in Figure 14. From the weight ratio, it can be seen that the weight ratio of each IMF component is different. The lower the weight ra-

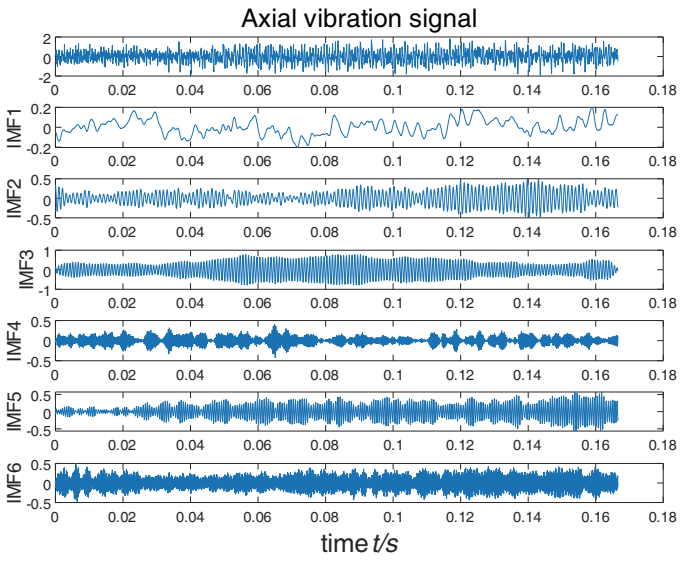

FIGURE 10. Decomposition diagram of wear axial of cage fracture.

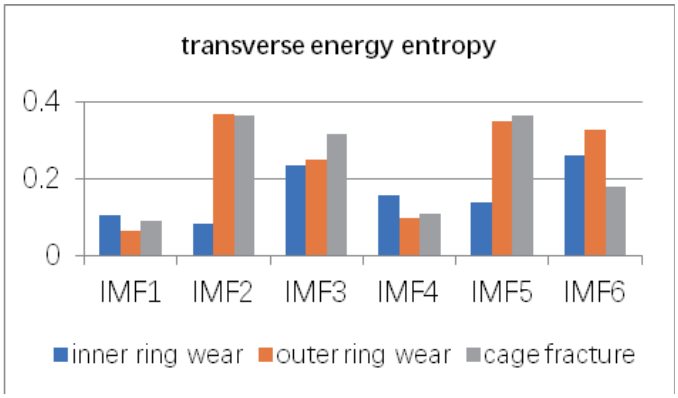

FIGURE 11. Comparison diagram of transverse energy entropy.

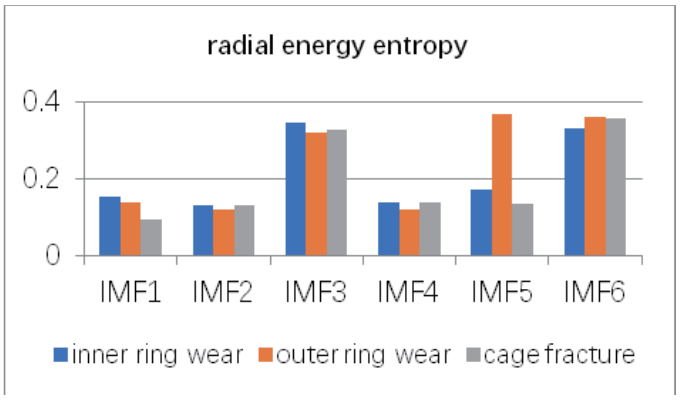

FIGURE 12. Comparison diagram of radial energy entropy.

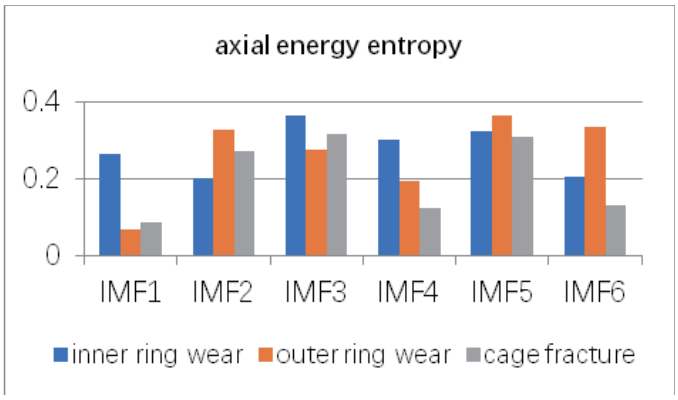

FIGURE 13. Comparison diagram of axial energy entropy.

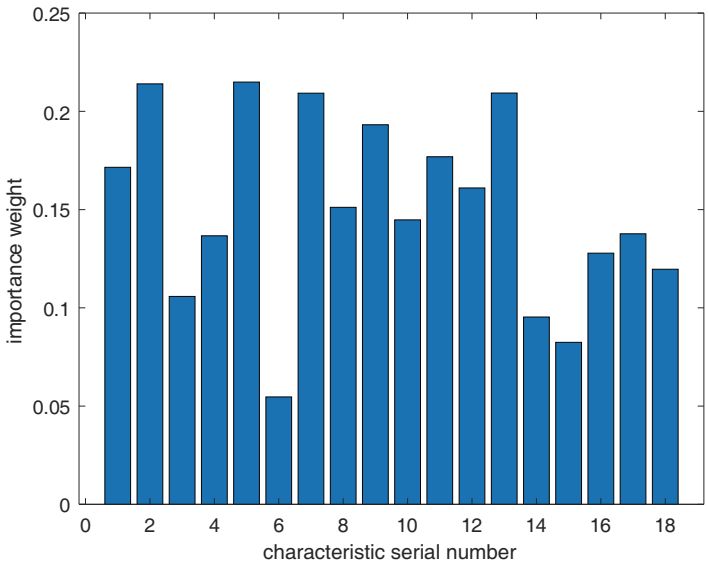

FIGURE 14. The weight proportion diagram of each IMF component.

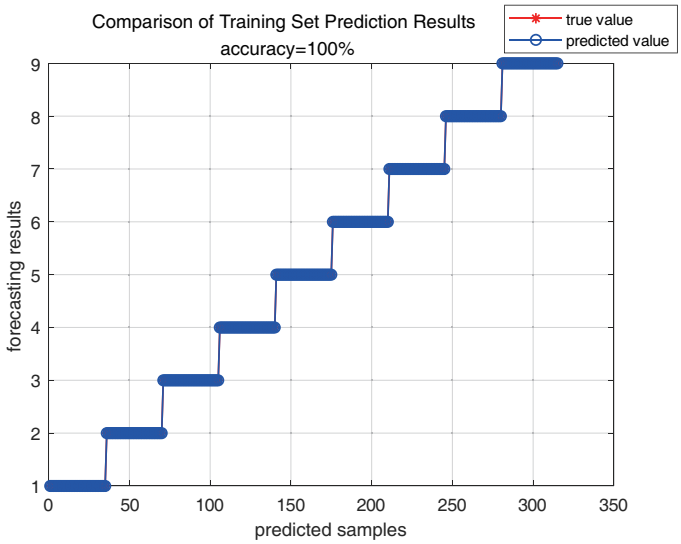

FIGURE 15. Training set classification results.

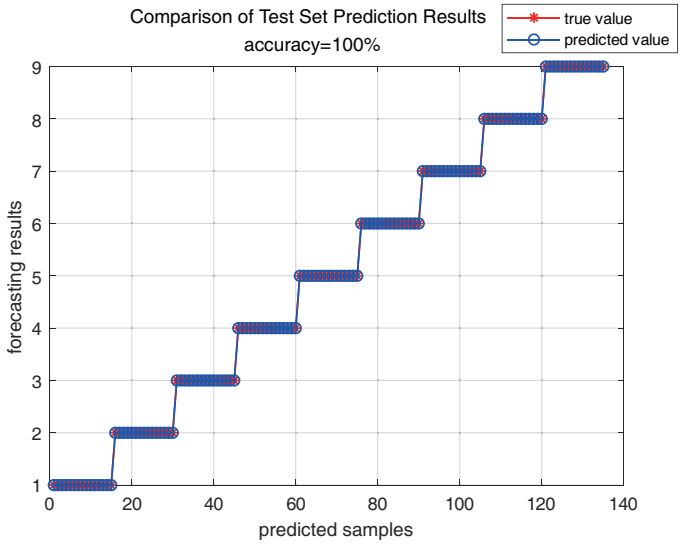

FIGURE 16. Test set classification results.

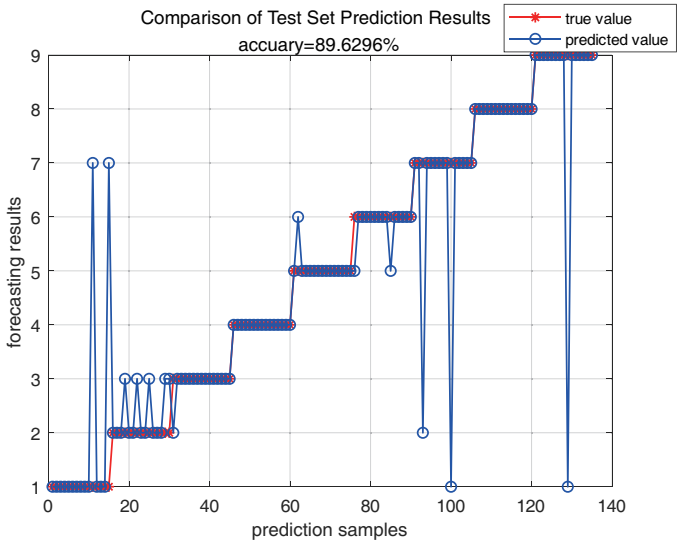

FIGURE 17. Classification results of transverse signals.

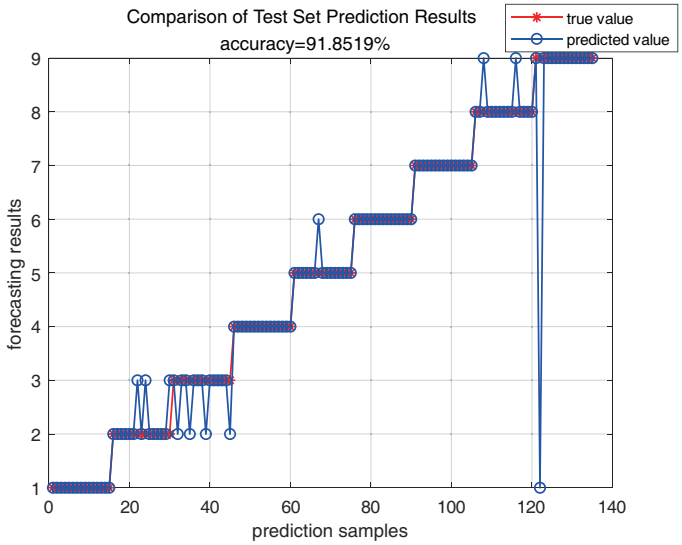

FIGURE 18. Classification results of three dimensional signals.

tio is, the less sensitive it is. The top 50% of the weight ratio is selected as the sensitive fault feature set. Finally, the sensitive fault feature set is input into CNN for fault diagnosis.

5.4. Fault Recognition

The fault set composed of sensitive fault feature vectors selected by ReliefF feature selection is input into the built CNN model for training and disrupted for testing. 50 samples were selected for each fault type, and a total of 450 samples were tested. The motor bearing fault test results are shown in Figures 15 and 16, respectively. In order to show the accuracy of this method, the accuracy of fault identification for collecting transverse vibration signals in a single dimension is shown in Figure 17. The fault recognition accuracy of the three-dimensional vibration signal is shown in Figure 18.

According to the previous training set and test set, it can be seen that the 315 samples of the training set are all accurate, and the 135 samples of the test set have only one prediction error, with an accuracy rate of 99%. The fault recognition accuracy of only collecting single dimension transverse signal is only 89%, and the fault recognition rate of three dimension signals is only 92%. After comparison, the fault recognition rate of the method of collecting three-dimensional vibration signals and selecting sensitive fault feature vectors by ReliefF algorithm is higher than that of the first two methods.

6. CONCLUSION

Aiming at the common faults of motor bearing: inner ring wear, outer ring wear, and cage fracture, this paper proposes a fault diagnosis method of motor bearing based on ReliefF feature selection. Firstly, the vibration signals of transverse, radial and axial dimensions are collected, and the collected vibration signals are decomposed by VMD, and the energy entropy of each decomposed IMF is calculated as the feature vector. In order to

improve the fault diagnosis rate, the ReliefF algorithm is used to calculate the feature weight of the feature vector, eliminate the feature vector with lower weight, and select the sensitive feature vector. Finally, it is input into the built CNN model for fault diagnosis output. In order to show the superiority of this method, the three-dimensional vibration signals of only collecting transverse vibration signals and not using ReliefF to screen out sensitive feature vectors are classified for fault diagnosis. The experimental results show that the correct rate of the test set of the transverse vibration signal is only 89%, and the correct rate of the test set of the three-dimensional vibration signal is slightly improved to 91%, while the correct rate of the sensitive feature vector of the three-dimensional vibration signal screened by ReliefF is greatly improved to 99%. This method not only retains the important feature vector, but also eliminates the influence of the irrelevant feature vector on the diagnosis result, which ensures the accuracy and shows the superiority of this method.

In the subsequent research, it will limited to not only the type of fault and the degree of load, but also the degree of fault diagnosis.

REFERENCES

[1] Shao, S., R. Yan, Y. Lu, P. Wang, and R. X. Gao, "DCNN-based multi-signal induction motor fault diagnosis," *IEEE Transactions on Instrumentation and Measurement*, Vol. 69, No. 6, 2658–2669, 2020.

[2] Yang, H., X. Li, and W. Zhang, "Interpretability of deep convolutional neural networks on rolling bearing fault diagnosis," *Measurement Science and Technology*, Vol. 33, No. 5, 055005, 2022.

[3] Feng, D. and Y. Li, "Research on intelligent diagnosis method for large-scale ship engine fault in non-deterministic environment," *Polish Maritime Research*, Vol. 24, No. s3, 200–206, 2017.

[4] Wang, B., "Design of motor condition detection and fault diagnosis system based on vibration characteristics," Qingdao University, 2021.

[5] Xie, G. N., Y. Tong, and W. B. Lu, "Application of wavelet in fault diagnosis of coal mining machine asynchronous motor," *Control Engineering*, Vol. 20, No. 4, 2013.

[6] Tao, Z., Z. Pang, M. Wang, *et al.*, "Early fault diagnosis method of rolling bearings based on VMD," *Journal of Beijing Institute of Technology*, Vol. 45, No. 2, 103–110, 2019.

[7] Jin, Z., P. Mu, Y. Zhang, *et al.*, "An improved VMD and its application in bearing fault diagnosis," *Mechanical Design and Manufacturing*, Vol. 2, 42–46, 2022.

[8] Yuan, L., D. Lian, X. Kang, Y. Chen, and K. Zhai, "Rolling bearing fault diagnosis based on convolutional neural network and support vector machine," *IEEE Access*, Vol. 8, 137 395–137 406, 2020.

[9] Korobovaen, , Sevalnevgs, Gromovvi, *et al.*, "Steels for the manufacture of roller bearings for special purposes," *Trudy VIAM*, No. 11, 105, 2021.

[10] Zhang, M., M. Yang, M. Yang, S. Li, *et al.*, "Influence of carbon and nitrides in high-nitrogen stainless bearing steel on mechanical properties," *Journal of Iron and Steel Research*, Vol. 24, No. 5, 18–23, 2012.

[11] Tang, G., L. Zhu, and X. Hu, "Rolling bearing fault diagnosis based on optimised VMD and deep confidence network," *Bearing*, No. 10, 47–53, 2020.

[12] Zhao, G. Q., Z. D. Jiang, C. Hu, Y. Gao, and G. Niu, "Bearing fault diagnosis based on wavelet packet energy entropy and DBN," *Journal of Electronic Measurement and Instrumentation*, Vol. 33, No. 2, 32–38, 2019.

[13] Ince, T., S. Kiranyaz, L. Eren, M. Askar, and M. Gabbouj, "Real-time motor fault detection by 1-D convolutional neural networks," *IEEE Transactions on Industrial Electronics*, Vol. 63, No. 11, 7067–7075, Nov. 2016.

[14] Tian, S. and Z. Kang, "Vibration analysis of circuit breaker mechanical failure based on improved variational modal decomposition and SVM," *Vibration and Shock*, Vol. 38, No. 23, 90–95, 2019.

[15] Zhang, L., J. Wang, Y. Zhao, *et al.*, "Relief-based combinatorial feature selection," *Journal of Fudan (Natural Science Edition)*, No. 5, 893–898, 2004.

[16] Li, K., Z. Wei, and H. Song, "Vehicle colour recognition based on squeezenet convolutional neural network," *Journal of Chang'an University (Natural Science Edition)*, Vol. 40, No. 4, 109–116, 2020.