Academic Journal of Science and Technology
ISSN: 2771-3032 | Vol. 9, No. 3, 2024

Bearing Fault Diagnosis Based on VMD-WPT-ViT for Wind Turbine

Hao Zhen[1], Shaoting Wang[2], Longfeng Han[3], Min Liu[3]

[1]Huaneng Beijing Thermal Power Co. Ltd., Beijing 100023, China
[2]State Grid Digital Technology Holding Co.,LTD, Beijing 100032, China
[3]School of Control and Computer Engineering, North China Electric Power University, Baoding 071003, China

Abstract: As an important part of wind turbine, the fault diagnosis of rolling bearings is helpful to maintain the stable operation of the fan. However, most of the existing fault diagnosis methods focus on one-dimensional time series data, which is not only difficult to extract multi-level feature information, but also fails to consider the noise problem in the signal. Therefore, to solve the above problems, a noise reduction algorithm combining variational mode decomposition (VMD) and wavelet packet threshold noise reduction (WPT) is proposed, and Vision-Transformer (ViT) model is combined to realize the fault diagnosis of rolling bearings of wind turbines. VMD is used to decompose the original signal and find out the noisy component. WPT is used to de-noise the component, and the de-noised component and other components are reconstructed to form a pure signal and converted into a two-dimensional image. Finally, a diagnostic model is built based on Vision-transformer. The experimental results show that the proposed method can effectively improve the diagnostic accuracy by extracting the advanced global features of the signal.

Keywords: Lrolling bearing; Two-dimensional gray map; Fault diagnosis; Vision-Transformer; VMD; Wavelet packet threshold denoising.

1. Introduction

In today's society, with the rapid development of the world economy, non-renewable energy has been over-exploited by human beings, resulting in an urgent inventory. Wind energy as a stable and clean renewable energy storage is rich, so wind power has become an important choice for energy development in recent years. According to the statistics of relevant departments, in 2021, China's installed capacity of wind turbines increased by 15,911 units, until the end of 2021, the total installed capacity of more than 170,000 units, compared with thermal power generation and hydroelectric power generation, wind power ranked third in the total installed capacity of the grid, accounting for 13.5%. In 2022, the newly installed capacity affected by the epidemic will be 22.52GW, down 8.83% year-on-year, but in terms of the bidding situation, 2023 May usher in the first peak in the history of wind power installed capacity [1][2]. Due to the harsh environment of the fan, it is easy to fail, and the bearing, as the main part of the fan, also has a high incidence of failure. Fault diagnosis is of great significance for the safe and stable operation of the fan and the reduction of later maintenance costs.

At present, the fault diagnosis methods of rolling bearings mainly include methods based on signal processing, model-based methods and deep learning methods [3]. However, in the actual situation, due to noise disturbance, modeling error and other reasons, it is difficult to establish an accurate model, which makes the fault diagnosis effect poor, relatively speaking, the latter two methods are more popular with researchers at home and abroad. In literature [4], the bearing vibration acceleration signal is converted into frequency domain signal by fast Fourier transform, and the signal features are extracted by 1D-CNN for fault diagnosis. In literature [5], EEMD was first used to decompose the original signal to extract its energy features, and then the feature set was input into the SVM model optimized by differential evolution algorithm to identify and monitor high-speed bearings. In literature [6], an improved coarse-grained multi-scale dispersal entropy is proposed. After the vibration data is decomposed by variational mode decomposition and the optimal IMF component is selected according to the mutual relation number, the nonlinear features of IMF component are extracted by this improved method and the fault diagnosis is performed by PNN. In literature [7], the optimal parameters of Morlet wavelet were determined by PSO, and combined envelope demodulation based on iterative two-mean improvement of kurtosis index was used to classify rolling bearing faults. In literature [8], principal component analysis is used to reduce the dimensionality of the original vibration data and extract corresponding features. Finally, deep belief network is used to realize fault diagnosis. In literature [9], multiple attention modules are introduced on the basis of one-dimensional convolutional neural networks to improve diagnosis accuracy by enhancing fault correlation features. Although the above research methods have obtained good diagnostic results, the bearing vibration data is inevitably polluted by noise in the collection process, and these noises will cover up the original signal data, and have a certain impact on the subsequent feature extraction and fault diagnosis.

In order to solve the noise problem of the original signal, literature [10] firstly uses CEEMDAN (adaptive noise complete set empirical Mode decomposition) to decompose the original signal, and combines wavelet threshold (WT) for noise reduction. After signal decomposition, CMSE is used to select noisy components, and then WT is used for noise reduction of these components. The vibration signal after noise reduction is reconstructed with other components, and the experiment shows that the method has a good effect of noise reduction. In literature [11], in order to solve the problem of noise interference, wavelet threshold de-noising

algorithm is used to de-noise the original signal, and finally the signal frequency is extracted by combining the analytic mode decomposition algorithm, thus improving the diagnostic accuracy. Literature [12] proposed a method combining EMD decomposition and probabilistic filtering technology to effectively remove electromagnetic interference of rolling body bearings. Although the above research carried out noise reduction processing on the original signal, the one-dimensional model was generally used for diagnosis, and the signal features could not be well extracted, thus reducing the diagnostic accuracy.

To solve these problems, this paper proposes a Wavelet Packet Threshold (WPT) algorithm based on Variational Mode Decomposition (VMD). First, VMD was used to decompose the original signal, and the IMF component containing noise was found by calculating the mutual relation number. In order to obtain the pure signal, the remaining components without noise were reconstructed with the IMF component treated with WPT noise reduction. As one-dimensional diagnostic model is difficult to train and feature extraction is single, this paper will also build Vision-Transformer model and input de-noised data into the model after converting it into two-dimensional grayscale images to extract multi-scale features to improve diagnostic accuracy. The feasibility and effectiveness of the proposed method are verified by fault diagnosis of different parts of rolling bearing damage in CWRU data set.

2. VMD-WPT Joint Noise Reduction Principle

2.1. VMD decomposition principle

Variational mode decomposition is a new mode decomposition algorithm proposed by Konstantin et al in 2014. Compared with the decomposition algorithms of EMD family, such as EMD, EEMD, CEEMDAN, ICEEMDAN, etc., VMD does not need to add Gaussian noise before decomposition. Although CEEMDAN and other algorithms add white noise of positive and negative sequences to cancel each other in the decomposition process, there are still some reconstruction errors.

Variational mode decomposition is a frequency decomposition estimation method [13], which is based on Wiener Filter, Hilbert transformer and heterodyne demodulation for signal decomposition. Using VMD to process signals can avoid disadvantages such as EMD family mode confusion and end effect [14]. Variational mode decomposition is an adaptive decomposition method, which can determine the number of IMF components required by the data according to the actual situation. The VMD algorithm decomposed the actual signal into several eigenmodal components through iterative search, and the bandwidth of these components in the frequency domain has specific sparse properties [15].

The decomposition process of VMD is actually the process of finding the optimal solution of the variational problem, which can be realized by solving the construction sum of the corresponding variational problem. The constrained variational problems corresponding to VMD decomposition are as follows:

$$\begin{cases} \min\limits_{\{u_k\},\{\omega_k\}} \left\{ \sum\limits_k \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 \right\} \\ s.t. \ \sum\limits_k u_k = f \end{cases} \tag{1}$$

Among, $\{u_k\} = \{u_1, \dots, u_k\}$ Is the decomposed IMF components, $\{\omega_k\} = \{\omega_1, \dots, \omega_k\}$ Represents the central frequency of each IMF component.

Each mode component can be obtained by solving the above formula.u_k.The expression formula of is:

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

Among,αIs a quadratic penalty parameter,λIs the Lagrange multiplication operator.Center frequency ω_k The expression formula of is:

$$\omega_k^{n+1} = \frac{\int_0^\infty \omega |\hat{u}_k(\omega)|^2 d\omega}{\int_0^\infty |\hat{u}_k(\omega)|^2 d\omega} \tag{3}$$

The specific steps are as follows:

Algorithm 2: Complete optimization of VMD

Initialize $\{\hat{u}_k^1\},\{\omega_k^1\},\hat{\lambda}^1$,n←0

Repeat

n ← n+1

for k=1: K do

Update $\hat{u}_k$ for all $\omega \geq 0$:

$$\hat{u}_k^{n+1}(\omega) \leftarrow \frac{\hat{f}(\omega) - \sum_{i<k} \hat{u}_i^{n+1}(\omega) - \sum_{i>k} \hat{u}_i^n(\omega) + \frac{\hat{\lambda}^n(\omega)}{2}}{1 + 2\alpha(\omega - \omega_k^n)^2} \tag{4}$$

ω_k:

$$\omega_k^{n+1} = \frac{\int_0^\infty \omega |\hat{u}_k^{n+1}(\omega)|^2 d\omega}{\int_0^\infty |\hat{u}_k^{n+1}(\omega)|^2 d\omega} \tag{5}$$

End for

Dual ascent for all $\omega \geq 0$:

$$\hat{\lambda}^{n+1}(\omega) \leftarrow \hat{\lambda}^n(\omega) + \tau(\hat{f}(\omega) - \sum_k \hat{u}_k^{n+1}(\omega)) \tag{6}$$

Until convergence: $\sum_k \left\| \hat{u}_k^{n+1} - \hat{u}_k^n \right\|_2^2 / \|\hat{u}_k^n\|_2^2 < \epsilon$ (7)

In formula (7),ϵ is the accuracy of discrimination.

2.2. Wavelet packet threshold denoising principle

In engineering applications, wavelet denoising and smoothing filter denoising are often used. Smoothing filter noise reduction is mainly achieved by fitting the low frequency part of the signal and smoothing out the high frequency part. If the noise is in the high frequency part of the signal, it can be removed by smoothing operation; otherwise, the noise will be left [16]. The wavelet denoising is actually to decompose the signal with multi-resolution, but due to the limitation of its function scale, the wavelet transform will only decompose the low frequency part, ignore the high

frequency part, and do not further decompose. The above two noise reduction methods do not process the high-frequency part of the signal, and can not remove the noise point well, which will greatly affect the diagnostic accuracy of the subsequent diagnostic model.

Wavelet packet threshold noise reduction algorithm is an optimization algorithm of wavelet noise reduction, which provides a more delicate and detailed noise reduction method. In signal processing, the algorithm will further decompose the part of the high-frequency signal that is not operated by the wavelet transform, and select the frequency band with high matching degree to the original signal spectrum features. This method will achieve better noise reduction effect for non-stationary signals. The effect diagrams of the two kinds of decomposition are shown in Figure 1 and Figure 2. It can be seen from the figure that the wavelet decomposition only processes the low-frequency part of the signal, while the wavelet packet decomposition processes the high-frequency part of the signal while the low-frequency part is decomposed.

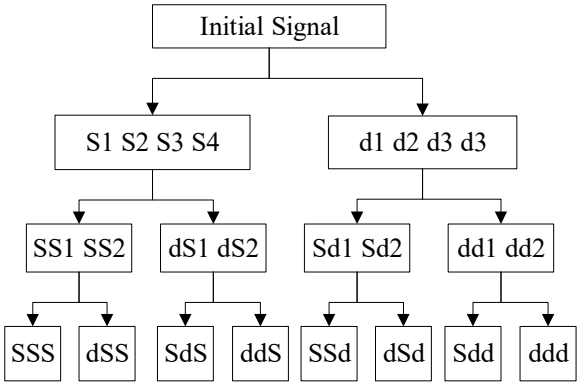

Figure 1. Wavelet packet decomposition diagram

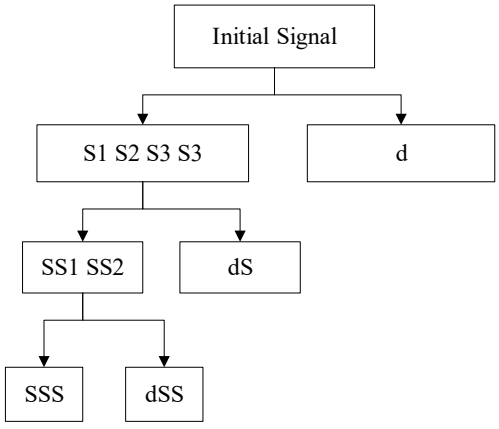

Figure 2. Wavelet decomposition diagram

Mathematically, the wavelet packet is composed of a set of linear functions. If it is assumed that $\varphi^0(t) = \varnothing(t)$and $\varphi^1 = \varphi(t)$are scale functions and wavelet functions respectively, $h(n)$ and $g(n)$ representing the coefficients of conjugate mirror filter banks, the wavelet packet can be expressed as the following recursive relation [17]:

$$\varphi^{2k}(t) = \sqrt{2}\sum_n h(n)\varphi^k(2t-n) \quad (8)$$

$$\varphi^{2k+1}(t) = \sqrt{2}\sum_n g(n)\varphi^k(2t-n) \quad (9)$$

The steps of wavelet packet threshold denoising are as follows:

(1) Select the appropriate wavelet basis function and the corresponding number of decomposition layers according to the characteristics of the original signal, and decompose it.

(2) The optimal wavelet basis is determined according to the given standard entropy.

(3) The optimal threshold is obtained through repeated tests, and the threshold is used to quantify the high-frequency coefficients of different scales.

(4) The low frequency coefficient and the high frequency coefficient after quantization are reconstructed by wavelet.

In the process of wavelet packet threshold denoising, the most important is the selection of threshold, it can be said that whether the selected threshold is appropriate will directly affect the quality of noise reduction.

3. Vision-Transformer Model

This paper adopts the signal-image conversion method proposed in literature [18], which randomly obtains fragment data of size M from the original signal and combines it into an image of M*M. The integer between 0 and 255 can be obtained from formula (10), and the value in the image after conversion in this paper can be obtained through this formula.

$$P(m,n) = round\{\frac{L[(m-1)\times M + n] - \min(L)}{\max(L) - \min(L)} \times 255 \quad (10)$$

Where, $P(m,n)$ represents the specific data of a certain position in the converted picture; The size of the converted two-dimensional gray image is M*M; L(i) represents the value of the i the data point in the sampled timing signal, whose total length is M2; round(x) is a rounding function that can round the data.

Before Vision-Transformer appeared, CNNs have always been the dominant model in Vision research. Both in terms of architecture scale and convolutional structure, CNNs provide strong feature extraction capability in deep learning vision research tasks. In the field of natural language processing, transformer, as a basic model, utilizes self-attention mechanism to greatly improve the performance of language modeling. It is the success of this model that has led people to gradually explore whether the transformer model can be used in the field of vision research. The 2020 vision transformer (VIT) model [19] can achieve global modeling capability by calculating global self-attention operations. ViT divides the input image into several patches (16*16), and then projects each Patch into Transformer as a vector of fixed length. The subsequent operation of encoder is exactly the same as that in original Transformer. However, because the ViT classifies the images, a special token is added to the input sequence, and the output corresponding to the token is the final output.

Vision-Transformer (ViT) network structure is shown in the following figure:

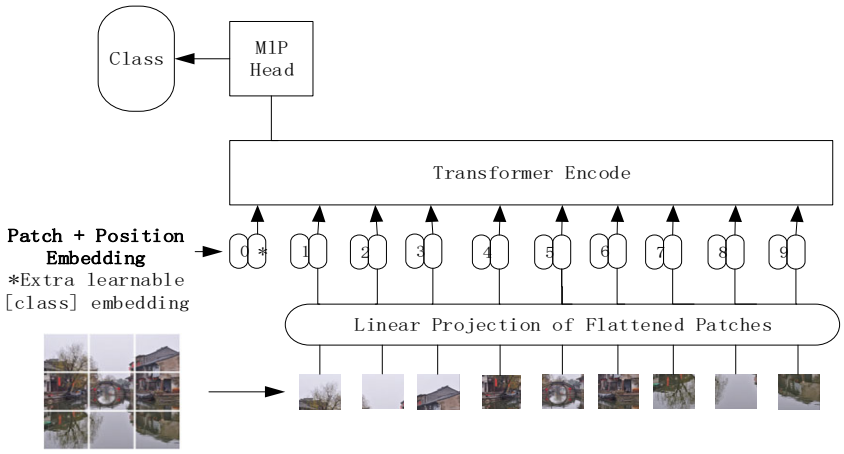

Figure 4. VIT flow frame diagram

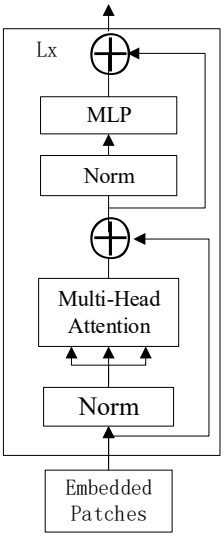

Figure 5. Transformer Encoder module

As can be seen from the figure above, the entire process of Vision-Transformer model framework is as follows: First of all, the picture (224*224) is divided into patches (16*16) of fixed size on the Patch embedding. Then each image will generate 224*224/16*16=196 patches, that is, the length of the input sequence is 196, and the dimension of each Patch is 16*16*3=768. The dimension of the linear projection layer is 768*N (N=768), so the dimension after the input passes through the linear projection layer is still 196*768, that is, there are 196 tokens, and the dimension of each token is 768. A special character class needs to be added, so the final dimension is 197x768. So far, a visual problem has been transformed into a seq2seq problem by patch embedding. After adding the location code, the location code has a total of N lines (N and the length of the input sequence), each line is a vector with the same dimension as the embedding dimension, after this operation, the dimension remains unchanged. It is then input into multiple blocks stacked by the transformer Encoder, and finally the output corresponding to the class is used as its final output after multiple self-attention.

ViT uses global relationship modeling to expand the receptive field of the image to obtain more context information, and uses multi-head self-attention mechanism to distinguish which features are more important to the task, making the output results of the model more explanatory. Experiment and discussion

4. Diagnostic Process

The main process of the model method proposed in this paper includes the pre-processing of data set noise reduction, signal-image conversion, and model training and verification. The process is as follows:

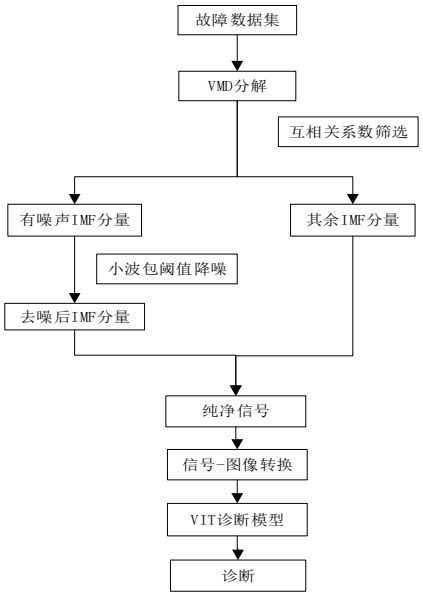

Figure 6. Diagnostic flow chart

Vmd-wpt-vit algorithm model firstly decomposed the vibration data of rolling bearing by VMD, calculated the interrelation number of each eigenmode component and the original signal data, selected the component with small interrelation number as the component containing noise, and denoised it by wavelet packet threshold denoising algorithm. After that, the denoised components are reconstructed with other components to obtain the denoised signal data, and then these data are converted into two-dimensional gray scale graphs and input into the ViT model for training. The model with optimal parameters is obtained through the training data set, and the test set data is input into the model for testing.

5. Experiment and Discussion

5.1. VMD-WPT Noise reduction

As an important part of the wind turbine, the rolling bearing's normal operation is very important for the fan. Therefore, the vibration data of rolling bearings in CWRU data set is selected for the experiment. The data set mainly includes rolling element (RE), inner ring (IR) and outer ring (OR) faults. Each type of failure has three different sizes of damage diameters of 0.018mm (0.007 in), 0.36mm (0.014 in) and 0.54mm (0.028 in) : The above 9 faults and the normal state of the bearing together constitute the 10 health states of the experiment. Each health state has 2000 samples, and each sample is obtained by collecting 50176 data points continuously at a collection frequency of 12Kw. Then, these samples are converted into two-dimensional gray graphs through VMD-WPT denoising and input into SwinT model for fault diagnosis.

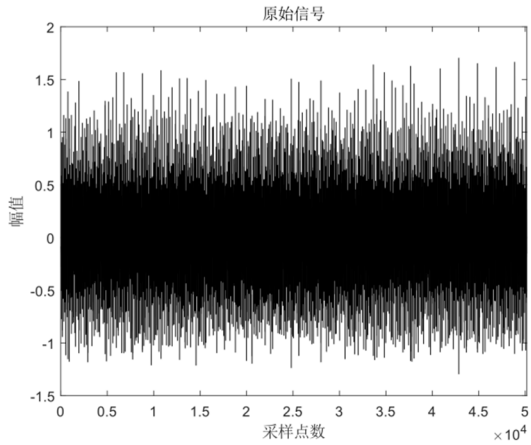

Figure 7. The damage size of the inner ring is 0.18

In this experiment, the characteristic parameters of VMD decomposition are as follows: penalty factor alpha=2000, decomposition mode number K=10, and initialization center frequency is uniformly distributed, i.e. init=1. The decomposed waveforms of each IMF component are shown in Figure 8.

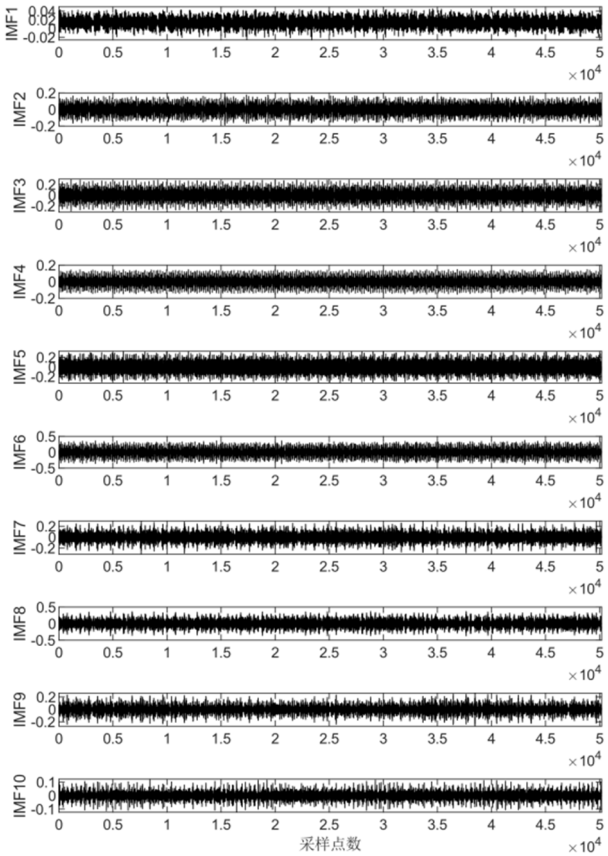

Figure 8. Waveform of IMF component

Equation (11) was used to calculate the number of correlations between each IMF component and the original vibration data, and the results are shown in Table 1.

$$r_i = \frac{\sum(x_i - \bar{x}_i)(y - \bar{y})}{\sum(x_i - \bar{x}_i)^2 \sum(y - \bar{y})^2} \tag{11}$$

Where, x_i represents each IMF component, and y represents the original vibration signal.

Table 1. Table of correlation coefficients of each IMF component

Modal components	IMF1	IMF2	IMF3	IMF4	IMF5
r_i	0.1059	0.1059	0.1059	0.1059	0.1059
Modal components	IMF6	IMF7	IMF8	IMF9	IMF10
r_i	0.4979	0.4979	0.4979	0.4979	0.4979

As can be seen from Table 1, there is a big difference between the correlation number between the original vibration signal and each IMF component. Among them, the correlation number of IMF2, IMF3, IMF4, IMF5, IMF6, IMF7, IMF8 and IMF9 are all greater than 0.2, which can be regarded as the dominant IMF component. The coefficients of IMF1 and IMF10 are both less than 0.2, and it is preliminarily concluded that these two IMF components contain more noise signals.

IMF1 and IMF10 were denoised by wavelet packet threshold and reconstructed with other IMF components to obtain pure signals. Figure 9 shows the waveform of pure vibration signal after noise reduction.

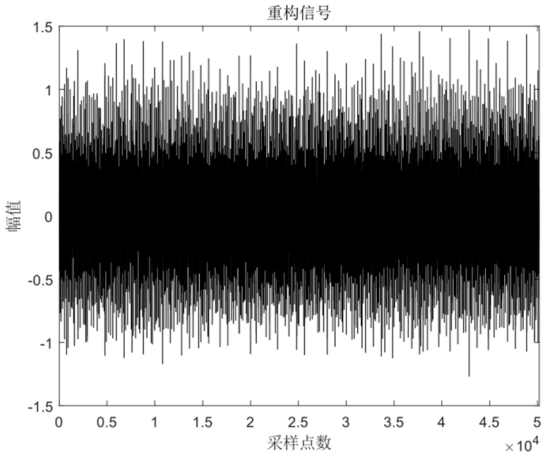

Figure 9. Waveform of vibration signal after noise reduction

In combination with FIG. 7 and FIG. 9, it can be seen that in terms of signal peak value and local waveform characteristics, the original vibration signal without noise reduction has a high similarity and coincidence degree with the pure signal after noise reduction. In order to further verify the similarity between the signal and the original signal after VMD-WPT combined noise reduction, the correlation number of the two signals of 10 types in the experiment was calculated, and the results are shown in Table 2.

Table 2. Table of correlation coefficients between various health States and original signals after noise reduction

Signal type	IR0.18	IR0.18	IR0.18	IR0.18	IR0.18
r_i	0.9886	0.9886	0.9886	0.9886	0.9886
Signal type	OR0.36	OR0.36	OR0.36	OR0.36	OR0.36
r_i	0.9632	0.9632	0.9632	0.9632	0.9632

It can be seen that after VMD-WPT denoising, the two signals still have a high correlation.

5.2. Signal to picture conversion

The vibration signals that have been denoised jointly by VMD-WPT are converted into two-dimensional gray-scale graphs using the method proposed in the above section. Each state contains 2000 picture samples with the size of 224*224, and a dataset with 20,000 picture samples is finally obtained. All kinds of gray-scale graphs are shown in Figure 10.

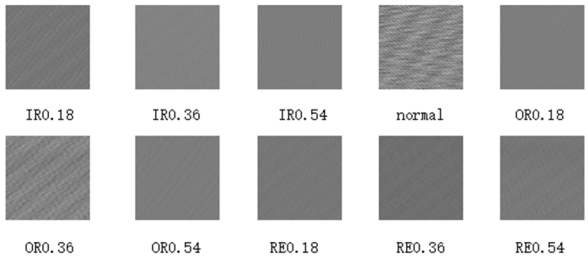

Figure 10. Gray scale of health status in 10

5.3. Model training and diagnosis

The data set is divided into the training set, the verification set and the test set according to the ratio of 8:1:1, then the training set and the verification set will use 18000 pictures to train the model, and the remaining 2000 pictures are the test set to test the model performance. Figure 11 shows the trend of accuracy over the number of iterations.

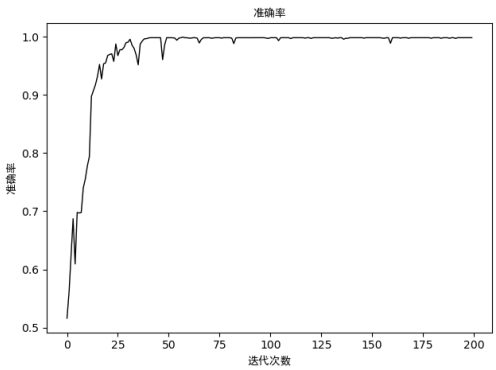

Figure 11. Accuracy curve

It can be seen that the accuracy rate is basically stable after 160 iterations, and the accuracy rate is close to 1, reaching 99.85%.

In order to better verify the effectiveness of this research method, set empirical mode factory-wavelet packet threshold denoising (EEMD-WPT) is used to pre-process fault data and input these data into Vision-Transformer model for comparative experiments. In addition, The data processed by VMD-WPT is also input into ResNet network to further verify the effectiveness of the proposed method. FIG. 12 to FIG. 14 show the confusion matrix obtained by various methods.

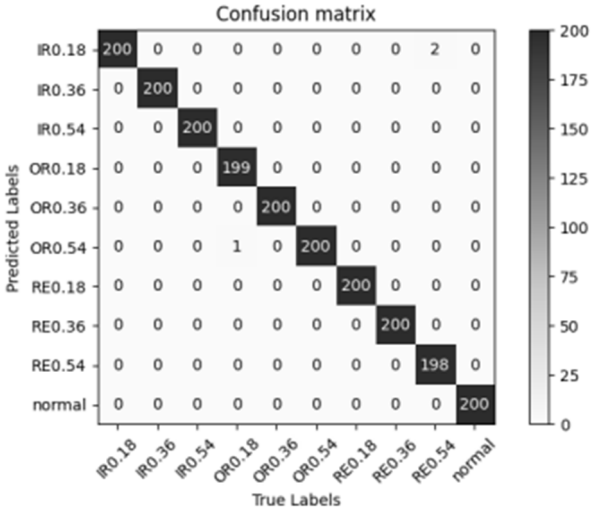

Figure 12. VMD-WPT-ViT combination

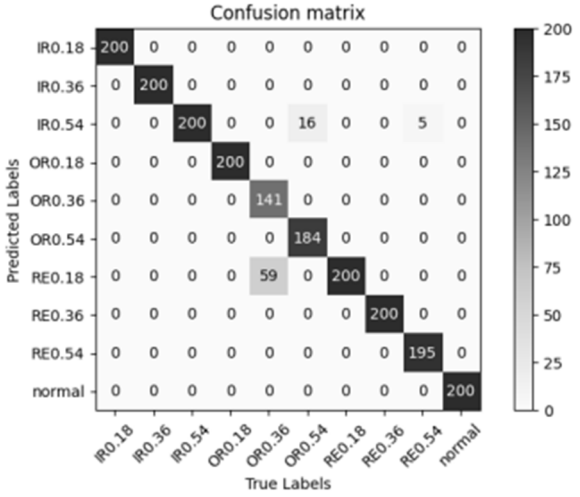

Figure 13. EEMD-WPT-ViT combination

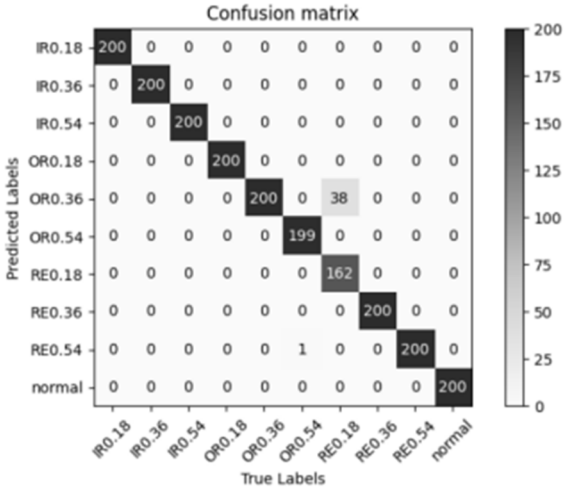

Figure 14. VMD-WPT-ResNet combination

As can be seen from Figure 12, there are only 3 diagnostic errors in VMD-WPT-ViT model, while the rest are diagnosed

correctly, and the diagnostic accuracy rate reaches 99.85%. There are 75 diagnostic errors in the EEMD-WPT-ViT model, and the diagnostic accuracy is 96%. It can be seen that using VMD decomposition for noise reduction can better remove noise. Although the diagnostic accuracy of VMD-WPT-ResNet model reaches 98.05%, it is still 1.7% lower than that of the model proposed in this paper. It is obvious that Vision-Transformer model can achieve better fault diagnosis results than ResNet network.

Finally, Accuracy, Precision, Recall and Specificity were selected to measure the performance of each specificity, as shown in Table 3. The accuracy, accuracy, recall and specificity of VMD-WPT-ViT model are higher than other models, which shows that the model algorithm has small error and high precision, and it is feasible and effective to apply it to the fault diagnosis of rolling bearings.

Table 3. Diagnostic results of each model

Diagnostic method	Accuracyt/%	Precision/%	Recall/%	Specificity/%
VMD-WPT-ViT	*99.85*	*99.85*	*99.85*	*99.98*
EEMD-WPT-ViT	*96*	*96.77*	*96*	*99.55*
VMD-WPT-RexNet	*98.05*	*98.35*	*98.05*	*99.78*

6. Conclusion

In order to improve the accuracy of rolling bearing fault diagnosis, a rolling bearing fault diagnosis method based on VMD-WPT-ViT model is proposed in this paper. Variational mode decomposition and wavelet packet threshold denoising algorithm are used to decompose the original vibration data and reduce the noise. The noise reduction method not only eliminates noise, but also effectively retains the characteristic information in the original vibration signal. And to a large extent help improve the accuracy of subsequent fault diagnosis; The one-dimensional time series data is converted into two-dimensional grayscale images, and a ViT-Transformer network is built to independently extract important features in two-dimensional images using the self-attention mechanism, which further improves the model diagnosis accuracy. Through comparative experiments, the diagnostic accuracy and performance of VMD-WPT-ViT model are more verified, and the effectiveness and feasibility of the proposed method in fault diagnosis are proved.

References

[1] ZANG Xiaosong. The installed capacity of wind power industry may usher in explosive growth in 2023 [N]. Securities Times,2023-01-03(A06).

[2] Wind Energy Professional Committee of China Renewable Energy Society. Statistical Briefing on China's wind power Hoisting Capacity in 2021 [J]. Wind Energy,2022(05):38-52.

[3] Liu Y, Li J, Li Q, et al. Transfer learning with inception ResNet-based model for rolling bearing fault diagnosis[J]. Journal of Advanced Mechanical Design, Systems, and Manufacturing,2022, 16(2): 1-19.

[4] Chen Yuhang, Li Zhengping, Xiao Lei. Fault Diagnosis of Roller Bearing in Spinning Frame based on FFT-1D-CNN [J]. Cotton Textile Technology,2023,51(01):16-21.

[5] Ye Kai, Huang Xuemei, Zhang Leian, et al. High-speed Bearing Fault Diagnosis of Wind Turbine Based on DE-SVM [J]. Machine Tool & Hydraulics,2022,50(18):153-157. (in Chinese)

[6] Tian Sai, Yao Bin, Chen Binqiang, et al. Rolling bearing Fault Diagnosis Method based on Morlet Wavelet and improved Kurtosis [J]. Tool Technology, 2002,56(10):141-146.

[7] Liu Bei, CAI Jianhua, Peng Ziqi. Fault Diagnosis Method of Rolling Bearing based on VMD-IMDE-PNN [J]. Noise and Vibration Control,202,42(05):96-101+133.

[8] Zhu J, Hu T, Jiang B, et al. Intelligent bearing fault diagnosis using PCA-DBN framework[J]. Neural Computing and Applications, 2020, 32(14): 10773-10781.

[9] Wang H, Liu Z, Peng D, et al. Understanding and learning discriminant features based on multiattention 1DCNN for wheelset bearing fault diagnosis[J]. IEEE Transactions on Industrial Informatics, 2019, 16(9): 5735-5745.

[10] CAI Gai-Po, Zhao Xiao-Tao, Hu Xian-Neng, et al. Research on De-noising Method of Cylinder Vibration Signal of Combined Ball Mill with CEEMDAN-Wavelet Threshold [J]. Mechanical Science and Technology, 2019,39(07):1077-1085.

[11] Jiang Lei, Zhu Mingsi, Li Feilong. Motor Bearing Fault Diagnosis Based on AMD and Wavelet Threshold Denoising [J]. Machine Design & Manufacture,2021(09):47-50.

[12] Kim S J, Kim K, Hwang T, et al. Motor-current-based electromagnetic interference de-noising method for rolling element bearing diagnosis using & emission sensors [J]. Journal of Measurement, 2022193:110912.

[13] Liu Changliang, Wu Yingjie, Zhen Chenggang. Fault Diagnosis of Rolling Bearings Based on variational Mode Decomposition and Fuzzy C-means Clustering [J]. Proceedings of the CSEE, 2015:3558-3565.

[14] Zi-Fei X, Min-Nan Y, Chun L. Application of the proposed optimized recursive variational mode decomposition in nonlinear decomposition[J]. Acta Physica Sinica, 2019, 68(23):238401

[15] Dragomiretskiy K, Zosso D. Variational mode decomposition [J]. IEEE Transactions on Signal Processing, 2013, 62(3): 531-544.

[16] Zhou Yina, LU Jingyi, Zhang Yong, Dong Hongli, Yang Dandi. VMD-SG-WT denoising method and its application in Chaos denoising [J]. Journal of Northeast Petroleum University, 2019,44(04):113-120+132+13.

[17] Dong Lichao, Guo Xingming, Zheng Yineng. Research on Wavelet packet denoising algorithm of heart sound signal based on CEEMD [J]. Journal of Vibration and Shock, 2019, 38(09): 192-198+222. (in Chinese)

[18] Wen L, Li X Y, GAO , et al. A new convolutional neural network-based data-driven fault diagnosis method[J]. IEEE Transactions on Industrial Electronics, 2018, 65(7): 5990-5998. (in Chinese)

[19] Dosovitskiy A, Beyer L, Kolesnikov A, et al. An image is worth 16x16 words: Transformers for image recognition at scale[J]. ar**v preprint ar**v:2010.11929, 2020.