Research Article

Wavelet Packet Transform-Assisted Least Squares Support Vector Machine for Gear Wear Degree Diagnosis

Hongmin Wang [ID] [1] **and Liang Chan** [ID] [1,2]

[1]*School of Automation, Harbin University of Science and Technology, Harbin 150080, China*
[2]*Heilongjiang Provincial Key Laboratory of Complex Intelligent System and Integration, Harbin 150080, China*

Correspondence should be addressed to Hongmin Wang; 13904810908@163.com

Wear degree detection of gears is an effective way to prevent faults. However, due to the interference of high-speed meshing vibration and environmental noise, the weak vibration signal generated by the gear is easily covered by the noise, which makes it difficult to detect the degree of wear. To address this issue, this paper proposes a novel gear wear degree diagnosis method based on local weighted scatter smoothing method (LOWESS), wavelet packet transform (WPT), and least square support vector machine (APSO-LSSVM) optimized by adaptive particle swarm algorithm. According to the low signal-to-noise ratio characteristic of gear vibration signal, LOWESS is first used to preprocess the signal spectrum. Then, the characteristic parameters used to characterize gear wear are extracted from different decomposition depths by WPT and, finally, combined with APSO-SVM to diagnose the degree of gear wear. Compared with the basic least squares support vector machine, the improved method has better performance in sample classification. The experimental results show that the method in this paper can effectively reduce the diagnosis error caused by background noise, and the diagnosis accuracy reaches 98.33%, which can provide a solution for the health status monitoring of gears.

1. Introduction

With the popularity of electric vehicles, the evaluation of mechanical durability (reliability) of electric vehicles has always been a research hotspot in the field of electric drive systems. As devices that transmit power or rotational motion, the health status of gears is closely related to the life of an electric vehicle's drivetrain. Therefore, there is a great need for advanced wear diagnosis technology to minimize unplanned downtime caused by gear wear and predict its future development trend so that corrective measures can be taken in time before any further damage occurs to the machine [1, 2].

In general, gear faults produce shocks, with the result that transient excitations can be observed in the vibration signal. However, local wear of gear tooth surfaces usually leads to faint transients in the signal, and the accuracy of the collected signal is greatly affected by background noise, measurement point location, and operating environment,

which makes it difficult to capture the intrinsic information about local defects of gears, resulting in the original fault detection method not applicable to the distinction of wear degree [3, 4]. Although the fault detection method has a certain manifestation of the degree of wear detection, it is still not enough to support the judgment of the degree of wear. Since the degree of wear detection can prevent the occurrence of fault, it is necessary to develop a new weak fault detection method that can effectively detect gear wear to learn the wear characteristics and accurately evaluate the health of the gear. The gearbox health monitoring technology based on vibration analysis is very effective for the diagnosis of gear wear, as the change of the vibration signal is a response to gear defects and wear growth. This study adopts the diagnosis method based on vibration analysis. Compared with the diagnosis after the fault, the wear diagnosis during the operation can realize the early warning of the fault and the tracking of the defect [5, 6].

At present, fault diagnosis of rotating machinery is usually divided into three steps: fault signal acquisition, fault feature extraction, and fault pattern identification. One of the most critical of these steps is fault feature extraction, which involves mapping the original vibration signal to relevant feature parameters to characterize the degree of wear of the gear. In order to realize the intelligent diagnosis of health status, in-depth research has been carried out in the field of fault diagnosis. For example, Islam et al. [7] focused on the process of feature extraction and recognition and provided azimuthal health status information by defining new evaluation indicators-defect rate and two-dimensional visualization of acoustic signals. Liang et al. [8] applied wavelet transform to extract time-frequency image features and verified the effectiveness of the method by two experiments. Yu [9] proposed a new time-frequency analysis method-transient extracting transform, which can effectively characterize and extract the transient components in the vibration signal of rotating machinery. Zhang et al. [10] used shift-invariant K-means singular value decomposition dictionary learning to detect early faults in gearbox bearings. Teng et al. [11] used the cepstrum method to distinguish the approximate frequency components and applied complex wavelet transform to detect the characteristics of weakly loaded faults buried in high energy. Babouri et al. [12] introduced an advanced signal processing method called cyclostationarity analysis. The experimental results prove that this method has the ability to diagnose defects in rotating machinery. Wang et al. [13] defined the cross-correlation kurtosis as an indicator to clarify and reconstruct the preprocessed signals by a stochastic resonance method, which is combined with a machine learning method based on ensemble bagged trees to detect early faults in bearings. Liu et al. [14] proposed an improved variational modal decomposition based method to distinguish planetary gear failures under different operating conditions. The above studies have proposed new methods or improved related algorithms for the early detection of rotating machinery, but the early health monitoring of gears under the background of strong noise still needs further research.

In order to solve the problem that strong noise background is not enough to realize online monitoring, this paper combines the locally weighted scatter smoothing method (LOWESS) with the wavelet packet transform (WPT) to smooth the noise and enhance the signal-to-noise ratio. The LOWESS can eliminate meaningless extreme points, maximize the elimination of noise effects, and protect the local integrity of the original signal. Algorithms based on wavelet packet transform can better describe the signal characteristics of early gear wear. Other signal decomposition algorithms (VMD, EMD, etc.) differ in that they are very sensitive to frictional vibrations in time series. Due to the high sensitivity of characteristic parameters, this method is particularly suitable for early monitoring of gear health status. Li et al. [15] used the second-generation redundant wavelet packet transform to extract statistical features, diagnosed gear faults through support vector machines, and applied this method to gearbox fault diagnosis. Derkacheva et al. [16] used the LOWESS algorithm to reduce the noise of satellite observation data to accurately measure the glacier speed. To process nonstationary gear vibration signals, Bafroui et al. [17] combined the resampling technique at constant angle increment with continuous wavelet transform and identified gear faults through MLP neural network. To eliminate the influence of noise, Lu et al. [18] applied the bootstrap resampling method to optimize the parameters of CEEMD and diagnosed bearing faults through support vector machines. Shao et al. [19] proposed a method for estimating the direction of arrival of a weak nonstationary signal, which utilizes the spatial time-frequency distribution of cross terms to solve the problem of weak signals that cannot be extracted under a noise background. It can be found from the above literature that a suitable preprocessing method helps a lot in the analysis of nonstationary signals.

For LS-SVM, there is a possibility of classifier performance degradation when dealing with certain highly correlated features of nonstationary signals. However, the least squares support vector machine classifier based on APSO optimization has good potential to classify any nonstationary signal, while the traditional LS-SVM-based fault diagnosis method has the disadvantages of easy failure in handling nonstationary vibration signals and large dependence on the classifier hyperparameters. In the literature, Wei et al. [20] applied moth-flame (MFO) optimized LS-SVM to detect rolling bearing faults. Dutta et al. [21] proposed a feature extraction framework based on the combination of multivariate empirical mode decomposition and phase space reconstruction and applied LS-SVM to classify EEG signals. Ma and Liu [22] proposed an intelligent optimal weighted LS-SVM identification method and applied APSO to optimize its parameters. Experiments proved that the method has the identification ability of nonlinear models. Wu [23] used the APSO algorithm to optimize SVM and applied the method to mixed model prediction.

In order to overcome the above problems, this paper proposes a WPT-assisted LS-SVM method for gear wear degree diagnosis. To solve the problems of noise interference and parameter adjustment for health assessment, an optimized method for signal decomposition and feature extraction was used. Then APSO is applied for the selection of the parameter optimal solution of LS-SVM. Finally, the extracted features were used to train LS-SVM to classify the health status of gear in the wear process. The novelty of the proposed method is the implementation of signal smoothing techniques to filter the noise and the APSO algorithm to select the optimal solution of the vector machine parameters to facilitate a high-quality and efficient training process.

2. Theoretical Background

2.1. Locally Weighted Scatter Smoothing Method. Locally weighted scatter smoothing method (LOWESS) [24] is a useful tool to view the relationship between two-dimensional variables. The main idea is to take a certain percentage of local data to fit a polynomial regression curve to observe patterns and trends in the local presentation of the data. In data smoothing, it works similarly to the moving average technique, where the value of each point within a specified

window is obtained by a weighted regression of data from neighboring points within the window.

Taking point x as the center, a fixed length of data is intercepted before and after, record $(x, y^\wedge)$ is the central value of the regression line, and $y^\wedge$ is the corresponding value of the curve after fitting. For all n data points, n weighted regression lines can be made, and the connection of the center value $y^\wedge$ of each regression line is the LOWESS curve of this piece of data. The definition of its loss function is as follows:

$$\sum_i w^{(i)} \left(y^i - \theta^T x^{(i)} \right)^2, \tag{1}$$

where $\theta^T x^{(i)}$ is the predicted value of each sample, the actual value is $y^{(i)}$, and the sample weight is $w^{(i)}$. Its mathematical expression is as follows:

$$w^{(i)} = \exp\left(-\frac{\left(x^{(i)} - x \right)^2}{2\tau^2} \right), \tag{2}$$

where τ is the attenuation factor, x is the sample to be predicted, and $x^{(i)}$ are the surrounding samples. When predicting the sample x, the farther the surrounding sample is from it, the smaller the weight of the surrounding sample is. When τ is smaller, the weight decays faster as the distance increases.

2.2. Wavelet Packet Transform. WPT [25] is an effective tool for dealing with nonstationary sequences, with the difference that the Fourier transform requires the signal to be transformed in either the time or the frequency domain. Wavelet analysis can decompose the signal in two scales of time domain and frequency domain at the same time; it can not only well portray the locality of the signal in time domain, but also reflect the locality of the signal in frequency domain, so it can focus on any detail of the object.

Wavelet packet transform (WPT) is an extension of the discrete wavelet transform. Its multiresolution analysis capability can further decompose the detailed information of the signal in the high-frequency region. WPT decomposes a signal into two subsignals, the approximation signal and the detail signal, and the wavelet packet decomposition tree formed is shown in Figure 1. The definition of the wavelet packet function is as follows:

$$W_{j,k}^n(t) = 2^{j/2} W^n \left(2^j t - k \right), \tag{3}$$

where j is the scaling parameter and k is the translation parameter. When $n = 0$ and $n = 1$, the first two wavelet packet functions $W_{0,0}^0(t) = \phi(t)$ and $W_{0,0}^1(t) = \psi(t)$ represent the scaling function and the mother wavelet function, respectively. When $n = 2, 3, ..., N$, the recursive relations of other wavelet packet functions are defined as follows:

$$W_{0,0}^{2n}(t) = \sqrt{2} \sum_k h(k) W_{1,k}^n (2t - k),$$
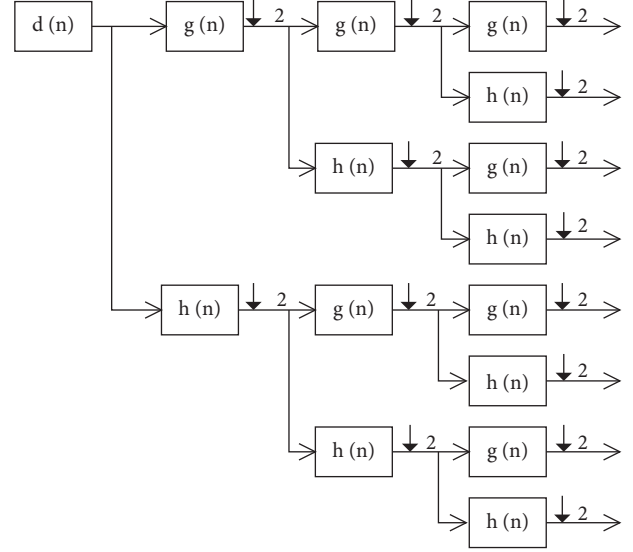$$W_{0,0}^{2n+1}(t) = \sqrt{2} \sum_k g(k) W_{1,k}^n (2t - k), \tag{4}$$

Figure 1: Wavelet packet decomposition tree of the signal.

where $h(k)$ and $g(k)$ are low-pass and high-pass filters associated with the scaling function and the mother wavelet function. The wavelet coefficients are obtained by the inner product of the signal $x(t)$ and the wavelet packet function, as shown in the following formula:

$$\Omega_j^n(k) = \langle x, W_{j,k}^n \rangle = \int_{-\infty}^{\infty} x(t) W_{j,k}^n(t) dt. \tag{5}$$

Each wavelet packet coefficient $\Omega_j^n(k)$ is a specific subspace of each frequency resolution level, which is related to the scaling parameter j and the oscillation parameter n. WPT performs complete decomposition on each node to produce two components, low-pass approximation and high-pass detail coefficients. Downsampling operation causes the signal to be decomposed from $\left\{ \Omega_0^1(k), k = 1, 2, ..., N \right\}$ to $\left\{ \Omega_j^n(k), k = 1, 2, ..., (N/2)^j \right\}$.

2.3. Least Square Support Vector Machine. Least squares support vector machine (LS-SVM) [26], as an improved algorithm that can improve the computing power and performance of the model, overcomes the deficiencies of long training time, randomness of training results, and overlearning compared with artificial neural networks, making the efficiency much higher.

The least squares support vector machine (LS-SVM) is an improvement of the support vector machine (SVM), with the difference that LS-SVM changes the inequality constraints in the original method into equality constraints. Least squares method is achieved by solving a set of linear equations, thus greatly facilitating the solution of Lagrange. The classification problem can be expressed as

$$\text{minimize}: L_{P_{LS-SVM}} = \frac{1}{2} w \cdot w + C \frac{1}{2} \sum_{i=1}^N \xi_i^2, \tag{6}$$

$$\text{subject to}: t_i (w \cdot \phi(i) + b) = 1 - \xi_i, \quad i = 1, ..., N,$$

where C is the regularization parameter. For nonlinearly separable samples, the optimization problem of LS-SVM is equivalent to solving the following dual optimization problem:

$$L_{D_{LS-SVM}} = \frac{1}{2} w \cdot w + C \frac{1}{2} \sum_{i=1}^{N} \xi_i^2 - \sum_{i=1}^{N} \alpha_i \left(t_i \left(w \cdot \phi(i) + b \right) - 1 + \xi_i \right),$$

(7)

where α_i is the Lagrange multiplier, derivation of the various variables of the Lagrange function; the optimality conditions of the above formula can be obtained as follows:

$$\frac{\partial L_{D_{LS-SVM}}}{\partial w} = 0 \longrightarrow w = \sum_{i=1}^{N} \alpha_i t_i \phi(X_i),$$

$$\frac{\partial L_{D_{LS-SVM}}}{\partial b} = 0 \longrightarrow \sum_{i=1}^{N} \alpha_i t_i = 0,$$

$$\frac{\partial L_{D_{LS-SVM}}}{\partial \xi_i} = 0 \longrightarrow \alpha_i = C\xi_i, \quad i = 1, ..., N,$$

$$\frac{\partial L_{D_{LS-SVM}}}{\partial \alpha_i} = 0 \longrightarrow t_i \left(w \cdot \phi(X_i) + b \right) - 1 + \xi_i = 0, \quad i = 1, ..., N.$$

(8)

The aforementioned equations can be equivalently written as

$$Z = \begin{bmatrix} t_1 \phi(X_i) \\ t_N \phi(X_N) \end{bmatrix},$$

$$\Omega_{LS-SVM} = ZZ^T,$$

(9)

where $\overrightarrow{1} = [1, 1, ..., 1]^T$, $T = [t_1, t_2, ..., t_N]^T$, and $\alpha = [\alpha_1, \alpha_2, ..., \alpha_N]^T$. For a new sample, the decision function of the LS-SVM classifier can be expressed as

$$f(X) = \text{sign} \left(\sum_{i=1}^{N} \alpha_i t_i K(X, X_i) + b \right).$$

(10)

The Lagrange multipliers α_i in LS-SVM are proportional to the training error ξ_i. The solution of linear equations makes the calculation efficiency much higher than the SVM model.

3. Optimization of LS-SVM and Diagnosis Model Description

3.1. Parameter Selection Based on Adaptive Particle Swarm Optimization.
Due to the use of LS-SVM for health status recognition, its parameters have a great impact on the performance of classification. We need to adopt an algorithm to calculate the optimal value of the parameter.

Particle swarm algorithm [27] is a heuristic swarm intelligence optimization algorithm, and its basic concept comes from the study of bird predation behavior. The algorithm calculates the fitness value of the sample data by the fitness function and finds the best position based on the position of the target point and the current position, as well as all particles in the whole population. The next position of

each particle is determined by its own motion experience and the motion experience of other particles, and the optimal solution is found through continuous iteration. The particle velocity and position update iteration are as follows:

$$v_i = w \cdot v_i + c_1 \text{rand0} \cdot \left(p\text{best}_i - x_i \right) + c_2 \cdot \text{rand0} \cdot \left(g\text{best}_i - x_i \right),$$

$$x_i = x_i + v_i,$$

(11)

where x_i is the current position of the particle, v_i is the particle velocity, w is the inertia factor, and c_1 and c_2 are the coefficients.

Particle swarm algorithm has a fast convergence speed, but it also has the disadvantages of easy premature convergence, low search accuracy, and low efficiency of late iteration. To this end, variation operations are introduced into the PSO algorithm to reinitialize certain variables with a certain probability, allowing the optimized particle swarm algorithm to jump out of the currently searched local optimum position and carry out the search in a larger space, thus increasing the possibility of the algorithm to find the optimum value in the space.

3.2. Diagnostic Model.
In order to better solve the problem of diagnosing gear wear diagnosis under complex conditions, a new diagnostic scheme based on WPT and LS-SVM is proposed in this paper. Figure 2 shows the overall framework of our proposed method. The method is divided into three main steps, namely, signal preprocessing and wavelet packet transformation, feature extraction, and early gear health status recognition:

Step1 : the vibration sensor is used to collect and store the gear vibration signal, and then the LOWESS algorithm is used to preprocess the original signal to obtain the signal data after noise reduction and smoothing.

Step2 : Wavelet packet transform is used to process vibration signals at different decomposition depths to improve the signal-to-noise ratio. Then, features are extracted from the obtained wavelet packet coefficients to obtain a dataset that can effectively describe the health status of the gears.

Step3 : LS-SVM is used to train the feature set and identify the health status of the gear.

After the gear vibration signal with noise is sampled by the vibration sensor, the noise is attached to the gear vibration signal by means of amplitude modulation, which results in extreme points on the original signal and causes interference in the subsequent process of health status identification. In order to achieve online detection, it is necessary to eliminate those meaningless extreme points to obtain real gear vibration data. For this purpose, we preprocess the spectrum of the signal by LOWESS. The LOWESS algorithm chooses a first-order polynomial model to replace local information and uses linear weighted least squares to match the polynomial. In the algorithm, we set 2%
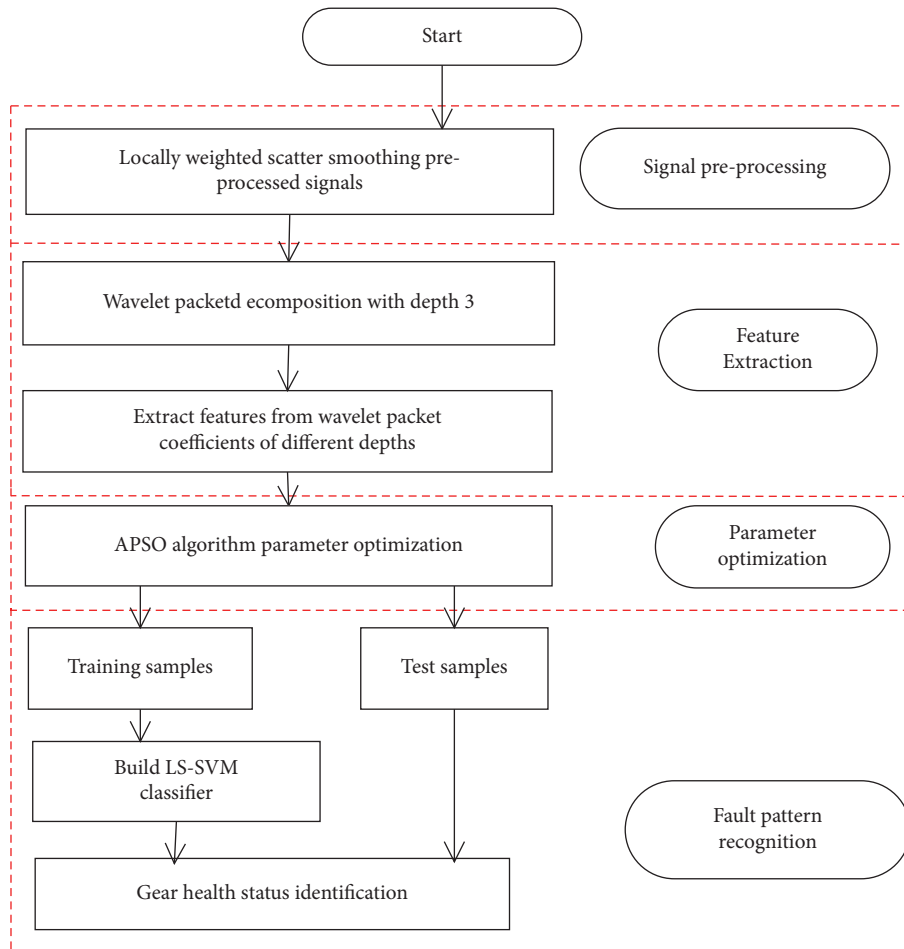
FIGURE 2: Framework of the proposed method.

of the signal length as the matching length. Experiments have proved that the 2% matching length not only guarantees the local characteristics of the signal's frequency spectrum but also eliminates meaningless extreme points.

The preprocessed signal considers the wavelet packet decomposition level from 1 to 3, and there will be 14 wavelet packet nodes. The db9 wavelet, as a wavelet basis function widely used for gear health condition monitoring, is used as a wavelet basis function for WPT filtering in this paper. The wavelet packet node with the maximum depth of 3 is plotted in Figure 3. To achieve early gear health status monitoring, 7 features were extracted from 14 wavelet packet nodes. Table 1 lists 7 feature parameters, including standard deviation, kurtosis, mean value, skewness, clearance factor, square root amplitude value, and crest factor. A total of 98 characteristic parameters are generated for LS-SVM training.

As mentioned earlier, in order to promote high-efficiency and high-speed training, it is necessary to select the best training parameters. The APSO algorithm not only converges to the global optimum quickly but also has good discrimination ability. The algorithm updates the position by updating the optimal position in the position experienced by the individual and the optimal position of the fitness searched by all the particles in the population and finds the optimal parameters by update iterations. It should be noted

that the parameters in the extracted feature set are highly correlated, and there are differences in the sensitivity contribution of different parameters to the classification. To maximize the implementation of efficient training, the APSO-LSSVM-based diagnostic model can accurately identify the health status of gears.

4. Construction of Experimental Platform and Description of Experimental Data

Gear with a worn tooth surface may cause a fatal accident. To avoid this situation, the defect of the gear must be found as early as possible in order to track the growth of the defect. It is important to improve the diagnosis of gear health status to minimize the economic loss due to downtime. To conduct evaluation experiments, a unipolar gearbox experiment platform was created to collect gear vibration signal data and simulate several common wear faults types in gearboxes to verify the effectiveness of the method.

4.1. Construction of the Experimental Platform. In order to verify the effectiveness of the proposed method, the experiment involves three faults, namely, gear wear, pitting, and cracks, and they were researched and
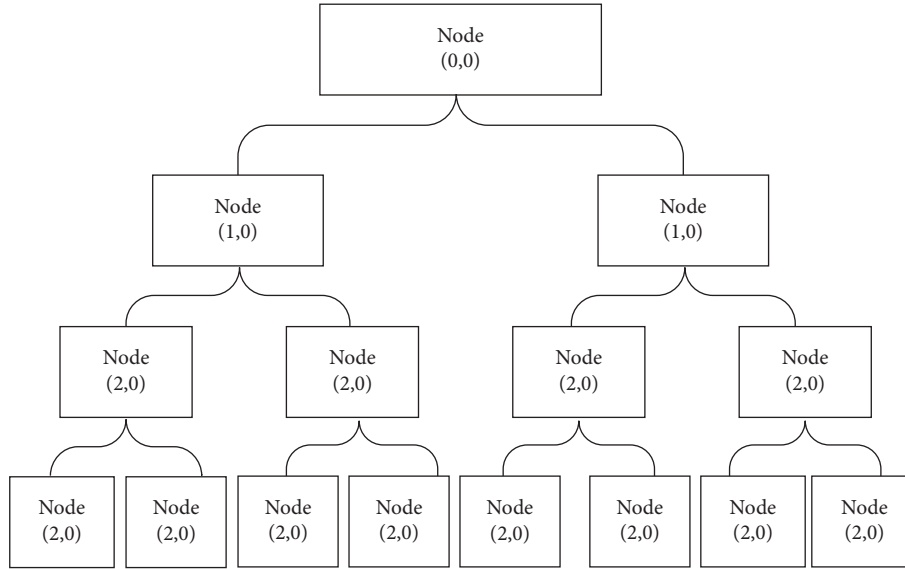
Figure 3: The wavelet packet node with the maximum depth of 3.

Table 1: The 7 feature parameters extracted.

Attributes	Calculation method	Attributes	Calculation method
Standard deviation	$\sqrt{\frac{1}{N-1}\sum_{i-1}^{N}(x_i-\overline{x})^2}$	Kurtosis	$\frac{1}{N}\sum_{i=1}^{N}x_i^4$
Mean value	$\lim_{N\longrightarrow\infty}\frac{1}{N}x_i(t)=E[X(t)]$	Skewness	$\frac{1}{N}\sum_{i=1}^{N}x_i^3$
Clearance factor	$(\max(|x_i|)/(1/N)\sum_{i=1}^{N}|x_i|^{(1/2)})\sqrt{b^2-4ac}$	Square root amplitude value	$\sqrt{(\sum_{i=1}^{N}x_i^2/N)}=\sqrt{(x_1^2+x_1^2+\ldots+x_N^2/N)}$
Crest factor	$(\max(|x_i|)/\sqrt{(1/N)\sum_{i=1}^{N}x_i^2})$	—	—

analyzed using the experimental test rig shown in Figure 4(a). The overall framework of the experimental test platform is shown in Figure 4(b). It can be noted that the experimental test platform mainly consists of three parts: gearbox, signal acquisition device, and signal storage device, where the gearbox mainly consists of a DC brushless motor, active wheel, driven wheel, and belt used to drive the gear. The signal acquisition device consists of a vibration sensor and tachometer, the vibration sensor is located at 0.5 cm above the vertical of the driving wheel, and a total of 6 types of gear vibration signals are collected. The gear vibration signal is sampled during gearbox operation and transferred to the computer for storage, with a sampling frequency of fs = 5120 Hz and a speed setting of 880r/min.

4.2. Experimental Data Description. All experimental data are obtained under the same operating conditions, and the gear speed and load are constant. The collected subhealthy gear vibration signals are all under the condition that the gear can still operate normally without breakage of teeth and other faults. The health status of the experimental gear is described in Table 2.

In this paper, six types of subhealth gear vibration signals are collected, which are divided into three different degrees of tooth surface wear and three different levels of tooth surface cracks. 120 samples are collected for each category of data, of which 60 are used for training and 60 are used for testing, and all samples are 720 in total, with category labels set to 1, 2, 3, 4, 5, and 6. The dataset description is as follows (Table 3.

5. Experimental Results and Discussion

By implementing LOWESS, the envelope spectrum of the preprocessed vibration signal is obtained as shown in Figure 5(b). Figure 5(a) shows the envelope spectrum of the original signal. In Figure 5(b), under the premise of protection of local structure of spectrum, most of the extreme points caused by noise amplitude modulation are eliminated. It also means that it is possible to evaluate the gear health in the early wear stage by monitoring the vibration changes of the preprocessed gear vibration signal.

Generally, the wear condition of the gears needs to be disassembled for inspection, so the current detection method is not conducive to the implementation of online detection. In this paper, to avoid excessive consumption of detection time, a large amount of collected data is used to provide a guarantee for the identification of gear health status. In order to prove the superiority of the proposed method, the improved LS-SVM and the basic LS-SVM model were tested according to the adaptive parameters.
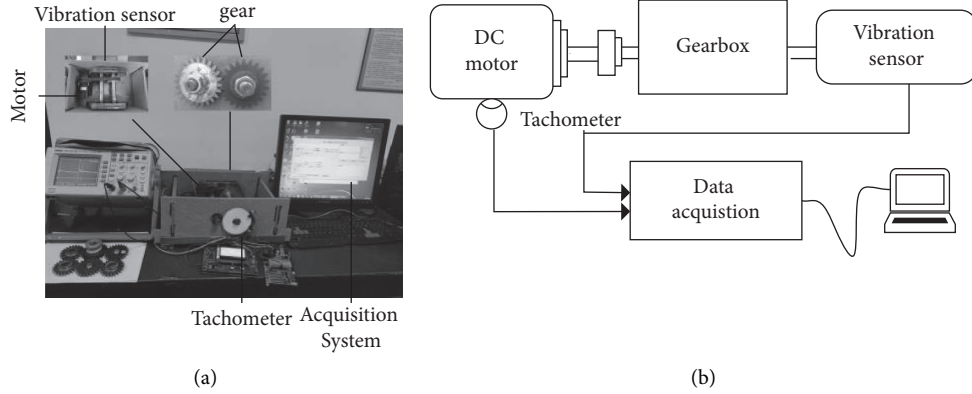
FIGURE 4: The gear test stand. (a) Front view of the experimental test setup. (b) Data acquisition system diagram.

TABLE 2: Gear health status description.

Gear category	Depth of wear	Gear category	Crack depth (mm)
Mild wear	The public law line shortens by 0.05	Mild crack	Crack depth 0.35
Moderate wear	Public law line shortening <0.1	Moderate crack	Crack depth 0.68
Heavy wear	Public law line is shortened by 0.1–0.2	Heavy crack	Crack depth 1.23

TABLE 3: Description of gear dataset.

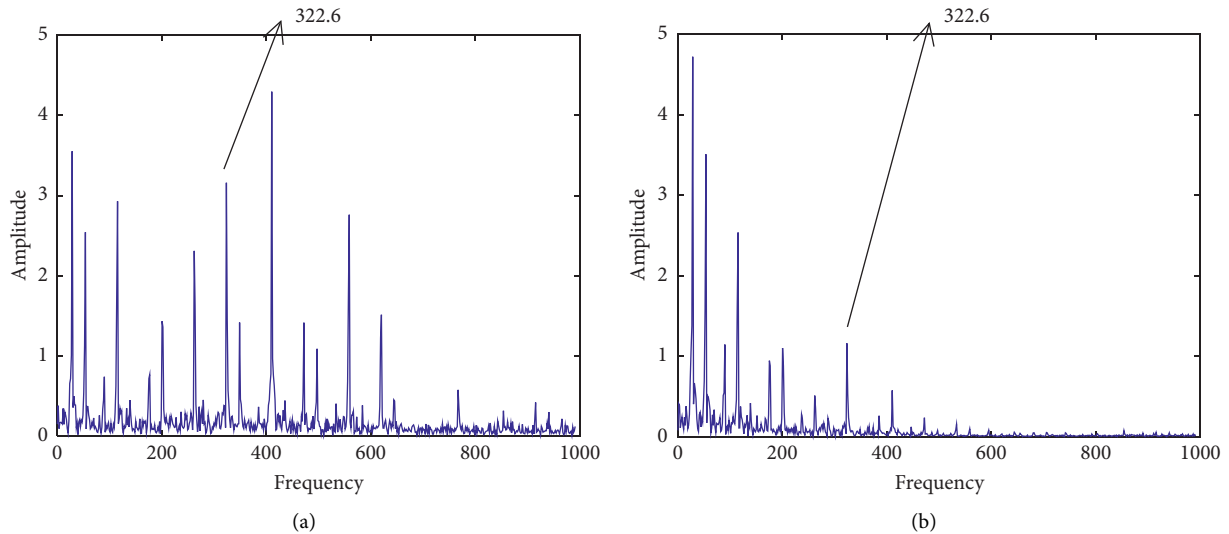| Health status	Number of training samples	Number of test samples	Label
Mild wear	60	60	1
Moderate wear	60	60	2
Heavy wear	60	60	3
Mild crack	60	60	4
Moderate crack	60	60	5
Heavy crack	60	60	6

FIGURE 5: Preprocessed signal diagram. (a) Original signal envelope spectrum. (b) Envelope spectrum after LOWESS.

We obtained the classification results of two different models of the test set data, as shown in Table 4.

It can be observed that when the same features are used as input, the diagnosis result of the proposed method is much better than the basic LS-SVM model. At the same time, it also shows that appropriate parameters can facilitate the high-precision and high-efficiency training of LS-SVM. Overall, the method used in this paper is more representative in the identification of health status of gears.

TABLE 4: Model diagnosis result.

Model	Number of correctly classified samples	Number of misclassified samples	Accuracy (%)
APSO-LSSVM	354	6	98.33
LS-SVM	255	105	70.83
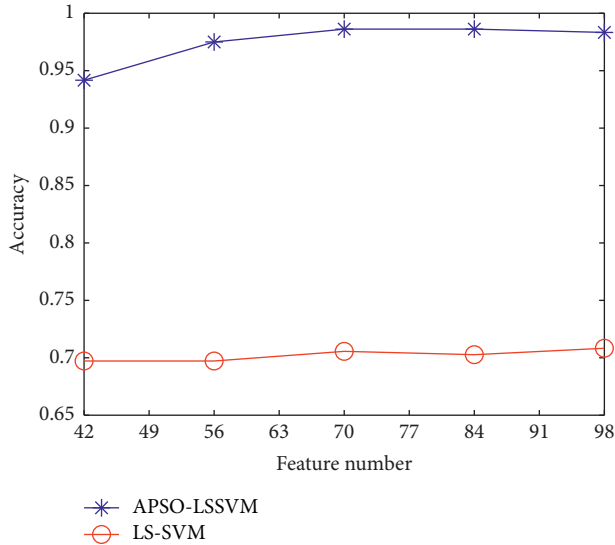
FIGURE 6: Accuracy versus number of features graph.

TABLE 5: Test results of APSO-LSSVM with different number of features.

Feature number	Label					
	1	2	3	4	5	6
84	100	96.67	100	100	95	100
70	100	96.67	100	100	95	100
56	100	90	100	100	95	100
42	100	90	100	100	95	80

Result (%) is the row label for the data rows above.

TABLE 6: Test results of LS-SVM with different number of features.

Feature number	Label					
	1	2	3	4	5	6
84	91.67	61.67	58.33	66.67	63.33	75
70	100	70	66.67	58.33	63.33	66.67
56	100	56.67	66.67	66.67	61.67	66.67
42	100	58.33	66.67	66.67	60	66.67

Result (%) is the row label for the data rows above.

Because the proposed model is a multiclassification model based on sensitive features, it is necessary to investigate the influence of the number of feature parameters on the classification results of the model. The features extracted by the wavelet packet transform, although highly correlated, suffer from the problem of feature redundancy. If a smaller number of features is chosen, the computational complexity will be reduced accordingly, and if a sufficient number of features are used, the recognition accuracy will be greatly improved. To find a balance between computational burden and recognition accuracy to ensure high accuracy, this article investigates the relationship between the number of features and accuracy, as shown in Figure 6, and with the increase in the number of features, the accuracy rate increases first, and the accuracy rate remains high when the number of features changes from 70 to 98. However, when the number of features drops to 56, the recognition rate decreases.

Tables 5 and 6 summarize the single fault diagnosis results of the proposed method and the comparison model. Obviously, the recognition rate of the method used in this paper is better than that of the comparison model, and the diagnostic results of samples belonging to different health states under the same number of features have achieved satisfactory performance.

For the data in Table 6, under the premise of the same number of features, the accuracy of label 2 and label 5 is lower, because different features have different sensitivity to classification. After comparing the misdiagnosed samples in label 2 and label 5, for adjacent fault categories, the difference in signal

part feature values is small, which affects the performance of the classifier. But in general, the diagnosis method in this paper can provide a solution for the diagnosis of gear health.

6. Conclusion

This paper introduces a new method for diagnosing the health status of gears. The main idea of this method is to use LOWESS to reduce noise and then extract feature parameters from the wavelet packet nodes in layers 0 to 3. Finally, the APSO optimized LS-SVM was used to identify 6 gear faults with different wear degree at the same motor speed. A large amount of experimental data verifies that the accuracy of this method for gear wear diagnosis is 98.33%. By comparing the experiments, we obtained the following conclusions:

(1) Under the same working conditions, the combination of LOWESS and WPT can effectively reduce noise interference and enhance the ability of certain features in classification

(2) The parameters selected by the APSO algorithm are optimal values, which can provide a solution for monitoring the health status of gears

(3) Compared with the basic LS-SVM, the method used in this article has better performance and higher accuracy

Data Availability

The codes used in this paper are available from the author upon request.

Conflicts of Interest

Acknowledgments

References

[1] M. Sun, C. Lu, Z. Liu, Y. Sun, H. Chen, and C. Shen, "Classifying, predicting, and reducing strategies of the mesh excitations of gear whine noise: a survey," *Shock and Vibration*, vol. 2020, Article ID 9834939, 20 pages, 2020.

[2] S. Wang, P. Niu, and Y. Guo, "A piecewise hybrid stochastic resonance method for early fault detection of roller bearings," *IEEE Access*, vol. 8, pp. 73320–73329, 2021.

[3] H. Li, Y. Xu, D. An, L. Zhang, S. Li, and H. Shi, "Application of a flat variational modal decomposition algorithm in fault diagnosis of rolling bearings," *Journal of Low Frequency Noise, Vibration and Active Control*, vol. 39, no. 2, pp. 331–351, 2020.

[4] M. J. Hasan, M. M. M. Islam, and J.-M. Kim, "Acoustic spectral imaging and transfer learning for reliable bearing fault diagnosis under variable speed conditions," *Measurement*, vol. 138, pp. 620–631, 2019.

[5] A. Silva, A. Zarzo, and J. M. Munoz-Guijosa, "Early fault detection of single-point rub in gas turbines with accelerometers on the casing based on continuous wavelet transform," *Journal of Sound and Vibration*, vol. 487, Article ID 115628, 2020.

[6] D. Wang, "K-nearest neighbors based methods for identification of different gear crack levels under different motor speeds and loads: Revisited," *Mechanical Systems and Signal Processing*, vol. 70-71, pp. 201–208, 2016.

[7] M. M. M. Islam and J.-M. Kim, "Automated bearing fault diagnosis scheme using 2D representation of wavelet packet transform and deep convolutional neural network," *Computers in Industry*, vol. 106, no. 1, pp. 142–153, 2019.

[8] P. Liang, C. Deng, J. Wu, and Z. Yang, "Intelligent fault diagnosis of rotating machinery via wavelet transform, generative adversarial nets and convolutional neural network," *Measurement*, vol. 159, Article ID 107768, 2020.

[9] G. Yu, "A concentrated time-frequency analysis tool for bearing fault diagnosis," *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 2, pp. 371–381, 2020.

[10] Z.-H. Zhang, J.-M. Ding, C. Wu, and J.-H. Lin, "Impulsive component extraction using shift-invariant dictionary learning and its application to gear-box bearing early fault diagnosis," *Journal of Central South University*, vol. 26, no. 4, pp. 824–838, 2019.

[11] W. Teng, X. Ding, X. Zhang, Y. Liu, and Z. Ma, "Multi-fault detection and failure analysis of wind turbine gearbox using complex wavelet transform," *Renewable Energy*, vol. 93, no. 1, pp. 591–598, 2016.

[12] M. K. Babouri, N. Ouelaa, T. Kebabse, and A. Djebala, "Application of the cyclostationarity analysis in the detection of mechanical defects: comparative study," *International Journal of Advanced Manufacturing Technology*, vol. 103, no. 5–8, pp. 1681–1699, 2019.

[13] S. Wang, P. Niu, and Y. Guo, "Early diagnosis of bearing faults using decomposition and reconstruction stochastic resonance system," *Measurement*, vol. 158, no. 1, Article ID 107709, 2020.

[14] L. Liu, L. Chen, and Z. Wang, "early fault detection of planetary gearbox based on acoustic emission and improved variational mode decomposition," *IEEE Sensors Journal*, vol. 21, no. 2, pp. 1735–1745, 2020.

[15] N. Li, R. Zhou, Q. Hu, and X. Liu, "Mechanical fault diagnosis based on redundant second generation wavelet packet transform, neighborhood rough set and support vector machine," *Mechanical Systems and Signal Processing*, vol. 28, pp. 608–621, 2012.

[16] A. Derkacheva, J. Mouginot, and R. Millan, "Data reduction using statistical and regression approaches for ice velocity derived by landsat-8, sentinel-1 and sentinel-2," *Remote Sensing*, vol. 12, no. 12, Article ID 1935, 2020.

[17] H. Heidari Bafroui and A. Ohadi, "Application of wavelet energy and Shannon entropy for feature extraction in gearbox fault detection under varying speed conditions," *Neurocomputing*, vol. 133, pp. 437–445, 2014.

[18] Y. Lu, R. Xie, and S. Y. Liang, "CEEMD-assisted kernel support vector machines for bearing diagnosis," *International Journal of Advanced Manufacturing Technology*, vol. 106, no. 7-8, pp. 3063–3070, 2020.

[19] S. Shao, A. Liu, C. Yu, H. Yang, Y. Li, and B. Li, "Spatial time-frequency distribution of cross term-based direction-of-arrival estimation for weak non-stationary signal," *EURASIP Journal on Wireless Communications and Networking*, vol. 2019, no. 1, Article ID 239, 2019.

[20] J. Wei, H. Huang, L. Yao, Y. Hu, Q. Fan, and D. Huang, "New imbalanced fault diagnosis framework based on Cluster-MWMOTE and MFO-optimized LS-SVM using limited and complex bearing data," *Engineering Applications of Artificial Intelligence*, vol. 96, Article ID 103966, 2020.

[21] S. Dutta, M. Singh, and A. Kumar, "Classification of non-motor cognitive task in EEG based brain-computer interface using phase space features in multivariate empirical mode decomposition domain," *Biomedical Signal Processing and Control*, vol. 39, pp. 378–389, 2018.

[22] L. Ma and X. Liu, "A novel APSO-aided weighted LSSVM method for nonlinear hammerstein system identification," *Journal of the Franklin Institute*, vol. 354, no. 4, pp. 1892–1906, 2017.

[23] Q. Wu, "Hybrid forecasting model based on support vector machine and particle swarm optimization with adaptive and Cauchy mutation," *Expert Systems with Applications*, vol. 38, no. 4, pp. 9070–9075, 2011.

[24] R. Andersen, "Nonparametric methods for modeling nonlinearity in regression analysis," *Annual Review of Sociology*, vol. 35, no. 1, pp. 67–85, 2009.

[25] R. Yan, R. X. Gao, and X. Chen, "Wavelets for fault diagnosis of rotary machines: a review with applications," *Signal Processing*, vol. 96, pp. 1–15, 2014.

[26] G. B. Huang, H. Zhou, X. Ding, and R. Zhan, "Extreme learning machine for regression and multiclass classification," *IEEE Transactions on Systems, Man, and Cybernetics - Part B: Cybernetics*, vol. 42, no. 2, 2012.

[27] Y. Zhang, S. Wang, and G. Ji, "A comprehensive survey on particle swarm optimization algorithm and its applications," *Mathematical Problems in Engineering*, vol. 2015, Article ID 931256, 38 pages, 2015.